

Alla mia famiglia

Indice

Introduzione	vi
1 Premessa	1
1.1 Alcune note sul ramo malattia	1
1.1.1 Il ramo malattia in Italia	3
1.2 La compagnia assicurativa	4
2 La valutazione del rischio in Solvency 2	8
2.1 Introduzione	8
2.2 Cenni al sistema Solvency 1	9
2.3 Il progetto Solvency 2	11
2.4 La Standard Formula	17
2.4.1 La Standard Formula secondo la Direttiva	17
2.4.2 I Quantitative Impact Studies	20
2.4.3 La Standard Formula secondo il QIS5	21
2.4.3.1 L'approccio Market Wide	22
2.4.3.2 Il SCR per la nostra compagnia secondo l'approccio Market Wide	26
2.4.3.3 L'approccio Undertaking Specific	29
2.4.3.4 Il SCR per la nostra compagnia secondo l'approccio Undertaking Specific	33

2.5	Qualche considerazione	35
3	La valutazione del rischio attraverso il CRM	39
3.1	Introduzione	39
3.2	Lo studio del costo singolo sinistro	43
3.2.1	L'adattamento a distribuzioni note	45
3.2.1.1	La stima della distribuzione Log-Normale	45
3.2.1.2	La stima della distribuzione Gamma	48
3.2.1.3	Esponenziale	51
3.2.1.4	La distribuzione di Weibull	52
3.2.1.5	La distribuzione Beta Generalizzata del Secondo tipo	53
3.2.1.6	Gaussiana Inversa	57
3.2.2	I criteri per la scelta di una distribuzione	58
3.2.2.1	Test di Kolmogorov-Smirnov	59
3.2.2.2	Test di Anderson-Darling	60
3.2.2.3	Test del χ^2	60
3.2.2.4	Test del Rapporto di Verosimiglianza	61
3.2.3	La scelta della distribuzione	62
3.2.4	Il Costo singolo sinistro senza un modello	63
3.3	Il numero dei sinistri	66
3.3.1	Poisson	69
3.3.2	Poisson Misturata	70
3.3.2.1	La Poisson Composta Gamma	73
3.3.3	Qualche precisazione sul metodo	74
3.4	La simulazione	75
4	Un approccio innovativo per le stime	81
4.1	Introduzione	81
4.2	Gli approcci classici	82

4.2.1	Il metodo dei momenti	82
4.2.2	Il metodo della massima verosimiglianza	83
4.3	Il <i>Minimum Distance Approach</i>	84
4.4	Note rispetto ai tre approcci	87
4.5	Alcune proprietà degli stimatori MDA	88
4.5.1	Un primo confronto	88
4.5.2	Un secondo confronto	90
4.5.2.1	Robustezza	91
4.5.2.2	Alcune distanze alternative	92
4.5.3	La valutazione dell'adattamento delle stime	97
4.6	La stima del costo singolo sinistro attraverso il MDA	102
4.6.1	Le stime sotto ipotesi di Log-Normalità	103
4.6.2	Le stime sotto ipotesi di Burr Inversa	104
4.7	La distribuzione del costo sinistri aggregato	107
5	Alcuni modelli composti	113
5.1	Introduzione	113
5.2	La distribuzione misturata	114
5.2.1	L'impostazione teorica	114
5.2.2	Massima Verosimiglianza o <i>Minimum Distance Approach</i>	117
5.2.2.1	La prima verifica	118
5.2.2.2	La seconda verifica	121
5.2.3	Le stime	122
5.3	La <i>spliced distribution</i>	125
5.3.1	L'impostazione teorica	125
5.3.2	Le stime	130
5.4	La valutazione del rischio	132

A	Un approccio individuale per la <i>frequency</i>	135
A.1	Poisson	136
A.2	Poisson Misturata	136
A.3	Poisson Composta	137
A.3.1	La Poisson Composta Gamma	138
A.3.2	La Poisson Composta Log-Normale	138
A.3.3	La Poisson Composta Gaussiana Inversa	138
A.4	La scelta della distribuzione	138
A.5	La simulazione	139
B	Incidenza del numero di simulazioni	142
B.1	La consistenza delle stime di media e varianza	143
B.2	La consistenza delle stime di VaR e Tail-VaR	146
B.3	Alcune verifiche empiriche	147
	Conclusioni	150
	Riferimenti bibliografici	153

Ringraziamenti

Desidero innanzitutto ringraziare il Professor Nino Savelli per avermi consigliato questo progetto di tesi e per avermi trasmesso, nel corso degli ultimi anni, la passione per le Scienze Attuariali.

Un ringraziamento sentito va al Professor Diego Zappa per avermi permesso, confrontandosi spesso con me, di sviluppare molti degli argomenti trattati e per avermi destato anche la passione per la Statistica.

Pur non potendoli nominare, desidero ringraziare coloro che mi hanno consentito l'accesso ai dati della compagnia assicurativa, necessari per il taglio applicativo dell'elaborato.

Voglio inoltre ringraziare con affetto la mia famiglia, in particolare i miei genitori che sono sempre stati un valido sostegno per me e per il mio percorso di studi.

Infine, ringrazio tutti i miei amici, incontrati prima e durante questi cinque anni universitari. In particolare ringrazio il gruppo di fedelissimi gallaratesi, coloro che hanno condiviso con me l'esperienza dell'appartamento, chi mi è stato accanto nel lavoro della consiglieri e (perché no?) tutti coloro che, chiedendomi un aiuto a preparare un esame, mi hanno negato la possibilità di disimparare.

Introduzione

Giuro su Apollo medico e Asclepio e Igea e Panacea e su gli dei tutti e le dee, chiamandoli a testimoni, di tener fede secondo le mie forze e il mio giudizio a questo giuramento e a questo patto scritto. Riterrò chi mi ha insegnato quest'arte pari ai miei stessi genitori e metterò i miei beni in comune con lui, e quando ne abbia bisogno lo ripagherò del mio debito e i suoi discendenti considererò alla stregua dei miei propri fratelli, e insegnerò loro quest'arte, se essi desiderano apprenderla, senza compensi né impegni scritti: trasmetterò gli insegnamenti scritti e verbali e ogni altra parte del sapere ai miei figli così come anche ai figli del mio maestro e agli allievi che hanno sottoscritto il patto e giurato secondo l'uso medicale, ma a nessun altro. Mi varrò del regime per aiutare i malati secondo le mie forze e il mio giudizio, ma mi asterrò dal recar danno e ingiustizia. Non darò a nessuno alcun farmaco mortale, neppure se richiesto, né mai proporrò un tale consiglio; ugualmente non darò alle donne pessari per provocare l'aborto. Preserverò pura e santa la mia vita e la mia arte. Non opererò neppure chi soffre del male della pietra, ma lascerò il posto a coloro che sono esperti di questa pratica. In qualsiasi casa andrò, io vi entrerò per aiutare i malati, astenendomi dal recare ingiustizia e danno, e specialmente da ogni atto di libidine su corpi di donne e uomini, liberi o schiavi. E quanto vedrò e udirò esercitando la mia professione, ed anche al di fuori di essa nei miei rapporti con gli uomini, se mai non debba essere divulgato attorno, ritenendolo alla stregua di un sacro segreto. Se dunque terrò fede a questo giuramento e non vi verrò meno, mi sia dato godere il meglio della vita della vita e dell'arte, tenuto da tutti e per sempre in onore. Se invece sarò trasgressore e spergiuro, mi colga il contrario di ciò.

Giuramento di Ippocrate

Con questo celebre giuramento, attribuito ad Ippocrate, i medici iniziano la propria professione da oltre 2300 anni. Chiunque ha potuto sperimentare e sa che la professione

del medico è molto diversa dalle altre e non occorre un grave sforzo per intuire che, per tutta una serie di fattori, la scelta di esercitarla non può essere una semplice risposta alla necessità di svolgere un'attività.

Pur con tutte le peculiarità che contraddistinguono questa professione ed anche se ciò non può costituire il fine per cui la si esercita, le prestazioni dei medici devono essere retribuite e gli oneri possono arrivare a cifre importanti ed insostenibili per moltissime persone.

In Italia, storicamente, il compito di permettere a tutti i cittadini l'accesso alle cure mediche è stato assolto in larghissima parte dal Sistema Sanitario Nazionale, erede, al pari degli istituti di previdenza, di una concezione dello Stato Sociale che affonda le sue radici nella tradizione cattolica.

La presenza di questo sistema fino ad oggi ha garantito le prestazioni mediche, ma per contro ha frenato lo sviluppo di forme alternative di finanziamento di queste spese, come potrebbe essere l'assicurazione del ramo malattia.

Tuttavia le assicurazioni che fanno riferimento a questo ramo sono rimaste molto utili soprattutto per i lavoratori autonomi, per quali, in caso di malattia, non è previsto un ente che provveda all'indennizzo per non aver potuto percepire uno stipendio.

Da ormai due decenni, diverse concause hanno messo in crisi lo Stato Sociale e ne hanno minato la sostenibilità. Dal lato della previdenza pubblica, per questa ragione, abbiamo assistito negli anni a diverse riforme soprattutto volte ad assicurarne la sostenibilità nel lungo periodo, riducendo l'ammontare delle prestazioni garantite. Abbiamo anche assistito, in risposta a questo, ad una forte crescita e ad uno sviluppo delle forme di previdenza complementare. Ma il Sistema Sanitario è sempre rimasto pressoché inalterato fino ad oggi.

Tuttavia, il Presidente del Consiglio dei Ministri, il Professor Mario Monti, ha dichiarato intervenendo alla presentazione del progetto del Centro di biotecnologia Ri.Med:

La crisi ha colpito tutti, il campo medico non è un'eccezione.

Le previsioni di crescita economica e quelle di invecchiamento della popolazione mostrano che la sostenibilità futura dei sistemi sanitari, incluso il

nostro Servizio Sanitario Nazionale, di cui andiamo fieri, potrebbe non essere garantita, se non si individueranno nuove modalità di finanziamento e di organizzazione dei servizi e delle prestazioni.

Mario Monti
(27 Novembre 2012)

Queste parole, oggi, appaiono lontane dall'essere veramente poste in essere, ma in un'ottica di lungo periodo non possiamo escludere che questo possa accadere.

Dal momento stesso a partire dal quale lo stato si dovesse fare da parte, possiamo intuire che saranno dei privati a prenderne il posto e possiamo pensare che il mondo assicurativo abbia una posizione privilegiata a questo fine, proprio come accade in altri contesti oggi diversi dal nostro. Non si può quindi escludere la possibilità che questo ramo conosca una forte crescita nel corso dei prossimi anni.

Abbiamo quindi deciso di implicarci con questa realtà e l'analisi che intendiamo condurre, l'oggetto del nostro elaborato, è un tema molto sentito e molto vasto sotto il profilo attuariale: la valutazione del rischio assicurativo.

Una compagnia assicurativa, per definizione, svolge come attività principale l'assunzione di rischi che vengono ceduti da una collettività di individui, gli assicurati. Uno dei cardini che rende l'attività assicurativa realmente praticabile è la compensazione, cioè il fatto che un portafoglio di rischi è relativamente meno rischioso di quanto lo sia uno preso singolarmente. Tuttavia, a meno di un portafoglio di dimensione illimitata ed in assenza di rischi sistematici, la rischiosità non si annulla.

Le passività di una compagnia di assicurazione sono quindi per natura rischiose ed un'analisi di valutazione del rischio è di vitale importanza, al fine di prendere decisioni di management e di assicurare la solvibilità. Attraverso questo tipo di analisi, una compagnia cerca infatti di valutare quale sarà la misura degli oneri a cui sarà soggetta nel futuro, con una data probabilità ed un orizzonte temporale.

Il passaggio obbligato per questa analisi, come per molte altre, consiste nella costruzione di modelli. Nel corso dell'elaborato quindi ne produrremo una grande ed eterogenea varietà, passando dai più semplici ai più raffinati.

I modelli che svilupperemo nascono per definizione come delle semplificazioni della realtà e nel corso dell'elaborato tenderemo a costruirne di sempre più complessi per includere un sempre maggior numero di fattori.

Su questo punto, tuttavia, potrebbe emergere un dubbio, sempre spontaneo quando si ha a che fare con dei modelli. Ci confrontiamo con una domanda posta dall'economista Ariel Rubinstein in [34], che si chiede:

What are we trying to accomplish as economic theorists? We essentially play with toys called models. We have the luxury of remaining children over the course of our entire professional lives and we are even well paid for it.

Certo l'idea di sviluppare dei modelli per il solo gusto di farlo è innegabilmente accattivante, ma pronunciamo subito un *caveat* a questo riguardo. Una valutazione di stampo attuariale ha tanto più senso quanto più emerge, nel confronto con la realtà, la sua correttezza. I nostri modelli nascono con il fine prettamente tecnico di valutare il rischio e, proprio per evitare che si rivelino dei giocattoli efficienti solo nelle teorie, li vedremo applicati ad un caso pratico, al portafoglio di una compagnia realmente esistente.

Sorge quindi un altro tema, strettamente correlato al precedente. Quanto più un modello è complicato, tanto più ci aspettiamo che possa meglio cogliere differenti aspetti della realtà che ci si propone di descrivere. Perché quindi dovremmo occuparci dei modelli più semplici e non ci dedichiamo direttamente allo sviluppo dei modelli che presenteremo negli ultimi capitoli?

La risposta è nuovamente da ritrovarsi nella natura dell'analisi che qui conduciamo. Per portare a termine questo elaborato abbiamo avuto a disposizione diverso tempo ed energie, poiché si tratta di un'analisi tecnica ma pur sempre conclusiva di un percorso di studi, ma se dovessimo applicare questi modelli ad una compagnia, avremmo scadenze molto più serrate e ci troveremmo di fronte ad un *trade-off* tra precisione del modello e tempistiche per lo sviluppo, in modo particolare se la compagnia è di piccole dimensioni.

Non esiste infatti un modello che possa sempre essere considerato migliore degli altri sotto ogni aspetto.

Da un lato un modello deve essere sufficientemente elaborato sotto il profilo statistico-attuariale da garantire un buon grado di affidabilità delle stime, dall'altro le esigenze della compagnia richiedono che il modello possa essere applicabile. Un attuario, in una compagnia, si trova quindi nella condizione di dover trovare un equilibrio tra queste due richieste. Per questo non avrebbe senso omettere i primi modelli, se non altro come occasione di confronto con i modelli successivi.

Entriamo quindi nel merito del contenuto dell'elaborato, delineandone, per sommi capi, la struttura.

Il primo capitolo inizia riportando alcune nozioni introduttive sul ramo malattia e sul mercato di tale ramo in Italia. Successivamente si focalizza sulla descrizione del portafoglio della compagnia assicurativa sul quale svilupperemo e vaglieremo i modelli per la valutazione del rischio.

Il secondo capitolo sarà incentrato sulla valutazione del rischio assicurativo secondo il modello introdotto dall'EIOPA al fine di implementare la direttiva della comunità europea comunemente chiamata Solvency 2. Ne discuteremo pertanto le scelte e le assunzioni prima di applicare il modello al portafoglio della nostra compagnia.

Nel terzo capitolo condurremo la valutazione del rischio attraverso il *Collective Risk Model*, concentrandoci quindi sulla ricerca di una distribuzione per *frequency* e *severity*. Per quanto riguarda il costo sinistri in particolare, analizzeremo molti differenti modelli distributivi parametrici e li confronteremo attraverso alcune statistiche e attraverso l'uso del confronto grafico con il QQ-plot. Svilupperemo anche un modello distributivo non parametrico.

Nel quarto, proseguendo il filone del *Collective Risk Model*, vaglieremo un approccio innovativo per la stima dei parametri delle distribuzioni ed un metodo per la valutazione dell'adattamento. Lo adotteremo quindi per le distribuzioni della *severity* selezionate in

precedenza e, sotto queste nuove ipotesi, valuteremo nuovamente il rischio.

Infine, nel quinto capitolo, ci focalizzeremo nuovamente sulla ricerca di una distribuzione per il costo dei sinistri. Questa volta non adotteremo le distribuzioni selezionate nel terzo capitolo, ma utilizzeremo delle distribuzioni complesse frutto di altre distribuzioni più semplici. Valuteremo quindi l'adattamento di una distribuzione misturata e di una distribuzione *spliced*. Per ognuno di questi capitoli, confronteremo i risultati ottenuti con quelli dei capitoli precedenti.

Capitolo 1

Premessa

1.1 Alcune note sul ramo malattia

Il nostro elaborato sarà incentrato sulla valutazione del rischio assicurativo nel ramo malattia e già da questa affermazione si rende necessaria una precisazione.

In Italia la giurisprudenza ha collocato le assicurazioni sulla salute storicamente all'interno dei rami danni, più precisamente i rami 1 e 2¹. Tuttavia sono spesso state prestate sotto forma di garanzie accessorie anche per i contratti di assicurazione sulla vita.

Benché l'assicurazione sia sempre caratterizzata, per sua stessa natura, dall'assunzione di un rischio da parte dell'assicuratore dietro il pagamento di un premio da parte dell'assicurato, esiste una netta differenza tra le compagnie che esercitano i rami vita e danni. Lo stesso Codice Civile, dando la nozione di assicurazione nell'Articolo 1882 è costretto a presentare una definizione che già di per sé fa trasparire una certa dualità dei contratti assicurativi. Esso, infatti, recita:

L'assicurazione è il contratto col quale l'assicuratore, verso pagamento di un premio, si obbliga a rivalere l'assicurato, entro i limiti convenuti, del

¹ramo 1 infortuni e ramo 2 malattia

danno ad esso prodotto da un sinistro, ovvero a pagare un capitale o una rendita al verificarsi di un evento attinente alla vita umana.

Questa profonda differenza tra le compagnie che si occupano del ramo vita e quelle invece che sono attive nel ramo danni si ripercuote anche all'interno del ramo malattia, l'oggetto del nostro studio. Le compagnie di assicurazione sulla vita, infatti, stipulano polizze pluriennali, gestendo i rischi che si assumono sulla base del concetto di *accumulazione*, formando le riserve matematiche. Le polizze sono in questo caso sviluppate in modo che l'equilibrio attuariale sia raggiunto nell'arco dell'intera vita dell'assicurato o comunque di un periodo lungo, per ciascun contratto. Prendendo ad esempio le polizze di tipo *Long Term Care*, non di rado riscontriamo aspetti tipici delle polizze sulla vita²:

- il valore del capitale (o della rendita) garantito in caso di perdita dell'autosufficienza può essere strettamente collegato all'ammontare dei premi versati e all'andamento di una gestione separata;
- può essere previsto un limite di età al di sotto del quale l'assicurato non ha diritto alle prestazioni anche nel caso in cui si verifichi il sinistro (ad esempio 65 anni);
- può essere prevista la possibilità, per l'assicurato che decida di smettere di versare i premi, di continuare a godere della garanzia, fermo restando che le prestazioni siano collegate alla posizione maturata con i premi versati;
- può essere prevista la possibilità del riscatto, proprio grazie alle riserve matematiche.

Vi sono poi le polizze sviluppate con tecniche proprie dei rami danni. Le compagnie che operano nel mercato con quest'ottica gestiscono i propri rischi sulla base del concetto di polizza a *ripartizione*. Le compagnie stipulano contratti in modo che l'equilibrio sia raggiunto annualmente, quindi, al pari delle polizze danni, non prevedranno l'accumulo di un capitale sotto forma di riserva matematica.

²Maggiori approfondimenti si possono ritrovare in [32].

Rimanendo nell'esempio delle polizze Long Term Care, ad esempio, la compagnia AXA propone una polizza gestita con i criteri dei rami danni, che non da alcuna possibilità di riscatto e le cui prestazioni sono indipendenti dal numero di anni trascorsi in assicurazione.

La tecnica di gestione del rischio a cui si affidano le compagnie che operano in questo modo è il *pooling*. Con esso si intende che, in un portafoglio, la rischiosità diminuisce, in termini relativi, al crescere della dimensione, quindi la rischiosità relativa di una compagnia sarà comunque minore della rischiosità cui è soggetto ogni singolo assicurato.

Nel contesto italiano, le polizze sulla salute gestite con le tecniche dei rami vita rappresentano sicuramente una realtà di rilievo, anche perché, grazie al sistema di welfare presente nella nostro Paese, il rischio legato alla malattia viene avvertito meno che in altri paesi, quindi una fetta consistente delle assicurazioni malattia è in realtà costituita da garanzie accessorie offerte dalle polizze vita.

Tuttavia le polizze che prenderemo in considerazione noi saranno le polizze gestite secondo i criteri dell'assicurazione danni, poiché il nostro elaborato si incentrerà anche su un caso pratico di stima del rischio assicurativo di una compagnia realmente esistente.

Presenteremo nel seguito alcune informazioni di rilievo riguardanti il ramo malattia e la compagnia in oggetto.

1.1.1 Il ramo malattia in Italia

Presentiamo qui qualche statistica e qualche indicatore tecnico del ramo malattia, ottenuto attraverso [15] e [1].

Il ramo malattia è il secondo previsto nell'ordinamento italiano tra i rami danni. Secondo le statistiche ANIA, i premi emessi nel corso del 2011 ammontavano a 2176 milioni di euro, con una crescita rispetto al 2010 dello 0.2%. Questo dato rappresenta un'anomalia ed un forte rallentamento, considerando che la crescita è stata del 4.95% annuo se si consideriamo gli ultimi 10 anni. Per il totale dei rami danni, la crescita media degli ultimi 10 anni è stata del 2.0% risultando quindi inferiore a quella del ramo

che stiamo considerando, ma l'ultima variazione si è attestata ad un +2.1%, quindi il rallentamento del ramo malattia non è correlato ad un rallentamento del mercato assicurativo danni.

Diamo anche un'ordine dimensionale del ramo malattia. Questo ramo ha un peso molto limitato all'interno dei rami danni, stimato dall'ISVAP pari al 6% nel 2011, ma possiamo osservare una crescita tendenziale di questa percentuale, tanto che nel 2006 era circa il 4.9%.

Il *combined ratio* del ramo è sistematicamente rimasto al di sopra del 100% dal 1998 al 2010, con una media del 102.6%, ma nel 2011 è stato osservato il valore 99.6%. Possiamo anche cercare di risalire a quali fattori abbiano contribuito a questa riduzione.

- Sul fronte dei sinistri, possiamo osservare che il *loss ratio* del 2011 è stato del 70.3%, mentre mediamente gli anni precedenti si è registrato il 75.7%.
- Anche l'*expense ratio* è diminuito, passando dal 28.0% dei cinque anni precedenti al 26.1%.

La causa che viene individuata dall'ISVAP per spiegare la riduzione degli oneri per i sinistri è da ritrovarsi nella *severity*. Viene infatti posto in evidenza che il costo medio per sinistro (pagato e riservato) del 2011 è stato di € 435, risultando del 12.2% inferiore rispetto all'anno precedente. A parziale compensazione di questo valore, tuttavia, si è osservata una forte crescita nel numero dei sinistri denunciati, pari al 12.9%³.

Come ultima statistica notiamo anche che il risultato del conto tecnico è sempre stato negativo tra il 2006 ed il 2010, sia al lordo sia al netto della riassicurazione, ma nel 2011 si è osservato un valore positivo.

1.2 La compagnia assicurativa

Abbiamo a nostra disposizione alcuni dati riguardanti una compagnia di assicurazione. Per ragioni di riservatezza e di confidenzialità non abbiamo la possibilità rivelarne il

³Quest'ultimo dato, tuttavia, riguarda i sinistri per anno di denuncia, non per anno di accadimento.

nome, ma possiamo dire che si tratta di una compagnia di medie dimensioni che opera da qualche anno nel mercato italiano delle assicurazioni danni.

Ci è stata data la possibilità di valutare il rischio di un sottoportafoglio della dimensione di alcuni milioni di euro, in termini di volume dei premi contabilizzati, e per farlo abbiamo avuto accesso alle informazioni riguardanti i sinistri liquidati e riservati tra il 01/01/2008 ed il 31/12/2011.

Il sottoportafoglio riguarda il solo lavoro diretto della compagnia, sottoscritto attraverso il canale *Broker*, riguardate quindi polizze di tipo *Business*, che quindi assicurano più persone contemporaneamente e possono generare molti sinistri nel corso di un anno.

Con riferimento al solo anno 2008, abbiamo avuto accesso anche alle informazioni per quanto riguarda la tipologia del sinistro liquidato⁴, ma questo non ci è stato possibile con riferimento agli altri anni di bilancio.

Rimanendo per quanto possibile generici, teniamo anche in considerazione qualche particolarità rispetto alla metodologia con cui vengono liquidati i sinistri. La compagnia collabora in senso stretto con una società di servizi che ha sviluppato una rete con molte strutture convenzionate, sparse su tutto il territorio italiano. Per cui, per la maggior parte dei sinistri che riguardano il rimborso di spese mediche, avviene la liquidazione in via diretta alle strutture, ferme restando le limitazioni contrattuali. Possiamo pensare che questo fattore costituisca una concausa di una particolarità che esporremo nel corso dei capitoli successivi. Rispetto a quanto riscontrato sull'intero mercato del ramo malattia, infatti, osserviamo per la nostra compagnia un costo sinistri medio minore.

Notiamo anche che la nostra analisi si dovrà fermare al rischio di premio e dovremo trascurare il rischio di riservazione, proprio a causa della limitata lunghezza temporale del nostro dataset⁵.

Ultima notazione preliminare che riteniamo opportuno produrre riguarda la coerenza

⁴Come vedremo nella Sezione 2.4.2, questa informazione è cruciale, poiché la Standard Formula adottata dal QIS5 prevede una suddivisione in *LoB* differente da quella adottata dal Codice Civile italiano. Il ramo malattia, in realtà, viene suddiviso in quattro diverse *LoB*.

⁵Notiamo che il triangolo di *run-off* che potremmo costruire sarebbe un semplice 4×4 che non ci permette una coerente valutazione della rischiosità.

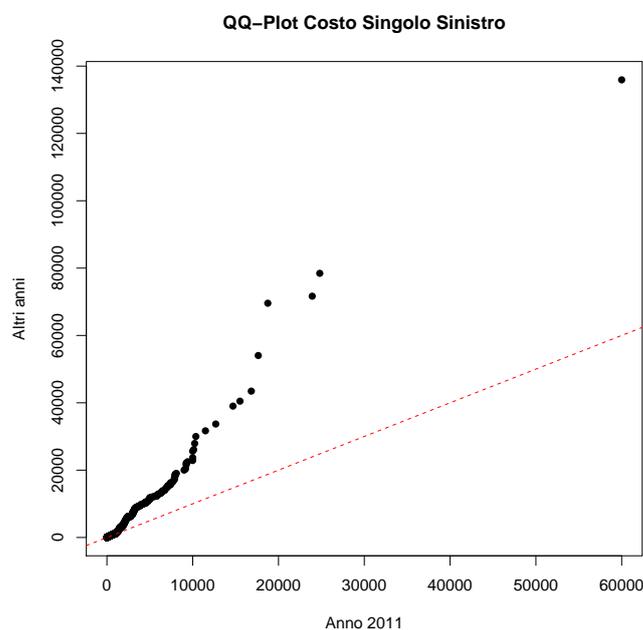


Figura 1.1: QQ-plot del costo sinistri

dei dati tra i vari anni di bilancio. Premettendo che il portafoglio è in fase di *start-up*, i volumi mostrano una forte crescita nell'arco del tempo, tanto che i premi contabilizzati nel 2011 sono più che raddoppiati rispetto al 2008. Ciò pone in dubbio l'assunzione di coerenza che ci risulta necessaria ai fini di una valutazione effettivamente rappresentativa della rischiosità del portafoglio.

Un'esemplificazione di questo viene data in Figura 1.1. Essa rappresenta il *QQ-plot* del costo dei singoli sinistri osservati nel 2011 messo a confronto con i costi rilevati nei tre anni precedenti⁶, con valori riportati alle grandezze monetarie del 2011.

Benché gli analoghi QQ-plot costruiti per gli altri anni di bilancio, che qui non riportiamo, mostrino delle distribuzioni sicuramente più coerenti, appare evidente come nell'anno 2011 qualcosa sia cambiato. Purtroppo non ci è stato possibile risalire alla causa di un simile cambiamento nel comportamento degli assicurati e ciò costituisce un grave pregiudizio alla possibilità per noi di condurre un'analisi esaustiva.

⁶Per rendere più evidente ciò che vogliamo mostrare non riportiamo nel grafico un costo considerato *outlier* verificatosi nel 2009.

In base ai dati in nostro possesso, comunque, valuteremo nei capitoli seguenti il rischio del portafoglio affrontando diversi approcci, dai più semplici proposti dai regolatori europei ai più complessi e raffinati.

Capitolo 2

La valutazione del rischio in Solvency 2

2.1 Introduzione

In questo capitolo, dopo una breve introduzione a Solvency 2, procederemo alla stima del rischio assicurativo attraverso i metodi proposti dall'EIOPA nelle specifiche tecniche del QIS5, il quinto studio d'impatto quantitativo condotto per stimare l'incidenza che avrebbe l'entrata in vigore della direttiva comunitaria Solvency II.

Ci concentreremo quindi in Sezione 2.4.3.2 a stimare il volume premi atteso per l'anno di bilancio 2012, che utilizzeremo all'interno delle sezioni successive per la stima del rischio attraverso il metodo proposto come riferimento per tutto il mercato assicurativo, denominato *Market wide*. Infine, sempre a partire dalla stima dei volumi, ci concentreremo sui metodi proposti come più aderenti all'effettiva rischiosità del portafoglio che stiamo considerando, denominati *Undertaking specific*.

2.2 Cenni al sistema Solvency 1

A causa delle peculiarità dell'attività assicurativa, il legislatore ha sviluppato una giurisprudenza che ne regola nel dettaglio ogni aspetto. Due delle caratteristiche tipiche delle le compagnie assicurative che occorre prendere in considerazione per comprendere alcune scelte di legislazione sono la così definita “inversione del ciclo monetario” e l'impossibilità di prevedere la dimensione degli oneri. Le compagnie, cioè, incassano i premi delle polizze prima di dover sostenere gli oneri legati ai sinistri da esse generati senza poterne conoscere con certezza la dimensione.

Tali caratteristiche rendono l'attività assicurativa particolarmente rischiosa e potenzialmente soggetta al *moral hazard* da parte degli assicuratori, poiché teoricamente essi potrebbero operare senza necessità di detenere mezzi propri, riversando quindi interamente il rischio sugli assicurati. I legislatori nazionali hanno quindi imposto alle compagnie, come vincolo per l'accesso all'attività assicurativa, di detenere un capitale adeguato a sostenere gli oneri prevedibili.

Il sistema di norme fino ad ora in vigore nell'Unione Europea discende dalle direttive di primo livello che davano l'indicazione, attraverso le regolamentazioni nazionali, di una misura di capitale minimo che dovesse essere detenuto dalle compagnie. Esso è calcolato in relazione all'ammontare dei premi emessi o dei sinistri liquidati per quanto riguarda il lavoro dei rami danni, mentre per i rami vita viene definito in funzione delle riserve matematiche e dei capitali sotto rischio.

Entrando nel dettaglio del requisito di capitale, definito Margine di Solvibilità, per le compagnie di assicurazione danni, esso è calcolato semplicemente come il massimo tra una quota dei premi emessi ed una quota dei sinistri liquidati. Più precisamente, indicando con B i premi emessi e con S i sinistri liquidati,

$$MMS = \alpha \cdot \max(\text{Margine premi}; \text{Margine sinistri}) \quad (2.1)$$

dove α indica il grado di conservazione,

$$\alpha = \max \left(50\%; \frac{S^{\text{Netto Rias}}}{S^{\text{Lordo Rias}}} \right) \quad (2.2)$$

e Margine premi e Margine sinistri sono calcolati come

$$\text{Margine premi} = \begin{cases} 18\%B & \text{se } B \leq \text{€ } 50\text{Mln} \\ 18\%\text{€ } 50\text{Mln} + 16\%(B - \text{€ } 50\text{Mln}) & \text{se } B > \text{€ } 50\text{Mln} \end{cases} \quad (2.3)$$

e

$$\text{Margine sinistri} = \begin{cases} 26\%S & \text{se } S \leq \text{€ } 35\text{Mln} \\ 26\%\text{€ } 35\text{Mln} + 23\%(S - \text{€ } 35\text{Mln}) & \text{se } S > \text{€ } 35\text{Mln} \end{cases} \quad (2.4)$$

Quindi, le uniche grandezze che indicano la rischiosità dell'impresa che sono considerate sono il fattore dimensionale e la quota di lavoro conservata¹.

Ovviamente, un requisito di capitale costruito in questo modo non risulta aderente e commisurato alla reale situazione di rischiosità dell'impresa, non prendendo in considerazione neppure la necessità di definire percentuali distinte per considerare la differente rischiosità dei diversi rami. Per compensare questa carenza del metodo di calcolo del Margine Minimo di Solvibilità è quindi richiesta alle compagnie una prudenzialità nelle stime delle proprie attività e passività, in modo da garantire comunque, in modo implicito, una maggiore sicurezza in termini di minore rischio di fallimento.

Tuttavia, una simile richiesta pregiudica la possibilità di comprendere quale sia il livello di solvibilità e l'effettiva esposizione ai rischi di una compagnia, per chiunque ne legga semplicemente il bilancio, che sia un potenziale investitore o che sia un assicurato e rischia di complicare anche il compito dell'autorità di vigilanza, limitandone l'efficacia degli interventi.

¹Le gli importi limite delle fasce vengono poi modificati con provvedimenti *una tantum* che tengono conto dell'inflazione, si veda ad esempio [20]

Alle compagnie assicurative, quindi, viene chiesto di redigere il proprio bilancio secondo criteri opposti a quelli previsti dai criteri contabili internazionali (IAS-IFRS), che propongono invece il criterio del *fair value* come fondamentale metodo di valutazione di tutte le attività e passività delle imprese.

È in questo contesto che prende piede il progetto di Solvency 2, che vedremo nel dettaglio nelle prossime sezioni. Esso si propone, almeno per quanto riguarda la stima dei rischi quantificabili cui sono soggette le imprese, innanzitutto come un riordino ed una richiesta di esplicitare quali siano le poste di bilancio che rispecchiano le reali aspettative di ciò che accadrà e quali invece le poste che implicitamente coprono il rischio che ciò che si realizzerà si discosti dalle aspettative.

2.3 Il progetto Solvency 2

Il progetto Solvency 2 si articola in tre pilastri fondamentali che si possono trovare negli articoli della Direttiva 2009/138/CE:

- **Il Primo Pilastro: i Requisiti Quantitativi** contiene le indicazioni quantitative ed i principi per la valutazione delle attività e le passività. In particolare a questo pilastro fa capo anche la scelta di identificare il capitale richiesto in funzione dell'effettiva rischiosità dell'impresa e dovrebbe essere tale da permettere all'impresa di tenere sotto controllo il rischio legato allo svolgimento dell'attività assicurativa.
- **Il Secondo Pilastro: i Requisiti Qualitativi** contiene i requisiti in termini di *governance* delle compagnie, di controllo interno dei rischi e le procedure di vigilanza. Con queste norme si cerca di controllare i rischi di tipo non quantificabile.
- **Il Terzo Pilastro: la Disciplina di Mercato** contenente le norme che definiscono quali informazioni circa la rischiosità e la situazione finanziaria debbano essere rese pubbliche al fine di permettere una più libera concorrenza.

Non ci dilungheremo oltre sul Secondo e sul Terzo Pilastro previsti dalla normativa, ma ci concentreremo sul Primo, proprio a causa delle scelte operate in merito al capitale richiesto.

La direttiva entra nel merito delle valutazioni di attività e passività dell'impresa² ed in particolare tratta del calcolo del Requisito Patrimoniale di Solvibilità, SCR, negli articoli da 100 a 127.

In particolare, il terzo comma dell'articolo 101 della direttiva recita:

Il requisito patrimoniale di solvibilità è calibrato in modo da garantire che siano presi in considerazione tutti i rischi quantificabili cui è esposta un'impresa di assicurazione o di riassicurazione. Esso copre l'attività esistente nonché le nuove attività che si prevede vengano iscritte nel corso dei dodici mesi successivi. Per quanto riguarda l'attività esistente, esso copre esclusivamente le perdite inattese.

Il requisito patrimoniale di solvibilità corrisponde al valore a rischio dei fondi propri di base dell'impresa di assicurazione o di riassicurazione soggetto ad un livello di confidenza del 99.5% su un periodo di un anno.

In altre parole l'SCR, così come previsto dalla direttiva, altro non è che una misura del rischio a cui sono soggette le compagnie di assicurazione e vengono anche identificati i tre fattori fondamentali a cui fa riferimento:

L'orizzonte temporale, scelto pari ad un anno;

Il metodo di calcolo, identificato come il *Value-at-Risk*, VaR;

Il livello di confidenza, definito come il 99.5%.

Si rende necessaria qualche considerazione nel merito di queste scelte. Con riferimento all'orizzonte temporale, la scelta di un solo anno, coniugata ad una probabilità

²È questa direttiva, ad esempio, che impone il calcolo delle riserve tecniche come somma delle due componenti *best estimate* e *risk margin*.

abbastanza bassa di fallimento, appare un buon compromesso. Tuttavia, se una compagnia assicurativa svolgesse la propria attività con lungimiranza, dovrebbe anche condurre delle analisi del proprio rischio su di un orizzonte pluriennale.

Ipotizziamo che una compagnia si trovi a dover liquidare più sinistri di quanti ne attendeva o che il costo sinistri si alzi rispetto alle aspettative. Per il primo anno, a meno di casi eccezionali, il SCR sarà sufficiente a non dichiarare il fallimento. Potrebbe anche essere sufficiente per gli anni successivi, ma, erodendosi, presto non sarebbe più sufficiente ad operare nel rispetto delle prescrizioni. Non vi è infatti alcuna garanzia circa una relazione tra il capitale sotto rischio ad un anno e quello pluriennale, soprattutto se il portafoglio di rischi della compagnia è particolarmente dinamico.

Una compagnia, per assicurare la continuità della propria attività, dovrebbe quindi detenere un capitale sufficiente a far fronte ai propri impegni oltre il singolo anno. Ovviamente si potrebbe richiedere un orizzonte temporale illimitato ed un livello di confidenza pari ad 1, ma questo porterebbe al risultato privo di ragionevolezza di un SCR infinito.

Tra la soluzione di non fare alcuna richiesta circa la valutazione del rischio pluriennale e la richiesta di un SCR infinito si può certamente trovare un buon compromesso. Una possibilità sarebbe quella di definire, con un livello di confidenza minore, un SCR con un orizzonte di alcuni anni e scegliere come SCR finale il maggiore tra i due. Orizzonte temporale e livello di confidenza di questo nuovo SCR andrebbero poi identificati in modo che la valutazione possa essere condotta correttamente ed in modo che la richiesta di capitale non sia eccessiva.

Alternativamente, questo nuovo SCR pluriennale potrebbe essere inteso come un meccanismo di tipo *early warning*, tale per cui, se la compagnia si trova in una situazione in cui non viene rispettato, debbano essere prese delle decisioni per tornare al di sopra della soglia, ma non stringenti quanto quelle che vengono imposte nel caso del non rispetto del requisito su base annuale. Tuttavia il legislatore non ha preso una decisione di questo tipo.

Vale poi sicuramente la pena spendere qualche parola a proposito dell'uso del VaR come misura di rischio, dandone innanzitutto una definizione. Indicando con \tilde{U} la riserva di rischio dell'impresa al termine dell'anno di esercizio e con u_α il percentile di ordine α , il VaR con livello di confidenza $1 - \alpha$ è definito come

$$\text{VaR}_{1-\alpha} = \mathbb{E}(\tilde{U}) - u_\alpha.$$

Ricordando, poi, dai noti modelli semplificati della Teoria del Rischio³ che la variabile aleatoria \tilde{U} , rimanendo sempre su un orizzonte annuale, è la variabile aleatoria costo aggregato dei sinistri \tilde{X} con il segno invertito e traslata orizzontalmente, vale

$$\text{VaR}_{1-\alpha} = x_{1-\alpha} - \mathbb{E}(\tilde{X})^4.$$

In [4] ritroviamo anche un'interpretazione del concetto di misura di rischio. Questo *paper* è scritto con un'ottica volta più ai rischi finanziari che non a quelli assicurativi, per cui il rischio viene definito come il “possibile valore futuro delle posizioni attualmente aperte”. Tuttavia ritroviamo alcuni aspetti che è opportuno considerare anche nel nostro campo. La misura di rischio $\rho(X)$, relativa al rischio X viene interpretata come:

When positive, the number $\rho(X)$ assigned by the measure ρ to the risk X will be interpreted as the minimum extra cash the agent has to add to the risky position X [...] to be allowed to proceed with his plans.

Diverse sono le proprietà che vengono richieste per le misure di rischio, che vengono definite assiomi di coerenza. Indicando con $\rho(\cdot)$ la misura di rischio, X ed Y due rischi ed α un numero reale, devono infatti valere:

Invarianza per traslazione, che richiede $\rho(X + \alpha) = \rho(X) - \alpha$;

Subadditività, per cui $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$;

³Al lordo della riassicurazione e se le spese vengono considerate deterministiche e pari ai caricamenti.

⁴Si veda, ad esempio, [35].

Omogeneità positiva, $\rho(\alpha \cdot X) = \alpha \cdot \rho(X)$;

Monotonicità, intendendo che se $X \leq Y$, $\rho(X) \geq \rho(Y)$;

Rilevanza, poiché se $X < 0$, $\rho(X) > 0$.

Il VaR rispetta questi assiomi, ma possiamo fare qualche considerazione in merito alla Monotonicità

Anche se il sistema Solvency 2 si è orientato verso una misura del rischio di tipo VaR, riteniamo opportuno notare che in altri contesti si è preferito l'uso di un altro indicatore. Ad esempio, nello Swiss Solvency Test e nel Risk Based Capital⁵ è stato preferito il Tail-VaR. Il Tail-VaR è definito come

$$\text{Tail-VaR}_{1-\beta} = \mathbb{E}(\tilde{U}) - \mathbb{E}(\tilde{U} | \tilde{U} < u_\beta) = \mathbb{E}(\tilde{X} | \tilde{X} > x_{1-\beta}) - \mathbb{E}(\tilde{X})$$

Esso, ragionando in termini di \tilde{X} , rappresenta quindi il valore atteso delle perdite considerate *tail*, cioè che superano il valore considerato come il percentile di ordine $1 - \beta$, a cui viene sottratto il valore atteso della variabile.

Ovviamente, se $1 - \beta$ viene scelto pari ad $1 - \alpha$, si ottiene certamente $\text{Tail-VaR}_{1-\beta} \gg \text{VaR}_{1-\alpha}$, ma, ad esempio, lo SST ha adottato il 99% come livello di confidenza.

Tuttavia possiamo notare due peculiarità di questi indicatori. Innanzitutto, considerando la variabile aleatoria $\tilde{\Xi}$, definita come $\{\Xi\} = \{X | X > x_{99\%}\}$, il $\text{Tail-VaR}_{99\%}$ ne rappresenta il valore atteso, mentre il $\text{VaR}_{99.5\%}$ la mediana, tenendo comunque presente che, ad entrambi, viene poi sottratto $\mathbb{E}(\tilde{X})$. Quindi, come noto dalla statistica descrittiva, considerando l'indice di asimmetria della variabile aleatoria Ξ , γ_Ξ ,

$$\gamma_\Xi > 0 \quad \iff \quad \text{Tail-VaR}_{99\%} > \text{VaR}_{99.5\%}.$$

⁵Le regolamentazioni rispettivamente in vigore in Svizzera e negli Stati Uniti

Pur non essendoci una certezza matematica in tal senso, notiamo che i casi in cui $\gamma_{\Xi} < 0$ siano da considerarsi eccezionali⁶, quindi generalmente il requisito di capitale calcolato con il Tail-VaR sarà maggiore di quello ottenuto mediante il VaR.

La seconda considerazione che facciamo è strettamente legata alla precedente. Sempre dalla Statistica Descrittiva sappiamo che, a differenza della media, la mediana è una statistica *robusta* rispetto ai valori estremi. Contestualizzando questa affermazione nel nostro ambito, possiamo affermare che il $\text{VaR}_{99.5\%}$ non dipende dalla forma che la distribuzione di \tilde{X} assume dopo il percentile di ordine 99.5%. Quindi, pensando a due imprese che abbiano distribuzioni del costo sinistri aggregato molto simili fino al percentile di ordine considerato, ma che oltre tale limite differiscano sensibilmente, avendo la prima una coda che si annulla molto più velocemente della seconda, i VaR delle due imprese sarebbero identici. Al contrario, il Tail-VaR della compagnia con la coda più pesante, risulterà maggiore di quello dell'altra compagnia.

In ultima analisi, nell'esempio prospettato, è indubbio che la prima compagnia sia soggetta ad una rischiosità minore di quanto non lo sia la seconda, ma ciò viene colto solo dal Tail-VaR, mentre il VaR le considera identiche. Dalla precedente considerazione possiamo trarre la conclusione che il Tail-VaR sia una misura del rischio più efficiente di quanto non lo sia il VaR, riuscendo a discriminare in base al rischio meglio di quanto non faccia quest'ultima misura. Quindi, riprendendo la definizione di Monotonicità per le misure di rischio, possiamo dichiarare che il Tail-VaR sia una misura più coerente.

In un'ottica di valutazione del rischio di un'impresa assicurativa, è auspicabile quindi prendere in considerazione anche questa misura, cosa di cui terremo conto nei capitoli successivi.

Da ultimo, possiamo fare qualche considerazione in merito alla scelta del livello di confidenza. Accettare lo 0.5% come probabilità che una compagnia fallisca su base annuale significa permettere che essa possa non avere sufficienti risorse a far fronte ai

⁶È possibile pensare che possa accadere una situazione simile solo in presenza di un sistema di limitazioni contrattuali o, se si considera il lavoro al netto della riassicurazione, di un trattato riassicurativo non proporzionale.

propri impegni per una volta ogni 200 anni. Da questo punto di vista, questo non appare problematico. Tuttavia, invertendo i fattori, significa che ogni anno dovrebbe fallire mediamente una compagnia su 200 e ciò è sicuramente più preoccupante.

Non ci aspettiamo che questo capiti veramente, poiché molte compagnie si doteranno probabilmente di un capitale superiore alla richiesta proprio per avere maggiore credibilità, tuttavia questa è la scelta del legislatore. In ogni caso notiamo che il giudizio di *rating* che viene assegnato ad una società con una probabilità di *default* simile è approssimativamente *BBB*.

2.4 La Standard Formula

2.4.1 La Standard Formula secondo la Direttiva

Tornando al testo della direttiva di Solvency 2, sempre alla sezione in cui si va a trattare del SCR, l'articolo 100 introduce alla possibilità di calcolare il requisito in base ad una Formula Standard, meglio specificata negli articoli da 103 a 111 e nell'Allegato IV

Essa viene descritta come una struttura modulare, in cui ogni modulo rappresenta un SCR necessario a coprire uno specifico rischio.

Sempre nella direttiva, nell'Articolo 103, si trova anche l'indicazione che il risultato finale deve essere costituito dalla somma del SCR di base (BSCR), del SCR per il rischio operativo (SCR_{OP}) e degli aggiustamenti per la capacità di assorbimento delle perdite dovute alle imposte differite ed alle riserve tecniche. Il BSCR, poi, deve essere frutto dell'aggregazione di (almeno) cinque sottomoduli, di cui tre legati direttamente al rischio di sottoscrizione, uno dei quali rappresentante il rischio per l'assicurazione salute. Da notare, con riferimento all'articolo 104 il comma 4, recitante

Ciascuno dei moduli di rischio [...] è calibrato utilizzando una misura di rischio del tipo VaR con un livello di confidenza del 99.5% su un periodo di un anno.

Laddove appropriato, gli effetti di diversificazione sono presi in considerazione nella struttura di ciascun modulo di rischio.

ed il comma 7, in cui troviamo

Previa approvazione da parte delle autorità di vigilanza, quando calcolano i moduli del rischio di sottoscrizione per l'assicurazione vita, per l'assicurazione non vita e per l'assicurazione malattia, le imprese di assicurazione e di riassicurazione possono sostituire, nell'ambito della formula standard, un sottoinsieme di parametri con parametri specifici dell'impresa interessata.

Tali parametri sono calibrati sulla base dei dati interni dell'impresa interessata o di dati che sono direttamente rilevanti per le operazioni di tale impresa tramite l'uso di metodi standardizzati.

In sede di concessione dell'approvazione, le autorità di vigilanza verificano la completezza, l'accuratezza e l'adeguatezza dei dati utilizzati.

In ogni singolo modulo devono quindi essere calibrati i parametri secondo i tre principi già espressi in precedenza, approccio che prende il nome di *market wide*, poiché la calibrazione dovrà essere coerente con l'intero mercato europeo. Tuttavia, alle compagnie viene data la possibilità di stimare da sé il proprio SCR utilizzando, pur rimanendo nella stessa struttura del calcolo dell'approccio *market wide*, parametri che risultano coerenti con il proprio portafoglio di rischi. Questo secondo approccio è definito *undertaking specific*.

Ultima scelta operata dal legislatore europeo all'interno della direttiva, che riteniamo opportuno notare, riguarda la scelta nell'aggregazione dei diversi moduli della formula standard.

Sebbene, come stabilito dall'Articolo 103, presentato in precedenza, il SCR finale sia una semplice somma, per ogni modulo e sottomodulo che compone il BSCR è stata scelta l'aggregazione lineare, come mostrato nell'Allegato IV. La formula che viene presentata

con riferimento ai moduli è:

$$\text{SCR} = \sqrt{\sum_{i,j} \text{Corr}_{i,j} \times \text{SCR}_i \times \text{SCR}_j},$$

dove i SCR_i rappresentano i SCR calcolati per i sottomoduli ed i $\text{Corr}_{i,j}$ indicano i coefficienti di correlazione lineare, direttamente ritrovabili nell'Allegato all'interno di una matrice simmetrica solo per quanto riguarda l'aggregazione che restituisce il valore del BSCR.

Anche questa scelta merita qualche considerazione, poiché discende da due assunzioni che vengono fatte implicitamente, necessarie all'adozione di questa formula:

1. La struttura di interdipendenze dei vari rischi è lineare;
2. Le distribuzioni di ogni singolo modulo e sottomodulo sono assunte Gaussiane.

La formula di aggregazione utilizzata è, infatti, la stessa nota in letteratura per calcolare i percentili di una variabile aleatoria frutto della somma di due variabili aleatorie aventi distribuzione Gaussiana, a partire dai percentili di queste ultime due e dalla matrice di correlazione lineare. La scelta di una struttura modulare di questo tipo porta indubbi vantaggi in termini di semplificazione dei calcoli⁷ e permette anche alle compagnie di scegliere di utilizzare l'approccio *undertaking specific*, o addirittura un modello interno, solo per alcuni moduli, mentre per gli altri decidere di mantenere l'approccio *market wide*.

La formula, sfruttando anche le proprietà dei percentili della distribuzione Gaussiana, rispetta anche uno dei requisiti fondamentali delle misure di rischio: la Subadditività, così come presentata in Sezione 2.3. Infatti, se $\text{Corr}_{i,j} < 1$, vale che il SCR di un modulo è minore della somma dei SCR dei suoi sottomoduli.

⁷Altri metodi noti in letteratura attuariale sono la *Common Shock Correlation* e l'uso delle *Copule*, ma entrambi necessitano comunque di ipotesi altrettanto forti se considerati un'ottica *market wide* o di approfonditi studi, troppo onerosi per una compagnia di piccole dimensioni, se si pensa ad un approccio *undertaking specific*.

Tuttavia, come vedremo nel seguito, i singoli moduli non sono costruiti a partire dall'ipotesi della distribuzione Gaussiana, pertanto la scelta di questo metodo per l'aggregazione non appare una scelta ottimale.

Inoltre, si è osservato empiricamente che spesso le interdipendenze tra i rischi non sono lineari e per questo sono stati studiati modelli che prendono in considerazione questo aspetto. Ad esempio è stato possibile includere forme diverse di dipendenza, come ad esempio le copule per la *tail dependency*. Tuttavia studiare un metodo che possa rivelarsi adatto a tutte le compagnie del mercato europeo non sarebbe possibile.

2.4.2 I Quantitative Impact Studies

Come abbiamo rilevato in precedenza, nella direttiva si trovano le indicazioni per quanto riguarda la calibrazione dei parametri e delle formule che andranno a comporre la Standard Formula, ma non appare il metodo preciso per il calcolo del SCR. Il progetto Solvency 2, infatti, prevede che l'implementazione avvenga tramite l'approccio Lamfalussy e, mentre l'emanazione della direttiva appartiene al primo livello, la calibrazione dei parametri appartiene al secondo e terzo livello.

Questo compito è stato quindi demandato al CEIOPS (ora EIOPA) che nel corso degli anni ha avviato pubbliche consultazioni e Studi di Impatto Quantitativo, i *QIS*.

Il CEIOPS, ha condotto cinque studi di impatto quantitativo tra il 2005 ed il 2011. Ogni volta era chiesto alle compagnie dell'Unione Europea di costruire un report che descrivesse la propria situazione se entrasse in vigore la normativa Solvency 2 e se dovessero rispondere ai requisiti di volta in volta ipotizzati dal CEIOPS stesso. Ciascun QIS si concludeva poi con un report di sintesi che evidenziasse i risultati ottenuti dai dati aggregati sull'intero mercato e per ogni singolo stato.

Non intendiamo dilungarci né sulla struttura né sui risultati dei primi quattro QIS, anche perché, per quanto riguarda molti aspetti, essi sono molto eterogenei e difforni da ciò che poi è entrato a far parte della direttiva⁸.

⁸Si pensi, ad esempio, che fino al QIS2 il target non era una misura di rischio di tipo VaR, ma

Ci concentriamo invece su quello che è l'approccio considerato nel QIS5 per la valutazione del rischio assicurativo, in particolare prendendo in considerazione il rischio di sottoscrizione del ramo salute.

2.4.3 La Standard Formula secondo il QIS5

Come nozione preliminare, notiamo che, per ogni modulo previsto dai QIS, il SCR può essere calcolato attraverso un approccio di tipo *scenario based* o di tipo *factor based*. Al primo approccio fanno riferimento ad esempio i sottomoduli del rischio di sottoscrizione per le assicurazioni vita, il cui SCR viene calcolato come la differenza tra il valore delle attività al netto delle passività, il NAV, calcolato secondo le aspettative e lo stesso calcolato nell'ipotesi che si sia verificato uno scenario inatteso, uno *stress*, che abbia fatto ridurre il valore delle voci dell'attivo o aumentare le passività. Tale approccio viene anche definito Δ NAV.

Il secondo approccio, invece, prevede di ottenere il SCR, a partire da delle formule chiuse che avranno come input dei parametri che, come accennato in precedenza, potranno essere calibrati direttamente dal CEIOPS o essere calcolati dalla singola impresa⁹. L'approccio *factor based* viene adottato, ad esempio, nel sottomodulo *Premium and Reserve* del rischio di sottoscrizione dell'assicurazione *Non-Life*.

Il modulo di nostro interesse, il modulo *Health*, è frutto dell'aggregazione di tre sottomoduli:

- Il primo ricomprende la sottoscrizione di rischi gestiti dalla compagnia con tecniche simili a quelle adottate dalle compagnie di assicurazione sulla vita, *SLT*;
- Il secondo attiene al rischio derivante dalla sottoscrizione delle polizze per le compagnie che utilizzano tecniche simili a quelle utilizzate dalle compagnie danni. *Non-SLT*;

Tail-VaR.

⁹si tratta degli approcci *market wide* e *undertaking specific* di cui si diceva nella Sezione 2.4.1.

- Il terzo fa riferimento ai rischi di natura catastrofale, di cui non ci occuperemo nel seguito.

Si noti che la distinzione tra il primo e il secondo sottomodulo ripercorre esattamente la distinzione esistente all'interno del ramo malattia come presentata in Sezione 1.1.

Il primo sottomodulo è poi costituito da altri sei sotto-sottomoduli, ciascuno dei quali è valutato mediante approccio scenario based e ripercorre gli stessi rischi considerati nel modulo *Life*: mortalità, longevità, disabilità e morbilità, riscatto, spese e revisione dei contratti.

Il secondo, invece, al pari del modulo *Non-Life*, risulta dall'aggregazione del modulo *Premium and Reserve* e del modulo *Lapse*. Mentre il sottomodulo *Lapse* fa riferimento ad un approccio scenario based, il *Premium and Reserve* è di tipo factor based.

I dati in nostro possesso, che ci permetteranno un'applicazione della standard formula, come riportato in Sezione 1.2, riguardano una compagnia che opera con tecniche proprie del ramo danni e non disponiamo dei dati necessari a valutare l'incidenza del rischio di riscatto, quindi la nostra analisi si concentrerà sul solo modulo *Premium and Reserve*, di cui vedremo nel seguito sia l'approccio market wide, sia gli undertaking specific.

2.4.3.1 L'approccio Market Wide

In [11] troviamo tutto ciò che ci serve per calcolare il requisito di capitale per il modulo Premium and Reserve.

In ultima analisi, il SCR per questo modulo viene definito come

$$\text{Health}_{\text{Premium\&Reserve}}^{\text{NonSLT}} = \rho(\sigma_{\text{NonSLT Health}}) \cdot V_{\text{NonSLT Health}}. \quad (2.5)$$

Diviene quindi necessario definire la grandezza $V_{\text{NonSLT Health}}$ che rappresenta un'indicazione del volume delle obbligazioni che la compagnia assume nelle assicurazioni sa-

lute e il parametro $\sigma_{\text{NonSLT Health}}$ che indicherà la rischiosità relativa del portafoglio, a cui dovrà essere applicata la funzione $\rho(x)$, di cui si dirà in seguito.

$\sigma_{\text{NonSLT Health}}$ Iniziamo con la valutazione della variabilità relativa.

Occorre innanzitutto notare, come accennato nella Sezione 1.1, che il QIS5 definisce diverse *Lines of Business* all'interno di quello che comunemente in Italia è definito ramo malattia. Sono infatti definite quattro differenti LoBs, presentate in Tabella 2.1 e a ciascuna viene associato un parametro σ per il rischio di tariffazione ed uno per il rischio di riservazione. La LoB *medical expense* ricomprende alcune garanzie offerte dalle polizze

LoB	Premium (σ_{pre})	Reserve (σ_{res})
Medical expense	$4\% \cdot NP_{lob}$	10%
Income protection	$8.5\% \cdot NP_{lob}$	14%
Workers' compensation	$5.5\% \cdot NP_{lob}$	11%
Non-proportional health reinsurance	17%	20%

Tabella 2.1: Parametri σ forniti dal QIS5 per l'approccio market wide

come può essere il rimborso spese mediche. In generale ad essa fanno riferimento tutte le prestazioni che la compagnia offre come rimborso per le spese sostenute dagli assicurati per le cure mediche o per gli esami diagnostici, basandosi sulle fatture presentate o in modo forfettario, a seconda della tipologia prescelta.

Per *income protection* si intendono le garanzie offerte agli assicurati come le diarie da ricovero o le diarie giornaliere che garantiscono un reddito a coloro che non sono in grado di lavorare a causa di malattia o di infortunio.

La *workers' compensation* ricomprende le garanzie offerte dalle prime due LoBs se la causa del sinistro è legata alla professione dall'assicurato. Sono quindi ricomprese le prestazioni conseguenti agli infortuni sul luogo di lavoro e alle malattie professionali. Questo ruolo, in Italia, è svolto dall'INAIL, quindi non lo considereremo nel seguito.

Infine, per *non-proportional (health) reinsurance*, si intendono i contratti assunti come riassicurazione la cui controparte è una compagnia che sottoscrive contratti appartenenti alle tre LoBs precedenti. Si noti che vengono considerati solo i contratti non proporzionali, poiché quelli invece proporzionali vengono assimilati direttamente alle tre LoBs precedenti, essendo soggetti alla medesima rischiosità relativa. In ogni caso la nostra compagnia non si occupa di polizze di questo tipo, quindi anche questa non sarà considerata nel seguito.

Notiamo anche che, con riferimento al solo rischio di tariffazione, le σ_{prem} sono poi moltiplicate per il fattore NP_{lob} . Esso riduce il fattore σ_{prem} tenendo conto della possibilità che la compagnia abbia ceduto parte del rischio attraverso dei trattati di riassicurazione non proporzionale, riducendo la rischiosità relativa del portafoglio. La nostra compagnia, tuttavia, non ha in essere alcun trattato di questo tipo, quindi non ci curiamo di questo fattore.

Il primo passo per ottenere $\sigma_{NonSLT\ Health}$ è quello di ricavare, all'interno di ogni LoB il fattori di volatilità $\sigma_{(lob)}$, aggregando i fattori del rischio di tariffazione e del rischio di riservazione, adottando una ponderazione che vedremo in seguito e una correlazione di 0.5. La formula sarà:

$$\sigma_{(lob)} = \frac{\sqrt{(\sigma_{(prem)}V_{(prem)})^2 + \sigma_{(prem)}\sigma_{(res)}V_{(prem)}V_{(res)} + (\sigma_{(res)}V_{(res)})^2}}{V_{(prem)} + V_{(res)}}. \quad (2.6)$$

I fattori $V_{(prem)}$ e $V_{(res)}$ adottati come ponderazione corrispondono poi agli indicatori della dimensione degli obblighi cui è sottoposta la compagnia, identificati rispettivamente con il volume dei premi attesi per l'anno successivo e con il volume della riserva sinistri. Entreremo maggiormente nel dettaglio di questi indicatori nel seguito.

Una volta identificate le $\sigma_{(lob)}$, attraverso la formula:

$$\sigma_{NonSLT\ Health} = \frac{\sqrt{\sum_{r \times c} 0.5 \cdot \sigma_r \cdot \sigma_c \cdot V_r \cdot V_c}}{\sum_r V_r}, \quad (2.7)$$

si ottiene finalmente $\sigma_{\text{NonSLT Health}}$. In essa r e c indicano le diverse LoBs, le correlazioni calibrate dal CEIOPS sono tutte pari a 0.5 e nuovamente è stata adottata una ponderazione, in questo caso attraverso le misure di volume V_{lob} , che definiremo in seguito.

$V_{\text{NonSLT Health}}$ Veniamo ora ai calcoli necessari per ottenere i fattori di ponderazione e per $V_{\text{NonSLT Health}}$.

Innanzitutto, $V_{(res)}$ viene definita come la misura di volume per il rischio di riserva, considerata per ogni LoB come l'ammontare sola porzione di *best estimate* della riserva sinistri. Questo calcolo dovrebbe essere fatto con gli importi al netto della riassicurazione, essendo poi previsto un differente modulo dedicato alla valutazione del rischio di controparte.

Invece $V_{(prem)}$, la misura di volume per il rischio di tariffazione viene ricavata dalla formula:

$$V_{(prem)} = \max \left(P_{lob}^{t,written}, P_{lob}^{t,earned}, P_{lob}^{t-1,written} \right) + P_{lob}^{PP}. \quad (2.8)$$

In essa, ricordando che anche in questo caso i calcoli devono essere svolti al netto della riassicurazione,

$P_{lob}^{t,written}$ rappresenta la stima dei premi che verranno contabilizzati nel corso del prossimo anno;

$P_{lob}^{t-1,written}$ indica i premi contabilizzati nel corso dell'anno appena trascorso;

$P_{lob}^{t,earned}$ rappresenta la stima dei premi di competenza per il prossimo anno;

P_{lob}^{PP} considera il valore attuale dei premi non ancora contabilizzati per i contratti che prevedono coperture che vanno oltre l'anno successivo.

Una volta ottenuti $V_{(prem)}$ e $V_{(res)}$, per ogni LoB si può procedere al calcolo di V_{lob} tenendo conto della diversificazione geografica. La formula adottata è:

$$V_{lob} = (V_{(prem)} + V_{(res)}) \cdot (0.75 + 0.25 \cdot DIV_{lob}). \quad (2.9)$$

DIV_{lob} è detto indice di diversificazione geografica e viene ricavato¹⁰ da

$$DIV_{lob} = \frac{\sum_j (V_{(prem,j)} + V_{(res,j)})^2}{\left(\sum_j (V_{(prem,j)} + V_{(res,j)})\right)^2}, \quad (2.10)$$

dove il pedice j indica la macro regione geografica¹¹ a cui fanno riferimento i volumi¹².

Infine, semplicemente sommando tutti i V_{lob} si ottiene $V_{NonSLT\ Health}$.

$\rho(\cdot)$ Infine, l'ultima parte della Formula 2.5 che ci resta da definire è la funzione $\rho(\cdot)$.

Essa è definita come:

$$\rho(x) = \frac{\exp\left(N_{0.995} \cdot \sqrt{\ln(x^2 + 1)}\right)}{\sqrt{x^2 + 1}} - 1, \quad (2.11)$$

in cui $N_{0.995}$ è il percentile di ordine 99.5% di una distribuzione Normale Standard.

La funzione $\rho(\cdot)$ è costruita in modo che, calcolata in σ , restituisce il valore del $VaR_{99.5\%}$ di una distribuzione Log-Normale, avente coefficiente di variabilità pari a σ e valore atteso unitario.

2.4.3.2 Il SCR per la nostra compagnia secondo l'approccio

Market Wide

Innanzitutto, come riportato nella Sezione 1.2, occorre ricordare che non abbiamo a disposizione né i dati per quanto riguarda la riassicurazione, né un database significativo

¹⁰Deve comunque essere posto pari ad 1 nel caso in cui si usi un approccio undertaking specific.

¹¹Seguendo segmentazione sviluppata dall'ONU, ne sono state individuate 18.

¹²L'intero nostro portafoglio è riferito a soli rischi sottoscritti nel mercato italiano, quindi in seguito non ci cureremo di questo fattore.

per quanto riguarda il rischio di riservazione. Inoltre, disponiamo dei dati da cui si può ricavare la suddivisione per LoB disposta dal QIS5 solo per quanto riguarda il 2008.

Ci potremo curare, quindi, del solo rischio di tariffazione, perciò non procederemo all'aggregazione dei fattori di rischiosità all'interno di ogni LoB, ma considereremo direttamente i coefficienti della prima colonna della Tabella 2.1 come fossero $\sigma_{(lob)}$, non utilizzando quindi la Formula 2.6.

In secondo luogo, le nostre valutazioni dovranno essere condotte al lordo della riassicurazione. Tuttavia, considerando che il trattato in essere è di tipo *quota share*, notiamo come questa distorsione non pregiudichi l'affidabilità della valutazione, se non per un fattore dimensionale. Questo perché tutti i NP_{lob} saranno comunque pari ad 1, ma non potremo valutare i premi al netto della riassicurazione.

Infine, rispetto alla suddivisione per LoB, abbiamo comunque certezza che la compagnia non operi nel *workers' compensation*, né che assuma rischi in riassicurazione. Condurremo quindi la valutazione nell'ipotesi che abbia un portafoglio composto esclusivamente da rischi appartenenti alla LoB *medical expense* e successivamente *income protection*. Da ultimo, basandoci sulla suddivisione del portafoglio osservata per l'anno 2008, ricaveremo una nuova valutazione del rischio.

Occorre quindi valutare la dimensione del portafoglio. Come detto in precedenza, esso è costituito unicamente da polizze sottoscritte per clienti di tipo business, prevedendo coperture standard dal primo gennaio al 31 dicembre. Questo costituisce una semplificazione non indifferente dei calcoli, poiché $P_{lob}^{t,earned}$ e P_{lob}^{PP} della Formula 2.8 potranno essere trascurati.

Tuttavia occorre concentrarci sulla stima di $P_{lob}^{t,written}$. Come si può notare dalla Figura 2.1, i premi mostrano una indubbia crescita nel corso degli anni, con un andamento che appare a prima vista esponenziale. È quindi possibile condurre una stima di $P_{lob}^{t,written}$ attraverso una regressione esponenziale, ottenendo € 26,076,525.90.

Questo valore, accettabile come prima stima, non può tuttavia essere considerato pienamente soddisfacente. Infatti, pur osservando una crescita quasi esponenziale nel

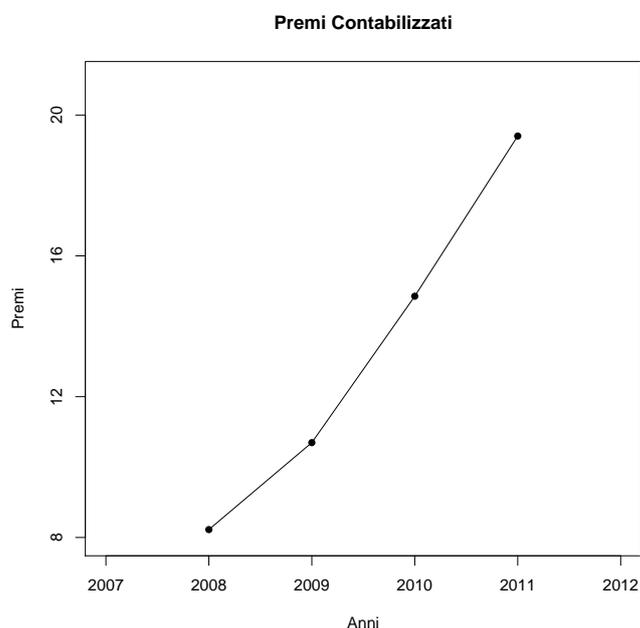


Figura 2.1: Premi lordi contabilizzati, dati in € Mln

valore dei premi emessi, se si va ad osservare il numero delle polizze per ogni anno, si osserva una curva assolutamente concava.

Andando quindi a valutare il premio medio per ogni singola polizza, notiamo un comportamento molto variabile: nel 2008 il premio medio era elevato, mentre si è sensibilmente ridotto di circa € 20,000 nel 2009 e nel 2010, per poi ritornare ai livelli precedenti nel 2012. Sorge quindi il dubbio che la compagnia abbia stariffato nel 2009 e nel 2010 per aumentare il numero di polizze e, osservando anche la serie storica dei *combined ratio* dell'intero portafoglio della compagnia riferito al solo ramo malattia¹³, otteniamo una ulteriore conferma di questo.

Per ottenere quindi un secondo possibile valore di $P_{lob}^{t,written}$ decidiamo di produrre una stima del costo medio per l'anno 2012 e moltiplicarla per il numero di polizze attese. Conduciamo quindi la stima attraverso i minimi quadrati ponderati, utilizzando il numero di polizze come pesi, e otteniamo la nuova stima del volume dei premi, pari a

¹³Dati Infobila 2012.

Ipotesi di LoB	Premi		
	€ 26,076,525.90	€ 18,701,390.10	€ 22,976,233.45
Medical expense	€ 2,805,751.05	€ 2,012,209.95	€ 2,472,169.47
Income protection	€ 6,253,214.67	€ 4,484,639.07	€ 5,509,756.96
Suddivisione 2008	€ 3,729,975.51	€ 2,675,039.13	€ 3,286,510.96

Tabella 2.2: SCR stimati con l'approccio market wide

€ 18,701,390.10.

Infine, per completezza, conduciamo anche una stima di un modello lineare ai minimi quadrati ponderati, sempre con i pesi descritti in precedenza, direttamente sui premi emessi nei quattro anni osservati. Tale modello restituisce come stima € 22,976,233.45.

Il prosieguo del metodo non ci permette ulteriori possibilità di *judgment*, quindi riportiamo in Tabella 2.2 le valutazioni del SCR con l'approccio market wide. Come descritto in precedenza conduciamo le valutazioni sotto tre ipotesi di suddivisione del portafoglio.

2.4.3.3 L'approccio Undertaking Specific

In [11] sono descritti tre metodi affinché le compagnie possano procedere ad una valutazione del proprio rischio più aderente al portafoglio effettivamente detenuto.

Ciascun metodo procede a calcolare delle $\sigma_{(prem)}$ e $\sigma_{(res)}$ per ogni LoB a partire dai dati osservati negli anni precedenti che poi saranno aggregate per ottenere una $\sigma_{\text{NonSLT Health}}$ da utilizzare, attraverso l'usuale $\rho(\sigma)$ per calcolare il SCR.

Ogni stima viene anche calibrata in modo che divenga una media ponderata tra i fattori market wide e undertaking sepecific. Ad esempio, considerando la stima di $\sigma_{(prem)}$, essa sarà una media ponderata tra un fattore $\sigma_{(prem,USP)}$ e il fattore dell'approccio market wide $\sigma_{(prem,M)}$. La formula adottata da [11] è:

$$\sigma_{(prem)} = c \cdot \sigma_{(prem,USP)} + (1 - c) \cdot \sigma_{(prem,M)},$$

Anni	5	6	7	8	9	≥ 10
c	34%	51%	67%	81%	92%	100%

Tabella 2.3: Valori assunti da c per le LoB del modulo Health

dove il fattore c varia in funzione della lunghezza della serie storica adottata come in Tabella 2.3, essendo la serie della durata di almeno cinque anni. Una simile correzione è messa in atto al fine di prendere in considerazione la credibilità della serie storica adottata, pur considerandola solo in termini di lunghezza temporale. Una serie storica più lunga viene considerata maggiormente rappresentativa del portafoglio e meno soggetta a variazioni della rischiosità contingenti.

Veniamo quindi ai tre metodi per stimare la rischiosità. Come in precedenza, ci occuperemo solo del fattore $\sigma_{(prem)}$, non avendo la possibilità di fare confronti coerenti per il reserve risk.

Il primo metodo Undertaking Specific Il primo metodo, il più semplice dei tre, punta a stimare $\sigma_{(prem,USP)}$ sotto l'ipotesi che la distribuzione delle perdite nell'anno y , \tilde{U}_Y sia

$$\tilde{U}_y \sim V_y \mu + \sqrt{V_y} \beta \tilde{\varepsilon}_y. \quad (2.12)$$

\tilde{U}_Y viene considerato quindi pari alla somma di una componente non aleatoria corrispondente al valore atteso delle perdite, μ indica infatti il *loss ratio* medio osservato e V_y i premi di competenza per l'anno y , e di una componente aleatoria. $\tilde{\varepsilon}_y$ viene assunta come una variabile aleatoria di tipo *white noise*, avente valore atteso nullo e deviazione standard unitaria, quindi si assume che la varianza di \tilde{U}_Y sia pari a $V_y \beta^2$. L'obiettivo diviene quindi la stima di β e il CEIOPS ne dà anche una formula chiusa, ottenuta con il metodo dei minimi quadrati:

$$\hat{\beta} = \sqrt{\frac{1}{N-1} \sum_y \frac{(U_y - V_y \hat{\mu})^2}{V_y}}, \quad (2.13)$$

dove $\hat{\mu} = \frac{\sum_y U_y}{\sum_y V_y}$.

Una volta ottenuto $\hat{\beta}$, è possibile stimare $\sigma_{(prem,USP)}$ attraverso la relazione:

$$\sigma_{(prem,USP)} = \frac{\hat{\beta}}{\sqrt{V_{(prem)}}}, \quad (2.14)$$

essendo $V_{(prem)}$ lo stesso descritto in Sezione 2.4.3.1.

Il secondo metodo Undertaking Specific Il secondo approccio, un po' più complicato del precedente, si contraddistingue dagli altri in particolare per due ipotesi.

L'ipotesi di base è sempre la Formula 2.12, ma viene corredata dalle assunzioni che la distribuzione di \tilde{U}_y sia una Log-Normale di parametri $S_y = \sqrt{\ln\left(1 + \frac{\beta^2}{V_y \mu^2}\right)}$ e $M_y = \ln(V_y \mu) - \frac{1}{2}S^2$ e che si possa utilizzare, per la stima di β e μ , il metodo della massima verosimiglianza.

Al posto della verosimiglianza, viene poi richiesto l'uso della Log-Verosimiglianza, opportunamente semplificata, secondo la Formula:

$$\ln L = \sum_y \left(-\ln(S_y) - \frac{(\ln(U_y) - M_y)^2}{2S_y^2} \right). \quad (2.15)$$

Massimizzandola rispetto a μ e β , si ottengono quindi le stime $\hat{\mu}$ e $\hat{\beta}$ e, mediante la Formula 2.14 si ottiene $\sigma_{(prem,USP)}$.

Il terzo metodo Undertaking Specific Il terzo metodo proposto dal CEIOPS è sicuramente il più elaborato e trae dichiaratamente spunto da un metodo adottato in Svizzera nello SST.

La relazione adottata per calcolare $\sigma_{(prem,USP)}$ è:

$$\sigma_{(prem,USP)} = \frac{\sqrt{\mathbb{V}(\tilde{S})}}{V_{(prem)}},$$

in cui \tilde{S} indica la variabile aleatoria Costo Aggregato dei Sinistri. Notevole è l'ipotesi adottata per la distribuzione di \tilde{S} : essa viene considerata come una somma di un numero aleatorio \tilde{N} di sinistri, ciascuno di importo aleatorio \tilde{X}_i . Si considera anche indipendenza tra \tilde{N} e \tilde{X}_i ¹⁴.

Altra assunzione che andrebbe fatta sarebbe la presenza di un fattore di disturbo $\tilde{\Theta} = [\tilde{\Theta}_1; \tilde{\Theta}_2]$, tale che,

$$\tilde{N}|\Theta_1 \sim Pois(\lambda(\Theta_1)) \quad \tilde{X}_i|\Theta_2 \sim F(\mu(\Theta_2); \sigma(\Theta_2)),$$

ma lo stesso documento [11], per semplificare, assume che solo \tilde{N} dipenda da Θ , che $\lambda(\Theta) = \lambda \cdot \Theta$ e che $\mathbb{E}(\Theta) = 1$. Grazie a queste assunzioni, vale

$$\mathbb{V}(S) = \mu^2 \lambda^2 \mathbb{V}(\Theta) + \lambda \mu^2 + \lambda \sigma^2, \quad (2.16)$$

noto risultato del CRM, nel caso in cui la distribuzione di \tilde{N} sia una Poisson Misturata.

Alle compagnie viene quindi richiesta la stima dei dati che vanno a comporre la Formula 2.16, in cui

- μ , valore atteso del costo singolo sinistro, equivale alla somma dei costi aggregati osservati per i diversi anni, divisi per il numero totale dei sinistri osservati.
- σ , la deviazione standard del costo, viene calcolata come la media delle deviazioni standard osservate in ogni anno.
- λ , numero atteso dei sinistri, si ottiene in due passaggi:
 - si divide la somma del numero totale dei sinistri per l'ammontare totale dei premi, corretto per l'inflazione, ottenendo un numero medio di sinistri per ogni Euro di premio

¹⁴Si noti come queste ipotesi corrispondono alle ipotesi del CRM, di cui tratteremo più approfonditamente nella Sezione 3.1.

- successivamente, si moltiplica il valore ottenuto per il $V_{(prem)}$ visto in precedenza, per riportare alle grandezze attese del portafoglio il valore ottenuto.

Viene anche data la possibilità alle compagnie di utilizzare un diverso sistema di ponderazione, se è ritenuto opportuno.

- $\mathbb{V}(\Theta)$, la varianza del fattore di disturbo, viene stimata attraverso una formula chiusa già commentata in [17]. Il CEIOPS, notando che il disturbo $\tilde{\Theta}$ è stato considerato nella Teoria del Rischio come un fattore di rischio sistematico che quindi agisce sull'intero mercato, rileva che sia una contraddizione la possibilità di procedere ad una stima undertaking specific di $\mathbb{V}(\Theta)$ e si riserva di fornirne delle stime market wide.

2.4.3.4 Il SCR per la nostra compagnia secondo l'approccio

Undertaking Specific

Veniamo dunque all'applicazione di quanto esposto nel paragrafo precedente per il portafoglio della nostra compagnia.

Rileviamo innanzitutto che, stando a quanto riportato in Tabella 2.3, poiché i dati a nostra disposizione descrivono un periodo della durata di soli quattro anni, i metodi undertaking specific non dovrebbero essere neppure presi in considerazione dalla compagnia ai fini del calcolo del SCR. Ciò, tuttavia, non nega la possibilità di condurre il calcolo, se non altro al fine di ottenere una stima della rischiosità ad uso interno.

Inoltre, con riferimento ai primi due metodi, occorre tenere presente che vengono fatte delle assunzioni, oltre ai metodi di stima di cui si accennava nei precedenti paragrafi.

Una di esse, comune ad entrambi, riguarda il fatto che sia il costo aggregato dei sinistri atteso, sia la sua variabilità siano proporzionali al volume dei premi. Non abbiamo elementi sufficienti a dimostrare se questa assunzione, necessaria per poter adottare la Formula 2.12, sia o meno forte considerando i nostri dati, quindi possiamo accettarla.

Tuttavia viene adottata un'altra ipotesi che ci fa propendere per un metodo differente dai primi due: il loss ratio atteso deve essere costante, cioè non vi devono essere delle variazioni nel tasso di premio. Come abbiamo rilevato in Sezione 2.4.3.2, questa non è un'assunzione realistica per i nostri dati.

Procedendo al calcolo secondo lo stesso criterio adottato in Sezione 2.4.3.2, utilizziamo nuovamente i dati al lordo della riassicurazione e produciamo in Tabella 2.4 le stime ottenute secondo le tre ipotesi per quanto riguarda il volume dei premi atteso.

Mostriamo differenti stime per il terzo metodo che si originano da due differenti stime del parametro λ . La prima è ottenuta adottando la ponderazione suggerita dal CEIOPS, mentre la seconda utilizzando il numero di polizze nei diversi anni, ma entrambe con la variabilità del fattore di disturbo stimata secondo il criterio descritto in [11].

Per il valore atteso e la deviazione standard della *severity* adottiamo i valori che si ottengono aggregando i costi osservati nel corso degli anni. Per questo è stata adottata l'ipotesi che il portafoglio resti stabile nell'arco dei diversi anni di bilancio. Come già abbiamo notato in precedenza questa costituisce un'ipotesi forte per l'analisi in cui ci stiamo cimentando.

Metodo	Premi		
	€ 26,076,525.90	€ 18,701,390.10	€ 22,976,233.45
Primo	€ 15,042,534.11	€ 13,151,890.28	€ 14,285,312.91
Secondo	€ 17,724,700.71	€ 15,554,436.51	€ 16,855,702.63
Terzo A	€ 19,467,609.59	€ 14,050,961.48	€ 17,190,650.55
Terzo B	€ 12,494,524.03	€ 13,267,420.54	€ 12,758,461.27

Tabella 2.4: SCR stimati con l'approccio undertaking specific

Rispetto all'ultimo punto della sezione precedente, proviamo anche ad identificare un differente valore della varianza del fattore di disturbo a partire dalle statistiche ANIA.

Dalla Teoria del Rischio sappiamo che, indicando con P i premi equi, con B i premi di tariffa, con X il costo aggregato dei sinistri, con c_z il CoV dei sinistri e con n il numero atteso dei sinistri vale, sotto le usuali ipotesi del CRM,

$$\sigma\left(\frac{X}{P}\right) = \sqrt{\frac{1 + c_z^2}{n} + \mathbb{V}(\Theta)}.$$

Metodo	Premi		
	€ 26,076,525.90	€ 18,701,390.10	€ 22,976,233.45
Terzo A	€ 7,014,054.45	€ 5,200,355.11	€ 6,252,210.96
Terzo B	€ 4,824,087.50	€ 4,959,180.75	€ 4,870,141.40

Tabella 2.5: SCR stimati con il terzo approccio undertaking specific, con i dati ANIA per la stima del fattore di disturbo

Quindi,

$$\lim_{n \rightarrow +\infty} \sigma\left(\frac{X}{P}\right) = \sqrt{\mathbb{V}(\Theta)}.$$

L'Appendice Statistica dell'ANIA ci permette di stimare $\sigma\left(\frac{X}{B}\right)$ ma, essendo $\frac{P}{B} = \frac{1-c}{1+\lambda}$, possiamo approssimare

$$\sqrt{\mathbb{V}(\Theta)} \simeq \sigma\left(\frac{X}{B}\right) \frac{1+\lambda}{1-c},$$

avendo supposto che i numeri ricompresi nelle statistiche ANIA siano abbastanza grandi da permettere l'uso del precedente limite.

Il valore di $\sigma(\Theta)$ calcolato secondo il metodo canonico è 0.2178, mentre, partendo dalle statistiche ANIA, è notevolmente inferiore e pari a 0.0319. Mediante questo nuovo valore possiamo stimare nuovamente il SCR come in Tabella 2.5 .

2.5 Qualche considerazione

Qualche commento in merito ai metodi adottati ed ai risultati ottenuti si rende quantomeno necessario.

Punto di forza dei metodi presentati è sicuramente la semplicità di calcolo. Per produrre le stime del SCR è stato infatti sufficiente, per la maggior parte dei calcoli¹⁵, l'uso di un semplice foglio Excel, dal momento che si tratta quasi interamente di formule chiuse. La Standard Formula ha quindi centrato l'obiettivo che la direttiva si poneva all'interno del Punto (19) delle considerazioni preliminari:

La presente direttiva non dovrebbe essere troppo onerosa per le imprese di assicurazione di piccole e medie dimensioni. Uno degli strumenti per conseguire tale obiettivo è la corretta applicazione del principio di proporzionalità. Tale principio dovrebbe valere sia per i requisiti imposti alle imprese di assicurazione e di riassicurazione, sia per l'esercizio dei poteri di vigilanza.

Tuttavia non possiamo non rilevare alcuni aspetti cruciali degli approcci proposti. Innanzitutto ci preme sottolineare che la funzione $\rho(\cdot)$ presuppone che la distribuzione delle perdite sia una Log-Normale, come già detto in Sezione 2.4.3.1. Questa è un'ipotesi forte, superata nella comune prassi attuariale, in cui troviamo l'uso di modelli molto più raffinati. Ciò significa che per tutte le compagnie viene adottata in modo univoco una scelta per quanto riguarda la forma della distribuzione del costo aggregato dei sinistri, senza tenere in considerazione quanto emerge dalle evidenze empiriche. Questo punto in particolare sarà oggetto di discussione nei capitoli successivi. L'ipotesi, inoltre, contrasta con la scelta adottata rispetto all'aggregazione dei moduli e sottomoduli, che invece presupponeva distribuzioni gaussiane.

In secondo luogo notiamo che, stante la formulazione 2.5, il valore atteso del costo sinistri viene considerato pari al volume dei premi. Non è quindi presa in considerazione la possibilità che la compagnia assuma i rischi con un tasso di premio che sia calcolato per ottenere un utile atteso. Questo fattore rappresenta una scelta prudenziale nel caso che abbiamo prospettato, quindi la problematicità riguarda solo il piano degli obiettivi che ci si era posti con il progetto Solvency 2.

¹⁵L'unico calcolo per cui abbiamo adottato R è stato il calcolo dei parametri della *severity* nel terzo metodo.

Tuttavia, se la compagnia si trova in una fase in cui ha preso la decisione di stariffare, ad esempio al fine di guadagnare quote di mercato, l'ammontare dei premi è inferiore al valore atteso del costo aggregato dei sinistri. Dal momento che il $\text{VaR}_{99.5\%}$, calcolato attraverso la formula $\rho(\cdot)$ è pari alla differenza tra il percentile di ordine 99.5% ed il valore atteso, adottarlo come valore del SCR significa sottostimare la rischiosità della compagnia.

Successivamente facciamo presente che uno dei risultati più importanti della Teoria del Rischio, che dà la giustificazione teorica della sostenibilità dell'attività assicurativa, è totalmente trascurato. Infatti, mentre la rischiosità relativa di un portafoglio dovrebbe ridursi al crescere dei volumi fino ad annullarsi¹⁶, ciò non viene preso in considerazione nella Standard Formula. Quindi, poiché i parametri in Tabella 2.1 sono frutto di una calibrazione adatta ad una compagnia di media dimensione, risultano assolutamente inadeguate per il nostro portafoglio. Per tale ragione otteniamo, dalle valutazioni market wide, dei SCR molto inferiori a quelli undertaking specific. Non possiamo quindi pensare che l'approccio market wide sia opportuno sul nostro portafoglio.

Infine, nell'approccio market wide si dichiara che i parametri σ sono stati calibrati per prendere in considerazione anche l'aleatorietà delle spese. Tuttavia, rileviamo che nei metodi undertaking specific, le stime sono fatte solo sui sinistri al netto delle spese. In [11] troviamo una risposta a questa incoerenza: viene richiesta l'assunzione che la volatilità dei sinistri e delle spese sia simile. Questa assunzione non è tuttavia verificabile per la nostra compagnia.

Veniamo anche a qualche commento relativo ai tre metodi undertaking specific.

I primi due metodi sono molto simili. Entrambi discendono dall'assunzione che il costo sinistri aggregato abbia distribuzione Log-Normale, solo che, mentre nel primo essa compare solo per giustificare l'uso della funzione $\rho(\cdot)$, nella seconda viene utilizzata

¹⁶A meno di un fattore di rischio sistematico, che ne diviene il limite inferiore, o a meno che i rischi di cui è composto il portafoglio siano pienamente e positivamente correlati, nel qual caso la variabilità relativa non cambia.

anche per la stima di $\hat{\beta}$, da cui poi si ottiene $\sigma_{(prem,USP)}$. La grave carenza di questi due metodi riguarda il fatto che essi non sfruttano appieno le informazioni contenute nei dati. Sebbene il costo aggregato sia frutto di frequency e severity, ciò non viene considerato al momento della stima della variabilità relativa.

Il terzo metodo adotta una differente ipotesi al fine di ottenere una stima di $\sigma_{(prem,USP)}$. La stima, infatti, prevede che la distribuzione del costo aggregato sia una Poisson Composta Generalizzata. Il metodo appare, quindi, molto più raffinato dei precedenti, tuttavia viene nuovamente adottata la funzione $\rho(\cdot)$, che quindi reintroduce l'ipotesi di Log-Normalità.

Notiamo anche che il metodo proposto per la stima di $\sigma(\Theta)$ sia inadeguato, poiché rimane legato solo alla volatilità del portafoglio, tradendo quello che sarebbe il presupposto originario del fattore di disturbo, cioè di rischiosità sistematica legata al mercato.

Pur essendo certi che non esista un metodo perfetto, che sia coerente per le valutazioni di qualunque compagnia, nei successivi capitoli condurremo delle valutazioni del rischio assicurativo con metodi che possano meglio adattarsi al portafoglio oggetto dei nostri studi.

Capitolo 3

La valutazione del rischio attraverso il CRM

3.1 Introduzione

L'obiettivo di questo capitolo sarà di costruire un modello per la stima del rischio legato al portafoglio della compagnia in esame attraverso il calcolo del requisito di capitale, su un orizzonte annuale. I dati a nostra disposizione contengono le informazioni dettagliate per quanto riguarda il numero di sinistri verificatisi ogni anno, con la distinzione, per ciascuno di essi, della polizza che ne ha dato origine e del costo di ogni singolo sinistro, in termini di importi liquidati e di importi riservati. Poiché la compagnia oggetto del nostro studio ha da poco iniziato la propria attività¹, ci potremo concentrare solo sul *premium risk*, quindi il dato che utilizzeremo rispetto al costo dei singoli sinistri sarà dato dalla somma degli importi liquidati nel corso dell'anno di calendario di accadimento del sinistro e dell'importo riservato alla fine dell'anno. All'opposto, le liquidazioni successive e gli adeguamenti della riserva sinistri, facendo capo al *reserve risk*, non verranno nel seguito considerati.

¹Le prime polizze emesse risalgono all'anno 2008.

Il nostro tentativo sarà quello di stimare il rischio a partire dalla stima della distribuzione della variabile aleatoria costo aggregato dei sinistri per l'anno 2012. Per sfruttare appieno le informazioni contenute nei dati a nostra disposizione, accantoniamo subito la possibilità di descrivere tale variabile con una distribuzione nota in letteratura: dovremo infatti comporre diversi fattori che coinvolgeranno diversi stadi di aleatorietà.

Il modello che ci appare più adeguato è, invece, il Collective Risk Model. In un modello di questo tipo, infatti, il costo sinistri aggregato \tilde{X}_{2012} sarà ricavato attraverso la relazione

$$\tilde{X}_{2012} = \sum_{j=0}^{\tilde{K}_{2012}} \tilde{Z}_{2012,j}, \quad (3.1)$$

e dipenderà quindi dalla scelta che opereremo di un opportuno modello per il numero dei sinistri \tilde{K}_{2012} e di uno per il costo singolo sinistro $\tilde{Z}_{2012,j}$.

L'ipotesi fondamentale che rende il CRM particolarmente attraente è l'assunzione di indipendenza tra le variabili \tilde{K}_{2012} e $\tilde{Z}_{2012,j}$. Sotto questa ipotesi, assolutamente verosimile per il mondo assicurativo, sarà possibile studiare \tilde{K}_{2012} e $\tilde{Z}_{2012,j}$ separatamente, con notevole semplificazione delle procedure.

Una volta identificate le distribuzioni di \tilde{K}_{2012} e $\tilde{Z}_{2012,j}$, la funzione di ripartizione di \tilde{X}_{2012} , $F_{\tilde{X}}(\cdot)$ sarà:

$$F_{\tilde{X}}(x) = P(\tilde{X} \leq x) = \sum_{k=0}^{\infty} P(\tilde{K} = k) P\left(\sum_{j=1}^k \tilde{Z}_j \leq x\right). \quad (3.2)$$

Indicando poi con $F_{\tilde{Z}_j}(\cdot)$ la funzione di ripartizione di \tilde{Z}_j , cioè $F_{\tilde{Z}_j}(z) = P(\tilde{Z}_j \leq z)$, possiamo riscrivere

$$P\left(\sum_{j=1}^k \tilde{Z}_j \leq x\right) = F_{\tilde{Z}_j}^{K*}(x), \quad (3.3)$$

dove $K*$ indica il prodotto di convoluzione² di ordine K .

Chiaramente, la ricerca di una formulazione per la funzione di ripartizione del costo

²Si richiama qui al concetto di convoluzione. Come definito in [7], date due funzioni $f(\cdot)$ e $g(\cdot)$

aggregato dei sinistri attraverso questa via appare molto onerosa.

Una seconda possibilità per descrivere la variabile aleatoria \tilde{X}_{2012} è rappresentata dall'uso delle funzioni generatrici di momenti³ e cumulanti⁴. Sfruttando le proprietà di queste funzioni e le ipotesi adottate in merito alla indipendenza ed identica distribuzione delle $\tilde{Z}_{2012,j}$, possiamo scrivere

$$M_{\tilde{X}}(s|\tilde{K} = k) = M_{\sum_{j=1}^k \tilde{Z}_j}(s) = \prod_{j=1}^k M_{\tilde{Z}_j}(s) = \left(M_{\tilde{Z}_j}(s)\right)^k, \quad (3.4)$$

avendo indicato con M . la funzione generatrice dei momenti.

integrabili in \mathbb{R} , il prodotto di convoluzione è la funzione definita come

$$(f * g)(x) = \int_{-\infty}^{+\infty} f(y)g(x-y)dy.$$

Si noti anche che nel caso in cui $f(\cdot)$ e $g(\cdot)$ siano definite solo su \mathbb{R}^+ , come nel nostro caso, dato che trattiamo con dei costi sinistri, vale

$$(f * g)(x) = \int_0^x f(y)g(x-y)dy.$$

Si tenga anche presente che il prodotto per convoluzione gode anche della proprietà associativa, per cui, indicando con $h(\cdot)$ una terza funzione, vale $(f * g * h)(x) = ((f * g) * h)(x)$.

³Richiamiamo al concetto di funzione generatrice dei momenti. La funzione generatrice dei momenti di una variabile aleatoria \tilde{Y} con funzione di densità $f(y)$ è definita come

$$M_{\tilde{Y}}(s) = \mathbb{E}(e^{ys}) = \int_{-\infty}^{+\infty} e^{ys}f(y)dy.$$

Vi sono diverse proprietà di questa funzione. In particolare ricordiamo la proprietà da cui prende il nome, cioè

$$M_{\tilde{Y}}^{(j)}(s) = \mathbb{E}(\tilde{Y}^j)$$

e la proprietà per cui

$$M_{\tilde{Y}_1 + \tilde{Y}_2}(s) = M_{\tilde{Y}_1}(s) \cdot M_{\tilde{Y}_2}(s)$$

se \tilde{Y}_1 e \tilde{Y}_2 sono due variabili aleatorie indipendenti.

⁴Diamo anche qualche nozione rispetto alla funzione generatrice dei cumulanti. La funzione generatrice dei momenti di una variabile aleatoria \tilde{Y} è definita come

$$\psi_{\tilde{Y}}(s) = \ln(M_{\tilde{Y}}(s)) = \ln(\mathbb{E}(e^{ys})).$$

Anche questa funzione gode di diverse proprietà. Indicando con k_i il cumulante i -esimo della distribuzione di \tilde{Y} , varrà

$$\psi_{\tilde{Y}}^{(j)}(s) = k_j.$$

Anche in questo caso vi la proprietà simila a quella di cui gode la funzione generatrice dei cumulanti per cui se \tilde{Y}_1 e \tilde{Y}_2 ,

$$\psi_{\tilde{Y}_1 + \tilde{Y}_2}(s) = \psi_{\tilde{Y}_1}(s) + \psi_{\tilde{Y}_2}(s).$$

Da essa quindi otteniamo

$$\begin{aligned}
 M_{\tilde{X}}(s) &= \mathbb{E} \left(\left(M_{\tilde{Z}_j}(s) \right)^k \right) = \mathbb{E} \left(M_{\tilde{X}}(s | \tilde{K} = k) \right) = \\
 &= \mathbb{E} \left(e^{k \ln(M_{\tilde{Z}_j}(s))} \right) = \mathbb{E} \left(e^{k(\psi_{\tilde{Z}_j}(s))} \right) = M_{\tilde{K}} \left(\psi_{\tilde{Z}_j}(s) \right), \quad (3.5)
 \end{aligned}$$

dove ψ indica la funzione generatrice dei cumulanti.

Una volta scelte opportunamente le distribuzioni di \tilde{Z}_j e di \tilde{K} , se esse ammettono funzioni generatrici dei momenti, sarà possibile ottenere la funzione generatrice dei momenti di \tilde{X} . Dal momento che una delle proprietà fondamentali delle funzioni generatrici dei momenti è che sono univocamente definite dalla distribuzione di probabilità e viceversa, sarà possibile identificare la distribuzione di \tilde{X} .

Anche questa soluzione non ci appare percorribile per due ragioni in particolare.

- Innanzitutto non è semplice passare da una funzione generatrice dei momenti ad una misura del rischio. Grazie a queste particolari funzioni, infatti, diviene molto più semplice il calcolo dei momenti delle distribuzioni, ma non dei percentili, quindi di VaR e Tail-VaR.
- Le funzioni generatrici dei momenti non sono definite per qualsiasi distribuzione⁵, quindi dovremmo decidere di escludere aprioristicamente alcune distribuzioni, in particolare per la *severity*.

Occorre quindi un'altra via per ottenere le caratteristiche della distribuzione della \tilde{X} una volta descritto il modello. Proprio a causa della sua complessità, il CRM mal si presta alla stima per via esplicita di una funzione di densità o di ripartizione, quindi siamo portati a scegliere un approccio che non la prevede.

L'approccio che invece sembra permetterci al contempo di valutare le caratteristiche della distribuzione di \tilde{X}_{2012} , sia in termini di percentili, sia in termini di momenti (ove

⁵Adotteremo nel seguito ad esempio la distribuzione di Burr Inversa per il costo singolo sinistro che non ammette questa funzione generatrice.

definiti), senza tuttavia dover passare per gli onerosi calcoli necessari ad una ricerca di una funzione di densità in modo esplicito, è il metodo simulativo Monte Carlo.

Il metodo Monte Carlo, infatti, permette di costruire la distribuzione di una variabile aleatoria frutto di diversi stadi di aleatorietà, come nel nostro caso \tilde{X}_{2012} , a partire da una serie di numeri pseudo-casuali generati da un algoritmo di tipo ricorsivo⁶. Una volta infatti trovati dei modelli che si adattino bene a descrivere le variabili aleatorie \tilde{K}_{2012} e $\tilde{Z}_{2012,j}$, sarà sufficiente simulare pseudo-casualmente un k_{2012} , successivamente ricavare un vettore di $z_{2012,j}$ di dimensione k_{2012} e sommarne ogni elemento per ottenere un valore di x_{2012} . Ripetendo l'estrazione un numero sufficientemente grande di volte si ottiene quello che può essere considerato come un campione casuale estratto dalla variabile aleatoria \tilde{X}_{2012} che può essere adottato per calcolare le stime di percentili, momenti, cumulanti ed altre statistiche significative per la variabile aleatoria.

Poiché le x_{2012} ottenute saranno considerate esattamente come un campione casuale per la stima delle statistiche significative, sono da ritenersi validi gli assunti ed i teoremi propri della statistica inferenziale, per cui, al crescere della dimensione del campione, ci si aspetta che la variabilità delle stime si riduca⁷.

3.2 Lo studio del costo singolo sinistro

Per trovare un adeguato modello per la descrizione della variabile aleatoria costo del j -esimo sinistro nell'anno 2012, $\tilde{Z}_{2012,j}$, prenderemo in considerazione due differenti vie.

La prima che percorreremo avrà come obiettivo la costruzione di un modello semplice. Essa, sicuramente la più immediata e la più praticata, consisterà nella scelta di una

⁶Per i nostri calcoli utilizzeremo il software R, generando numeri pseudo casuali attraverso l'algoritmo denominato "Marsenne-Twister", avente un periodo di $2^{19937} - 1$. Questo algoritmo genera numeri pseudo-casuali distribuiti uniformemente nell'intervallo $[0, 1]$, numeri che saranno utilizzati, attraverso diversi metodi, per l'estrazione di campioni da distribuzioni note.

⁷Ad esempio, essendo μ la media di una popolazione e σ^2 la varianza, se si estrae un campione casuale semplice di dimensione n dalla popolazione, lo stimatore media campionaria $\bar{\mu}(n)$, vale che $\mathbb{V}(\bar{\mu}(n)) = \frac{\sigma^2}{n}$, decrescente rispetto ad n . Avremo modo di affrontare più nel dettaglio questo argomento nell'Appendice B

famiglia di distribuzioni nota in letteratura. Si andranno a stimare i parametri di alcune di esse per poi scegliere in modo opportuno quella che meglio si adatta ai dati.

Non ci aspettiamo, in questo caso, di trovare un modello che replichi perfettamente l'aleatorietà dei nostri dati, poiché il numero di parametri che prenderemo di volta in volta in considerazione sarà quantomeno esiguo se confrontato con la mole di dati a nostra disposizione. Requisito fondamentale per la costruzione di un modello, che è stata la difficoltà di molti studiosi prima di noi, sarà quindi la parsimonia nella scelta del numero di parametri. Implicitamente, stabiliamo che i nostri dati non siano pienamente credibili, fissando, di volta in volta, la forma che un modello debba avere. Questa scelta, sebbene non venga neppure messa in discussione dalla maggior parte degli studiosi, potrebbe stridere con il senso comune, ma ci sono due validi motivi per cui decidiamo di percorrerla:

- Evitiamo di correre il rischio della sovrapparametrizzazione, cioè di includere troppi parametri nel nostro modello. Se infatti prevediamo più parametri di quanti ne siano necessari, il modello si adatterà molto bene ai dati osservati, ma risulterà poco o nulla predittivo per quanto riguarda i futuri sinistri.
- Non vi è alcuna ragione valida per poter assumere un valore come il massimo possibile per un sinistro. Tutti i nostri sinistri infatti hanno un costo inferiore a € 132208.60, ma è presente un *outlier* che fa registrare un valore di € 1322862. Scegliendo una distribuzione che abbia come supporto \mathbb{R}^+ , assumiamo che sia possibile, seppure improbabile, che accadano sinistri di importo superiore a qualunque limite.

La seconda via che seguiremo mirerà alla costruzione di un modello che si adatti molto bene ai dati senza tuttavia puntare ad utilizzare una distribuzione esplicita. Utilizzeremo quindi il metodo dell'Ogiva come descritto in [18].

Assunzione fondamentale che facciamo rispetto alla distribuzione dei costi dei singoli sinistri è che non vi siano interdipendenze di alcun genere tra essi, quindi assu-

miamo che tutti i $\tilde{Z}_{i,j}$ siano indipendenti ed identicamente distribuiti⁸ e nel seguito li indicheremo semplicemente con \tilde{Z} . Ci concentreremo quindi sullo studio delle $\tilde{Z}_{i,j}$ con $i = 2008, \dots, 2011$ per poter descrivere le $\tilde{Z}_{2012,j}$.

Per ragioni di coerenza dovremo portare i nostri costi sinistri alle grandezze monetarie del 2011. Poiché nei nostri dati si osserva un andamento del costo medio, nel corso degli anni, molto erratico e tendenzialmente decrescente, non possiamo ricavare da essi un vettore di inflazione utilizzabile. Decidiamo quindi di adottare l'inflazione osservata dall'ISTAT come variazione dell'Indice Nazionale dei Prezzi al Consumo per l'Intera Collettività.

3.2.1 L'adattamento a distribuzioni note

Nel seguito iniziamo con la stima dei parametri di alcune distribuzioni di probabilità note in letteratura. Due sono i requisiti irrinunciabili perché una distribuzione di probabilità possa essere presa in considerazione per descrivere un costo sinistri: deve essere continua e deve assegnare probabilità nulle ai valori negativi. In letteratura attuariale sono usualmente proposti molti modelli, dai più semplici ai più complessi, ma noi ci limiteremo a mostrarne alcuni. Per una rassegna dei più noti si veda [24].

Condurremo le stime sia tenendo conto del valore *outlier* sia escludendolo.

3.2.1.1 La stima della distribuzione Log-Normale

Il primo modello che prendiamo in considerazione è la distribuzione Log-Normale. Secondo questo modello la funzione di densità della variabile \tilde{Z} dipende dai parametri μ e σ^2 risulta

$$f_{\tilde{Z}}(z) = \frac{1}{\sqrt{2\pi\sigma z}} e^{-\frac{1}{2} \frac{(\ln(z)-\mu)^2}{\sigma^2}} \quad z \in \mathbb{R}^+. \quad (3.6)$$

Tale modello viene spesso utilizzato nella pratica attuariale per diverse ragioni che comprendono la semplicità nell'utilizzo e la così definita "pesantezza della coda".

⁸A meno di una riscalatura dovuta all'inflazione, come descriveremo in seguito.

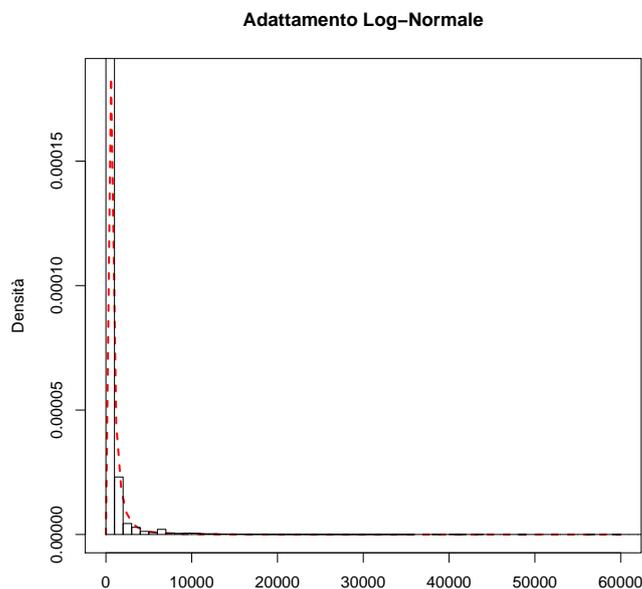


Figura 3.1: Istogramma vs Log-Normale

Per quanto riguarda la prima ragione, se si tiene presente che

$$\tilde{Z} \sim LNorm(\mu, \sigma^2) \iff \ln(\tilde{Z}) \sim Norm(\mu, \sigma^2),$$

appare evidente la semplicità al momento della stima dei parametri e durante la simulazione. Unico appunto che rappresenta un limite di questo modello riguarda i tempi di simulazione: pur ottenendo le stime dei parametri $\hat{\mu}$ e $\hat{\sigma}^2$ utili per \tilde{Z} , in letteratura non si trova alcuna indicazione della distribuzione di $\sum_{j=0}^K \tilde{Z}_j$ sebbene si possano trovare delle formule di approssimazione⁹. Per quanto riguarda invece la “pesantezza della coda”, questa distribuzione assegna probabilità non nulle anche a valori particolarmente elevati¹⁰, proprietà sicuramente apprezzabile nel campo assicurativo, poiché durante la simulazione permette il verificarsi degli eventi estremi.

⁹Si vedano in proposito [43] e [14]

¹⁰Si ricordi, a titolo esemplificativo, che il percentile di ordine 99.5 corrisponde circa a tre volte il valore dello scarto quadratico medio.

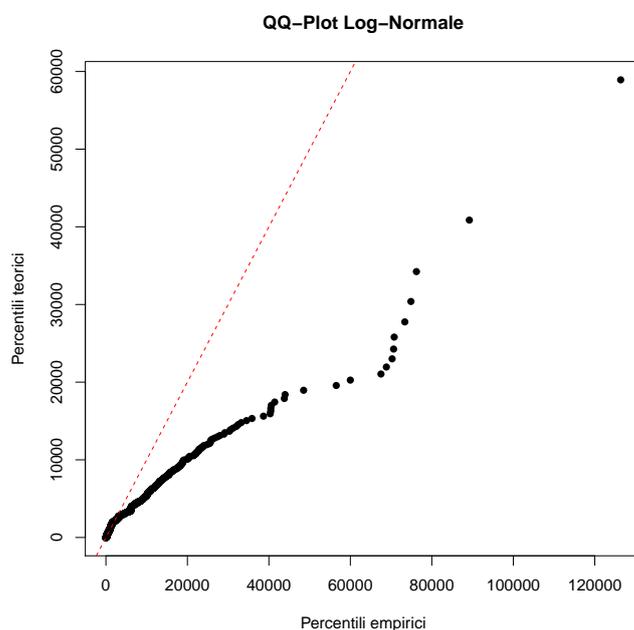


Figura 3.2: QQ-plot Log-Normale

La stima dei parametri secondo il metodo della massima verosimiglianza, escludendo l'*outlier*, produce come risultato $\hat{\mu} = 4.41$ e $\hat{\sigma}^2 = 1.49$ ¹¹. Sotto questa stima la distribuzione assume una forma abbastanza differente dall'istogramma dei nostri dati, come appare in Figura 3.1. Procediamo quindi, in Figura 3.2, alla costruzione di un QQ-plot che metta a confronto i percentili della distribuzione teorica con quelli della distribuzione empirica.

I grafici prodotti ci mostrano come l'adattamento della distribuzione Log-Normale sia poco efficiente. In particolare possiamo notare dalla Figura 3.2 come la distribuzione Log-Normale teorica si adatti male soprattutto per quanto riguarda i valori estremi, generalmente la parte più delicata in un modello. Si può osservare come i punti siano sistematicamente al di sotto della linea che indica la bisettrice del primo quadrante, denotando una sottostima della probabilità di accadimento di eventi estremi da parte del modello teorico. Per tali ragioni procederemo alla stima dei parametri di altre distri-

¹¹Invece, includendo l'*outlier* le stime si modificano solo leggermente divenendo $\hat{\mu} = 4.41$ e $\hat{\sigma}^2 = 1.50$.

buzioni note in letteratura, per poi concentrarci sulla scelta del modello attraverso un confronto basato su alcune statistiche di adattamento.

3.2.1.2 La stima della distribuzione Gamma

La distribuzione Gamma è una distribuzione continua, il cui supporto corrisponde a $(0, +\infty)$, descritta dai due parametri α e β . Secondo questo modello, \tilde{Z} ha funzione di densità

$$f_{\tilde{Z}}(z) = \frac{1}{\beta^\alpha \Gamma(\alpha)} z^{(\alpha-1)} e^{-\left(\frac{z}{\beta}\right)} \quad z \in \mathbb{R}^+, \quad (3.7)$$

dove $\Gamma(\cdot)$ è la nota funzione gamma definita come $\Gamma(x) = \int_0^{+\infty} t^{(x-1)} e^{-t} dt$. Anche in questo caso è possibile fare delle stime di massima verosimiglianza per $\hat{\alpha}$ e $\hat{\beta}$, che restituiscono rispettivamente i valori 0.571 e $2.23E - 03$ ¹².

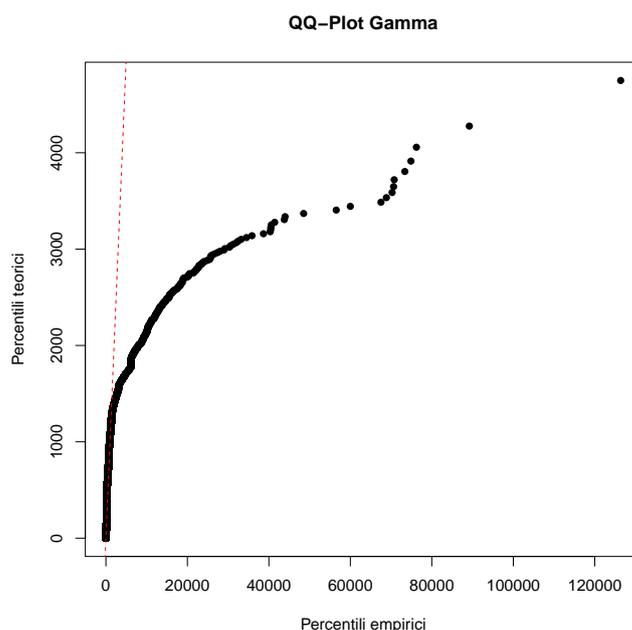


Figura 3.3: QQ-plot Gamma

La Figura 3.3 mostra come anche questo modello sottostimi la probabilità di accadimento degli eventi estremi, ma in modo ancora più marcato rispetto alla Log-Normale.

¹²Con l'outlier divengono 0.538 e $2.03E - 03$.

Per tale regione proveremo a verificare se vi sia un migliore adattamento rispetto a due modelli che prendono piede dalla distribuzione Gamma.

La distribuzione Gamma Generalizzata Il primo modello che prendiamo in considerazione è la distribuzione Gamma Generalizzata così come descritta in [40]. Si tratta di una generalizzazione a tre parametri della distribuzione, secondo cui la funzione di densità diviene

$$f_{\tilde{Z}}(z) = |p| \frac{1}{a^{p\nu} \Gamma(\nu)} z^{(p\nu-1)} e^{-\left(\frac{z}{a}\right)^p} \quad z \in \mathbb{R}^+. \quad (3.8)$$

Si noti come i parametri ν ed a corrispondano rispettivamente ai parametri α e β della distribuzione Gamma, mentre il parametro p rappresenta la generalizzazione. Con questa distribuzione gli usuali algoritmi per la stima dei parametri falliscono nella ricerca dell'ottimo a causa della natura altamente non lineare della funzione di verosimiglianza. Tuttavia, sempre in [40] si trova anche un metodo per la stima dei parametri secondo il metodo dei momenti. Una volta definiti con μ_Y , σ_Y^2 e γ_Y rispettivamente la media, la varianza e l'asimmetria della distribuzione empirica di $Y = \ln(Z)$, si procede alla stima di ν attraverso la relazione

$$-|\gamma_Y| = \frac{\frac{d^3 \ln(\Gamma(\hat{\nu}))}{d\hat{\nu}^3}}{\left(\frac{d^2 \ln(\Gamma(\hat{\nu}))}{d\hat{\nu}^2}\right)^{\frac{3}{2}}}. \quad (3.9)$$

E, successivamente, alla stima di \hat{a} e \hat{p} attraverso

$$\hat{p} = \frac{\left(\frac{d^2 \ln(\Gamma(\hat{\nu}))}{d\hat{\nu}^2}\right)^{\frac{1}{2}}}{\sigma_Y} \quad (3.10)$$

$$\hat{a} = e^{\left(\mu_Y - \frac{1}{\hat{p}} \frac{d \ln(\Gamma(\hat{\nu}))}{d\hat{\nu}}\right)}. \quad (3.11)$$

Essendo quindi $\mu_Y = 4.41$, $\sigma_Y^2 = 1.495$ e $\gamma_Y = -0.116$ sui nostri dati, otteniamo le stime $\hat{\nu} = 9.09$, $\hat{a} = 6.58\text{E} - 03$ e $\hat{p} = 2.28\text{E} - 01$ usando le Formule 3.9, 3.10 e 3.11¹³.

¹³ $\hat{\nu} = 9.16$, $\hat{a} = 6.07\text{E} - 03$ e $\hat{p} = 2.27\text{E} - 01$ considerando tutti i valori.

Anche in questo caso produciamo il QQ-plot, in Figura 3.4.

La distribuzione Log-Gamma Prendiamo quindi in considerazione la distribuzione Log-Gamma¹⁴. Tale distribuzione segue la filosofia della distribuzione Log-Normale, ovvero $\tilde{Z} \sim LGamma(\alpha, \beta)$ se, e solo se, $\tilde{Y} \sim LGamma(\alpha, \beta)$ con $\tilde{Y} = \ln(\tilde{Z})$.

Un problema che sorge con questa distribuzione e che la rende molto meno verosimile rispetto ad una distribuzione Log-Normale sta nel fatto che, mentre il supporto di quest'ultima è $(0, +\infty)$, il supporto della distribuzione Log-Gamma é $(1, +\infty)$. Infatti, se il logaritmo è una funzione del tipo

$$\ln : (0, +\infty) \mapsto (-\infty, +\infty) \quad (3.12)$$

e la funzione di ripartizione della distribuzione Normale Φ è del tipo

$$\Phi : (-\infty, +\infty) \mapsto (0, 1), \quad (3.13)$$

sostituendo la 3.12 nella 3.13 si osserva come il supporto della distribuzione Log-Normale sia \mathbb{R}^+ .

Tuttavia la funzione di ripartizione della distribuzione Gamma Γ è del tipo

$$\Gamma : (0, +\infty) \mapsto (0, 1), \quad (3.14)$$

quindi la distribuzione Log-Gamma considera ammissibili solo valori che hanno come immagine, per la funzione logaritmo, dei valori non negativi, vale a dire che la realizzazione di valori appartenenti all'intervallo $(0, 1)$ è assunta come impossibile¹⁵. Per ovviare a tale problema occorre riscalarare i nostri dati sommando il valore 1 a ciascuno di essi.

Procediamo quindi alla stima dei parametri della distribuzione Log-Gamma.

¹⁴Si noti che non utilizziamo la distribuzione Log-Gamma descritta in [21] appartenente alle distribuzioni Extreme Value, ma utilizziamo quanto descritto in [26].

¹⁵Un tale problema è di portata non indifferente se si considera che riguarda 2144 occorrenze nel nostro dataset.

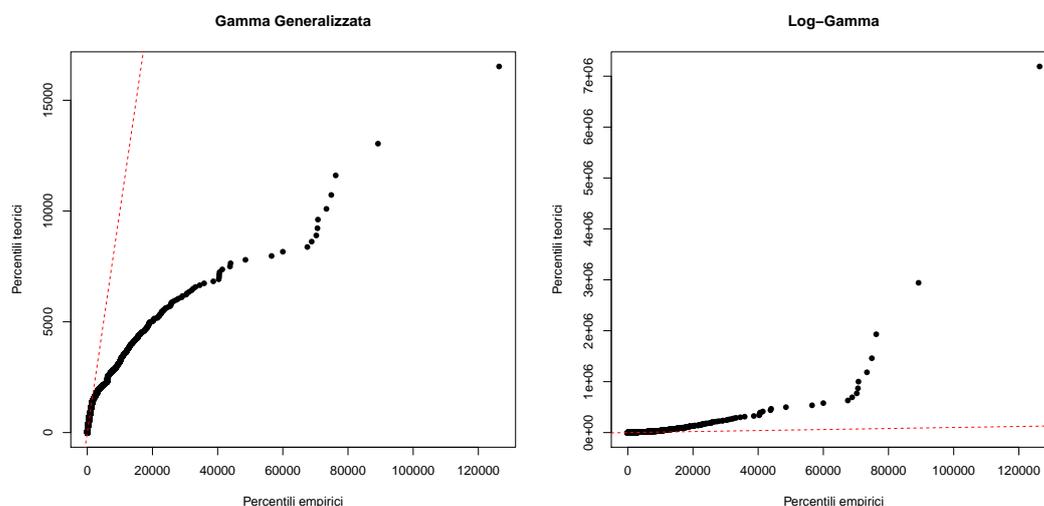


Figura 3.4: QQ-plot Gamma Generalizzata e Log-Gamma

Come è possibile osservare dalla Figura 3.4 né la distribuzione Gamma Generalizzata, né la distribuzione Log-Gamma garantiscono un buon adattamento ai dati a nostra disposizione. La prima infatti risente, come la distribuzione Gamma Semplice, di una sottostima del rischio di eventi estremi, una eccessiva leggerezza della coda, mentre la seconda del problema opposto.

3.2.1.3 Esponenziale

Il più semplice modello che prendiamo in considerazione è la distribuzione esponenziale.

Secondo questo modello, la funzione di densità delle \tilde{Z} dovrebbe essere

$$f_{\tilde{Z}}(z) = \frac{e^{-\frac{z}{\alpha}}}{\alpha} \quad z \in \mathbb{R}^+, \tag{3.15}$$

dipendente dal solo parametro α , che assume valore di $3.635\text{E-}03^{16}$, stimato con il metodo della massima verosimiglianza. L'adattamento, come è lecito aspettarsi per una distribuzione ad un solo parametro non è assolutamente buono, come vedremo nelle sta-

¹⁶ $\hat{\alpha} = 3.451\text{E} - 03$ senza escludere l'*outlier*.

tistiche proposte in Sezione 3.2.1.4, ma lo abbiamo presentato per introdurre alla sezione successiva.

3.2.1.4 La distribuzione di Weibull

Un'altra nota distribuzione che rappresenta una generalizzazione a parametri τ e θ della della distribuzione esponenziale è la distribuzione di Weibull. Si tratta di una distribuzione nata per descrivere il punto di rottura di materiali utilizzati nel campo aeronautico¹⁷ e successivamente applicata per un'ampia gamma di applicazioni. In particolare una variabile aleatoria \tilde{X} ha distribuzione di Weibull se $\tilde{Y} = \left(\frac{\tilde{X}}{\theta}\right)^\tau$ segue la distribuzione esponenziale come in Formula 3.15 con parametro $\alpha = 1$. Pertanto la funzione di densità della distribuzione Weibull sarà

$$f_{\tilde{X}}(x) = \frac{\tau}{\theta} \left(\frac{x}{\theta}\right)^{(\tau-1)} e^{-\left(\frac{x}{\theta}\right)^\tau} \quad x \in \mathbb{R}^+. \quad (3.16)$$

Si noti che si può ulteriormente generalizzare la Formula 3.16 sostituendo $x - \xi_0$ ad x aggiungendo quindi il parametro di localizzazione ξ_0 , come in [22]. Tuttavia nel seguito non ci occuperemo di una tale generalizzazione, non essendo necessaria per i nostri dati. Applicando infine la Formula 3.16 a \tilde{Z} nella forma equivalente utilizzata in [24] si ha

$$f_{\tilde{Z}}(z) = \frac{\tau}{\theta^\tau} z^{(\tau-1)} e^{-\left(\frac{z}{\theta}\right)^\tau} \quad z \in \mathbb{R}^+, \quad (3.17)$$

e le stime di massima verosimiglianza divergono $\hat{\tau} = 0.653$ e $\hat{\theta} = 169.87$ ¹⁸.

In Figura 3.5 mostriamo il QQ-plot ottenuto che mostra come il modello sia inadeguato nonostante la generalizzazione apportata.

¹⁷Si veda in proposito [22].

¹⁸Senza restrizioni divergono $\hat{\tau} = 0.647$ e $\hat{\theta} = 169.65$.

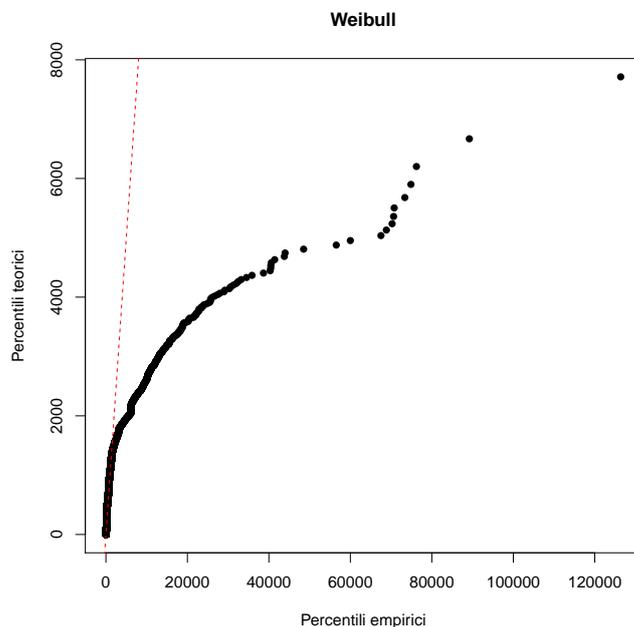


Figura 3.5: QQ-plot Weibull

3.2.1.5 La distribuzione Beta Generalizzata del Secondo tipo

La distribuzione Beta generalizzata del Secondo tipo è una generalizzazione a quattro parametri della distribuzione Beta semplice. Tuttavia, mentre la distribuzione Beta ha supporto l'intervallo $[0, 1]$, la Beta Generalizzata rispetta il requisito che avevamo posto in precedenza di avere come supporto $(0, +\infty)$. Secondo tale modello, la funzione di densità di \tilde{Z} , di parametri a, p, q e b è

$$f_{\tilde{Z}}(z) = \frac{az^{ap-1}}{b^{ap}B(p, q) \left[1 + \left(\frac{z}{b}\right)^{ap+q}\right]} \quad z \in \mathbb{R}^+, \quad (3.18)$$

dove $B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$.

Questa famiglia di distribuzioni appare sicuramente molto attraente, flessibile ed elegante nel *fitting* dei dati, tuttavia, come ci ricorda [23] questo modello potrebbe essere affetto da sovrapparametrizzazione.

Anche in questo caso adottiamo il metodo della massima verosimiglianza per la stima

dei parametri, che ci restituisce i valori $\hat{a} = 5.85$, $\hat{p} = 0.139$, $\hat{q} = 0.201$ e $\hat{b} = 122.05^{19}$ e, come in precedenza, osserviamo il QQ-plot proposto in Figura 3.6.

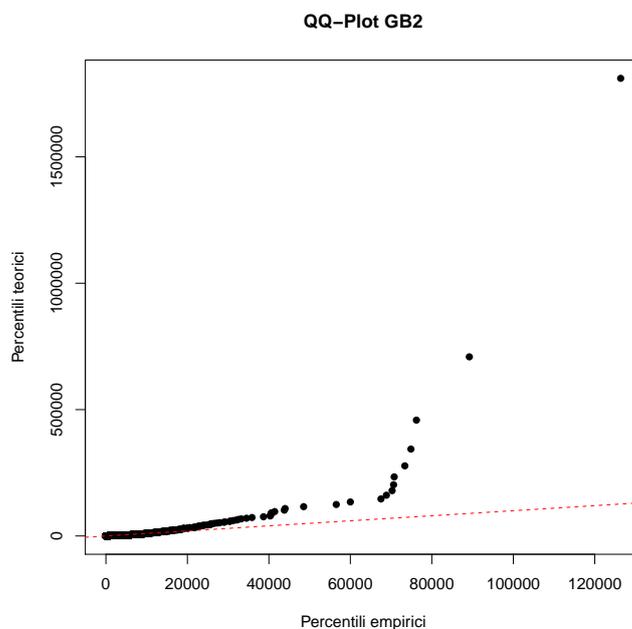


Figura 3.6: QQ-plot Beta Generalizzata del Secondo tipo

Sebbene sulla coda l'adattamento risulti molto approssimativo, possiamo osservare come per buona parte della distribuzione i percentili teorici ed empirici combacino. Ciò ci fa pensare che un modello simile possa essere adeguato a descrivere i nostri dati. Osserviamo quindi ciò che accade se prendiamo in considerazione dei modelli appartenenti a questa famiglia che si configurano come casi particolari della Beta Generalizzata.

La distribuzione di Burr Il primo caso particolare di cui ci occupiamo è la distribuzione di Burr, detta anche Singh-Maddala. La distribuzione che prendiamo in considerazione, la Burr tipo XII, è stata introdotta per la prima volta in [8] come una delle distribuzioni che, indicando con F la funzione di ripartizione, soddisfano l'equazione

¹⁹ $\hat{a} = 5.97$, $\hat{p} = 0.136$, $\hat{q} = 0.197$ e $\hat{b} = 122.10$ nel caso in cui non si adottino restrizioni.

differenziale

$$F'(x) = F(x)[1 - F(x)]g(x), \quad (3.19)$$

tratto distintivo della classe di distribuzioni appartenenti al così definito Sistema di Burr.

Come accennato tale distribuzione può essere vista come un caso particolare della distribuzione Beta Generalizzata del Secondo tipo, in cui il parametro p è pari ad 1. Utilizzando quindi la notazione adottata in [24], la funzione di densità di \tilde{Z} sarà

$$f_{\tilde{Z}}(z) = \frac{\alpha\gamma\left(\frac{z}{\theta}\right)^\gamma}{z\left[1 + \left(\frac{z}{\theta}\right)^\gamma\right]^{\alpha+1}} \quad z \in \mathbb{R}^+ \quad (3.20)$$

in cui $\theta = b$, $\gamma = a$ e $\alpha = q$. Le stime di massima verosimiglianza sono $\hat{\theta} = 118.79$, $\hat{\gamma} = 1.158$ e $\hat{\alpha} = 1.301$ ²⁰.

La distribuzione di Burr Inversa Un'altra distribuzione che può essere vista come un caso particolare della Beta Generalizzata di Secondo tipo è la Burr Inversa, detta anche Dagum. La funzione di densità in questo modello²¹ sarà,

$$f_{\tilde{Z}}(z) = \frac{\tau\gamma\left(\frac{z}{\theta}\right)^{\gamma\tau}}{z\left[1 + \left(\frac{z}{\theta}\right)^\gamma\right]^{\tau+1}} \quad z \in \mathbb{R}^+. \quad (3.21)$$

Anche questa distribuzione rispetta l'Equazione 3.19 ed, essendo stata descritta in [8], è anche nota come la distribuzione Burr III²².

In Figura 3.7 sono proposti i QQ-plot della distribuzione di Burr e della Burr Inversa. Da un rapido confronto possiamo osservare come siano simili tra loro. La distribuzione di Burr Inversa, tuttavia appare leggermente più fedele ai dati empirici.

²⁰Con le corrispondenti $\hat{\theta} = 123.55$, $\hat{\gamma} = 1.146$ e $\hat{\alpha} = 1.335$ nel già citato caso.

²¹si adotta ancora la notazione di [24], in cui $\theta = b$, $\gamma = a$ e $\tau = p$.

²²Essa si configura come il caso della Beta Generalizzata di Secondo tipo in cui il parametro q è pari ad 1 e viene definita Burr Inversa perché

$$X \sim InvBurr(\gamma, \theta, \tau) \iff \frac{1}{X} \sim Burr\left(\gamma, \frac{1}{\theta}, \tau\right).$$

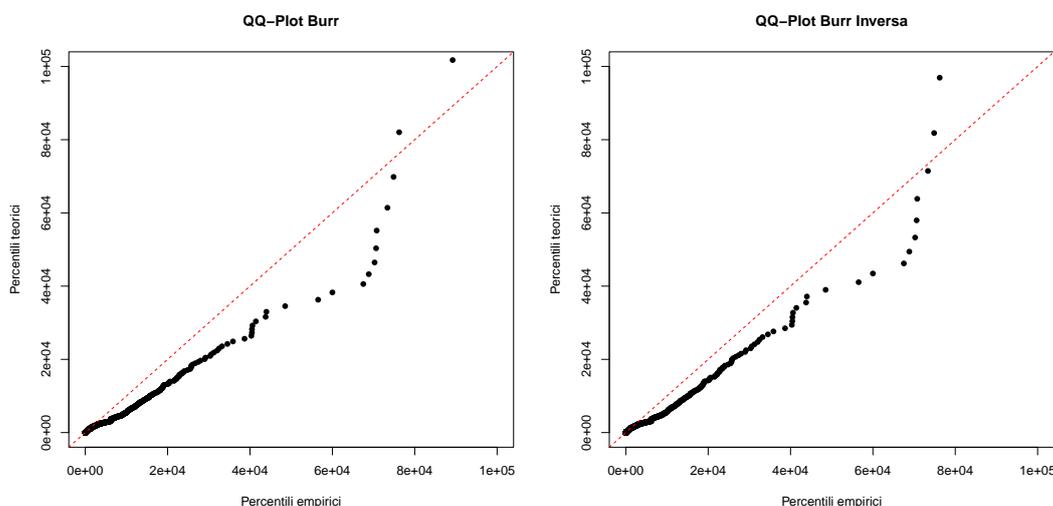


Figura 3.7: QQ-plot Burr e Burr Inversa

Log-Logistica L'ultima distribuzione appartenente a questa famiglia che prendiamo in considerazione è la distribuzione Log-Logistica, detta anche distribuzione di Fisk. Si tratta del caso particolare in cui una distribuzione di Burr abbia $\alpha = 1$, che equivale al caso in cui una distribuzione di Burr Inversa abbia $\tau = 1$. Pertanto la funzione di densità prevista da questo modello a due parametri sarà:

$$f_{\hat{Z}}(z) = \frac{\gamma \left(\frac{z}{\hat{\theta}}\right)^\gamma}{z \left[1 + \left(\frac{z}{\hat{\theta}}\right)^\gamma\right]^2} \quad z \in \mathbb{R}^+. \quad (3.22)$$

In questo caso le stime dei parametri divengono $\hat{\gamma} = 1.253$ e $\hat{\theta} = 86.83^{23}$.

Come si può osservare in Figura 3.8 anche questo modello non si discosta molto dai precedenti analizzati all'interno della Sezione 3.2.1.5, ma, a differenza degli altri, ha un adattamento troppo povero per quanto riguarda la coda. Ciò può essere interpretato come un segnale che i modelli precedenti non sono sovrapparametrizzati, infatti, come suggerito da [23] il parametro α che andiamo a vincolare nella distribuzione Log-Logistica è quello che controlla la coda destra della distribuzione.

²³ Appaiono molto simili in questo caso le stime condotte sull'intero set di dati e risultano $\hat{\gamma} = 1.253$ e $\hat{\theta} = 86.87$.

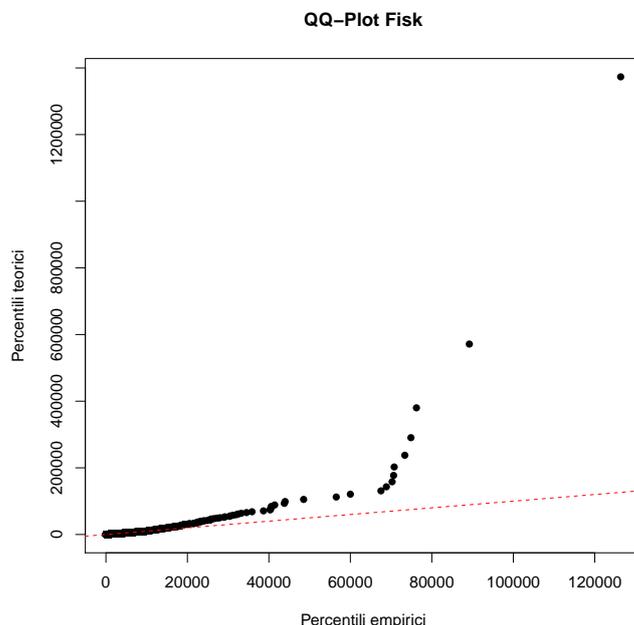


Figura 3.8: QQ-plot Log-Logistica

3.2.1.6 Gaussiana Inversa

Prendiamo ora in considerazione la distribuzione Gaussiana Inversa, anche conosciuta come distribuzione di Wald. Come descritto in [22] ed in [16] tale distribuzione nasce per descrivere, per una particella che segua un Moto Browniano, il tempo che essa impiega a coprire una fissata distanza²⁴. Utilizzando quindi la notazione di [24], la funzione di densità sarà:

$$f_{\bar{Z}}(z) = \left(\frac{\tau}{2\pi z^3}\right)^{\frac{1}{2}} e^{-\frac{\tau}{2z}\left(\frac{z-\mu}{\mu}\right)^2} \quad z \in \mathbb{R}^+. \tag{3.23}$$

²⁴All’opposto, la distribuzione Gaussiana descrive la distanza coperta dalla particella, fissato il tempo. Come descritto in [42], indicando quindi con y il tempo, con x la distanza percorsa, con ω la velocità e con β la costante di diffusione, tale che risulti βy pari alla varianza, le funzioni di densità sono

$$f_X(x) = \frac{1}{\sqrt{2\pi\beta y}} e^{-\frac{1}{2}\frac{(x-\omega y)^2}{\beta y}} \quad \text{e} \quad f_Y(y) = \frac{x}{\sqrt{2\pi\beta y^3}} e^{-\frac{1}{2}\frac{(x-\omega y)^2}{\beta y}}.$$

Le funzioni generatrici dei cumulanti sono quindi:

$$\psi_X(s) = y \left(-\omega s + \frac{1}{2}\beta s^2\right) = yG(s) \quad \text{e} \quad \psi_Y(s) = x \frac{\left(\omega - \sqrt{\omega^2 + 2\beta s}\right)}{\beta} = xG^{-1}(s),$$

proprietà da cui deriva il nome “Gaussiana Inversa”.

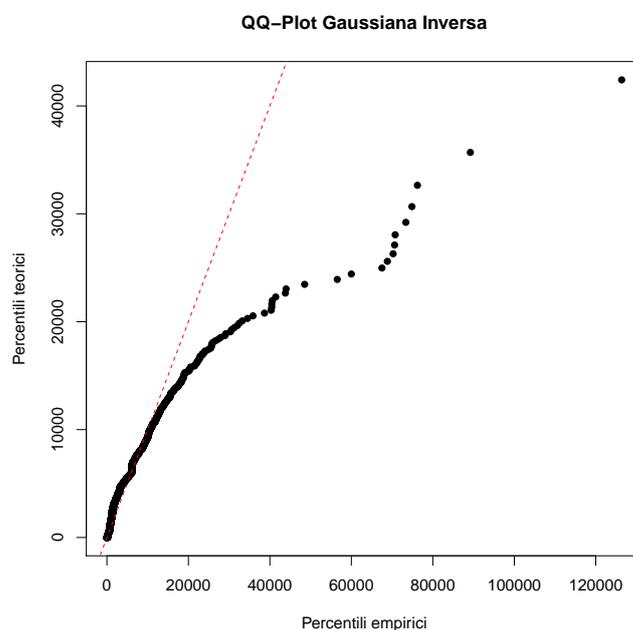


Figura 3.9: QQ-plot Gaussiana Inversa

Avendo ottenuto come stime di massima verosimiglianza dei parametri i valori $\hat{\tau} = 19.21$ $\hat{\mu} = 262.98$, proponiamo in Figura 3.9 il QQ-plot

3.2.2 I criteri per la scelta di una distribuzione

Dopo aver ottenuto le stime e le prime analisi qualitative dei vari modelli nelle precedenti sezioni, procediamo con un'analisi statistica per la selezione del migliore. Consideriamo le metodologie suggerite in [24] per scegliere la distribuzione da adottare per simulare il costo del singolo sinistro.

Il primo approccio è basato sul giudizio e sull'esperienza. Premesso che nel nostro caso il modello non è né aprioristicamente determinato dalla provenienza stessa dei dati, né suggerito da qualche caso esemplare di successo rilevante, i due casi più semplici prospettati in [24], occorre prendere una decisione per quanto riguarda l'obiettivo che vogliamo raggiungere. Il modello che selezioneremo, come già sottolineato in precedenza, verrà utilizzato per condurre una simulazione di tipo Monte Carlo e sappiamo che uno

dei modelli più spesso adottati in letteratura per modellare il costo sinistri (singolo, ma talvolta anche aggregato²⁵) è la Log-Normale. In effetti, la distribuzione Log-Normale non mostra un pessimo adattamento ai dati in nostro possesso. Decidiamo quindi di adottare questa distribuzione come primo modello.

Il secondo approccio, sicuramente più quantitativo e meccanico prevede che le statistiche di quattro test di ipotesi siano usate come punteggi per stabilire quale sia la scelta migliore. Questi test sono stati costruiti per verificare l'appartenenza di due campioni alla medesima popolazione o, più comunemente per verificare l'adattamento di una distribuzione a dei dati empirici. Questi test assumono come H_0 "Il campione proviene dalla distribuzione considerata" e H_1 "Il campione non proviene dalla distribuzione considerata". Occorre comunque tener presente che le statistiche possono dare dei risultati contraddittori o comunque non perfettamente allineati, poiché i test possono considerare aspetti differenti dell'adattamento²⁶ ed è a questo punto che entra nuovamente in gioco l'esperienza.

3.2.2.1 Test di Kolmogorov-Smirnov

La prima statistica che consideriamo proviene dal test noto col nome di Kolmogorov-Smirnov. Il test si basa sulla statistica

$$D = \max_{z \in S_Z} |F_n(z) - F^*(z)|, \quad (3.24)$$

dove $F_n(z)$ indica la funzione di ripartizione empirica e $F^*(z)$ la teorica. La distribuzione di probabilità della statistica D viene discussa in [25], in cui si trova anche una dimostrazione che non dipende dalla $F^*(z)$, e successivamente tabulata in [39]. La statistica D , evidentemente, sarà tanto maggiore quanto più sarà differente la distribuzione empirica da quella teorica, tuttavia lo stesso peso viene assegnato a tutti gli scostamenti, sia che

²⁵Ad esempio nella Standard Formula prevista dal QIS5 per il calcolo dell'SCR.

²⁶ad esempio vedremo in Sezione 3.2.2.1 un test basato sullo scostamento massimo, mentre in Sezione 3.2.2.2 un test basato sullo scostamento medio.

siano sulla coda, sia nel corpo della distribuzione e ciò rischia di ridurre la potenza del test. Ad esempio, ipotizziamo di dover verificare l'adattamento di due differenti distribuzioni teoriche $F^{*(1)}$ e $F^{*(2)}$ ad un campione empirico ed ipotizziamo che la prima abbia un buon adattamento per quanto riguarda il corpo della distribuzione, ma sia completamente fuorviante per quanto riguarda la coda della distribuzione, mentre $F^{*(2)}$ abbia un ottimo adattamento sulla coda, ma abbia scostamenti simili o di poco superiori per quanto riguarda il corpo. È probabile che $D^{(1)} \leq D^{(2)}$ poiché è ragionevole aspettarsi che gli scostamenti sul corpo siano in valore assoluto maggiori degli scostamenti sulla coda e saremmo quindi portati a preferire $F^{*(1)}$, ma nel caso ipotizzato $F^{*(2)}$ dovrebbe essere la distribuzione migliore.

3.2.2.2 Test di Anderson-Darling

Il secondo test che consideriamo è quello introdotto in [2] e poi ripreso in [3] che adotta come statistica

$$A^2 = n \int_0^{+\infty} \frac{[F_n(z) - F^*(z)]^2}{F^*(z)[1 - F^*(z)]} dF^*(z). \quad (3.25)$$

A differenza del test precedente, questo non assegna lo stesso peso a tutti i valori e vengono considerati tutti gli scostamenti, non solo il massimo, quindi viene dato un peso maggiore rispetto al test precedente a quelli relativi alla coda.

3.2.2.3 Test del χ^2

Un altro test che consideriamo è il test del χ^2 . Per misurare gli scostamenti tra la distribuzione empirica e la teorica secondo questo test è innanzitutto necessario dividere i nostri dati in k classi e calcolare, su ognuna di esse, la frequenza relativa p_j delle nostre osservazioni. Successivamente si calcola p_j^* , la probabilità che la distribuzione teorica

assegna ad ognuna delle classi considerate. Il test prende quindi in esame la statistica

$$\chi^2 = \sum_{j=1}^k \frac{n(p_j^* - p_j)^2}{p_j^*}, \quad (3.26)$$

dove n indica il numero di osservazioni, distribuita secondo una variabile χ^2 con un numero di gradi di libertà pari al numero dei parametri della distribuzione sottratti a n .

Il test del χ^2 , a differenza dei precedenti, permette qualche discrezionalità mentre viene costruito, poiché non vi è alcun vincolo per quanto riguarda la scelta degli estremi delle classi da utilizzare. Decidiamo quindi di fissare gli estremi in modo che le frequenze assolute teoriche di ogni singola classe siano costanti e pari a circa 100 osservazioni. Modificheremo quindi gli estremi delle classi per le diverse distribuzioni che considereremo.

3.2.2.4 Test del Rapporto di Verosimiglianza

Infine, l'ultimo test che consideriamo è il Likelihood Ratio test. In esso le ipotesi sono differenti e sono H_0 : "I dati provengono dalla distribuzione $F^{*(1)}$ " e H_1 : "I dati provengono dalla distribuzione $F^{*(2)}$ ", dove tuttavia $F^{*(1)}$ deve essere un caso particolare di $F^{*(2)}$. Si adotta come statistica

$$T = 2 \ln \left(\frac{L_2}{L_1} \right) = 2(\ln(L_2) - \ln(L_1)), \quad (3.27)$$

dove L_2 e L_1 indicano i valori della funzione di verosimiglianza calcolati per i due modelli. Si dimostra anche che $T \sim \chi^2(\nu = k)$ con k pari alla differenza tra il numero di parametri dei due modelli. Tuttavia, come punteggio per la nostra scelta non considereremo la statistica T , ma solo il valore assunto dalla funzione di verosimiglianza.

3.2.3 La scelta della distribuzione

Proponiamo in Tabella 3.1 il prospetto dei valori assunti dalle varie statistiche presentate in precedenza per le distribuzioni descritte.

	K-S	A-D	χ^2	(Log)-Verosim.
Log-Normale	0.0761	977.79	745623.88	-568061.59
Gamma	0.1972	Inf	786958.96	-586745.40
Gamma Gen.	0.0913	Inf	749976.84	-567545.52
Log-Gamma	0.1033	2549.43	763289.80	-577131.60
Esponenziale	0.3068	Inf	805712.31	-603234.24
Weibull	0.1302	Inf	759255.64	-575233.47
Beta Gen. 2	0.0564	Inf	745230.64	-562835.83
Burr	0.0676	408.50	756933.53	-564725.10
Burr Inversa	0.0651	328.47	736093.32	-564209.81
Log-Logistica	0.0670	475.77	759523.08	-565126.14
Gaussiana Inversa	0.2627	Inf	797737.82	-597224.92

Tabella 3.1: Valori delle statistiche dei modelli

Ordiniamo anche, in Tabella 3.2, le distribuzioni secondo le statistiche considerate come punteggi, in modo da avere un impatto visivo più significativo.

Notiamo che le distribuzioni della famiglia Beta si comportano molto bene sui nostri dati. In particolare notiamo che la Burr Inversa compare ai primi posti di tutte le graduatorie ed è la migliore per quanto riguarda il test di Anderson-Darling, quello che dà maggior importanza ai valori della coda della distribuzione. Tuttavia osserviamo anche che la Beta Generalizzata si comporta meglio della Burr Inversa per quanto riguarda il test di Kolmogorov-Smirnov e la verosimiglianza, ma la statistica A^2 del test di Anderson-Darling non converge. Eseguendo il test del rapporto di verosimiglianza otteniamo il valore $T = 2747.975^{27}$, ciononostante due fattori ci fanno propendere per la scelta della distribuzione di Burr Inversa:

²⁷Come descritto in Sezione 3.2.1.5 la Burr Inversa è un caso particolare della Beta Generalizzata del Secondo Tipo, per cui il test può essere condotto.

- la statistica di Anderson-Darling ci indica un adattamento migliore della Burr Inversa;
- l'algoritmo da cui generiamo numeri pseudo-casuali dalla distribuzione Beta Generalizzata del Secondo Tipo, con i parametri che abbiamo stimato, restituisce un valore infinito in circa quattro casi su 10000, rendendoci impossibile il calcolo delle statistiche in base alle quali valutare il rischio.

Decidiamo quindi di adottare la distribuzione Burr Inversa per le simulazioni che condurremo nella Sezione 3.4.

	K-S	A-D	χ^2	(Log)-Verosim.
1	Beta Gen. 2	Burr Inversa	Burr Inversa	Beta Gen. 2
2	Burr Inversa	Burr	Beta Gen. 2	Burr Inversa
3	Log-Logistica	Log-Logistica	Log-Normale	Burr
4	Burr	Log-Normale	Gamma Gen.	Log-Logistica
5	Log-Normale	Log-Gamma	Burr	Gamma Gen.
6	Gamma Gen.	Gamma	Weibull	Log-Normale
7	Log-Gamma	Gamma Gen.	Log-Logistica	Weibull
8	Weibull	Esponenziale	Log-Gamma	Log-Gamma
9	Gamma	Weibull	Gaussiana Inversa	Gamma
10	Gaussiana Inversa	Beta Gen. 2	Gamma	Gaussiana Inversa
11	Esponenziale	Gaussiana Inversa	Esponenziale	Esponenziale

Tabella 3.2: Ranking dei modelli

3.2.4 Il Costo singolo sinistro senza un modello

In [18] troviamo un metodo molto semplice per costruire una distribuzione che abbia un buon adattamento ai dati: si tratta dell'Ogiva. Per dare un'intuizione di come questo metodo operi, ipotizziamo che la variabile \tilde{X} sia una variabile aleatoria e che a nostra disposizione ci sia solo il campione $\{x_1, x_2, x_3, x_4, x_5\}$, estratto casualmente, i cui dati sono stati approssimati all'intero superiore.

In Figura 3.10 riportiamo la funzione di ripartizione empirica $F_n(x)$ costruita sul campione a nostra disposizione. Sicuramente la $F_n(x)$ può essere considerata una stima

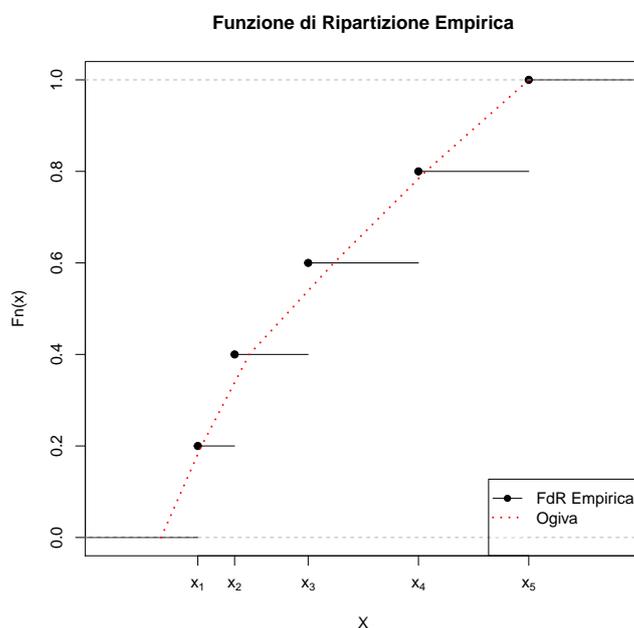


Figura 3.10: Funzione di Ripartizione Empirica della variabile X

di $F(x)$, la vera funzione di ripartizione di X , ma è preferibile renderla in qualche modo più regolare, in particolare se si intende assumere che la distribuzione vera sia di tipo continuo. Il modo più semplice di farlo è quello di tracciare una spezzata di k segmenti che passi attraverso la ripartizione empirica incontrando i punti $(c_i, F_n(c_i))$ $i = 0, \dots, k$, dove $c_0 \leq x_1$, $c_k \geq x_5$, $F_n(c_0) = 0$ e $F_n(c_k) = 1$. Questa spezzata, che chiameremo $H(x)$ prende il nome di Ogiva e la sua derivata prima $h(x)$ è una stima della funzione di densità.

Si noti che per la costruzione di questo modello non viene in alcun modo vincolata la scelta delle distanze tra i punti intermedi, tuttavia possiamo intuire come questa discrezionalità sia molto delicata, poiché implicita al modello vi è l'ipotesi forte che ci sia equidistribuzione tra due c_i .

Da un lato quindi potremmo essere portati a scegliere delle differenze infinitesime tra i punti intermedi, certi di ottenere un ottimo adattamento ai nostri dati, ma dall'altro, il principio di parsimonia ci impone di utilizzare un limitato numero di segmenti. Decisione

che prendiamo sarà comunque quella di definire i c_i non in termini di distanze tra essi, ma in termini di distanze tra i valori di $F_n(c_i)$.

Per dare un'idea della potenzialità di questo metodo osserviamo cosa accade se selezioniamo i c_i in modo che l'incremento nella probabilità sia pari a $1E - 04$. La statistica D del test di Kolmogorov-Smirnov assume il valore di $5E - 04$, con un p -value pari ad 1. In Figura 3.11 mostriamo il QQ-plot di questo modello.

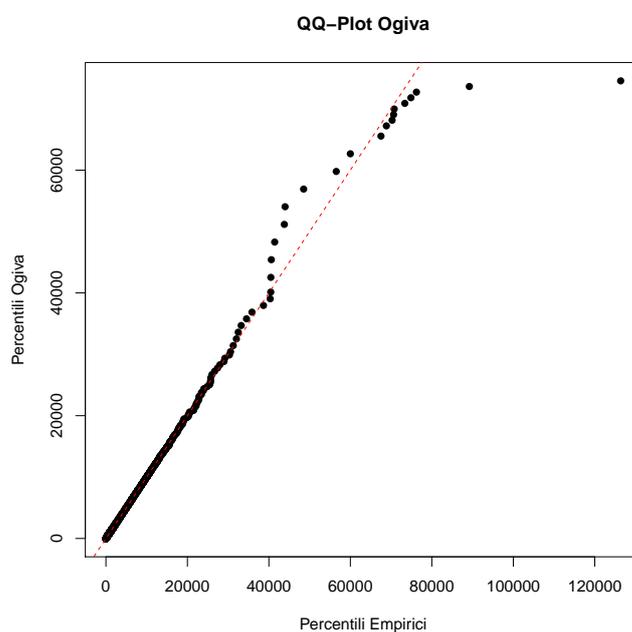


Figura 3.11: QQ-plot Ogiva con incremento $1E - 04$

Dall'altro lato, tuttavia, ridurre il numero di punti intermedi va ad erodere significativamente l'adattamento del modello, soprattutto per quanto concerne la coda destra della distribuzione. Costruendo infatti l'Ogiva con incrementi dell'ordine di 0.01, si ottiene un valore 0.0259 per la statistica D , ma si osserva un errore non trascurabile per quanto riguarda la coda. Ciò appare naturale se si considera che il 99° percentile della nostra distribuzione empirica è € 3104.154, quindi si assume che ci sia equidistribuzione nell'intervallo [€ 3104.154, € 135910.4], ipotesi del tutto inopportuna.

Decidiamo quindi di ridurre ulteriormente il numero di punti fino al 99° percentile,

adottando come step di probabilità 0.065, ma aumentandoli oltre tale limite, con uno step che si riduce progressivamente dallo 0.2% allo 0.02% e scegliendo $c_k = \text{€ } 1$. Otteniamo così un modello in cui $D = 0.0609$, il cui QQ-plot viene proposto in Figura 3.12 da cui si può notare che l'adattamento, anche per quanto riguarda la coda, è molto buono.

Tuttavia, la carenza di questo modello riguarda la limitazione che viene imposta sulla coda. Tale limite costituisce un'ipotesi molto forte anche se la nostra analisi tratta polizze che per natura prevedono limitazioni contrattuali ben definite. Dimostrazione di questa affermazione è il fatto di aver osservato un evento estremamente differente dagli altri per dimensione.

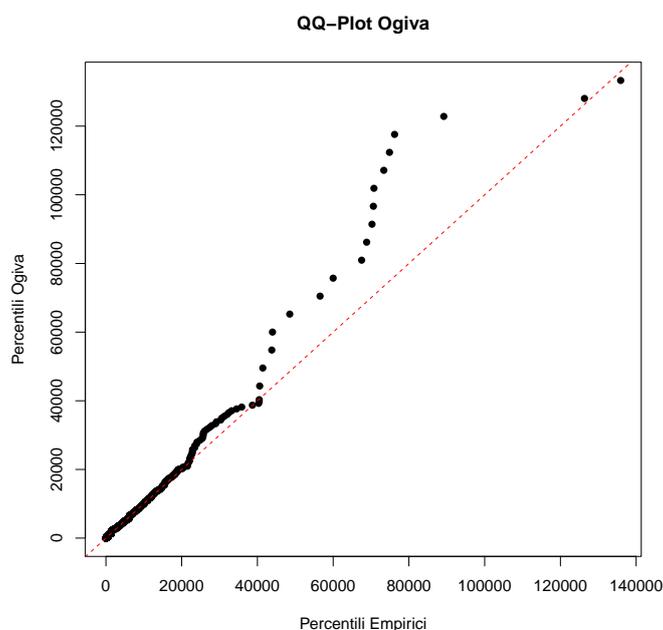


Figura 3.12: QQ-plot Ogiva con incrementi differenti

3.3 Il numero dei sinistri

Il secondo passo per la costruzione del Collective Risk Model consiste nello studio della variabile aleatoria numero dei sinistri nell'anno 2012, \tilde{K}_{2012} . Prima di cercare un

modello appropriato per la descrizione di \tilde{K}_{2012} è opportuno osservare alcune particolarità dei dati a nostra disposizione. Riportiamo in Tabella 3.3 i dati relativi al numero di polizze in essere nei vari anni ed al numero di sinistri osservati²⁸.

	2008	2009	2010	2011
Polizze	116	197	263	274
Sinistri	16385	21587	29379	23816
Frequenza Sinistri	141.25	109.58	111.71	86.92
Varianza numero sinistri	959405.23	961093.79	948929.65	964348.06

Tabella 3.3: Riassuntivo numero sinistri, al netto dei sinistri IBNR

Riportiamo i dati al netto dei sinistri *Incurred, But Not Reported* al fine di ottenere un quadro analitico più coerente. Essendo infatti il portafoglio in fase di *start-up*, se includessimo gli IBNR, avremmo una frequenza sinistri per l'anno 2011 influenzata dai tardivi degli anni precedenti, mentre per il 2008, il dato sarebbe identico. Terremo comunque conto, nel seguito, dei sinistri di questa tipologia.

Un altro fattore di cui terremo conto sarà la stima del numero di polizze in essere nell'anno 2012, ma, per non complicare inutilmente il modello, decidiamo per una stima di tipo deterministico. Come in Figura 3.13 i modelli che appaiono più verosimili per l'evoluzione del numero di polizze sono una parabola ed un'esponenziale. Tuttavia, poiché non c'è alcun segno che le polizze stiano diminuendo, ma anzi sembra che si stiano assestando, propendiamo per il modello esponenziale, che ci restituisce la stima per il 2012 di 278 polizze.

Per poter correttamente sviluppare un modello che possa descrivere i dati relativi al numero dei sinistri nel nostro portafoglio, dobbiamo innanzitutto chiederci come potremmo rendere i dati in nostro possesso più coerenti gli uni con gli altri. Per il 2008, come abbiamo detto, disponiamo di informazioni più mature circa lo sviluppo dei sinistri in termini di IBNR, ma questo non ci è possibile con riferimento al 2011, ultimo anno

²⁸Ricordiamo che la clientela che qui trattiamo non è di tipo *retail*, quindi delle frequenze sinistri così elevate sono da considerarsi coerenti.

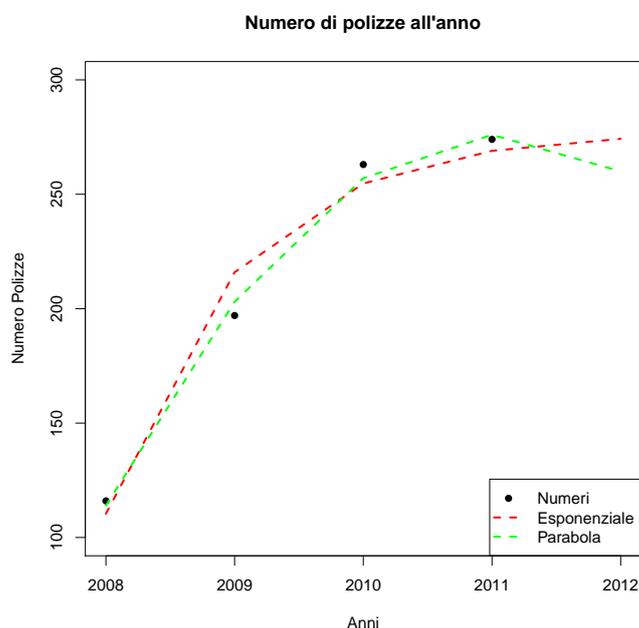


Figura 3.13: Numero di polizze in essere

	0	1	2	3
2008	16385	7465	58	1
2009	21587	6447	118	
2010	29379	10113		
2011	23816			

Tabella 3.4: Triangolo dei sinistri denunciati, valori incrementali

di cui abbiamo le statistiche. Per ottenere un calcolo di quale possa essere il numero atteso dei sinistri per le generazioni più giovani, raccogliamo i nostri dati in un triangolo di *run-off*.

Il metodo che ci appare più opportuno per la stima del numero di sinistri generati in un'anno è il metodo *chain ladder*. Possiamo calcolare i *link ratios* attraverso il metodo classico, per cui, a partire dal triangolo dei sinistri cumulati, $m_j = \frac{\sum_{i=0}^{t-j} n_{i,j}}{\sum_{i=0}^{t-j} n_{i,j-1}}$. Otteniamo così i valori dei *link ratios* riportati in Tabella 3.5 e in Tabella 3.6 procediamo alla stima di quale sia il numero dei sinistri attesi per tutte le generazioni.

m_1	m_2	m_3
1.35671	1.00339	1.00004

Tabella 3.5: *Link ratios*

	0	1	2	3
2008	16385	23850	23908	23909
2009	21587	28034	28152	28153.18
2010	29379	39492	39625.96	39627.62
2011	23816	32311.48	32421.09	32422.45

Tabella 3.6: Triangolo dei sinistri denunciati con stima dell'evoluzione, valori cumulati

I dati ottenuti, tuttavia, non sono ancora del tutto coerenti, poiché, come abbiamo potuto osservare, il numero di polizza, nel corso degli anni, cambia notevolmente.

Decidiamo quindi di portare le quattro osservazioni del numero sinistri aggregato alle grandezze attese per il 2012. Avendo un'aspettativa per quanto riguarda il numero delle polizze del 2012, possiamo condurre una semplice riscalatura su base moltiplicativa, ottenendo, come nuovi numeri dei sinistri, i valori in Tabella 3.7.

Anni	2008	2009	2010	2011
Polizze	116	197	263	274
\tilde{K}	57299.16	39728.86	41887.75	32895.77

Tabella 3.7: Numeri dei sinistri riportati alle grandezze del 2012

L'ipotesi classica di base che viene adottata rispetto al numero dei sinistri in un CRM è che siano distribuiti come una Poisson. Essa, poi, viene usualmente generalizzata introducendo un fattore di disturbo.

3.3.1 Poisson

La distribuzione di Poisson è una distribuzione discreta, avente per supporto \mathbb{N} , definita da un solo parametro che ne indica sia il valore atteso, sia la varianza, la cui

funzione di probabilità è

$$P(\tilde{K} = k) = \frac{e^{-n} n^k}{k!} \quad k \in \mathbb{N}. \quad (3.28)$$

Il nostro prossimo passo, quindi, è rappresentato dalla stima di n , che indicherà il numero dei sinistri attesi per il 2012. Potremmo condurne la stima come semplice media dei numeri riportati in Tabella 3.7, ma daremmo la stessa credibilità a tutti i valori, anche se gli ultimi sono più consolidati dei precedenti, essendo frutto di un numero maggiore di polizze.

Decidiamo quindi per una media ponderata con il numero di polizze, per dare più credibilità ai dati più consolidati, ottenendo $\tilde{n} = 40592.01$.

Notiamo tuttavia un risultato che non ci permette di accettare pienamente il modello. Otteniamo infatti che $\sigma_{\tilde{K}}^2 = 77171566.98^{29}$, ben differente da \tilde{n} . Dobbiamo quindi prendere in considerazione un differente modello.

3.3.2 Poisson Misturata

Il modello che prendiamo in considerazione è un modello di Poisson misturato. Tratteremo più nel dettaglio i modelli di questo tipo nel corso del Capitolo 5, ma per ora basti considerare che introduciamo una variazione rispetto al parametro della distribuzione di Poisson.

La variazione è un fattore di disturbo moltiplicativo che va a modificare il parametro della distribuzione, Poisson in questo caso, con l'aspettativa che riesca a cogliere l'aleatorietà dovuta all'anno di calendario.

Entrando nel dettaglio di cosa significhi l'introduzione di un fattore di disturbo, richiamiamo dei concetti ben descritti in [12]. Il processo aleatorio che descrive nel corso degli anni la variabile \tilde{K} è costituito da diversi fattori:

1. *trend*, cioè variazioni tendenziali di lungo periodo;

²⁹Il dato è nuovamente stato prodotto considerando una ponderazione con il numero di polizze.

2. cicli di lungo periodo;
3. oscillazioni stagionali o comunque di breve periodo³⁰;
4. la pura fluttuazione casuale.

Mentre l'ultimo di questi fattori è già descritto dal modello di Poisson Semplice, i primi tre incidono principalmente sulla frequenza sinistri, quindi sul parametro n della distribuzione. Possiamo anche ipotizzare che, in un'ottica di breve periodo come la nostra, i primi due fattori possano essere considerati deterministici. Inoltre, poiché la nostra serie storica è costituita da appena quattro dati, non ci è possibile cogliere l'effetto di questi due fattori che saranno quindi inclusi nella stima di n . Tuttavia il terzo fattore dovrà essere in qualche modo considerato nel nostro modello.

I dati a nostra disposizione per conoscere il comportamento di questo fattore sono solo le quattro osservazioni della variabile \tilde{K} , opportunamente aggiustate come descritto, ma grazie ad esse possiamo produrre una stima di quale possa essere la deviazione standard del disturbo, che considereremo aleatorio e lo chiameremo \tilde{q} .

Per essere più precisi in merito al modello che adottiamo, partiamo considerando la nuova funzione di probabilità di \tilde{K} condizionata all'evento $\tilde{q} = q$. Essa sarà

$$P(\tilde{K} = k | \tilde{q} = q) = \frac{e^{-nq}(nq)^k}{k!} \quad k \in \mathbb{N}, q \in \mathbb{R}^+. \quad (3.29)$$

Da questa possiamo passare alla distribuzione di \tilde{K} attraverso

$$P(\tilde{K} = k) = E_{\tilde{q}} \left(P(\tilde{K} = k | \tilde{q} = q) \right) = \int_0^{+\infty} \frac{e^{-nq}(nq)^k}{k!} dH(q), \quad (3.30)$$

dove $H(q)$ indica la funzione di ripartizione di \tilde{q} .

Innanzitutto poniamo due condizioni su \tilde{q} necessarie affinché sia coerente con la scelta di adottare un fattore moltiplicativo:

³⁰La differenza tra questo fattore ed il precedente è l'autocorrelazione. Se consideriamo una serie storica dei numeri dei sinistri osservati su base annuale, il contributo delle oscillazioni stagionali sull'autocorrelazione sarà nullo, mentre sarà rilevante quello dei cicli di lungo periodo.

- la prima sarà che $\tilde{q} \geq 0$, altrimenti si otterrebbe il parametro della Poisson pari a $n \cdot \tilde{q} < 0$, divenendo privo di significato;
- la seconda sarà $\mathbb{E}(\tilde{q}) = 1$, questo per assicurare che il valore atteso di \tilde{K} resti pari a n .

Procediamo come segue. La funzione generatrice dei momenti di un processo di Poisson Semplice è:

$$M_{\tilde{K}}(s) = e^{n(e^s-1)}, \quad (3.31)$$

che, per un processo di Poisson composto con la variabile aleatoria \tilde{q} , diviene

$$M_{\tilde{K}}(s) = \mathbb{E}_{\tilde{q}}(M_{\tilde{K}}(s|\tilde{q} = q)) = \mathbb{E}_{\tilde{q}}\left(e^{nq(e^s-1)}\right) = M_{\tilde{q}}(n(e^s - 1)) \quad (3.32)$$

da cui, quindi, passando per semplicità alla generatrice dei cumulanti,

$$\psi_{\tilde{K}}(s) = \psi_{\tilde{q}}(n(e^s - 1)). \quad (3.33)$$

Da questa relazione possiamo quindi ricavare la relazione che ci permetterà di calcolare $\sigma_{\tilde{q}}$, infatti

$$\begin{aligned} \sigma_{\tilde{K}}^2 &= \psi_{\tilde{K}}''(s=0) = [\psi_{\tilde{q}}(n(e^s - 1))]''_{s=0} = \\ &= \left[\psi_{\tilde{q}}'(n(e^s - 1))(ne^s) + (ne^s)^2 \psi_{\tilde{q}}''(n(e^s - 1)) \right]_{s=0} = n + n^2 \sigma_{\tilde{q}}^2. \end{aligned} \quad (3.34)$$

Possiamo utilizzare quindi la Formula 3.34 per ottenere una stima attraverso il metodo dei momenti di $\sigma_{\tilde{q}}$. Attraverso quindi n e $\sigma_{\tilde{K}}^2$ prodotti nella sezione precedente, arriviamo ad una stima di $\sigma_{\tilde{q}}$ pari a 0.2164.

3.3.2.1 La Poisson Composta Gamma

Per la scelta della distribuzione di \tilde{q} le uniche condizioni necessarie sono la non negatività e il fatto che il valore atteso debba essere unitario.

Per il resto non abbiamo imposizioni particolari. Potremmo scegliere un modello distributivo opportunamente calibrato o una distribuzione in forma di tabella.

Tuttavia, nella pratica attuariale, un modello che viene spesso adottato a questo scopo è la distribuzione Gamma.

Non disponiamo di dati a sufficienza per costruire un metodo di selezione di una distribuzione come abbiamo fatto per la severity, quindi non possiamo supportare una particolare scelta distributiva.

L'unica cosa che possiamo fare è una semplice scelta e decidiamo quindi di adottare l'ipotesi distributiva Gamma per il fattore di disturbo, affidandoci alla comune prassi attuariale.

Come descritto in precedenza, la prima necessità nella scelta dei parametri della Gamma sarà di avere, secondo la parametrizzazione che abbiamo adottato, $\alpha = \frac{1}{\beta}$, al fine di assicurare $\mathbb{E}(\tilde{q}) = 1$.

Essendo, quindi, $\mathbb{V}(x) = \alpha\beta^2$ se $X \sim \text{Gamma}(\alpha, \beta)$, otteniamo $\hat{\beta} = 0.0468$ e $\hat{\alpha} = 21.363$.

Tuttavia una distribuzione di Poisson, composta con un fattore di disturbo moltiplicativo distribuito come una Gamma avente valore atteso unitario, altro non è che una distribuzione Binomiale Negativa (o Pólya), quindi

$$P(\tilde{K} = k) = \binom{\alpha + k - 1}{k} p^\alpha (1 - p)^k \quad k \in \mathbb{N}, \quad (3.35)$$

dove α è lo stesso parametro della Gamma e $p = \frac{\alpha}{\alpha + n}$ ³¹. Nel nostro caso, quindi si avrà $p = 5.260E - 4$.

³¹Si veda, ad esempio, [12] per una dimostrazione di questa affermazione.

Questa proprietà, sul piano teorico, potrebbe apparire poco importante, poiché sia la Poisson composta Gamma, sia la Binomiale Negativa sono distribuzioni a due parametri e possiamo stabilire una corrispondenza biunivoca tra i primi ed i secondi.

Tuttavia, sul piano computazionale, per condurre una simulazione Monte Carlo sulla prima è necessario coinvolgere due stadi di aleatorietà, mentre per la seconda ne è sufficiente uno solo, con notevole risparmio in termine di tempi di elaborazione.

3.3.3 Qualche precisazione sul metodo

In queste sezioni abbiamo valutato l'aspettativa rispetto al numero dei sinistri per l'anno 2012. Il modello qui sviluppato, tuttavia, non descrive solamente il numero atteso, ma replica il comportamento della variabile aleatoria \tilde{K} , e questo è un requisito fondamentale, poiché il nostro obiettivo è possibilità di adottarlo in sede di simulazione del costo aggregato dei sinistri.

Tuttavia ci potremmo chiedere se non sia una forzatura adottare questo modello per simulare il numero dei sinistri senza considerare una differenziazione dei sinistri tardivi.

Per rispondere notiamo due fattori.

	2008	2009	2010	2011
2008	234.36	174.76	197.63	42
2009		373.31	271.16	300.33
2010			336.03	317.05
2011				174.63

Tabella 3.8: Costi medi dei sinistri osservati per anno di accadimento ed anno di denuncia

In Tabella 3.8 sono riassunti i costi medi dei sinistri, con l'indicazione degli anni in cui sono avvenuti e degli anni in cui sono stati denunciati.

I sinistri tardivi appaiono avere un costo medio minore di quanto non lo abbiano i sinistri denunciati nell'anno di accadimento.

Tuttavia, ipotizziamo che questa tipologia di sinistri non abbia una forte incidenza sulla distribuzione della *severity*, poiché saranno frutto delle medesime polizze che

hanno dato origine agli altri sinistri. Possiamo quindi accettare come semplificazione quest'assunzione con riferimento ai sinistri IBNR.

Un'altra precisazione si rende necessaria in merito alla lunghezza della serie storica da adottare. Abbiamo utilizzato una serie storica costituita dai quattro anni di bilancio dei quali abbiamo le osservazioni e ci chiediamo se questo costituisca o meno un problema per le stime. Dal punto di vista strettamente statistico, sappiamo che tanto maggiore è la dimensione campionaria, tanto più accurate saranno le stime prodotte. Per poter quindi condurre una stima efficiente, in particolare della variabilità da imputare al fattore di disturbo dovremmo disporre di una serie molto più lunga.

Tuttavia, nel nostro contesto, un dato relativo ad un portafoglio vecchio di alcuni anni potrebbe essere del tutto insignificante se, ad esempio, il portafoglio della compagnia è particolarmente dinamico. Tale dato, infatti, sarebbe relativo ad un portafoglio diverso da quello del quale staremmo valutando il rischio.

L'approccio adottato, comunque è funzionale ad un approccio di tipo collettivo, alla base stessa della *ratio* del CRM. Nell'Allegato A, invece valuteremo un differente approccio.

3.4 La simulazione

Manca ora l'ultimo passo necessario per lo studio del costo sinistri aggregato: la simulazione Monte Carlo.

Come abbiamo spiegato in Sezione 3.1, il metodo Monte Carlo è necessario, poiché la distribuzione del costo aggregato dei sinistri \tilde{X}_{2012} che andremo a stimare sarà frutto della composizione di diverse variabili aleatorie, quindi il problema risulta troppo complicato per un approccio di tipo classico.

Innanzitutto, occorre simulare il numero dei sinistri dalla variabile \tilde{K}_{2012} . Non adottiamo il doppio stadio di aleatorietà simulando il valore del fattore di disturbo q dalla

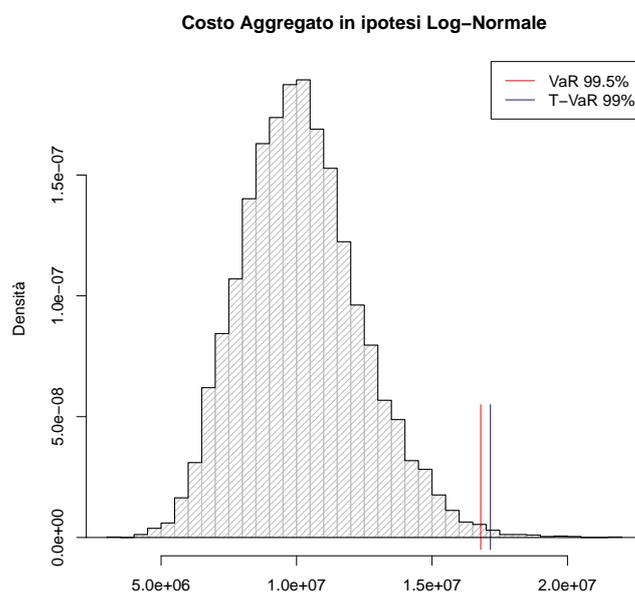


Figura 3.14: Istogramma del Costo Aggregato 2012 sotto l'ipotesi $\tilde{Z}_{2012,j} \sim \text{LNorm}$

variabile Gamma da utilizzare per misturare il processo Poisson, ma come descritto in Sezione 3.3.2, simuliamo direttamente dalla Binomiale Negativa.

Invece, per quanto riguarda il costo singolo sinistro, conduciamo tre diverse simulazioni in modo da poter confrontare i risultati. La prima ipotizzerà la distribuzione Log-Normale per $\tilde{Z}_{2012,j}$, successivamente ipotizzeremo la distribuzione Burr Inversa ed infine l'Ogiva.

Simulando 10000 possibili valori per il numero dei sinistri, otteniamo un campione sufficientemente ampio per la stima delle statistiche di nostro interesse per la distribuzione di \tilde{X}_{2012} .

Nelle Figure 3.14, 3.15 e 3.16 mostriamo le distribuzioni simulate del costo aggregato dei sinistri \tilde{X}_{2012} secondo le tre differenti ipotesi, mentre in Tabella 3.9 mostriamo i valori ottenuti per alcune statistiche di rilievo.

Anche avvalendoci dei dati ANIA, non siamo stati in grado di identificare un caricamento di sicurezza empirico, quindi i VaR saranno calcolati come differenza tra il

percentile ed il valore atteso. La stessa considerazione vale anche per i Tail-VaR.

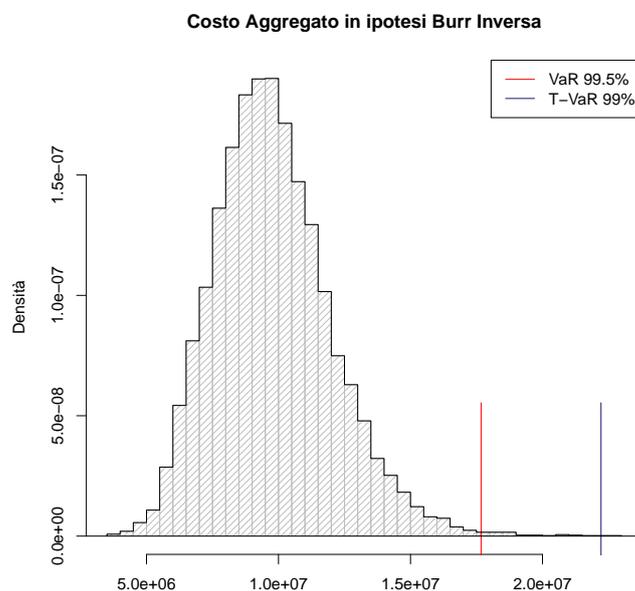


Figura 3.15: Istogramma del Costo Aggregato 2012 sotto l'ipotesi $\tilde{Z}_{2012,j} \sim InvBurr$

Le tre distribuzioni ottenute sono abbastanza simili per quanto riguarda alcuni aspetti, ma differiscono sensibilmente su altri.

Analizzando innanzitutto il valore atteso, esso appare molto simile per le prime due distribuzioni, mentre se si adotta l'Ogiva per la *severity*, si ottiene un valore nettamente più alto. Ciò è spiegabile se si tiene presente che l'Ogiva ha un buon adattamento anche per quanto riguarda la coda della distribuzione, considerando gli eventi estremi più probabili di quanto non si faccia nelle altre due ipotesi³².

Passando ai coefficienti di variabilità relativa, quelli ottenuti sotto l'ipotesi di Log-Normalità e sotto l'Ogiva appaiono tra loro abbastanza simili e molto contenuti, mentre per quanto riguarda la distribuzione ottenuta prevedendo che \tilde{Z} si distribuisca come una Burr Inversa mostra un valore significativamente differente dalle altre. Lo stesso si può

³²In Tabella 3.10, considerando i valori attesi, notiamo anche come l'unico che si modifica significativamente, risentendo del valore *outlier* sia quello calcolato sotto l'ipotesi dell'Ogiva.

dire anche osservando l'indice di asimmetria. Questo ha una spiegazione di cui parleremo nel seguito.

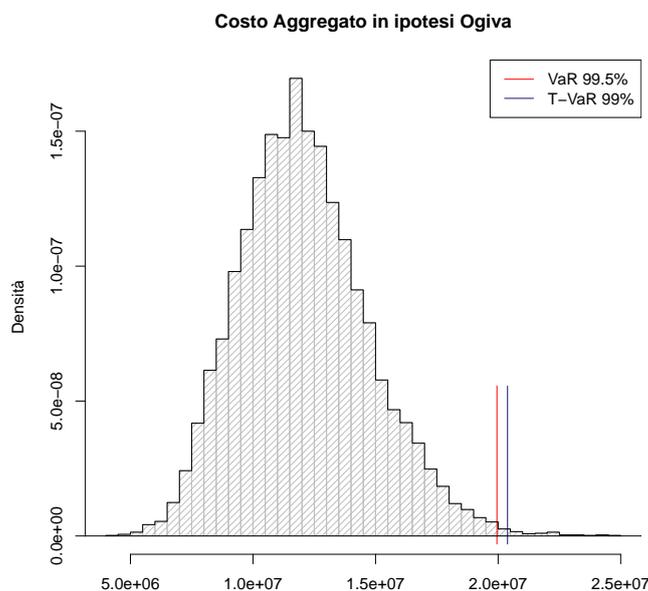


Figura 3.16: Istogramma del Costo Aggregato 2012 sotto l'ipotesi $\tilde{Z}_{2012,j} \sim Ogiva$

	Log-Normale	Burr Inversa	Ogiva
$\mathbb{E}(\tilde{X})$	10.201	9.855	12.087
$\sigma(\tilde{X})$	2.207	3.201	2.634
CoV	0.216	0.325	0.218
$\gamma(\tilde{X})$	0.456	27.254	0.460
VaR 99.5%	6.596	7.819	7.865
T-VaR 99%	6.955	12.363	8.298
T-VaR/VaR	1.054	1.581	1.055

Tabella 3.9: Statistiche significative del costo aggregato dei sinistri, valori in milioni di euro

Possiamo anche osservare come questi indicatori influenzino il calcolo del Tail-VaR: mentre nelle due distribuzioni poco asimmetriche VaR e Tail-VaR, pur discostandosi, rimangono dello stesso ordine di grandezza, nella terza la differenza è molto più marcata.

Per maggiore precisione, notiamo che osservando il rapporto tra VaR e Tail-VaR, che il secondo appare circa del 5% più alto del primo se le ipotesi sono la Log-Normale o l'Ogiva, mentre, nel caso della Burr Inversa, la differenza è quasi del 60%.

	Log-Normale	Burr Inversa	Ogiva
$\mathbb{E}(\tilde{X})$	10.211	9.887	16.994
$\sigma(\tilde{X})$	2.209	3.239	4.288
CoV	0.216	0.328	0.252
$\gamma(\tilde{X})$	0.456	27.848	0.489
VaR 99.5%	6.603	7.843	12.823
T-VaR 99%	6.962	12.513	13.547
T-VaR/VaR	1.054	1.595	1.056

Tabella 3.10: Statistiche significative del costo aggregato dei sinistri includendo l'*outlier*, valori in milioni di euro

Passiamo quindi ai confronti tra le Tabelle 3.9 e 3.10, in cui ripetiamo l'elaborazione adottando le stime ottenute includendo il valore *outlier*. Notiamo come l'unica distribuzione per cui si osservano differenze degne di nota è l'Ogiva. Aumentano, infatti, oltre al valore atteso, anche tutti gli indici di rischio, in particolare VaR e Tail-VaR.

Possiamo quindi chiederci quale sia la stima più coerente per la valutazione del rischio assicurativo tra quelle proposte.

L'Ogiva è sicuramente una soluzione accattivante, a causa della semplicità con cui può essere sviluppata e dell'estrema duttilità, dovuta al fatto di essere un modello non parametrico. Tuttavia, le assunzioni di equidistribuzione, soprattutto sulla coda destra, non hanno alcuna fondatezza e abbiamo visto che distorcono le valutazioni.

La Burr Inversa aveva un buon adattamento ai dati della nostra *severity*, ma vi è un grave pregiudizio all'affidabilità della stima. Il momento secondo, infatti, è definito solo se $\gamma > 2$, mentre altrimenti diverge. Le nostre stime sono $\hat{\gamma} = 1.482$ senza il valore *outlier*, mentre includendolo abbiamo $\hat{\gamma} = 1.481$. In entrambi i casi, quindi la varianza diverge e, di conseguenza ci possiamo aspettare che i requisiti di capitale, pur convergendo ad un valore al crescere del numero di simulazioni, assumano un valore poco

credibile. Approfondiremo questo problema nell'Appendice B³³.

Per quanto riguarda la stima sotto l'ipotesi di Log-Normalità, non abbiamo questo tipo di problemi, ma la povertà dell'adattamento ai dati empirici renderebbe rischiosa l'adozione di questa ipotesi.

Proponiamo anche un ultimo paragone: stando al CRM, attraverso alle Formule 3.5 e 3.33, possiamo ottenere una formulazione della varianza del costo sinistri aggregato. In particolare si può ottenere che

$$\mathbb{V}(\tilde{X}) = na_{2,z} + n^2 m^2 \sigma_q^2. \quad (3.36)$$

Sui nostri dati empirici, quindi, risulta $\mathbb{V}(\tilde{X}) = \text{€ } 2.434$ milioni. Questo valore ci mostra che in realtà il modello che adotta l'ipotesi distributiva della Log-Normale per il costo del singolo sinistro sottostima la rischiosità effettiva.

³³In esso infatti valuteremo la consistenza delle stime ottenute sotto le diverse ipotesi.

Capitolo 4

Un approccio innovativo per le stime

4.1 Introduzione

Nel precedente capitolo abbiamo costruito un modello per la valutazione del rischio assicurativo che permettesse un miglior adattamento ai dati a nostra disposizione, sfruttando tutte le informazioni in essi contenute. In questo modo abbiamo ottenuto un modello molto più raffinato della Standard Formula di Solvency 2 vista nel Capitolo 2, seppure più complesso.

Tuttavia, i QQ-plot mostrati nel precedente capitolo hanno rivelato come, per il costo singolo sinistro, non ci sia stato possibile ottenere un buon adattamento ad una distribuzione nota in letteratura, almeno per quanto riguarda la coda destra della distribuzione.

In questo capitolo, dopo alcune note introduttive, valuteremo la possibilità di condurre le stime attraverso approcci differenti da quelli adottati in precedenza, facendoci in particolare guidare dall'approccio presentato in [37], e produrremo delle nuove valutazioni del rischio.

4.2 Gli approcci classici

Nel precedente capitolo l'adattamento ai nostri dati di alcuni modelli parametrici noti in letteratura è stato condotto attraverso due approcci differenti.

Per la stima dei parametri dei modelli che dovevano descrivere il costo singolo sinistro ci siamo avvalsi del metodo della massima verosimiglianza, mentre per il numero dei sinistri abbiamo adottato il metodo dei momenti.

Prima di introdurre un differente approccio appare opportuno rimarcare alcuni aspetti nozionistici di quanto finora utilizzato.

4.2.1 Il metodo dei momenti

Il metodo dei momenti è uno dei metodi più elementari per la stima dei parametri di una distribuzione. Se abbiamo a nostra disposizione un campione X di n osservazioni appartenenti ad una popolazione avente, per ipotesi, una distribuzione la cui Funzione di Ripartizione sia $F(x|\underline{\theta})$, dove $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_p)$ indica il vettore dei p parametri incogniti della distribuzione, possiamo procedere alla stima di $\underline{\theta}$ come segue.

Occorre innanzitutto ricavare i primi p momenti dall'origine¹ della distribuzione, in funzione dei parametri, tali che $\mu'_k(\underline{\theta}) = \mathbb{E}(X^k|\underline{\theta})$, con $k = 1, 2, \dots, p$.

Si calcolano successivamente i primi k momenti empirici sul campione X , tali che $\hat{\mu}'_k = \frac{\sum_{j=1}^n x_j^k}{n}$.

La stima attraverso il metodo dei momenti, quindi, è quel $\hat{\underline{\theta}}$ che risolve il sistema di p equazioni in p incognite

$$\mu'_k(\underline{\theta}) = \hat{\mu}'_k \quad \text{con } k = 1, 2, \dots, p. \quad (4.1)$$

L'idea di fondo, quindi, è di costruire un modello che abbia gli stessi primi p momenti del campione osservato. Questo approccio, tuttavia, presenta diversi problemi.

¹Viene necessariamente ipotizzato che essi siano definiti per la distribuzione.

In primo luogo, i momenti di ordine superiore a p non sono in alcun modo considerati e anche scegliendo di condurre la stima adottando momenti di ordine superiore a costo di trascurarne alcuni di ordine inferiore non risolve il problema.

In secondo luogo, nel caso in cui il campione provenga da una distribuzione troncata, il metodo diviene molto più complesso ed è addirittura irrisolvibile nel caso in cui l'origine delle osservazioni sia un miscuglio di differenti popolazioni.

4.2.2 Il metodo della massima verosimiglianza

Un metodo più raffinato per la stima dei parametri di una distribuzione è il metodo della massima verosimiglianza.

È innanzitutto opportuno introdurre il concetto di funzione di verosimiglianza. Date n variabili aleatorie campionarie X_1, X_2, \dots, X_n , se ne può indicare con $f_{\underline{X}}(x_1, x_2, \dots, x_n | \underline{\theta})$ la funzione di densità (o di probabilità, se si considerano delle variabili discrete) congiunta, dato un vettore di parametri $\underline{\theta} \in \Theta$. La funzione di verosimiglianza è la stessa funzione di densità congiunta, ma vista tuttavia nell'ottica opposta a quella cui si è abituati.

Se infatti la funzione di densità mette in relazione un vettore di variabili aleatorie ad un valore corrispondente alla densità di probabilità, avendo fissato i parametri, la funzione di verosimiglianza ha per dominio Θ , mentre le variabili aleatorie sono considerate fissate. Cioè, indicando con $L(\cdot)$ la funzione di verosimiglianza,

$$L(\underline{\theta} | x_1, x_2, \dots, x_n) = f_{\underline{X}}(x_1, x_2, \dots, x_n | \underline{\theta}). \quad (4.2)$$

Nel caso in cui le variabili aleatorie X_1, X_2, \dots, X_n siano indipendenti ed identicamente distribuite e abbiano funzione di densità $f_X(x | \underline{\theta})$, vale

$$L(\underline{\theta} | x_1, x_2, \dots, x_n) = \prod_{i=1}^n f_{X_i}(x_i | \underline{\theta}). \quad (4.3)$$

Per presentare lo stimatore procediamo come in [27]. Sotto alcune opportune condizioni di regolarità della distribuzione di X , si può dimostrare che, indicato con $\underline{\theta}_0$ il valore vero del parametro,

$$P [L(\underline{\theta}_0|\underline{x}) > L(\underline{\theta}|\underline{x})] \rightarrow 1 \text{ per } n \rightarrow \infty. \quad (4.4)$$

Lo stimatore di massima verosimiglianza sarà quindi il $\hat{\underline{\theta}} \in \Theta$ tale che, una volta osservato il campione \underline{x} ,

$$L(\hat{\underline{\theta}}|\underline{x}) = \max_{\underline{\theta} \in \Theta} L(\underline{\theta}|\underline{x}). \quad (4.5)$$

L'idea che sta all'origine dello stimatore di massima verosimiglianza è che i parametri incogniti di una distribuzione debbano essere trovati in modo che assegnino la massima densità di probabilità (o probabilità) possibile alle osservazioni. In altre parole, si cercano i parametri che rendano le n osservazioni appartenenti ad una variabile casuale il più possibile verosimili.

4.3 Il *Minimum Distance Approach*

Un metodo per la stima dei parametri delle distribuzioni differente dai due finora adottati è rappresentato dal *Minimum Distance Approach*, MDA.

Il MDA è un metodo per la stima dei parametri che si pone come obiettivo la ricerca della distribuzione che abbia la forma il più possibile simile alla forma della distribuzione empirica dei dati osservati.

Questo obiettivo viene perseguito minimizzando la distanza $d(\cdot, \cdot)$ calcolata tra la funzione di ripartizione empirica del vettore delle nostre osservazioni $F_n(\underline{X})$ e la funzione di ripartizione adottata come ipotesi distributiva, con incognito parametro $\underline{\theta}$, $F^*(\underline{X}|\underline{\theta})$. Quindi lo stimatore MDA è il $\bar{\underline{\theta}}$ tale che

$$d(F_n(\underline{X}), F^*(\underline{X}|\bar{\underline{\theta}})) = \min_{\underline{\theta} \in \Theta} d(F_n(\underline{X}), F^*(\underline{X}|\underline{\theta})). \quad (4.6)$$

L'ipotesi distributiva non sarà quindi l'unica scelta che deve essere operata dal ricercatore, come accade per gli altri due approcci, ma occorre anche scegliere una opportuna funzione di distanza $d(\cdot, \cdot)$ da adottare come funzione di perdita.

Come testimoniato in [30] sono molte le funzioni di perdita adottate in letteratura. Noi ne prenderemo in considerazione inizialmente tre.

La prima di esse, di cui già abbiamo utilizzato una formulazione simile all'interno della Sezione 3.2.2.1, è la distanza di Kolmogorov ponderata. Se $K(\cdot)$ e $L(\cdot)$ indicano due funzioni di ripartizione, essa è definita come

$$d(K(\cdot), L(\cdot)) = D(K(\cdot), L(\cdot)) = \max_{a \in \mathbb{R}} |K(a) - L(a)| \cdot \psi(L(a)), \quad (4.7)$$

dove $\psi(\cdot)$ è una funzione di ponderazione².

La seconda, anch'essa già utilizzata, ma in Sezione 3.2.2.2, corrisponde alla distanza di Anderson-Darling, essendo

$$d(K(\cdot), L(\cdot)) = A^2(K(\cdot), L(\cdot)) = \int_{-\infty}^{+\infty} (K(a) - L(a))^2 \psi(L(a)) da. \quad (4.8)$$

Tale distanza, se si assume che $\psi(\cdot) \equiv 1$, diviene la statistica chiamata Cramér-Von Mises, che utilizzeremo nel seguito.

Infine, l'ultima distanza che consideriamo sarà in funzione della distanza di Minkowski di primo ordine, detta anche Distanza Manhattan:

$$d(K(\cdot), L(\cdot)) = MH(K(\cdot), L(\cdot)) = \mathbb{E}_{a \in \mathbb{R}} (|K(a) - L(a)|). \quad (4.9)$$

Un'altra decisione che lo statistico deve prendere riguarda la formulazione che si intende adottare per la funzione di ripartizione empirica in base alla quale calcolare la distanza. Questa decisione è cruciale, poiché incide direttamente sulle stime che verranno prodotte.

²Se $\psi(\cdot) \equiv 1$ si ritorna alla distanza come utilizzata nel test.

Potremmo essere portati ad adottare la classica formulazione per $F_n(x)$, che prevede

$$F_n(x) = \begin{cases} 0 & \text{se } x < x_1 \\ F_1 & \text{se } x_1 \leq x < x_2 \\ \dots & \\ F_{n-1} & \text{se } x_{n-1} \leq x < x_n \\ 1 & \text{se } x \geq x_n \end{cases},$$

dove F_i indica l' i -esima frequenza relativa cumulata.

Questa formulazione è ottima per la statistica descrittiva, ma noi ci troviamo in un'ottica inferenziale. In statistica descrittiva il fatto che $F_n(x_n)$ sia pari ad uno non rappresenta un problema, poiché ci si propone semplicemente di dare una fedele rappresentazione dei dati. In statistica inferenziale, tuttavia, significa implicitamente assumere che la probabilità che si realizzi un evento x_r maggiore di x_n sia nulla. Questo, in generale nella statistica inferenziale ed in particolare nella nostra condizione di ricerca della distribuzione del costo sinistri, costituisce un'ipotesi forte ed inaccettabile. In assenza di limitazioni contrattuali, per quanto improbabile, non possiamo assumere impossibile un sinistro di tipo catastrofe.

Pertanto modifichiamo la formulazione della ripartizione empirica per ovviare a questo inconveniente. In tutte le stime che seguiranno useremo $F'_n(x) = F_n(x) - \frac{1}{2n}$ come ripartizione, anche se la indicheremo, per semplicità, con $F_n(x)$.

Questa nuova formulazione ci permette di avere $F_n(x_n)$ minore di uno, ma al contempo $F_n(x_1)$ maggiore di zero. Oltre alla necessità teorica che ci ha condotto a questa scelta, queste due condizioni sono necessarie affinché le distanze che presenteremo in Sezione 4.5.2.2 siano calcolabili ed utilizzabili come funzioni obiettivo per una stima.

4.4 Note rispetto ai tre approcci

Lo stimatore di massima verosimiglianza gode di alcune proprietà studiate dalla statistica inferenziale che lo rendono particolarmente apprezzato a livello teorico.

Sempre in [27] troviamo le dimostrazioni di alcune di esse. Adottando alcune ipotesi di regolarità più stringenti rispetto a quelle adottate per la 4.4, si dimostra che:

1. esistono, con probabilità uno per $n \rightarrow \infty$, le soluzioni $\hat{\theta}$ che massimizzano la verosimiglianza, che convergono in probabilità al parametro θ_0 , quindi sono asintoticamente corretti;
2. se la soluzione è unica, lo stimatore è consistente;
3. se lo stimatore è consistente, $\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N\left(\underline{0}, (\mathbf{I}_n(\theta))^{-1}\right)$, cioè è anche asintoticamente efficiente;
4. se lo stimatore efficiente secondo la definizione di Rao-Cramér esiste, il sistema di verosimiglianza ha una soluzione unica coincidente con lo stimatore efficiente.

Come notato in [29], anche gli stimatori ottenuti con il metodo dei momenti godono delle proprietà asintotiche di correttezza e consistenza, anche se non necessariamente dell'efficienza.

Queste proprietà assicurano una solida base teorica all'uso degli stimatori di massima verosimiglianza, ma mal si adattano al nostro caso.

La teoria adotta lo stimatore a partire da un campione di valori osservati che provengono da una popolazione più ampia, avendo come orizzonte la stima dei parametri di quest'ultima.

Nel mondo assicurativo, tuttavia, i dati che osserviamo non rappresentano un campione proveniente da una popolazione più ampia, ma semmai la popolazione stessa. Uno stimatore più accurato potrebbe essere uno che, mantenendo inalterato il valore atteso, ricerchi una forma il più possibile simile alla vera distribuzione. In quest'ottica lo stimatore MDA acquista un ruolo privilegiato.

Inoltre, per una compagnia assicurativa, è importante condurre una valutazione del proprio rischio che sia coerente e che rispecchi quantomeno i valori osservati per quanto riguarda il costo atteso del singolo sinistro. In questo, la coerenza può essere raggiunta proprio attraverso l'imposizione di un vincolo per quanto riguarda il valore atteso.

Oltre alla considerazione precedente, che testimonia una maggior raffinatezza dello stimatore MDA sotto il profilo teorico, nel nostro contesto, possiamo notare anche che esso gode di una maggiore robustezza rispetto ai valori *outlier*.

Come descritto in [30] la maggior parte delle misure di distanza adottabili come funzioni di perdita sono fortemente sensibili alla presenza di valori *outlier*, ciononostante, anche attraverso simulazioni di tipo Monte Carlo, si può mostrare come lo stimatore MDA sia molto meno condizionato da questo fattore.

Nelle sezioni successive daremo quindi un'idea di quanto finora descritto a livello teorico riguardo alle proprietà degli stimatori MDA.

4.5 Alcune proprietà degli stimatori MDA

4.5.1 Un primo confronto

Per esporre le proprietà degli stimatori MDA iniziamo con un semplice esempio. Simuliamo un campionamento di 1000 valori da una variabile aleatoria \tilde{X} avente distribuzione Log-Normale di parametri $\mu = 0$ e $\sigma = 1$. Valore atteso e varianza di \tilde{X} saranno dunque pari rispettivamente a 1.648721 e 4.670774³.

Osserviamo tuttavia che la media campionaria assume il valore di 1.648269 e la varianza campionaria (non corretta) 4.594602.

Le stime dei parametri della distribuzione, ottenute con il metodo della massima verosimiglianza, ci restituiscono dei valori attraverso cui otteniamo che il valore atteso

³Si ricorda che se $X \sim LNorm(\mu, \sigma)$, vale

$$\mathbb{E}(\tilde{X}) = e^{(\mu + \frac{1}{2}\sigma^2)} \quad \mathbb{V}(\tilde{X}) = e^{(2\mu + \sigma^2)} \cdot (e^{\sigma^2} - 1).$$

della distribuzione dovrebbe essere 1.649884, mentre la varianza 4.673349. La stima si pone quindi tra i valori osservati ed i valori veri.

Se conduciamo invece la stima secondo il metodo MDA, otteniamo i dati riportati in Tabella 4.1. In essa riportiamo le stime di media e varianza di \tilde{X} condotte con le tre differenti funzioni di distanza presentate nella sezione precedente. Inoltre, seguendo il pensiero che sta alla base dell'approccio, conduciamo anche la stima ponendo il vincolo che il valore atteso stimato sia pari alla media campionaria.

		K-S	MH	C-vM
Liberi	$\mathbb{E}(\tilde{X})$	1.649100	1.651380	1.651317
	$\mathbb{V}(\tilde{X})$	4.660238	4.691830	4.688330
Vincolati	$\mathbb{V}(\tilde{X})$	4.647705	4.656012	4.651010

Tabella 4.1: Valori stimati attraverso il MDA

Come si può notare, i valori ricavati con il MDA con il valore atteso vincolato sono più simili ai dati della distribuzione empirica che non a quelli veri della popolazione originaria, a differenza di quanto accade con le stime di massima verosimiglianza. Quanto affermato può non apparire evidente dai pochi dati presentati⁴, quindi replichiamo lo stesso esperimento 5000 volte per campioni da 10000 osservazioni ciascuno e valutiamo gli scostamenti nelle stime di medie e varianze.

		MLE	K-S	MH	C-vM
$\mathbb{E}(\tilde{X})$	val. popolazione	4.146E-04	5.352E-04	5.239E-04	5.140E-04
	val. campione	8.484E-04	1.051E-03	9.855E-04	1.027E-03
$\mathbb{V}(\tilde{X})$	val. popolazione	3.839E-02	5.713E-02	5.737E-02	5.543E-02
	val. campione	2.544E-01	2.889E-01	3.077E-01	2.991E-01

Tabella 4.2: *Bias* tra le stime libere ed i valori veri

Nelle Tabelle 4.2 e 4.3 riportiamo il *bias* dei valori ottenuti con gli stimatori rispetto ai valori osservati sul campione e rispetto ai valori veri. Lo abbiamo calcolato, in maniera

⁴Avendo inoltre condotto una sola stima, non abbiamo alcuna significatività statistica.

un po' impropria, non sui parametri canonici della distribuzione Log-Normale, ma sui valori di media e varianza, per meglio esplicitare ciò che intendiamo. Confermando quanto già conosciamo dalla teoria, il metodo della massima verosimiglianza appare indubbiamente essere il migliore in ciascuno dei casi considerati, ma possiamo osservare come il *bias* degli stimatori MDA si avvicini a quello del ML nel caso in cui si considerino gli scostamenti dai valori campionari. Questa è una caratteristica del MDA.

		MLE	K-S	MH	C-vM
$\mathbb{E}(\tilde{X})$	val. popolazione	4.741E-04	4.741E-04	4.741E-04	4.741E-04
	val. campione	0	0	0	0
$\mathbb{V}(\tilde{X})$	val. popolazione	4.282E-02	5.719E-02	5.216E-02	5.255E-02
	val. campione	2.573E-01	2.845E-01	3.028E-01	2.975E-01

Tabella 4.3: *Bias* tra le stime vincolate e i valori veri

L'esempio finora trattato dimostra che le stime condotte attraverso il MDA, sotto il profilo della distorsione non reggono il confronto con quanto si può ottenere attraverso il metodo della massima verosimiglianza. Ciò costituisce sicuramente un limite, anche perché è sul criterio della distorsione che si basa buona parte della statistica classica, ma l'intuizione è che le stime MDA riescano a cogliere le caratteristiche del campione osservato -e non della popolazione di provenienza- meglio di quanto non si possa ottenere attraverso la massima verosimiglianza. Tuttavia ancora non possiamo esplicitare l'intuizione per cui queste caratteristiche sono convenienti nel nostro contesto. Per farlo proponiamo un confronto visivo nella sezione successiva.

4.5.2 Un secondo confronto

Abbiamo potuto osservare come lo stimatore MDA si comporti nel caso di una variabile Log-Normale con caratteristiche abbastanza *standard*. Ci chiediamo tuttavia cosa accada nel caso in cui si decida di prendere in considerazione una variabile più simile a quella oggetto del nostro studio.

Rimanendo in un esempio semplice, poiché questa sezione è volta a mostrare le proprietà del MDA prima di concentrarsi sull'effettiva valutazione del rischio, mostriamo cosa accade nel caso in cui i campioni provengano da una Gamma, ma noi ipotizziamo che la distribuzione sia Log-Normale. Scegliamo anche dei parametri per la Gamma simili a quelli stimati nel capitolo precedente, vale a dire $\alpha = 0.54$ e $\beta = 0.002$.

4.5.2.1 Robustezza

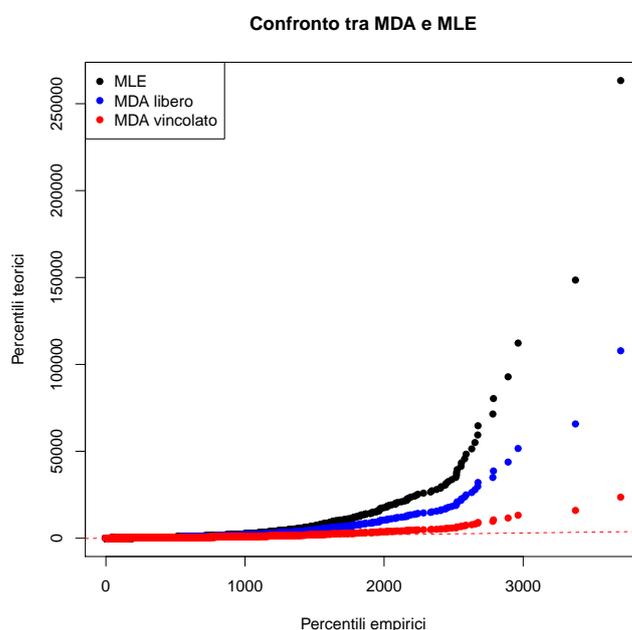


Figura 4.1: Confronto tra le stime ottenute con il MLE e con il MDA (con distanze K-S)

Il grafico ottenuto in Figura 4.1 mostra il QQ-plot che si ottiene conducendo la stima attraverso la massima verosimiglianza e il MDA, vincolato o libero, che adotta la distanza di Kolmogorov-Smirnov come funzione di perdita⁵.

L'adattamento non sembra buono in nessuno dei tre casi, come peraltro era prevedibile date le premesse adottate. Tuttavia, appare evidente come il MDA riesca a trovare

⁵Adottando una delle altre due distanze si ottengono risultati del tutto analoghi, che tuttavia non presentiamo

una distribuzione con un adattamento nettamente migliore del MLE, in particolare nel caso in cui si sia imposto il vincolo.

Osservando i risultati, possiamo quindi dedurre che il MDA ha una maggiore robustezza rispetto al MLE per quanto riguarda anche l'ipotesi iniziale sulla distribuzione. Nel nostro caso questa proprietà è di primaria importanza, poiché la distribuzione del costo sinistri non ci può essere nota a priori.

Abbiamo quindi la certezza che, se anche faremo un errore nell'identificazione del modello distributivo, il MDA ci garantirà un adattamento migliore, per la sua stessa costruzione, di altri metodi.

4.5.2.2 Alcune distanze alternative

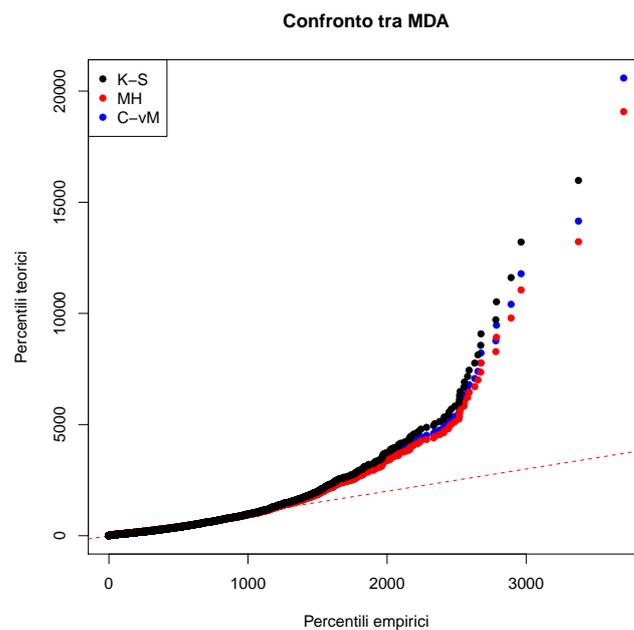


Figura 4.2: Confronto tra le stime MDA ottenute con le diverse funzioni di distanza

Proponiamo in Figura 4.2 un confronto tra gli adattamenti ottenuti con le tre distanze descritte in precedenza, rimanendo nello stesso esempio di campionamento da Gamma e *fitting* di Log-Normale.

In base al QQ-plot la stima migliore appare quella prodotta adottando la distanza Manhattan. Tuttavia non possiamo essere certi che questa affermazione sia corretta per tutti i possibili casi, ma possiamo trarre qualche considerazione dallo stesso metodo di costruzione delle distanze.

In [41] troviamo una discussione in merito alla potenza dei test di ipotesi per la valutazione dell'adattamento di alcune distribuzioni teoriche ad alcuni campioni. Tra gli altri, vengono considerati i test di ipotesi basati sulle distanze di Kolmogorov-Smirnov e sulla distanza di Cramér-von Mises.

In esso vengono riportate delle simulazioni di tipo Monte Carlo da cui si osserva come il test basato sulla distanza di Kolmogorov-Smirnov sia meno potente di quello basato sulla statistica di Cramér-von Mises.

Il risultato non ci sorprende, poiché la distanza di Kolmogorov-Smirnov prende in considerazione solo lo scostamento massimo tra le due funzioni di ripartizione e non tutti gli scostamenti.

Una statistica costruita in questo modo non permette una valutazione pienamente efficiente della distanza tra due funzioni di ripartizione, poiché non considera congiuntamente tutti gli scostamenti. Al contrario la distanza di Cramér-von Mises considera una media di tutti gli scostamenti, per questo risulta più efficiente.

In [41] non viene presa in considerazione la distanza di tipo Manhattan, ma possiamo comunque fare qualche considerazione in merito. Parimenti alla distanza di Cramér-von Mises, sintetizza tutti gli scostamenti, ma non viene adottata la potenza di secondo ordine che implicitamente assegna maggior peso agli scostamenti più rilevanti. Nella distanza Manhattan, quindi, viene data la stessa importanza a ciascuno di essi.

Tuttavia, in sede di stima dei parametri non possiamo aprioristicamente stabilire quale sia la migliore, ma possiamo aspettarci che il MDA con le distanze Cramér-von Mises e Manhattan ricercheranno una distribuzione che ben si adatta dove i dati sono concentrati, trascurando invece le code della distribuzione. Al contrario, il MDA con la distanza di Kolmogorov-Smirnov potrebbe ricercare una distribuzione che si adatti

meno bene dove i dati sono concentrati, pur avendo uno scostamento massimo minore.

Dal QQ-plot proposto in Figura 4.2 notiamo come le tre distanze, nel caso considerato producano un adattamento simile, anche se la distanza di Manhattan apparentemente restituisca una stima migliore.

Il nostro obiettivo, tuttavia, rimane quello di ottenere un modello distributivo che possa replicare la distribuzione del costo singolo sinistro al fine di valutare il rischio assicurativo, quindi abbiamo necessità di porre grande attenzione alla coda della distribuzione, attenzione che non viene adeguatamente prestata in queste stime, come visivamente si intuisce dal QQ-plot.

È quindi possibile pensare ad una differente funzione di distanza da adottare per l'applicazione del metodo. Un modo per pensare di dare maggior rilievo ai dati appartenenti alle code di una distribuzione potrebbe essere quello di adottare una distanza relativa.

Diamo un'intuizione del vantaggio che deriverebbe dall'uso di una distanza simile. Prendiamo in considerazione due Funzioni di Ripartizione continue $F(\cdot)$ e $G(\cdot)$ aventi come supporto $S^{(F)}$ e $S^{(G)}$ e ipotizziamo che

$$G(x) = \begin{cases} 0 & \text{se } F(x) \leq 0 \\ F(x) + 0.1\% & \text{se } 0 \leq F(x) \leq 99.9\% \\ 1 & \text{se } F(x) \geq 99.9\% \end{cases}.$$

Esisteranno quindi due punti $a, b \in S^{(G)} \subset S^{(F)}$ tali per cui $F(a) = 0.1\%$, $G(a) = 0.2\%$, $F(b) = 50.1\%$ e $G(b) = 50.2\%$.

Le distanze di Cramér-von Mises e Manhattan tenderanno rispettivamente a $(0.1\%)^2$ e 0.1% , assegnando lo stesso peso a tutti gli scostamenti, mentre la distanza di Kolmogorov-Smirnov varrà esattamente 0.1% . Tuttavia, se consideriamo il rapporto fra i valori assunti dalle Funzioni di ripartizione in a ed in b , notiamo che $\frac{G(b)}{F(b)} = 1.002$, mentre $\frac{G(a)}{F(a)} = 2$. In altre parole, $G(\cdot)$ assegna ai valori della coda sinistra della distribuzione una proba-

bilità di accadimento doppia rispetto a $F(\cdot)$, mentre nel corpo della distribuzione ciò appare trascurabile. Ci appare quindi opportuno prendere in considerazione anche una distanza che, basandosi sul rapporto tra le funzioni, permetta di cogliere questi aspetti.

Introduciamo quindi una prima funzione di distanza relativa, $R1$, che adotteremo poi nelle nostre stime.

$$d(F(\cdot), G(\cdot)) = R1(F(\cdot), G(\cdot)) = \mathbb{E}_{a \in \mathbb{R}} \left(\left| \ln \frac{F(a)}{G(a)} \right| \right). \quad (4.10)$$

In essa proponiamo un opportuno uso del logaritmo e del valore assoluto, in modo che sia invariante rispetto alla scelta della prima distribuzione⁶.

Una simile funzione di distanza, tuttavia, è prevedibile che assegni maggiore peso agli scostamenti della coda sinistra della distribuzione, ma concretamente noi siamo maggiormente interessati alla coda destra. Proponiamo quindi una seconda funzione di distanza $R2$ basata, invece che sulle funzioni di ripartizione, sul loro complemento ad uno, cioè sulle retro-ripartizioni. Essa sarà

$$d(F(\cdot), G(\cdot)) = R2(F(\cdot), G(\cdot)) = \mathbb{E}_{a \in \mathbb{R}} \left(\left| \ln \frac{1 - F(a)}{1 - G(a)} \right| \right). \quad (4.11)$$

Ultimo passo che facciamo, considerato che $R2$ è soggetta al medesimo problema di $R1$, seppure con riferimento all'altra coda, è la proposizione di una terza misura, $R3$, che sia la semplice media delle due precedenti

$$d(F(\cdot), G(\cdot)) = R3(F(\cdot), G(\cdot)) = \frac{R1(F(\cdot), G(\cdot)) + R2(F(\cdot), G(\cdot))}{2}. \quad (4.12)$$

Allo stesso modo, riproponendo la differenza che intercorre tra i diversi tipi di distanze finora considerati, è possibile definire anche $R4$, $R5$ e $R6$ basando ciascuno dei precedenti

⁶Infatti, utilizzando le proprietà dei logaritmi,

$$\left| \ln \frac{F(a)}{G(a)} \right| = |\ln(F(a)) - \ln(G(a))| = |\ln(G(a)) - \ln(F(a))| = \left| \ln \frac{G(a)}{F(a)} \right|.$$

sul rapporto massimo invece che sul rapporto medio.

$$d(F(\cdot), G(\cdot)) = R4(F(\cdot), G(\cdot)) = \max_{a \in \mathbb{R}} \left(\left| \ln \frac{F(a)}{G(a)} \right| \right). \quad (4.13)$$

$$d(F(\cdot), G(\cdot)) = R5(F(\cdot), G(\cdot)) = \max_{a \in \mathbb{R}} \left(\left| \ln \frac{1 - F(a)}{1 - G(a)} \right| \right). \quad (4.14)$$

$$d(F(\cdot), G(\cdot)) = R6(F(\cdot), G(\cdot)) = \frac{R4(F(\cdot), G(\cdot)) + R5(F(\cdot), G(\cdot))}{2}. \quad (4.15)$$

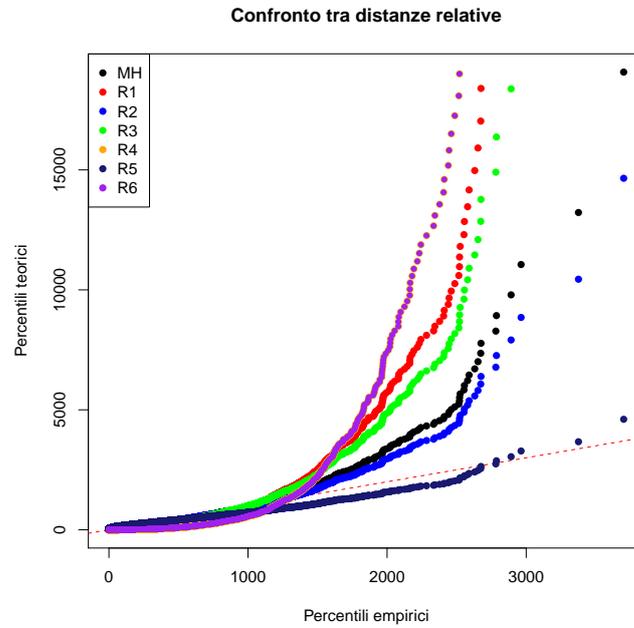


Figura 4.3: Confronto tra le stime MDA Manhattan e distanze relative

Le sei funzioni di distanza che abbiamo proposto sono strettamente correlate alla distanza di Hellinger. In [6] essa viene definita come

$$d(F(\cdot), G(\cdot)) = HD(F(\cdot), G(\cdot)) = 2\mathbb{E}_{a \in \mathbb{R}} \left(F(a)^{\frac{1}{2}} - G(a)^{\frac{1}{2}} \right)^2. \quad (4.16)$$

Pur avendo formalizzato la distanza R1 a partire da un ragionamento diverso rispetto

a quello sottostante la distanza di Hellinger, la generalizzazione delle due è la medesima:

$$\mathbb{E}_{a \in \mathbb{R}} |\Psi(F(a)) - \Psi(G(a))|^r,$$

che dipende dalla potenza r e dalla funzione monotona $\Psi(\cdot)$ applicata alle due ripartizioni considerate. $R1$ adotta $\Psi(\cdot) = \ln(\cdot)$ e $r = 1$, mentre nella distanza di Hellinger $\Psi(\cdot) = (\cdot)^{\frac{1}{2}}$ e $r = 2$. É poi possibile pensare ad una ulteriore generalizzazione includendo anche una ponderazione, ma non intendiamo in questa sede occuparci di questo.

Proponiamo quindi in Figura 4.3 un confronto tra i QQ-plot di comparazione tra i dati ottenuti con campionamento dalla variabile casuale Gamma e le distribuzioni ottenute con il MDA adottando la distanza di Manhattan e le sei distanze relative.

Ad un primo sguardo possiamo notare che le distanze $R2$ ed $R5$ mostrano un migliore adattamento ai dati anche rispetto alla distanza di Manhattan, presa come paragone. Notiamo anche che le distanze $R4$ ed $R6$ vengono ottimizzate da valori praticamente identici⁷.

Apparentemente abbiamo quindi trovato un metodo che si comporta meglio del MDA con le usuali distanze, ma in Figura 4.4 proponiamo un dettaglio dello stesso grafico precedente, in cui si mostra ciò che accade per i percentili compresi tra 0 e 1000.

Da questo particolare appare come in realtà la distanza Manhattan si comporti molto meglio per quanto riguarda il corpo della distribuzione. Ci troviamo quindi in una situazione di indecisione da cui potremo uscire solo dotandoci di un criterio univoco per la valutazione dell'adattamento dei QQ-plot.

4.5.3 La valutazione dell'adattamento delle stime

Nella Sezione 3.2.3 abbiamo verificato la bontà dell'adattamento delle stime di alcune distribuzioni note in letteratura attraverso le statistiche di alcuni test di ipotesi.

⁷Infatti $\hat{\mu}_{R4} = 1.912589$, mentre $\hat{\mu}_{R6} = 1.912572$

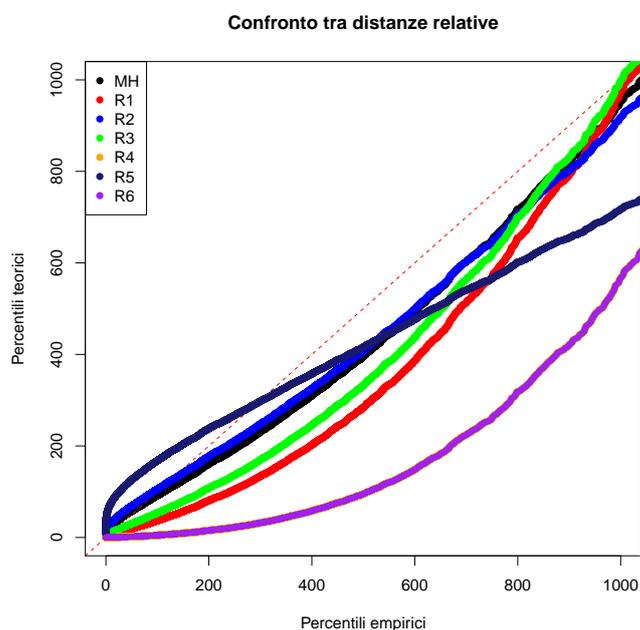


Figura 4.4: Particolare del confronto tra le stime MDA Manhattan e distanze relative

In quel contesto cercavamo, dato un certo metodo di stima, di trovare il modello distributivo che meglio si comportava, valutandolo attraverso alcuni fattori chiave.

Ora, tuttavia, fissato un certo modello distributivo, stiamo cercando di stabilire quale possa essere il miglior metodo di stima, quindi il problema si modifica sostanzialmente, in modo particolare considerando che il metodo di stima al vaglio è il MDA. Quest'approccio ci permette di dare origine, per ciascuna delle statistiche considerate, ad uno stimatore che le utilizzi come funzioni di distanza-obiettivo da minimizzare⁸. Ovviamente se valutassimo la bontà di adattamento attraverso le stesse funzioni utilizzate come funzioni di perdita, il risultato sarebbe tautologico, a meno di errori evidenti⁹.

Già nel corso Capitolo 3 avevamo adottato il QQ-plot come strumento per una prima valutazione dell'adattamento, ma qui ci proponiamo di utilizzarlo in maniera sistematica. Un primo ragionamento che si potrebbe fare prende piede dal riconoscimento di quale

⁸Abbiamo già descritto la possibilità di adottare la distanza di Kolmogorov-Smirnov come funzione di perdita.

⁹Da questo non sfugge neppure il metodo della massima verosimiglianza, poiché, come notato ad esempio in [6], anch'esso può essere visto come un caso particolare di MDA

sia il QQ-plot ideale: indiscutibilmente un QQ-plot in cui tutti i punti giacessero sulla bisettrice del primo e quarto quadrante sarebbe considerato frutto di una stima perfetta. Si potrebbe quindi condurre la regressione ai minimi quadrati ordinari condotta sui punti del QQ-plot e verificare lo scostamento di intercetta e coefficiente angolare empirici da quelli della bisettrice.

L'ipotesi è sicuramente accattivante ed indubbiamente semplice da realizzare in termini di computazione, ma pone immediatamente un dubbio, da cui ne scaturisce un secondo.

Il primo, di soluzione non banale e neppure trascurabile è se sia più corretto condurre la regressione dei percentili empirici sui teorici o viceversa. Il secondo è se effettivamente abbia senso ipotizzare una relazione di interdipendenze tra i percentili delle due distribuzioni.

Non possiamo quindi essere pienamente soddisfatti da questo primo approccio. Come seconda possibilità possiamo basarci sulla nota formula della distanza punto-retta¹⁰, calcolando le distanze tra i punti e la bisettrice. Il vantaggio, in questo caso, sarebbe quello di considerare le distanze in termini di segmenti perpendicolari alla bisettrice, quindi indipendenti dalla scelta di quale distribuzione adottare come ascissa e quale come ordinata. Anche in questo caso si apre un dubbio, una volta ottenute le distanze, su come se ne possa ottenere una sintesi. Si potrebbe infatti decidere di farne una semplice somma, ma anche la somma dei quadrati sarebbe coerente e ricondurrebbe alla distanza euclidea. Già fermandoci a queste due ci rendiamo conto che abbiamo introdotto due ordinamenti differenti, ma potremmo introdurne molti altri, ciascuno con una propria coerenza. Benché riconosciamo che avere delle alternative sia un valore, come già discusso nel Capitolo 3, in questo modo si vanno ad aprire troppe possibilità e non avremmo ragione di preferirne una piuttosto che un'altra.

L'ultimo approccio per la valutazione che consideriamo può essere visto come una

¹⁰Dato un punto P di coordinate (x_0, y_0) ed una retta r avente equazione $ax + by + c = 0$, la distanza è pari a $\frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}$

sintesi dei due precedenti. Esso prevede innanzitutto di ricalcolare le coordinate dei punti attraverso una rotazione del piano cartesiano, in modo che la bisettrice del primo e quarto quadrante vada a coincidere con l'asse delle ascisse¹¹ e successivamente condurre su di essi una regressione lineare. Tuttavia, la regressione lineare che meglio si presta al nostro scopo non è la regressione ai minimi quadrati ordinari, ma semmai la regressione ai minimi scarti assoluti.

Dalle nozioni basilari della statistica descrittiva, sappiamo che una delle proprietà di cui gode media aritmetica è di minimizzare la somma degli scarti al quadrato di una distribuzione, quindi possiamo intuire che i parametri della regressione condotta con i minimi quadrati ordinari siano in qualche modo correlati ad essa.

Al contrario sappiamo che una delle proprietà della mediana è di minimizzare la somma degli scarti assoluti di una distribuzione, funzione obiettivo della regressione che stiamo considerando.

La regressione ai minimi quadrati è uno strumento che produce stime basate indifferentemente su tutti i dati a disposizione, anche sui valori estremi. Tuttavia, nel nostro contesto, ci aspettiamo che i valori su cui vogliamo condurre la regressione siano fortemente eteroschedastici, a causa dei valori dei percentili *tail*, quindi il rischio è che le stime siano distorte.

Al contrario, la regressione che ci proponiamo di adottare gode della proprietà della robustezza rispetto ai valori *outlier* tipica della mediana, proprietà che ci garantisce una maggior efficienza delle stime.

Una volta ottenuta la regressione, attraverso il valore del coefficiente angolare e la somma degli scostamenti (assoluti) dall'asse delle ascisse potremo fare gli opportuni confronti tra i differenti metodi.

Proponiamo un semplice esempio per meglio comprendere il criterio presentato. Simuliamo tre campioni X , Y e Z , i primi due provenienti da una distribuzione Uniforme

¹¹La rotazione del piano è ottenibile attraverso la nota matrice $R = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$: un punto di coordinate (x_0, y_0) assumerà le coordinate $(x'_0, y'_0)' = R \times (x_0, y_0)'$. Nel nostro caso è necessaria l'adozione di $\theta = -\frac{\pi}{4}$.

di parametri 0 e 1 ed il terzo proveniente da una distribuzione Gaussiana con valore atteso 0 e varianza 1. Se utilizziamo la regressione lineare con l'intercetta vincolata a zero, otteniamo come coefficiente angolare $6.858E-04$ tra X e Y , mentre tra X e Z 0.353364 . Se poi osserviamo la somma dei residui al quadrato, otteniamo nel primo caso 0.08128182 , mentre nel secondo 38470.4 , quindi, come ci si poteva aspettare il criterio adottato riconosce che i primi due campioni provengono da due popolazioni più simili di quanto non lo sia quella di provenienza del terzo. Senza necessità di dilungarci in ulteriori commenti al risultato ottenuto, abbiamo mostrato come attraverso questo criterio si possa stabilire, ordinalmente, quale sia il modello o il metodo di stima che meglio si adatta.

	Libere		Vincolate	
	Coefficiente	Scostamento	Coefficiente	Scostamento
ML	0.4904	3.305E+06	-0.08318	1.024E+06
K-S	0.2685	1.799E+06	-0.06993	5.263E+05
MH	0.2673	1.696E+06	-0.06015	4.375E+05
C-vM	0.2649	1.705E+06	-0.06393	4.668E+05
R1	0.9631	2.750E+07	-0.08355	1.136E+06
R2	0.0351	4.618E+05	-0.05547	3.630E+05
R3	0.6997	5.889E+06	-0.08245	8.545E+05
R4	0.9999	1.137E+09	-0.13220	2.121E+06
R5	0.1438	1.142E+06	-0.08755	5.775E+05
R6	-0.8829	1.662E+06	-0.13220	2.121E+06

Tabella 4.4: Confronto tra i metodi di stima: coefficienti angolari e medie degli scostamenti quadri

Possiamo quindi produrre, come ulteriore esempio, la Tabella 4.4 in cui sono raccolti i coefficienti angolari e le medie degli scostamenti assoluti calcolati sulle stime ottenute per i diversi metodi, liberi e con valore atteso vincolato, per il caso prospettato nella Sezione 4.5.2.1.

Possiamo notare come le stime ottenute attraverso il metodo della massima verosimiglianza siano peggiori rispetto a quanto ottenuto con gli approcci MDA con le distanze

classiche. Allo stesso tempo notiamo anche come lo stimatore R^2 risulti particolarmente efficiente.

Pertanto, nel seguito, lo terremo in considerazione per le stime condotte sui dati veri a nostra disposizione.

4.6 La stima del costo singolo sinistro attraverso il MDA

Una volta mostrate alcune proprietà rilevanti del MDA possiamo ritornare al nostro problema di partenza conducendo nuovamente la stima della distribuzione del costo singolo sinistro.

Prima di procedere, tuttavia, decidiamo di svolgere un'ulteriore pulizia del database. Per far comprendere il motivo per cui abbiamo questa esigenza rileviamo che, ad esempio, per il 2011 osserviamo che per 1045 sinistri l'importo pagato e riservato è di € 1, mentre l'importo € 120 è stato associato a 33 sinistri nel 2008, 3484 nel 2009, 6213 nel 2010 e 210 nel 2011. In generale possiamo notare che, all'interno dei differenti anni di bilancio, i valori di alcuni importi siano particolarmente ricorrenti.

È lecito supporre che la riserva che la compagnia mette a bilancio per alcuni sinistri sia frutto di stime basate sull'esperienza maturata su sinistri di tipologia simile e non caso per caso. Questa prassi è necessaria per una compagnia, soprattutto con riferimento ai sinistri di piccolo importo per cui non è possibile disporre di una base di informazioni dettagliate prima che avvenga la liquidazione. È poi possibile, con riferimento ai sinistri di importo maggiore, che intervengano anche le limitazioni contrattuali nella definizione del valore.

Come in precedenza, non ci è possibile risalire con certezza alla causa di quanto osservato, ma dobbiamo tenerne conto nella nostra analisi statistica. Dobbiamo quindi necessariamente ipotizzare la presenza di un processo discreto associato al processo continuo del costo singolo sinistro.

Dopo aver quindi "pulito" il database da questi valori procederemo nuovamente alla

stima dei parametri di alcune distribuzioni anche secondo il metodo della massima verosimiglianza. Nel precedente capitolo avevamo selezionato la distribuzione Log-Normale e la distribuzione di Burr Inversa come le migliori per descrivere i nostri dati, le adotteremo quindi nuovamente per le stime delle sezioni successive.

4.6.1 Le stime sotto ipotesi di Log-Normalità

	Libere		Vincolate	
	Coefficiente	Scostamento	Coefficiente	Scostamento
ML	-0.0211	4.175E+06	0.0637	4.243E+06
K-S	-0.0580	4.835E+06	0.0476	3.633E+06
MH	-0.0721	5.023E+06	0.0592	4.108E+06
C-vM	-0.0666	4.947E+06	0.0543	3.880E+06
R1	-0.0121	4.212E+06	0.0674	4.308E+06
R2	-0.0330	4.385E+06	0.0476	3.621E+06
R3	-0.0270	4.338E+06	0.0647	4.272E+06
R4	0.9901	2.449E+09	0.0322	2.797E+06
R5	-0.3396	8.277E+06	-0.1904	1.473E+07
R6	0.3627	1.465E+07	0.0470	5.246E+06

Tabella 4.5: Confronto tra le stime sotto ipotesi di Log-Normalità

La Tabella 4.5 raccoglie le nuove statistiche che abbiamo deciso di considerare per valutare la bontà di adattamento dei modelli. Stando al coefficiente angolare la stima migliore è quella prodotta dallo stimatore R4 nel caso in cui si ponga il vincolo sul valore atteso, mentre nel caso in cui si lo si lasci libero da R1. Tuttavia rileviamo anche che, per quanto riguarda gli stimatori senza vincoli, a K-S, R2 ed R3 sono associati degli scostamenti simili a quelli di R1.

Si rende opportuno un altro confronto, quindi proponiamo, in Figura 4.5 il confronto tra i QQ-plot delle tre differenti stime. In esso appare evidente come i quattro stimatori producano stime pressoché identiche, nonostante le diversità di approccio, ma ognuno di essi sottostima la probabilità degli eventi *tail*, essendo i tracciati al di sotto della bisettrice.

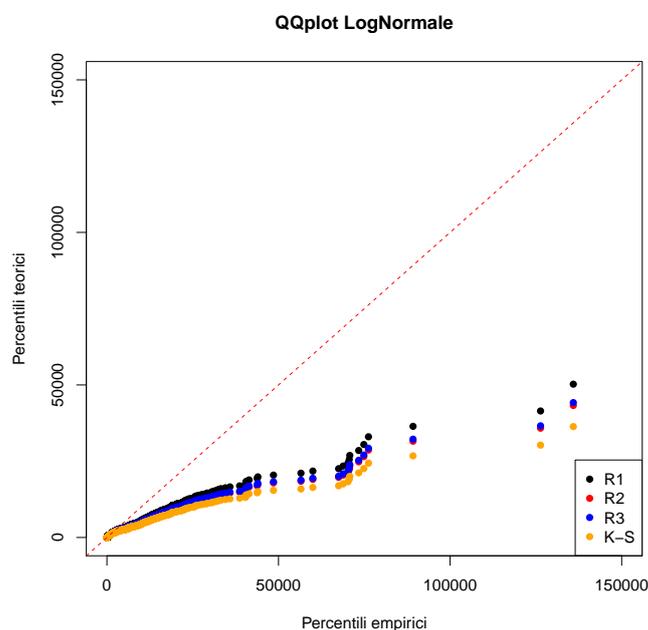


Figura 4.5: Confronto tra le stime MDA ottenute con R1, R2, R3 e K-S

Dal punto di vista invece delle stime con valore atteso vincolato, la stima migliore viene prodotta da R4, sia considerando il coefficiente angolare, sia gli scarti. Possiamo altresì notare che R2 e K-S restituiscono delle statistiche di adattamento simili, per questo proponiamo il QQ-plot di confronto in Figura 4.6.

Possiamo vedere come, almeno per i percentili inferiori a 150000, R4 si comporti effettivamente meglio degli altri.

Nelle considerazioni conclusive di questo capitolo, quindi, utilizzeremo le stime prodotte da R4 vincolato e da R1 libero per le simulazioni al fine di ottenere la stima di VaR e Tail-VaR.

4.6.2 Le stime sotto ipotesi di Burr Inversa

Come nella sezione precedente, la Tabella 4.6 raccoglie le statistiche di adattamento, ma in questo caso verifichiamo il comportamento degli stimatori sotto l'ipotesi della distribuzione di Burr Inversa.

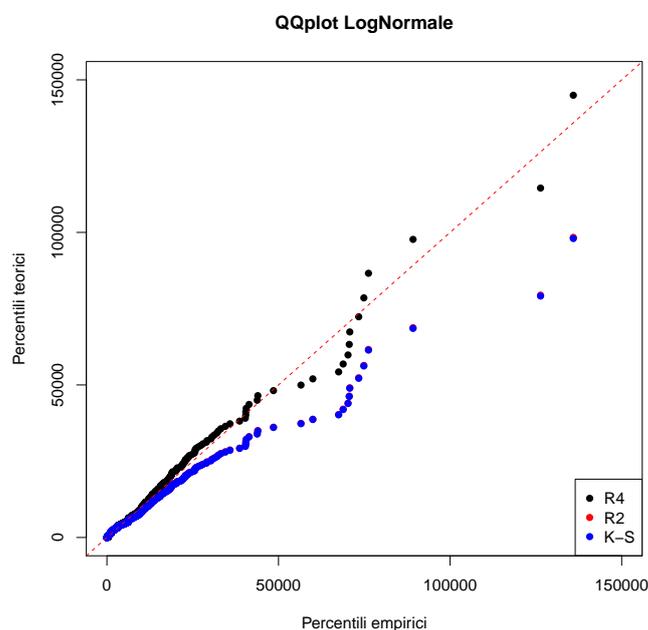


Figura 4.6: Confronto tra le stime MDA ottenute con R4, R2 e K-S vincolati

Stando al coefficiente angolare, lo stimatore MDA che adotta le distanze di Manhattan è quello che raggiunge l'adattamento migliore, ma notiamo che R6 appare migliore sotto il profilo degli scostamenti.

Proponiamo in Figura 4.7 il QQ-plot che mostra i confronti tra le due stime. Da esso possiamo vedere come il comportamento di R6 sia ben diverso da MH, ma il primo ha un adattamento migliore per la coda della distribuzione. Decidiamo quindi di adottare le stime prodotte da R6.

Dal lato delle stima con valore atteso vincolato, invece, occorre innanzitutto fare una doverosa precisazione. Il valore atteso della distribuzione di Burr Inversa, infatti, se e solo se $\gamma > 1$ vale:

$$\mathbb{E}(z) = -\frac{\theta\Gamma\left(-\frac{1}{\gamma}\right)\Gamma\left(\frac{1}{\gamma} + \tau\right)}{\gamma\Gamma(\tau)},$$

divergendo altrimenti. Pertanto in alcuni casi abbiamo dovuto porre $\gamma > 1$ come vincolo.

Le statistiche ottenute sono state riassunte nella parte destra della Tabella 4.8. La

	Libere		Vincolate	
	Coefficiente	Scostamento	Coefficiente	Scostamento
ML	0.0367	5.799E+06	-0.0659	3.033E+06
K-S	0.0731	7.460E+06	-0.0441	2.704E+06
MH	0.0053	3.481E+06	-0.0461	2.715E+06
C-vM	0.9841	2.537E+08	-0.0449	2.701E+06
R1	0.0291	5.272E+06	-0.0544	2.838E+06
R2	0.0085	4.044E+06	-0.0384	2.619E+06
R3	0.0096	4.367E+06	-0.0494	2.762E+06
R4	-0.2930	7.670E+06	0.0024	3.459E+06
R5	-0.1406	4.178E+06	-0.0418	4.233E+06
R6	0.0227	3.433E+06	-0.0034	3.163E+06

Tabella 4.6: Confronto tra le stime sotto ipotesi di Burr Inversa

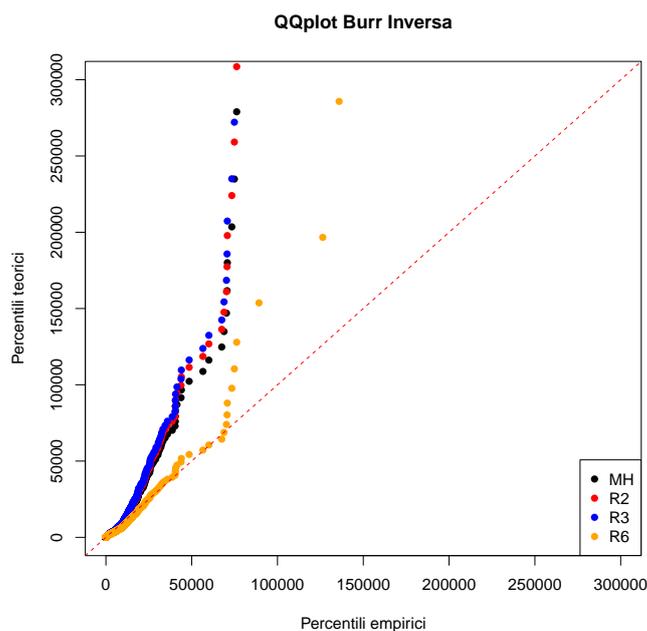


Figura 4.7: Confronto tra le stime MDA ottenute con MH, R2 ed R3 liberi

stima migliore viene ricavata dal metodo R4, osservando i coefficienti angolari, ma gli scostamenti prodotti dalle tre distanze classiche sono inferiori, ponendoci nella medesima indecisione di prima.

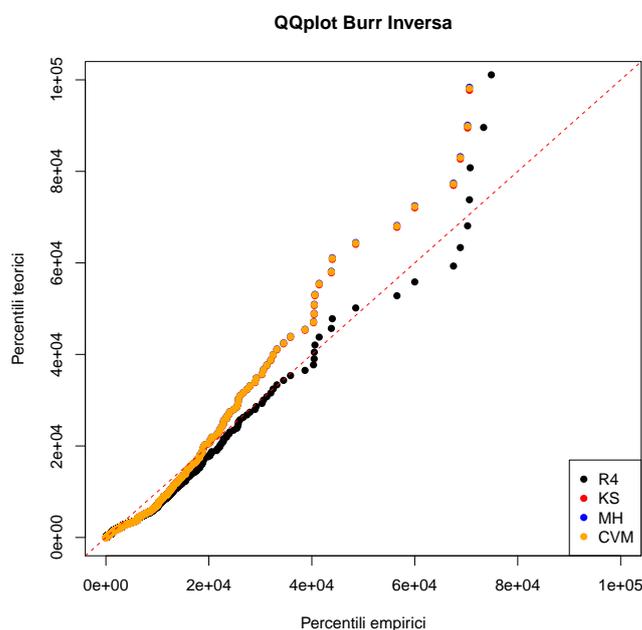


Figura 4.8: Confronto tra le stime ottenute con la massima verosimiglianza ed R5 vincolati

Proponiamo nuovamente, in Figura 4.8 un confronto tra le stime ottenute con i quattro metodi. Possiamo notare che R4 si adatta molto meglio delle altre sulla coda e il criterio del coefficiente angolare riesce a cogliere questa disparità. Al contrario, il miglior adattamento delle distanze classiche dove i dati sono concentrati, influisce sugli scostamenti medi, riducendoli. La scelta ricade comunque sulla stima prodotta da R4.

4.7 La distribuzione del costo sinistri aggregato

Anche in questo caso costruiamo la nostra distribuzione del costo aggregato dei sinistri attraverso un approccio simulativo di tipo Monte Carlo. Come nel capitolo precedente, dopo aver simulato il numero dei sinistri, simuliamo per ciascuno di essi anche il costo a partire da una delle quattro ipotesi che abbiamo deciso di adottare nella sezione precedente.

Due sono le differenze, oltre alle ipotesi sulla distribuzione del singolo sinistro, che

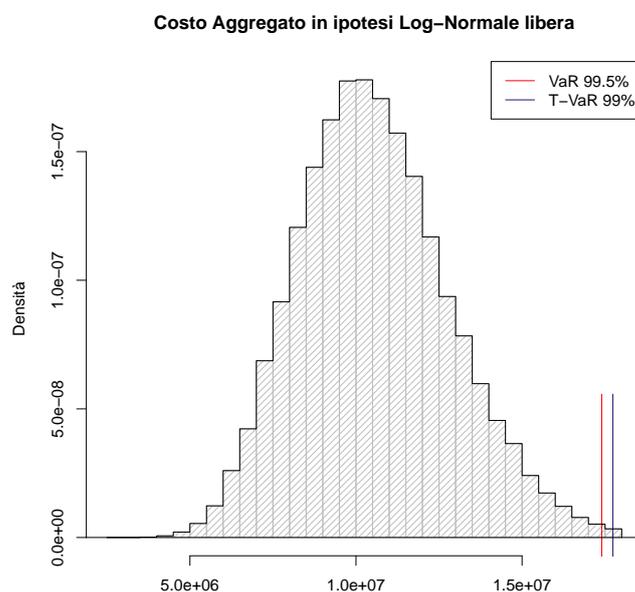


Figura 4.9: Distribuzione del costo sinistri aggregato

tuttavia differenziano questa simulazione dalla precedente:

- il numero di replicazioni;
- la presenza di un secondo processo, oltre a quello continuo, nella definizione del costo sinistri.

Nel precedente capitolo l'algoritmo che ci permetteva di simulare il costo sinistri a partire dall'Ogiva non era nativo del *software* che abbiamo a disposizione, ragion per cui le replicazioni occupavano un tempo macchina rilevante, condizione che ci aveva fatto decidere di produrne solo 10000. Per ragioni di coerenza avevamo poi deciso di adottare lo stesso limite anche per gli altri casi considerati. In questo caso, invece, tutte le funzioni che consideriamo sono interne, quindi abbiamo deciso di aumentare il numero di simulazioni portandolo a 100000.

Nella Sezione 4.6 avevamo descritto e posto rimedio ad un problema di coerenza interna dei dati. Siccome la pulizia riguarda circa il 30% dei nostri sinistri ogni anno,

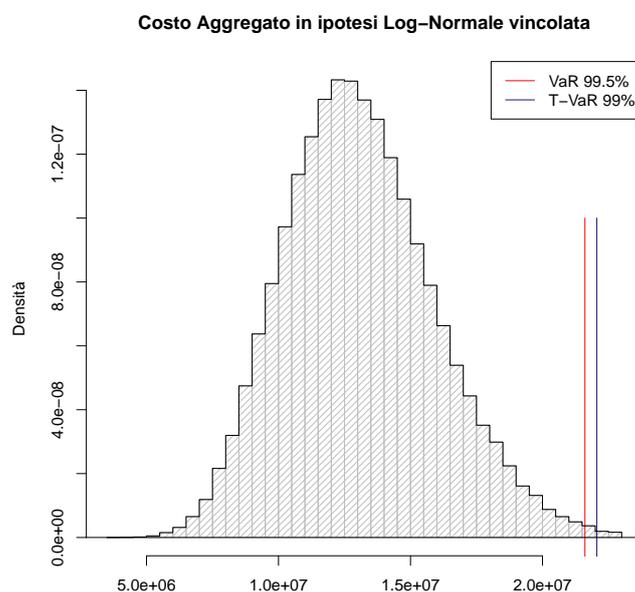


Figura 4.10: Distribuzione del costo sinistri aggregato

essa non è assolutamente trascurabile. Per rendere nella simulazione questo problema decidiamo di campionare casualmente, direttamente dal vettore degli importi che sono stati esclusi, il 30% dei sinistri, generando il restante 70% dalle distribuzioni oggetto delle nostre quattro ipotesi.

Otteniamo quindi i dati raccolti in Tabella 4.7. Vediamo come, nel caso in cui si vincoli il valore atteso del costo singolo sinistro, il valore atteso del costo aggregato sia maggiore e con esso anche la valutazione del rischio.

Allo stesso tempo, notiamo che la simulazione condotta adottando la distribuzione di Burr Inversa per il costo singolo sinistro porta a delle stime della variabilità e dell'asimmetria molto maggiori rispetto a quanto prospettato nell'ipotesi di Log-Normalità. La causa è la medesima rilevata nel precedente capitolo.

Benché, infatti, la condizione $\gamma > 1$ sia sufficiente per assicurare che il valore atteso del costo singolo sinistro esista finito nell'ipotesi di distribuzione di Burr Inversa, se passiamo al momento secondo la condizione diviene $\gamma > 2$. Non essendo quest'ultima

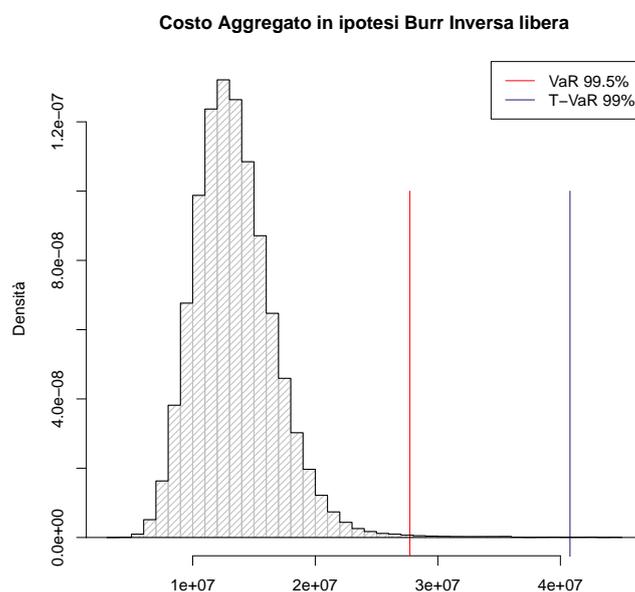


Figura 4.11: Distribuzione del costo sinistri aggregato

rispettata¹², $\sigma(\tilde{X})$ e CoV divergono, trascinando le valutazioni del rischio a valori privi di senso. Questo problema non è assolutamente risolvibile attraverso l'imposizione, come vincolo, $\gamma > 2$, poiché gli algoritmi restituiscono soluzioni d'angolo. Considerazioni analoghe possono essere fatte anche per quanto riguarda l'indice di asimmetria.

Al contrario, le stime prodotte sotto ipotesi di Log-Normalità sono comparabili con quelle prodotte in precedenza. Avendo noi incluso nella nostra stima anche il valore *outlier*, esse sono confrontabili con quelle riportate in Tabella 3.10.

In particolare notiamo che la stima condotta attraverso lo stimatore libero restituisce una valutazione della rischiosità perfettamente in linea con la precedente, con delle lievi differenze trascurabili.

Questo ci lascia molto sorpresi poiché, per questa nuova stima abbiamo condotto la simulazione attraverso un processo con differenze di rilievo rispetto a quanto praticato nel capitolo precedente.

¹²Le nostre stime sono $\hat{\gamma} = 1.2908$ e $\hat{\gamma} = 1.2432$ rispettivamente per la stima libera e vincolata.

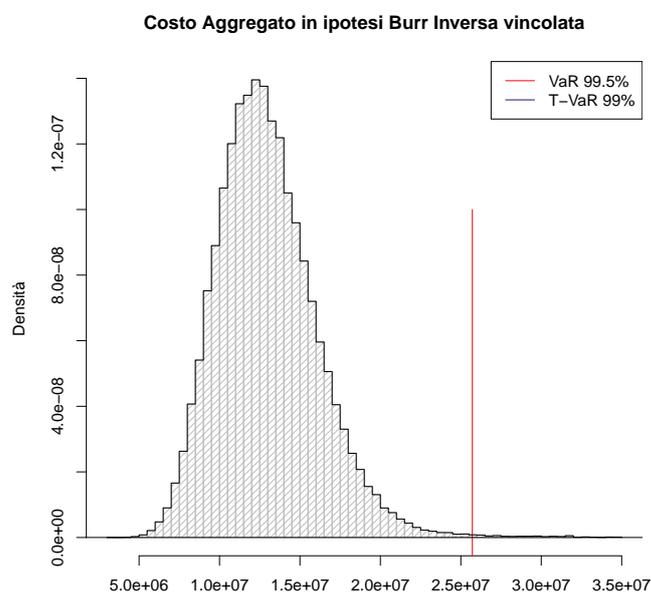


Figura 4.12: Distribuzione del costo sinistri aggregato

Pur avendo infatti introdotto un secondo processo, come descritto in precedenza, abbiamo ottenuto una differenza di circa € 350000 per quanto riguarda il valore atteso, ma di soli € 200000 circa per VaR e Tail-VaR. Oltre alla variabilità, infatti, anche l'indice di asimmetria si avvicina particolarmente a quello del capitolo precedente, discostandosi di appena 0.02.

Passando poi al confronto tra la distribuzione ottenuta nel precedente capitolo e la distribuzione con il vincolo sul valore atteso del costo singolo sinistro ciò che ci lascia sorpresi è innanzitutto la dimensione della differenza tra i valori attesi delle due distribuzioni, essendo di quasi tre milioni di euro.

Anche la rischiosità assoluta appare maggiore, ma il CoV rimane pressoché inalterato, essendo appunto il valore atteso maggiore. Tuttavia la nuova distribuzione risulta meno asimmetrica.

Notiamo infine che VaR e Tail-VaR sono maggiori in modo rilevante, essendo maggiori rispettivamente del 28.6% e del 28.4% se confrontati alle precedenti stime.

	Log-Normale		Burr Inversa	
	Libera	Vincolata	Libera	Vincolata
$\mathbb{E}(\tilde{X})$	10.583	13.117	13.626	13.112
$\sigma(\tilde{X})$	2.301	2.874	9.002	7.543
CoV	0.217	0.219	0.661	0.575
$\gamma(\tilde{X})$	0.433	0.435	115.454	106.613
VaR 99.5%	6.816	8.492	14.092	12.604
T-VaR 99%	7.153	8.943	27.157	23.432
T-VaR/VaR	1.049	1.053	1.927	1.859

Tabella 4.7: Statistiche significative del costo aggregato dei sinistri, valori in milioni di euro

Lo stesso paragone, con le medesime sottolineature, può essere presentato tra le stime condotte con e senza il vincolo sul valore atteso prodotte in questo capitolo, essendo la nuova stima senza vincolo del tutto affine alla precedente.

Infine, osservando i rapporti tra VaR e Tail-VaR, notiamo che, sotto l'ipotesi di Log-Normalità, sono pressoché identici a quelli ottenuti nel capitolo precedente, mentre per la Burr Inversa il divario si amplifica, divenendo quasi doppio.

Capitolo 5

Alcuni modelli composti

5.1 Introduzione

Il passo successivo che compiamo all'interno del nostro elaborato riguarda nuovamente la ricerca di una distribuzione del costo del singolo sinistro che meglio possa replicare i dati a nostra disposizione.

Nel Capitolo 3 abbiamo adottato il classico metodo della stima di massima verosimiglianza come punto di partenza, ma non siamo stati in grado, pur confrontando una grande varietà di soluzioni, di ottenere un buon *fitting* per i dati del costo singolo sinistro.

Successivamente, nel Capitolo 4 abbiamo valutato la possibilità di utilizzare un metodo differente per la stima e ci siamo orientati verso il *Minimum Distance Approach*. Tuttavia, non abbiamo ancora ottenuto un modello che sia in grado di cogliere la coerenza dei dati replicando accuratamente sia quelli di piccolo importo sia quelli di diverse migliaia di euro.

In questo capitolo proveremo a raggiungere questo obiettivo costruendo noi stessi una distribuzione, a partire da distribuzioni note in letteratura.

Nella parte iniziale di questo capitolo assumeremo che la distribuzione del costo singolo sinistro sia in realtà una mistura di due differenti distribuzioni aventi il medesimo supporto.

Successivamente ci concentreremo su un modello di *spliced distribution*, dove cioè le distribuzioni di partenza sono tronche e coprono supporti disgiunti.

Chiuderemo il capitolo ritornando alla valutazione del rischio attraverso l'usuale approccio simulativo.

5.2 La distribuzione misturata

5.2.1 L'impostazione teorica

Fino ad ora abbiamo ricercato una distribuzione del costo singolo sinistro nell'ipotesi che i dati su cui conduciamo la nostra analisi fossero coerenti gli uni con gli altri. Abbiamo quindi ricercato una distribuzione che li descrivesse, provando con casi parametrici e non e con diversi metodi per la stima.

Qui, in un certo senso, ci comportiamo come se facessimo cadere questa ipotesi. In realtà questa affermazione è vera solo in apparenza e ne vedremo il motivo.

Un manuale che si accosta alla teoria delle distribuzioni misturate è [28] anche se non vengono viste con diretto riferimento al contesto assicurativo. In esso ritroviamo una formalizzazione del problema che qui riproponiamo.

Si considera una popolazione X costituita da m sottopopolazioni, le componenti della popolazione, indicizzate come $j = 1, 2, \dots, m$. Estruendo un campione di dimensione n dalla popolazione X , possiamo ottenere i dati (X_i, J_i) , dove $i = 1, 2, \dots, n$ e J_i indica il gruppo di appartenenza di X_i .

Se si estrae un campione dalla j_0 -esima componente, si potrà chiaramente indicare con $f(x|J = j_0)$ la funzione di densità condizionata e, a partire da questa si può definire la densità congiunta attraverso il teorema di Bayes, risultando

$$f(x, J = j) = f(x|J = j) \cdot P(J = j). \quad (5.1)$$

Quanto finora presentato non costituisce nulla di nuovo, ma il problema del modello

misturato sorge se non si può disporre di informazioni riguardo al gruppo J ovvero se si cerca una distribuzione della variabile X senza vincoli riguardo la componente di appartenenza.

La distribuzione di X sarà quindi

$$f(x) = \sum_{j=1}^m f(x|J=j) \cdot P(J=j), \quad (5.2)$$

ed essendo $\sum_{j=1}^m P(J=j) = 1$, essa rappresenta il valore atteso delle $f(x|J=j)$.

Quanto presentato fino ad ora rappresenta un modello in cui J è discreto e finito, ma il ragionamento è il medesimo anche nel caso in cui sia continuo, sostituendo la sommatoria con l'integrale e la funzione di probabilità con la densità. Allo stesso modo, possiamo anche considerare una popolazione X discreta solo scambiando la funzione di densità con la probabilità.

Possiamo rileggere con questo sguardo i modelli che abbiamo vagliato trattando del numero dei sinistri. Avevamo adottato una popolazione X discreta rappresentante il numero di sinistri generati dall'aggregato delle nostre polizze e una suddivisione in un numero infinito di componenti, ottenuta attraverso l'uso di un fattore di disturbo continuo.

Ci eravamo quindi posti nella condizione di sottopopolazioni aventi ciascuna una distribuzione condizionata appartenente alla medesima famiglia, avendo come unica differenza il parametro, modificato da una variabile aleatoria continua.

Tornando ad un esempio generale, possiamo anche pensare al caso in cui ad essere differenti tra le sottopopolazioni non siano solo i parametri, ma lo stesso modello distributivo.

Ciò che potrebbe destare qualche dubbio è l'ipotesi che viene fatta a livello teorico rispetto alla provenienza dei dati da diverse sottopopolazioni. Questo appare privo di senso, poiché il database contiene le statistiche del solo ramo malattia, quindi i dati dovrebbero avere una certa coerenza interna.

Quello che assumiamo, quindi è che ogni sinistro possa comportarsi, con una certa probabilità, secondo una delle distribuzioni che prenderemo come ipotesi. In un certo senso è come se stessimo adottando semplicemente una nuova ipotesi distributiva ottenuta dalla mistura di diverse popolazioni.

Le conclusioni e la metodologia di stima, comunque, sono i medesimi dell'approccio di [28], ma vi è questa sottile differenza a livello teorico.

Altrimenti, nuovamente tornando all'ipotesi della suddivisione in popolazioni, potremmo pensare che vi siano più popolazioni a seconda della tipologia del sinistro. Tuttavia, in questo caso, troveremmo più affine un differente approccio, che vedremo nella Sezione 5.3.

In questa sezione osserveremo quindi come si adatta un modello in cui X è continuo, dovendo adattarsi alla variabile aleatoria costo singolo sinistro, mentre J è discreto. In particolare scegliamo un modello distributivo costruito come la mistura di due distribuzioni Log-Normali, lasciando incogniti e quindi cercando di stimare sia la proporzione sia i parametri.

Quindi, se $f_1(x|\mu_1, \sigma_1)$ e $f_2(x|\mu_2, \sigma_2)$ sono due funzioni di densità di due variabili casuali Log-Normali e se π è la proporzione di dati provenienti dalla prima componente, la funzione di densità della mistura sarà

$$f(x|\mu_1, \sigma_1, \mu_2, \sigma_2, \pi) = \pi \cdot f_1(x|\mu_1, \sigma_1) + (1 - \pi) \cdot f_2(x|\mu_2, \sigma_2). \quad (5.3)$$

Partendo da questa è possibile definire la funzione di ripartizione come

$$\begin{aligned} F(x|\mu_1, \sigma_1, \mu_2, \sigma_2, \pi) &= \int_0^{+\infty} f(x|\mu_1, \sigma_1, \mu_2, \sigma_2, \pi) dx = \\ &= \int_0^{+\infty} (\pi \cdot f_1(x|\mu_1, \sigma_1) + (1 - \pi) \cdot f_2(x|\mu_2, \sigma_2)) dx = \pi \cdot F_1(x|\mu_1, \sigma_1) + (1 - \pi) \cdot F_2(x|\mu_2, \sigma_2). \end{aligned} \quad (5.4)$$

Vediamo quindi come sarà possibile la stima dei parametri di questa distribuzione.

5.2.2 Massima Verosimiglianza o *Minimum Distance Approach*

Nuovamente dobbiamo considerare questa domanda che già avevamo proposto nel capitolo precedente.

Abbiamo potuto infatti osservare, attraverso due semplici esempi, che lo stimatore ML è il migliore in termini di efficienza nel caso in cui l'ipotesi distributiva iniziale sia corretta, mentre il MDA si comporta meglio se l'ipotesi è errata.

Entrambi i risultati erano prevedibili, date le note proprietà dello stimatore ML e data la stessa idea che ha guidato alla formulazione del MDA, ma non si poteva essere aprioristicamente certi del secondo.

Anche in questo caso vogliamo vedere come si comportano i due approcci poiché nutriamo qualche dubbio sulle effettive possibilità che il MDA sia efficace.

Innanzitutto rileviamo che entrambi sono applicabili, potendo contare sulla Formula 5.3 per la computazione della densità, necessaria allo stimatore ML, e sulla Formula 5.4 per la ripartizione, utile per il MDA. Inoltre il valore atteso sarà

$$\mathbb{E}(\tilde{Z}) = \pi \cdot e^{(\mu_1 + \frac{1}{2}\sigma_1^2)} + (1 - \pi) \cdot e^{(\mu_2 + \frac{1}{2}\sigma_2^2)}, \quad (5.5)$$

che si può facilmente scrivere in funzione, ad esempio, di μ_1 ottenendo

$$\mu_1 = \ln \left(\frac{\mathbb{E}(\tilde{Z}) - (1 - \pi) \cdot e^{(\mu_2 + \frac{1}{2}\sigma_2^2)}}{\pi} \right) - \frac{1}{2}\sigma_1^2. \quad (5.6)$$

Grazie quindi alla Formula 5.6 possiamo anche produrre le stime vincolate senza necessità di complicare ulteriormente il modello.

Considerato quindi che si possono ottenere gli stimatori, ciò su cui vogliamo indagare è se esista una dimensione campionaria oltre la quale lo stimatore MDA risulti inadeguato. Temiamo infatti che, per campioni di grandi dimensioni, lo stimatore MDA non riesca a cogliere e a definire le due distribuzioni oggetto della mistura. Procederemo quindi nel seguito attraverso due verifiche.

5.2.2.1 La prima verifica

L'esperimento che decidiamo di condurre si articola come segue:

1. si campionano pseudo-casualmente n valori da una mistura di due Log-Normali con parametri noti;
2. si conducono le stime dei parametri attraverso i metodi che si intende valutare;
3. si replica la stima k volte;
4. dopo aver deciso la dimensione dello *step* s si ritorna al primo punto e ricomincia sostituendo $n + s$ ad n e si ripete fino al raggiungimento della dimensione campionaria obiettivo;
5. si calcola la distorsione rispetto ai parametri noti.

Questo semplice schema ci permette di verificare la possibilità che gli stimatori possano rivelarsi inadeguati al crescere della dimensione campionaria.

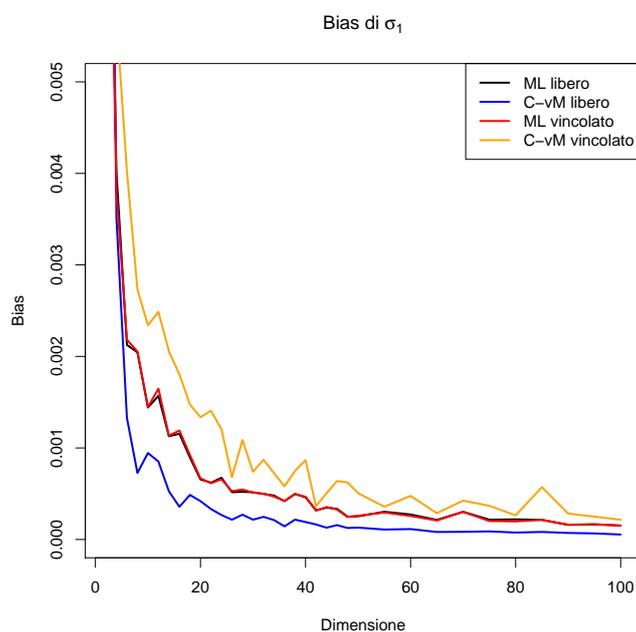
Come abbiamo visto nel precedente capitolo, il metodo migliore per valutare le *performances* di uno stimatore MDA, è un metodo non parametrico basato sul QQ-plot. Esso, infatti, ben si adatta alla filosofia sottostante all'approccio, dato che valuta l'adattamento alla forma della distribuzione.

Tuttavia questo non ci è possibile quando si produce una così grande quantità di stime, quindi ci dobbiamo affidare al classico *bias*.

In particolare decidiamo di confrontare le *performances* dello stimatore C-vM al confronto con lo stimatore ML, sia ponendo il vincolo per quanto riguarda il valore atteso, sia lasciandolo libero.

Restano solo da definire i valori da assegnare ai parametri della distribuzione e ai parametri di controllo dello schema che ci siamo prefissi.

Abbiamo scelto i primi affinché le componenti che costituiscono la mistura siano distanti e riconoscibili, ma non troppo e che siano entrambe numerose. Abbiamo quindi scelto $\pi = 0.75$, $\mu_1 = 1$, $\sigma_1 = 1.5$, $\mu_2 = 4$ e $\sigma_2 = 1.5$.

Figura 5.1: Bias delle stime di σ_1

Data la dimensione del nostro dataset, siamo preoccupati che la soglia critica dopo la quale lo stimatore MDA risulti inefficiente sia al di sotto di 100000 unità statistiche, pertanto questo valore diviene la nostra dimensione obiettivo.

Certamente scegliere un k abbastanza elevato darebbe un'ottima significatività statistica all'esperimento ed uno step s piccolo ci permetterebbe di individuare con precisione la soglia critica, se ne esiste una.

Tuttavia, il tempo macchina per un esperimento simile è davvero molto rilevante, quindi abbiamo dovuto limitare il numero delle replicazioni ed il numero di step.

Una sola di queste replicazioni necessita infatti di diversi passaggi che per un computer risultano particolarmente onerosi. Non ci riferiamo solo alla generazione di valori pseudo-casuali, che in questo caso rappresenta il problema minore, ma agli stessi algoritmi di stima.

Per tutte le stime, infatti, occorre ottimizzare una funzione non invertibile a quattro o cinque parametri¹, ricorrendo quindi ad algoritmi iterativi. In particolare, essendo necessariamente $0 \leq \pi \leq 1$ dobbiamo anche ricorrere ad un algoritmo che ci consenta di imporre dei vincoli sui parametri e la scelta ricade su quello proposto in [9]. Questo algoritmo di ottimizzazione, chiaramente, risulta via via sempre più oneroso al crescere della dimensione campionaria, richiedendo tempi sempre maggiori per l'ottimizzazione.

Decidiamo quindi di adottare $k = 100$ ed $n = 2000$, crescendo con uno step s pari a 2000 fino al raggiungimento della dimensione campionaria di 50000, passando poi ad uno step di 5000 fino alla soglia obiettivo.

Il risultato fugge i nostri dubbi circa l'esistenza della soglia critica al di sotto della dimensione campionaria di 100000 unità. Proponiamo, a titolo esemplificativo, il grafico riportato in Figura 5.1 che rappresenta l'andamento del bias al crescere della dimensione campionaria.

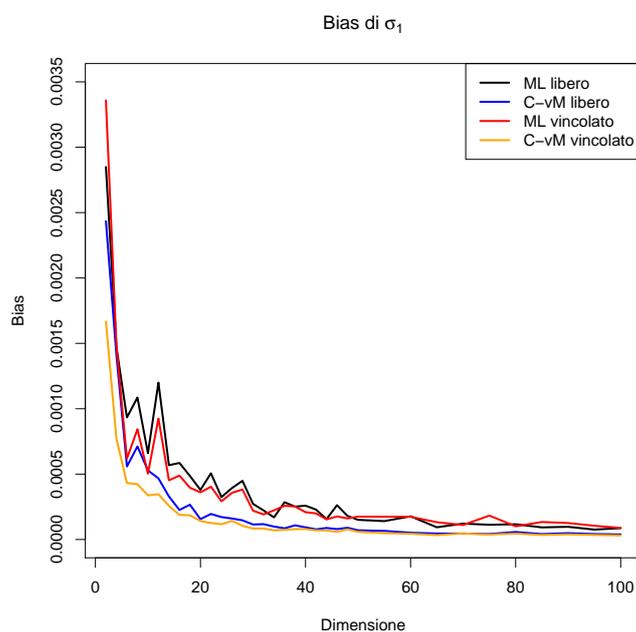
Possiamo infatti notare che il bias tende ad annullarsi per n crescente per tutti i metodi considerati. Rimaniamo sorpresi dal fatto che la distorsione delle stime appaia inferiore per lo stimatore C-vM senza vincoli rispetto ad entrambi gli stimatori ML, ma possiamo sospettare che questo sia dovuto all'algoritmo di stima che evidentemente è più potente per l'ottimizzazione della distanza C-vM rispetto alla (log)verosimiglianza.

Si tratta sicuramente di un buon metodo per il nostro scopo, ma possiamo anche chiederci il motivo per cui non abbiamo riscontrato la criticità. Le risposte più verosimili possono essere:

- l'inesistenza di tale soglia;
- l'esistenza della soglia, ma solo oltre $n = 100000$.

In entrambi i casi potremmo allora riproporre come in precedenza lo stimatore MDA, poiché questo limite non rappresenta un problema per noi.

¹Quattro se la stima è vincolata e cinque se è libera.

Figura 5.2: Bias delle stime di σ_1

Tuttavia, possiamo anche supporre che la causa di entrambe le risposte possibili sia da ricercarsi nella definizione dei parametri iniziali della popolazione. Le due sottopopolazioni, come abbiamo detto, sono abbastanza distanti² e abbastanza numerose. Questo potrebbe incidere rendendo il problema troppo semplice da risolvere e quindi eliminando o spostando indefinitamente in là quello che potrebbe essere un limite naturale.

5.2.2.2 La seconda verifica

Date le premesse fatte alla prima verifica, riproponiamo lo stesso percorso, ma adottiamo dei differenti parametri.

Quanto ottenuto nella sezione precedente può rappresentare una calibrazione dell'algoritmo che quindi siamo pronti ad utilizzare per mettere alla prova gli stimatori in una situazione di *stress*.

²Stando ai parametri adottati, il valore atteso della prima è 8.37, mentre per la seconda 168.17.

Scegliamo quindi come parametri $\pi = 0.95$, $\mu_1 = 4$, $\sigma_1 = 1.5$, $\mu_2 = 7$ e $\sigma_2 = 1.5$, lasciando invariati i parametri di controllo dello schema.

Possiamo vedere chiaramente che le due sottopopolazioni sono ben più distanti che in precedenza, ma una delle due risulta molto più piccola dell'altra.

Inoltre, per rendere più stabile l'algoritmo riparametrizziamo le nostre funzioni di distanza in funzione di $\text{logit}(\pi)$, μ_1 , $\log(\sigma_1)$, μ_2 , $\log(\sigma_2)$, in modo che i nuovi parametri non siano soggetti a vincoli.

Il risultato è del tutto simile al precedente e mostriamo, come in precedenza, il solo grafico del bias delle stime del parametro σ_1 .

Dal grafico possiamo dedurre che gli stimatori vincolati sono meno distorti di quelli liberi e che, come in precedenza, non vi è alcuna evidenza empirica della presenza della soglia critica con i parametri adottati.

Pertanto, possiamo passare alla stima dei parametri della distribuzione della *severity* attraverso questi stimatori.

5.2.3 Le stime

Conduciamo quindi le stime dei parametri della distribuzione del costo singolo sinistro. Lo facciamo adottando come stimatori quello di massima verosimiglianza, il MDA basato sulla distanza di Cramer-von Mises, R2 ed R5, così come presentati nella Sezione 4.5.2.2.

	π	μ_1	σ_1	μ_2	σ_2
ML	0.286	4.65	2.07	4.41	1.33
C-vM	0.499	4.4	2.03	4.53	1.09
R2	0.475	4.39	2.04	4.54	1.11
R5	0.423	-5.33	4.62	4.96	1.63

Tabella 5.1: Stime vincolate del miscuglio di Log-Normali

Ciò che otteniamo è quanto riportato nella Tabella 5.1 con riferimento alle stime vincolate, mentre le stime libere sono in Tabella 5.2.

	π	μ_1	σ_1	μ_2	σ_2
ML	0.544	4.43	1.23	4.54	1.91
C-vM	0.339	4.59	1.02	4.39	1.79
R2	0.804	4.38	1.31	5.09	2.08
R5	0.269	-1.66	2.58	5.05	1.67

Tabella 5.2: Stime libere del miscuglio di Log-Normali

Possiamo osservare che le stime ottenute imponendo il vincolo sul valore atteso sono tutte più affini tra loro di quanto non lo siano le stime libere.

Vediamo che tra esse si discosta la stima dei parametri della prima componente ottenuta con R5 e la stima di π ricavata con la massima verosimiglianza. Tuttavia, dato che le due componenti sono molto simili possiamo immaginare che la stima ML non sia molto distante dalle successive due.

Conducendo invece le stime senza il vincolo sul valore atteso, invece, otteniamo dei valori molto diversi gli uni dagli altri.

Notiamo comunque che nessuno stimatore ha prodotto una stima simile a quelle che abbiamo considerato nella sezione precedente, in cui avevamo ipotizzato una componente molto più piccola e distante dall'altra.

Procediamo quindi alla valutazione di quanto ottenuto come abbiamo descritto nella Sezione 4.5.3.

	Libere		Vincolate	
	Coefficiente	Scostamento	Coefficiente	Scostamento
ML	0.00324	1.094E+06	0.01357	1.795E+06
C-vM	-0.02222	1.949E+06	0.00503	1.303E+06
R2	0.00772	2.549E+06	0.00592	1.295E+06
R5	0.14153	4.281E+06	-0.03551	3.347E+06

Tabella 5.3: Confronto tra le stime sotto ipotesi di mistura di Log-Normali

La Tabella 5.3 raccoglie le statistiche d'adattamento delle nuove stime. La lettura di

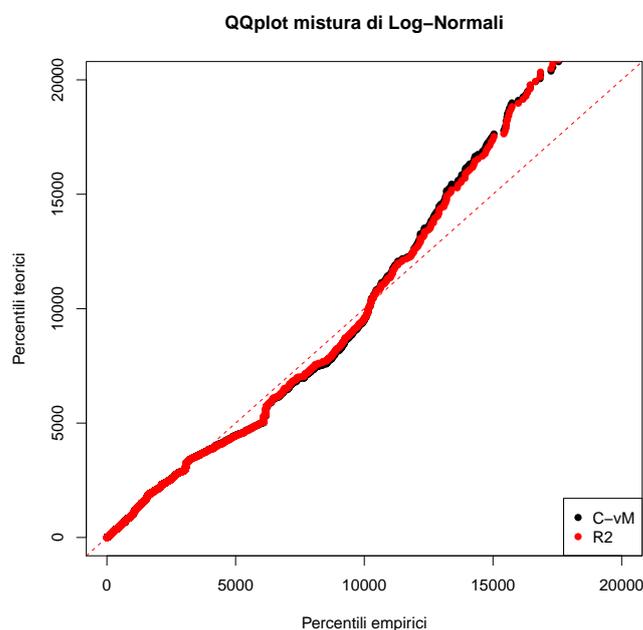


Figura 5.3: QQ-plot

questo strumento ormai ci appare chiara, essendo del tutto affine alla Tabella 4.4 vista in precedenza. Possiamo vedere come, per quanto riguarda le stime libere, la migliore sia prodotta dallo stimatore di massima verosimiglianza, sia considerando gli scostamenti, sia il coefficiente angolare.

Tuttavia, osservando i risultati delle stime con il vincolo, possiamo vedere che la situazione non è definita univocamente come nell'altro caso. Ci troviamo in una situazione di incertezza, da cui possiamo uscire solo attraverso un confronto grafico, proposto in Figura 5.3.

Posto che le due stime risultano molto vicine, come già avevamo potuto notare nella Tabella 5.1, possiamo vedere che i tracciati del QQ-plot sono pressoché identici. Appare comunque lievemente migliore il tracciato della stima R2, che soprattutto sulla coda destra.

Possiamo anche vedere un confronto con i valori degli scostamenti e dei coefficienti ottenuti con quelli riportati nelle Tabelle 4.5 e 4.6.

Questo risultato, rispetto ai valori della Tabella 4.5, ci conferma semplicemente che abbiamo condotto correttamente la stima. Il modello al vaglio e la distribuzione Log-Normale, infatti, non sono altro che due modelli annidati, il primo essendo una generalizzazione dell'altro. Chiaramente, ponendo $\pi = 1$ si ritorna al modello distributivo della Log-Normale, quindi se avessimo ottenuto delle valutazioni peggiori delle statistiche che consideriamo, evidentemente avremmo sbagliato su qualche passaggio.

Rispetto invece alla Tabella 4.6 otteniamo invece conferma che il modello di mistura si adatta meglio al confronto con la distribuzione Burr Inversa. Le uniche statistiche che ci danno un risultato opposto, infatti, sono quelle legate allo stimatore R5, ma abbiamo già avuto modo di mostrare quanto quest'ultimo sia particolare.

Possiamo quindi concludere che sia opportuno condurre la valutazione del *premium risk* anche con questo nuovo modello per la distribuzione della *severity*. Adotteremo quindi le stime prodotte con la massima verosimiglianza senza vincoli e dal MDA con distanza R2 con il valore atteso vincolato.

5.3 La *spliced distribution*

5.3.1 L'impostazione teorica

In questa sezione veniamo ad un ultimo modello per la distribuzione del costo singolo sinistro. Come il modello che abbiamo discusso nella sezione precedente, si tratta di una nuova distribuzione costruita a partire da altre distribuzioni più semplici, ma cambia completamente l'impostazione del problema.

Ad esempio in [10] troviamo la descrizione di un modello di questo tipo. Gli autori rilevano che, nel contesto attuariale, molto spesso sono utilizzate per modellare il costo sinistri le distribuzioni di Pareto e la Log-Normale. Il vantaggio della prima distribuzione è quello di avere una coda destra molto pesante, permettendo di replicare meglio dell'altra la probabilità di accadimento degli eventi *tail*. Tuttavia, la funzione di densità della Pareto è monotona decrescente, forma che mal si adatta ai dati della coda destra

della distribuzione, i così definiti *attritional claims*, cioè i sinistri di piccolo importo che usualmente costituiscono la stragrande maggioranza delle osservazioni.

Al contrario, la densità della Log-Normale non è monotona decrescente³, permettendo un miglior *fitting* sugli *attritional claims*, ma non sui sinistri *tail*.

Il passo che quindi compiono gli autori di [10] è quello di unire le due distribuzioni a formarne una nuova, la *spliced distribution*.

Formalmente, nel caso in cui si adottino due distribuzioni, aventi rispettivamente funzioni di densità $f_1(\cdot)$ e $f_2(\cdot)$, la nuova funzione di densità $f(\cdot)$ sarà definita come:

$$f(x) = \begin{cases} c \cdot f_1(x) & x \leq \theta \\ c \cdot f_2(x) & x > \theta \end{cases}. \quad (5.7)$$

In [10] si decide anche di adottare una Log-Normale per $f_1(\cdot)$ ed una Pareto a due parametri, con locazione pari al θ stesso per $f_2(\cdot)$. Come Pareto, in particolare, viene adottata la formulazione classica definita come:

$$f_2(x) = \frac{\alpha \theta^\alpha}{x^{\alpha+1}} \quad (5.8)$$

Vi è una differenza abissale che distingue questo approccio dalla mistura presentata nella sezione precedente: nella mistura i supporti delle distribuzioni che uniamo possono essere sovrapposti⁴, mentre in questo caso si identificano dei valori soglia che dividono i supporti delle distribuzioni.

Chiaramente, se andremo ad utilizzare delle distribuzioni che abbiano per definizione supporto illimitato, come ad esempio la Log-Normale e la Pareto, le dovremo opportunamente troncare e dovremo quindi ricavare il valore c in modo che sia garantito il rispetto

³Si ricorda, infatti, che la moda della Log-Normale di parametri μ e σ è pari a $e^{\mu-\sigma^2}$.

⁴Abbiamo visto il caso di due Log-Normali, aventi quindi supporto comune $(0, +\infty)$.

della condizione

$$\int_{-\infty}^{+\infty} f(x)dx = c \cdot \left(\int_{-\infty}^{\theta} f_1(x)dx + \int_{\theta}^{+\infty} f_2(x)dx \right) = 1, \quad (5.9)$$

mentre la condizione $f(x) \geq 0$ risulta rispettata per ogni possibile valore di x a patto che le $f_i(x)$ siano funzioni di densità e che $c \geq 0$. Il problema diviene quindi a quattro parametri.

Una ulteriore coppia di vincoli che viene imposta in [10], essendo richiesto

$$f(\theta^-) = f(\theta^+), \quad f'(\theta^-) = f'(\theta^+). \quad (5.10)$$

Essi servono a garantire la continuità e la differenziabilità di $f(\cdot)$ nel punto di soglia θ . Questa nuova imposizione vincola quindi due parametri in funzione degli altri, lasciando due soli gradi di libertà.

Questi ultimi due vincoli sono desiderabili per ottenere una maggior raffinatezza del modello sotto un profilo di mera eleganza, ma non sono condizioni necessarie affinché $f(\cdot)$ sia una funzione di densità.

Dal punto di vista teorico, infatti, gli unici requisiti necessari sono la condizione 5.9 e $f(\cdot) \geq 0$.

Anche sotto il profilo dell'applicazione per cui utilizzeremo la nostra nuova funzione non riscontreremmo alcun problema se queste condizioni non fossero valide. Per il campionamento da una variabile casuale continua, infatti, si utilizza la funzione di ripartizione che risulterebbe, in ogni caso, continua, crescente e con codominio pari all'intervallo $(0, 1)$. L'unico aspetto apparentemente problematico sarebbe la presenza di un punto angoloso, ma sappiamo che, trattandosi appunto di una variabile continua, la probabilità di un punto è nulla.

Inoltre, in [31], troviamo anche un altro commento a questo modello. L'imposizione dei vincoli di continuità e differenziabilità si ripercuote sulle proporzioni di mistura, di modo che il 39.21% della popolazione proverrà dalla Log-Normale e che il valore soglia

θ corrisponde sempre al percentile di ordine 64.514 di questa distribuzione.

Per superare questo problema, in [38] viene proposta una ulteriore generalizzazione del modello che adotta un fattore moltiplicativo per ponderare le due distribuzioni, grazie al quale, anche imponendo le nuove condizioni, viene meno la determinatezza della proporzione.

La seconda generalizzazione che viene proposta riguarda la distribuzione di Pareto. Fino ad ora, infatti era stata adottata la Pareto classica, la cui funzione di densità è descritta nella Formula 5.8. Qui, invece, viene adottata la Pareto Tipo II, nota anche come distribuzione Lomax, la cui funzione di densità è

$$f_2 = \frac{\alpha s^\alpha}{(x + s)^{\alpha+1}}. \quad (5.11)$$

Inoltre, essendo α e s parametri rispettivamente di forma e scala, si riscrive includendo anche la soglia θ come parametro di localizzazione, in modo che il contenuto della parentesi al denominatore sia $x - \theta + s$.

In [31] troviamo una prosecuzione di questo filone di ricerca, attraverso ulteriori generalizzazioni del modello. Ritroviamo, nel sommario, una frase che ci permette di cogliere la novità introdotta:

instead of using a single threshold value applying uniformly to the whole data set, the model proposed in the present paper allows for heterogeneity with respect to the threshold and let it vary among observations⁵.

In esso, infatti, troviamo un modello che prevede un'aleatorietà per il livello soglia. Per la precisione sono state adottate due possibili soluzioni per descrivere la variabile $\tilde{\theta}$: adottare una distribuzione Gamma o adottare una distribuzione Log-Normale.

⁵“Invece di adottare un valore soglia unico da applicare uniformemente all'intero *data set*, il modello proposto nel presente articolo consente l'eterogeneità rispetto alla soglia e permette che vari tra le osservazioni”.

Fino a qui alcuni modelli presentati in letteratura sul tema delle distribuzioni *spliced*, con riferimento ad un contesto simile al nostro. Occorre ora chiedersi quale potrebbe essere un modello idoneo alle nostre valutazioni.

Qualunque modello altro non è se non una rappresentazione semplificata della realtà ed il fine per cui viene sviluppato è che sia possibile comprenderne un aspetto nuovo.

Nel campo di applicazione di questi modelli, inoltre, essi acquistano tanto più valore quanto più sono fruibili e fonte di informazioni altrimenti inconoscibili.

Traiamo spunto da queste due affermazioni e ci poniamo in un contesto concreto. Ad esempio la nostra compagnia si potrebbe chiedere quale sia il trattato di riassicurazione più efficiente per la riduzione del proprio rischio. Per valutare, in particolare, un trattato di tipo *excess of loss*, dovrebbe definire il valore da adottare come priorità. Avendo condotto uno studio del proprio *database* sinistri attraverso uno dei modelli presentati in [38], adottare la stima di θ potrebbe essere una scelta ottimale, proprio perché divide i sinistri *attritional* da quelli di portata maggiore, i più rischiosi.

Al contrario, i modelli proposti in [31] non permettono questo tipo di ragionamento, pur essendo molto più raffinati.

Per quanto riguarda invece il modello inizialmente proposto in [10], la problematica sollevata rispetto alla stretta relazione che intercorre tra la stima di θ ed i parametri della Log-Normale, non lo rende affidabile in quest'ottica.

Adottiamo quindi un modello simile a quello previsto in [38], seppure più semplice.

Il nostro modello sarà costituito da una distribuzione Log-Normale fino alla soglia θ e da una Pareto di Tipo II oltre tale limite. Tuttavia, a differenza del modello che si ritrova in [38], considereremo entrambe le distribuzioni troncate, quindi non inseriremo il θ come parametro di localizzazione.

Indicando quindi con θ la soglia, $f_1(\cdot)$ e $F_1(\cdot)$ rispettivamente funzione di densità e di ripartizione di una Log-Normale e con $f_2(\cdot)$ e $F_2(\cdot)$ di una Pareto Tipo II, il nostro modello sarà:

$$f(x) = \begin{cases} \pi(\theta) \cdot \frac{f_1(x)}{F_1(\theta)} & \text{se } 0 < x \leq \theta \\ (1 - \pi(\theta)) \cdot \frac{f_2(x)}{(1 - F_2(\theta))} & \text{se } x > \theta \end{cases}. \quad (5.12)$$

$\pi(\theta)$ sarà scelto in modo da rispettare innanzitutto la condizione 5.9, quindi, poiché le due componenti sono già normalizzate ad uno, sarà sufficiente che sia $0 \leq \pi(\theta) \leq 1$. Lo scegliamo in funzione di θ nel senso che corrisponderà alla proporzione dei nostri dati empirici che staranno al di sotto della soglia.

Un modello simile, a differenza di quanto prodotto in [10], sacrifica le condizioni di continuità e di differenziabilità per conservare tutti e cinque i gradi di libertà, ma al contempo si mantiene più semplice di quelli presentati negli altri due lavori.

Come ultima analisi ci preme notare che un modello molto simile è già stato adottato all'interno del nostro elaborato. Il modello che avevamo introdotto nel Capitolo 3 come modello non parametrico chiamato Ogiva, altro non era se non un modello di questa tipologia. A differenza di quello attualmente al vaglio, tuttavia, le distribuzioni di cui era costituito erano delle distribuzioni Uniformi ed erano molte più delle due che adottiamo ora.

5.3.2 Le stime

Come in precedenza conduciamo sia la stima con il MDA, con distanza di Cramer-von Mises, sia attraverso la massima verosimiglianza.

Mentre nelle stime condotte nel corso del presente e del precedente capitolo abbiamo potuto condurre le stime con e senza il vincolo rispetto al valore atteso, in questo caso non ci è possibile.

In precedenza eravamo infatti riusciti a ricavare in forma esplicita una funzione che legasse uno dei parametri agli altri ed alla media osservata. Questo, avendo a che fare con delle funzioni troncate, non è un problema dalla soluzione banale, quindi non abbiamo la possibilità di introdurre il vincolo.

Come ulteriore considerazione, mentre per la computazione della verosimiglianza possiamo avvalerci della Formula 5.12, la funzione di ripartizione, necessaria alla distanza di Cramér-von Mises, sarà definita come:

$$F(x) = \begin{cases} \pi(\theta) \cdot \frac{F_1(x)}{F_1(\theta)} & \text{se } 0 < x \leq \theta \\ \pi(\theta) + (1 - \pi(\theta)) \cdot \frac{F_2(x) - F_2(\theta)}{(1 - F_2(\theta))} & \text{se } x > \theta \end{cases}. \quad (5.13)$$

Le stime, condotte attraverso i due metodi restituiscono i risultati riportati in Tabella 5.4.

	θ	$\pi(\theta)$	μ	σ	α	s
ML	13.577	0.113	4.460	1.641	1.259	115.725
C-vM	13.154	0.110	4.407	1.643	1.348	128.795

Tabella 5.4: Stime vincolate del miscuglio di Log-Normali

Possiamo notare che le stime risultano vicine per entrambi i metodi. Dato che le funzioni di cui andiamo a ricercare l'ottimo globale sono altamente non lineari, ricche di ottimi locali e costruite su cinque parametri, non possiamo avere una certezza assoluta riguardo ai risultati prodotti. Oltre al solito algoritmo di stima, per limitare la possibilità di essere incorsi in un punti di ottimo locale, abbiamo prodotto diverse stime, ma i valori che ci restituiscono i valori minimi per le funzioni obiettivo sono quelli riportati in tabella.

Per correttezza e completezza informativa rileviamo che la stima C-vM è stata ottenuta impostando i valori di ottimo ML come parametri di inizializzazione. Tuttavia, a partire da essi, la funzione di distanza è stata ridotta dal valore di circa 1.0721 a circa 0.7209 nel corso di oltre 2000 iterazioni. La stima può pertanto essere considerata affidabile.

Con riferimento a quanto detto nella sezione precedente, circa la possibilità di adottare il valore soglia θ come priorità, rileviamo che non avrebbe senso un trattato di riassicurazione *Excess of Loss* che andrebbe ad incidere quasi sul 90% dei sinistri, tuttavia, nel corso delle elaborazioni abbiamo rilevato per entrambi i metodi un minimo

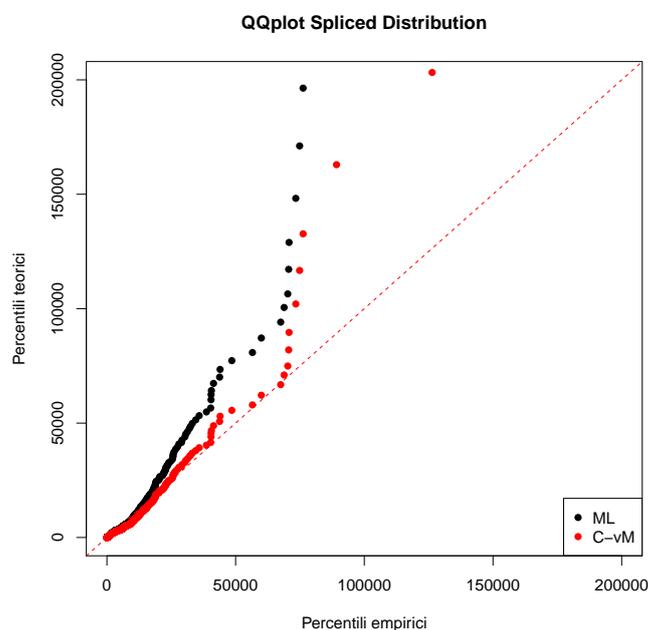


Figura 5.4: QQ-plot

locale in corrispondenza di $\theta \sim 1600$, cioè con $\pi(\theta) \sim 95\%$. Questo sarebbe chiaramente un livello più credibile sotto questo punto di vista.

Produciamo, in Figura 5.4 il confronto tra i QQplot delle due stime ottenute.

In questo caso non abbiamo dubbi nella scelta della distribuzione. Vediamo infatti che, sebbene le stime abbiano prodotto dei risultati molto vicini, la distribuzione ottenuta con il MDA si adatta molto meglio ai nostri dati empirici.

Valuteremo quindi la rischiosità della compagnia simulando il costo dei sinistri attraverso la distribuzione identificata con questa nuova stima.

5.4 La valutazione del rischio

Come nei capitoli precedenti, anche in questo caso adottiamo l'approccio di tipo Monte Carlo per la stima della distribuzione del costo sinistri aggregato.

Questa volta, tuttavia, adottiamo le medesime decisioni prese nel capitolo preceden-

te, con riferimento al numero di replicazioni del processo ed alla presenza del secondo processo discreto.

	Mistura		<i>Spliced</i>
	Libera	Vincolata	
$\mathbb{E}(\tilde{X})$	12.026	13.116	12.320
$\sigma(\tilde{X})$	2.638	2.908	6.546
CoV	0.219	0.222	0.531
$\gamma(\tilde{X})$	0.441	0.452	67.516
VaR 99.5%	7.869	8.693	14.876
T-VaR 99%	8.230	9.126	26.999
T-VaR/VaR	1.046	1.050	1.815

Tabella 5.5: Statistiche significative del costo aggregato dei sinistri, valori in milioni di euro

I valori che otteniamo con questa nuova stima sono pressoché in linea con quelli ottenuti nelle elaborazioni del precedente capitolo, per quanto riguarda la distribuzione misturata, ma differiscono rispetto alla distribuzione *spliced*.

Notiamo infatti che i valori ottenuti con la prima di queste nuove ipotesi si situano tra quelli ottenuti con i modelli che adottavano la Log-Normale nel capitolo precedente, mentre con la seconda si ottiene una valutazione del rischio superiore di circa il 2.4% maggiore rispetto a quella ottenuta con la distribuzione vincolata ma non misturata.

In particolare, osserviamo che la seconda distribuzione, pur avendo un coefficiente di variabilità relativa ed un'asimmetria comparabili a quelli della prima. Tuttavia, la stima del VaR 99.5% appare di circa il 10.4% maggiore, mentre la stima del Tail-VaR di quasi il 10.9% maggiore.

Con riferimento alla distribuzione *spliced*, otteniamo invece dei valori del tutto differenti. Tuttavia, anche in questo caso rileviamo che, con le stime che abbiamo ottenuto, il momento secondo della distribuzione Pareto diverge. Abbiamo infatti adottato $\alpha = 1.348$, ma la varianza della distribuzione assume valore finito solo se $\alpha > 2$.

Considerato che il 90% circa dei valori proviene da questa, che peraltro è troncata sulla coda sinistra, anche il momento secondo della distribuzione *spliced* diverge. Anche

in questo caso, quindi, ricadiamo nelle stesse problematiche che abbiamo visto in precedenza con i modelli che adottavano la distribuzione di Burr Inversa per il costo dei sinistri.

Tratteremo questa tematica nell'Appendice B.

Appendice A

Un approccio individuale per la *frequency*

Nella Sezione 3.3 abbiamo sviluppato un modello per la simulazione del numero dei sinistri a partire da un approccio collettivo, fondamento del CRM. In questa appendice, invece, vagliamo alcuni modelli costruiti su base individuale con lo stesso fine. Andremo quindi a ricercare una distribuzione adatta a descrivere il numero dei sinistri generati da una singola polizza.

Iniziamo riprendendo in considerazione i dati della Tabella 3.3. Si può notare come il numero dei sinistri sia particolarmente volatile sia considerando i diversi anni, sia rimanendo all'interno del medesimo anno.

Tuttavia rileviamo anche quanto siano simili le misure di volatilità all'interno dei diversi anni, fatto che ci dimostra una coerenza tra i dati nel corso del tempo

Infine osserviamo quanto sia significativa la differenza tra media e varianza del numero dei sinistri per polizza, fattore che rende inverosimile un modello di tipo Poisson semplice.

Cominceremo comunque nella Sottosezione A.1 con un modello di questo tipo, per poi costruirne delle varianti.

A.1 Poisson

Il primo modello che consideriamo è la distribuzione di Poisson. Come detto in precedenza non ci aspettiamo che possa avere un buon adattamento, ma si tratta comunque del punto di partenza per sviluppare degli altri modelli più raffinati.

Proprietà molto importante per questa distribuzione è l'additività: se $X \sim Pois(n_X)$ e $Y \sim Pois(n_Y)$, vale che $(X + Y) \sim Pois(n_X + n_Y)$ se vi è indipendenza tra \tilde{X} e \tilde{Y} .

Grazie a questa proprietà possiamo spostare l'attenzione sul numero di sinistri generati da ogni singola polizza.

Indicheremo quindi con \tilde{K} e con \mathbf{n} le grandezze riferite alle singole polizze e varrà quindi $\tilde{K} = \sum \tilde{K} \sim Pois(n = \sum \mathbf{n})$, sotto le opportune ipotesi.

La stima di \hat{n} condotta sui dati a nostra disposizione risulta 111.939 per ogni singola polizza, ma appaiono evidenti le carenze di questo modello. A titolo esemplificativo basta pensare che, considerando 274 polizze, come nel 2011, $P(\tilde{K} < 24000 | n = 274\hat{n}) = 1.525E-231 \sim 0$, ma il dato empirico è stato $k_{2011} = 23816$. Risulta quindi evidente la necessità di modificare il modello.

A.2 Poisson Misturata

Introduciamo quindi, come nella Sezione 3.3.2 un fattore di disturbo moltiplicativo \tilde{q} che vada a modificare il parametro della nostra distribuzione

Senza per ora adottare un'ipotesi distributiva per \tilde{q} , possiamo stimare le quattro osservazioni della variabile \tilde{q} come i rapporti tra il numero medio di sinistri per polizza osservato all'interno di ogni anno e il numero medio nell'arco dei quattro anni. L'assunzione sottostante è che la frequenza sinistri osservata in ogni anno sia pari a $q\mathbf{n}$.

Proponiamo quindi la Tabella A.1 che raccoglie le medesime statistiche della Tabella 3.3, ma calcolate sui dati standardizzati rispetto q che abbiamo calcolato. Osserviamo nuovamente un dato per cui non possiamo essere pienamente soddisfatti: la varianza al-

l'interno dei gruppi si discosta molto dalla frequenza media. Questo è un chiaro segnale che il modello debba essere in qualche modo modificato.

	2008	2009	2010	2011
Polizze	116	197	263	274
Sinistri	12450	21141	28216	29374
Frequenza Sinistri	107.33	107.31	107.29	107.20
Varianza numero sinistri	97800.52	109188.50	85594.30	113310.91
CV numero sinistri	2.91	3.08	2.73	3.14

Tabella A.1: Riassuntivo numero sinistri depurato dal fattore di disturbo

A.3 Poisson Composta

Andremo a questo punto a verificare l'adattamento di alcuni modelli, noti in letteratura anche con il nome di Poisson Sovradisperse. Ciascuno di essi assume l'esistenza di una variabile aleatoria $\tilde{\Theta}$ che agisca come disturbo moltiplicativo sul valore di n , ragion per cui, come in precedenza, si cercherà di stimare i parametri delle distribuzioni $f_{\tilde{\Theta}}(\theta)$ in modo che $\mathbb{E}(\tilde{\Theta}) = 1$.

Condurremo le stime attraverso il metodo dei momenti, avvalendoci della Formula 3.34, opportunamente rivista adottando la notazione $\tilde{\Theta}$ e $\tilde{\kappa}$.

La famiglia di distribuzioni è chiamata Sovradispersa perché se $\sigma_{\tilde{\Theta}}^2$ è maggiore di 0, $\sigma_{\tilde{\kappa}}^2$ diviene maggiore di n . Questo è esattamente quanto accade per i nostri dati riferiti al numero dei sinistri, ragion per cui riteniamo che questa famiglia di modelli possa avere un buon adattamento.

Prendiamo in considerazione in particolare tre modelli di Poisson Sovradisperse che si differenziano per la scelta della distribuzione di $\tilde{\Theta}$, poiché gli unici requisiti sono quelli che sono stati imposti per i fattori di disturbo nella Sezione 3.3.2, cioè che il valore atteso deve essere unitario e deve essere assicurata la non negatività.

A.3.1 La Poisson Composta Gamma

La prima distribuzione $F_{\tilde{\Theta}}(\theta)$ che consideriamo è la Gamma, che già avevamo adottato anche nell'approccio collettivo.

Come descritto in precedenza, questo tipo di mistura da origine ad una distribuzione Binomiale Negativa. Procedendo quindi alla stima secondo il metodo dei momenti condotta sui nostri dati, otteniamo $\hat{p} = 1.0498\text{E} - 03$ e $\hat{\alpha} = 0.113^1$.

A.3.2 La Poisson Composta Log-Normale

Consideriamo in questo caso un modello descritto anche in [13]. Questo modello considera che il fattore di disturbo $\tilde{\Theta}$ sia distribuito come una Log-Normale così come descritta nella Sezione 3.2.1.1.

Per assicurare che valga $\mathbb{E}(\tilde{\Theta}) = 1$, è necessario che $\mu = -\frac{\sigma^2}{2}$. Inoltre, sapendo che $\sigma_{\tilde{\Theta}}^2 = e^{\sigma^2} - 1$, utilizzando la Formula 3.34, ricaviamo la stima $\hat{\sigma}^2 = 2.28$ con il metodo dei momenti.

A.3.3 La Poisson Composta Gaussiana Inversa

Infine, l'ultimo modello che consideriamo prevede che $\tilde{\Theta}$ sia distribuito come una Gaussiana Inversa, come descritta nella Sezione 3.2.1.6. I parametri stimati, utilizzando nuovamente la Formula 3.34 e sapendo che $\sigma_{\tilde{\Theta}}^2 = \frac{\mu^3}{\tau}$, sono $\hat{\mu} = 1$ e $\hat{\tau} = 1.013\text{E} - 01$.

A.4 La scelta della distribuzione

È quindi necessario un esame approfondito di queste distribuzioni. Riportiamo in Tabella A.2 i valori delle funzioni di verosimiglianza, dell'Akaike Information Criterion e del Bayesian Information Criterion.

Il guadagno, in termini di bontà dell'adattamento, che possiamo osservare nel passaggio dal modello semplice definito dal solo parametro n all'aggiunta di un secondo

¹Il parametro n sarebbe invece pari a 107.26, valore coerente con quelli ottenuti in precedenza.

parametro è sicuramente notevole e non può essere trascurato, mentre appare evidente come i tre modelli a due parametri siano molto affini.

La nostra scelta si restringe quindi a due modelli: la Poisson Composta Gamma e la Poisson Gaussiana Inversa. Il secondo modello viene preso in considerazione per dei valori riportati in Tabella A.2, mentre il primo non viene escluso anche a causa di quanto esposto in Sezione 3.3.2.1 a proposito dei vantaggi in termini di tempi di elaborazione.

	(Log)Verosimiglianza	AIC	BIC
Semplice	-135317.70	270637.41	270642.15
Gamma	-4886.17	9776.33	9785.82
Log-Normale	-4928.20	9860.41	9869.90
Gaussiana Inversa	-4859.92	9723.84	9733.33

Tabella A.2: Adattamento dei modelli di tipo Poisson Composta al numero dei sinistri

Per valutare quale sia il miglior modello, conduciamo un'estrazione pseudo-casuale di 10000 costi aggregati dei sinistri adottando entrambe le distribuzioni. Ciò significa che, trattandosi di modelli individuali, dobbiamo estrarre per ogni replicazione 278 numeri pseudo-casuali da ciascuna e sommarli. Dalla Pólya otteniamo un campione il cui Campo di variazione è [9993, 68770], mentre per la Poisson-Gaussiana Inversa osserviamo [147, 1690661] che risulta poco credibile. Appare quindi evidente che la distribuzione migliore per modellare il numero dei sinistri sia la Pólya.

A.5 La simulazione

Riproponiamo qui le stime dei costi aggregati dei sinistri ottenute nei vari capitoli. Per completezza evidenziamo alcuni fattori.

- Per la scelta della distribuzione del primo fattore di disturbo \tilde{q} abbiamo adottato l'Uniforme. Empiricamente, otteniamo dalle quattro osservazioni che $\sigma_{\tilde{q}}^2 = 0.0432$ e, essendo $\mathbb{V}(\tilde{q}) = \frac{(a-b)^2}{12}$, possiamo impostare la seconda equazione del sistema, ottenendo quindi $a = 0.640$ e $b = 1.360$.

- Per il calcolo di $\sigma_{\tilde{q}}^2$ non abbiamo adottato alcuna ponderazione, poiché questo disturbo è inteso come rischio sistematico dell'intero mercato.
- Proprio perché \tilde{q} indica un rischio sistematico, dovrà agire allo stesso modo sull'intero portafoglio, quindi per ogni replicazione simuleremo un unico valore q che andrà a disturbare allo stesso modo tutte le 278 estrazioni. Per questo disturbo scegliamo la riparametrizzazione della Binomiale Negativa che utilizza direttamente il valore atteso μ oltre ad α . Varrà quindi $p = \frac{\alpha}{(\alpha+\mu)}$.
- Le stime che produciamo sono al netto dei sinistri tardivi, che non abbiamo considerato nello sviluppo di questo modello. Le stime non saranno quindi direttamente comparabili.

	Log-Normale	Burr Inversa	Ogiva
$\mathbb{E}(\tilde{X})$	7.522	7.228	8.912
$\sigma(\tilde{X})$	2.088	2.249	2.491
CoV	0.278	0.311	0.280
$\gamma(\tilde{X})$	0.463	2.355	0.463
VaR 99.5%	5.974	6.917	7.089
T-VaR 99%	6.185	8.809	7.388

Tabella A.3: Statistiche significative del costo aggregato dei sinistri sotto le ipotesi al vaglio nel Capitolo 3, senza il sinistro *outlier*, valori in milioni di euro

	Log-Normale	Burr Inversa	Ogiva
$\mathbb{E}(\tilde{X})$	7.528	7.277	12.323
$\sigma(\tilde{X})$	2.091	2.272	3.900
CoV	0.278	0.312	0.316
$\gamma(\tilde{X})$	0.539	2.529	0.606
VaR 99.5%	6.193	7.096	12.193
T-VaR 99%	6.466	9.207	12.614

Tabella A.4: Statistiche significative del costo aggregato dei sinistri sotto le ipotesi al vaglio nel Capitolo 3, considerando anche il sinistro *outlier*, valori in milioni di euro

	Log-Normale		Burr Inversa	
	Libera	Vincolata	Libera	Vincolata
$\mathbb{E}(\tilde{X})$	7.766	9.626	9.992	9.616
$\sigma(\tilde{X})$	2.146	2.679	7.217	6.101
CoV	0.276	0.278	0.722	0.635
$\gamma(\tilde{X})$	0.493	0.496	124.315	112.078
VaR 99.5%	6.195	7.725	12.027	10.920
T-VaR 99%	6.451	8.073	22.400	19.475

Tabella A.5: Statistiche significative del costo aggregato dei sinistri sotto le ipotesi al vaglio nel Capitolo 4, valori in milioni di euro

	Mistura		<i>Spliced</i>
	Libera	Vincolata	
$\mathbb{E}(\tilde{X})$	8.824	9.626	9.053
$\sigma(\tilde{X})$	2.457	2.705	8.545
CoV	0.278	0.281	0.944
$\gamma(\tilde{X})$	0.496	0.518	133.672
VaR 99.5%	7.144	7.878	11.852
T-VaR 99%	7.403	8.266	23.608

Tabella A.6: Statistiche significative del costo aggregato dei sinistri sotto le ipotesi al vaglio nel Capitolo 5, valori in milioni di euro

Appendice B

Incidenza del numero di simulazioni

In questa appendice intendiamo mettere in risalto un aspetto molto importante dell'utilizzo del metodo Monte Carlo. Questo metodo è di vitale importanza quando ci si occupa di problemi come il nostro, in cui appare estremamente complessa, se non addirittura impossibile, la ricerca di una soluzione per via analitica. Grazie ad esso, infatti, troviamo un valore approssimato della soluzione che ricerchiamo.

Tuttavia possiamo intuire che l'approssimazione sarà tanto più affine al vero ma incognito valore obiettivo delle nostre ricerche, quanto più alto sarà il numero di replicazioni che adotteremo. Ci chiediamo quindi quale sia l'incidenza del numero di replicazioni sulla nostra stima.

Proviamo a fare qualche semplice ragionamento in merito alla consistenza delle stime. Adottiamo, a questo fine, le ipotesi classiche del CRM, vale a dire l'indipendenza tra il numero ed il costo dei sinistri e tra i costi sinistri presi a coppie. Riutilizziamo quindi anche la notazione del Capitolo 3 e adottiamo per la variabile \tilde{K} il modello sviluppato nella Sezione 3.3.

B.1 La consistenza delle stime di media e varianza

Proviamo innanzitutto a valutare la consistenza delle stime del valore atteso e della varianza delle distribuzioni che abbiamo sviluppato nel corso dei capitoli.

Trattiamo inizialmente della stima del valore atteso della distribuzione. Ipotizziamo di aver ottenuto un campione casuale semplice di H possibili costi sinistri aggregati attraverso una simulazione Monte Carlo: da essi possiamo calcolare la media campionaria come $\bar{X} = \frac{\sum_{i=1}^H X_i}{H}$.

Stante poi l'ipotesi di indipendenza tra le differenti X_i , possiamo rievocare le note relazioni alla base della statistica inferenziale, tali per cui,

$$\mathbb{V}(\bar{X}) = \mathbb{V}\left(\frac{\sum_{i=1}^H X_i}{H}\right) = \frac{H\mathbb{V}(X_i)}{H^2} = \frac{\mathbb{V}(X_i)}{H}. \quad (\text{B.1})$$

Già da questo possiamo trarre la conclusione che la media campionaria è uno stimatore consistente per il valore atteso, se e solo se la $\mathbb{V}(\tilde{X})$ esiste finita.

Il medesimo ragionamento, condotto sullo stimatore della varianza campionaria corretta $S_c^2 = \frac{\sum_{i=1}^H (X_i - \bar{X})^2}{H-1}$, appare più complesso, soprattutto se non ricorriamo al teorema del limite centrale. Indicando $Y_i = X_i - \mu$, dove μ indica il vero valore atteso della variabile aleatoria \tilde{X} , possiamo scrivere

$$S_c^2 = \frac{H \sum_{i=1}^H Y_i^2 - \left(\sum_{i=1}^H Y_i\right)^2}{H(H-1)} \quad (\text{B.2})$$

e, passando ai quadrati,

$$S_c^4 = \frac{H^2 \left(\sum_{i=1}^H Y_i^2\right)^2 - 2H \sum_{i=1}^H Y_i^2 \left(\sum_{i=1}^H Y_i\right)^2 + \left(\sum_{i=1}^H Y_i\right)^4}{H^2(H-1)^2}, \quad (\text{B.3})$$

da cui, passando infine al valore atteso,

$$\mathbb{E}(S_c^4) = \frac{H^2 \mathbb{E} \left(\left(\sum_{i=1}^H Y_i^2 \right)^2 \right) - 2H \mathbb{E} \left(\sum_{i=1}^H Y_i^2 \left(\sum_{i=1}^H Y_i \right)^2 \right) + \mathbb{E} \left(\left(\sum_{i=1}^H Y_i \right)^4 \right)}{H^2(H-1)^2}, \quad (\text{B.4})$$

Stanti quindi le ipotesi di indipendenza tra Y_i ed Y_j se $i \neq j$ e di identica distribuzione, possiamo scrivere quanto segue:

$$\mathbb{E} \left(\left(\sum_{i=1}^H Y_i^2 \right)^2 \right) = \mathbb{E} \left(\sum_{i=1}^H (Y_i^2)^2 \right) + \mathbb{E} \left(\sum_{i=1}^H \sum_{j \neq i} Y_i^2 Y_j^2 \right) = H\mu_4 + H(H-1)\mu_2^2, \quad (\text{B.5})$$

con riferimento al primo fattore, indicando con μ_i il momento centrale i -esimo della variabile \tilde{X} .

Per il secondo,

$$\begin{aligned} & \mathbb{E} \left(\left(\sum_{i=1}^H Y_i^2 \right) \left(\sum_{i=1}^H Y_i \right)^2 \right) = \\ & \mathbb{E} \left(\left(\sum_{i=1}^H Y_i^2 \right) \left(\sum_{i=1}^H Y_i^2 \right) \right) + \mathbb{E} \left(\left(\sum_{i=1}^H Y_i^2 \right) \left(\sum_{i=1}^H \sum_{i \neq j} Y_i Y_j \right) \right) = \\ & \mathbb{E} \left(\left(\sum_{i=1}^H Y_i^2 \right)^2 \right) + 0, \end{aligned} \quad (\text{B.6})$$

infine, per il terzo,

$$\begin{aligned}
 \mathbb{E}\left(\left(\sum_{i=1}^H Y_i\right)^4\right) &= \\
 \mathbb{E}\left(\left(\sum_{i=1}^H Y_i^2 + \sum_{i=1}^H \sum_{i \neq j} Y_i Y_j\right)^2\right) &= \\
 \mathbb{E}\left(\left(\sum_{i=1}^H Y_i^2\right)^2 + \left(\sum_{i=1}^H \sum_{i \neq j} Y_i Y_j\right)^2 + 2\left(\sum_{i=1}^H Y_i^2\right)\left(\sum_{i=1}^H \sum_{i \neq j} Y_i Y_j\right)\right) &= \\
 \mathbb{E}\left(\left(\sum_{i=1}^H Y_i^2\right)^2 + \left(\sum_{i=1}^H \sum_{i \neq j} \sum_{k=1}^H \sum_{l \neq k} Y_i Y_j Y_k Y_l\right) + 2\left(\sum_{i=1}^H Y_i^2\right)\left(\sum_{i=1}^H \sum_{i \neq j} Y_i Y_j\right)\right) &= \\
 \mathbb{E}\left(\left(\sum_{i=1}^H Y_i^2\right)^2\right) + 2\mathbb{E}\left(\sum_{i=1}^H \sum_{i \neq j} Y_i^2 Y_j^2\right) + 2 \cdot 0 &= \\
 H\mu_4 + 3H(H-1)\mu_2^2. &
 \end{aligned} \tag{B.7}$$

Utilizzando i risultati ottenuti nelle Formule B.5, B.6 e B.7 nella Formula B.4, otteniamo che

$$\mathbb{E}(S_c^4) = \frac{(H-1)\mu_4 + (H^2 - 2H + 3)\mu_2^2}{H(H-1)}, \tag{B.8}$$

da cui

$$\mathbb{V}(S_c^2) = \mathbb{E}(S_c^4) - \mu_2^2 = \frac{\mu_4 - \mu_2^2}{H} + \frac{2}{H(H-1)}\mu_2^2. \tag{B.9}$$

Abbiamo quindi la certezza che anche la stima della varianza ottenuta attraverso la varianza campionaria corretta è consistente e converge per $H \rightarrow +\infty$. Tuttavia, a differenza della media campionaria, deve essere definito, oltre al momento centrale secondo, anche il momento centrale quarto di \tilde{X} .

Vediamo quindi cosa accade nel nostro modello, sfruttando le ipotesi del CRM. Come descritto in [35], la funzione generatrice dei cumulanti di \tilde{X} sotto le ipotesi che abbiamo adottato è:

$$\psi_{\tilde{X}}(s) = \psi_{\tilde{q}}(nM_{\tilde{Z}}(s) - n), \tag{B.10}$$

da cui, attraverso le derivate si ottiene che

$$\mathbb{V}(\tilde{X}) = \psi''_{\tilde{z}}(s=0) = n^2 m^2 \sigma_{\tilde{q}}^2 + n a_{2,\tilde{z}} \quad (\text{B.11})$$

dove m ed $a_{2,\tilde{z}}$ indicano rispettivamente il momento primo ed il momento secondo del costo singolo sinistro. Possiamo quindi essere sicuri che, se esistono finiti tutti i fattori che risultano dalla Formula B.11, lo stimatore del valore atteso sarà consistente.

Sul fronte della varianza, indicando con k_i il cumulante di ordine i -esimo e ricordando che

$$\mu_4 = k_4 + 3k_2^2,$$

possiamo nuovamente basarci sulle derivate della funzione generatrice dei cumulanti per verificarne le condizioni per la convergenza asintotica. Il risultato mostra che le condizioni necessarie e sufficienti sono l'esistenza dei primi quattro cumulanti della variabile \tilde{q} e dei primi quattro momenti dall'origine di \tilde{Z} .

Passando quindi ai modelli che abbiamo discusso nei capitoli precedenti, la consistenza di queste stime è garantita per tutti i modelli che abbiamo sviluppato basati sulla distribuzione Log-Normale e sull'Ogiva, ma non per quelli basati sulla distribuzione di Burr Inversa e sulla distribuzione *spliced*.

B.2 La consistenza delle stime di VaR e Tail-VaR

Tuttavia le nostre valutazioni del rischio sono basate sul VaR e sul Tail-VaR, che quindi coinvolgono i percentili del costo aggregato dei sinistri, non direttamente i momenti.

In questo caso richiamiamo semplicemente un teorema, formalizzato in [19] la cui dimostrazione può essere ritrovata in [5] per quanto riguarda il VaR.

Per presentarlo indichiamo con V_α e T_α i veri valori di VaR e Tail-VaR di ordine α , mentre con \hat{V}_α^H e \hat{T}_α^H le stime ottenute attraverso il metodo Monte Carlo con H

replicazioni.

Si dimostra che, se esiste un $\epsilon > 0$ tale che la funzione di densità $f_{\tilde{X}}(x)$ sia continua e differenziabile per ogni valore $x \in (V_\alpha - \epsilon, V_\alpha + \epsilon)$, allora, per $H \rightarrow \infty$, $\hat{V}_\alpha^H \rightarrow V_\alpha$ e $\hat{T}_\alpha^H \rightarrow T_\alpha$ con probabilità uno.

H	1000	5000	10000	50000	100000	500000
$\mathbb{E}(\tilde{X})$	10.202	10.202	10.211	10.223	10.225	10.221
$\sigma(\tilde{X})$	2.131	2.195	2.209	2.211	2.221	2.215
CoV	0.209	0.215	0.216	0.216	0.217	0.217
$\gamma(\tilde{X})$	0.501	0.445	0.456	0.423	0.432	0.429
VaR 99.5%	5.861	6.612	6.603	6.510	6.559	6.594
T-VaR 99%	6.596	6.972	6.962	6.844	6.893	6.875

Tabella B.1: Statistiche significative del costo aggregato dei sinistri sotto l'ipotesi di distribuzione Log-Normale per la *severity* per dimensioni campionarie crescenti, valori in milioni di euro

Se poi $\mathbb{E}(\tilde{X}^2) < \infty$, valgono

$$\frac{f_{\tilde{X}}(V_\alpha)}{\alpha(1-\alpha)}\sqrt{H}\left(\hat{V}_\alpha^H - V_\alpha\right) \xrightarrow{d} N(0,1)$$

e

$$\frac{(1-\alpha)}{\sqrt{\mathbb{V}\left(\left[\tilde{X} - V_\alpha\right]^+\right)}}\sqrt{H}\left(\hat{T}_\alpha^H - T_\alpha\right) \xrightarrow{d} N(0,1).$$

Nel nostro caso abbiamo la variabile \tilde{X} che è una sommatoria di variabili continue, quindi possiamo essere certi della prima tesi del teorema. Tuttavia, abbiamo visto che per i parametri che abbiamo stimato per la distribuzione di Burr Inversa, $a_{2,\tilde{Z}}$ diverge e con esso anche $\mathbb{E}(\tilde{X}^2)$.

B.3 Alcune verifiche empiriche

Verifichiamo empiricamente cosa accade sul nostro database. Da un primo confronto possiamo vedere che le stime prodotte sotto l'ipotesi di Log-Normalità, proposte in

Tabella B.1, sono pressoché stabili per tutte le grandezze che consideriamo. Invece, per la stima ottenuta con la Burr Inversa, i risultati sono differenti, come si osserva dalla Tabella B.2.

Possiamo infatti identificare una tendenza a crescere di $\sigma(\tilde{X})$ e soprattutto di $\gamma(\tilde{X})$, come ci si poteva aspettare. Allo stesso modo possiamo vedere che le valutazioni del rischio cambiano notevolmente, ma, in accordo con quanto esposto, non si osserva un trend di crescita o di decrescita.

H	1000	5000	10000	50000	100000	500000
$\mathbb{E}(\tilde{X})$	9.908	9.867	9.887	9.894	9.884	9.877
$\sigma(\tilde{X})$	2.679	2.483	3.239	3.487	3.046	2.905
CoV	0.270	0.252	0.328	0.352	0.308	0.294
$\gamma(\tilde{X})$	4.579	2.903	27.848	42.938	33.886	26.914
VaR 99.5%	8.511	8.087	7.843	7.580	7.550	7.489
T-VaR 99%	13.597	10.744	12.513	11.982	11.008	10.887

Tabella B.2: Statistiche significative del costo aggregato dei sinistri sotto l'ipotesi di distribuzione di Burr Inversa per la *severity* per dimensioni campionarie crescenti, valori in milioni di euro

Tuttavia ciò che veramente ci interessa è la variabilità di queste stime. La Tabella B.3 riporta i risultati di un altro esperimento che abbiamo condotto. Abbiamo proceduto come segue:

- abbiamo simulato 100 campioni di dimensione $H = 1000$ dalle due popolazioni in oggetto;
- abbiamo calcolato le usuali statistiche su ciascuno di essi;
- abbiamo calcolato la deviazione standard delle stime ottenute;
- successivamente abbiamo ripetuto la stessa procedura per campioni di dimensione $H = 10000$.

Possiamo osservare che le deviazioni standard di VaR e Tail-VaR si riducono al crescere del numero delle simulazioni, in accordo con il teorema presentato. Siamo

H	Log-Normale		Burr Inversa	
	1000	10000	1000	10000
VaR 99.5%	4.68E+05	1.52E+05	9.40E+05	3.10E+05
T-VaR 99%	4.26E+05	1.45E+05	6.84E+06	1.52E+06

Tabella B.3: Deviazione standard di VaR e Tail-VaR sotto le ipotesi di distribuzione Log-Normale e Burr Inversa per la *severity* per dimensioni campionarie di 1000 e di 10000 unità

infatti certi che, se $H \rightarrow \infty$, essi tenderanno a convergere al vero valore, pur non avendo informazioni riguardo alla distribuzione asintotica degli errori delle valutazioni ottenute a partire dalla seconda distribuzione. Notiamo tuttavia che le deviazioni delle stime della prima distribuzione sono sistematicamente al di sotto delle corrispettive della seconda distribuzione.

Conclusioni

L'aspetto caratterizzante dell'attività assicurativa è l'assunzione dei rischi cui sono sottoposti degli individui dietro il pagamento di una remunerazione, il premio. Per una compagnia assicurativa la gestione del rischio rappresenta il *core business* ed il primo passo atto a renderne possibile il controllo è la valutazione.

Per poter operare, una compagnia deve prestare a questa attività grande attenzione e le stime devono essere prodotte con oggettività in modo da garantirne l'affidabilità.

Arriviamo quindi a raccogliere i punti principali che abbiamo trattato nel corso dell'elaborato anche con l'intento di confrontare un'ultima volta i modelli sviluppati.

Nel corso del primo capitolo abbiamo valutato il rischio assicurativo attraverso i modelli proposti dall'EIOPA. Abbiamo potuto osservare che il modello che viene proposto come standard porta ad una valutazione dell'ordine di alcuni milioni di euro. Avevamo anche notato che i parametri utilizzati fanno riferimento ad una compagnia di medie dimensioni, mentre la nostra è di piccole dimensioni, quindi la valutazione del rischio era da ritenersi inadeguata.

Passando poi ai metodi che adottavano i parametri calibrati non direttamente dall'EIOPA, ma in modo che fossero più aderenti alla rischiosità della compagnia, avevamo ottenuto delle valutazioni molto differenti, in genere superiori alla decina di milioni di euro, ottenendo conferma dell'inadeguatezza del metodo vagliato in precedenza.

Tuttavia, l'obiettivo di lungo periodo del progetto Solvency 2 è di fare in modo che ogni compagnia assicurativa sviluppi un proprio modello per la valutazione del rischio aderente alle proprie caratteristiche. Per questo nella prosecuzione ci siamo cimentati

con la costruzione di un modello simile.

Ciò a cui abbiamo dato maggiore peso nel corso degli ultimi tre capitoli è stata una tematica che rappresenta un problema tecnico di difficile soluzione quando si conduce un'analisi come quella che abbiamo presentato: la ricerca di una opportuna distribuzione per il costo sinistri.

Abbiamo quindi adottato dapprima la metodologia classica della stima attraverso la massima verosimiglianza, selezionando due distribuzioni e la stima attraverso un metodo non parametrico. Abbiamo potuto osservare che la scelta della distribuzione della *severity* gioca un ruolo cruciale nella valutazione, riscontrando grandi differenze in termini di valutazione del rischio al variare delle ipotesi adottate.

Successivamente abbiamo adottato l'approccio della distanza minima per la stima dei parametri delle distribuzioni del costo sinistri, sviluppando anche delle funzioni di distanza alternative rispetto a quelle classiche. Abbiamo anche osservato l'incidenza che ha l'imposizione di un vincolo sul valore atteso dei sinistri che renda la distribuzione più coerente con i dati empirici. Questo approccio, noto da decenni alla statistica, rappresenta una novità nella pratica attuariale che sta iniziando da poco a prendere piede. Pertanto esso costituisce un nuovo campo di sviluppo per la scienza attuariale.

Da ultimo abbiamo condotto nuovamente l'analisi adottando l'approccio di due modelli che fanno riferimento a distribuzioni composte per il costo sinistri: la distribuzione misturata e la distribuzione *spliced*. Essi sono modelli ancora più raffinati dei precedenti, essendo in particolare il primo una generalizzazione di quanto adottato in precedenza con l'adozione della distribuzione Log-Normale.

Una compagnia assicurativa, dovendo prestare la massima attenzione ad analisi di valutazione del rischio, dovrebbe vagliare quale metodo meglio si adatta alla propria situazione. Per quanto riguarda la compagnia che ci ha permesso di utilizzare il proprio database, fermo restando che gli ultimi modelli presentati sono, dal punto di vista metodologico, i più raffinati, non possiamo stabilire con certezza quali siano i migliori in termini di applicabilità rispetto all'intera compagnia, non solo al portafoglio su cui

abbiamo lavorato.

Tuttavia possiamo fare qualche suggerimento. La *standard formula* proposta dall'EIOPA è stata calibrata avendo come obiettivo la valutazione della rischiosità di una compagnia mediana nel mercato europeo. Benché possa semplificare di molto i calcoli, la formula è troppo poco versatile per cogliere l'effettiva rischiosità del portafoglio su cui abbiamo condotto le analisi, soprattutto a causa del fattore dimensionale. L'auspicio quindi è che si scelga almeno uno dei modelli che abbiamo trattato nel seguito e che se ne confrontino i risultati.

Bibliografia

- [1] Appendice statistica. Relazione tecnica, ANIA, 2012.
- [2] T. W. Anderson e D. A. Darling. Asymptotic theory of certain “Goodness of Fit” criteria based on stochastic processes. *The Annals of Mathematical Statistics*, 23(2): pp. 193–212, 1952.
- [3] T. W. Anderson e D. A. Darling. A test of goodness of fit. *Journal of the American Statistical Association*, 49(268): pp. 765–769, 1954.
- [4] P. Artzner, F. Delbaen, J.-M. Eber, e D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3): pp. 203–228, 1999.
- [5] R. R. Bahadur. A note on quantiles in large samples. *The Annals of Mathematical Statistics*, 37(3): pp. 577–580, 1966.
- [6] A. Basu, H. Shioya, e C. Park. *Statistical Inference: The Minimum Distance Approach*. Monographs on Statistics and Applied Probability. CRC Press, 2011. ISBN 978-1-420-09965-2.
- [7] M. Bramanti, C. D. Pagani, e S. Salsa. *Matematica. Calcolo infinitesimale e algebra lineare*. Zanichelli, seconda edizione, 2004. ISBN 978-8-808-07547-5.
- [8] I. W. Burr. Cumulative frequency functions. *The Annals of Mathematical Statistics*, 13(2): pp. 215–232, 1942.

- [9] R. Byrd, P. Lu, J. Nocedal, e C. Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal of Scientific Computing*, 16(5): pp. 1190–1208, 1995.
- [10] Kahadawala C. e Malwane M. A. A. Modeling actuarial data with a composite lognormal-pareto model. *Scandinavian Actuarial Journal*, 2005(5): pp. 321–334, 2005.
- [11] CEIOPS. *QIS5 Technical Specifications, Annex to Call for Advice from CEIOPS on QIS5*. Brussels, 2010.
- [12] C. D. Daykin, T. Pentikäinen, e M. Pesonen. *Practical Risk Theory for Actuaries*. Chapman & Hall, London, 1994. ISBN 978-0-412-42850-0.
- [13] M. Denuit, X. Maréchal, S. Pietrebois, e J. F. Walhim. *Actuarial Modelling of Claim Counts*. John Wiley & Sons, Chichester, 2007. ISBN 978-0-470-02677-9.
- [14] D. Dufresne. Sums of lognormals. *Actuarial Research Clearing House, SOA*, 1, 2009.
- [15] Servizio Rapporti Internazionali e Studi Sezione Studi e Statistica. Statistiche andamento alcuni rami elementari. Relazione tecnica, ISVAP, Roma, 2012.
- [16] J. L. Folks e R. S. Chhikara. The inverse gaussian distribution and its statistical application—a review. *Journal of the Royal Statistical Society. Series B (Methodological)*, 40(3): pp. 263–289, 1978.
- [17] A. Gisler. *The Insurance Risk in the SST and in Solvency II: Modelling and Parameter Estimation*. 39th ASTIN Colloquium, Helsinki, Giugno 2009.
- [18] R. V. Hogg e S. A. Klugman. *Loss Distributions*. John Wiley & Sons, New York, prima edizione, 1984. ISBN 978-0-471-87929-9.
- [19] J. L. Hong e L. Guangwu. Monte carlo estimation of value-at-risk, conditional value-at-risk and their sensitivities. pp. pp. 95–107, 2011.

- [20] ISVAP. Provvedimento n. 2768 del 29 dicembre 2009.
- [21] N. L. Johnson, S. Kotz, e N. Balakrishnan. *Continuous Univariate Distributions*, volume 2. John Wiley & Sons, New York, seconda edizione, 1995. ISBN 978-0-471-58494-0.
- [22] N. L. Johnson, S. Kotz, e N. Balakrishnan. *Continuous Univariate Distributions*, volume 1. John Wiley & Sons, New York, seconda edizione, 1995. ISBN 978-0-471-58495-9.
- [23] C. Kleiber e S. Kotz. *Statistical Size Distributions in Economics and Actuarial Sciences*. John Wiley & Sons, New York, 2003. ISBN 978-0-471-15064-0.
- [24] S. A. Klugman, H. H. Panjer, e G. E. Willmot. *Loss Models: from Data to Decisions*. John Wiley & Sons, Hoboken, terza edizione, 2008. ISBN 978-0-470-18781-4.
- [25] A. N. Kolmogorov. Sulla determinazione empirica di una legge di distribuzione. *Giornale dell' Istituto Italiano degli Attuari*, 4: pp. 83–91, 1933.
- [26] J. F. Lawless. Inference in the generalized gamma and log gamma distributions. *Technometrics*, 22(3): pp. 409–419, 1980.
- [27] E. L. Lehmann. *Theory of Point Estimation*. John Wiley & Sons, New York, 1983. ISBN 978-0-471-05849-1.
- [28] B. G. Lindsay. *Mixture Models: Theory, Geometry, and Applications*. NSF-CBMS Regional Conference Series in Probability and Statistics. Institute of Mathematical Statistics, 1995. ISBN 978-0-940-60032-4.
- [29] U. Magagnoli. Materiali per la lezioni di Statistica II. 2008.
- [30] W. C. Parr e W. R. Schucany. Minimum distance and robust estimation. *Journal of the American Statistical Association*, 75(371): pp. 616–624, 1980.

- [31] M. Pigeon e M. Denuit. Composite LognormalPareto model with random threshold. *Scandinavian Actuarial Journal*, 2011(3): pp. 177–192, 2011.
- [32] E. Pitacco. *Modelli attuariali per le assicurazioni sulla salute*. CERAP. EGEA, Milano, 1995. ISBN 978-8-823-80286-5.
- [33] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. ISBN 3-900051-07-0.
- [34] A. Rubinstein. Dilemmas of an economic theorist. *Econometrica*, 74(4): pp. 865–883, 2006.
- [35] N. Savelli e G. P. Clemente. Dispense di teoria del rischio I. Università Cattolica del Sacro Cuore, 2010.
- [36] N. Savelli e G. P. Clemente. Dispense di teoria del rischio II. Università Cattolica del Sacro Cuore, 2012.
- [37] N. Savelli, G. P. Clemente, e D. Zappa. *Modelling and calibration for non-life underwriting risk: from empirical data to risk capital evaluation*. ASTIN Mexico Colloquia, Città del Messico, Ottobre 2012.
- [38] D. P. M. Scollnik. On composite lognormal-pareto models. *Scandinavian Actuarial Journal*, 2007(1): pp. 20–33, 2007.
- [39] N. Smirnov. Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*, 19(2): pp. 279–281, 1948.
- [40] E. W. Stacy e G. A. Mihram. Parameter estimation for a generalized gamma distribution. *Technometrics*, 7(3): pp. 349–358, 1965.
- [41] M. A. Stephens. EDF Statistics for Goodness of Fit and some comparisons. *Journal of the American Statistical Association*, 69(347): pp. 730–737, 1974.
- [42] M. C. K. Tweedie. Inverse statistical variates. *Nature*, 155(3938): p. 453, 1945.

- [43] J. Wu, N. Mehta, e J. Zhang. A flexible lognormal sum approximation In *Proceedings of IEEE Global Telecommunications Conference*. A cura di GLOBECOM, volume 6, 2005.

Elenco delle figure

1.1	QQ-plot del costo sinistri	6
2.1	Premi lordi contabilizzati, dati in € Mln	28
3.1	Istogramma vs Log-Normale	46
3.2	QQ-plot Log-Normale	47
3.3	QQ-plot Gamma	48
3.4	QQ-plot Gamma Generalizzata e Log-Gamma	51
3.5	QQ-plot Weibull	53
3.6	QQ-plot Beta Generalizzata del Secondo tipo	54
3.7	QQ-plot Burr e Burr Inversa	56
3.8	QQ-plot Log-Logistica	57
3.9	QQ-plot Gaussiana Inversa	58
3.10	Funzione di Ripartizione Empirica della variabile X	64
3.11	QQ-plot Ogiva con incremento $1E - 04$	65
3.12	QQ-plot Ogiva con incrementi differenti	66
3.13	Numero di polizze in essere	68
3.14	Istogramma del Costo Aggregato 2012 sotto l'ipotesi $\tilde{Z}_{2012,j} \sim \text{LNorm}$	76
3.15	Istogramma del Costo Aggregato 2012 sotto l'ipotesi $\tilde{Z}_{2012,j} \sim \text{InvBurr}$	77
3.16	Istogramma del Costo Aggregato 2012 sotto l'ipotesi $\tilde{Z}_{2012,j} \sim \text{Ogiva}$	78

4.1	Confronto tra le stime ottenute con il MLE e con il MDA (con distanze K-S)	91
4.2	Confronto tra le stime MDA ottenute con le diverse funzioni di distanza	92
4.3	Confronto tra le stime MDA Manhattan e distanze relative	96
4.4	Particolare del confronto tra le stime MDA Manhattan e distanze relative	98
4.5	Confronto tra le stime MDA ottenute con R1, R2, R3 e K-S	104
4.6	Confronto tra le stime MDA ottenute con R4, R2 e K-S vincolati	105
4.7	Confronto tra le stime MDA ottenute con MH, R2 ed R3 liberi	106
4.8	Confronto tra le stime ottenute con la massima verosimiglianza ed R5 vincolati	107
4.9	Distribuzione del costo sinistri aggregato	108
4.10	Distribuzione del costo sinistri aggregato	109
4.11	Distribuzione del costo sinistri aggregato	110
4.12	Distribuzione del costo sinistri aggregato	111
5.1	Bias delle stime di σ_1	119
5.2	Bias delle stime di σ_1	121
5.3	QQ-plot	124
5.4	QQ-plot	132

Elenco delle tabelle

2.1	Parametri σ forniti dal QIS5 per l'approccio market wide	23
2.2	SCR stimati con l'approccio market wide	29
2.3	Valori assunti da c per le LoB del modulo Health	30
2.4	SCR stimati con l'approccio undertaking specific	34
2.5	SCR stimati con il terzo approccio undertaking specific, con i dati ANIA per la stima del fattore di disturbo	35
3.1	Valori delle statistiche dei modelli	62
3.2	Ranking dei modelli	63
3.3	Riassuntivo numero sinistri, al netto dei sinistri IBNR	67
3.4	Triangolo dei sinistri denunciati, valori incrementali	68
3.5	<i>Link ratios</i>	69
3.6	Triangolo dei sinistri denunciati con stima dell'evoluzione, valori cumulati	69
3.7	Numeri dei sinistri riportati alle grandezze del 2012	69
3.8	Costi medi dei sinistri osservati per anno di accadimento ed anno di denuncia	74
3.9	Statistiche significative del costo aggregato dei sinistri, valori in milioni di euro	78
3.10	Statistiche significative del costo aggregato dei sinistri includendo l' <i>outlier</i> , valori in milioni di euro	79
4.1	Valori stimati attraverso il MDA	89

4.2	<i>Bias</i> tra le stime libere ed i valori veri	89
4.3	<i>Bias</i> tra le stime vincolate e i valori veri	90
4.4	Confronto tra i metodi di stima: coefficienti angolari e medie degli scostamenti quadri	101
4.5	Confronto tra le stime sotto ipotesi di Log-Normalità	103
4.6	Confronto tra le stime sotto ipotesi di Burr Inversa	106
4.7	Statistiche significative del costo aggregato dei sinistri, valori in milioni di euro	112
5.1	Stime vincolate del miscuglio di Log-Normali	122
5.2	Stime libere del miscuglio di Log-Normali	123
5.3	Confronto tra le stime sotto ipotesi di miscuglio di Log-Normali	123
5.4	Stime vincolate del miscuglio di Log-Normali	131
5.5	Statistiche significative del costo aggregato dei sinistri, valori in milioni di euro	133
A.1	Riassuntivo numero sinistri depurato dal fattore di disturbo	137
A.2	Adattamento dei modelli di tipo Poisson Composta al numero dei sinistri	139
A.3	Statistiche significative del costo aggregato dei sinistri sotto le ipotesi al vaglio nel Capitolo 3, senza il sinistro <i>outlier</i> , valori in milioni di euro . .	140
A.4	Statistiche significative del costo aggregato dei sinistri sotto le ipotesi al vaglio nel Capitolo 3, considerando anche il sinistro <i>outlier</i> , valori in milioni di euro	140
A.5	Statistiche significative del costo aggregato dei sinistri sotto le ipotesi al vaglio nel Capitolo 4, valori in milioni di euro	141
A.6	Statistiche significative del costo aggregato dei sinistri sotto le ipotesi al vaglio nel Capitolo 5, valori in milioni di euro	141

B.1	Statistiche significative del costo aggregato dei sinistri sotto l'ipotesi di distribuzione Log-Normale per la <i>severity</i> per dimensioni campionarie crescenti, valori in milioni di euro	147
B.2	Statistiche significative del costo aggregato dei sinistri sotto l'ipotesi di distribuzione di Burr Inversa per la <i>severity</i> per dimensioni campionarie crescenti, valori in milioni di euro	148
B.3	Deviazione standard di VaR e Tail-VaR sotto le ipotesi di distribuzione Log-Normale e Burr Inversa per la <i>severity</i> per dimensioni campionarie di 1000 e di 10000 unità	149