



Università degli Studi di Trieste

DIPARTIMENTO DI SCIENZE ECONOMICHE, AZIENDALI, MATEMATICHE E STATISTICHE
BRUNO DE FINETTI
CORSO DI LAUREA MAGISTRALE IN SCIENZE STATISTICHE E ATTUARIALI

TESI DI LAUREA MAGISTRALE

Reverse sensitivity analysis: alcune applicazioni assicurative e finanziarie

Laureando:
Bettini Alberto

Relatore:
Prof. Pietro Millosovich

Alla mia famiglia e Cinzia

Indice

Introduzione	1
1 Stress testing: il settore assicurativo e bancario	5
1.1 I rischi assicurativi e finanziari	6
1.2 La direttiva Solvency II	7
1.2.1 Valutazione di attività e passività	8
1.2.2 I requisiti di capitale	10
1.2.3 Own Risk and Solvency Assesment	12
1.3 La direttiva Basel III	13
1.3.1 Dettagli sul primo pilastro	14
1.3.2 Alcune delle criticità di Basilea II	16
1.4 Stress Testing, una panoramica	18
1.4.1 I soggetti: imprese e supervisor	19
1.4.2 Analisi di sensitività	20
1.4.3 Un confronto tra settore assicurativo e bancario	22
1.4.4 EIOPA Stress Test 2018	23
2 Il modello per l'analisi di sensitività	27
2.1 Misure di rischio	29
2.2 Reverse sensitivity analysis	31
2.2.1 Vincolo espresso dalla misura stressata	33
2.2.2 Vincolo espresso in termini di VaR	35
2.2.3 Vincolo espresso in termini di VaR e ES	40
2.3 Algoritmo di simulazione Monte Carlo	41
2.4 Misurare la sensitività	42
3 Reverse sensitivity analysis: un portafoglio di rendite vitalizie	47
3.1 Struttura generale del modello stocastico	50
3.2 Il modello stocastico per la mortalità	53
3.2.1 Ipotesi numeriche comuni e calcolo del premio	58

3.3	Modello 1. Mortalità e tassi deterministici	59
3.3.1	Analisi di sensitività	59
3.3.2	Modello 1 - Uno stress alternativo, VaR e ES	66
3.3.3	Modello 1 - Stress avverso all'assicuratore	68
3.4	Modello 2. Scenari multipli per la mortalità	72
3.4.1	Analisi di sensitività	73
3.5	Modello 3. Scenari multipli per la mortalità con tasso aleatorio	78
3.5.1	Analisi di sensitività	80
3.6	Modello 4. Approccio pienamente stocastico	87
3.6.1	Modello 4 - analisi di sensitività	87
4	Reverse sensitivity analysis: il rischio di credito	91
4.1	Modelli Mistura	94
4.1.1	Modello mistura bernoulliano	96
4.1.2	Modello Beta-Binomiale	98
4.2	Struttura generale del modello stocastico	100
4.3	Modello di base	102
4.3.1	Analisi di sensitività - Modello base	106
4.4	Modelli alternativi	110
4.4.1	L'effetto dell'esposizione	111
4.4.2	La dipendenza within	112
4.4.3	La dipendenza between	114
4.4.4	Loss given default stocastico	117
	Conclusione	121
A	Nozioni fondamentali per l'analisi di sensitività	123
A.1	Funzione inversa generalizzata	124
A.2	Derivata di Radon-Nikodym	126
A.3	Divergenza di Kullback - Leibler	129
A.4	Relazioni di dipendenza tra variabili aleatorie	131
A.4.1	Copule	134
A.4.2	Misure di dipendenza	139
B	Modelli stocastici per la mortalità	145
B.1	Il modello di Lee - Carter	145
B.2	Estensioni al modello di Lee-Carter	148
B.2.1	Esempi di modelli della famiglia GAPC	151
B.3	Un confronto sui dati italiani 1950 - 2014	154
	Bibliografia	159

Introduzione

Lo sviluppo di modelli quantitativi è di fondamentale importanza nell'era moderna per permettere agli operatori economici di effettuare scelte razionali. L'incremento della potenza computazionale ha reso più semplice l'implementazione di articolati sistemi di calcolo che consentono di sintetizzare efficacemente grandi quantità di informazioni. Questi modelli hanno raggiunto una complessità tale da rendere spesso molto difficoltosa una profonda comprensione del loro comportamento. Per analizzare efficacemente le relazioni tra le variabili del modello e poter conseguentemente trarne informazioni utili, è necessario pertanto impiegare opportune metodologie. In quest'ottica, un approccio largamente impiegato è l'analisi di sensitività. Con questa espressione si intende lo studio di come l'incertezza dell'output di un modello matematico si possa attribuire alle varie fonti di aleatorietà in input. Nell'approccio classico, si impone un cambiamento all'input del modello attraverso l'impiego di un determinato scenario e si individua la variazione dell'output dovuta a tale cambiamento. Alternativamente, si può impiegare una logica invertita, in cui si individua lo stress su uno o più input che comporta una prefissata variazione dell'output. Si parla in questo caso di reverse stress testing. Per effettuare operativamente l'analisi è necessaria in entrambi casi la riesecuzione del modello considerando le nuove ipotesi. Si comprende quindi che l'analisi di sensitività può risultare molto onerosa se l'esecuzione di un modello richiede lunghi tempi computazionali. Nonostante ciò, queste tecniche hanno assunto negli ultimi anni un ruolo predominante nella fase di costruzione e validazione di un modello.

I modelli interni impiegati in ambito assicurativo e bancario per la valutazione dei requisiti patrimoniali sono usualmente complessi al punto da soffrire

dei problemi sopra descritti. In questo ambito, alla luce delle nuove direttive europee, è usuale l'impiego di misure di rischio, come ad esempio il Value at Risk o l'Expected Shortfall, in qualità di importanti variabili decisionali e in particolare per il calcolo dei requisiti di capitale.

Il tema centrale della trattazione è un particolare metodo di analisi di sensitività, denominato *reverse sensitivity analysis*. Tale approccio è stato introdotto in Pesenti et al. (2019) e si inserisce nell'ambito del *reverse stress testing*. In questo contesto, un modello è individuato da un vettore aleatorio che descrive i fattori di rischio in input e da una funzione che mappa tale vettore sull'asse reale, determinando così l'output aleatorio del modello. Come nel *reverse stress testing* classico, si parte da uno stress fissato della variabile in output, e in particolare da una variazione critica di una misura di rischio associata a tale variabile. La novità dell'approccio consiste nel determinare lo scenario più plausibile, nel senso che minimizza l'entropia relativa rispetto a uno scenario base, che è consistente con lo stress imposto sulla variabile output. L'obiettivo principale è quindi comprendere l'impatto che lo scenario di stress comporta sui fattori di rischio, individuandone i più influenti. Contrariamente ad altre tecniche, un vantaggio importante di tale metodologia è che la sua implementazione non richiede la riesecuzione del modello.

La *reverse sensitivity analysis* è sviluppata nella trattazione in due distinti macroesempi. È fondamentale osservare che l'obiettivo generale non concerne la descrizione delle caratteristiche dei modelli impiegati nei due esempi, bensì lo studio della *reverse sensitivity analysis* e delle sue potenzialità. Il primo esempio considera la gestione, da parte di un assicuratore, di un portafoglio di rendite vitalizie omogenee sull'arco temporale di sopravvivenza degli assicurati. In questo caso, si assume come variabile output il fondo finale di portafoglio dell'assicuratore. I fattori di rischio aleatori sono individuati dall'effetto della mortalità (che include il rischio di longevità) e quello legato all'aleatorietà dei rendimenti delle attività che l'assicuratore ha posto a copertura della riserva di portafoglio. L'obiettivo è comprendere come uno scenario di stress avverso all'assicuratore, espresso tramite un incremen-

to della misura di rischio calcolata sul fondo finale, impatti sui fattori di rischio, ricavandone così un ranking. L'analisi è sviluppata attraverso quattro modelli innestati l'uno dentro l'altro. Complessivamente si è compreso che l'effetto legato al tasso di rendimento (rischio sistematico) è in generale più rilevante di quello legato alla mortalità (rischio di processo e sistematico).

Nel secondo esempio è analizzato un portafoglio di prestiti di un istituto bancario soggetti a rischio di insolvenza. Assumendo che la banca abbia suddiviso le proprie controparti in tre tranches in base alla loro qualità creditizia, si considera, come variabile output, la perdita aleatoria totale. Ognuna delle tre tranches è individuata da diverse esposizioni, probabilità di default e correlazioni. I fattori di rischio sono, con riferimento ad ogni tranche, i numeri di aleatori di insolvenze, le variabili latenti che determinano le probabilità di insolvenza e le perdite medie. L'obiettivo è in questo caso comprendere quale delle tre tranches sia complessivamente più responsabile per lo stress, e individuarne le cause. Un aspetto chiave del modello è la sua complessa struttura di dipendenza. I singoli default all'interno delle tranches sono tra loro dipendenti (dipendenza within), ma lo sono anche i numeri complessivi di default per tranche (dipendenza between). A partire da un modello base di riferimento, si sviluppano quattro studi mirati ad esaminare singoli aspetti del problema. L'indagine permette di apprezzare come la dipendenza, in particolare quella within, giochi un ruolo fondamentale per l'analisi di sensitività.

La tesi comprende quattro capitoli e due appendici. Nel primo capitolo è presentata una panoramica sulle direttive europee che sovrintendono il sistema assicurativo e quello bancario con l'obiettivo di comprendere la complessità di calcolo che queste comportano e il ruolo dello stress testing. Il secondo capitolo introduce la reverse sensitivity analysis nelle sue diverse forme, e le relative proprietà. Si considerano inoltre due misure di sensitività che consentono di quantificare numericamente l'impatto che lo stress ha avuto sui fattori di rischio. Nel terzo e quarto capitolo sono presentati i due esempi sopracitati. La tesi si conclude con due appendici nelle quali sono introdotti i fondamenti teorici per la comprensione di alcuni degli argomenti trattati.

Capitolo 1

Stress testing: il settore assicurativo e bancario

Il primo capitolo, di natura introduttiva, ha l'obiettivo di delineare gli aspetti fondamentali delle direttive vigenti che stabiliscono i requisiti di capitale per le imprese di assicurazione (Solvency II) e gli istituti bancari (Basel III). Particolare attenzione è posta sulla tecnica dello stress testing per comprendere lo sviluppo che essa ha conseguito negli ultimi anni e il suo ruolo nelle normative. Il capitolo è strutturato in quattro sezioni. La prima descrive, in termini generali, la centralità del concetto di rischio nelle due normative. Le due sezioni successive riguardano rispettivamente Solvency II e Basel III, si forniscono alcune nozioni chiave per comprendere la struttura delle due normative e inquadrarne gli scopi. In quest'ambito si presta particolare attenzione ai requisiti patrimoniali che imprese di assicurazione e istituti bancari devono detenere per fronteggiare le perdite future. Nell'ultima sezione è presentata una panoramica sullo stress testing. Si descrivono, in primo luogo, i diversi obiettivi della metodologia individuati a seconda del soggetto che implementa le analisi (i singoli istituti o i supervisori). Successivamente, nell'ambito dell'analisi di sensitività, si introduce, in termini generali, l'approccio denominato reverse stress testing, che è l'argomento centrale della tesi. Il capitolo si conclude con una breve descrizione di un esempio di esercizio di stress testing per le imprese di assicurazione.

1.1 I rischi assicurativi e finanziari

L'attività di banche e imprese di assicurazione è completamente permeata da incertezza a causa della natura del loro ciclo produttivo. Per questa ragione, tali soggetti economici devono sviluppare opportune metodologie statisticomatematiche finalizzate alla persecuzione di strategie economiche per la gestione dei rischi a cui sono esposte. In questo ambito, per la definizione di rischio, si usa distinguere tra rischi puri e rischi speculativi. I primi riguardano eventi le cui conseguenze possono impattare negativamente sull'abilità di un soggetto economico di conseguire obiettivi e strategie pianificati; essi sono pertanto pertinenti solo ad eventi avversi. I rischi speculativi sono invece collegati ad eventi che possono comportare, per il soggetto economico preso in considerazione, sia conseguenze negative che positive. Il Risk Management è la disciplina che si occupa di identificare, quantificare e descrivere le tipologie di rischio (in particolare i rischi puri) a cui l'impresa è potenzialmente esposta. Sebbene sia in primo luogo interesse delle imprese affrontare opportunamente i rischi per garantire una sana e corretta gestione, in Europa sono attualmente vigenti una normativa per le imprese di assicurazione (Solvency II) e una per le banche (Basel III) che sovrintendono, regolano e controllano il sistema di gestione dei rischi dei singoli istituti. Le due normative, seppur diverse in molti aspetti, hanno comuni intenti. In entrambe le normative il concetto di rischio è centrale, le imprese devono dotarsi di organizzazione interna e mezzi propri per valutare e monitorare nel tempo la propria posizione di rischiosità. Un obiettivo comune di carattere generale è l'unificazione del mercato europeo. Esso è messo in atto spingendo le imprese ad adottare comuni approcci per renderle confrontabili in modo trasparente. In secondo luogo vi è l'intento di tutelare i consumatori imponendo alle imprese precisi vincoli nella valutazione delle poste a bilancio. In particolare, sia le imprese di assicurazione che gli istituti bancari, devono detenere un ammontare di capitale proprio (chiamato capitale di solvibilità) in modo da poter fronteggiare future perdite con un alto livello di probabilità. Nelle due prossime sezioni si esporranno sinteticamente alcuni degli aspetti principali delle due normative.

1.2 La direttiva Solvency II

Per le imprese di assicurazione è attiva in Europa, dal 1 gennaio 2016, una nuova direttiva chiamata Solvency II (ufficialmente Direttiva 2009/138/CE, Commissione Europea,(2009)) che ha rivisitato profondamente la precedente vigilanza prudenziale. L'idea centrale della normativa è la persecuzione di un approccio basato sul rischio (risk based). Le imprese devono tenere in considerazione e gestire quindi tutte le fonti di incertezza alle quali sono esposte, in particolare esse devono fornire un resoconto sulla quantità di capitale necessaria (Solvency Capital Requirement) a mantenerle solventi in caso di avverarsi di scenari estremi. L'attenzione della nuova normativa non è legata esclusivamente ai criteri quantitativi per il calcolo del requisito di solvibilità, ma è più in generale rivolta al complesso sistema di regole a presidio della stabilità delle imprese; è chiaro quindi come sia necessario e indispensabile per queste imprese adottarsi di complessi modelli per le valutazioni finanziarie ed attuariali. Si rammenti che prima del 2016 era attiva in Europa dagli anni 70' un'altra direttiva comunemente denominata Solvency I. Il passaggio alla nuova normativa è stato motivato dall'incapacità di Solvency I di

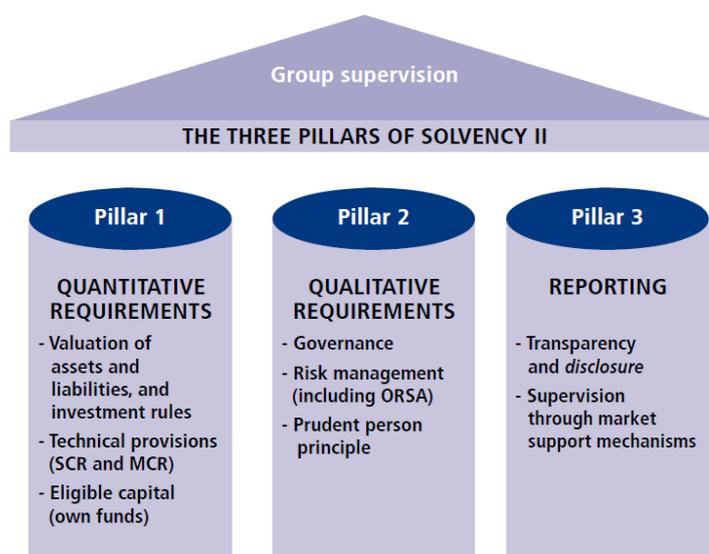


Figura 1.1: La struttura a tre pilastri della normativa europea Solvency II (fonte: IVASS).

rappresentare correttamente il rischio a cui le imprese di assicurazione sono esposte. A titolo di esempio, il capitale di solvibilità era valutato tenendo in considerazione esclusivamente i rischi di natura assicurativa e non quelli finanziari (rischio di credito, di mercato, etc.) che le imprese di assicurazione devono affrontare.

La direttiva Solvency II si sviluppa su tre pilastri (Figura 1.1). Nel primo pilastro sono definiti i requisiti di carattere quantitativo. In particolare sono presentate le metodologie per la valutazione del capitale di solvibilità e delle riserve tecniche. Il secondo pilastro concerne il sistema dei controlli interni. La direttiva impone alle imprese di dotarsi di un opportuno sistema di governance interno per una corretta gestione dei rischi a cui sono potenzialmente esposte. Nel terzo pilastro sono delineate infine le metodologie attraverso le quali l'impresa deve comunicare con l'esterno. Si impone all'impresa di essere completamente trasparente sia nei confronti degli istituti di vigilanza sia nei confronti di sottoscrittori e azionisti. Procedendo dal primo pilastro descriviamo brevemente alcuni aspetti fondamentali della normativa.

1.2.1 Valutazione di attività e passività

Per la valutazione delle poste a bilancio (attività e passività) il principio generale richiesto dalla direttiva è che le procedure per la stima siano consistenti con il mercato. Le attività e le passività devono ammontare dunque all'importo al quale potrebbero essere scambiate sul mercato tra parti consapevoli e consenzienti sotto situazioni di mercato non stressate. La Figura 1.2 rappresenta, in forma schematizzata, lo stato patrimoniale di un'impresa di assicurazioni in Solvency II. Per quanto concerne le attività è evidente dunque che, essendo la maggior parte di esse scambiate regolarmente nei mercati, possono essere inserite nel bilancio osservando semplicemente, nell'istante d'interesse, il valore delle medesime (approccio mark-to-market). Per alcune tipologie di attività per le quali non è presente un esplicito valore di mercato si può impiegare un opportuno modello per la valutazione (approccio mark-to-model). Come per ogni valutazione, la regola generale, in ogni

caso, è quella di seguire l'approccio della persona prudente. Per la valutazione market-consistent delle riserve tecniche, che rappresentano per l'assicuratore la principale voce delle sue passività e non sono scambiate sui mercati, Solvency II delinea una precisa e semplice regola. Il valore delle riserve deve essere pari all'ammontare che l'assicuratore dovrebbe pagare qualora egli cedesse l'intera passività ad una controparte sul mercato. Operativamente, per la valutazione è necessario scomporre la riserva in due componenti. La prima componente cosiddetta Best Estimate rappresenta la miglior valutazione (nel senso più realistica possibile) che l'assicuratore può effettuare per la propria posizione debitoria. Pertanto questa valutazione, effettuate tramite modelli attuariali, non deve contenere nessun marginale prudenziale. Da aggiungersi a questa vi è una seconda componente chiamata Risk Margin che rappresenta il premio (compenso teorico) che l'assicuratore richiede per mantenere un margine di solvibilità a fronte degli impegni assunti. Per la valutazione del Risk Margin si impiega l'approccio del costo del capitale. La valutazione deve dunque corrispondere al costo, in termini di capitale, necessario per costituire fondi propri tali da far fronte alle obbligazioni dell'assicuratore per tutta la loro durata di vita (ottica di run-off).

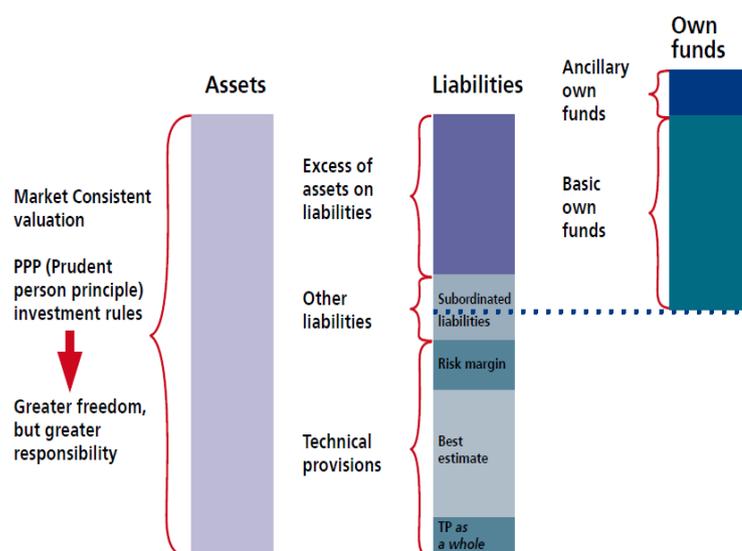


Figura 1.2: Struttura dello stato patrimoniale in Solvency II (fonte: IVASS).

Dalla differenza tra attività e passività si delineano i fondi propri dell'assicuratore. Una componente fondamentale dei fondi propri dell'assicuratore è costituita dal capitale di solvibilità.

1.2.2 I requisiti di capitale

Solvency II delinea due distinti requisiti di capitale: il Minimum Capital Requirement (MCR) e il Solvency Capital Requirement (SCR). Il Solvency Capital Requirement è determinato come capitale economico che le imprese devono detenere per garantire, con una probabilità del 99.5%, di onorare gli obblighi assunti nei confronti delle controparti nei successivi dodici mesi. Operativamente, per la valutazione di tale quantità, la direttiva propone diversi approcci (si veda la Figura 1.3 per una rappresentazione schematica in ordine di complessità). Il primo è un approccio semplificato (cd. formula standard) di tipo modulare; con l'utilizzo di opportuni parametri calibrati sul mercato europeo (EIOPA (2010)). In alternativa le imprese di assicurazione possono sviluppare un modello interno (completo o parziale) calibrato sulle proprie specificità. In tutti i casi, la stima del SCR si basa sul Value at Risk. Per dettagli su questa misura di rischio si veda la prima sezione del prossimo capitolo. Il modello interno, che descrive quantitativamente tutti i rischi a cui l'impresa è esposta, può essere impiegato solo previa approvazione da par-

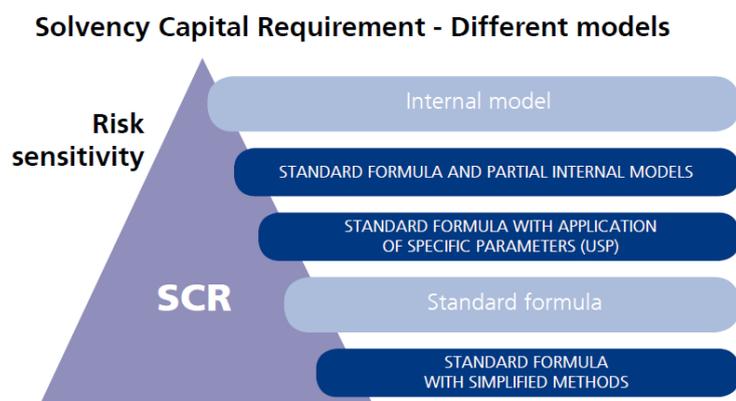


Figura 1.3: I differenti approcci per la valutazione del Solvency Capital Requirement (fonte: IVASS).

te dell'istituto di vigilanza. Si tenga in considerazione che, come è evidente, l'approccio del modello interno individua al meglio le specificità dell'impresa; la costruzione di questo è però in generale estremamente onerosa e per istituti di dimensioni ridotte generalmente non realizzabile. Qualora la formula standard venga ritenuta non adeguata (essa ha in generale molte limitazioni) e l'assicuratore non disponga dei mezzi economici necessari per sviluppare un modello interno, sono possibili soluzioni intermedie. L'assicuratore può infatti impiegare nella standard formula, al posto dei parametri calibrati sul mercato europeo, gli Undertaking Specific Parameters ovvero dei parametri calibrati internamente dall'impresa. Si rammenti ad ogni modo che, qualsiasi approccio si scelga, il Solvency Capital Requirement deve tener conto di tutte le fonti di rischio a cui l'impresa è esposta. Per un'impresa di assicurazione, oltre ai rischi tipici della sua attività (rischi assicurativi), sono presenti per Solvency II anche tutti i rischi legati al mercato finanziario (rischi finanziari) e infine quelli operativi (Figura 1.4). Si noti che in generale, seppur secondariamente, sono presenti anche altre tipologie di rischi che non consideriamo in questa trattazione. In generale la valutazione del SCR, anche tramite approccio standard, è un processo laborioso e complesso, per questa ragione esso deve essere obbligatoriamente calcolato con cadenza annuale. Per garantire e monitorare lo stato di salute dell'impresa è stato introdotto il Minimum Capital Requirement (MCR) che, a differenza del SCR, va calcolato con cadenza trimestrale. Operativamente esso si basa su un'ottica Value

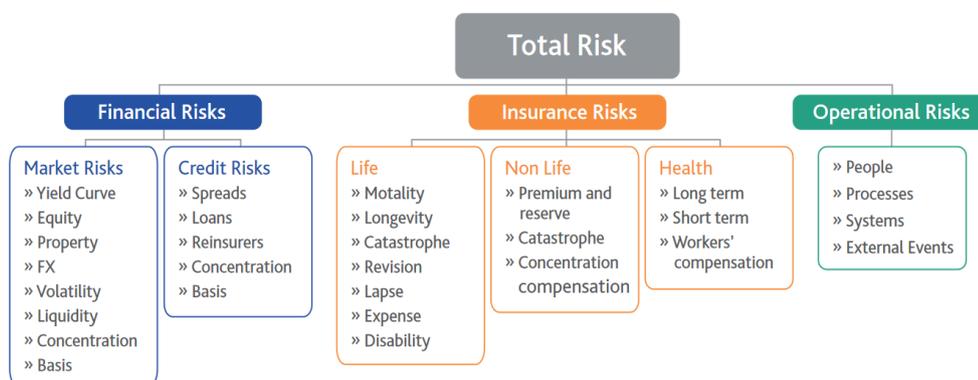


Figura 1.4: I principali rischi che deve affrontare un'impresa di assicurazioni (fonte: Moody's Analytics).

at Risk al 85% e consiste in una percentuale dei fondi propri che dipende da alcuni parametri; è compreso in genere tra il 25% e il 45% del SCR. Qualora l'impresa non detenga un capitale sufficiente da coprire il requisito minimo è considerata essere in grave pericolo di fallimento.

1.2.3 Own Risk and Solvency Assesment

Il secondo pilastro di Solvency concerne il complesso sistema di governance aziendale di cui l'impresa deve dotarsi per poter far fronte ai rischi assunti. L'ORSA costituisce una parte fondamentale del secondo pilastro; esso comprende l'insieme di processi che permette all'assicuratore di analizzare le specificità dei rischi che deve fronteggiare in modo da poter prendere decisioni strategiche. L'ORSA rappresenta pertanto, in primo luogo, uno strumento imprescindibile per il risk management dell'impresa e inoltre ha un ruolo fondamentale anche per il controllo delle valutazioni e dei processi da parte degli organismi di vigilanza. Tra le analisi più importanti che Solvency II richiede alle imprese nell'ORSA ci sono: una definizione precisa dell'appetito al rischio dell'impresa, il calcolo dei requisiti patrimoniali tramite formula standard o modello interno (o approcci intermedi), la proiezione del bilancio aziendale su un intervallo di tre - cinque anni sia per le attività che per passività basato su diversi scenari (stress testing) e l'analisi della qualità dei dati impiegata (data quality). Per una visione schematica si veda la Figura 1.5. Si osservi che, qualunque approccio venga impiegato per il calcolo del SCR, l'istituto deve anche effettuare una valutazione sulla solidità delle ipotesi sot-



Figura 1.5: I processi di ORSA (fonte: Moody's Analytics).

tostanti al modello. La pratica dello stress testing permea completamente nel tessuto di ORSA in quanto l'assicuratore la impiega sia per valutare i requisiti di capitale sia nella proiezione del bilancio con i vari scenari futuri. Ad integrazione dei rischi di natura quantitativa presenti nel requisito patrimoniale è possibile effettuare un'analisi qualitativa sulla materialità dei rischi a cui l'azienda è esposta. Si rammenti che, per ogni valutazione, vanno discusse in modo approfondito le ipotesi sottostanti al calcolo e giustificate in ogni aspetto le procedure di stima. Le valutazioni devono inoltre essere coerenti alle strategie aziendali e con l'appetito al rischio.

1.3 La direttiva Basel III

In risposta alle evidenti problematiche del sistema bancario evidenziate dalla crisi finanziaria del 2007/08 il Comitato di Basilea, che sotto la supervisione della Banca dei Regolamenti Internazionali (BIS) monitora il sistema bancario europeo, ha proposto di riformare la normativa allora vigente (Basel II) instaurando una nuova direttiva denominata Basel III (BCBS (2010)). Nonostante il comitato non abbia in realtà capacità di regolamentazione autonoma, quasi tutti gli enti supervisor degli stati membri accolgono gli accordi tenendo conto delle proprie specificità. La stesura del progetto Basel III è iniziata nel 2010, a seguito di ripetuti rinvii, la direttiva entrerà in vigore gradualmente a partire dal 2022. L'obiettivo generale della normativa è incrementare la solidità finanziaria del sistema bancario imponendo precisi vincoli quantitativi che incrementino la liquidità delle banche e ne riducano la leva finanziaria. Come si è menzionato in precedenza, un'impresa di assicurazione è esposta ai rischi assicurativi, finanziari e operativi (Figura 1.4). Una banca, di conto suo, è esposta esclusivamente agli ultimi due. Le principali criticità per il risk management delle banche sono dunque rappresentate dal rischio di credito, di mercato e i rischi operativi. Si discuterà di alcune delle differenze tra le due tipologie di istituti nella sezione 1.4.3. Operativamente, lo scopo della normativa consiste nel rafforzare i punti deboli di Basel II senza alternarne drasticamente la struttura. Pertanto, in linea con Basel II, la direttiva Basel III è scomponibile in tre pilastri, si veda a tal fine la

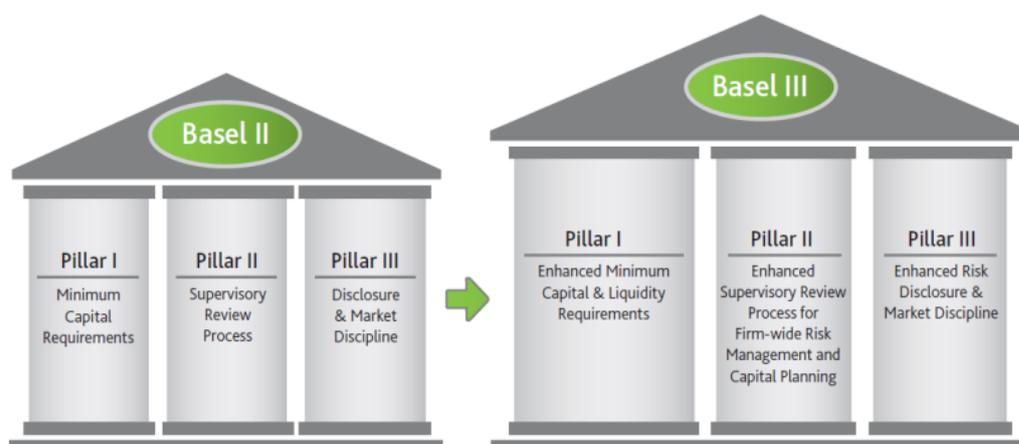


Figura 1.6: Schema della struttura delle direttive europee Basel II e Basel III.

Figura 1.6. Il primo pilastro concerne i requisiti patrimoniali minimi richiesti alle banche a fronte dei rischi suddetti. In quest'ambito, oltre a richiedere requisiti più stringenti, sono stati introdotti in Basel III anche vincoli concernenti la leva finanziaria e la liquidità minima che l'istituto deve detenere. Il secondo pilastro delinea, in primo luogo, i processi interni tramite i quali una banca deve monitorare l'adeguatezza del capitale allocato e inoltre le metodologie che gli organi di vigilanza possono impiegare per verificare i processi dei singoli istituti. Il terzo pilastro, come in Solvency II, riguarda la disciplina di mercato e la comunicazione con l'esterno, con particolare riferimento agli istituti di vigilanza. Analizziamo ora alcuni aspetti fondamentali del primo pilastro.

1.3.1 Dettagli sul primo pilastro

Introduciamo in questa sezione alcuni dei principali aspetti quantitativi che il primo pilastro di Basel III richiede con riferimento ai requisiti patrimoniali, alla leva finanziaria e alla liquidità degli istituti. Il requisito di capitale è basato sul rapporto tra due quantità: fondi propri e Risk-Weighted Assets (RWA). Il rapporto è soggetto a diverse limitazioni inferiori in base al capitale (numeratore) che si sta considerando. A seconda della qualità del capitale il patrimonio di una banca può essere scomposto in Tier 1 Capital

e Tier 2 Capital. Il Tier 1 Capital è individuato dal cosiddetto Core Capital che consiste in capitale azionario e dalle riserve palesi (retained earnings). Questa tipologia di capitale, a differenza del Tier 2 Capital, può essere in generale impiegato dalla banca per coprire eventuali perdite senza dover necessariamente compromettere la sua attività. Il Tier 2 Capital è considerato essere più rischioso del Tier 1 in quanto composto principalmente da attività più complesse e meno liquide. Esiste inoltre una classe di capitale aggiuntiva chiamata Additional Tier 1 Capital consistente in titoli perpetui (senza scadenza). Le attività finanziarie della banca sono considerate in virtù della loro rischiosità nel RWA. Tale quantità può essere in generale valutata tramite approcci standard o modelli interni e deve riflettere la complessiva rischiosità delle attività dell'istituto tenendo adeguatamente conto delle tre fonti principali di rischio: rischio di credito, rischio di mercato e rischio operativo. L'idea centrale che guida tale valutazione è che qualora la banca detenga attività più sicure, essa ha bisogno di meno capitale per far fronte agli impegni assunti. Considerato il rapporto tra capitale e RWA, i vincoli minimi imposti da Basel III, sono i seguenti: 4.5% con riferimento al solo capitale Tier 1 (minimum Common Equity Tier 1), 6% considerando Tier 1 e Additional Tier 1 Capital (minimum Tier 1 Capital) e infine 8% per tutte le classi di capitale (Overall Regulatory Capital). In generale, nel passaggio tra Basel II a Basel III le percentuali sono aumentate (ad esempio il minimum Common Equity Tier 1 ammontava a 4% in Basel II); sono stati introdotti inoltre ulteriori ammortizzatori da aggiungere ai vincoli appena citati. Il primo è mandatario e consiste in un aggiuntivo 2.5% del RWA (capital conservation buffer) e dunque porta la limitazione globale concernente il Capitale Tier 1 (senza considerare quello aggiuntivo) a 7%. Il secondo, volto a far fronte alle crisi cicliche, è di natura discrezionale.

Gli istituti bancari fanno estensivamente ricorso al debito per finanziarsi, in quanto ciò risulta conveniente sia in termini di capitale di rischio sia per i vantaggi fiscali che esso genera. Come si è evidenziato durante la crisi finanziaria, in conseguenza a gravi svalutazioni delle attività delle banche, l'impiego di leve finanziarie eccessive può generare forti problematiche. Fi-

nanziare le proprie attività con debito permette in generale di moltiplicare ricavi e perdite. Pertanto, qualora il costo del finanziamento superi i ricavi che la banca riesce a generare dalle attività finanziate, l'istituto consegue gravi perdite. Per limitare l'impiego di eccessive leve finanziarie, Basell III ha introdotto la seguente limitazione: il capitale di tipo Tier 1 deve costituire almeno il 3% dell'esposizione totale a cui l'istituto è soggetto. In ultima istanza, Basel III ha introdotto anche un vincolo per la gestione del rischio di liquidità. In particolare è richiesto agli istituti bancari di detenere attività liquide di alta qualità per far fronte ai cash flow in uscita pianificati per i futuri 30 giorni.

1.3.2 Alcune delle criticità di Basilea II

Gli istituti bancari operano storicamente suddividendo le proprie attività in due macroclassi: il portafoglio bancario e il portafoglio di negoziazione. Questa distinzione, in passato piuttosto opaca, comporta importanti conseguenze sia dal punto di vista del bilancio sia per la gestione dei rischi. Come si è evidenziato dal comportamento di molti istituti, ciò ha permesso l'instaurarsi di condizioni per le quali fosse conveniente, in termini di risparmio di capitale, la migrazione di un'attività da un portafoglio ad un altro (BCBS (2013)). Si parla in questo caso di arbitraggi normativi, in quanto si sfruttano le lacune delle direttive per trarre vantaggi economici. Per questa ragione un obiettivo importante delle nuove direttive di Basilea è stato tracciare delle linee guida per rendere più chiara la distinzione tra le due categorie e pertanto più trasparente il confronto tra gli istituti. Una comprensione approfondita delle differenze tra i due portafogli comporterebbe necessariamente una discussione tecnica al di là degli obiettivi del capitolo. In termini generali, il portafoglio di negoziazione comprende le attività, solitamente con alto grado di liquidità, che la banca considera disponibili ad essere scambiate sul mercato. Gli obiettivi preposti per tali attività possono essere molteplici, tra cui ad esempio: generare profitti da movimenti della curva dei tassi e impiegare le attività per consolidare posizioni di hedging. Tipicamente le attività del portafoglio di negoziazione sono valutate a valore di mercato. Il portafoglio

bancario, d'altro canto, contiene le attività destinate ad essere tenute in portafoglio fino a scadenza, come ad esempio (solitamente) i prestiti. Questo tipo di attività sono valutate a costo storico. La semplicistica dicotomia che questa differenziazione rappresenta può comportare gravi errori nella descrizione della rischiosità delle attività.

In termini generali è evidente che il portafoglio di negoziazione sia più esposto al rischio di mercato mentre il portafoglio bancario al rischio di credito. La crisi finanziaria del 2007/08 ha dimostrato quanto gli istituti bancari fossero sottocapitalizzati anche a causa delle lacune delle direttive che concedevano arbitraggi normativi. Un primo aspetto è dovuto, come già si menzionava, alla diffusa pratica di trasferire ad esempio titoli cartolarizzati complessi (come i CDO) fortemente esposti al rischio di credito, dal portafoglio bancario a quello di negoziazione, con conseguente alleggerimento della quantità di capitale di solvibilità. Si rammenti infatti che per i due portafogli gli approcci impiegati in Basel II per la valutazione sono in generale diversi. Le conseguenze di questa deficienza hanno comportato la necessità da parte del comitato di apportare opportuni accorgimenti a Basel II introdotti nella cosiddetta direttiva Basel 2.5 (2009). I più importanti di questi sono relativi al rischio di credito del portafoglio di negoziazione che in precedenza veniva generalmente ignorato. Il nuovo approccio affianca alla valutazione del Value at Risk (99%) della profit/loss aleatoria a 10 giorni (che cattura il rischio di mercato), un'ulteriore valutazione che consideri la perdita aleatoria a un anno dovuta a default delle controparti e migrazioni di rating (per catturare il rischio di credito). Basilea III ha accolto le novità introdotte dalla versione 2.5, estendendone alcuni aspetti. Un aspetto molto interessante è attualmente ancora in discussione è la sostituzione del Value at Risk, come misura di rischio di riferimento, con l'Expected Shortfall per tenere meglio conto della coda superiore della distribuzione della perdita. Per un approfondimento teorico e operativo su questa tematica si veda Chang et al. (2017).

1.4 Stress Testing, una panoramica

Sebbene gli stress test abbiano assunto, negli ultimi anni, un ruolo centrale nella realtà operativa del mondo finanziario e assicurativo globale, non si pervengono definizioni consistenti che siano in grado di sintetizzare efficacemente il significato di queste pratiche senza creare fraintendimenti. Nell'ambito della tesi si impiegherà il termine stress test con connotazione generale e non tecnica qualora si faccia riferimento a qualsiasi tipo di analisi nella quale, attraverso l'impiego di un modello statistico con precise ipotesi sottostanti, si vuole comprendere la capacità di un sistema di fronteggiare un possibile scenario futuro. Gli scenari che si possono utilizzare in uno stress test sono di diversa natura. Si parla di scenari baseline qualora essi rappresentino una futura situazione economica ritenuta verosimile. Molto spesso le analisi più interessanti vengono riscontrate impiegando scenari avversi, ovvero utilizzando ipotetiche situazioni economiche future (fittizie o derivate dal passato) le quali ci si aspetta comportino una perdita per l'istituzione finanziaria presa in considerazione. In Figura 1.7 è presentato uno schema logico che descrive il ruolo generale dello stress testing all'interno del sistema aziendale. Gli obiettivi delle analisi possono essere in generale molteplici a seconda dei soggetti che le implementano. Analizziamo nella prossima sezione due distinti punti di vista quello delle imprese e quello dei supervisori.

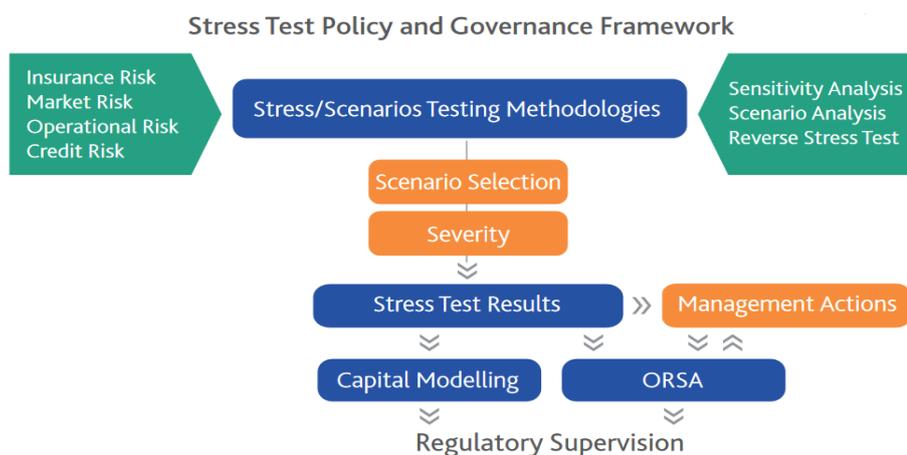


Figura 1.7: Schema logico sul ruolo dello stress testing (fonte: Moody's Analytics).

1.4.1 I soggetti: imprese e supervisori

A seconda del soggetto che conduce le analisi (imprese o supervisori) gli stress test possono essere impiegati per diversi scopi. In questo ambito si distingue in analisi (o test) di tipo bottom-up e top-down. Le analisi bottom-up sono effettuate dai singoli istituti tramite approcci, tecniche e modelli sviluppati internamente. Pertanto questo tipo di analisi, anche se possono essere richieste ed analizzate dai supervisori, sono specifiche per le singole imprese. Nelle analisi top-down, il supervisore, fissando scenari futuri e metodologie, effettua uno studio a livello globale senza coinvolgere le singole imprese. Analizziamo, in maggior dettaglio, i due punti di vista a confronto: supervisori e imprese.

Dal punto di vista di un ente supervisore, le analisi sono principalmente finalizzate a comprendere la robustezza del sistema economico nel suo complesso ad un determinato shock. Per ottenere informazioni utili in quest'ottica sono necessari studi di sistema che tengano in considerazione le interazioni che possono instaurarsi tra le imprese che operano nel mercato (logica top-down). Queste analisi vanno al di là della somma dei risultati dei singoli istituti e, vista la complessa struttura dei sistemi economici, non sono in generale di semplice attuazione. Si preferisce pertanto un approccio alternativo nel quale si impiegano informazioni concernenti le singole imprese e le si aggrega per ottenere risultati a livello di intero sistema (logica bottom-up). Le analisi bottom-up sono lo strumento fondamentale tramite il quale gli enti di supervisione esercitano la propria attività di controllo. È essenziale che gli stress test in questo ambito siano costruiti in modo modulare per tener conto delle dimensioni delle varie imprese operanti nel sistema. In conclusione al capitolo si presenterà un esempio di stress test bottom-up per le imprese di assicurazione europee. L'impiego simultaneo di entrambi i tipi di analisi consente ai supervisori di comprendere la natura del sistema economico sia come composto da singole entità che nel suo complesso. Questo tipo di conoscenza è in generale essenziale per gestire al meglio le crisi sistemiche. L'implementazione degli stress test è attuata dagli enti supervisori svilup-

pando opportuni scenari futuri che possono coincidere con situazioni già riscontrate nel passato o alternativamente avere una natura completamente ipotetica.

Per le imprese gli stress test rappresentano, in prima istanza, uno degli strumenti principali per le analisi quantitative del risk management. Sebbene, come si è detto, molti esercizi di studio siano imposti dai supervisori è interesse in primo luogo delle imprese comprendere la propria posizione di rischio per determinare opportune strategie e azioni di risk management. Nell'implementazione degli stress test le imprese devono tenere in considerazione molti aspetti. Per comprendere al meglio le vulnerabilità a cui sono esposte, le analisi devono essere effettuate in modo da essere compatibili con il business model aziendale e lo specifico appetito al rischio.

1.4.2 Analisi di sensitività

Una classica metodologia largamente utilizzata nell'ambito dello stress testing è l'analisi di sensitività (sensitivity analysis). Con questa espressione si intende lo studio di come l'incertezza dell'output di un modello matematico si possa attribuire alle varie fonti di aleatorietà in input, in particolare si parla di analisi di incertezza (uncertainty analysis) qualora si focalizzi l'attenzione sull'aspetto quantitativo del problema. Si tenga in considerazione che il termine può in generale essere fuorviante in quanto in alcuni ambiti viene impiegato con un'accezione più ristretta concernente le sole analisi che considerano variazioni di un singolo fattore di rischio in input. Gli obiettivi di questo tipo di studio sono molteplici, ad esempio: comprendere le complesse relazioni sussistenti tra le variabili in input e in output, identificare i fattori di rischio più influenti (factor prioritisation), controllare la robustezza dell'output del modello a variazioni anche poco rilevanti dell'input e per individuare aree del modello che necessitano miglioramenti.

Si usa distinguere, nell'ambito dell'analisi di sensitività, tra metodologie locali e globali. Le metodologie locali si focalizzano su un particolare punto nel

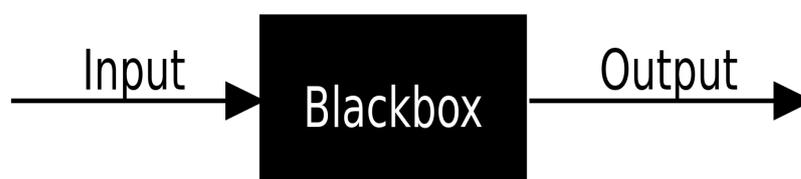


Figura 1.8: Schema logico di un modello a scatola nera.

dominio dello spazio degli input e si è in generale interessati a capire come a partire da questo punto si comporti l'output (usualmente attraverso l'impiego di derivate parziali). I metodi globali d'altro canto si concentrano non su un singolo punto bensì su un range di valori nello spazio dei fattori di input. In generale, per valutazioni in ambito stocastico (come nel caso di questa tesi), si utilizzano questo tipo di metodologie. Per questo tipo di analisi è presente in letteratura una mole enorme di differenti tecniche e approcci dovuti alla vasta varietà di discipline interessate e ai diversi fini per il quale il modello è costruito, si veda a tal fine Saltelli et al. (2008) e Borgonovo e Plischke (2016).

Nell'approccio classico all'analisi di sensitività si impone un cambiamento all'input del modello attraverso l'impiego di un determinato scenario (solitamente avverso) e si va ad individuare la variazione dell'output dovuta a tale cambiamento. Il successo di queste pratiche deriva dalla possibilità di studiare con semplicità il funzionamento di modelli complessi (come quelli impiegati in ambito assicurativo e bancario). Tale complessità impedisce in generale di analizzare il comportamento del modello tramite semplice intuizione; è necessaria pertanto una metodologia operativa. Prendendo a prestito una terminologia della teoria dei sistemi, si può pensare al modello come una scatola nera (black box); esso è dunque descrivibile solo tramite input e output e il suo preciso funzionamento interno non è visibile con chiarezza (Figura 1.8) I modelli interni impiegati dalle imprese di assicurazione oppure dagli istituti bancari, possono essere visti sotto quest'ottica. Questi modelli sono infatti essenzialmente il risultato dell'aggregazione di diverse linee di business, ognuna generalmente composta da un numero molto elevato di diverse posizioni.

Con il termine *reverse stress testing* si intende in generale una procedura finalizzata all'esplicita identificazione degli scenari (intesi come stress in input) che comportano una prefissata variazione dell'output. Nei modelli impiegati in ambito economico si considerano variazioni dell'output che possono compromettere un'impresa al punto da farle perdere credibilità e fiducia nel mercato. Il successo di questa metodologia, è cresciuto notevolmente negli ultimi anni in quanto rappresenta una logica alternativa alle procedure tradizionali. Per effettuare le analisi l'aspetto essenziale è definire diversi tipi di fallimenti (punti di rottura) del business model aziendale per andare a testare possibili scenari futuri che possono comportare il raggiungimento di tali fallimenti. A titolo di esempio si può pensare al fallimento del business model aziendale come all'incapacità di un'impresa di pagare dividendi agli azionisti per un lungo periodo di tempo. Nell'ambito di questa tesi si proporrà uno studio prima teorico (Capitolo 2) e poi operativo (Capitoli 3 e 4) di un particolare tipo di *reverse stress testing* denominato *reverse sensitivity analysis*.

1.4.3 Un confronto tra settore assicurativo e bancario

Con riferimento alle direttive Solvency II e Basel III le pratiche di stress testing costituiscono un elemento essenziale per le valutazioni quantitative delle posizioni di rischio. Molti degli approcci standard per le valutazioni, come ad esempio la formula standard per la valutazione del SCR in Solvency II, sono fondamentalmente basati su un'ottica di stress testing. La sostanziale differenza tra il sistema bancario e quello assicurativo in termini di stress testing è dovuta ai contenuti delle analisi. Le imprese di assicurazione sono soggette ad una peculiare esposizione nei confronti dei rischi assicurativi a cui le banche non sono esposte. I principali rischi di natura assicurativa per il settore Life sono rappresentati dal rischio di mortalità e il rischio di riscatto mentre per il settore Non-Life dal rischio di riservazione e dal rischio catastrofale. La presenza di queste fonti aggiuntive di rischio per le imprese di assicurazioni non comporta esclusivamente la necessità di

dover affrontare un maggior numero di analisi ma anche il problema di comprendere come i rischi di natura assicurativa siano correlati con gli altri. Si noti che operativamente il problema dell'aggregazione e diversificazione dei rischi è un aspetto centrale nelle valutazioni. È pertanto evidente che per effettuare stress test le imprese di assicurazione devono affrontare un livello di complessità aggiuntivo rispetto le banche.

Un'altra differenza importante è legata alla durata degli impegni dei due istituti. Un'impresa di assicurazioni che opera nel settore Life a fronte di premi incassati dispone di passività su orizzonti temporali molto lunghi. In contrasto, una banca finanzia con passività a breve termine (depositi) le sue attività a lungo termine (mutui e prestiti). Questo aspetto delinea due tipologie distinte di rischio a cui sono esposti i due istituti. Per le imprese di assicurazione (operanti nel settore Life) è necessario investire in attività con scadenze molto lunghe per coprire le posizioni di debito descritte dalle riserve matematiche, ciò incrementa generalmente rischi concernenti oscillazioni della curva dei tassi. Una banca è invece più esposta a problemi concernenti il rischio di credito, eccessive leve finanziarie e di carenza di liquidità.

1.4.4 EIOPA Stress Test 2018

Nel 2018, sotto la supervisione di EIOPA, è stato effettuato il quarto esercizio di stress test con l'obiettivo di valutare la robustezza del mercato assicurativo europeo a futuri scenari avversi (EIOPA (2018)). In particolare sono stati calibrati scenari economici tali da poter innescare rischi sistemici nel settore finanziario capaci di impattare anche sull'economia reale. L'indagine è stata effettuata su un campione di 42 imprese operanti nel mercato europeo, selezionate in modo da rappresentare un campione significativo.

L'esercizio è condotto con una logica bottom-up, dunque le analisi sono state effettuate dalle singole imprese impiegando ciascuna i propri mezzi. EIOPA ha fissato tre distinti scenari futuri di riferimento, si veda la Figura 1.9. L'obiettivo è valutare l'impatto che questi scenari hanno sul bilancio di gruppo delle singole imprese, con particolare riferimento ai fondi propri, e ai requisiti di solvibilità. In presenza di situazioni stressate agiscono simultaneamente

eventi tra loro dipendenti, ogni scenario comporta dunque la presenza di più fattori di rischio stressati contemporaneamente. Il primo scenario (yield curve up) è descritto da un improvviso incremento della curva dei tassi privi di rischio (shift parallelo) assieme ad un incremento inflativo. L'incremento dei tassi comporta la necessità per l'assicuratore di dover gestire un alto numero di riscatti nel suo portafoglio di polizze vita. In ultima istanza, a causa dell'aumento dell'inflazione di settore, l'assicuratore deve anche fronteggiare un problema concernente un insufficiente livello di riseriazione nel settore Non-Life. Il secondo scenario (yield curve down) è individuato da un protratto periodo di stagnazione dei tassi d'interesse, operativamente individuato da un decremento parallelo della curva dei tassi privi di rischio. In aggiunta, si assume che lo sviluppo di nuova tecnologia in campo medico incrementi significativamente la durata media di vita. L'assicuratore deve quindi ridefinire le proprie ipotesi per descrivere la mortalità (best estimate). Il terzo scenario è di natura catastrofale. Si immagina pertanto che, in un breve lasso temporale, si verifichino diverse catastrofi naturali che colpiscono diverse aree in Europa (2 terremoti, 4 uragani e 2 alluvioni); l'assicuratore deve pertanto risarcire i danni causati senza poter impiegare efficacemente trattati riassicurativi. I risultati aggregati delle analisi sono rappresentati nella Figura 1.10. Per il primo scenario il livello delle attività, che originariamente rappresentava il 109.5 % delle passività, si riduce al 107.6%. L'aumento della curva dei

SCENARIOS



3 scenarios, two encompassing a combination of market and insurance specific risks and one natural catastrophe scenario.

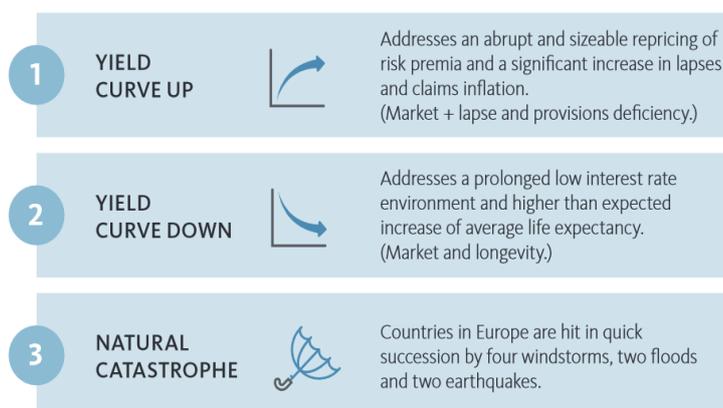


Figura 1.9: I tre scenari selezionati da EIOPA per l'analisi (fonte: EIOPA).

tassi riduce drasticamente il valore di mercato degli attivi dell'assicuratore. L'aumento della curva impatta anche decrementano le riserve dell'assicuratore, ma l'effetto non è abbastanza forte da compensare il primo. Molto spesso per rappresentare in termini relativi la solvibilità di un impresa si utilizza il Solvency Ratio. Esso è definito come il rapporto tra fondi propri e SCR espresso in termini percentuali. Se il ratio è inferiore a 100% l'assicuratore non ha sufficienti fondi propri per coprire il proprio requisito di solvibilità. Il livello medio del Solvency Ratio europeo pre-stress è pari a 202.4%. Conseguentemente allo stress implicato dal primo scenario esso si riduce a 145%. Il secondo scenario, è quello con conseguenze più gravi. Il decremento della curva dei tassi impatta incrementando le riserve dell'assicuratore che portano il livello globale di attività a 106.7% rispetto le passività. Una componente critica per l'assicuratore in questo scenario è individuata dai prodotti assicurativi con rendimenti garantiti che, anche in presenza di tassi molto bassi, devono sempre fornire al sottoscrittore il rendimento minimo contrattuale. Il Solvency Ratio si riduce in questo caso a 137.4%. In ultima istanza, lo scenario catastrofe ha un impatto nettamente più contenuto rispetto ai primi due. Ciò è principalmente dovuto alla presenza dei trattati riassicurativi.



Figura 1.10: Sintesi dei risultati del EIOPA Stress Test 2018 (fonte: EIOPA).

Capitolo 2

Il modello per l'analisi di sensitività

Viene presentato in questo capitolo il framework per l'analisi di sensitività introdotto in Pesenti et al. (2019) chiamato reverse sensitivity analysis. Consideriamo un decision maker, ad esempio un risk manager o un'autorità di supervisione, interessato all'impiego di un modello stocastico a supporto di una decisione strategica. Nell'ambiente standard utilizzato in letteratura per l'analisi di sensitività di tipo globale (Borgonovo and Plischke (2016), Saltelli et al. (2008)) si introduce, in prima istanza, un vettore aleatorio che comprende i fattori di rischio in input al modello stocastico. A questo vettore è assegnata una misura di probabilità iniziale (baseline) che rappresenta interpretativamente l'attuale stato di conoscenza del decision maker, ed una funzione (molto complessa) che lo trasforma in una singola variabile aleatoria, che è l'output del modello. La variabilità dell'output, in relazione a cambiamenti nei fattori di rischio, è l'oggetto della trattazione. La reverse sensitivity analysis è condotta imponendo uno stress sulla variabile in output, stress espresso tipicamente in termini di una misura di rischio; in particolare si considereranno il Value at Risk e l'Expected Shortfall. Si immagina che una variazione, in particolare un incremento, del valore della misura di rischio associata alla variabile in output comporti una situazione problematica per il decision maker; egli vuole quindi comprendere come lo stress impatti

sul modello, derivando quali sono i fattori di rischio che ne sono i maggiori responsabili.

Per comprendere un'applicazione dell'ambiente stocastico appena descritto si consideri, ad esempio, il caso di un risk manager di una multinazionale del settore assicurativo o bancario interessato a studiare un particolare modulo del modello interno che l'impresa utilizza per valutare la propria posizione di solvibilità. La perdita aleatoria (ad esempio in un anno) che l'impresa deve fronteggiare per il modulo considerato, è esprimibile in funzione dei fattori di rischio (risk driver) a cui quel settore è soggetto. La misura di probabilità baseline individua, in questo ambito, la distribuzione di probabilità che è impiegata dall'impresa per la descrizione probabilistica dei fattori di rischio. La posizione di solvibilità dell'impresa è usualmente valutata a partire da una misura di rischio calcolata sulla perdita aleatoria, ad esempio in ambito assicurativo tramite il Value at Risk. Il risk manager può dunque studiare, impiegando la reverse sensitivity analysis, l'impatto sulla distribuzione iniziale che una variazione critica della misura di rischio calcolata sulla perdita comporta derivando un ranking dei fattori di rischio che ne sono responsabili. Alternativamente si può considerare ad esempio un regulator che vuole analizzare la robustezza di un campione di imprese ad una particolare struttura di stress in linea con quella appena descritta. Come si analizzerà in dettaglio successivamente, l'implementazione del framework per la reverse sensitivity analysis non comporta la necessità di risimulare gli scenari sotto la misura stressata. Ciò è particolarmente importante per un regulator che, quasi sempre, potrebbe avere accesso solo ad un limitato set di simulazioni e non a tutto il modello. Questo aspetto può avere inoltre rilevanza anche qualora le simulazioni siano computazionalmente molto esigenti.

Il capitolo è strutturato in quattro sezioni. Nella prima sezione si introduce la fondamentale nozione di misura di rischio, considerando in particolare i due casi sopra menzionati. Nella seconda sezione viene presentato il modello per la reverse sensitivity analysis (Pesenti et al. (2019)), e sono esposti tre risultati importanti per lo studio, di cui i primi due, con le dimostrazioni.

Nella terza sezione è presentato l'approccio (algoritmo) per l'implementazione della reverse sensitivity analysis tramite simulazione Monte Carlo. Infine, nell'ultima sezione del capitolo, vengono introdotte due particolari misure di sensitività legate all'ambiente della reverse sensitivity analysis che consentono di valutare numericamente l'importanza (in termini di impatto sulla distribuzione causato dallo stress) di ogni fattore di rischio in input.

2.1 Misure di rischio

Le misure di rischio sono uno strumento fondamentale nell'ambito del risk management in quanto consentono di ridurre un fenomeno complesso come quello rappresentato da una variabile aleatoria ad un singolo numero reale. Nei loro confronti si è sviluppato negli ultimi anni un forte interesse dovuto all'importanza che queste misure hanno nella valutazione dei requisiti di capitale sia nelle normative europee delle imprese di assicurazioni (*VaR*) sia per quelle delle banche (*ES*). A causa del ruolo centrale che esse ricoprono, una parte prominente della letteratura (ad esempio Artzner et al. (1999), Föllmer and Schied (2011) e Acerbi (2002)) si è occupata estensivamente delle loro proprietà. Introduciamo ora la definizione generale di misura di rischio.

Considerato uno spazio misurabile (Ω, \mathcal{A}) , sia \mathcal{X} uno spazio lineare di variabile aleatorie definite su Ω e contenente le variabili aleatorie costanti. Una misura di rischio è un funzionale ρ definito su \mathcal{X} a valori in \mathbb{R} . Interpretativamente, una misura di rischio è un'applicazione che consente di mappare una variabile aleatoria sulla retta reale, fornendone quindi una sintesi. Prima di introdurre il Value at Risk e l'Expected Shortfall esponiamo alcune proprietà importanti che una misura di rischio dovrebbe possedere (Artzner et al.(1999)).

1. Subadditività: $\rho(X + Y) \leq \rho(X) + \rho(Y)$, per ogni $X, Y \in \mathcal{X}$.
2. Positiva omogeneità: $\rho(\lambda X) = \lambda\rho(X)$, per ogni $\lambda \geq 0$, $X \in \mathcal{X}$.

3. Monotonia: $\rho(X) \leq \rho(Y)$, per ogni $X, Y \in \mathcal{X}$, $X \leq Y$.
4. Invarianza per traslazioni: $\rho(X+m) = \rho(X)+m$, per ogni $m \in \mathbb{R}$, $X \in \mathcal{X}$.

Una misura di rischio che soddisfa queste 4 proprietà si dice coerente. Introduciamo ora il Value at Risk.

Definizione

Sia P una misura di probabilità definita su (Ω, \mathcal{A}) e $X \in \mathcal{X}$ una variabile aleatoria sul medesimo spazio. Fissato $\alpha \in (0, 1)$, si definisce Value at Risk di X a livello di confidenza α la seguente quantità:

$$VaR_\alpha^P(X) = \inf\{x \in \mathbb{R} : P(X \leq x) \geq \alpha\} \equiv F_X^{-1}(\alpha). \quad (2.1)$$

Nella definizione il simbolo F_X^{-1} indica la funzione inversa generalizzata. Il Value at Risk di X al livello $\alpha \in (0, 1)$ coincide dunque, per definizione, con la funzione inversa generalizzata di X calcolata in α . Per un approfondimento sulle proprietà della funzione inversa generalizzata si veda l'Appendice A.1. Si osservi che nella definizione non ci sono vincoli sul segno della variabile aleatoria X , è considerato quindi un generico rischio di tipo speculativo (la variabile aleatoria che lo descrive può assumere determinazioni sia positive che negative). Procediamo ora introducendo un'altra importante misura di rischio, l'Expected Shortfall.

Definizione

Sia P una misura di probabilità definita su (Ω, \mathcal{A}) e $X \in \mathcal{X}$ una variabile aleatoria sul medesimo spazio. Fissato $\alpha \in [0, 1)$, si definisce Expected Shortfall di X a livello di confidenza α la seguente quantità:

$$ES_\alpha^P(X) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_u^P(X) du.$$

A differenza del Value at Risk, l'Expected Shortfall tiene in considerazione l'intera coda della distribuzione di X . Si può infatti dimostrare che,

se la distribuzione di X è continua, allora l'Expected Shortfall può essere rappresentato alternativamente come:

$$ES_{\alpha}^P(X) = E_P(X|X > VaR_{\alpha}^P(X)).$$

Si tratta quindi della speranza matematica dei valori che eccedono il Value at Risk al medesimo livello di confidenza. L'Expected Shortfall è inoltre una misura di rischio coerente, mentre il Value at Risk non lo è in quanto non soddisfa, in generale, la proprietà di subadditività.

2.2 Reverse sensitivity analysis

Presentiamo, in prima istanza, il classico ambiente utilizzato nell'ambito dell'analisi di sensitività di tipo globale (Borgonovo & Plischke (2016), Saltelli et al. (2008)). Considerato uno spazio misurabile (Ω, \mathcal{A}) , indichiamo con \mathcal{P} l'insieme di tutte le misure di probabilità definite su tale spazio. Sia $\mathbf{X} = (X_1, \dots, X_n)$ un vettore aleatorio (a valori quindi in \mathbb{R}^n), che rappresenta i fattori di rischio in input del modello. Sia inoltre $g : \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione $(\mathcal{A}, \mathcal{B})$ -misurabile, chiamata funzione di aggregazione, che applicata al vettore aleatorio dei fattori in input \mathbf{X} restituisce la variabile aleatoria $Y = g(\mathbf{X})$ che rappresenta l'output del modello ed è la variabile aleatoria di interesse per il problema. Il comportamento di Y al variare dei fattori di rischio \mathbf{X} è l'oggetto di interesse nell'analisi di sensitività. In ambito finanziario ad esempio Y potrebbe rappresentare la perdita aleatoria a cui è esposta un'impresa che opera sul mercato a seguito di una determinata operazione. Consideriamo una misura di probabilità $P \in \mathcal{P}$, che determina in particolare la distribuzione iniziale di \mathbf{X} , ed è chiamata misura di probabilità baseline. Interpretativamente, P traduce l'attuale valutazione probabilistica disponibile per il vettore aleatorio \mathbf{X} . La tripletta (\mathbf{X}, g, P) viene chiamata modello baseline. Con riferimento ad una qualsiasi distribuzione alternativa $Q \in \mathcal{P}$, chiamiamo la tripletta (\mathbf{X}, g, Q) modello alternativo e Q misura di probabilità alternativa. Denotiamo infine con riferimento ad una fissata $Q^{\zeta} \in \mathcal{P}$, Q^{ζ} assolutamente continua rispetto a P , con $\zeta = \frac{dQ^{\zeta}}{dP}$ la derivata di

Radon-Nikodym (RN) di Q^ζ rispetto a P , che ricordiamo essere una variabile aleatoria non negativa tale che $E_P(\zeta) = 1$. Operativamente la derivata di RN consente di “distorcere” la probabilità baseline per passare ad una probabilità aggiornata, per dettagli si veda l'Appendice A.2.

Il punto di partenza della reverse sensitivity analysis è la definizione di uno stress sulla distribuzione dell'output Y che immaginiamo sia problematico per il decisore. Chiamiamo quindi la tripletta (\mathbf{X}, g, Q) modello stressato, con $Q \in \mathcal{P}$ misura di probabilità stressata, il modello per il quale, sotto a Q , la variabile aleatoria Y rispetta un insieme di vincoli probabilistici definiti dallo stress imposto e inoltre Q ha divergenza di Kullback-Leibler¹ minima rispetto a P . In altri termini, la misura di probabilità stressata è la soluzione del seguente problema di minimo:

$$\min_{Q \in \mathcal{P}} D_{KL}(Q||P), \text{ sotto ai vincoli definiti dallo stress.} \quad (2.2)$$

Si noti che in generale il passaggio dalla misura di probabilità baseline a quella stressata può indurre una nuova distribuzione e struttura di dipendenza tra i fattori di rischio in input. Analizzeremo nel seguito tre diversi risultati inquadrabili nella logica (2.2) nei quali lo stress imposto è di natura diversa; in tutti i casi si fornirà una soluzione esplicita al problema di ottimo attraverso l'espressione della derivata di RN. Una volta individuata, attraverso uno dei prossimi risultati, la misura stressata soluzione del problema di minimo, si può impiegare per condurre diverse tipologie di analisi. L'obiettivo è in generale comprendere quali fattori di rischio abbiano subito uno stress più evidente nel passaggio dalla misura di probabilità baseline a quella stressata. Procediamo esponendo i tre teoremi nelle successive tre sottosezioni. Il secondo e il terzo sono di fondamentale importanza nella trattazione e saranno estensivamente impiegati nei prossimi due capitoli.

¹La divergenza di Kullback-Leibler di P da Q rappresenta una misura della discrepanza tra le due misure di probabilità; per la definizione formale e proprietà strutturali si veda l'Appendice A.3

2.2.1 Vincolo espresso dalla misura stressata

Affrontiamo il Problema (2.2) nel caso in cui i vincoli definiti dallo stress siano espressi tramite la misura di probabilità stressata Q . In particolare consideriamo, in prima istanza, degli insiemi disgiunti $B_i \in \mathcal{B}$, $i = 1, \dots, I$. A partire da questi si possono individuare gli eventi $\{Y \in B_i\} = \{\omega \in \Omega : Y(\omega) \in B_i\}$, che rappresentano interpretativamente valori dell'output "fuori controllo"; ad esempio in ambito finanziario, se interpretiamo Y come perdita, gli insiemi B_i individuano zone di perdita estrema. Enunciamo ora formalmente il Teorema 1, che è un caso particolare del Teorema 3.1 di Csiszar (1975).

Teorema 1

Siano $B_1, \dots, B_I \in \mathcal{B}$ tali che $B_i \cap B_j = \emptyset$, per ogni $i \neq j$, e inoltre $P(Y \in B_i) > 0$, $i = 1, \dots, I$. Fissati i numeri reali $\alpha_1, \dots, \alpha_I \geq 0 : \sum_{i=1}^I \alpha_i \leq 1$, esiste un'unica soluzione al seguente problema di minimo:

$$\min_{Q \ll P} D_{KL}(Q \| P), \quad \text{sub : } Q(Y \in B_i) = \alpha_i \quad i = 1, \dots, I. \quad (2.3)$$

La derivata di RN della soluzione è siffatta:

$$\zeta = \sum_{i=0}^I \frac{\alpha_i}{P(Y \in B_i)} \mathbb{1}_{\{Y \in B_i\}}, \quad \text{ove } \alpha_0 = 1 - \sum_{i=1}^I \alpha_i \text{ e } B_0 = \left(\bigcup_{i=1}^I B_i \right)^c.$$

Si noti che la derivata di RN soluzione del problema di ottimo è una funzione costante a tratti di Y . Passando quindi dalla misura di probabilità baseline P a quella stressata Q , i valori di Y all'interno del medesimo insieme B_i ricevono la stessa "distorsione" probabilistica. In particolare, se $\alpha_i > P(Y \in B_i)$ e dunque $\zeta > 1$, i valori di Y in B_i , sotto la misura stressata Q soluzione del problema, subiscono un incremento probabilistico. I valori di Y che non appartengono all'insieme B_i subiscono un decremento. Procediamo con la dimostrazione.

Dimostrazione.

Osserviamo, in primo luogo, che ζ è una derivata di RN. Risulta infatti banalmente essere non negativa e inoltre realizza $E_P(\zeta) = 1$:

$$\begin{aligned} E_P(\zeta) &= \int_{\Omega} \zeta dP = \int_{\Omega} \sum_{i=0}^I \frac{\alpha_i}{P(Y \in B_i)} \mathbf{1}_{\{Y \in B_i\}} dP = \\ &= \sum_{i=0}^I \frac{\alpha_i}{P(Y \in B_i)} \int_{\Omega} \mathbf{1}_{\{Y \in B_i\}} dP \\ &= \sum_{i=0}^I \frac{\alpha_i}{P(Y \in B_i)} \int_{\{Y \in B_i\}} dP = \\ &= \sum_{i=0}^I \alpha_i = 1. \end{aligned}$$

La misura stressata Q , corrispondente alla soluzione del problema di minimo, rispetta inoltre il vincolo $Q^\zeta(Y \in B_i) = \alpha_i$, $i = 0, \dots, I$:

$$\begin{aligned} Q^\zeta(Y \in B_i) &= E_P(\zeta \mathbf{1}_{\{Y \in B_i\}}) = \int_{\{Y \in B_i\}} \sum_{j=0}^I \frac{\alpha_j}{P(Y \in B_j)} \mathbf{1}_{\{Y \in B_j\}} dP = \\ &= \sum_{j=0}^I \frac{\alpha_j}{P(Y \in B_j)} \int_{\{Y \in B_i\}} \mathbf{1}_{\{Y \in B_j\}} dP = \frac{\alpha_i}{P(Y \in B_i)} P(Y \in B_i) = \alpha_i. \end{aligned}$$

Consideriamo ora un'altra arbitraria derivata di RN ξ tale che $Q^\xi(Y \in B_i) = \alpha_i$ per $i = 0, \dots, I$. Risulta che (sfruttando la disuguaglianza di Jensen, ricordando che la funzione $x \log x$ è convessa):²

$$\begin{aligned} D_{KL}(Q^\xi \| P) &= E_P(\xi \log(\xi)) = \sum_{i=0}^I E_P(\xi \log(\xi) | Y \in B_i) P(Y \in B_i) \\ &\geq \sum_{i=0}^I E_P(\xi | Y \in B_i) \log(E_P(\xi | Y \in B_i)) P(Y \in B_i) \end{aligned}$$

²Si rammenti la definizione di speranza matematica di una variabile aleatoria X condizionata ad un evento A tale che $P(A) > 0$: $E_P(X|A) = \frac{E_P(X \mathbf{1}_A)}{P(A)}$.

$$\begin{aligned}
&= \sum_{i=0}^I \frac{\mathbb{E}_P(\xi \mathbf{1}_{\{Y \in B_i\}})}{P(Y \in B_i)} \log \left(\frac{\mathbb{E}_P(\xi \mathbf{1}_{\{Y \in B_i\}})}{P(Y \in B_i)} \right) P(Y \in B_i) \\
&= \sum_{i=0}^I Q^\zeta(Y \in B_i) \log \left(\frac{Q^\zeta(Y \in B_i)}{P(Y \in B_i)} \right) \\
&= \sum_{i=0}^I \alpha_i \log \left(\frac{\alpha_i}{P(Y \in B_i)} \right) = D_{KL}(Q^\zeta \| P).
\end{aligned}$$

Segue quindi la tesi, ricordando che la divergenza di Kullback-Leibler è strettamente convessa (Appendice A.3.1). ■

Si osservi che nella dimostrazione si è impiegata esclusivamente la convessità della funzione $x \log x$. Il risultato è quindi banalmente estendibile sostituendo alla divergenza di Kullback-Leibler una qualsiasi f-divergenza (vedi Appendice A.3).

2.2.2 Vincolo espresso in termini di VaR

Consideriamo il Problema (2.2) nel caso in cui i vincoli definiti dallo stress siano espressi in termini di Value at Risk. In generale questo caso è diverso da quello del Problema (2.3). Osserviamo però che, se la funzione di ripartizione di Y sotto la misura baseline P è crescente, allora un vincolo espresso in termini di Value at Risk del tipo $\text{VaR}_\alpha^Q(Y) = q$, individua univocamente due boreliani disgiunti e quindi il problema di ottimizzazione vincolata è un caso particolare del precedente. Enunciamo il Teorema 2.

Teorema 2

Siano $0 < \alpha < 1$ e $q \in \mathbb{R} : \text{VaR}_\alpha^P(Y) < q < \sup Y$. Consideriamo il seguente problema di minimo:

$$\min_{Q \ll P} D_{KL}(Q \| P), \quad \text{sub} : \text{VaR}_\alpha^Q(Y) = q. \quad (2.4)$$

Il problema ammette unica soluzione se e solo se $P(q - \varepsilon < Y < q) > 0$ per ogni $\varepsilon > 0$; la derivata di RN della soluzione è siffatta:

$$\zeta = \frac{\alpha}{P(Y < q)} \mathbb{1}_{\{Y < q\}} + \frac{1 - \alpha}{P(Y \geq q)} \mathbb{1}_{\{Y \geq q\}}. \quad (2.5)$$

La derivata di RN soluzione del Problema (2.4) assume esclusivamente due valori. Essa è in questo caso una funzione non decrescente di Y ; ciò significa che sotto la misura di probabilità stressata Q soluzione del problema, lo scenario ($Y \geq q$) riceve un incremento probabilistico ($\zeta > 1$). Si osservi inoltre che q non può in generale essere scelto arbitrariamente in quanto deve rispettare la condizione $P(q - \varepsilon < Y < q) > 0$, per ogni $\varepsilon > 0$. Se questa condizione non è verificata, come potrebbe accadere nel caso in cui Y fosse discreta, il problema rimane significativo qualora si sostituisca il vincolo $VaR_\alpha^Q(Y) = q$ con $Q(Y < q) = \alpha$, utilizzando il risultato del Teorema 1. Dimostriamo il Teorema 2.

Dimostrazione.

Sia $q : P(q - \varepsilon < Y < q) > 0$, per ogni $\varepsilon > 0$; verifichiamo in primo luogo che ζ è una derivata di RN per la quale risulta: $VaR_\alpha^{Q^\zeta}(Y) = q$. Banalmente è non negativa e inoltre si ha:

$$\begin{aligned} E_P(\zeta) &= \int_\Omega \zeta dP = \int_\Omega \frac{\alpha}{P(Y < q)} \mathbb{1}_{\{Y < q\}} + \frac{1 - \alpha}{P(Y \geq q)} \mathbb{1}_{\{Y \geq q\}} dP \\ &= \frac{\alpha}{P(Y < q)} \int_\Omega \mathbb{1}_{\{Y < q\}} dP + \frac{1 - \alpha}{P(Y \geq q)} \int_\Omega \mathbb{1}_{\{Y \geq q\}} dP \\ &= \alpha + 1 - \alpha = 1. \end{aligned}$$

La derivata di RN ζ soddisfa:

$$q = VaR_\alpha^{Q^\zeta}(Y) = \inf\{x \in \mathbb{R} : Q^\zeta(Y \leq x) \geq \alpha\}.$$

Infatti, in primo luogo:

$$Q^\zeta(Y \leq q) = E_P(\zeta \mathbb{1}_{\{Y \leq q\}}) = \int_{\{Y \leq q\}} \left(\frac{\alpha}{P(Y < q)} \mathbb{1}_{\{Y < q\}} + \frac{1 - \alpha}{P(Y \geq q)} \mathbb{1}_{\{Y \geq q\}} \right) dP$$

$$\geq \int_{\{Y < q\}} \frac{\alpha}{P(Y < q)} dP = \alpha.$$

Inoltre assunto arbitrariamente $\hat{q} < q$, si ha:

$$\begin{aligned} Q^\zeta(Y \leq \hat{q}) &= E_P(\zeta \mathbf{1}_{\{Y \leq \hat{q}\}}) = \int_{\{Y \leq \hat{q}\}} \frac{\alpha}{P(Y < q)} \mathbf{1}_{\{Y < q\}} + \frac{1 - \alpha}{P(Y \geq q)} \mathbf{1}_{\{Y \geq q\}} dP \\ &= \frac{\alpha}{P(Y < q)} \int_{\{Y \leq \hat{q}\}} \mathbf{1}_{\{Y < q\}} dP = \frac{\alpha}{P(Y < q)} P(Y \leq \hat{q}) < \alpha, \end{aligned}$$

ove l'ultima disequazione è una conseguenza dell'ipotesi $P(q - \varepsilon < Y < q) > 0$, per ogni $\varepsilon > 0$. Quindi q è il più piccolo elemento per cui vale la relazione $Q^\zeta(Y \leq x) \geq \alpha$.

Consideriamo ora un'altra derivata di RN $\xi = \frac{dQ^\xi}{dP}$ per la quale vale $VaR_\alpha^{Q^\xi}(Y) = q$. La divergenza di Kullback-Leibler di Q^ξ rispetto a P è: (si utilizza la disuguaglianza di Jensen):

$$\begin{aligned} D_{KL}(Q^\xi \| P) &= E_P(\xi \log(\xi)) \\ &= E_P(\xi \log(\xi) | Y < q) P(Y < q) + E_P(\xi \log(\xi) | Y \geq q) P(Y \geq q) \\ &\geq E_P(\xi | Y < q) \log(E_P(\xi | Y < q)) P(Y < q) + E_P(\xi | Y \geq q) \log(E_P(\xi | Y \geq q)) P(Y \geq q) \\ &= \frac{E_P(\xi \mathbf{1}_{\{Y < q\}})}{P(Y < q)} \log\left(\frac{E_P(\xi \mathbf{1}_{\{Y < q\}})}{P(Y < q)}\right) P(Y < q) + \\ &+ \frac{E_P(\xi \mathbf{1}_{\{Y \geq q\}})}{P(Y \geq q)} \log\left(\frac{E_P(\xi \mathbf{1}_{\{Y \geq q\}})}{P(Y \geq q)}\right) P(Y \geq q) \\ &= Q^\xi(Y < q) \log\left(\frac{Q^\xi(Y < q)}{P(Y < q)}\right) + Q^\xi(Y \geq q) \log\left(\frac{Q^\xi(Y \geq q)}{P(Y \geq q)}\right) \\ &= k(Q^\xi(Y < q), P(Y < q)), \end{aligned}$$

con

$$k(x, y) = x \log\left(\frac{x}{y}\right) + (1 - x) \log\left(\frac{1 - x}{1 - y}\right), \quad x, y \in (0, 1).$$

Osserviamo che si ha:

$$\frac{\partial k(x, y)}{\partial x} = \log\left(\frac{x}{y}\right) - \log\left(\frac{1 - x}{1 - y}\right) < 0 \Leftrightarrow x < y \quad (2.6)$$

e quindi la funzione $x \rightarrow k(x, y)$, fissato y , è non crescente sull'intervallo $(0, y)$. Sussiste inoltre la seguente catena di disuguaglianze:

$$Q^\xi(Y < q) \leq \alpha \leq P(Y \leq VaR_\alpha^P(Y)) \leq P(Y < q). \quad (2.7)$$

Per provare la prima disuguaglianza di (2.7) procediamo per assurdo. Sia $Q^\xi(Y < q) > \alpha$, vale allora la seguente relazione:

$$\lim_{r \rightarrow q^-} Q^\xi(Y \leq r) = Q^\xi(Y < q) > \alpha,$$

Tale relazione implica l'esistenza di $r' < q : Q^\xi(Y \leq r') \geq \alpha$ che viola la definizione di q . La seconda disuguaglianza di (2.7) discende dalla definizione di $VaR_\alpha^P(Y)$ (Appendice A.1.1), infine la terza è una banale conseguenza dell'ipotesi $VaR_\alpha^P(Y) < q$.

Pertanto, in luce di (2.7), per la divergenza di Kullback-Leibler di Q^ξ da P si ha:

$$\begin{aligned} D_{KL}(Q^\xi \| P) &\geq k(Q^\xi(Y < q), P(Y < q)) \geq k(\alpha, P(Y < q)) \\ &= \alpha \log \left(\frac{\alpha}{P(Y < q)} \right) + (1 - \alpha) \log \left(\frac{1 - \alpha}{P(Y \geq q)} \right) \\ &= D_{KL}(Q^\xi \| P). \end{aligned}$$

L'unicità è una conseguenza della stretta convessità della divergenza di Kullback-Leibler (Appendice A.3.1).

Per completare la dimostrazione, mostriamo che se la condizione $P(q - \varepsilon < Y < q) > 0$, per ogni $\varepsilon > 0$ non è verificata allora il problema non ammette soluzione. Esista dunque $\varepsilon > 0 : P(q - \varepsilon < Y < q) = 0$. Se $P(Y = q) = 0$ allora, per l'assoluta continuità di Q rispetto a P , si ha $Q(q - \varepsilon < Y \leq q) = 0$ e quindi il problema non ammette soluzione in quanto il vincolo $VaR_\alpha^Q(Y) = q$ non può essere rispettato. Sia dunque $P(Y = q) > 0$ e sia ξ una derivata di RN per la quale $VaR_\alpha^{Q^\xi}(Y) = q$. Indicati con $r = Q^\xi(Y \leq q)$ e con $p = P(Y \leq q)$, consideriamo la divergenza di Kullback-Leibler di Q^ξ rispetto

a P , ragionando come in precedenza si trova:

$$\begin{aligned} D_{KL}(Q^\xi \| P) &= E_P(\xi \log(\xi) | Y \leq q)p + E_P(\xi \log(\xi) | Y > q)(1-p) \\ &\geq r \log\left(\frac{r}{p}\right) + (1-r) \log\left(\frac{1-r}{1-p}\right) = D_{KL}(Q^{\xi^r} \| P), \end{aligned}$$

ove si è posto $\xi^u = \frac{dQ^{\xi^u}}{dP} = \frac{u}{p} \mathbb{1}_{\{Y \leq q\}} + \frac{1-u}{1-p} \mathbb{1}_{\{Y > q\}}$, con $0 \leq u \leq 1$.

La famiglia di derivate di RN ξ^u soddisfa $VaR_\alpha^{Q^{\xi^u}}(Y) = q$ se e solo se si ha:

$$\alpha \leq u < \alpha \frac{p}{P(Y < q)}. \quad (2.8)$$

In quanto, per rispettare $VaR_\alpha^{Q^{\xi^u}}(Y) = q$, in primo luogo deve risultare:

$$Q^{\xi^u}(Y \leq q) \geq \alpha \Leftrightarrow E_P(\xi^u \mathbb{1}_{\{Y \leq q\}}) \geq \alpha \Leftrightarrow u \geq \alpha,$$

e inoltre, assunto arbitrariamente $q' < q$:

$$\begin{aligned} Q^{\xi^u}(Y \leq q') < \alpha &\Leftrightarrow E_P(\xi^u \mathbb{1}_{\{Y \leq q'\}}) < \alpha \\ &\Leftrightarrow \frac{u}{p} P(Y \leq q') < \alpha \Leftrightarrow u < \alpha \frac{p}{P(Y \leq q')}, \end{aligned}$$

da cui, osservato che $P(q - \varepsilon < Y < q) = 0$, dall'arbitrarietà di q' discende la relazione (2.8). Alla luce di (2.8) il problema di ottimo vincolato (2.4) si riduce a minimizzare $D_{KL}(Q^{\xi^u} \| P)$ sotto la condizione (2.8). Come funzione di u la divergenza di Kullback-Leibler di Q^{ξ^u} da P è una funzione non crescente sull'intervallo $(0, p]$ (vedi (2.6)); pertanto questo problema di ottimo vincolato non ha soluzione essendo che $\alpha \frac{p}{P(Y < q)} < p$. Concludendo, il problema di ottimo vincolato (2.4) ammette soluzione se e solo se la condizione $P(q - \varepsilon < Y < q) > 0$, per ogni $\varepsilon > 0$, è verificata. ■

Si osservi che, anche in questo caso, nella dimostrazione si è impiegata esclusivamente la convessità della funzione $x \log x$, il risultato è quindi banalmente estendibile sostituendo alla divergenza di Kullback-Leibler una qualsiasi f-divergenza (vedi Appendice A.3).

2.2.3 Vincolo espresso in termini di VaR e ES

Affrontiamo in questa sezione il Problema (2.2) con vincoli espressi in termini di Value at Risk e Expected Shortfall simultaneamente in modo da stressare l'intera coda della distribuzione della variabile aleatoria output Y . Enunciamo quindi il Teorema 3, per la dimostrazione si veda Pesenti et al. (2019).

Teorema 3

Siano $0 < \alpha < 1$ e $q, s \in \mathbb{R}$ per i quali si abbia: $VaR_\alpha^P(Y) < q < s < \sup Y$. Ipotizziamo che la funzione generatrice dei momenti della variabile aleatoria $Y|Y > q$ esista in un intorno di 0 e che risulti $E_P(Y|Y > q) < s$. Si consideri il seguente problema di ottimo:

$$\min_{Q \ll P} D_{KL}(Q||P), \quad \text{sub} : VaR_\alpha^Q(Y) = q, \quad ES_\alpha^Q(Y) = s. \quad (2.9)$$

Definiamo gli insiemi $A_1 = \{Y \geq q\}$ e $A_2 = \{Y > q\}$ e denotiamo con θ_i^* , $i = 1, 2$, l'unica soluzione positiva della seguente equazione:

$$E_P((Y - s)e^{\theta(Y-q)}|A_i) = 0. \quad (2.10)$$

Esiste allora, in ognuno dei seguenti casi, soluzione unica al problema di minimo vincolato:

1. $P(q - \varepsilon < Y < q) > 0$, per ogni $\varepsilon > 0$ e inoltre $E_P(e^{\theta_1^*(Y-q)}|A_1) \leq \frac{P(A_1^c)/P(A_1)}{\alpha/(1-\alpha)}$.
2. $P(Y = q) > 0$, $P(q - \varepsilon < Y < q) = 0$ per qualche $\varepsilon > 0$ e infine $E_P(e^{\theta_2^*(Y-q)}|A_2) \geq \frac{P(A_2^c)/P(A_2)}{\alpha/(1-\alpha)}$.

La derivata di RN della soluzione è siffatta:

$$\zeta_i = \frac{\alpha}{P(A_i^c)} \mathbf{1}_{A_i^c} + \frac{1 - \alpha}{E_P(e^{\theta_i^*(Y-q)} \mathbf{1}_{A_i})} e^{\theta_i^*(Y-q)} \mathbf{1}_{A_i}, \quad i = 1, 2. \quad (2.11)$$

Si noti in primo luogo che, confrontando il Teorema 3 con il Teorema 2 dove lo stress è imposto esclusivamente a livello di Value at Risk, aggiungendo

l'Expected Shortfall il problema ha soluzione anche nel caso di distribuzioni discrete. Si osservi che la condizione sulla funzione generatrice dei momenti implica l'impossibilità di scegliere indipendentemente q e s . In questo caso la derivata di RN ha un andamento esponenziale per valori di Y che eccedono q , diversamente dal Teorema 2, dove era costante. Gli eventi che interessano la coda destra della distribuzione di Y subiscono, sotto la misura stressata, un incremento probabilistico tanto più forte quanto più sono estremi.

2.3 Algoritmo di simulazione Monte Carlo

Nei tre Teoremi analizzati si è sempre ricavata un'espressione esplicita per la derivata di RN soluzione del problema di ottimo vincolato. Ciò consente l'implementazione del cambio di misura (da P a Q) nell'ambito della simulazione Monte Carlo. La derivata di RN è in generale una funzione di Y nel senso che: $\zeta(\omega) = \eta(Y(\omega))$, $\omega \in \Omega$ per qualche funzione $\eta : \mathbb{R} \rightarrow \mathbb{R}$ (ad esempio nel caso del Teorema 2, η è una funzione non decrescente di Y). Operativamente, per implementare la reverse sensitivity analysis in ambito simulativo, si può quindi seguire una logica di questo tipo:

1. Effettuare M simulazioni (scenari) $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(M)}$ di \mathbf{X} sotto la misura di probabilità iniziale P .
2. Calcolare $y^{(k)} = g(\mathbf{x}^{(k)})$, $k = 1 \dots, M$.
3. Disponendo quindi di $y^{(1)}, \dots, y^{(M)}$ e di η (che è in generale diversa nei 3 problemi) si può costruire la derivata di RN tramite: $\zeta^{(k)} = \eta(y^{(k)})$, $k = 1, \dots, M$.
4. Ricordando che per la definizione della derivata di RN si ha $Q(A) = E_P(\zeta \mathbf{1}_A)$, per effettuare valutazioni sotto alla probabilità stressata Q in ambito simulativo si può ricorrere a:

$$F_Y^Q(y) = Q(Y \leq y) = \frac{1}{M} \sum_{k=1}^M \zeta^{(k)} \mathbf{1}_{\{y^{(k)} \leq y\}}, \quad y \in \mathbb{R} \quad (2.12)$$

$$F_{X_i}^Q(x) = Q(X_i \leq x) = \frac{1}{M} \sum_{k=1}^M \zeta^{(k)} \mathbb{1}_{\{x_i^{(k)} \leq x\}}, \quad x \in \mathbb{R}, \quad i = 1, \dots, n \quad (2.13)$$

5. Tramite lo studio delle distribuzioni stressate del punto 4 si possono ricavare importanti informazioni sul ruolo dei fattori di rischio in input.

Il numero totale di simulazioni da effettuare sotto alla probabilità baseline P è pari a $M(n+1)$, numero che potrebbe diventare molto alto qualora in presenza di un elevato numero di fattori di rischio si utilizzino molti scenari. Si osservi infine che l'approccio proposto consente di effettuare le simulazioni sotto alla probabilità stressata Q senza dover risimulare gli scenari, ciò potrebbe essere di fondamentale importanza qualora le simulazioni siano computazionalmente esigenti.

2.4 Misurare la sensitività

Una volta ottenuta la distribuzione della variabile aleatoria output Y sotto la misura stressata Q , soluzione di uno dei problemi di ottimo vincolato introdotti in precedenza o in un altro modo, si possono indagare molti aspetti del modello per comprendere l'effetto dello stress. Molto spesso, oltre ad analizzare l'impatto dello stress sulla distribuzione della variabile output, è interessante studiare il comportamento dei fattori di rischio. Con riferimento a questi, un primo aspetto da analizzare è legato alle singole distribuzioni marginali. Questa analisi può essere effettuata informalmente tramite un confronto grafico tra le funzioni di ripartizione o quelle di densità (o probabilità) sotto le due misure oppure confrontando numericamente i momenti (o altri indici). Un'altra interessante indagine che può essere eseguita sui fattori di rischio concerne la loro struttura di dipendenza. Si rammenti infatti che, tramite il passaggio da P a Q , essa viene in generale alterata. Per studiare questo aspetto in modo semplice, si può ad esempio calcolare la matrice di correlazione, sotto le due misure di probabilità, per dettagli si veda l'Appendice A.4.

Qualora si fosse interessati, come spesso accade nella pratica, a valutare quale dei fattori di rischio sia più influente per il modello stocastico (ovvero quale è più impattato dal passaggio da P a Q o equivalentemente quale è più sensibile allo stress), nel caso in cui il numero di fattori di rischio non sia particolarmente elevato si può ricorrere ad un'analisi di tipo grafico che può fornire importanti informazioni preliminari. Se si fosse però interessati a misurare formalmente l'impatto che lo stress ha comportato sulla distribuzione dei fattori di rischio (o se il numero di fattori di rischio è troppo elevato per effettuare proficuamente un'analisi di tipo grafico) è necessario introdurre un'opportuna misura di sensitività.

Consideriamo uno spazio di probabilità (Ω, \mathcal{A}, P) . Data una variabile aleatoria V definita su Ω , denotiamo con $V_{|W}, V_{|W^\dagger}$ le variabili aleatorie che soddisfano $V_{|W} \stackrel{d}{=} V_{|W^\dagger} \stackrel{d}{=} V$ e tali che la coppia $(V_{|W}, W)$ sia comonotona e che la coppia $(V_{|W^\dagger}, W)$ sia contromonotona³. Segue banalmente dalla relazione (iii) del Teorema A.4.3 sul rapporto tra comonotonia, contromonotonia e covarianza che, per ogni variabile aleatoria V' tale che $V' \stackrel{d}{=} V$, si ha:

$$E_P(WV_{|W^\dagger}) \leq E_P(WV') \leq E_P(WV_{|W}). \quad (2.14)$$

Definiamo ora una misura di sensitività legata al particolare framework della reverse sensitivity analysis (Pesenti et al. (2019)).

Definizione

Sia Q^ξ una misura di probabilità alternativa a P con associata derivata di RN: $\xi = \frac{dQ^\xi}{dP}$. La sensitività di una variabile aleatoria Z al cambio di misura è definita come:

$$S(Z, \xi) = \begin{cases} \frac{E_P(Z\xi) - E_P(Z)}{\max_{\psi \stackrel{d}{=} \xi} E_P(Z\psi) - E_P(Z)}, & \text{se } E_P(Z\xi) \geq E_P(Z) \\ -\frac{E_P(Z\xi) - E_P(Z)}{\min_{\psi \stackrel{d}{=} \xi} E_P(Z\psi) - E_P(Z)}, & \text{se } E_P(Z\xi) < E_P(Z). \end{cases}$$

³Per dettagli sulle nozioni di comonotonia e contromonotonia si veda l'Appendice A.4.

ove valgono le convenzioni $\pm\frac{\infty}{\infty} = \pm 1$ e $\frac{0}{0} = 0$.

In entrambi i casi il numeratore riflette la differenza tra il valore atteso della distribuzione prima e dopo lo stress (si rammenti che $E_P(Z\xi) = E_{Q^\xi}(Z)$). Si osservi inoltre che, per la relazione (2.14), risulta:

$$\arg \max_{\psi \stackrel{d}{=} \xi} E_P(Z\xi) = \xi_{|Z}, \quad \arg \min_{\psi \stackrel{d}{=} \xi} E_P(Z\xi) = \xi_{|Z^\dagger}.$$

Pertanto, il denominatore standardizza la quantità a numeratore assicurando che $\mathcal{S} \in [-1, 1]$. Ciò consente di approcciarsi alla valutazione di \mathcal{S} in un contesto di simulazione Monte Carlo in modo semplice; è infatti sufficiente, per determinare rispettivamente $\xi_{|Z}$ e $\xi_{|Z^\dagger}$, riarrangiare il campione di ξ in modo che esso sia rispettivamente ordinato allo stesso modo o al contrario di quello di Z . Con riferimento specifico al modello di reverse sensitivity (2.2) introduciamo quindi due misure di sensitività (Pesenti et al. (2019)).

Definizione

Sia Q^ζ una misura di probabilità alternativa con $\zeta = \frac{Q^\zeta}{dP} = \eta(Y)$, per una funzione non decrescente η . A fronte dei fattori di rischio in input X_i , $i = 1, \dots, n$ e dell'output Y definiamo le seguenti due misure di rischio:

$$\Gamma_i = \mathcal{S}(X_i, \zeta), \quad \Delta_i = \mathcal{S}(Y, \zeta_{|X_i}).$$

La quantità Γ_i è detta reverse sensitivity measure. Specularmente la quantità Δ_i è detta forward sensitivity measure.

In questo caso ζ è da interpretarsi come derivata di RN soluzione di uno dei problemi di ottimo vincolato enunciati in precedenza. Per ogni fattore di rischio in input X_i la relativa reverse sensitivity measure Γ_i riflette quanto la media di questo fattore di rischio è impattata dal cambio di misura (in termini relativi). Si osservi che Γ_i quantifica l'impatto che uno stress effettuato sulla variabile aleatoria in output Y ha sul fattore di rischio X_i e non viceversa. Potrebbe accadere un particolare fenomeno (molto raro) per il quale la relazione inversa non è realizzata; si parla allora di dissonanza

probabilistica (Cooke & van Noortwijk, 1999). Questa osservazione giustifica l'introduzione, a fianco della misura Γ_i di Δ_i . La definizione della forward sensitivity measure è simile alla reverse ma l'attenzione è posta sulla variazione del valore atteso della variabile aleatoria in output Y . Si considera quindi, l'impatto che il massimo cambiamento di misura nel fattore di rischio in input X_i ha sull'output Y . L'analisi simultanea delle due misure permette di rilevare il fenomeno di dissonanza probabilistica. Si osservi che le due misure di sensitività introdotte possono essere equivalentemente espresse in termini di covarianza. Ad esempio per la reverse sensitivity measure, si ha:

$$\Gamma_i = \mathcal{S}(X_i, \zeta) = \begin{cases} \frac{\text{Cov}(X_i, \zeta)}{\max_{\psi \stackrel{d}{=} \zeta} \text{Cov}(X_i, \psi)}, & \text{se } E_P(X_i \zeta) \geq E_P(X_i), \\ -\frac{\text{Cov}(X_i, \zeta)}{\min_{\psi \stackrel{d}{=} \zeta} \text{Cov}(X_i, \psi)}, & \text{se } E_P(X_i \zeta) < E_P(X_i). \end{cases}$$

Valori di Γ_i positivi indicano dunque che, sotto la misura di probabilità stressata, il valore atteso del fattore di rischio X_i è maggiore rispetto a quello sotto la misura baseline. Viceversa per valori di Γ_i negativi. Enunciamo ora alcune proprietà di cui godono le due misure di rischio (per la dimostrazione si veda Pesenti et al. (2019)).

Proprietà 2.4.1

1. $\Gamma_i, \Delta_i \in [-1, 1]$.
2. $\Gamma_i = \Delta_i = 0$ se X_i e Y sono stocasticamente indipendenti.
3. $\Gamma_i = \Delta_i = 1$ se (X_i, Y) è comonotona.
4. $\Gamma_i = \Delta_i = -1$ se (X_i, Y) è contromonotona.
5. $\Gamma_i, \Delta_i \geq 0$ se (X_i, Y) sono positively quadrant dependent (PQD) ⁴.
6. $\Gamma_i, \Delta_i \leq 0$ se (X_i, Y) sono negatively quadrant dependent (NQD).

⁴Vedi Appendice A.4.

Pertanto le due misure di sensitività sono sempre comprese nell'intervallo $[-1, 1]$. La comonotonia e la contromonotonia rappresentano le relazioni di dipendenza stocastica più estreme e sono rispettivamente individuate dai valori 1 e -1 degli indici. Se le due variabili aleatorie X_i e Y sono stocasticamente indipendenti allora le misure di sensitività sono nulle. Si osservi che le relazioni di PQD e NQD implicano rispettivamente correlazione (sia lineare che sui ranghi) positiva e negativa. Alla luce di queste osservazioni, le misure di sensitività Γ_i e Δ_i , possono essere intese anche come misure di dipendenza tra variabili aleatorie.

Le misure di sensitività introdotte permettono di quantificare la differenza in termini di valore atteso tra la distribuzione alternativa e quella baseline, qualora si fosse interessati in una qualche altra proprietà della distribuzione, come ad esempio alla coda, le definizioni di Γ_i e Δ_i possono essere estese andando a considerare particolari trasformazioni monotone dei fattori di rischio. Selezionata quindi un'opportuna funzione u non decrescente si possono considerare le quantità: $\mathcal{S}(u(X_i), \zeta)$ e $\mathcal{S}(u(Y), \zeta_{|X_i})$; dal momento che la coppia $(u(X_i), X_i)$ è comonotona, le interpretazioni rimangono inalterate.

Una particolare trasformazione u che si può considerare è la seguente:

$$u_v(X_i) = (X_i - F_{X_i}^{-1}(v))_+ - (F_{X_i}^{-1}(1-v) - X_i)_+, \quad 0.5 \leq v < 1$$

Se $v = 0.5$, allora u_v coincide con la funzione identica e quindi $\mathcal{S}(u_{0.5}(X_i), \zeta) = \Gamma_i$ e $\mathcal{S}(u_{0.5}(Y), \zeta_{|X_i}) = \Delta_i$. Se $v > 0.5$, la funzione u_v è pari a 0 se $X_i \in [F^{-1}(1-v), F^{-1}(v)]$ e linearmente decrescente altrimenti. Pertanto, all'aumentare di v , si enfatizza il peso delle code.

Capitolo 3

Reserve sensitivity analysis: un portafoglio di rendite vitalizie

In questo capitolo si applicherà il framework introdotto e analizzato nel capitolo precedente ad un modello stocastico per la descrizione di un portafoglio di rendite vitalizie. Per modellare il fondo il portafoglio si seguirà l'approccio proposto da Olivieri e Pitacco (2003). Le rendite vitalizie rappresentano un classico prodotto delle assicurazioni vita e dei fondi pensione. Tipicamente, si tratta di un prodotto per il quale, a fronte di un premio iniziale pagato dall'assicurato, l'assicuratore si impegna a garantire un'entrata periodica per tutta la sua durata di vita residua; al decesso nulla è dato agli eredi. La rendita è basata sul meccanismo di mutualità, in base al quale le riserve matematiche rilasciate dagli assicurati deceduti sono ripartite per far fronte agli impegni nei confronti di coloro che sono ancora in vita. La possibilità di riscattare il prodotto non è in generale concessa per evitare il fenomeno dell'antiselezione. Convertire quindi un importo monetario in rendita vitalizia è una decisione irreversibile, per questa ragione la rendita è considerata un asset di tipo illiquido. Al fine di renderla meno rigida e più appetibile sul mercato, molto spesso è possibile acquistare diverse strutture di garanzie e opzioni aggiuntive.

Analizziamo ora, assumendo il punto di vista dell'assicuratore, i diversi tipi di rischio a cui è soggetto questo prodotto. Per maggiori dettagli si veda

Pitacco et al. (2009).

1. Rischio di mortalità

Il rischio di mortalità è legato alla durata aleatoria di vita degli assicurati in portafoglio e può essere scomposto in due tipologie distinte di rischio. La prima è dovuta alla fluttuazione del numero aleatorio di decessi attorno al proprio valore atteso causata dalla variabilità intrinseca della grandezza aleatoria. Tale rischio è da considerarsi un rischio di processo e può essere diversificato incrementando la grandezza del portafoglio di rendite. La seconda tipologia è dovuta ad una deviazione sistematica dal numero atteso; è quindi il rischio causato da un'errata stima del trend di mortalità futuro. Si parla in particolare di rischio di longevità qualora ci si riferisca al trend di medio-lungo periodo. Questo tipo di rischio è particolarmente rilevante nell'ambito di prodotti con garanzie che coprono lunghe durate temporali come nel caso delle rendite vitalizie. Il rischio di longevità è di natura sistematica, non può pertanto esser diversificato incrementando la grandezza del portafoglio.

2. Rischio di investimento

Il rischio di investimento è dovuto all'incertezza dei mercati finanziari nei quali l'assicuratore investe. Può essere scomposto in rischio di mercato (o tasso), rischio di credito e altri. Il rischio di mercato in questo caso si esplica nella possibilità che, a fronte di oscillazioni del prezzo dei titoli scambiati sul mercato, il valore di mercato degli attivi a copertura delle riserve matematiche sia insufficiente a far fronte agli impegni assunti. Il rischio di credito si origina invece dalla possibilità che l'istituto emittente dei titoli, che l'assicuratore ha acquistato a copertura del suo impegno futuro, sia insolvente. Potrebbero essere considerati molte altri rischi legati ai mercati finanziari, per una discussione più approfondita si veda Olivieri e Pitacco (2003).

Per far fronte alle varie tipologie di rischio descritte, l'assicuratore deve quindi disporre di efficaci strumenti di valutazione che gli consentano di valutare adeguatamente i suoi impegni futuri e monitorarli nel tempo. A tale fine, immaginiamo che egli sia interessato alla costruzione di un modello interno

(parziale) in quanto sia consapevole che, per una corretta gestione del rischio di mortalità e di quello finanziario, sia necessario un approccio stocastico. Il suo obiettivo è, in generale, utilizzare il modello per determinare opportune decisioni strategiche individuando pertanto le azioni di risk management da intraprendere. In questo ambito, definita una struttura probabilistica realistica per descrivere il valore aleatorio delle passività, la solvibilità dell'assicuratore è intesa come la sua capacità di fronteggiare la propria posizione debitoria aleatoria con un elevato grado di probabilità. Da un punto di vista operativo è necessario scegliere quali risultati di portafoglio considerare esattamente e che orizzonte temporale fissare.

Definiamo la probabilità di rovina per l'assicuratore. Siano al tempo t : Z_t il valore del fondo di portafoglio e \mathcal{Y}_t il valore dell'impegno aleatorio residuo dell'assicuratore. Fissata una probabilità ε (il cui complemento a 1 rappresenta il grado di solvibilità che si vuole imporre), diremo che l'assicuratore è solvibile al tempo t , a livello $1 - \varepsilon$, se e solo se:

$$P(Z_t - \mathcal{Y}_t \geq 0) \geq 1 - \varepsilon.$$

In particolare, in seguito, considereremo la solvibilità dell'assicuratore in $t = n$, ove n indica l'epoca tale che $\mathcal{Y}_n = 0$, ovvero l'impegno residuo dell'assicuratore è nullo (non ci sono più assicurati nel portafoglio). In questo caso, si ha pertanto:

$$P(Z_n \geq 0) \geq 1 - \varepsilon. \tag{3.1}$$

Per una discussione approfondita su questa definizione si fa nuovamente riferimento a Olivieri e Pitacco (2003). Tramite la condizione (3.1) possiamo determinare, utilizzando una procedura di simulazione, l'ammontare di capitale (margine iniziale), definito precisamente nel seguito, che l'assicuratore deve allocare al portafoglio per essere solvibile ad un assegnato livello di probabilità.

3.1 Struttura generale del modello stocastico

Considereremo nel seguito quattro diversi modelli stocastici, innestati l'uno dentro l'altro, per descrivere il portafoglio di rendite vitalizie dell'assicurato. A partire da un (semplice) modello base si procede un passo alla volta aggiungendo o modificando opportune ipotesi o caratteristiche per rendere lo studio sempre più realistico ed interessante. Delineiamo in questa sezione le caratteristiche generali comuni ai modelli del capitolo. Successivamente, introducendoli, discuteremo le peculiarità delle specifiche ipotesi caso per caso.

Prima di procedere descrivendo la struttura probabilistica per la mortalità degli assicurati, introduciamo la classica notazione impiegata nella matematica attuariale per le rendite vitalizie. A partire da una base biometrica (tavola di sopravvivenza o modello parametrico) si possono ricavare le probabilità di sopravvivenza per gli assicurati. Indichiamo, con riferimento ad un assicurato in vita ad età x , con p_x e ${}_h p_x$ le probabilità che esso sia in vita rispettivamente all'età $x+1$ e $x+h$. Assumiamo, come spesso si fa nella matematica attuariale, un tasso di interesse i , riferito ad un periodo temporale di un anno, costante e deterministico. Con riferimento ad un assicurato in vita all'età x , la quantità ${}_h E_x = (1+i)^{-h} {}_h p_x$ rappresenta il valore attuale atteso (valore attuariale) di un'unità monetaria pagata nel caso in cui l'assicurato sia ancora in vita all'età $x+h$. Una rendita vitalizia consiste, come si è detto, in una sequenza di importi pagati ad ogni epoca contrattualmente stabilita, qualora l'assicurato sia in vita in quell'epoca. Il suo valore attuariale è quindi dato da (nel caso sia ad esempio anticipata e di rata unitaria):

$$\ddot{a}_x = \sum_{h=0}^{\omega-x-1} {}_h E_x,$$

ove con ω si indica la più piccola età intera non raggiungibile in vita (età estrema).

Consideriamo ora un portafoglio di rendite vitalizie immediate di rata co-

stante R contenente N_0 assicurati di età iniziale x . Con riferimento all'epoca t , indichiamo con N_t , $t = 1, \dots, \omega - x - 1$ il numero aleatorio di assicurati in vita in quell'epoca a fronte degli N_0 assicurati in vita all'epoca iniziale e con \tilde{i}_t , $t = 1, \dots, \omega - x$, la variabile aleatoria che rappresenta il tasso di interesse (semplice) che descrive globalmente le performance finanziarie, sul periodo temporale $(t - 1, t)$, degli attivi che l'assicuratore ha posto a copertura delle riserve. Nell'istante iniziale l'assicuratore incassa, a fronte degli N_0 contratti, tutti i premi unici e un'istante dopo inizia a pagare le rendite anticipate. L'equazione ricorsiva che fornisce il valore del fondo di portafoglio Z_t nel tempo è la seguente:

$$Z_t = (Z_{t-1} - N_{t-1}R)(1 + \tilde{i}_t), \quad t = 1, \dots, \omega - x. \quad (3.2)$$

Il valore iniziale del fondo è: $Z_0 = N_0 \Pi + M$, ove con M indichiamo il margine iniziale che l'assicuratore alloca al fondo per limitare la probabilità di rovina al livello ε e $\Pi = R \ddot{a}_x$ è, per il principio di equità, il premio unico pagato dal generico assicurato all'epoca iniziale.

Per la descrizione della mortalità degli assicurati del portafoglio si fa in generale riferimento ad un modello che tiene simultaneamente in considerazione sia l'età dell'individuo che l'anno di calendario. La collettività degli assicurati è considerata essere omogenea, dunque in un fissato anno di calendario hanno tutti la medesima età e sono pertanto soggetti alla stessa probabilità di decesso. Impiegheremo pertanto nel seguito la notazione compatta $q_{y,t}$ per indicare la probabilità di decesso per un individuo di età y nell'anno di calendario $\hat{t} + t$, ove \hat{t} rappresenta un anno di calendario di riferimento. Indicato con $D_t = N_t - N_{t+1}$ il numero aleatorio di decessi tra le epoche t e $t + 1$, $t = 0, \dots, \omega - x - 1$ ($N_{\omega-x} = 0$), consideriamo la seguente struttura probabilistica:

$$D_t | N_t, \tilde{q}_{x+t,t} \sim \text{Binom}(N_t, \tilde{q}_{x+t,t}), \quad t = 0, \dots, \omega - x - 1, \quad (3.3)$$

ove $\tilde{q}_{x+t,t}$ è la variabile aleatoria che rappresenta la probabilità di decesso entro un anno per un individuo di età $x + t$ nell'anno di calendario $\hat{t} + t$. Si noti la simultanea presenza del rischio di processo e quello sistematico. Il primo è individuato dall'ipotesi binomiale mentre il secondo dall'aleatorietà delle probabilità di decesso. Si osservi inoltre che anche l'aleatorietà del tasso d'interesse è da considerarsi un rischio sistematico. L'espressione (3.3) sarà impiegata in ogni modello per la simulazione della mortalità del portafoglio. In taluni modelli, la variabile aleatoria $\tilde{q}_{x+t,t}$ sarà da considerarsi degenerare, individuando pertanto, al variare di x , un'unica tavola di sopravvivenza. Similmente considereremo in alcuni modelli \tilde{i}_t degenerare e costante nel tempo.

Descriviamo ora come il framework per l'analisi di sensitività, introdotto nel capitolo precedente, è adattato al modello stocastico che descrivere l'evolversi nel tempo del fondo di un portafoglio di rendite. Come variabile aleatoria di interesse (output del modello) consideriamo il valore finale del fondo $Y = Z_n$. In alcuni casi, per motivazioni interpretative, considereremo al posto di Z_n il suo opposto; questa scelta sarà dettagliatamente spiegata nel seguito. Si noti che la variabile aleatoria Z_n ha un ruolo fondamentale in quanto la probabilità di rovina è definita a partire da questa quantità. Si ha quindi:

$$Y = Z_n = g(\mathbf{X}),$$

per una funzione g individuata dall'equazione ricorsiva (3.2) che descrive l'andamento nel tempo del fondo di portafoglio. Il vettore aleatorio \mathbf{X} includerà di volta in volta le variabili aleatorie di interesse per il problema, cioè i fattori di rischio che descrivono il fenomeno della mortalità e l'andamento dei tassi di interesse. Siamo quindi interessati a comprendere come uno stress espresso in termini di misura di rischio sul valore finale aleatorio del fondo impatti sulla distribuzione dei fattori di rischio del modello che sono sostanzialmente individuati dal fenomeno di mortalità e dai rendimenti che l'assicuratore conseguente sugli attivi posti a copertura delle riserve matematiche. Le particolari condizioni di stress considerate sono sempre discusse all'interno dei singoli modelli.

3.2 Il modello stocastico per la mortalità

Descriviamo in questa sezione il modello stocastico impiegato per la descrizione della mortalità degli assicurati. Sebbene sarà impiegato pienamente esclusivamente nel Modello 4, gli scenari per la descrizione della mortalità degli assicurati utilizzati negli altri modelli sono individuati, come vedremo, a partire da esso. Consideriamo dati relativi alla popolazione italiana (maschile) relativi all'intervallo temporale 1950 - 2014 (fonte: Human Mortality Database). Per la valutazione probabilistica del fenomeno della mortalità della popolazione impieghiamo il modello di Cairns, Blake e Dowd quadratico con effetto di coorte (M7) con ipotesi binomiale (per dettagli su questa scelta, e per un'analisi comparativa con altri modelli su dati italiani, si veda l'Appendice B). Il modello ha pertanto la seguente struttura (si rimanda nuovamente all'Appendice per definizioni e dettagli sulle quantità presenti)¹:

$$\begin{aligned}
 (1). \quad & D_{x,t} \sim \text{Binom}(E_{x,t}^0, \tilde{q}_{x,t}), \\
 (2). \quad & \text{logit}(\tilde{q}_{x,t}) = \eta_{x,t}, \\
 (3). \quad & \eta_{x,t} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + ((x - \bar{x}) - \tilde{\sigma}_x^2)\kappa_t^{(3)} + \gamma_{t-x}.
 \end{aligned} \tag{3.4}$$

Analizziamo la struttura del previsore $\eta_{x,t}$ in (3.4). A differenza di alcuni altri modelli, gli elementi del previsore che dipendono dall'età x sono funzioni fissate. L'effetto temporale è invece descritto da tre componenti $\kappa_t^{(1)}$, $\kappa_t^{(2)}$ e $\kappa_t^{(3)}$ che sono rispettivamente moltiplicate per un termine costante, uno lineare e uno quadratico dell'età; è presente infine un ultimo termine γ_{t-x} per descrivere l'effetto di coorte. Sebbene i dati disponibili comprendano l'intervallo di età 0 – 110, consideriamo, per stimare il modello, l'intervallo 60 – 89. Per età maggiori di 89 l'esperienza empirica evidenzia che i dati sono in generale poco affidabili soprattutto per epoche passate. Molto spesso in letteratura (Villegas et al. (2018), Currie (2014)) si preferisce dunque non impiegare dati riguardanti età estreme (maggiori di 90). Stimiamo il modello M7 utilizzando dati concernenti la popolazione italiana riferiti all'intervallo

¹Si osservi che, per semplicità di notazione, in questa sezione indichiamo l'anno di calendario con t .

temporale 1950 – 2014 e esaminiamone le componenti tramite dei grafici, si veda a tal fine la Figura 3.1. Nella colonna a sinistra si possono visualizzare gli effetti prefissati legati all'età. Nella colonna di destra ci sono i rispettivi valori stimati degli effetti temporali $\kappa_t^{(i)}$, $i = 1, 2, 3$ e dell'effetto coorte γ_{t-x} . Per epoche precedenti al 2015, le curve rappresentano i valori dei parametri

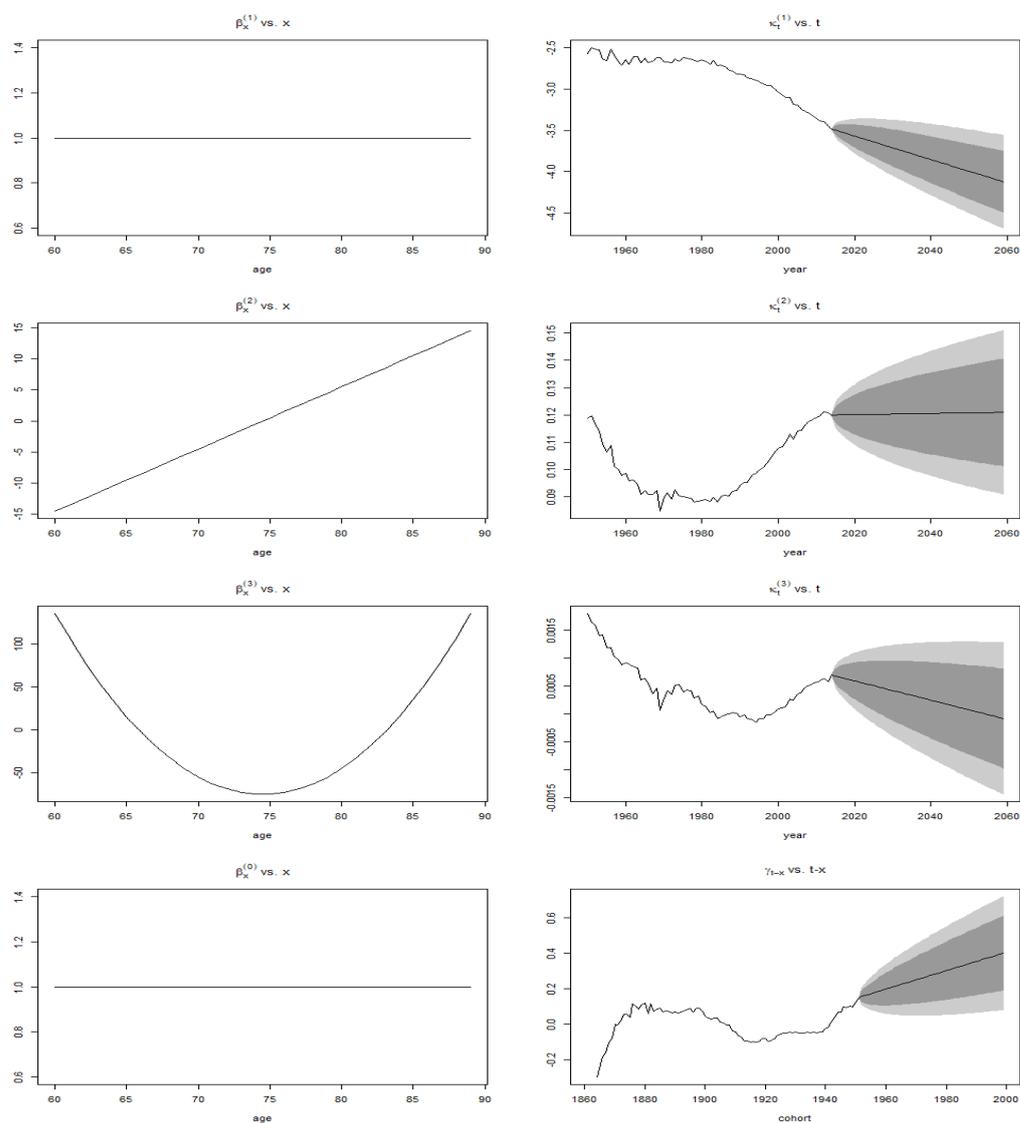


Figura 3.1: Componenti del modello M7: sulla colonna di sinistra $\beta_x^{(i)}$ che nel modello sono prefissati. Sulla colonna di destra i rispettivi effetti temporali $\kappa_t^{(i)}$ e l'effetto coorte γ_{t-x} .

stimati dal modello. A partire dal 2015 si può apprezzare il comportamento del modello in previsione. Le zone ombreggiate rappresentano gli intervalli di previsione rispettivamente al 80% (scura) e al 95% (chiara). La linea interna alla zona scura è la previsione centrale e fornisce indicazioni sul trend del fenomeno aleatorio. Si rammenti che, per le componenti legate all'effetto temporale, la previsione è operata tramite un processo stocastico di passeggiata casuale multivariata mentre per l'effetto di coorte è impiegato un modello ARIMA(1,1,0), per dettagli si veda l'Appendice B.

Sebbene sia interessante studiare le singole componenti del previsore (3.4) del modello, non è semplice, a partire da queste, comprendere intuitivamente il comportamento complessivo delle probabilità di morte vista la complessa struttura sottostante. Presentiamo dunque un altro grafico (Figura 3.2) che rappresenta il comportamento su tutto l'arco temporale 1950 – 2059 di $\tilde{q}_{65,t}$ (sinistra) e $\tilde{q}_{75,t}$ (destra). Per epoche precedenti al 2015 la linea nera rappresenta i valori stimati dal modello mentre la linea blu i valori empirici (ottenuti rapportando i decessi per quell'età e anno di origine sul numero iniziale di esposti al rischio). Come si vede in entrambi i grafici, l'accostamento

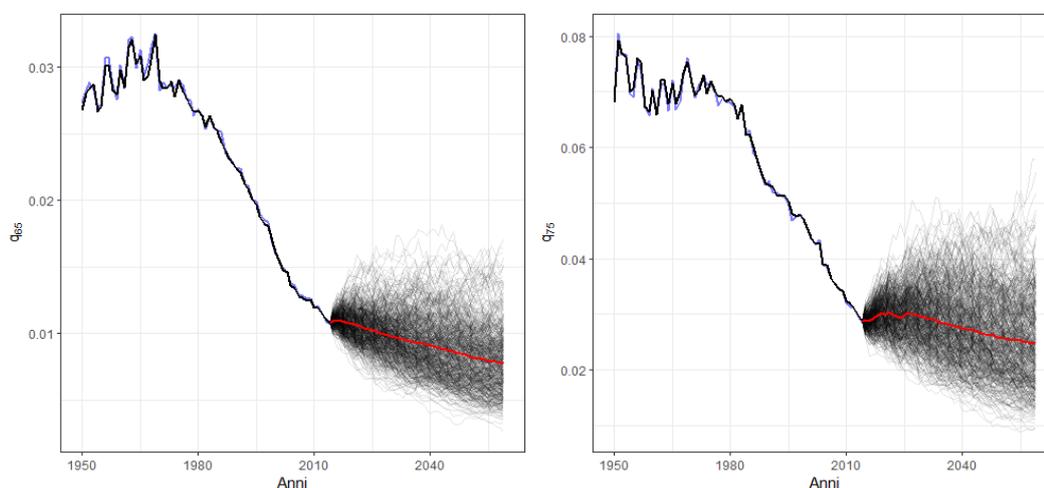


Figura 3.2: A sinistra: andamento per $t = 1950 - 2059$ di $\tilde{q}_{65,t}$. Per epoche precedenti al 2015 la linea nera rappresenta i valori stimati dal modello mentre quella blu i valori empirici (grezzi); per epoche successive al 2014 sono rappresentate in nero le prime 500 simulazioni ottenute dal modello M7 e in rosso la media di queste. A destra: medesima descrizione del grafico a sinistra per $\tilde{q}_{75,t}$.

tra i valori stimati e quelli empirici è molto buono. Per epoche successive al 2015 sono rappresentate in entrambi i casi le prime 500 traiettorie simulate (di un totale di 500 000) per comprendere il comportamento del modello in previsione. La linea rossa rappresenta la loro media. Come si può osservare il trend, come è naturale aspettarsi, è decrescente.

Studiamo ora come impiegare il modello M7 per descrivere la particolare mortalità degli assicurati. Successivamente, nei quattro modelli che introdurremo, si considererà sempre una collettività omogenea di assicurati aventi all'epoca iniziale, anno di riferimento $\hat{t} = 2015$, nella quale acquistano la rendita vitalizia immediata, tutti età x pari a 65. Si rammenti che il modello, come si è detto, considera l'intervallo di età 60 – 89. Per ottenere le probabilità di morte per età maggiori di 89 impiegheremo una semplice interpolazione lineare considerando 110 come età estrema. A partire dalla collettività presente all'epoca iniziale, per effettuare le simulazioni, si otterranno, diverse storie di mortalità impiegando la relazione (3.3). Operativamente effettuare una singola simulazione dal modello M7 comporta ottenere determinazioni dalla seguente matrice di variabili aleatorie:

$\tilde{Q}_{65,2015}$	$\tilde{Q}_{65,2016}$	$\tilde{Q}_{65,2017}$...	$\tilde{Q}_{65,2058}$	$\tilde{Q}_{65,2059}$
$\tilde{Q}_{66,2015}$	$\tilde{Q}_{66,2016}$	$\tilde{Q}_{66,2017}$...	$\tilde{Q}_{66,2058}$	$\tilde{Q}_{66,2059}$
$\tilde{Q}_{67,2015}$	$\tilde{Q}_{67,2016}$	$\tilde{Q}_{67,2017}$...	$\tilde{Q}_{67,2058}$	$\tilde{Q}_{67,2059}$
...
$\tilde{Q}_{108,2015}$	$\tilde{Q}_{108,2016}$	$\tilde{Q}_{108,2017}$...	$\tilde{Q}_{108,2058}$	$\tilde{Q}_{108,2059}$
$\tilde{Q}_{109,2015}$	$\tilde{Q}_{109,2016}$	$\tilde{Q}_{109,2017}$...	$\tilde{Q}_{109,2058}$	$\tilde{Q}_{109,2059}$

Tabella 3.1: Tabella contenente le variabili aleatorie che descrivono la mortalità per individui di età $x = 65, \dots, 109$ e anno di calendario $t = 2015, \dots, 2059$. La diagonale individua le variabili aleatorie d'interesse per descrivere la mortalità della collettività omogenea di assicurati di età 65 nel 2015.

Nel nostro caso particolare, nel quale siamo interessati alla sola mortalità della coorte di individui che hanno compiuto 65 anni nel 2015, le uniche variabili aleatorie d'interesse per il problema sono quelle della diagonale della matrice (in verde). Le determinazioni simulate di queste variabili aleatorie sono implementate attraverso la relazione (3.3) per ottenere la mortalità degli assicurati. Come si menzionava in precedenza, impiegheremo in modo

completo la struttura appena descritta esclusivamente nel Modello 4. Negli altri casi si considereranno opportuni scenari di mortalità derivanti dalle simulazioni di M7. In particolare nel Modello 1 considereremo un singolo scenario per la descrizione della mortalità individuato a partire dalla media empirica dei valori simulati dalla diagonale della tabella. Nei Modelli 2 e 3 si considereranno tre distinti scenari: lo scenario centrale è quello impiegato nel Modello 1, e a questo si aggiungono uno scenario con probabilità di decesso aggravate e uno con probabilità di decesso ridotte, individuati da particolari quantili empirici dei valori simulati dalla diagonale della tabella. In Figura 3.3 sono rappresentate le prime 1000 simulazioni dalla diagonale della Tabella 3.1 che rappresentano ognuna una diversa proiezione delle probabilità di decesso per la collettività di assicurati. La traiettoria verde individua la mortalità media, e sarà impiegata nel Modello 1 singolarmente e nei Modelli 2 e 3 come scenario medio. Le traiettorie rossa e blu rappresentano rispettivamente il quantile 0.9 e 0.1 e saranno impiegate per rappresentare i due scenari alternativi. Indichiamo le probabilità di morte implicate dalle tre traiettorie rispettivamente con: $q_x^{(med)}$, $q_x^{(max)}$ e $q_x^{(min)}$.

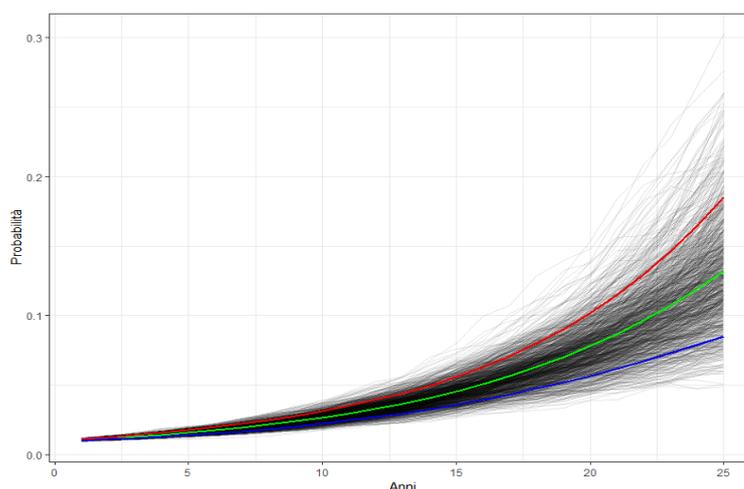


Figura 3.3: Le prime 1000 simulazioni per la mortalità degli assicurati della collettività. In verde la traiettoria individuata dalla media empirica ($q_x^{(med)}$) dell'intero set di 500 000 simulazioni, in rosso il quantile empirico al livello 0.9 ($q_x^{(max)}$) e in blu il quantile empirico al livello 0.1 ($q_x^{(min)}$).

3.2.1 Ipotesi numeriche comuni e calcolo del premio

Descriviamo brevemente in questa sottosezione le ipotesi numeriche comuni ai vari modelli e il metodo di calcolo del premio equo. Per ogni modello considerato nel seguito si impiegano 500 000 simulazioni. Le particolari condizioni di stress e i fattori di rischio \mathbf{X} considerati sono peculiari ad ogni modello e dunque descritti caso per caso. Il marginale iniziale M è diverso in ogni modello ma calibrato sempre con $\varepsilon = 2.5\%$. Un aspetto fondamentale da tenere in considerazione concerne il metodo tramite il quale i modelli sono annidati uno dentro l'altro. L'obiettivo generale di questa operazione è consentire un confronto tra questi e dunque comprendere il ruolo dei fattori di rischio. In ognuno dei modelli che sono considerati nel seguito si fa riferimento alla omogenea collettività di $N_0 = 1000$ assicurati che all'epoca iniziale $t = 0$ hanno età $x = 65$ con riferimento all'anno di calendario $\hat{t} = 2015$ (hanno compiuto 65 anni nel 2015). Per quanto concerne il fenomeno di mortalità si sono descritte in precedenza le peculiarità dei vari modelli. Il tasso di interesse è considerato nei primi due modelli deterministico e costante: $\tilde{i}_t = i = 3\%$. Successivamente si introduce un modello stocastico per la descrizione del comportamento del tasso, che mantiene 3% come tasso medio. Alla luce di queste osservazioni, per la valutazione del premio equo, che è effettuata in ambito attuariale tenendo solo conto di valori attesi (si veda (3.1)), si può considerare sempre quella del Modello 1. Fissata la rata di rendita R pari a 100, il premio equo è $\Pi = 1\,444,5$. Nella Tabella 3.2 si propone uno schema logico che individua la natura delle variabili aleatorie considerate nei vari modelli.

	Mortalità	Tasso	Fattori di rischio
Modello 1	Deterministico - 1 scenario	Deterministico	N_t
Modello 2	Stocastico - 3 scenari	Deterministico	$N_t S$
Modello 3	Stocastico - 3 scenari	Stocastico	$N_t S \tilde{i}_t$
Modello 4	Pienamente stocastico	Stocastico	$N_t \tilde{q}_{x+t,t} \tilde{i}_t$

Tabella 3.2: Schema riassuntivo sulla natura delle variabili input dei modelli del capitolo. S è l'ente aleatorio che descrive i tre scenari di mortalità.

3.3 Modello 1: mortalità e tassi deterministici

Introduciamo in questa sezione il primo modello stocastico per la descrizione del portafoglio di rendite dell'assicuratore; come fattori di rischio in input consideriamo esclusivamente $\mathbf{X} = (N_1, \dots, N_{\omega-x-1})$, facendo riferimento al solo rischio di processo. I tassi di mortalità futuri e i tassi di interesse sono deterministici: l'incertezza del modello è quindi individuata solo dalla struttura binomiale in (3.3). In tale struttura le variabili aleatorie $\tilde{q}_{x+t,t}$ sono da considerarsi degeneri, si ha dunque $\tilde{q}_{x+t,t} = q_x^{(med)}$.

3.3.1 Analisi di sensitività

Affrontiamo in questa sezione l'analisi di sensitività per il Modello 1. Il margine iniziale di portafoglio, fissata la probabilità di rovina pari a 2.5%, è $M = 31\,000$. Consideriamo uno stress sul Value at Risk del valore del fondo finale di portafoglio Y (in linea con il Problema (2.4)). In particolare, sia esso siffatto:

$$\alpha = 0.9, \quad q = VaR_\alpha^Q(Y) = (1 + stress)VaR_\alpha^P(Y), \quad (3.5)$$

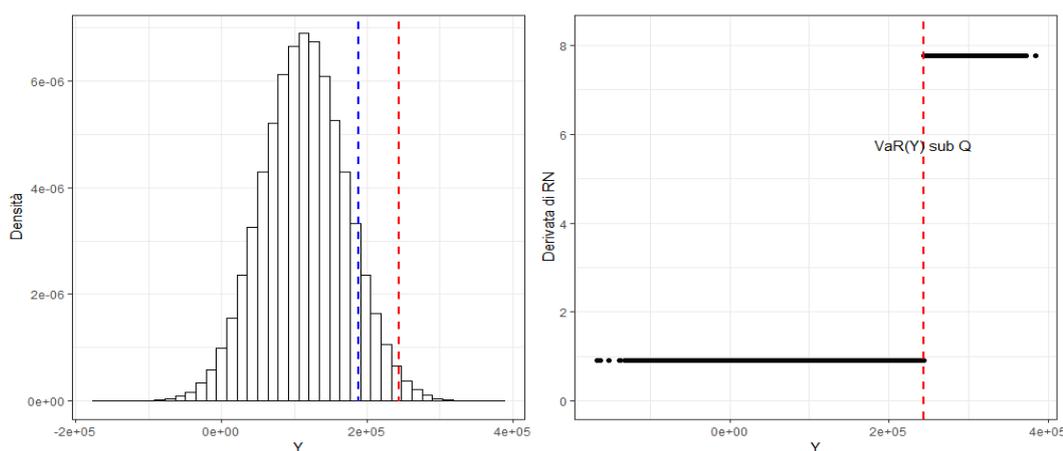


Figura 3.4: A sinistra: istogramma di Y ottenuto dai valori simulati del fondo finale. La linea blu coincide con il livello del $VaR_\alpha^P(Y)$, la linea rossa coincide invece $VaR_\alpha^Q(Y)$. A destra: derivata di RN ξ soluzione del Problema (2.4) sotto il vincolo espresso da (3.5).

dove poniamo $stress = 0.3$. Si osservi che, come menzionato nel Capitolo 2, qualora si impieghi il Teorema 2 in ambito discreto, è necessario sostituire il vincolo $VarR_\alpha^Q(Y) = q$ con $Q(Y \leq q) = \alpha$. In tutto il resto del capitolo non sorgeranno questo genere di problematiche in quanto si impiegherà sempre uno stress in linea con il Problema (2.9). Si può notare in Figura 3.4 (sinistra) che, come si è già detto, è stato allocato al fondo un capitale tale da rendere la probabilità di rovina (fondo finale negativo) pari a 2.5%. Si osservi che in questo caso, diversamente dall'interpretazione generale impiegata in Pesenti et al. (2019), lo stress che stiamo così imponendo rappresenta uno scenario “positivo” per l'assicuratore, ci aspettiamo dunque di osservare una maggiore mortalità nel portafoglio sotto la nuova probabilità; analizzeremo successivamente anche il caso opposto, più interessante per lo studio, ove lo stress rappresenta uno scenario avverso all'assicuratore. Nella Figura 3.4 (destra) si può osservare come la misura di probabilità stressata Q enfatizzi il peso su valori grandi del fondo (maggiore di q). Pertanto, sotto tale misura, l'assicuratore dispone di un fondo finale più consistente. Proseguiamo confrontando, tramite le funzioni di ripartizione simulate, le due distribuzioni del fondo finale prima e dopo lo stress. La funzione di ripartizione di Y , sotto la misura stressata Q soluzione del Problema (2.4) con vincolo dato

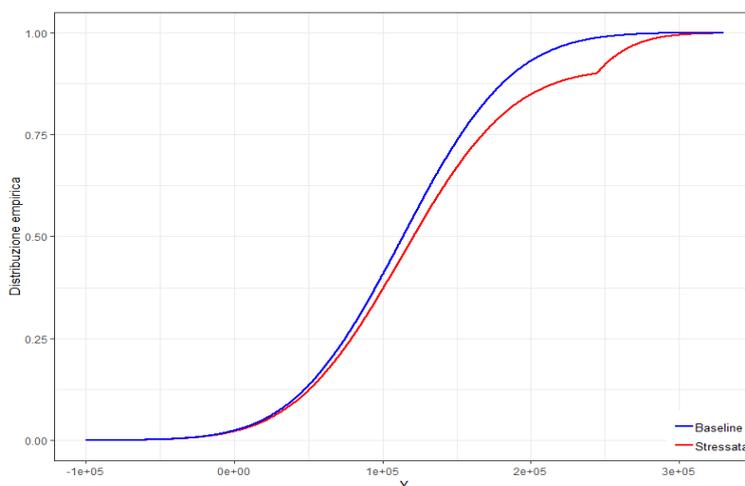


Figura 3.5: Confronto tra le funzioni di ripartizioni empiriche (simulate) sotto la probabilità baseline P (in blu) e stressata Q (in rosso).

dalla (3.5), è sempre minore a quella sub P (Figura 3.5), ad indicare che la prima domina stocasticamente al primo ordine la seconda.

Concentriamoci ora sul comportamento dei fattori di rischio. Ricordiamo che N_t è la variabile aleatoria che descrive il numero di assicurati del portafoglio in vita all'epoca t , a partire dai N_0 individui iniziali. Prima di analizzare come la distribuzione di N_t sia impattata dal passaggio dalla misura baseline a quella stressata, osserviamo graficamente come sono in generale strutturate le distribuzioni dei numeri aleatori N_t . Visualizziamo a tal fine, ad esempio, la funzione di ripartizione empirica e la funzione di probabilità di N_1 (Figura 3.6) sotto la misura baseline. I valori simulati di N_1 appartengono all'insieme di numeri interi compresi tra 971 e 1000. La distribuzione presenta una leggera asimmetria negativa, valori molto vicini a 1000 traducono una mortalità molto bassa e sono poco probabili come lo sono valori prossimi a 980 che traducono una forte mortalità. Analizziamo cosa succede alla distribuzione di N_1 conseguentemente allo stress (3.5), si veda a tal fine la Figura 3.7. Dalla funzione di probabilità (destra) appare chiaro che la distribuzione di N_1 sotto Q (rosso) presenta un peso maggiore sui valori minori di 988 e un peso minore su quelli superiori, a rappresentare che la mortalità sotto la misura stressata è più forte rispetto al caso baseline. Si preferirà nel seguito

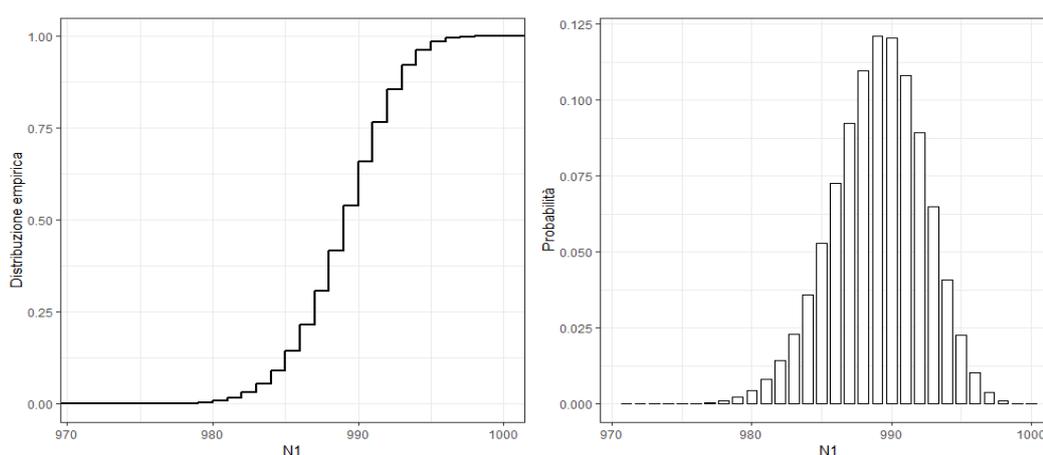


Figura 3.6: A sinistra: funzione di ripartizione empirica di N_1 . A destra: funzione di probabilità di N_1 , rappresentata tramite un grafico a barre.

impiegare esclusivamente la funzione di probabilità per il confronto in quanto è di più immediata interpretazione.

Procedendo, studiamo come si comporta lo stress al variare di t , confrontando graficamente la funzione di probabilità di N_t , $t = 1, 5, 10, 15, 20, 25, 30$ e 35 sotto le due misure di probabilità. L'obiettivo è in generale comprendere se, attraverso un'analisi di tipo grafico, si riesce a cogliere quali delle variabili aleatorie è più impattata dallo stress. Si noti che qualora il numero di valori assunti sia elevato (maggiore di 40), la rappresentazione tramite grafico a barre non è efficace per cogliere il comportamento di N_t , la sostituiamo pertanto con una curva continua, ombreggiando le aree dove la distribuzione sotto la misura baseline è maggiore di quella sotto la misura stressata in blu e il viceversa in rosso. Tutti i grafici della Figura 3.8 mostrano che dopo lo stress la distribuzione di N_t presenta sempre una maggiore mortalità (la coda sinistra ha sempre un peso maggiore sotto la misura stressata); osservando con attenzione si può notare il comportamento nel tempo di questo fenomeno. In particolare nei primi quattro grafici l'impatto è crescente, successivamente è decrescente. L'ultimo grafico, che mostra la distribuzione simulata del numero aleatorio di individui in vita a 100 anni, presenta uno stress sulla coda

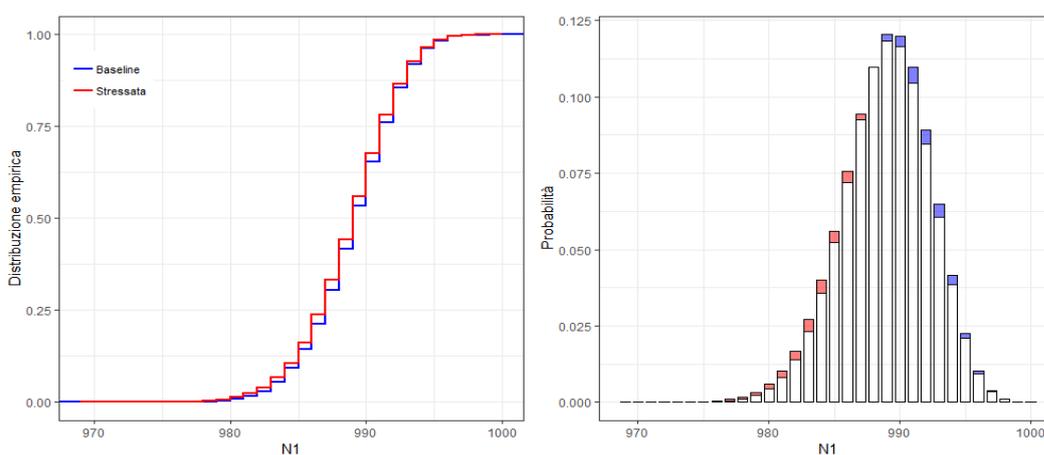


Figura 3.7: A sinistra: confronto tra funzione di ripartizione empirica di N_1 sub P (blu) e Q (rosso). A destra: confronto tra funzione di probabilità di N_1 , rappresentata tramite un grafico a barre, sub P (blu) e Q (rosso). I due grafici sono sovrapposti, pertanto qualora le punte delle barre sono di colore rosso la misura stressata supera la baseline, viceversa se di colore blu.

sinistra insignificante. In generale non è interessante analizzare il comportamento del modello per età maggiori di 100 in quanto pochi individui riescono ad arrivare in vita a quell'età. Riassumendo, appare evidente che lo stress (3.5) abbia impattato sulla distribuzione di N_t sempre nella stessa direzione e che esso sia crescente fino N_{15} e dopo decresca. Tramite analisi grafica, si è quindi delineato un possibile ordine di importanza dei fattori di rischio per questo primo modello. Vedremo nel seguito se le misure di sensitività, introdotte nella sezione 2.4, confermano o meno quanto finora analizzato.

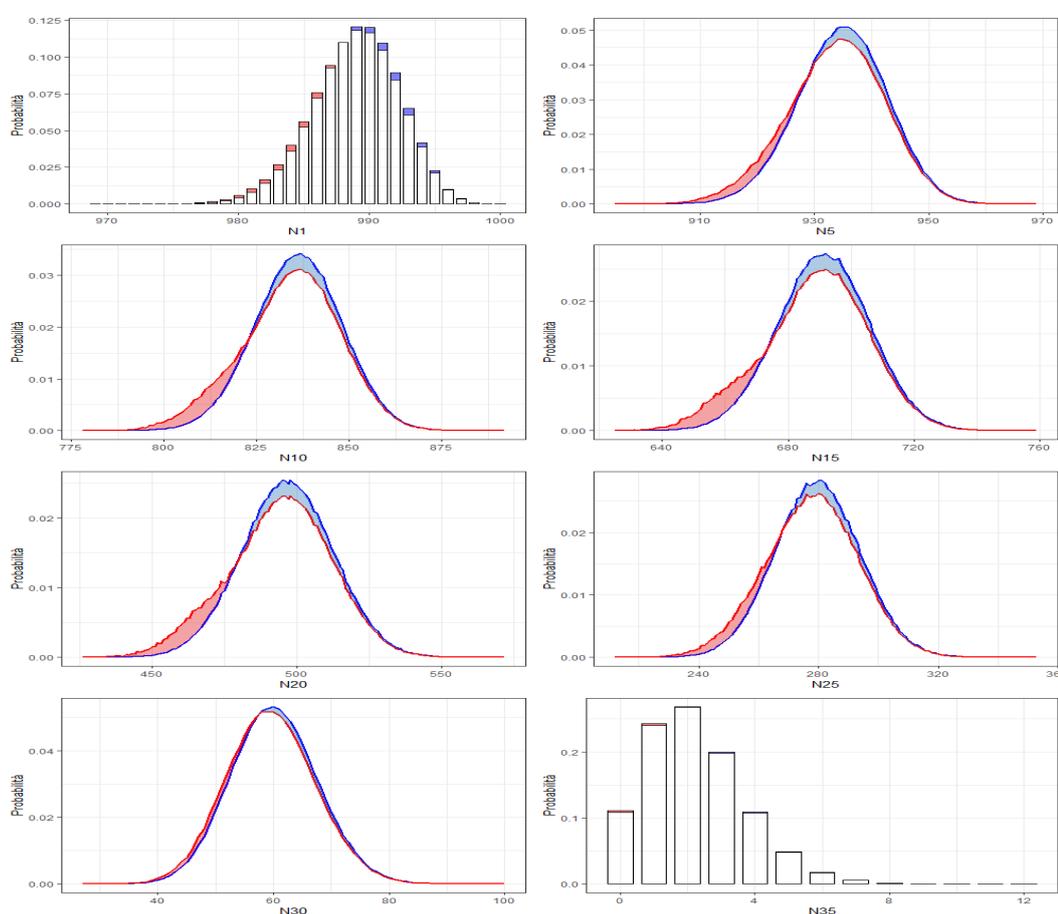


Figura 3.8: Confronto tra le funzioni di probabilità di N_t per $t = 1, 5, 10, 15, 20, 25, 30$ e 35 . La funzione di probabilità di N_t sotto al misura di probabilità baseline è in blu, quella sotto la misura stressata in rosso. Il colore delle zone ombreggiate mostra quale delle due distribuzioni è maggiore dell'altra per quel tratto di curva e servono a cogliere meglio l'entità dello stress.

Studiamo ora un altro aspetto, abbiamo fino a questo punto mantenuto la variabile *stress* costante e pari a 0.3. È interessante andare a studiare al variare dello stress come si comporta la distribuzione dei fattori di rischio in input N_t . Consideriamo due livelli di *stress* alternativi pari rispettivamente a: 0.1 e 0.5 (per il caso 0.3 si confronti con gli esempi precedenti). Controlliamo graficamente il comportamento di N_t per $t = 1, 15, 30$ in Figura 3.9. Si può osservare che, all'aumentare dello stress, l'impatto sulla distribuzione di N_t aumenta drasticamente; particolarmente rilevante è quello su N_{15} , ciò in generale non stupisce in quanto all'aumentare dello stress si richiede un fondo finale tale da aver associato un Value at Risk maggiore e ciò comporta maggior mortalità. Come già commentato in precedenza l'impatto dello stress ha un comportamento nel tempo prima crescente e poi decrescente. Procediamo calcolando le misure di sensitività introdotte nel capitolo precedente. Consideriamo nuovamente lo stress iniziale pari a 0.3. Confrontiamo

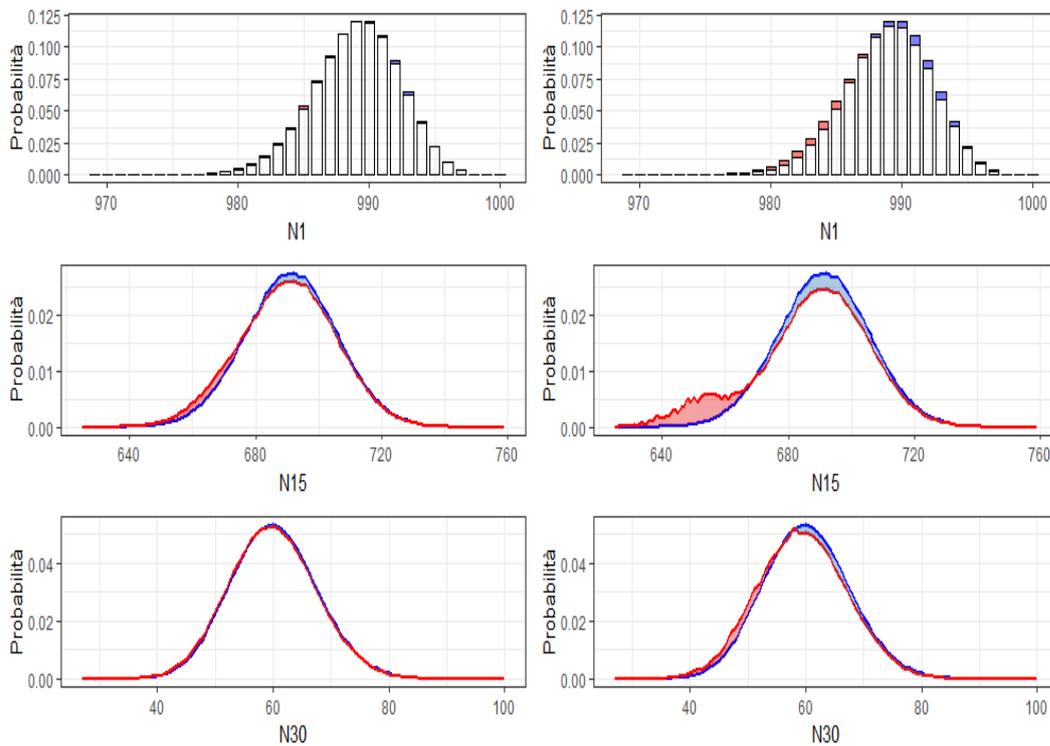


Figura 3.9: Confronto tra distribuzione di N_1 , N_{15} e N_{30} prima (in blu) e dopo (in rosso) lo stress. A sinistra la variabile *stress* è fissata pari a 0.1, a destra 0.5.

la reverse sensitivity measure e la forward sensitivity measure dei fattori di rischio del nostro modello, in particolare consideriamo N_t , $t = 1, \dots, 35$. Dato l'ampio numero di fattori di rischio si decide di impiegare un grafico a barre per visualizzare l'andamento delle due misure. Come si può osservare dalla Figura 3.10 a sinistra la misura di sensitività Γ_{N_t} conferma quanto già detto sul comportamento dello stress. I valori delle misure di sensitività sono negativi in quanto il valore atteso di N_t sotto la misura stressata è sempre minore di quello sotto la misura baseline. In particolare l'impatto del cambiamento di misura è crescente fino circa $t = 15$ e poi decresce. Il grafico della misura forward Δ_i , avendo il medesimo profilo dell'altro, non evidenzia problemi di dissonanza probabilistica. Anche nei prossimi modelli, non ci sarà mai evidenza di questo fenomeno e preferiremo pertanto non riportare il grafico della forward sensitivity measure.

In ultima istanza, studiamo il comportamento della dipendenza tra i fattori di rischio del modello, in particolare come lo stress impatti sulla sua struttura. Dalla definizione della distribuzione di N_t (si veda (3.2)) si comprende che la distribuzione di $N_{t+1}|N_t, N_{t-1}, \dots, N_1$ è uguale alla distribuzione di $N_{t+1}|N_t$ in quanto ognuno degli N_t dipende dagli altri solo tramite il precedente. Per comprendere come lo stress ha impattato sulla struttura di dipendenza di (N_1, \dots, N_{35}) studiamo allora il comportamento della

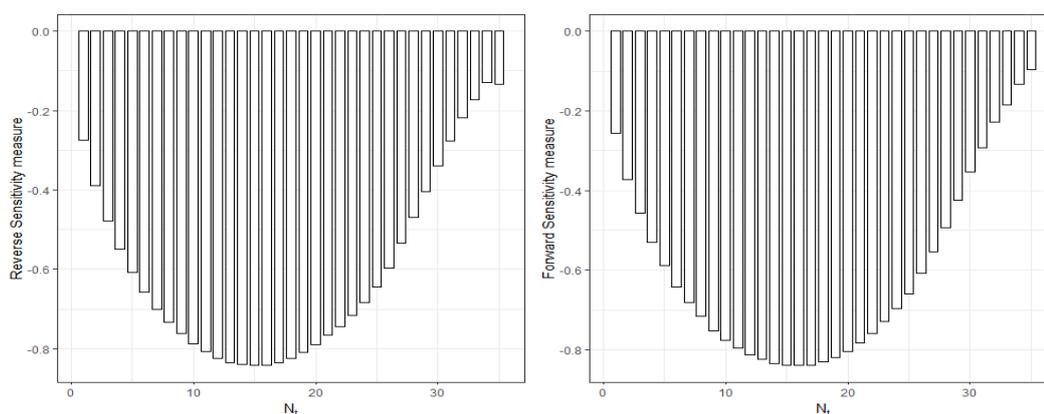


Figura 3.10: A sinistra: reverse sensitivity measure (Γ_{N_t}). A destra: forward sensitivity measure (Δ_{N_t})

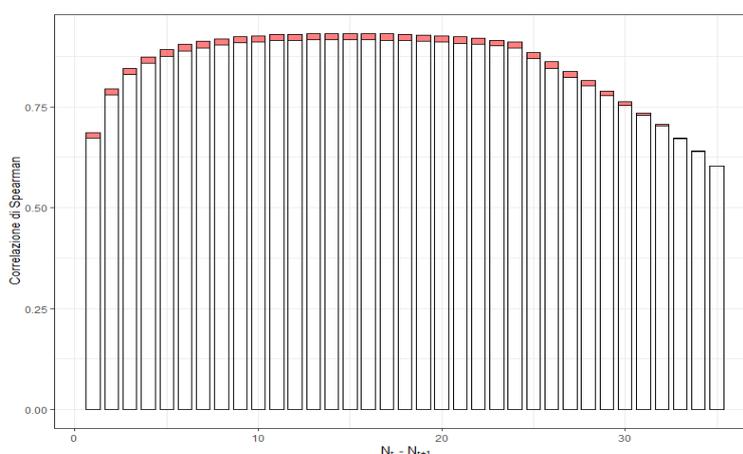


Figura 3.11: Confronto tra la correlazione di Spearman per le prime 36 coppie consecutive di N_t , in blu sub P , in rosso sub Q .

correlazione delle coppie consecutive prima e dopo lo stress. In particolare consideriamo la correlazione sui ranghi (di Spearman) e confrontiamo il comportamento di $Cor^S(N_t, N_{t+1})$ per $t = 1, \dots, 35$ tramite un grafico. Per una giustificazione di questo approccio si veda l'Appendice A.4.2. Come si può osservare dalla Figura 3.11 la correlazione dopo lo stress è aumentata per ogni coppia (N_t, N_{t+1}) , ed è massima in corrispondenza ai fattori di rischio più rappresentativi. Confrontando le Figure 3.10 e 3.11 si comprende che la struttura di dipendenza, seppur importante, non individua l'andamento della reverse sensitivity measure.

3.3.2 Modello 1 - Uno stress alternativo, VaR e ES

Nel modello appena descritto lo stress (3.5) è espresso esclusivamente in termini di Value at Risk. Una situazione di stress alternativa è proposta nel Teorema 3 dove vengono contemporaneamente stressati sia il Value at Risk che l'Expected Shortfall. Il vincolo che la probabilità stressata Q deve rispettare in questo caso interessa tutta la coda della distribuzione.

Consideriamo nuovamente il Modello 1 con le medesime ipotesi numeriche introdotte in precedenza; in linea con il Problema (2.9) introduciamo la

seguinte struttura per lo stress (fissato $\alpha = 0.9$):

$$\begin{aligned} q &= VaR_\alpha^Q(Y) = (1 + stress_{VaR}) VaR_\alpha^P(Y), \\ s &= ES_\alpha^Q(Y) = (1 + stress_{ES}) ES_\alpha^P(Y). \end{aligned} \tag{3.6}$$

Lo stress (3.6), dove vengono contemporaneamente stressati sia il Value at Risk che l'Expected Shortfall, è quindi un'alternativa al (3.5). Analizziamo, brevemente, come si comporta il Modello 1 con questa nuova struttura di stress. Fissiamo i valori $stress_{VaR} = 0.1$ e $stress_{ES} = 0.15$. La Figura 3.13 (sinistra) permette di apprezzare la diversa struttura dello stress rispetto la distribuzione di Y . Nella Figura (destra) si può analizzare come sono pesati, dalla misura di probabilità stressata, i valori di Y maggiori di q . Studiamo ora come la nuova struttura di stress impatta sulla distribuzione di Y tramite un confronto tra funzioni di ripartizione, si veda a tal fine la Figura 3.13. Valgono qualitativamente le medesime considerazioni fatte nel caso precedente con l'accortezza di notare che, in questo caso, dove lo stress colpisce tutta la coda della distribuzione, non si nota lo scalino presente nel Figura 3.5. Osserviamo, in conclusione di questo esempio, che confrontando i valori as-

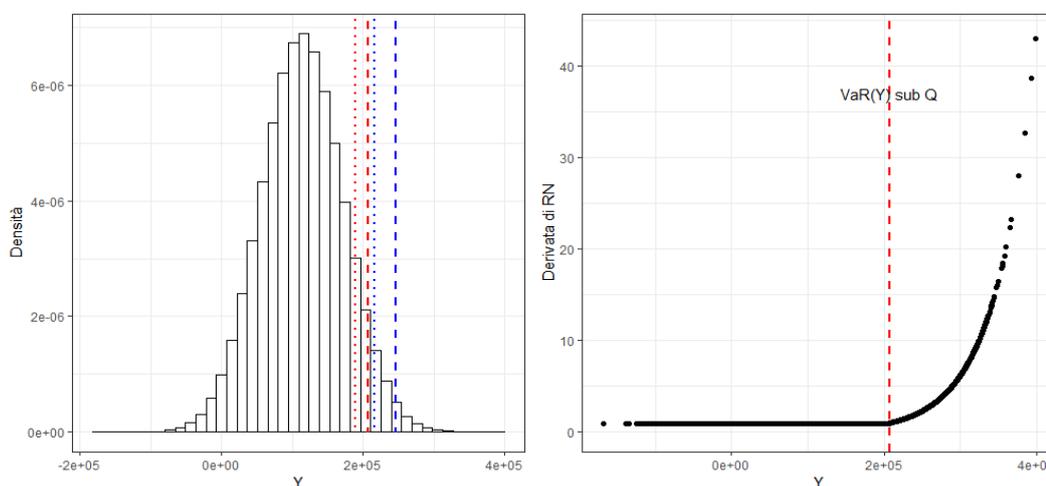


Figura 3.12: A sinistra: istogramma di Y ottenuto dai valori simulati del fondo finale. La linea rossa a punti coincide con il livello del $VaR_\alpha^P(Y)$ mentre quella a tratteggio con invece $q = VaR_\alpha^Q(Y)$. La linea blu a punti coincide con il livello del $ES_\alpha^P(Y)$ mentre quella a tratteggio con invece $s = ES_\alpha^Q(Y)$. A destra: derivata di RN ξ simulata soluzione del Problema (2.9) sotto al vincolo espresso da (3.6).

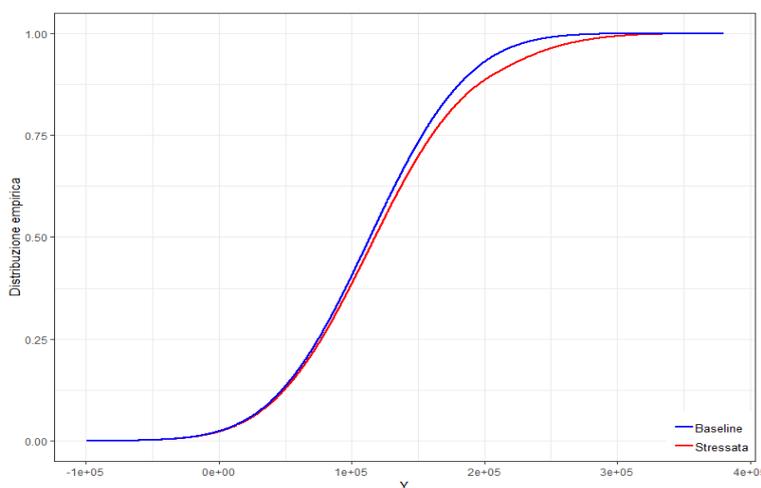


Figura 3.13: Confronto tra le funzioni di ripartizioni empiriche (simulate) sotto la misura di probabilità baseline P (in blu) e stressata Q (in rosso).

sunti dalla misura di sensitività Γ_{N_t} in questo caso e il precedente non si notano particolari differenze (la Figura non è riportata). Dunque, nonostante lo stress (3.6) sia in generale diverso da quello (3.5), l’impatto, in termini di reverse sensitivity measure è molto simile. Lo stress produce il medesimo ranking tra i fattori di rischio.

3.3.3 Modello 1 - Stress avverso all’assicuratore

Nei modelli che abbiamo considerato fino a questo momento, lo stress, sia nel caso (3.5) che nel (3.6), ha un effetto “positivo” per l’assicuratore in quanto, sotto la misura stressata Q , egli dispone in sostanza di un fondo finale più consistente. Nella pratica, molto spesso, si è interessati a casi in cui lo stress abbia un effetto “negativo”; si noti che in generale il risk management è finalizzato alla gestione e al controllo degli scenari avversi. Alla luce di tale considerazione, mantenuta inalterata la struttura del Modello 1, consideriamo come variabile aleatoria output del modello stocastico l’opposto del fondo finale, ovvero:

$$Y = -Z_n = g(N_1, \dots, N_{\omega-x}). \quad (3.7)$$

Questa posizione consente di analizzare, come vedremo, un portafoglio di rendite ove lo stress rappresenta uno scenario avverso all'assicuratore. Interpretativamente possiamo immaginare che l'assicuratore sia interessato a studiare, tramite un modello interno parziale, la struttura probabilistica del suo portafoglio di rendite sotto una particolare struttura di stress che comporti uno scenario avverso alla sua gestione.

Procediamo con un esempio numerico per comprendere le conseguenze di questa diversa assunzione iniziale. Consideriamo un nuovo stress che, come nell'esempio precedente, coinvolge sia il Value at Risk che l'Expected Shortfall, sia esso siffatto:

$$\begin{aligned} \alpha &= 0.9, & q &= Var_{\alpha}^Q(Y) = 0, \\ stress_{ES} &= 0.2, & s &= ES_{\alpha}^Q(Y) = E(Y|Y > q) (1 + stress_{ES}). \end{aligned} \tag{3.8}$$

Osserviamo che, sotto queste condizioni, sono verificate le ipotesi del Teorema 3. Per comprendere la natura dello stress (3.8), si osservi nella Figura 3.14 l'istogramma della distribuzione di Y e la derivata di RN soluzione del Problema (2.9) per questo caso. La Figura 3.14 è in linea con la 3.12 in

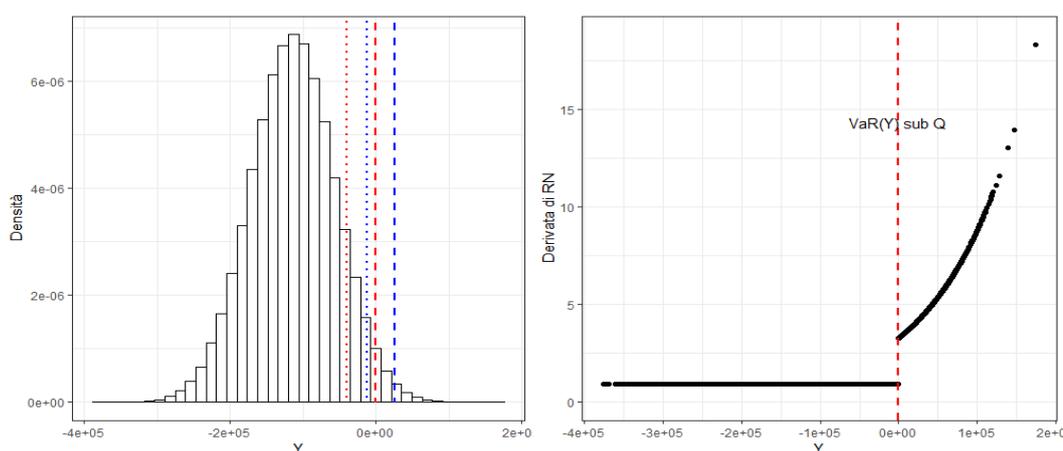


Figura 3.14: A sinistra: istogramma di Y ottenuto dai valori simulati del fondo finale. La linea rossa a punti coincide con il livello del $Var_{\alpha}^P(Y)$ mentre quella a tratteggio con invece $q = Var_{\alpha}^Q(Y)$. La linea blu a punti coincide con il livello del $ES_{\alpha}^P(Y)$ mentre quella a tratteggio con invece $s = ES_{\alpha}^Q(Y)$. A destra: derivata di RN ξ simulata soluzione del Problema (2.9) sotto al vincolo espresso da (3.8).

quanto stiamo impiegando in entrambi i casi uno stress ricavato dal Problema (2.9). La differenza sostanziale giace nella variabile aleatoria output che è stata cambiata di segno. Lo stress (3.8) ha un'importante interpretazione in quanto comporta che, sotto la probabilità stressata Q , il Value at Risk sia pari a 0, che rappresenta per l'assicuratore la soglia tra ricavo e perdita finale. In questo caso l'Expected Shortfall sub Q è quindi interpretabile come perdita media per l'assicuratore (condizionatamente all'informazione che si verifichi una perdita). Alla luce della sua natura interpretativa si continuerà a considerare lo stress (3.8) anche nei prossimi modelli.

Particolarmente interessante è notare come si comporta lo stress (3.8) sui fattori di rischio. Prendiamo ad esempio inizialmente in considerazione la variabile aleatoria N_1 . Diversamente dalle casistiche precedenti, l'impatto agisce in direzione opposta (si veda la Figura 3.15). La funzione di ripartizione di N_1 sotto la misura stressata Q è sempre minore di quella sotto la misura baseline P (diversamente da quanto accadeva nel caso della Figura 3.7). Per visualizzare meglio il fenomeno, si osservi il confronto tra la funzione di probabilità prima e dopo lo stress (3.8). La distribuzione di N_1 sotto Q assegna un peso maggiore ad un maggior numero di individui in vita rispetto a quella sub P , ciò comporta in generale che l'assicuratore dovrà pagare più

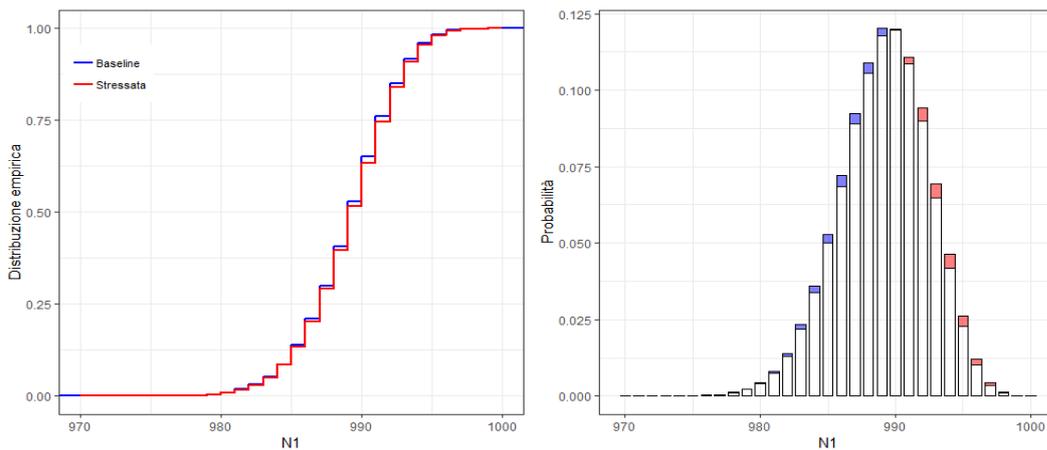


Figura 3.15: A sinistra: confronto tra funzione di ripartizione empirica di N_1 sotto la misura di probabilità baseline (blu) e quella stressata (rosso). A destra: confronto tra funzione di probabilità di N_1 sotto la misura di probabilità baseline (blu) e quella stressata (rosso).

rate di rendita e quindi un abbassamento del livello del fondo. Si realizza dunque uno scenario avverso all'assicuratore.

Consideriamo ora il caso generale per i fattori di rischio N_t . Prendiamo come riferimento nuovamente i valori di t pari a 1, 5, 10, 15, 20, 25, 30 e 35. Confrontando la Figura 3.16 con la 3.8, si nota che l'impatto dello stress agisce, ad ogni epoca, in direzioni opposte nei due casi. Sotto la misura stressata Q , ad ogni epoca t , si ha una sopravvivenza maggiore nel portafoglio di assicurati. Come nel caso base, riesce anche questa volta che l'impatto dello stress sia crescente nei primi quattro grafici, e poi scemi lentamente. La variabile aleatoria N_{35} non sembra essere impattata dallo stress. Per confermare o

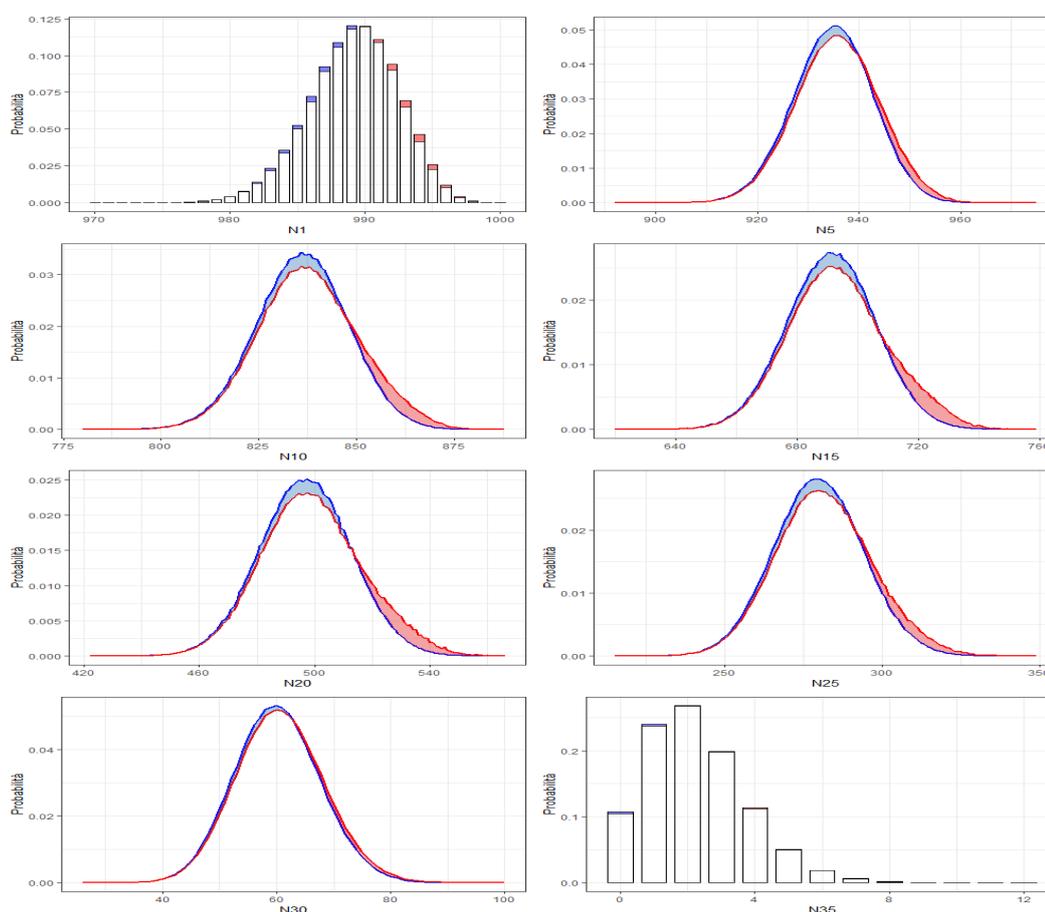


Figura 3.16: Confronto tra funzioni di probabilità prima e dopo lo stress (3.8) per $N_1, N_5, N_{10}, N_{15}, N_{20}, N_{25}, N_{30}$ e N_{35} . La distribuzione sotto la misura baseline è rappresentata in blu, quella sotto al misura stressata in rosso.

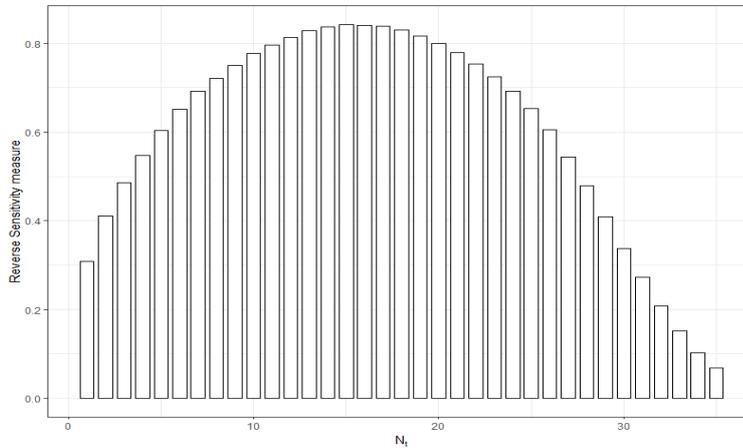


Figura 3.17: Reverse sensitivity measure Γ_{N_t} , per $t = 1, \dots, 35$.

smentire questa analisi, si può considerare la reverse sensitivity measure Γ_{N_t} (Figura 3.17). Si può osservare dalla Figura che la misura di sensibilità, anche se in questo caso è sempre positiva, ha un profilo molto simile in termini assoluti a quello della Figura 3.10. Pertanto si può concludere che i fattori di rischio che maggiormente causano l'incremento o il decremento del capitale sono gli stessi.

3.4 Modello 2: scenari multipli per la mortalità

I modelli considerati sin qui hanno, in sostanza, un'unica fonte di incertezza descritta dalla struttura binomiale (3.3), che riflette il rischio di fluttuazioni casuali attorno il valore atteso causate dalla variabilità intrinseca a N_t . Per estendere il modello considerando anche il rischio di longevità (rischio di natura sistematica) e preservandone la struttura, consideriamo tre scenari di mortalità distinti. L'obiettivo di questa estensione è in generale tener conto del rischio di longevità coerentemente con i fenomeni di espansione e rettangolarizzazione, per dettagli si veda Pitacco et al. (2009). Consideriamo come scenario centrale (*med*) quello utilizzato in precedenza, e i due scenari alternativi introdotti in sezione 3.2 (vedi Figura 3.3). uno scenario in cui si ha una riduzione della durata media di vita (*max*) e uno in cui si ha un aumento della durata media di vita (*min*). Si osservi che i due scenari

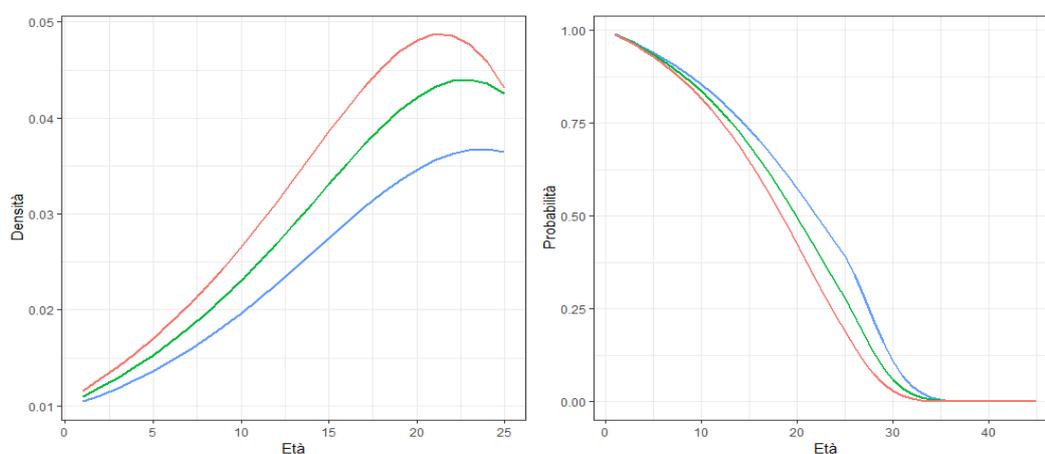


Figura 3.18: Confronto mortalità dei tre scenari. A sinistra: curva dei decessi (parte non interpolata). A destra: funzione di sopravvivenza. In verde lo scenario *med*, in blu lo scenario *min* e in rosso lo scenario *max*.

min e *max* sono scelti equiprobabili; ciò permette di effettuare un confronto tra questo modello e il precedente. Nella Figura 3.18 possiamo osservare il comportamento della mortalità nei tre scenari; in particolare si osservi che dal passaggio da *max* a *min* si possono cogliere entrambi i fenomeni già menzionati (nel grafico di destra) in quanto la curva in primo luogo trasla verso età più avanzate (espansione) e in secondo luogo è più concentrata intorno alla moda (rettangolarizzazione). Ciò comporta un incremento della durata di vita e inoltre sempre più individui decedono all'età modale. Nel grafico a sinistra si è troncato l'asse delle ascisse a $t = 25$ per non visualizzare l'effetto dell'interpolazione lineare.

3.4.1 Analisi di sensitività

Immaginiamo che l'assicuratore disponga di un portafoglio di rendite vitalizie analogo a quello utilizzato nel Modello 1 soggetto però a due fonti di incertezza. Oltre all'aleatorietà già analizzata nei casi precedenti indotta dalla struttura binomiale (3.3), consideriamo, per descrivere i tre possibili scenari che determinano la sopravvivenza del portafoglio di assicurati, un ente aleatorio S che assume tre determinazioni *min*, *med* e *max* con relative probabilità ρ_{min} , ρ_{med} e ρ_{max} (sotto la misura baseline P). La mortalità

in questo modello stocastico è dunque descritta nuovamente dalla relazione (3.3) ove però, diversamente dal Modello 1, $\tilde{q}_{x+t,t}$ è una variabile aleatoria che assume determinazioni diverse a seconda dello scenario S :

$$\tilde{q}_{x+t,t} = \begin{cases} q_x^U, & \text{se } S = \text{max}, & \rho_{\text{max}} = P(S = \text{max}) = 0.2, \\ q_x^M, & \text{se } S = \text{med}, & \rho_{\text{med}} = P(S = \text{med}) = 0.6, \\ q_x^L, & \text{se } S = \text{min}, & \rho_{\text{min}} = P(S = \text{min}) = 0.2. \end{cases} \quad (3.9)$$

Si noti che l'incertezza di $\tilde{q}_{x+t,t}$ opera a livello della generica simulazione. Dunque, fissata una simulazione, la coorte di N_0 individui iniziali è soggetta al medesimo scenario di mortalità; quale questo scenario sia dipende da S attraverso le probabilità ρ_{min} , ρ_{med} e ρ_{max} . Nel Modello 2 la variabile aleatoria output è:

$$Y = -Z_n = g(N_1, \dots, N_{\omega-x-1}, S) \quad (3.10)$$

Come è evidente dalla (3.10), ci mettiamo nuovamente nell'ottica di costruire uno stress che rappresenti una situazione avversa all'assicuratore. Come nei casi precedenti immaginiamo inoltre che l'assicuratore calibri un margine iniziale da allocare al portafoglio per limitare la probabilità di rovina al 2.5%.

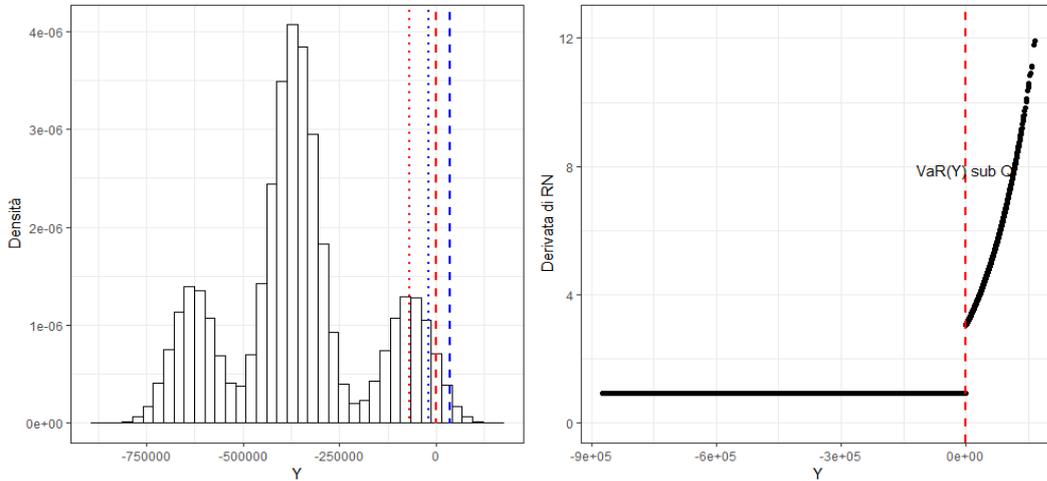


Figura 3.19: A sinistra: istogramma di Y ottenuto dai valori simulati del fondo finale. La linea rossa a punti coincide con il livello del $\text{VaR}_\alpha^P(Y)$ mentre quella a tratteggio con invece $q = \text{Var}_\alpha^Q(Y)$. La linea blu a punti coincide con il livello del $\text{ES}_\alpha^P(Y)$ mentre quella a tratteggio con invece $s = \text{ES}_\alpha^Q(Y)$. A destra: derivata di RN ξ simulata soluzione del Problema (2.9) sotto al vincolo espresso da (3.8).

Tale margine è in questo caso pari a $M = 100\,500$ ovvero circa 3 volte quello del Modello 1. Introdurre un nuovo fattore di rischio ha pertanto costretto l'assicuratore ad allocare un margine maggiore per far fronte alla perdita aleatoria.

Consideriamo per lo stress la medesima struttura (3.8) già utilizzata nel caso precedente ove si stressa contemporaneamente sia il Value at Risk che l'Expected Shortfall, in particolare si rammenti che in (3.8) si è posto il vincolo $q = VaR_\alpha^Q(Y) = 0$; ciò comporta che sotto la misura di probabilità stressata Q valori di Y maggiori del Value at Risk comportano per l'assicuratore una perdita finale. Procediamo analizzando la distribuzione di probabilità della variabile di interesse. La Figura 3.19 a sinistra fornisce indicazioni di come le nuove ipotesi probabilistiche per la mortalità di portafoglio hanno impattato sulla distribuzione di Y ; in particolare si noti l'andamento trimodale.

Procediamo tramite un confronto grafico tra le funzioni di ripartizione di Y prima e dopo lo stress, si veda a tal fine la Figura 3.20 a sinistra. La forma della funzione di ripartizione di Y è in questo caso molto diversa dalle precedenti a causa dell'introduzione dei tre scenari; si notino in particolare i tre scalini in corrispondenza ai valori di Y dove si ha massa probabilistica (vedi Figura 3.19)). Come negli altri casi la funzione di ripartizione sotto la

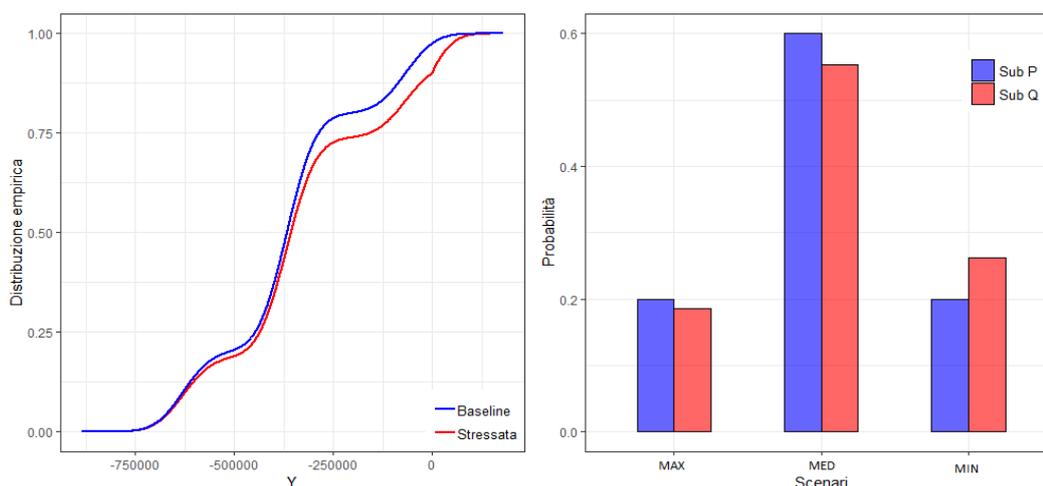


Figura 3.20: A sinistra: confronto tra funzioni di ripartizione di Y sotto le due misure. A destra: confronto tra probabilità dei tre scenari di mortalità sotto le due misure.

probabilità stressata Q è sempre minore di quella sotto la probabilità baseline P . In presenza di scenari multipli che descrivono la sopravvivenza degli assicurati particolarmente interessante è l'analisi dell'effetto che lo stress ha sulle probabilità dei tre scenari date in input. Nella Figura 3.20 a destra, rammentando che lo stress ha un impatto negativo sulla gestione del fondo, non stupisce che la probabilità sotto la misura stressata Q dello scenario *min* (quello con massima longevità) è aumentata da 0.2 a 0.262. La probabilità degli altri due scenari è invece diminuita. In particolare per lo scenario *max* da 0.2 a 0.185 e per *med* da 0.6 a 0.553.

Procediamo l'analisi tramite lo studio dell'impatto che lo stress ha comporta-

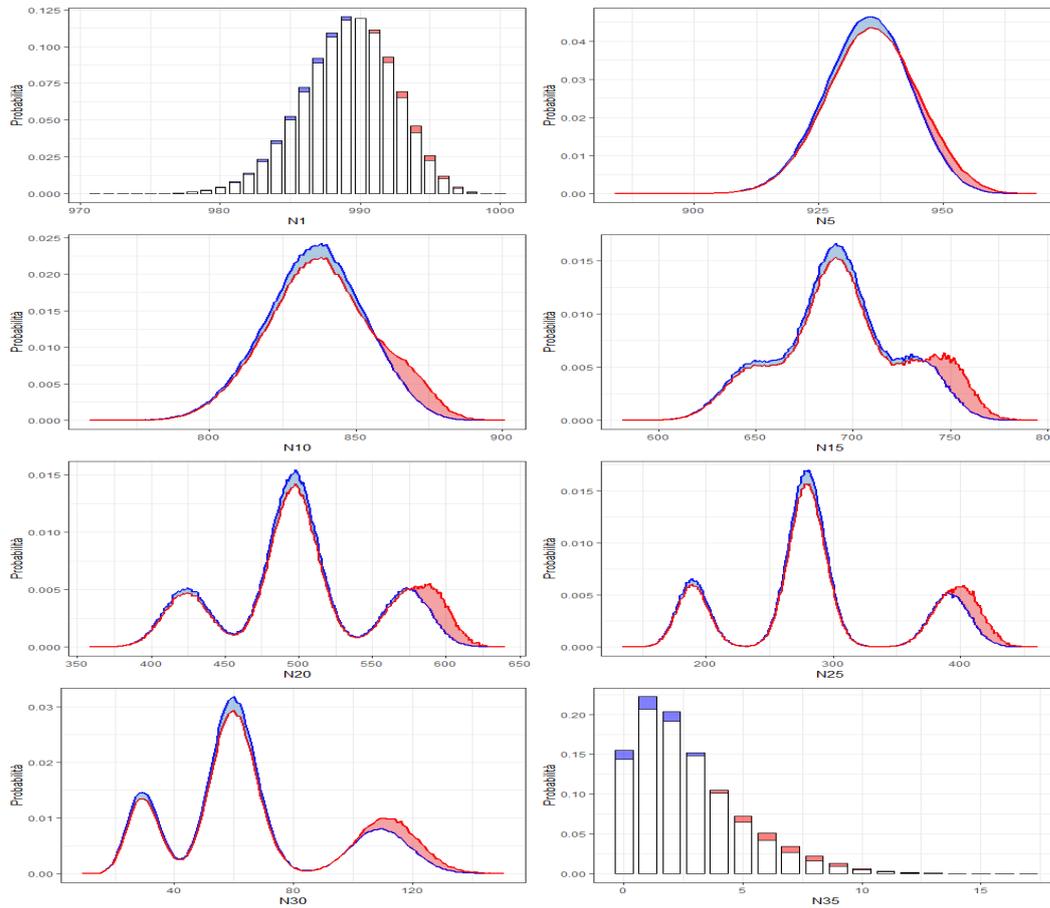


Figura 3.21: Confronto tra le funzioni di probabilità di N_t per $t = 1, 5, 10, 15, 20, 25, 30$ e 35 . La funzione di probabilità di N_t sotto al misura di probabilità baseline è in blu, quella sotto la misura stressata in rosso.

to sulle distribuzioni degli N_t . Come nel Modello 1, prendiamo in considerazione il campione con $t = 1, 5, 10, 15, 20, 25, 30$ e 35 (Figura 3.21). Osserviamo in primo luogo l'effetto che la presenza dei tre scenari comporta sui numeri aleatori N_t . Procedendo da $t = 1$ a $t = 25$ la distribuzione si dilata sempre di più fino ad acquisire una forma trimodale ben delineata. Successivamente appare ricompattarsi fino a ripresentare un'unica moda ad epoche estreme (si rammenti che N_{35} è il numero aleatorio di individui in vita a 100 anni). Lo stress colpisce sempre nella medesima direzione: sotto la nuova misura si ha maggiore sopravvivenza. In particolare si può notare che, qualora siano presenti le tre mode, è solo quella relativa allo scenario *min* ad avere un incremento probabilistico. Le altre presentano un decremento. Non è in questo caso semplice analizzare il comportamento dello stress per $t = 10, \dots, 25$, per comprendere quale dei fattori di rischio sia più impattato dal cambio di misura e ricavarne un ranking. Concludiamo l'analisi del Modello 2 misurando la sensitività dei fattori di rischio. La Figura 3.22 presenta i valori della reverse sensitivity measure Γ_{N_t} per $t = 1, \dots, 35$. La Figura ci consente di comprendere meglio quanto evidenziato dalla 3.21. In primo luogo osserviamo che l'andamento di Γ_{N_t} , al variare di t , è campanulare, prima crescente e poi decrescente. Le variabili aleatorie N_t per $t = 15, \dots, 25$ (circa) sono tutte fortemente impattate dal cambio di misura quasi allo stesso modo a

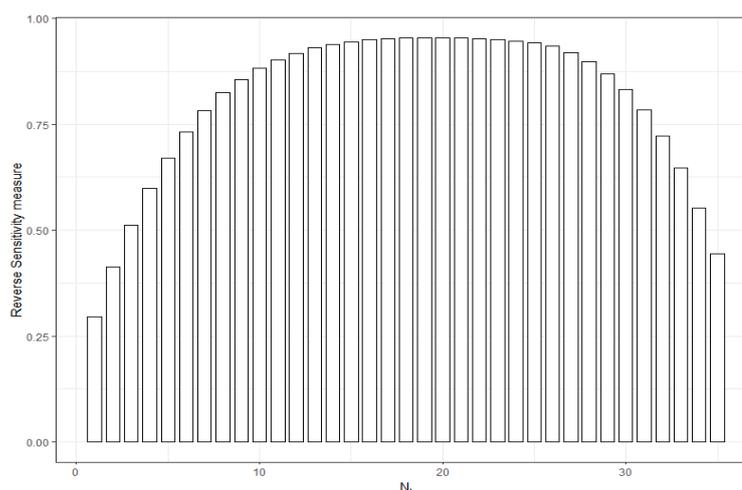


Figura 3.22: Reverse sensitivity measure (Γ_{N_t}), per $t = 1, \dots, 35$.

differenza di quanto avveniva nel Modello 1 (vedi Figura 3.10) . Infine si può osservare che ad età estreme l'effetto dello stress, come già evidenziato dalla Figura 3.21, scemi drasticamente.

In questo modello, come si è descritto nella prima parte, sono presenti sia il rischio di processo (la struttura binomiale) che quello sistematico (i tre scenari aleatori). In presenza di più scenari per la descrizione della mortalità del portafoglio si è interessati a misurare anche la sensitività degli scenari al cambio di misura. Ciò permetterebbe in generale di comprendere quale delle due fonti di rischio è complessivamente più importante. Si osservi però che le variabili aleatorie N_t inglobano entrambi gli effetti che andrebbero incorporati. Il problema in questo caso è che, tutte le volte che per Y si realizza un valore maggiore del suo Value at Risk sotto la misura stressata, necessariamente ciò è causato da una simulazione di mortalità dallo scenario *max*. Ciò è equivalente ad osservare che le determinazioni della derivata di RN associate agli scenari *min* e *med* sono costanti, non permettendo pertanto di studiare il comportamento di N_t condizionatamente allo scenario. Queste considerazioni fanno in ogni caso comprendere che la variabile S è sicuramente il fattore di rischio preponderante per questo modello.

3.5 Modello 3: scenari multipli per la mortalità con tasso aleatorio

Consideriamo in questa sezione, un nuovo modello per la valutazione del portafoglio di rendite. Mantenendo inalterata la struttura per la descrizione della sopravvivenza degli assicurati del portafoglio rispetto quella del Modello 2 (impiegando pertanto i tre scenari *min*, *med* e *max* già analizzati), introduciamo un modello stocastico che descriva l'evolversi nel tempo del tasso di interesse. Il tasso di interesse rappresenta in generale un'importante fonte di incertezza in quanto, come è noto, una forte componente della rischiosità di un portafoglio di rendite vitalizie è dovuta alle performance degli investimenti finanziari. Originariamente il modello di Vasicek (1977) era impiegato per

descrivere l'andamento del tasso privo di rischio. In questo capitolo, come in Olivieri e Pitacco (2003), sarà utilizzato per descrivere l'andamento, su tutto l'arco temporale di riferimento, degli attivi che l'assicuratore ha collocato a copertura delle riserve matematiche. Indicato con r_t il tasso spot istantaneo in t , il modello di Vasicek descrive l'evolversi nel tempo di r_t attraverso la seguente equazione differenziale stocastica:

$$dr_t = \alpha(\beta - r_t)dt + \sigma dW_t, \quad (3.11)$$

ove $\{W_t\}$ è un processo di Wiener standard, e β, α e σ sono parametri reali positivi: β rappresenta il livello medio del tasso spot istantaneo a lungo termine, α esprime la velocità con la quale il processo converge a β e infine σ descrive la volatilità del processo. All'epoca iniziale il livello del tasso è noto e pari a r_0 . In seguito, per la simulazione delle traiettorie del tasso d'interesse, si impiegherà il seguente risultato relativo al modello di Vasicek (per dettagli si veda Brigo e Mercurio (2006)). Condizionatamente all'informazione disponibile all'epoca s , con $s \leq t$, r_t è così distribuito:

$$r_t \sim N \left(r_s e^{-\alpha(t-s)} + \beta(1 - e^{-\alpha(t-s)}), \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha(t-s)}) \right). \quad (3.12)$$

A partire da r_t , ricaviamo (in via approssimata) i tassi semplici aleatori per le epoche $t = 1, \dots, \omega - x$ tramite l'espressione $\tilde{i}_t = e^{r^{t-1}} - 1$. Questi sono successivamente impiegati nell'equazione ricorsiva del fondo di portafoglio (3.2). Si noti che il primo tasso semplice \tilde{i}_1 è dunque un numero certo. All'epoca iniziale, il tasso medio relativo all'epoca futura t è individuato quindi dalla seguente espressione:

$$E(r_t) = r_0 e^{-\alpha t} + \beta(1 - e^{-\alpha t}). \quad (3.13)$$

Si osservi che si sta assumendo, come è ragionevole, l'indipendenza tra il fenomeno stocastico che descrive la mortalità del portafoglio degli assicurati e quello che descrive l'evolversi nel tempo del tasso di interesse. In questo modello, dove sono presenti contemporaneamente tre importanti fonti di

incertezza è di fondamentale importanza rilevare quale di queste tre sia preponderante sulle altre. In particolare, introducendo il tasso aleatorio, si può ora studiare come si comportano nel tempo sopravvivenza e tasso per comprendere quale dei due sia più importante in un determinato istante futuro. Procediamo con l'analisi di sensitività.

3.5.1 Analisi di sensitività

Immaginiamo che l'assicuratore disponga di un portafoglio di rendite vitalizie analogo a quello utilizzato nel Modello 2 nel quale però impiega un modello stocastico per la descrizione dell'andamento del tasso di interesse. La variabile aleatoria Y output del modello è:

$$Y = -Z_n = g(S, N_1, \dots, N_{\omega-x-1}, \tilde{i}_2, \dots, \tilde{i}_{\omega-x}) \quad (3.14)$$

Il vettore aleatorio dei fattori di rischio \mathbf{X} contiene dunque anche i tassi semplici aleatori ad ogni epoca \tilde{i}_t . Per quanto concerne il modello di Vasicek consideriamo i seguenti parametri: $\alpha = 0.1$, $\sigma = 0.007$, $\beta = 0.02956$ e per il tasso di interesse iniziale poniamo $r_0 = 0.02956$. Si noti che il tasso istantaneo iniziale r_0 è pari a β , pertanto la traiettoria media implicata dal modello è costante e pari a 3% (vedi (3.13)). Questa scelta permette, come si era fatto nel Modello 2, di confrontare questo modello con i precedenti.

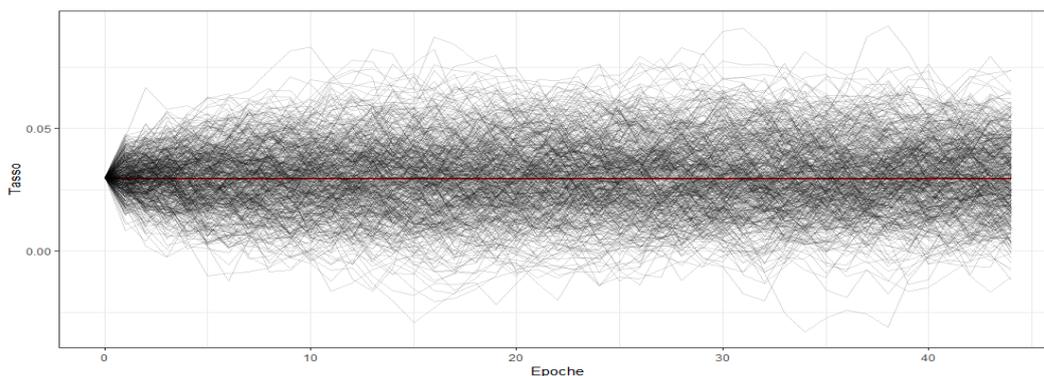


Figura 3.23: Traiettorie delle prime 500 simulazioni del tasso semplice impiegato nell'equazione del fondo 3.2. La traiettoria rossa indica il tasso medio implicato dal modello.

Per comprendere il comportamento dell'equazione (3.11) analizziamo in un grafico le prime 500 traiettorie dei tassi d'interesse. Come si può osservare dalla Figura 3.23, il modello di Vasicek non pone restrizioni sul segno di r_t (e dunque di \tilde{i}_t) che può diventare anche negativo. Questo aspetto, ritenuto in passato un limite del modello, è considerato oggi, a causa dei livelli odierni dei tassi di interesse, uno dei punti di forza che lo contraddistinguono. Come già fatto in precedenza, l'assicuratore alloca al portafoglio un margine iniziale per limitare la probabilità di rovina al 2.5%, che riesce pari a $M = 269\,000$. Si rammenti che nel Modello 2 era 100\,500, quindi, come nel passaggio tra il Modello 1 e il 2, considerare un'ulteriore fonte rischio costringe l'assicuratore ad allocare un margine maggiore.

Alla luce della posizione (3.14), siamo quindi interessati ad analizzare come uno stress che rappresenta uno scenario avverso all'assicuratore impatti sul modello stocastico introdotto. Per lo stress impieghiamo nuovamente la struttura (3.8). Si rammenti che lo stress interessa sia il Value at Risk che l'Expected Shortfall, e si impone in particolare che il Value at Risk sotto la misura stressata sia pari a 0. La presenza del modello stocastico per il

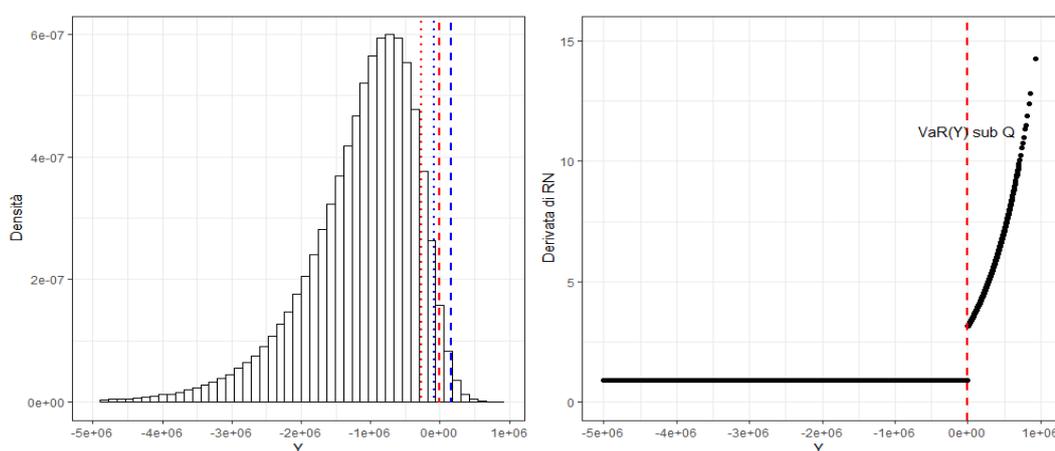


Figura 3.24: A sinistra: istogramma di Y ottenuto dai valori simulati del fondo finale. La linea rossa a punti coincide con il livello del $Var_{\alpha}^P(Y)$ mentre quella a tratteggio con invece $q = Var_{\alpha}^Q(Y)$. La linea blu a punti coincide con il livello del $ES_{\alpha}^P(Y)$ mentre quella a tratteggio con invece $s = ES_{\alpha}^Q(Y)$. A destra: derivata di RN ξ soluzione del Problema (2.9) sotto al vincolo espresso da (3.8).

tasso d'interesse compatta la massa probabilistica della distribuzione di Y , non sono pertanto più presenti le tre mode (si confronti la Figura 3.24 con la Figura 3.19). La distribuzione presenta inoltre una marcata asimmetria negativa.

Procediamo con il confronto tra le funzioni di ripartizione di Y sotto le due misure, si veda a tal fine la Figura 3.25 a sinistra. Come osservato dall'istogramma, la funzione di ripartizione di Y sotto la misura baseline non presenta i tre scalini che si osservavano nel Modello 2. Valgono in generale i medesimi commenti già proposti in precedenza. Nella Figura 3.25 a destra, è analizzato l'impatto che lo stress ha sulla probabilità dei tre scenari. Come nel caso precedente, lo scenario *min* è l'unica a presentare un incremento probabilistico. In particolare si passa da 0.2 a 0.236. Per quanto concerne gli altri due scenari, nel caso *max* si passa da 0.2 a 0.186 mentre per *med* da 0.6 a 0.578. Si osservi che, rispetto al caso precedente, l'impatto dello stress sugli scenari è dunque più contenuto a causa della presenza del tasso aleatorio che assorbe una parte della distorsione probabilistica. Nel seguito, quantificheremo tramite misura di rischio l'importanza della variabile di scenario per lo stress (3.8). Analizziamo l'impatto che lo stress ha avuto sugli altri fattori di rischio in input. In questo caso si può prendere in considera-

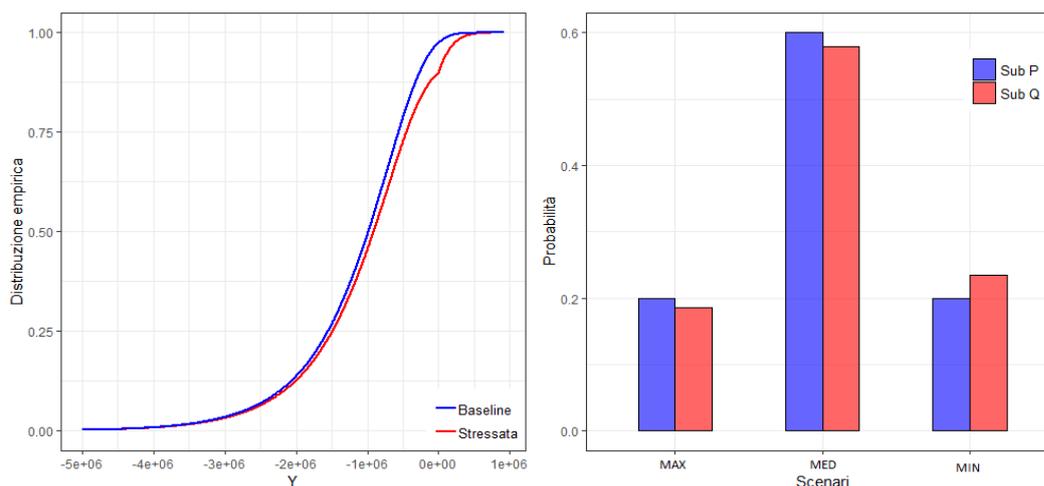


Figura 3.25: A sinistra: confronto tra funzioni di ripartizione di Y sotto le due misure di probabilità. A destra: confronto tra probabilità dei tre scenari di mortalità sotto le due misure di probabilità.

zione non solo l'effetto, al variare di t , sulle variabili aleatorie N_t ma anche sulle \tilde{i}_t . Non presentiamo il grafico dello stress sui singoli N_t in quanto è simile a quello del Modello 2 (Figura 3.21) anche se l'effetto dello stress è più contenuto; valutiamo invece l'impatto sulle variabili aleatorie \tilde{i}_t , si veda a tal fine la Figura 3.26. Nella Figura è rappresentato il comportamento a confronto della funzione di ripartizione di \tilde{i}_t , $t = 2, 5, 10, 15, 20, 25, 30$ e 35 , sotto la probabilità baseline P e quella stressata Q . Come si può osservare la funzione di ripartizione sotto la misura baseline è sempre minore di quella sotto la misura stressata ad indicare che lo stress comporta in genera-

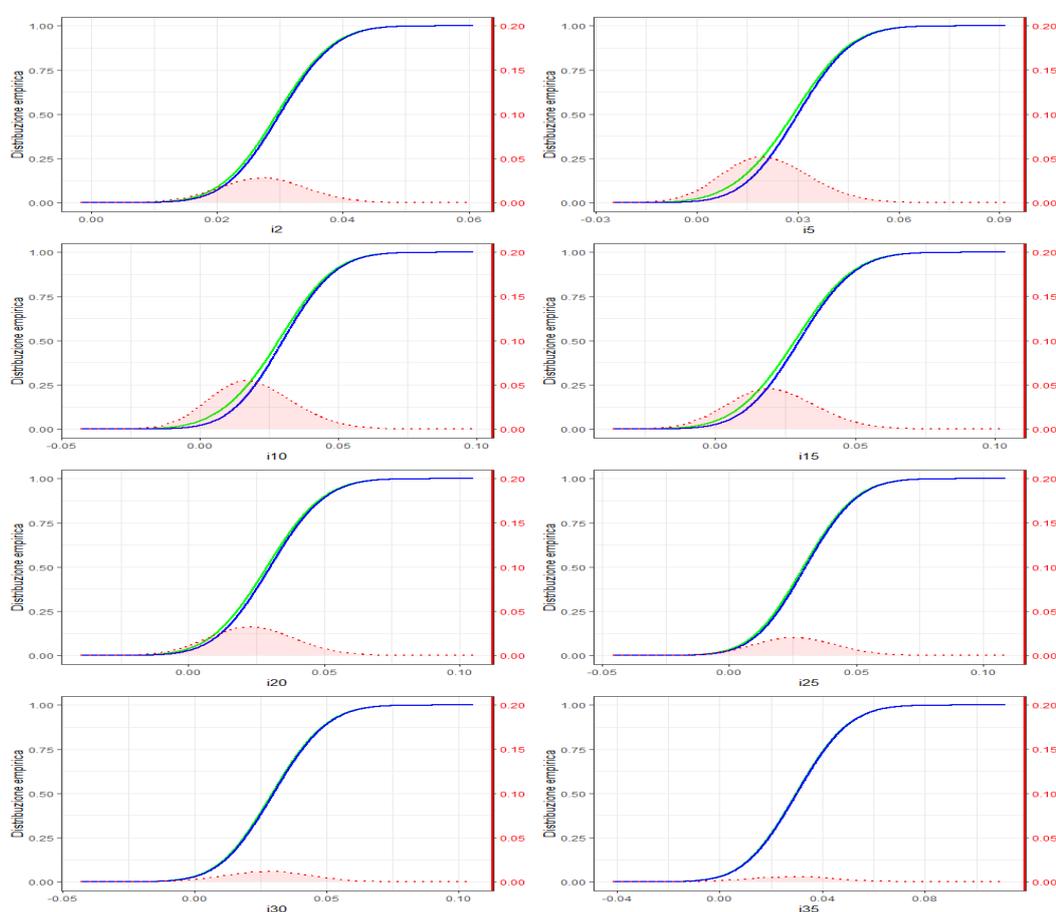


Figura 3.26: Confronto tra le funzioni di ripartizione di \tilde{i}_t per $t = 2, 5, 10, 15, 20, 25, 30$ e 35 . La funzione di ripartizione di \tilde{i}_t sotto al misura di probabilità baseline è in blu, quella sotto la misura stressata in verde. L'area rossa ombreggiata indica la differenza tra le due funzioni di ripartizione (sub Q - sub P) ed è riscalata sull'asse delle ordinate rosso a destra del grafico.

le tassi di interesse minori (l'assicuratore consegue realizzazioni minori sulle performance finanziarie degli attivi che ha allocato a copertura della riserva matematica). In particolare, si può osservare un comportamento dapprima crescente (fino $t = 10$) e poi lentamente decrescente, e ad epoche estreme non sembra esserci sostanziale differenza nella distribuzione prima e dopo lo stress.

Confrontiamo ora tramite un grafico il comportamento della reverse sensitivity measure al variare di t per N_t (effetto complessivo mortalità) e \tilde{i}_t (effetto complessivo tasso). Si osservi che, in generale, quando studiamo il comportamento della variabile aleatoria N_t , stiamo implicitamente tenendo conto simultaneamente sia del rischio di processo legato all'aleatorietà intrinseca all'effetto mortalità (nel nostro caso generata dalla binomiale (3.3)) che all'effetto dovuto alla presenza degli scenari (rischio sistematico). Come abbiamo già analizzato, lo stress comporta tassi in generale più bassi e dunque la reverse sensitivity measure legata a \tilde{i}_t è negativa per ogni t . Nella Figura 3.27 la reverse sensitivity measure per il tasso è considerata in valore assoluto per ottenere un confronto efficace.

La Figura conferma quanto detto sull'andamento nel tempo dei due profili di stress. Per quanto concerne il tasso si ha un profilo prima fortemente

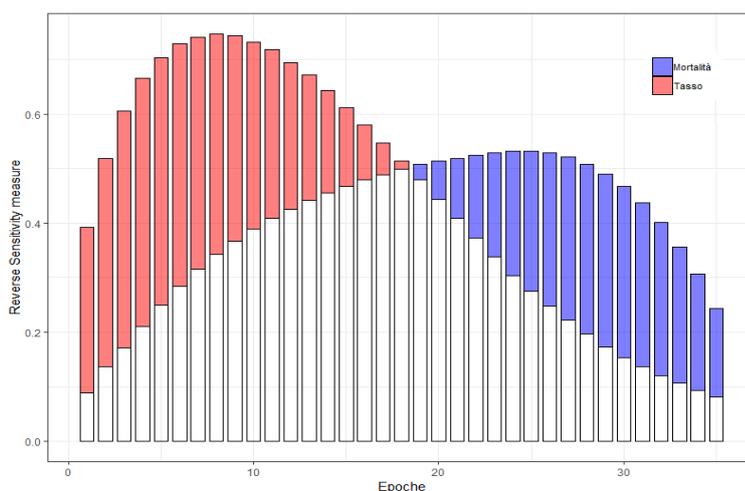


Figura 3.27: Confronto tra reverse sensitivity measure di N_t (rosso) e \tilde{i}_{t+1} (blu).

crescente, fino a $t = 8$, e poi decrescente. Per la mortalità il profilo della reverse sensitivity measure appare crescente fino a $t = 24$, e successivamente decresce. Ad ogni modo è evidente che il fattore di rischio predominante ad ogni epoca precedente a $t = 19$ è il tasso d'interesse mentre successivamente è più importante l'effetto della mortalità.

Concludiamo lo studio del Modello 3 considerando un approccio sintetico per la valutazione dell'importanza delle fonti di rischio. La Figura 3.27 ci fornisce indicazioni sull'importanza dell'effetto dovuto alla mortalità e al tasso per ogni epoca $t = 1, \dots, 35$. Con riferimento al medesimo intervallo temporale, indicata con $T = 35$ l'ultima epoca di riferimento, introduciamo le seguenti due variabili aleatorie:

$$\tilde{T} = \frac{1}{N_0} \sum_{t=0}^T N_t, \quad \tilde{i} = \left(\prod_{t=1}^{T+1} (1 + \tilde{i}_t) \right)^{\frac{1}{T+1}} - 1. \quad (3.15)$$

Dunque, \tilde{T} fornisce una sintesi per le variabili aleatorie N_t , e rappresenta, con riferimento ai primi 35 anni di assicurazione, il numero aleatorio di anni vissuti dalla collettività di N_0 assicurati per assicurato. La variabile aleatoria \tilde{i} , che sintetizza le \tilde{i}_t , è il tasso annuo semplice che, se costante sul periodo $[0, T]$, garantisce lo stesso livello di capitalizzazione della sequenza $\tilde{i}_1, \dots, \tilde{i}_{T+1}$. Con riferimento a queste due variabili aleatorie è possibile calcolare la reverse sensitivity measure per comprendere l'effetto complessivo di mortalità (rischio di processo e sistematico) e tasso di interesse (rischio sistematico). Si ha:

$$\Gamma_{\tilde{T}} = \mathcal{S}(\tilde{T}, \zeta) = 0.515, \quad \Gamma_{\tilde{i}} = \mathcal{S}(\tilde{i}, \zeta) = -0.696$$

Possiamo dunque concludere che, analizzato complessivamente, l'effetto dovuto al tasso è più importante di quello della mortalità. Osserviamo che, considerando nuovamente le espressioni in (3.15), ma con riferimento ad intervalli

temporali troncati, individuati dai primi 10 e 25 anni, si ottiene:

$$\begin{aligned} 10 \text{ anni : } & \Gamma_{\tilde{T}} = 0.314, \quad \Gamma_{\tilde{z}} = -0.835, \\ 25 \text{ anni : } & \Gamma_{\tilde{T}} = 0.504, \quad \Gamma_{\tilde{z}} = -0.817. \end{aligned}$$

Pertanto considerati intervalli temporali più brevi, a partire dall'epoca iniziale, la differenza tra le due misure di sensitività si acuisce sempre di più (vedi Figura 3.27).

Per misurare la sensitività della variabile di scenario S , che non è di tipo numerico, è necessario effettuare un'opportuna trasformazione. Consideriamo la variabile aleatoria \tilde{M} così definita:

$$\tilde{M} = h(S) = \begin{cases} t_{min} & P(S = min) = \rho_{min} = 0.2, \\ t_{med} & P(S = med) = \rho_{med} = 0.6, \\ t_{max} & P(S = max) = \rho_{max} = 0.2, \end{cases}$$

ove t_{min} , t_{med} e t_{max} indicano rispettivamente la durata media di vita troncata in T di un generico assicurato dall'epoca iniziale, nello scenario min , med e max . Per essa la reverse sensitivity measure è pari a:

$$\Gamma_{\tilde{M}} = \mathcal{S}(\tilde{M}, \zeta) = 0.623.$$

Misurata complessivamente, la sensitività del rischio di longevità, anche se minore di quella del tasso, è molto alta.

Descrivere l'effetto del rischio di longevità impiegando i tre scenari comporta, in primo luogo, una significativa semplificazione del fenomeno di mortalità stocastico. Inoltre, selezionare e assegnare coerentemente una distribuzione iniziale sulla terna di scenari è un'operazione delicata. Queste possibili critiche fanno comprendere che per uno studio più completo della sensitività rispetto il rischio di longevità è necessario un modello pienamente stocastico come quello descritto nella sezione 3.2.

3.6 Modello 4: approccio pienamente stocastico

Introduciamo in questa sezione l'ultimo modello per la descrizione del portafoglio di rendite vitalizie. Consideriamo nuovamente per il tasso di rendimento aleatorio il modello di Vasicek (con medesimi parametri) ma presentiamo una nuova struttura per la descrizione della mortalità del portafoglio di assicurati. Nei Modelli 2 e 3 si sono considerati, per la descrizione della mortalità del portafoglio, tre distinti scenari futuri. Nel Modello 4 si impiega in modo completo il modello M7 analizzato in 3.2. Implicitamente ciò è equivalente a considerare infiniti possibili scenari futuri.

3.6.1 Modello 4 - analisi di sensitività

Ipotizziamo, come negli altri casi, che l'assicuratore allochi un margine iniziale per limitare la probabilità di rovina al 2.5%, esso è pari a $M = 273\,000$, poco più alto del margine del Modello 3. Questo risultato fa intuire che, dal punto di vista della rischiosità, i Modelli 3 e 4 hanno strutture simili. Ciò in generale non deve sorprendere, in quanto l'unica differenza tra tali modelli è rappresentata dall'effetto della mortalità, ove si rammenti che i tre

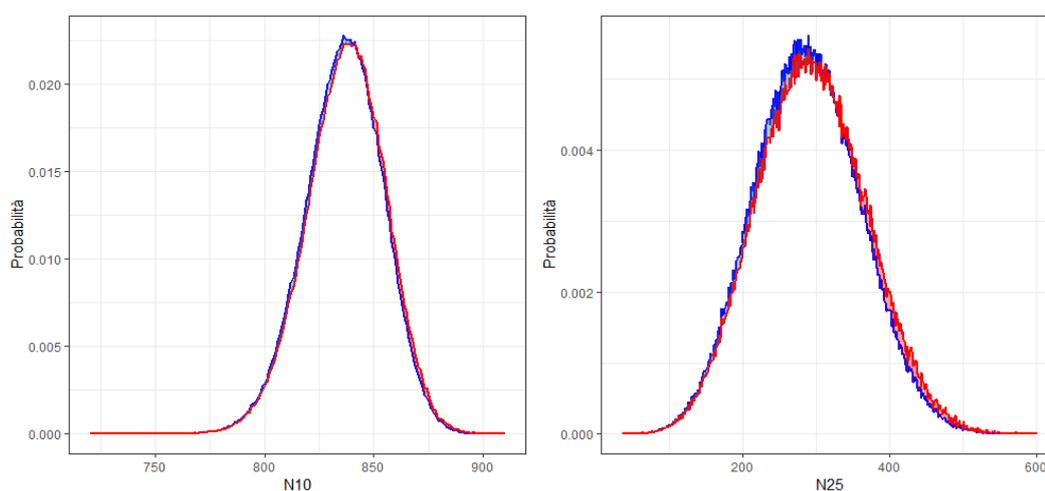


Figura 3.28: Confronto tra le funzioni di probabilità di N_t per $t = 10$ (sinistra) e $t = 25$ (destra). La funzione di probabilità di N_t sotto la misura di probabilità baseline è in blu, quella sotto la misura stressata in rosso.

scenari impiegati nel Modello 3 sono calibrati sui quantili 10 e 90 di quelli del Modello 4. Nell'analisi di sensitività si analizzerà pertanto se i tre scenari colgono adeguatamente o meno la variabilità del modello M7. Per il Modello 4, la variabile aleatoria output Y output è:

$$Y = -Z_n = g(N_1, \dots, N_{\omega-x-1}, \tilde{i}_2, \dots, \tilde{i}_{\omega-x}, \tilde{q}_{x,0}, \dots, \tilde{q}_{\omega-x-1, \omega-x-1}) \quad (3.16)$$

Il vettore aleatorio dei fattori di rischio \mathbf{X} contiene in questo Modello (al posto della variabile di scenario S dei modelli precedenti) le probabilità di decesso aleatorie della coorte di assicurati $\tilde{q}_{x+t,t}$. Impieghiamo nuovamente la struttura dello stress (3.8). A differenza degli esempi precedenti, si decide di non riportare il grafico della distribuzione di Y (e derivata di RN) in quanto è molto simile a quello del Modello 3. Analizziamo l'impatto che lo stress ha avuto sui fattori di rischio in input. Per quanto concerne l'effetto della mortalità, si veda la Figura 3.28, dove si può osservare che le variabili aleatorie N_t non presentano più un andamento trimodale (vedi Figura 3.21). L'entità dello stress è crescente passando da $t = 10$ a $t = 25$, ma comunque modesta. Come negli altri casi, sotto la misura stressata Q si fornisce maggior peso ad un maggior numero di sopravvissuti (scenario avverso all'assicuratore). L'analisi dell'impatto per il tasso di interesse non è riportata in quanto molto

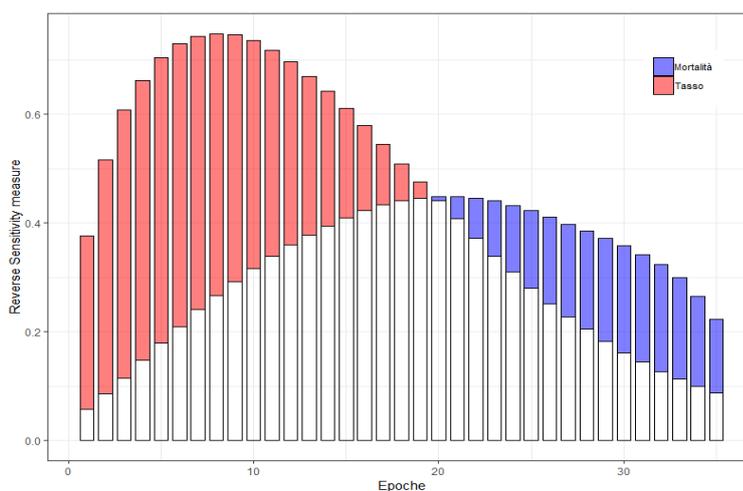


Figura 3.29: Confronto tra reverse sensitivity measure di N_t (rosso) e \hat{i}_{t+1} (blu).

simile a quella della Figura 3.26.

Confrontiamo ora, tramite la reverse sensitivity measure, l'impatto dello stress su tasso e mortalità, si veda a tal fine la Figura 3.29. Anche in questo caso, consideriamo la sensitività del tasso in termini assoluti (è negativa) per effettuare efficacemente il confronto grafico. Confrontando le Figure 3.29 e 3.27 l'andamento generale appare simile, con una importante differenza. Nel Modello 4, quando la sensitività della mortalità prevale sul tasso (da $t = 20$), l'effetto riesce decrescente.

Per misurare quale fattore di rischio ha complessivamente più importanza per il Modello 4, calcoliamo la reverse sensitivity measure con riferimento alle due quantità aleatorie \tilde{T} e \tilde{i} introdotte in (3.15). Per esse si ha in questo caso:

$$\Gamma_{\tilde{T}} = \mathcal{S}(\tilde{T}, \zeta) = 0.446, \quad \Gamma_{\tilde{i}} = \mathcal{S}(\tilde{i}, \zeta) = -0.703$$

Confrontando i risultati con quelli del Modello 3, l'effetto complessivo della mortalità è diminuito. Ciò è imputabile all'eccessivo peso (20 %) dato ai due scenari estremi *min* e *max*. Per misurare il solo rischio sistematico legato alla mortalità, introduciamo la seguente quantità:

$$\tilde{e} = \sum_{h=1}^T h \tilde{p}_{65},$$

ove ${}_h\tilde{p}_{65}$ è la variabile aleatoria che rappresenta la probabilità, per un assicurato in vita a 65 anni, di essere ancora in vita dopo h anni. La variabile aleatoria \tilde{e} misura dunque la durata media di vita a età 65 troncata in T di un generico individuo della collettività. Con riferimento a questa si ha:

$$\Gamma_{\tilde{e}} = \mathcal{S}(\tilde{e}, \zeta) = 0.427.$$

L'effetto del rischio sistematico legato alla mortalità è dunque inferiore a quello misurato tramite le variabili aleatorie \tilde{T} e \tilde{i} . La sensitività del rischio di mortalità è quasi interamente dovuta alla sua componente sistematica.

Capitolo 4

Reverse sensitivity analysis: il rischio di credito

L'obiettivo di questo capitolo è l'applicazione del framework concernente la reverse sensitivity analysis, introdotto nel secondo capitolo, ad un modello per la gestione di un portafoglio di prestiti soggetti a rischio di insolvenza. Come si è già evidenziato nel primo capitolo, gli operatori economici, finanziandosi a vicenda, si aprono alla possibilità che la controparte non faccia fronte globalmente, o in parte, agli impegni assunti. Questa evenienza genera un rischio denominato rischio di credito (o di controparte). In un'accezione più ampia, il rischio di credito si esplica in un peggioramento della qualità creditizia (credit worthiness) della controparte; ciò comporta, in generale, una diminuzione del valore del credito e quindi una sofferenza per il creditore.

Per gli istituti bancari il rischio di credito rappresenta, nella maggior parte dei casi, la fonte di rischio con maggior peso sulla gestione. A conferma di questo aspetto, il rischio di credito, come visto nel Capitolo 1, ha un ruolo centrale nella normativa Basel III. Anche per le imprese di assicurazione, seppur in peso minore rispetto alle banche, il rischio di credito ha un ruolo influente. Prima dell'introduzione di Solvency II, nella valutazione dei requisiti di capitale non era richiesto alle imprese di assicurazione di tener conto di nessun rischio di natura finanziaria. Oggi, è ormai evidente, che per le

imprese di assicurazione i rischi di natura finanziaria stanno assumendo nel tempo un ruolo di sempre maggior rilievo. In letteratura, alla luce della rilevanza che il rischio di credito ha assunto negli ultimi decenni, si è sviluppato un forte interesse che ha permesso la nascita di un vero e proprio settore di studio, in cui si possono distinguere due grandi filoni. Il primo concerne modelli quantitativi per la gestione del rischio di credito finalizzati ad ottenere la distribuzione della perdita aleatoria complessiva su un prefissato arco temporale (credit risk management). A partire da questa distribuzione, usualmente, si ottengono poi i requisiti di capitale per coprire la perdita ad un fissato livello di confidenza. Un secondo filone riguarda valutazioni di prodotti finanziari in presenza di rischio di credito, e in particolare di derivati sul credito. In questo caso è necessaria l'implementazione di un modello stocastico a tempo continuo, in quanto il pay-off di molti prodotti finanziari di questo tipo dipende dall'esatto istante del default. Ci concentreremo esclusivamente sui modelli del primo tipo.

Sebbene sia presente in letteratura una vasta varietà di modelli per il credit risk management, molti di questi hanno caratteristiche di analogia e possono essere visti come casi particolari di un'unica struttura generale. Costruire un modello per il credit risk management implica necessariamente dover affrontare alcune problematiche. La prima, e forse più importante di queste, è dovuta alla scarsità di dati pubblici disponibili per le valutazioni. Un altro problema centrale è la descrizione della complessa struttura di dipendenza delle variabili aleatorie che descrivono i default delle singole imprese. Come vedremo dettagliatamente, questo aspetto, che è il cuore dei modelli per il credit risk management, ha un peso molto importante sulla distribuzione della perdita aleatoria globale; più la dipendenza tra i rischi è forte più la coda destra della distribuzione diventa spessa e lunga. A titolo di esempio si veda la Figura 4.1 nella quale sono confrontate le distribuzioni delle perdite complessive per due portafogli di rischi omonogenei con medesime probabilità di default ma diversa struttura di dipendenza. Si considera, in questo caso, un istituto finanziario che ha segmentato il proprio portafoglio in tre tranches, ognuna con un'opportuna probabilità di default e struttura di di-

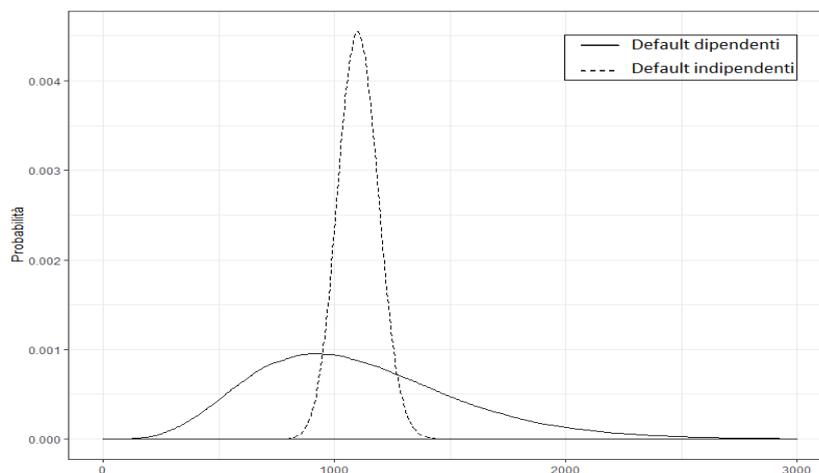


Figura 4.1: Confronto tra la distribuzione della perdita complessiva impiegando modelli con medesima probabilità di default ma diversa struttura di dipendenza. La linea continua è la distribuzione della perdita nel caso i default siano tra loro dipendenti la linea tratteggiata nel caso siano indipendenti. I dati impiegati sono quelli della sezione 4.3 (non è stata impiegata la copula tra le tranches).

pendenza. La perdita totale considera il comportamento complessivo delle tre tranches. Come si può vedere, nel caso di indipendenza tra i default la distribuzione della perdita è molto concentrata. Nel caso di default dipendenti la distribuzione presenta invece asimmetria e coda destra più marcata. Come vedremo successivamente, per descrivere la dipendenza stocastica tra i default si impiegherà una struttura su due livelli. Ciò implicherà la presenza di una dipendenza più forte che andrà quindi ad incrementare il peso della coda destra della distribuzione. In generale, dal punto di vista del risk management, le distribuzioni fortemente asimmetriche con code pesanti sono da trattare con estrema cautela. Per comprendere come è possibile indurre in modo semplice una struttura di dipendenza tra i default del portafoglio procediamo esponendo i modelli mistura. Nel corso del capitolo, per l'esposizione dell'apparato teorico, si fa riferimento a McNeil et al. (2015).

4.1 Modelli Mistura

Nei modelli mistura le variabili aleatorie che descrivono l'evento default per le varie controparti sono ipotizzate dipendere dalle realizzazioni di altre variabili economiche di riferimento. Si suppone in particolare che i default siano stocasticamente indipendenti condizionatamente alle determinazioni di queste variabili di riferimento. Questa tipologia di dipendenza stocastica è ragionevole in quanto ci si aspetta che il comportamento di alcune grandezze macroeconomiche impatti allo stesso modo sulla distribuzione dei default delle singole controparti; ciò comporta l'instaurarsi di un particolare tipo di dipendenza stocastica tra queste ultime. Tale approccio introduce quindi esplicitamente un rischio di natura sistematica nel modello. Molto spesso, inoltre, si fa riferimento a portafogli con un numero elevato di posizioni e pertanto studiare le complesse relazioni sussistenti individualmente tra queste è un'operazione difficilmente realizzabile. È importante osservare che questa famiglia di modelli non rappresentano una classe a sé stante ma generalizzano quasi tutti i modelli impiegati nell'industria. In particolare molti dei cosiddetti modelli soglia (threshold models) possono essere rappresentati equivalentemente come modelli mistura, si veda a tal fine McNeil et al. (2015) per una discussione approfondita. Prima di introdurre la struttura dei modelli mistura presentiamo l'ambiente di lavoro generale.

Introduciamo uno spazio di probabilità (Ω, \mathcal{A}) , e sia P è la misura di probabilità baseline definita su tale spazio. Consideriamo un operatore economico (ad esempio un istituto bancario) avente un portafoglio di m posizioni di credito con relative esposizioni su un orizzonte temporale T . Indichiamo con S_i , $i = 1, \dots, m$, la variabile aleatoria dicotomica che indica se, al tempo T , l' i -esimo debitore è andato incontro a default ($S_i = 1$) oppure no ($S_i = 0$). S_i è quindi l'indicatore del default dell' i -esimo debitore. Sia $\mathbf{S} = (S_1, \dots, S_m)^\top$ il vettore aleatorio avente le S_i come marginali. In questo ambito, indicato con \bar{p}_i la probabilità che l' i -esimo debitore sia insolvente in T , ovvero $\bar{p}_i = P(S_i = 1) = E_P(S_i)$, la correlazione tra gli indicatori di insolvenza S_i e

S_j , nota anche come default correlation, è:

$$\rho(S_i, S_j) = \frac{E(S_i S_j) - \bar{p}_i \bar{p}_j}{\sqrt{(\bar{p}_i - \bar{p}_i^2)(\bar{p}_j - \bar{p}_j^2)}}.$$

Sia M la variabile aleatoria che conta il numero di default, cioè $M = \sum_{i=1}^m S_i$. La perdita causata dal default dell' i -esima controparte (loss given default) è pari a $\delta_i e_i$, dove, con riferimento all' i -esima controparte, $e_i \geq 0$ indica l'esposizione complessiva e $\delta_i \in [0, 1]$ la percentuale di questa che viene persa conseguentemente al default. A partire da queste quantità è possibile definire l'effettiva perdita globale L come:

$$L = \sum_{i=1}^m \delta_i e_i S_i. \quad (4.1)$$

L è la variabile aleatoria d'interesse per questo problema e rappresenterà l'output del modello di sensitività.

Per semplificare l'analisi introduciamo un'importante ipotesi, che permette di tradurre l'idea intuitiva di omogeneità tra le controparti. Supponiamo dunque che il vettore aleatorio \mathbf{S} sia scambiabile, ovvero:

$$\mathbf{S} = (S_1, \dots, S_m) \stackrel{d}{=} (S_{\Pi(1)}, \dots, S_{\Pi(m)}) \text{ per ogni } \Pi,$$

ove $\Pi : \{1, \dots, m\} \rightarrow \{1, \dots, m\}$ indica una permutazione di $1, \dots, m$ e $\stackrel{d}{=}$ indica l'uguaglianza in distribuzione. Si rammenti che ciò comporta che, per ogni $k \in \{1, \dots, m\}$, gli $\binom{m}{k}$ possibili sottovettori k -dimensionali sono ugualmente distribuiti. Questa ipotesi, che semplifica notevolmente il problema, permette di introdurre una notazione semplificata che impiegheremo nel corso del capitolo. Indichiamo dunque con π_k la probabilità di default congiunta di k controparti, ovvero:

$$\pi_k = P(S_{i_1} = 1, \dots, S_{i_k} = 1), \{i_1, \dots, i_k\} \in \{1, \dots, m\}, 1 \leq k \leq m.$$

Nel caso di $k = 1$, utilizzeremo al posto di π_1 semplicemente il simbolo π . Si noti che queste probabilità sono valutate sotto la misura baseline P . Conseguono banalmente le seguenti relazioni, valide per ogni $i \neq j$:

$$\mathbb{E}_P(S_i) = \mathbb{E}_P(S_i^2) = P(S_i = 1) = \pi,$$

$$\text{Var}_P(S_i) = \mathbb{E}_P(S_i^2) - \mathbb{E}_P(S_i)^2 = \pi - \pi^2,$$

$$\mathbb{E}_P(S_i S_j) = P(S_i = 1, S_j = 1) = \pi_2,$$

$$\text{Cov}_P(S_i, S_j) = \mathbb{E}_P(S_i S_j) - \mathbb{E}_P(S_i)\mathbb{E}_P(S_j) = \pi_2 - \pi^2.$$

Dunque, la default correlation è comune ad ogni coppia (S_i, S_j) ed è pari a:

$$\rho_S = \rho(S_i, S_j) = \frac{\pi_2 - \pi^2}{\pi - \pi^2}.$$

Introduciamo ora un particolare modello di tipo mistura: il modello mistura bernoulliano. Il modello viene presentato in termini generali, inizialmente senza l'ipotesi di scambiabilità, per comprendere l'esatto ruolo e le semplificazioni che l'ipotesi comporta.

4.1.1 Modello mistura bernoulliano

Considerato un numero naturale $l < m$, sia $\Psi = (\Psi_1, \dots, \Psi_l)^\top$ un vettore aleatorio l -dimensionale. Il vettore aleatorio \mathbf{S} è distribuito come un modello mistura bernoulliano con fattore Ψ se esistono m funzioni $f_i : \mathbb{R}^p \rightarrow [0, 1]$ tali che, le marginali di \mathbf{S} sono, condizionatamente a Ψ , variabili aleatorie indipendenti, che soddisfano:

$$P(S_i = 1 \mid \Psi = \boldsymbol{\psi}) = f_i(\boldsymbol{\psi}).$$

Pertanto, assunto $\mathbf{s} = (s_1, \dots, s_m)^\top \in \{0, 1\}^m$, la legge congiunta di \mathbf{S} condizionata a Ψ si fattorizza nel seguente modo:

$$P(\mathbf{S} = \mathbf{s} \mid \Psi = \boldsymbol{\psi}) = \prod_{i=1}^m f_i(\boldsymbol{\psi})^{s_i} (1 - f_i(\boldsymbol{\psi}))^{1-s_i}.$$

Si osservi che, se le funzioni f_i coincidono, $f_i = f_1$ per ogni i , allora il modello soddisfa l'ipotesi di scambiabilità. Nel seguito si considererà un modello di questo tipo. In molti casi operativi non sono disponibili utili informazioni per la calibrazione di un modello con più fattori. Si considera pertanto un modello con $l = 1$. Indichiamo, per semplicità, con $H = f_1(\Psi)$ il fattore a cui ci si condiziona e con F_H la sua funzione di ripartizione. Si osservi che, per la proprietà iterativa della speranza matematica condizionata, condizionare a Ψ o ad $H = f_1(\Psi)$ è equivalente. Subordinatamente all'evento $(H = h)$, il numero aleatorio di default M ha quindi una distribuzione binomiale con parametri h e m :

$$P(M = k \mid H = h) = \binom{m}{k} h^k (1 - h)^{m-k}.$$

Per ottenere la distribuzione non condizionata di M è necessario integrare rispetto alla variabile condizionante, e dunque:

$$P(M = k) = \int_0^1 \binom{m}{k} h^k (1 - h)^{m-k} dF_H(h). \quad (4.2)$$

Alla luce delle ipotesi fatte, si possono calcolare le probabilità di default non condizionate π e π_k come momenti della distribuzione di H . Infatti, in primo luogo, si ha¹:

$$\pi = E_P(S_i) = E_P(E_P(S_i|H)) = E_P(H),$$

ed in secondo luogo risulta, per $k = 2, \dots, m$:

$$\begin{aligned} \pi_k &= P(S_i = 1, \dots, S_k = 1) = E_P(S_1, \dots, S_k) = E_P(E_P(S_1, \dots, S_k|H)) = \\ &= E_P(E_P(S_1|H) \cdots E_P(S_k|H)) = E_P(H^k). \end{aligned}$$

Pertanto, specificata la distribuzione F_H , le probabilità di default (di una o più controparti) non condizionate coincidono con i momenti di H . Osserviamo infine che la covarianza tra S_i e S_j , con $i \neq j$, è data dalla varianza di H

¹La seconda uguaglianza è un'applicazione della proprietà iterativa del valore atteso condizionato.

ed è dunque sempre non negativa:

$$\text{Cov}_P(S_i, S_j) = \pi_2 - \pi^2 = \text{Var}_P(H) \geq 0.$$

Per specificare completamente il modello occorre ancora assegnare distribuzione misturante F_H . Considereremo nella prossima sezione il caso in cui H è distribuita come una Beta. Il modello che si ottiene è denominato Beta-Binomiale, e rappresenta un'approssimazione del famoso modello CreditRisk⁺ introdotto da Credit Suisse Financial Products nel 1997, per dettagli si veda McNeil et al. (2015).

4.1.2 Modello Beta-Binomiale

Nel modello Beta-Binomiale si assume che $H \sim \text{Beta}(a, b)$; si rammenti che la funzione di densità della distribuzione Beta è siffatta:

$$g(h) = \frac{1}{\beta(a, b)} h^{a-1} (1-h)^{b-1}, \quad a, b > 0, \quad 0 < h < 1,$$

ove $\beta(a, b)$ denota la funzione beta di Eulero². Con questa particolare ipotesi è possibile ricavare un'espressione esplicita per tutte le quantità introdotte nella precedente sezione. Partiamo dalle probabilità di default non condizionate. Si ha:

$$\begin{aligned} \pi_k &= \mathbb{E}_P(H^k) = \frac{1}{\beta(a, b)} \int_0^1 h^k h^{a-1} (1-h)^{b-1} dh \\ &= \frac{1}{\beta(a, b)} \int_0^1 h^{k+a-1} (1-h)^{b-1} dh = \frac{\beta(a+k, b)}{\beta(a, b)} \quad 1 \leq k \leq m. \end{aligned}$$

Per la funzione di beta di Eulero, si può dimostrare, a partire dalla sua rappresentazione tramite funzione Gamma, che sussiste la seguente relazione ricorsiva, valida per $a > 0$, $b > 0$:

$$\beta(a+1, b) = \frac{a}{a+b} \beta(a, b).$$

² $\beta(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt$, con $x > 0, y > 0$. Essa è legata alla nota funzione Gamma dalla seguente relazione: $\beta(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$.

Pertanto, si ottiene la seguente espressione esplicita per la probabilità di default non condizionata:

$$\pi_k = \prod_{j=0}^{k-1} \frac{a+j}{a+b+j}, \quad 1 \leq k \leq m. \quad (4.3)$$

In particolare, si ha $\pi_1 = \pi = \frac{a}{a+b}$ e $\pi_2 = \pi \frac{a+1}{a+b+1}$. Come osservato nella sezione precedente, a partire da π e π_2 , è possibile individuare anche la default correlation, che riesce pari a:

$$\rho_S = \frac{\pi_2 - \pi^2}{\pi - \pi^2} = \frac{1}{a+b+1}. \quad (4.4)$$

Per la variabile aleatoria M , che conta il numero aleatorio di default tra le m controparti, si dice, in questo caso, che ha una distribuzione Beta-Binomiale con parametri m, a e b ; utilizzeremo nel seguito la notazione compatta: $M \sim \text{BBN}(m, a, b)$. È possibile, a partire dalla (4.2), pervenire ad un'espressione esplicita per la funzione di probabilità di M :

$$\begin{aligned} P(M = k) &= \int_0^1 \binom{m}{k} h^k (1-h)^{m-k} dF_H(h) \\ &= \binom{m}{k} \frac{1}{\beta(a, b)} \int_0^1 h^k (1-h)^{m-k} h^{a-1} (1-h)^{b-1} dh \\ &= \binom{m}{k} \frac{1}{\beta(a, b)} \int_0^1 h^{k+a-1} (1-h)^{m-k+b-1} dh \\ &= \binom{m}{k} \frac{\beta(a+k, b+m-k)}{\beta(a, b)}, \quad 0 \leq k \leq m. \end{aligned}$$

Si osservi che impiegando congiuntamente le relazioni (4.3) e (4.4), è possibile individuare univocamente a partire da π e ρ_S , i parametri a e b della distribuzione beta.

4.2 Struttura generale del modello stocastico

Consideriamo un istituto bancario (si potrebbe considerare alternativamente anche un'impresa di assicurazioni) avente un portafoglio di m posizioni di credito nei confronti di altrettante controparti finanziarie su un orizzonte temporale di un anno. Immaginiamo che la banca, tramite un sistema di rating interno o esterno, segmenti le m posizioni in tre macroclassi (tranche) in modo che all'interno di queste vi sia un certo grado di omogeneità. Indichiamo le tre tranche rispettivamente con le lettere maiuscole: A, B e C (nel seguito, in taluni casi, per snellire la notazione impiegheremo la lettera minuscola z per indicare le generica tranche). Siano m_A, m_B e m_C le numerosità delle controparti finanziarie nelle tre tranche suddette, dunque $m = m_A + m_B + m_C$. Interpretativamente si può immaginare che nella tranche A vi siano controparti con una probabilità di default molto bassa (rating alto). La banca, nei confronti di queste, è disposta ad esporsi su importi maggiori. Nella tranche B vi sono controparti con con probabilità di default medie; necessariamente il portafoglio della banca contiene molte controparti in questa tranche in quanto sono le più comuni sul mercato. In ultimo, nella tranche C, le controparti hanno una probabilità di default molto più alta, la banca è esposta su importi più ridotti. Indichiamo con $S_{i,z}$, ove $z = A, B, C$ e $i = 1, \dots, m_z$, la variabile aleatoria dicotomica che assume valore 1 qualora l' i -esima controparte della tranche z vada incontro a default. Siano poi M_A, M_B e M_C rispettivamente i numeri aleatori di default nelle tre macroclassi, per cui riesce $M_z = \sum_{i=1}^{m_z} S_{i,z}$, $z = A, B, C$. Per ciascuna di queste variabili aleatorie ipotizziamo una distribuzione Beta-Binomiale (quindi per ogni tranche immaginiamo di applicare il modello della sezione 4.1.2) con specifici parametri:

$$M_z \sim \text{BBN}(m_z, a_z, b_z).$$

Indichiamo con H_z , la corrispondente variabile aleatoria misturante, che è distribuita, per $z = A, B$ e C , come una beta. Ipotizziamo inoltre che vi sia una particolare struttura di dipendenza stocastica che descrive il vettore aleatorio $\mathbf{H} = (H_A, H_B, H_C)$, specificata tramite una copula (si discuterà di

questo punto in dettaglio nel seguito). Si osservi che, come si è già detto, dalle equazioni del modello beta binomiale relative alla probabilità di default non condizionata (4.3) e alla default correlation (4.4), si possono ricavare, a partire dai parametri a e b , univocamente π e ρ_S (e viceversa). Pertanto scriveremo $M_z \sim \text{BBN}(m_z, \pi_z, \rho_z) \equiv \text{BBN}(m_z, a_z, b_z)$.

La perdita globale L è, in questo caso, individuata dalle somma delle tre perdite globali per le macrosezioni: L_A, L_B e L_C . Ciascuna di queste perdite è così definita:

$$L_z = \sum_{i=1}^{m_z} \tilde{\delta}_{i,z} e_{i,z} S_{i,z}, \quad z = A, B, C,$$

ove indichiamo, con riferimento all' i -esima controparte della tranche z , con $\tilde{\delta}_{i,z}$ la variabile aleatoria (loss given default) che rappresenta la percentuale dell'esposizione monetaria che la banca perde conseguentemente al default di questa (vale 0 se il default non avviene) e con $e_{i,z}$ l'esposizione monetaria. Nel seguito, per motivi di semplicità, l'esposizione monetaria dipenderà esclusivamente da z e non da i , pertanto si immagina che e_z rappresenti l'esposizione monetaria media all'interno della tranche z . Per modellare $\tilde{\delta}_{i,z}$ si faranno opportune ipotesi che riflettono la qualità del credito delle controparti nelle tranche.

Con riferimento al framework per l'analisi di sensitività, considereremo, in questo ambito, come variabile aleatoria output di riferimento la perdita globale L (quindi L corrisponde alla variabile aleatoria Y del Capitolo 2). Il vettore dei fattori di rischio \mathbf{X} contiene le variabili aleatorie M_z, H_z e $\tilde{\delta}_{i,z}$. La funzione di aggregazione g è individuata dalle relazioni suddette.

Si osservi che il modello appena presentato comporta una dipendenza stocastica su due livelli. Il primo livello è individuato dalla dipendenza all'interno delle tranche (within); questa è implicata fundamentalmente dal particolare modello mistura bernoulliano che stiamo considerando e, come visto, nel caso scambiabile, è sintetizzata in ogni tranche da un unico coefficiente di correlazione ρ . Vi è inoltre un altro livello di dipendenza, quello tra le tranche (between), che, come si menzionava, sarà modellato tramite una copula. La copula descriverà la struttura di dipendenza tra le variabili aleatorie H_z del

modello, inducendo pertanto una particolare struttura di dipendenza anche sui numeri aleatori di default. Si osservi che, in termini di rischiosità, la differenza tra le variabili aleatorie H_z e le M_z è data esclusivamente dalla struttura binomiale. Sarà interessante nel seguito comprendere come si comporta la struttura di dipendenza complessiva del modello conseguentemente allo stress. L'obiettivo generale dell'analisi è comunque comprendere come, a partire dallo stress, sono impattate le distribuzioni dei fattori di rischio in input, ricavandone un ranking ed individuando pertanto la tranche più influente nell'analisi. Procediamo ora esponendo il modello di riferimento; a partire da questo si attueranno successivamente quattro analisi finalizzate ad indagare alcuni aspetti del modello.

4.3 Modello di base

Introduciamo in questa sezione il modello base per la descrizione del rischio di credito della banca presa in considerazione. In questo modello consideriamo le variabili aleatorie H_z e M_z , $z = A, B, C$. Dunque si ha:

$$L = g(\mathbf{X}) = g(H_A, H_B, H_C, M_A, M_B, M_C).$$

Le variabili aleatorie $\tilde{\delta}_{i,z}$ sono in questo caso da considerarsi degeneri e non dipendenti da i , assumono pertanto, per ogni tranche, un'unica determinazione che indichiamo con δ_z , $z = A, B, C$. Sia $m = 10\,000$ la numerosità delle controparti finanziarie della banca. La seguente tabella contiene le caratteristiche delle tranche.

	m_z	π_z	ρ_z	δ_z	e_z	$m_z e_z$
A	2 500	0.04%	0.04%	25%	80	200 000
B	5 000	0.97%	0.44%	37.5%	25	125 000
C	2 500	5.03%	1.328%	50%	10	25 000

Tabella 4.1: I dati per il modello base. Esposizioni in migliaia di euro.

Dalla tabella si osservi in primo luogo che il numero di controparti della tranche B è il doppio rispetto le altre due tranche. I dati concernenti probabilità

di default e correlazioni sono stati presi da McNeil et al. (2015) e sono calibrati tramite dati storici relativi ai default in Standard & Poor's. A riflettere il peggioramento della qualità del credito, sia la probabilità di default che la correlazione sono crescenti passando da A a C. Inoltre anche i parametri δ_z , che rappresentano la percentuale dell'esposizione che viene perduta conseguentemente al default, hanno il medesimo comportamento. Le quantità e_z (in migliaia) rappresentano, per ogni tranche, la media delle esposizioni delle controparti. La banca ha quindi esposizioni maggiori su controparti con qualità creditizia maggiore. L'ultima colonna mostra l'esposizione complessiva, in migliaia di euro, per le tre tranche. L'ultimo essenziale ingrediente per completare il modello per il rischio insolvenza impiegato successivamente nell'analisi di sensitività è la specificazione della copula C per determinare la struttura di dipendenza tra le tre tranche. A tal fine, consideriamo ini-

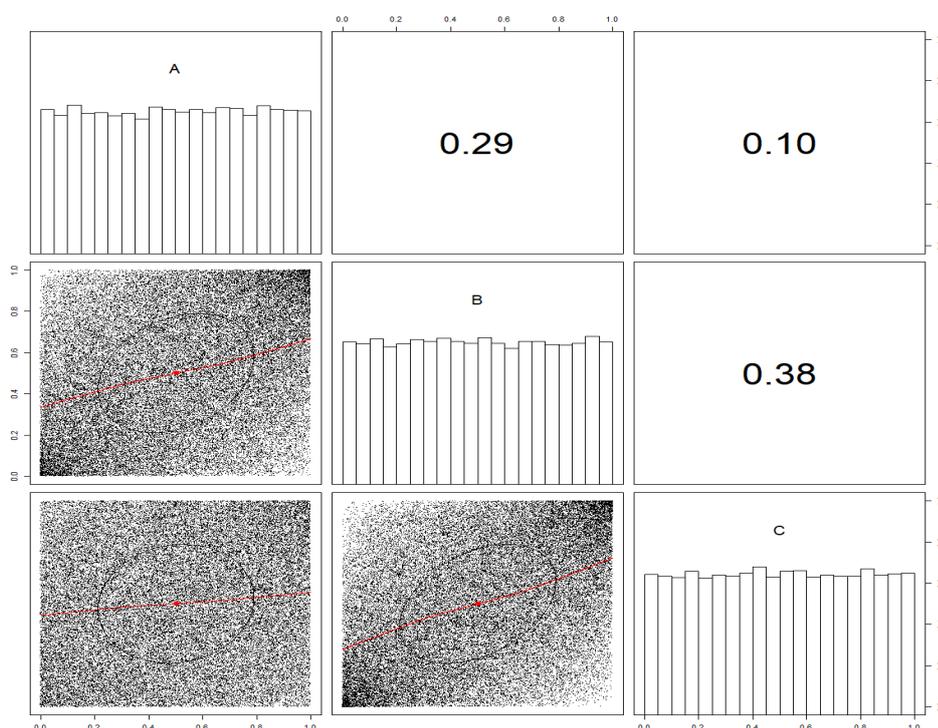


Figura 4.2: In Figura sono rappresentate le marginali della copula gaussiana impiegata (ci si riferisce ai primi 50 000 valori). In alto a destra, la correlazione di Spearman per le tre coppie.

zialmente una copula gaussiana tridimensionale, che agisce sulle tre variabili aleatorie H_z del modello (dalle quali tramite una binomiale, sono simulati, per ogni tranche, i numeri di default). La copula gaussiana C descrive dunque la struttura di dipendenza del vettore aleatorio $\mathbf{H} = (H_A, H_B, H_C)$. Le correlazioni tra le tranche sono: $\rho_{AB} = 0.3$, $\rho_{AC} = 0.1$, $\rho_{BC} = 0.4$. Impiegare una copula per la descrizione della dipendenza tra le tre distribuzioni beta induce una struttura di correlazione anche tra le M_z , che analizzeremo in seguito. Fissate a monte un numero di simulazioni pari a 500 000, la Figura 4.2 delinea la struttura simulata delle marginali della copula e la loro correlazione per i primi 50 000 valori. A partire dalla copula, si è simulato il vettore aleatorio \mathbf{H} , che ha appunto come struttura di dipendenza la copula gaussiana. Nella Figura 4.3, si sono rappresentati i primi 50 000 valori per le distribuzioni marginali di \mathbf{H} . Dalla Figura si può osservare che i coefficienti

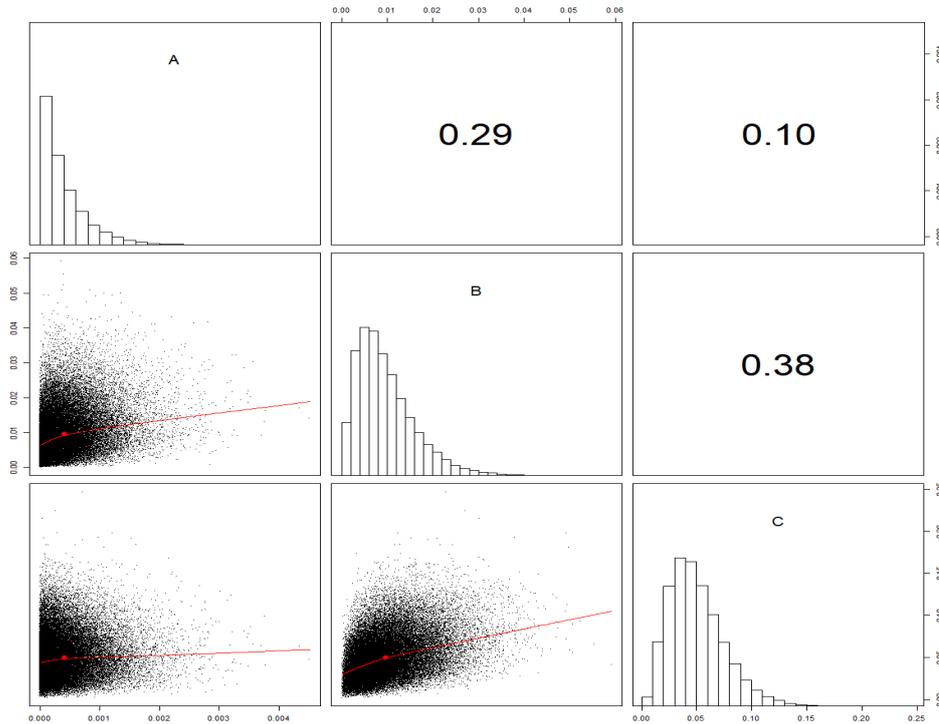


Figura 4.3: In Figura sono rappresentate le distribuzioni marginali (beta) simulate del vettore aleatorio \mathbf{H} (ci si riferisce ai primi 50 000 valori). In alto a destra, la correlazione di Spearman per le tre coppie.

di correlazione di Spearman sono i medesimi della copula, e si possono inoltre analizzare le tre distribuzioni beta impiegate per rappresentare il rischio sistematico di default nelle tranches. Si rammenti che, come si è già menzionato, i parametri delle tre distribuzioni beta sono univocamente determinati a partire dai dati sulle probabilità di default π_z e delle default correlation ρ_z . Prima di procedere con l'analisi dell'impatto dello stress, analizziamo brevemente le distribuzioni simulate per le variabili aleatorie M_z , le marginali del vettore aleatorio \mathbf{M} . In generale esse sono i fattori di rischio interpretativamente più interessanti. Procediamo inizialmente con una tabella riassuntiva. Come si

	Min	I Quartile	Mediana	Media	III Quartile	Max
M_A	0.0	0.0	1.0	1.0	2.0	21.0
M_B	0.0	24.0	41.0	48.5	65.0	365.0
M_C	0.0	79.0	116.0	125.8	162.0	612.0

Tabella 4.2: Alcune statistiche di riferimento per il confronto tra le distribuzioni simulate delle variabili aleatorie M_A , M_B e M_C .

può osservare, le tre distribuzioni simulate sono molto diverse. In particolare, M_A assume valori estremamente più piccoli rispetto M_B e M_C : su 2500 controparti in media solo una va incontro a default. Osserviamo inoltre che

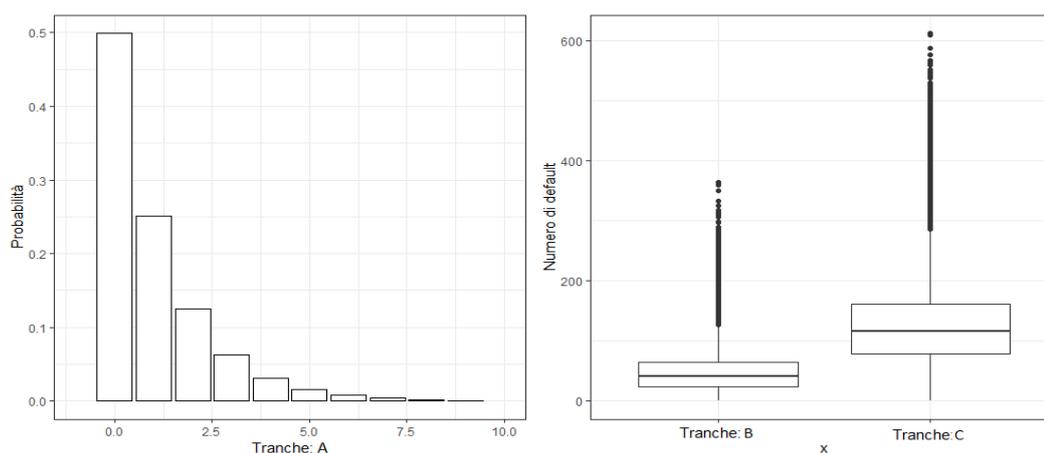


Figura 4.4: A sinistra: grafico a barre ottenuto dai valori simulati di M_A (si noti che è l'asse delle ascisse è troncato al valore 10). A destra: boxplot dei valori simulati di M_B e M_C .

tutte e tre le distribuzioni presentano code molto pesanti (il valore massimo è molto maggiore del terzo quartile). Per visualizzare graficamente al meglio l'andamento delle tre distribuzioni presentiamone un grafico. Sebbene siano tutte e tre variabile aleatorie a supporto discreto, utilizziamo per M_B e M_C , al posto del grafico a barre, i boxplot, in quanto il loro range di variazione è troppo ampio (coda superiore pesante). Si nota dalla Figura 4.4 che, in tutti e tre i casi, ma con maggiore enfasi in M_B e M_C , le code delle distribuzioni sono pesanti. Ciò in generale non stupisce, alla luce del meccanismo aleatorio che le genera.

4.3.1 Analisi di sensitività - Modello base

Introduciamo per l'analisi di sensitività uno stress che, in linea con il Teorema 3 del Capitolo 2, coinvolga Value at Risk ed Expected Shortfall.

$$\begin{aligned} stress_{VaR} &= 10\%, & q &= VaR_\alpha^Q(L) = VaR_\alpha^P(L) (1 + stress_{VaR}), \\ stress_{ES} &= 15\%, & s &= ES_\alpha^Q(L) = ES_\alpha^P(L) (1 + stress_{ES}), \end{aligned} \quad (4.5)$$

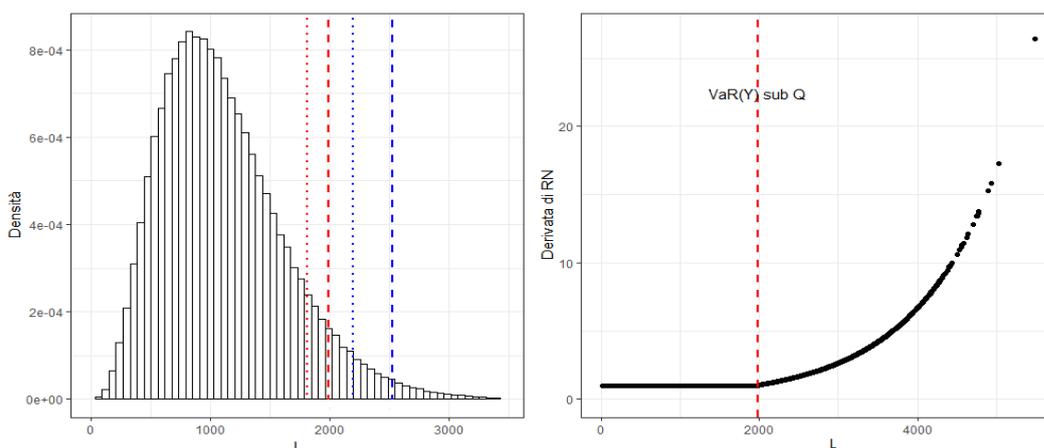


Figura 4.5: A sinistra: istogramma di L ottenuto dai valori simulati della perdita globale. La linea rossa a punti coincide con il livello del $VaR_\alpha^Q(L)$ mentre quella a tratteggio con invece $q = VaR_\alpha^P(L)$. La linea blu a punti coincide con il livello del $ES_\alpha^Q(L)$ mentre quella a tratteggio con invece $s = ES_\alpha^P(L)$. A destra: derivata di RN ξ soluzione del Problema (2.9) sotto al vincolo espresso da (4.5).

con $\alpha = 0.9$. La struttura dello stress implica pertanto che stiamo considerando una situazione avversa alla banca. Lo scenario stressato è uno scenario che impatta negativamente sulla gestione e ci aspettiamo dunque che, sotto la misura stressata, il numero dei default nelle varie tranche aumenti. L'obiettivo è comprendere quale delle tre tranche sia più impattata dallo stress, e come e di quanto cambiano le distribuzioni dei fattori di rischio sotto la misura di probabilità stressata Q . Sarà interessante anche studiare come viene impattata la loro struttura di correlazione. Conseguentemente all'andamento delle tre variabili aleatorie che descrivono il numero dei default nelle varie tranche, anche la distribuzione di L (perdita totale) mostra asimmetria (Figura 4.5 a sinistra). La derivata di RN simulata soluzione del Problema

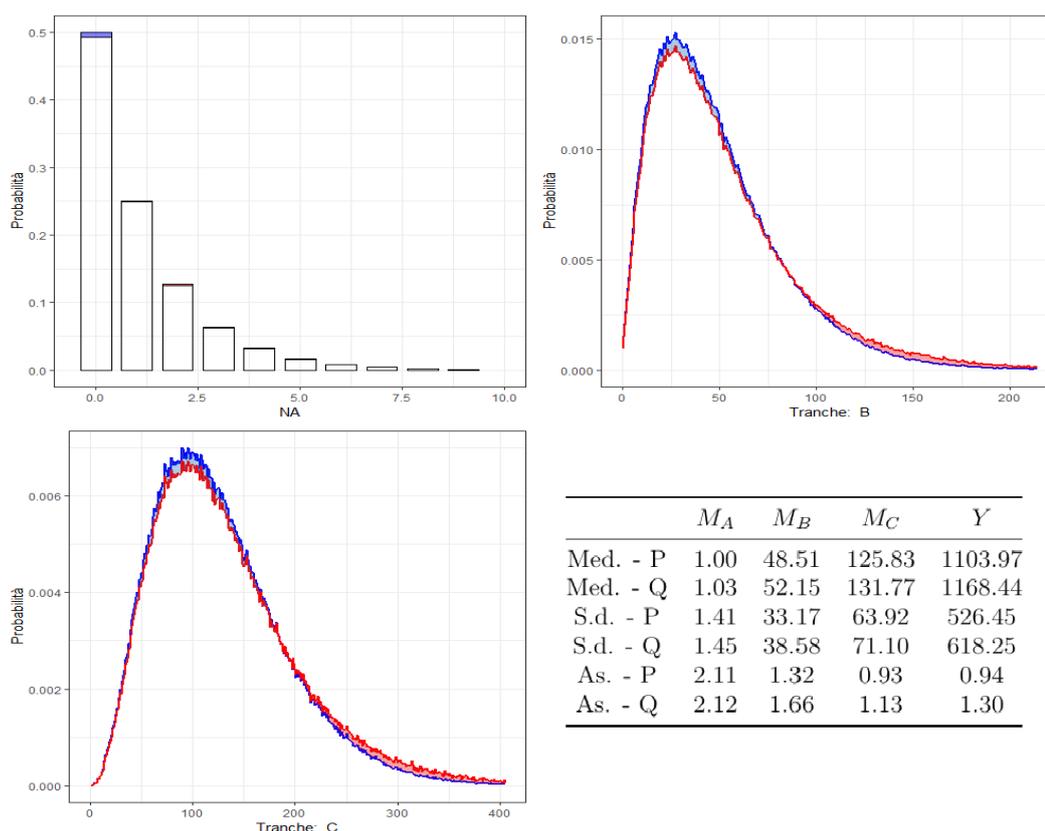


Figura 4.6: In alto a sinistra: confronto della distribuzione di M_A sotto le due misure. In alto a destra, di M_B ; in basso a sinistra, di M_C . In tutti i confronti il colore blu rappresenta la misura baseline mentre il rosso quella stressata. In basso a destra: tabella riassuntiva. Med = Media, S.d = deviazione standard, As = asimmetria.

(2.9) sotto ai vincoli imposti dalla (4.5) ha un andamento esponenziale (Figura 4.5 a destra). Analizziamo ora l'impatto dello stress sulle distribuzioni dei fattori di rischio, incominciando dagli M_z . Si veda a tal fine la Figura 4.6. Il grafico a barre, in alto a sinistra, confronta la distribuzione di M_A sotto la misura baseline (in blu) e quella stressata (in rosso). Sebbene si noti che la probabilità di non avere default sotto la misura baseline sia maggiore rispetto a quella sotto la misura stressata, la distribuzione non sembra essere impattata in modo sensibile. Considerazioni diverse valgono per le tranche B e C. In questo caso si noti che si è preferito utilizzare una rappresentazione alternativa per cogliere meglio il fenomeno vista l'estensione del supporto. In

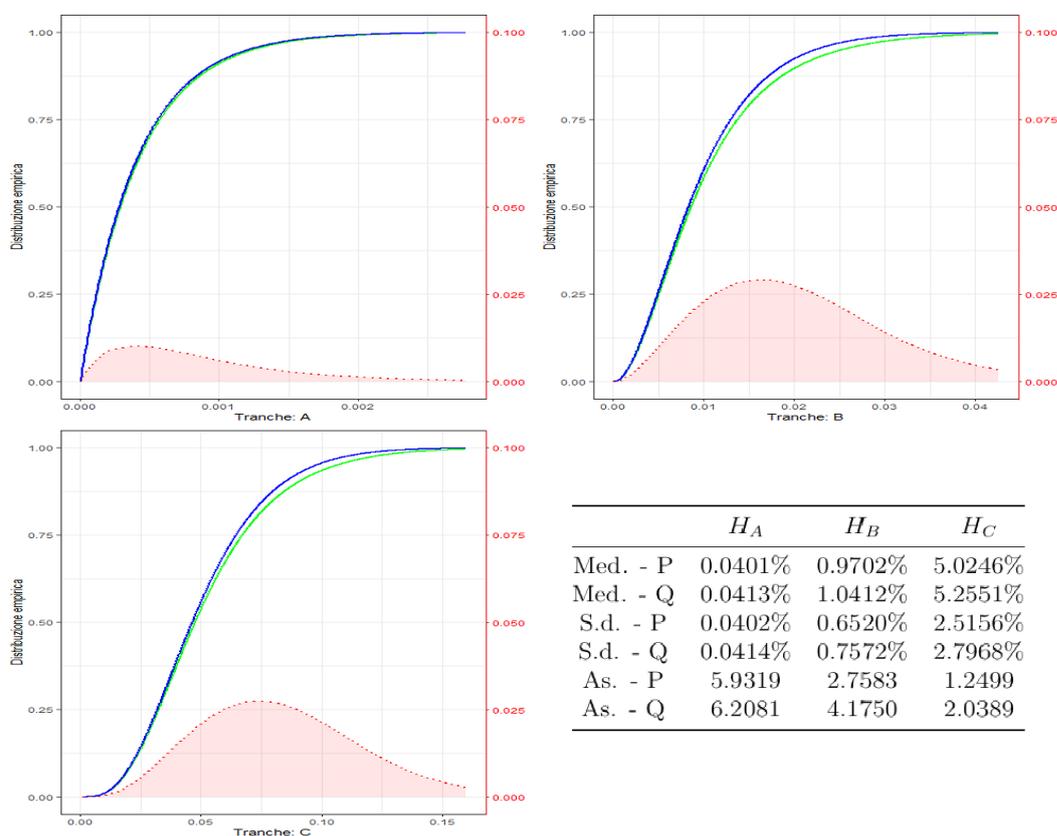


Figura 4.7: Confronto tra le funzioni di ripartizione di H_A (in alto a sinistra), H_B (in alto a destra) e H_C (in basso a sinistra) sotto le due misure. In ogni grafico la linea blu rappresenta la funzione di ripartizione sotto la misura baseline mentre la verde quella sotto la misura stressata. L'area rossa ombreggiata individua la differenza tra le due ($P-Q$) riscalata sull'asse delle ordinate a destra del grafico. In basso a destra: tabella riassuntiva. Med = Media, S.d = deviazione standard, As = asimmetria.

entrambi i casi, lo stress colpisce la distribuzione aumentando la probabilità di un elevato numero di default. Non appare chiaro quale dei due fattori di rischio sia più impattato. In basso a destra, nella tabella riassuntiva, si può apprezzare il comportamento di alcune statistiche. La media sotto la misura di probabilità stressata è sempre maggiore rispetto a quella sotto la misura baseline. Anche la deviazione standard e la simmetria aumentano. Consideriamo ora i fattori di rischio H_z , $z = A, B, C$ e procediamo con la medesima analisi, si veda a tal fine la Figura 4.7. Il risultato è in linea con quanto già analizzato per M_z . La variabile aleatoria H_A è poco impattata dallo stress, mentre le variabili H_B e H_C lo sono molto di più. Si rammenti che il valore atteso della variabile aleatoria H_z individua la probabilità di default di una singola controparte della tranche z . Dalla tabella in Figura 4.7 si può analizzare dunque come la probabilità di default cambi nel passaggio dalla misura baseline a quella stressata.

Per quantificare l'impatto, e dunque comprendere quale delle tre tranche sia più importante per la rappresentazione del rischio di default del portafoglio, possiamo utilizzare la reverse sensitivity measure introdotta nel Capitolo 2. Con riferimento a tutti i fattori di rischio, si ha in questo caso:

$$\begin{aligned}\Gamma_{M_A} &= \mathcal{S}(M_A, \zeta) = 0.15, & \Gamma_{H_A} &= \mathcal{S}(H_A, \zeta) = 0.19 \\ \Gamma_{M_B} &= \mathcal{S}(M_B, \zeta) = 0.82, & \Gamma_{H_B} &= \mathcal{S}(H_B, \zeta) = 0.81 \\ \Gamma_{M_C} &= \mathcal{S}(M_C, \zeta) = 0.76, & \Gamma_{H_C} &= \mathcal{S}(H_C, \zeta) = 0.75.\end{aligned}$$

Pertanto, la tranche più impattata è la B, segue a poca distanza la C, mentre, come si era osservato dai grafici, la tranche A è poco impattata dallo stress. Si osservi che la sensitività delle variabili aleatorie H_z e M_z è molto simile. Ciò in generale non stupisce in quanto H_z ha un ruolo fondamentale nella definizione di M_z . Questo risultato fa comprendere che da un punto di vista della rischiosità complessiva la struttura binomiale, ovvero il rischio di processo, ha poco impatto, mentre la parte di rischio sistematico, spiegata da H_z , riveste notevole importanza. Nel seguito, per individuare il ranking tra i fattori di rischio e conseguentemente la tranche più importante per l'analisi,

possiamo analizzare direttamente M_z tralasciando H_z .

Studiamo, in ultima istanza, il comportamento del coefficiente di correlazione di Spearman per comprendere come la struttura di dipendenza (between) del modello sia stata colpita dallo stress (4.5). Analizziamo a tal fine al posto della copula su \mathbf{H} , la struttura di dipendenza del vettore aleatorio \mathbf{M} , in quanto, come si è visto esso è il fattore di rischio più interessante (e interpretativo). La seguente tabella mostra che, sotto la misura stressata Q la correlazione (di Spearman) è più alta tra ogni coppia di tranche. Non deve

sub P	M_A	M_B	M_C	sub Q	M_A	M_B	M_C
M_A	1.00	0.18	0.06	M_A	1.00	0.20	0.08
M_B	0.18	1.00	0.37	M_B	0.20	1.00	0.42
M_C	0.06	0.37	1.00	M_C	0.08	0.42	1.00

Tabella 4.3: A sinistra: matrice di correlazione di Spearman per \mathbf{M} sotto la misura baseline P . A destra: medesima matrice sotto la misura stressata Q .

stupire che sotto la misura baseline i coefficienti di correlazione non sono uguali a quelli inseriti in input sulla copula tridimensionale di \mathbf{H} ; è infatti opportuno osservare che a partire dalle beta vengono simulati i default tramite la distribuzione binomiale e ciò ovviamente impatta sulla correlazione finale tra le tranche.

4.4 Modelli alternativi

Studiamo in questa sezione alcuni modelli alternativi per comprendere gli effetti delle varie ipotesi e dei parametri del modello base e metterli tra loro a confronto. Consideriamo in particolare quattro studi distinti, ognuno mirato a comprendere una particolare caratteristica del modello base. Ad ognuno di questi dedichiamo una sottosezione.

4.4.1 L'effetto dell'esposizione

Per comprendere il ruolo dell'esposizione nel modello procediamo introducendo due portafogli alternativi nei quali, mantenendo costanti gli importi monetari medi e_z per tranches, si considerano diverse ipotesi per le relative numerosità m_z . In ognuno dei casi, per valutare l'effetto delle nuove ipotesi, si procederà calcolando la reverse sensitivity measure di M_A, M_B e M_C per disporre di una misura sintetica per il confronto. La Tabella 4.4 descrive le caratteristiche dei due portafogli impiegati nell'analisi.

	m_z	e_z	$m_z e_z$		m_z	e_z	$m_z e_z$
A	3 333	80	266 640	A	7 000	80	560 000
B	3 333	25	83 325	B	2 000	25	50 000
C	3 333	10	33 330	C	1 000	10	10 000

Tabella 4.4: A sinistra: portafoglio 1. A destra: portafoglio 2

Si osservino, in generale, le differenze tra i due portafogli e quello utilizzato nel modello base (Tabella 4.1). Nel portafoglio 1 si è fissata uguale numerosità nelle tre tranches, dunque: $m_A = m_B = m_C = 3333$. Nell'ultima colonna, che indica l'esposizione complessiva $m_z e_z$ per la tranche z , si può osservare che in questo caso nella tranche A si ha più del doppio di esposizione complessiva rispetto le altre due tranches aggregate (nel caso base è 33% in più). A partire da questo portafoglio, preservando le altre ipotesi, procediamo effettuando il medesimo stress (4.5) e calcoliamo la reverse sensitivity measure per le tre macrosezioni. Si ha:

$$\Gamma_{M_A} = \mathcal{S}(M_A, \zeta) = 0.13,$$

$$\Gamma_{M_B} = \mathcal{S}(M_B, \zeta) = 0.63,$$

$$\Gamma_{M_C} = \mathcal{S}(M_C, \zeta) = 0.90.$$

L'effetto della diversa ipotesi di numerosità ha impattato sul ranking dei fattori di rischio. Per quanto concerne il fattore M_A esso rimane sempre quello di nettamente minor rilevanza. Cambia il ruolo di M_B e M_C , in particolare, in questo caso, il fattore di rischio più rilevante è M_C e non più

M_B . Ciò fa in generale comprendere che vi è, come è naturale pensare, un effetto importante legato all'esposizione complessiva per le classi. Seppur con questo portafoglio la tranche A aveva complessivamente un'esposizione molto più alta delle altre due, il rischio che essa comporta continua ad essere poco significativo. Nel portafoglio 2 si estremizza questo aspetto andando a considerare un portafoglio nel quale la tranche A ha dieci volte (circa) l'esposizione delle altre due sommate. L'obiettivo è capire se, con questa ipotesi estrema, l'effetto dell'esposizione nella tranche A è abbastanza forte da far guadagnare a questa la prima posizione nel ranking tra le tranche. Calcoliamo dunque, mantenendo inalterate le altre ipotesi e utilizzando la stessa struttura per lo stress, la reverse sensitivity measure impiegando il portafoglio 2. Il risultato è ora il seguente:

$$\Gamma_{M_A} = \mathcal{S}(M_A, \zeta) = 0.36,$$

$$\Gamma_{M_B} = \mathcal{S}(M_B, \zeta) = 0.80,$$

$$\Gamma_{M_C} = \mathcal{S}(M_C, \zeta) = 0.69.$$

Nonostante la misura di sensitività per la variabile aleatoria M_A abbia ricevuto un incremento considerevole rispetto al caso base (triplicando circa), essa rimane comunque la meno importante per lo stress. Γ_{M_A} è in questo caso circa la metà di Γ_{M_C} . La probabilità di default (e la correlazione) per la tranche A sono così basse rispetto le altre due che, anche con un'esposizione estrema, essa non riesce ad acquisire il ruolo di tranche con rischiosità predominante nel modello.

4.4.2 La dipendenza within

Per esaminare il ruolo della dipendenza within nel modello base consideriamo tre situazioni alternative. Rammentiamo che, come ampiamente discusso in precedenza, la struttura di dipendenza within nel modello Beta-Binomiale è completamente individuata dalla default correlation. Nel nostro modello, che considera tre distinte tranche, essa è pertanto assegnata tramite la terna di parametri ρ_A, ρ_B e ρ_C . La prima struttura alternativa che si può conside-

rare presuppone l'indipendenza stocastica tra le variabili aleatorie indicatrici che descrivono il default all'interno di ciascuna tranche. Nella Figura 4.1 si era osservato il comportamento della perdita totale di portafoglio qualora si considerasse l'ipotesi di indipendenza stocastica tra gli indicatori di default. Osserviamo che, in primo luogo, l'ipotesi di indipendenza stocastica tra i default è in generale priva di significato economico ed è molto difficile che possa essere accolta. In secondo luogo, con questa ipotesi, non è possibile affrontare l'analisi di sensitività con la struttura di stress considerata per il modello base. Ciò è dovuto alla quasi assenza di dispersione di L ; anche incrementando di solo 1% il Value at Risk sotto la misura baseline si ottiene un valore maggiore del massimo valore simulato per L . Non è possibile pertanto applicare il Teorema 2.2.3.

Gli altri due casi sono individuati dai seguenti set di parametri, nei quali, mantenendo inalterata la probabilità di default, si considerano diverse default correlation:

	π_z	ρ_z		π_z	ρ_z
A	0.04%	0.5%	A	0.04%	5%
B	0.97%	0.5%	B	0.97%	0.44%
C	5.03%	0.5%	C	5.03%	1.328%

Tabella 4.5: A sinistra: dipendenza within 1. A destra: dipendenza within 2

Si confrontino le due strutture per la dipendenza within con quella del modello base. Nel primo caso si considera una default correlation costante per le tre tranche e pari a 0.5%. Ciò comporta che per la tranche A la default correlation è aumentata di più di un fattore dieci, per la tranche B è rimasta quasi invariata, e infine per la tranche C è circa dimezzata. Nella seconda struttura tutti i parametri sono rimasti costanti fatta eccezione di ρ_A che è stato incrementato drasticamente fino a 5%. L'obiettivo è analizzare se, estremizzando il valore della default correlation nella tranche A , essa possa diventare la tranche di maggiore influenza per il modello. Mantenendo inalterati gli altri parametri e impiegando il medesimo stress, calcoliamo la reverse sensitivity measure per le variabili aleatorie M_z nei due casi (Tabella

4.6).

	Dipendenza within 1	Dipendenza within 2
Γ_{M_A}	0.32	0.84
Γ_{M_B}	0.91	0.72
Γ_{M_C}	0.57	0.66

Tabella 4.6: Reverse sensitivity measure di M_A, M_B e M_C con riferimento alle due strutture di dipendenza within.

Il risultato è interessante. Nel primo caso, sebbene il ranking sia rimasto lo stesso, la nuova ipotesi colpisce fortemente le misure di sensitività. Ciò fa comprendere che la dipendenza within ha un ruolo molto importante nel modello (abbiamo già analizzato le gravi conseguenze della sua rimozione). Nel secondo caso, l'incremento della default correlation per la tranche A ha incrementato il valore di Γ_{M_A} al punto da rendere la tranche A la più importante per il modello nel complesso. Si era osservato nella precedente sottosezione che, estremizzare l'effetto dell'esposizione, non permetteva alla variabile aleatoria M_A di assumere il ruolo di predominante. Alla luce di questa osservazione, possiamo concludere che, l'effetto della dipendenza within è più importante dell'effetto dell'esposizione.

4.4.3 La dipendenza between

Analizziamo in questa sottosezione l'impatto, in termini di reverse sensitivity measure, di diverse strutture di dipendenza between. Si rammenti che la dipendenza between è definita dalla copula che descrive la dipendenza stocastica tra le componenti del vettore $\mathbf{H} = (H_A, H_B, H_C)$. Come si è analizzato nel modello base ciò ha un diretto impatto sulla struttura di dipendenza del vettore $\mathbf{M} = (M_A, M_B, M_C)$. Per descrivere la dipendenza stocastica del vettore aleatorio \mathbf{H} consideriamo le seguenti strutture alternative (vedi Appendice A.4.1):

- I. Indipendenza stocastica tra tranche (copula di indipendenza).
- II. Copula gaussiana con parametri: $\rho_{AB} = 0.3$, $\rho_{BC} = 0.4$, $\rho_{AC} = 0.1$ (caso base).

- III. Copula gaussiana con parametri: $\rho_{AB} = \rho_{BC} = \rho_{AC} = 0.5$.
- IV. Copula t-student con 10 gradi di libertà. Coefficienti di correlazione come nel caso base.
- V. Copula t-student con 3 gradi di libertà. Coefficienti di correlazione come nel caso base.
- VI. Copula Gumbel con parametro $\theta = 1.5$.
- VII. Copula Gumbel con parametro $\theta = 10$.

Per le sette strutture di dipendenza proposte, con riferimento alle variabili aleatorie H_z , $z = A, B$ e C , calcoliamo il coefficiente di dipendenza di coda superiore per ogni possibile coppia. I valori sono presentati nella Tabella 4.7.

λ_u	I	II	III	IV	V	VI	VII
(H_A, H_B)	0	0	0	0.033	0.216	0.413	0.928
(H_A, H_C)	0	0	0	0.012	0.145	0.413	0.928
(H_B, H_C)	0	0	0	0.052	0.261	0.413	0.928

Tabella 4.7: Coefficienti di dipendenza di coda superiore per le sette strutture di dipendenza considerate.

Mantenendo inalterate le restanti ipotesi del modello base e la struttura dello stress (4.5), calcoliamo la reverse sensitivity measure per i tre fattori di rischio M_A, M_B e M_C . Rappresentiamo i risultati dell’analisi nella Tabella 4.8 .

	I	II	III	IV	V	VI	VII
Γ_{M_A}	0.02	0.15	0.35	0.19	0.26	0.58	0.72
Γ_{M_B}	0.70	0.81	0.84	0.84	0.86	0.92	0.99
Γ_{M_C}	0.60	0.76	0.81	0.80	0.84	0.90	0.99

Tabella 4.8: Confronto della reverse sensitivity measure di M_A, M_B e M_C con le sette diverse strutture di dipendenza between.

Procediamo commentando i risultati della tabella a partire dal caso base. Il

modello I suppone che i numeri aleatori di default tra le tre tranches siano tra loro stocasticamente indipendenti. Se si confronta il risultato concernente la reverse sensitivity measure sotto questa ipotesi (I), e quello del caso base (II), si può osservare che per tutte le tranches la sensitività aumenta. La presenza della copula gaussiana del caso base, aumenta dunque la sensitività delle tre tranches, preservandone il ranking. Si osservi che, sotto l'ipotesi di indipendenza, il valore di Γ_{M_A} è prossimo a 0. Procedendo con la struttura di dipendenza III, come ci si poteva aspettare, le tre misure di sensitività aumentano. La più sensibile in questo caso è la tranche A , la cui sensitività raddoppia circa. Ad ogni modo, come analizzato alla fine della prima Appendice, si rammenti che la copula gaussiana implica la presenza di indipendenza asintotica sulle code (vedi Tabella 4.7). Una struttura di copula più appropriata per questo tipo di modelli dovrebbe considerare la possibile presenza di dipendenza di coda. Consideriamo a tal fine la copula t-student e quella Gumbel. La struttura IV suppone che la dipendenza del vettore aleatorio \mathbf{H} sia descritta da una copula t-student tridimensionale con medesimi parametri di correlazione rispetto il caso base e 10 gradi di libertà. In questo caso la sensitività per le tranches B e C è molto simile a quella precedente (III), mentre per la tranche A è drasticamente più bassa. Pertanto, la presenza della dipendenza di coda superiore implicata dalla struttura di copula t-student, aumenta la sensitività (si confronti con II); ma agire semplicemente sulle correlazioni (da II a III) ha comunque un impatto più forte. Diminuendo i gradi di libertà della copula t-student (V), la dipendenza di coda superiore aumenta e anche la sensitività (rispetto IV). La copula Gumbel, dipendendo solo dal parametro θ induce una correlazione costante tra le tranches. Con $\theta = 1.5$ essa è circa pari a 0.5, si confronti dunque la struttura VI alla III. La sensitività è molto più alta, ciò è dovuto alla forte dipendenza di coda superiore di questo tipo di copula (vedi Tabella 4.7). Aumentando il parametro da 1.5 a 10 (VII), la dipendenza di coda aumenta, rendendo M_B e M_C (quasi) comonotone alla derivata di RN soluzione del problema di minimo vincolato.

4.4.4 Loss given default stocastico

Introduciamo in questa sezione una nuova fonte di rischio dovuta all'aleatorietà delle variabili aleatorie $\tilde{\delta}_{i,z}$. Si rammenti che $\tilde{\delta}_{i,z}$ descrive, con riferimento alle tranche $z = A, B$ e C , la percentuale del credito che si perde a seguito del default dell' i -esima controparte. Considereremo nel seguito le variabili indipendenti dagli eventi default. Questa scelta, seppur largamente impiegata nell'industria, è in generale criticabile in quanto la dottrina economica suggerisce anche per queste, come per le probabilità di default, presenza di dipendenza da variabili macroeconomiche (McNeil et al. (2015)). Consideriamo le seguenti quantità aleatorie:

$$\tilde{\delta}_z = \begin{cases} \frac{1}{M_z} \sum_{i=1}^{M_z} \tilde{\delta}_{i,z} & M_z > 0 \\ 0 & M_z = 0 \end{cases} \quad z = A, B, C.$$

Le variabili $\tilde{\delta}_z$ rappresentano, con riferimento alla tranche z , il loss given default medio. A partire da queste, per mantenere la coerenza con le posizioni impiegate nel modello base, si definisce:

$$L_z = e_z M_z \tilde{\delta}_z, \quad z = A, B, C,$$

che rappresenta la perdita aleatoria complessiva per la tranche z . La variabile output del modello L è la somma delle tre L_z . Essa dipende dunque, in questo caso, dal seguente vettore dei fattori di rischio:

$$L = g(\mathbf{X}) = g(H_A, H_B, H_C, M_A, M_B, M_C, \tilde{\delta}_A, \tilde{\delta}_B, \tilde{\delta}_C).$$

Si rammenti che nel modello base le variabili $\tilde{\delta}_{i,z}$ erano considerate degeneri e non dipendenti da i , pari rispettivamente a 0.25, 0.375 e 0.5. Consideriamo, per effettuare un'estensione coerente, per ogni tranche z , una distribuzione beta con media pari al valore puntuale impiegato nel modello base e varianza crescente al peggiorarsi della qualità creditizia. Nella Tabella 4.9 sono descritte le caratteristiche delle distribuzioni impiegate.

	α	β	Media	Varianza
Tranche A	1.2	3.6	0.25	0.032
Tranche B	1.2	2	0.375	0.056
Tranche C	1.2	1.2	0.5	0.0735

Tabella 4.9: Con riferimento alle tre tranches, i parametri, media e varianza delle variabili aleatorie $\tilde{\delta}_{i,z}$.

Mantenuta la medesima struttura per lo stress rispetto al modello base, implementiamo l'analisi di sensitività. Si osservi che, alla luce di quanto descritto, la simulazione del modello avviene in modo annidato. In primo luogo si simulano i valori per le M_z , a partire da questi si simulano le $\tilde{\delta}_{i,z}$ per $z = A, B, C$ e $i = 1, \dots, M_z$ e si costruiscono quindi le $\tilde{\delta}_z$. La Tabella 4.10 confronta alcune caratteristiche della distribuzione di $\tilde{\delta}_z$ sotto le due misure.

	$\tilde{\delta}_A$	$\tilde{\delta}_B$	$\tilde{\delta}_C$
Med. - P	0.24982	0.37500	0.49996
Med. - Q	0.25393	0.37576	0.50003
S.d. - P	0.15450	0.04880	0.02913
S.d. - Q	0.15592	0.04924	0.02878
As. - P	0.85212	0.30747	0.00068
As. - Q	0.77580	0.41416	-0.00369

Tabella 4.10: Confronto distribuzioni di $\tilde{\delta}_z$, $z = A, B$ e C , sotto le due misure di probabilità. Med = Media, S.d = deviazione standard, As = asimmetria.

Il valore atteso per ognuna delle tre tranches è più alto sotto la misura stressata. Lo stress comporta dunque un recupero creditizio minore. Per misurare la sensitività complessiva dei fattori di rischio e ricavarne un ranking, calcoliamo la reverse sensitivity measure, si ha:

$$\begin{aligned} \Gamma_{M_A} &= \mathcal{S}(M_A, \zeta) = 0.142, & \Gamma_{\tilde{\delta}_A} &= \mathcal{S}(\tilde{\delta}_A, \zeta) = 0.113 \\ \Gamma_{M_B} &= \mathcal{S}(M_B, \zeta) = 0.809, & \Gamma_{\tilde{\delta}_B} &= \mathcal{S}(\tilde{\delta}_B, \zeta) = 0.126 \\ \Gamma_{M_C} &= \mathcal{S}(M_C, \zeta) = 0.739, & \Gamma_{\tilde{\delta}_C} &= \mathcal{S}(\tilde{\delta}_C, \zeta) = 0.024. \end{aligned}$$

Nonostante le distribuzioni impiegate in input presentino una marcata dispersione, l'impatto dello stress sulle variabili aleatorie che descrivono il loss given default medio è in generale debole. La presenza dell'aleatorietà delle variabili che descrivono il loss given default riduce di poco la sensitività delle M_z senza comprometterne il ranking.

Conclusione

I modelli interni impiegati in ambito assicurativo e bancario per la valutazione dei requisiti di solvibilità sono complessi al punto da necessitare di opportune metodologie per comprendere in modo approfondito le relazioni tra i fattori di rischio in input e l'output del modello. Il framework per l'analisi di sensitività, introdotto in Pesenti et al. (2019), fornisce un approccio estremamente flessibile per indagare la sensibilità di un modello stocastico complesso. Nella trattazione sono stati selezionati due esempi operativamente significativi nella realtà assicurativa e bancaria sui quali si è condotta la reverse sensitivity analysis. Lo studio ha permesso di identificare efficacemente un ranking per importanza tra i fattori di rischio e dunque comprendere il ruolo delle variabili impiegate nel modello. I modelli utilizzati, seppur complessi, si basano su alcune ipotesi semplificatrici necessarie per supplire alla mancanza di dati. In un approccio più avanzato, ove fossero disponibili dati provenienti dall'industria, si potrebbero impiegare le medesime tecniche per comprendere il ruolo esatto dei fattori di rischio e conseguentemente trarre opportune strategie di risk management.

Tutte le analisi effettuate nel corso della trattazione sono state eseguite impiegando il software statistico R. Rendere accessibile il codice e le funzioni utilizzate, tramite la costruzione di un specifico pacchetto, potrebbe essere un possibile progetto futuro utile e vantaggioso per altri utilizzatori interessati all'implementazione della reverse sensitivity analysis.

Appendice A

Nozioni fondamentali per l'analisi di sensitività

In questa prima Appendice sono proposte definizioni, proprietà e richiami finalizzati ad una comprensione più completa degli argomenti trattati nella tesi. Si tenga in considerazione che l'obiettivo dell'Appendice non è descrivere esaurientemente i vari argomenti ma fornire le basi per un successivo approfondimento. L'Appendice è strutturata in quattro sezioni. La prima sezione concerne le proprietà della funzione inversa generalizzata, uno degli strumenti più importanti del risk management quantitativo. Sono esposti inoltre due risultati fondamentali: il metodo dell'inversione e la trasformazione integrale di probabilità. La seconda introduce un importante risultato della teoria della misura: il Teorema di Radon-Nikodym. Sono successivamente esposte alcune proprietà della relativa derivata. Nella terza sezione si introduce la divergenza di Kullback-Leibler (entropia relativa), che permette di definire una nozione di discrepanza tra due misure di probabilità. L'ultima sezione, la più ampia, è dedicata alla descrizione della struttura di dipendenza tra variabili aleatorie. Sono introdotti, in primo luogo, i concetti di comonotonia e contromonotonia impiegati nel Capitolo 2. Successivamente è definita la nozione di copula, accompagnata dal celeberrimo Teorema di Sklar. Si conclude l'Appendice introducendo e confrontando alcune misure di dipendenza: la correlazione lineare, la correlazione di Spearman e i

coefficienti di dipendenza di coda.

A.1 Funzione inversa generalizzata

La funzione inversa generalizzata o funzione quantile ha un ruolo centrale nel risk management quantitativo, ciò è principalmente dovuto all'importanza che questa ricopre nella definizione di alcune misure di rischio e per il metodo dell'inversione che è il concetto alla base della simulazione stocastica. La funzione inversa generalizzata è definita come il quantile sinistro a livello α ($0 < \alpha < 1$) della distribuzione di X , ovvero:

$$F_X^{-1}(\alpha) \equiv F_X^{-1,-}(\alpha) = \inf\{x \in \mathbb{R} : P(X \leq x) \geq \alpha\}. \quad (\text{A.1})$$

Si può definire anche specularmente il quantile destro a livello α , ($0 < \alpha < 1$):

$$F_X^{-1,+}(\alpha) = \inf\{x \in \mathbb{R} : P(X \leq x) > \alpha\}.$$

Si osservi che vale sempre: $F_X^{-1}(\alpha) \leq F_X^{-1,+}(\alpha)$ e l'uguaglianza si verifica se la funzione di ripartizione F_X è crescente in $F_X^{-1}(\alpha)$. Enunciamo ora alcune importanti proprietà di cui gode la funzione inversa generalizzata, per maggiori dettagli si veda McNeil et al. (2005) e Dhaene et al. (2006).

Proprietà della funzione inversa generalizzata A.1.1

- (i) È monotona non decrescente, continua a sinistra.
- (ii) $F_X(F_X^{-1}(\alpha)) \geq \alpha$, per ogni $\alpha \in (0, 1)$; vale l'uguaglianza se e solo se F_x è continua in $x = F_X^{-1}(\alpha)$.
- (iii) $F_X^{-1}(F_X(x)) \leq x$, per ogni $x \in \mathbb{R}$.
- (iv) $F_X^{-1}(\alpha) \leq x \Leftrightarrow \alpha \leq F_X(x)$, per ogni $x \in \mathbb{R}$, per ogni $\alpha \in (0, 1)$; in particolare vale l'uguaglianza se F_X è crescente e continua. In tale caso la funzione inversa coincide con la funzione inversa generalizzata.

- (v) Tratti di discontinuità di F_X corrispondono a tratti di costanza di F_X^{-1} e viceversa.
- (vi) $F_{g(X)}^{-1}(\alpha) = g(F_X^{-1}(\alpha))$, per ogni g funzione non decrescente e continua a sinistra, per ogni $0 < \alpha < 1$.

Presentiamo brevemente un risultato molto interessante concernente la funzione inversa generalizzata: il metodo dell'inversione.

Metodo dell'inversione A.1.2

Sia F una funzione di ripartizione. Considerata una variabile aleatoria U con distribuzione uniforme sull'intervallo $(0, 1)$, ovvero $U \sim \text{Unif}(0, 1)$, si ha:

$$P(F^{-1}(U) \leq x) = F(x), \text{ per ogni } x \in \mathbb{R}.$$

Equivalentemente, si può considerare una variabile aleatoria X con F come funzione di ripartizione. Si ha allora che: $F^{-1}(U) \stackrel{d}{=} X$, per ogni $U \sim \text{Unif}(0, 1)$.

Alla luce di questo risultato, per generare un campione pseudo-casuale da una qualsiasi variabile aleatoria è sufficiente disporre della funzione inversa generalizzata e di un algoritmo per la simulazione di valori pseudo-casuali dalla distribuzione uniforme sull'intervallo $(0, 1)$.

Alla fine di questo capitolo, nella sezione concernente le copule, faremo riferimento ad un altro importante risultato, noto in letteratura come trasformazione integrale di probabilità. Esso è descrivibile essenzialmente come l'opposto del metodo dell'inversione. Entrambi i risultati menzionati discendono dalle proprietà della funzione inversa generalizzata A.1.1

Trasformazione integrale di probabilità A.1.3

Considerata una variabile aleatoria X con funzione di ripartizione F continua, allora la variabile aleatoria Y , definita come $Y = F(X)$, è distribuita uniformemente sull'intervallo $(0, 1)$.

A.2 Derivata di Radon-Nikodym

Il Teorema di Radon-Nikodym rappresenta una pietra miliare nella teoria della misura ed è il risultato alla base della teoria della speranza condizionata. Per introdurlo, forniamo preliminarmente alcune importanti definizioni; per maggiori dettagli sul Teorema di Radon-Nikodym, sulla nozione di derivata e le sue proprietà si veda Billingsley (1995), Holzer (2007) e Ash (1972). Consideriamo uno spazio misurabile (Ω, \mathcal{A}, m) , ove con Ω indichiamo l'insieme ambiente (non vuoto), con \mathcal{A} una σ -algebra¹ su Ω e con m una misura² definita sulla σ -algebra \mathcal{A} . Introduciamo due definizioni.

Definizione

Una misura m definita su (Ω, \mathcal{A}) si dice:

- finita se $m(\Omega) < +\infty$.
- σ -finita se esiste una partizione numerabile $(A_n)_{n \in \mathbb{N}}$ di Ω tale che $m(A_n) < +\infty$, per ogni n .

Osserviamo che ogni misura finita è banalmente σ -finita ma non è vero il viceversa. Le misure di probabilità sono un esempio di misure finite per le quali si ha $m(\Omega) = 1$. La misura di Lebesgue unidimensionale λ è invece un esempio di misura σ -finita ma non finita. Procediamo ora con un'importante definizione concernente la relazione tra due misure, essa ha un ruolo centrale nel successivo Teorema di Radon-Nikodym.

Definizione

Considerate due misure ν, μ definite su (Ω, \mathcal{A}) , diremo che ν è assolutamente continua rispetto a μ , e scriviamo $\nu \ll \mu$, se:

$$\mu(A) = 0 \Rightarrow \nu(A) = 0, \text{ per ogni } A \in \mathcal{A}.$$

¹Famiglia di sottoinsiemi di 2^Ω contenente l'insieme ambiente e chiusa per complementazione e unione numerabile.

²Applicazione di \mathcal{A} in $[0, +\infty]$ che si annulla sull'insieme vuoto ed è numerabilmente additiva.

Qualora risulti anche $\mu \ll \nu$ diremo che le due misure sono equivalenti.

Quindi, assunta una misura μ , la misura ν è assolutamente continua rispetto a μ se concorda sugli insiemi di misura nulla. Introduciamo ora la nozione di derivata di Radon-Nikodym attraverso il seguente Teorema che rappresenta un risultato fondamentale nell'ambito della teoria della misura.

Teorema di Radon-Nikodym A.2.1

Consideriamo uno spazio misurabile (Ω, \mathcal{A}) e due misure σ -finite ν, μ definite su tale spazio. Se $\nu \ll \mu$ allora esiste una funzione $(\mathcal{A}, \mathcal{B})$ -misurabile^{3 4} $f : \Omega \rightarrow [0, +\infty)$, per la quale si ha:

$$\nu(A) = \int_A f d\mu, \text{ per ogni } A \in \mathcal{A}.$$

La funzione f si chiama derivata di Radon-Nikodym (RN) di ν rispetto μ ed è unica a meno di insiemi μ -trascurabili. Per essa è impiegata la notazione $f = \frac{d\nu}{d\mu}$. Prima di interpretare il Teorema nell'ambiente relativo al calcolo delle probabilità enunciamo alcune proprietà di cui gode la derivata di RN.

Proprietà della derivata di Radon-Nikodym A.2.2

Consideriamo tre misure σ -finite ν, μ e λ su (Ω, \mathcal{A}) . Sussistono le seguenti proprietà:

(i) Se $\nu \ll \mu \ll \lambda$, allora: $\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu} \frac{d\mu}{d\lambda}$, a meno di insiemi λ -trascurabili.

(ii) Se ν e μ sono equivalenti allora: $\frac{d\mu}{d\nu} = \left(\frac{d\nu}{d\mu} \right)^{-1}$, a meno di insiemi μ -trascurabili.

³La σ -algebra di Borel, indicata con \mathcal{B} , è la minima σ -algebra definita sull'insieme dei numeri reali che contiene tutti gli intervalli.

⁴Considerata, accanto alla σ -algebra \mathcal{A} su Ω , una σ -algebra \mathcal{A}' su un insieme non vuoto Ω' , l'applicazione $f : \Omega \rightarrow \Omega'$ è $(\mathcal{A}, \mathcal{A}')$ -misurabile se $f^{-1}(A') \in \mathcal{A}$ per ogni $A' \in \mathcal{A}'$.

(iii) Se $\mu \ll \lambda$ e $g : \Omega \rightarrow \mathbb{R}$, è una funzione μ -integrabile allora:

$$\int_{\Omega} g d\mu = \int_{\Omega} g \frac{d\mu}{d\lambda} d\lambda$$

Consideriamo da ora in avanti esclusivamente misure di probabilità, cioè, come si è già menzionato, misure che valgono 1 sull'insieme ambiente (sono quindi σ -finite). Assunta una misura di probabilità P , lo spazio (Ω, \mathcal{A}, P) è in questo caso detto spazio di probabilità. Consideriamo una variabile aleatoria⁵ X non negativa (a meno di insiemi trascurabili) definita sullo spazio introdotto e tale che: $E_P(X) = 1$; definiamo, a partire da questa, una funzione d'insieme Q su \mathcal{A} siffatta:

$$Q(A) = E_P(\mathbf{1}_A X) = \int_A X dP, \quad A \in \mathcal{A}. \quad (\text{A.2})$$

Q è una misura di probabilità su \mathcal{A} tale che $Q \ll P$, infatti si ha in primo luogo: $Q(\Omega) = E_P(\mathbf{1}_{\Omega} X) = E_P(X) = 1$. Sussiste inoltre la numerabile additività in quanto, considerata la successione di insiemi a due a due disgiunti $(A_n)_{n \in \mathbb{N}}$, risulta:

$$\begin{aligned} Q\left(\bigcup_{n \geq 1} A_n\right) &= E_P(\mathbf{1}_{\{\cup_{n \geq 1} A_n\}} X) = E_P\left(\sum_{n \geq 1} \mathbf{1}_{A_n} X\right) \\ &= \sum_{n \geq 1} E_P(\mathbf{1}_{A_n} X) = \sum_{n \geq 1} Q(A_n). \end{aligned}$$

L'assoluta continuità di Q rispetto a P è infine una banale conseguenza della definizione.

Se a partire dalla misura P e dalla variabile aleatoria X si definisce una nuova misura Q tramite la (A.2) si ha che essa è una misura di probabilità tale che $Q \ll P$; il Teorema di Radon-Nikodym dimostra che vale anche il viceversa. Pertanto, date due misure di probabilità P e Q tali che Q sia assolutamente continua rispetto a P il Teorema garantisce l'esistenza di una

⁵Una qualunque applicazione $f : \Omega \rightarrow \mathbb{R}$ che sia $(\mathcal{A}, \mathcal{B})$ -misurabile.

variabile aleatoria non negativa X tramite la quale è possibile calcolare Q attraverso P sfruttando la relazione (A.2). La derivata di RN ($X = \frac{dQ}{dP}$) consente quindi di esprimere valutazioni probabilistiche relative alla misura di probabilità Q tramite la misura di probabilità P .

A.3 Divergenza di Kullback - Leibler

A fronte della nozione di derivata di RN affrontata nella sezione precedente, introduciamo in questa sezione una definizione che ha un ruolo fondamentale per il modello di analisi di sensitività affrontato nel secondo capitolo. La divergenza di Kullback-Leibler (o entropia relativa), definita qui di seguito, è stata introdotta in Kullback & Leibler (1951) e rappresenta interpretativamente una “misura” della discrepanza tra due misure di probabilità.

Definizione

Fissato uno spazio misurabile (Ω, \mathcal{A}) , consideriamo due misure di probabilità P e Q definite su tale spazio. La divergenza di Kullback-Leibler (KL) di Q rispetto a P è così definita:

$$D_{KL}(Q\|P) = \int_{\Omega} \log \left(\frac{dQ}{dP} \right) dQ = E_Q \left(\log \left(\frac{dQ}{dP} \right) \right), \quad \text{se } Q \ll P$$

mentre è pari a $+\infty$ se la condizione $Q \ll P$ non è rispettata. Si osservi che la definizione può essere proposta in un'altra chiave (valore atteso rispetto a P) in quanto, alla luce di A.2.2 (iii), si ha:

$$D_{KL}(Q\|P) = \int_{\Omega} \frac{dQ}{dP} \log \left(\frac{dQ}{dP} \right) dP = E_P \left(\frac{dQ}{dP} \log \left(\frac{dQ}{dP} \right) \right). \quad (\text{A.3})$$

La posizione (A.3) è impiegata molto spesso nel secondo capitolo. Osserviamo che, dalla definizione, risulta banalmente che la divergenza di Kullback-Leibler non è simmetrica; è sufficiente infatti considerare due misure di probabilità tali che $Q \ll P$ ma non $P \ll Q$, in questo caso la divergenza di Q rispetto a P è un numero reale finito mentre quella di P rispetto a Q non è finita. Pertanto la divergenza di Kullback-Leibler non è una metri-

ca. Enunciamo ora alcune importanti proprietà di cui gode la divergenza di Kullback-Leibler; per maggiori dettagli si veda Cover e Thomas (2012) e Kullback (1959).

Proprietà della divergenza di Kullback-Leibler A.3.1

Considerate due misure di probabilità P e Q tali che $Q \ll P$, valgono, per la divergenza di Kullback-Leibler di Q rispetto a P , le seguenti proprietà strutturali:

- (i) $D_{KL}(Q\|P) \geq 0$.
- (ii) $D_{KL}(Q\|P) = 0$ se e solo se $Q = P$.
- (iii) $D_{KL}(Q\|P)$ è strettamente convessa in Q .

Osserviamo infine che la divergenza non gode in generale neppure della disuguaglianza triangolare. Alla luce di quanto descritto, la divergenza di Kullback-Leibler è una premetrica che genera, ad ogni modo, una topologia sullo spazio delle distribuzioni di probabilità.

La divergenza di Kullback-Leibler fa parte di una classe più ampia di divergenze chiamate f -divergenze introdotte in Ali e Silvey (1966). Per un'assegnata funzione $f : [0, +\infty) \rightarrow [0, +\infty)$ convessa tale che $f(1) = 0$, la f -divergenza di Q rispetto a P è definita come:

$$D_f(Q\|P) = \int_{\Omega} f\left(\frac{dQ}{dP}\right) dP.$$

La divergenza di Kullback-Leibler è dunque una f -divergenza con $f(x) = x \log(x)$, ove si considera che $0 \log(0) = 0$.

A.4 Relazioni di dipendenza tra variabili aleatorie

Fissato uno spazio di probabilità (Ω, \mathcal{A}, P) , tutte le successive quantità aleatorie si considereranno definite su questo spazio. Introduciamo le nozioni di comonotonia e contromonotonia per una coppia⁶ aleatoria (X, Y) ; esse traducono il più estremo tipo di dipendenza rispettivamente positiva e negativa per una coppia aleatoria con fissate marginali (McNeil et al. (2005)).

Definizione

Una coppia di variabili aleatorie (X, Y) si dice comonotona (contromonotona) se esiste una variabile aleatoria Z e due funzioni non decrescenti (non crescenti) g, h tali che: $(X, Y) \stackrel{d}{=} (g(Z), h(Z))$.

Si possono dimostrare alcune importanti condizioni equivalenti alla definizione di comonotonia e a quella di contromonotonia.

Teorema di caratterizzazione comonotonia A.4.1

Data una coppia aleatoria (X, Y) , con funzione di ripartizione congiunta $F_{X,Y}$ e marginali F_X e F_Y , sono equivalenti le seguenti 6 proposizioni, (per le ultime due è necessaria un'ipotesi aggiuntiva) :

- (i) (X, Y) è comonotona.
- (ii) Per ogni $x, y \in \mathbb{R}$, $F_{X,Y}(x, y) = \min\{F_X(x), F_Y(y)\}$.
- (iii) $(X, Y) \stackrel{d}{=} (F_X^{-1}(U), F_Y^{-1}(U))$ ove U è una variabile aleatoria distribuita uniformemente sull'intervallo $(0, 1)$.
- (iv) Esiste $B \in \mathcal{A}$ con $P(B) = 1$, tale che per ogni $\omega_1, \omega_2 \in B$, si ha:

$$[X(\omega_1) < X(\omega_2) \Rightarrow Y(\omega_1) \leq Y(\omega_2)]$$

⁶La nozione di comonotonia può essere estesa in generale a vettori aleatori, quella di contromonotonia riguarda invece esclusivamente coppie aleatorie.

Se inoltre F_X, F_Y sono funzioni continue le 4 proposizioni suddette sono equivalenti anche alle seguenti:

$$(v) Y = F_Y^{-1}(F_X(X)).$$

(vi) Esiste una funzione g non decrescente tale che: $Y = g(X)$ con probabilità 1.

Si osservi che per qualunque coppia aleatoria (anche non comonotona) sussiste la relazione $F_{X,Y}(x, y) \leq \min\{F_X(x), F_Y(y)\}$; il Teorema ci mostra quindi che l'uguaglianza vale se e solo se (X, Y) è comonotona. Di particolare rilevanza interpretativa sono le proposizioni (iv) e (vi). La (iv) indica che la coppia aleatoria (X, Y) è comonotona se e solo se al crescere di una variabile cresce anche l'altra; la (vi) indica invece che, nel caso di funzioni di ripartizioni continue, la comonotonia è condizione sufficiente e necessaria per l'esistenza di una funzione non decrescente che trasforma X in Y ⁷. Anche per la nozione di contromonotonia si può fornire un simile Teorema di caratterizzazione.

Teorema di caratterizzazione contromonotonia A.4.2

Data una coppia aleatoria (X, Y) , con funzione di ripartizione congiunta $F_{X,Y}$ e marginali F_X e F_Y , sono equivalenti le seguenti 6 proposizioni, (per le ultime due è necessaria un'ipotesi aggiuntiva) :

(i) (X, Y) è contromonotona.

(ii) $(X, Y) \stackrel{d}{=} (F_X^{-1}(U), F_Y^{-1}(1 - U))$ ove U è una variabile aleatoria distribuita uniformemente sull'intervallo $(0, 1)$.

(iii) Per ogni $x, y \in \mathbb{R}$, $F_{X,Y}(x, y) = \max\{F_X(x) + F_Y(y) - 1, 0\}$.

(iv) Esiste $B \in \mathcal{A}$ con $P(B) = 1$, tale che per ogni $\omega_1, \omega_2 \in B$ si ha:

$$[X(\omega_1) < X(\omega_2) \Rightarrow Y(\omega_1) \geq Y(\omega_2)].$$

⁷Si parla in letteratura in questo caso di dipendenza monotona tra X e Y .

Se inoltre F_X, F_Y sono funzioni continue le 4 proposizioni suddette sono equivalenti anche alle seguenti:

$$(v) Y = F_Y^{-1}(1 - F_X(X)).$$

(vi) Esiste una funzione g non crescente tale che: $Y = g(X)$ con probabilità 1.

Nel caso della contromonotonia, risultano essere validi i medesimi commenti del caso della monotonia con opportuni accorgimenti. Ad esempio nel caso della relazione (vi), osserviamo che la condizione di contromonotonia è condizione sufficiente e necessaria per l'esistenza di una funzione non crescente che trasforma una variabile aleatoria nell'altra. Prima di introdurre un'altra nozione di dipendenza tra variabili aleatorie, enunciamo il seguente Teorema che fornisce indicazioni sul rapporto tra la nozione di comonotonia e quella di correlazione.

Comonotonia, contromonotonia e covarianza A.4.3

Consideriamo una coppia aleatoria (X, Y) . Sussistono le seguenti relazioni:

$$(i) (X, Y) \text{ è comonotona} \Rightarrow \text{Cov}(X, Y) \geq 0$$

$$(ii) (X, Y) \text{ è contromonotona} \Rightarrow \text{Cov}(X, Y) \leq 0$$

(iii) Considerate due coppie aleatorie (X^*, Y^*) e (X', Y') con medesime marginali di (X, Y) sia (X^*, Y^*) comonotona e (X', Y') contromonotona. Sussiste la seguente relazione:

$$\text{Cov}(X', Y') \leq \text{Cov}(X, Y) \leq \text{Cov}(X^*, Y^*).$$

Si osservi che per le relazioni (i) e (ii) non vale il viceversa. La relazione (iii) implica inoltre che, per ogni coppia aleatoria (X, Y) si ha:

$$E(X^*Y^*) \geq E(XY). \tag{A.4}$$

Le relazioni di comonotonia e contromonotonia, come si è detto, rappresentano il tipo di dipendenza rispettivamente positiva e negativa, più estrema. In

termini generali la dipendenza positiva e negativa tra due variabili aleatorie è individuata dalla seguente definizione:

Definizione

Una coppia di variabili aleatorie (X, Y) si dice *positively quadrant dependent* (PQD) se:

$$P(X \leq x, Y \leq y) \geq P(X \leq x)P(Y \leq y).$$

Specularmente, essa si dice *negatively quadrant dependent* (NQD) se:

$$P(X \leq x, Y \leq y) \leq P(X \leq x)P(Y \leq y).$$

Si osservi che, in relazione ad una coppia aleatoria, le relazioni di comonotonia e contromonotonia implicano rispettivamente PQD e NQD.

A.4.1 Copule

Consideriamo un vettore aleatorio $\mathbf{X} = (X_1, \dots, X_d)$ definito sullo spazio di probabilità (Ω, \mathcal{A}, P) . Il comportamento del vettore è descritto dalla sua distribuzione congiunta $F_{\mathbf{X}}$. Questa può essere vista come la composizione di due elementi. Il primo è descritto dal comportamento delle distribuzioni marginali F_i , $i = 1, \dots, d$. Il secondo dalla struttura di dipendenza tra le sue componenti. L'approccio delle copule consente di scorporare i due effetti, isolandone la struttura di dipendenza. Come vedremo, ciò permette, in generale, di avere una comprensione più profonda sulla struttura di dipendenza stocastica di un vettore aleatorio. Operativamente, le copule sono molto spesso impiegate nel risk management qualora si fosse interessati a costruire, a partire dalle marginali, un modello multivariato con una fissata struttura di dipendenza (come fatto, ad esempio, nel Capitolo 4). Procediamo con la definizione di copula.

Definizione

Una funzione $C : [0, 1]^d \rightarrow [0, 1]$ si dice *copula d-dimensionale* se C è la funzione di ripartizione congiunta di un vettore aleatorio $\mathbf{X} = (X_1, \dots, X_d)$, le

cui marginali sono distribuite uniformemente sull'intervallo $(0, 1)$.

Prima di procedere con l'esposizione del Teorema di Sklar (1959) che permette di comprendere la fondamentale importanza delle copule, evidenziamo alcune proprietà strutturali. Se una funzione C è una copula, valgono le seguenti tre proprietà, che forniscono una caratterizzazione della copula traducendo la definizione in termini analitici.

- (i) $C(u_1, \dots, u_d)$ è non decrescente in ogni componente u_i .
- (ii) $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$, per ogni $i \in \{1, \dots, d\}$, $u_i \in [0, 1]$
- (iii) Per ogni $\mathbf{a}, \mathbf{b} \in \mathbb{R}^d$ tale che $a_i \leq b_i$ per ogni i , si ha:

$$\sum_{i_1=1}^2 \dots \sum_{i_d=1}^2 (-1)^{i_1+\dots+i_d} C(u_{1,i_1}, \dots, u_{d,i_d}) \geq 0,$$

ove $u_{j,1} = a_j$ e $u_{j,2} = b_j$, con $j \in \{1, \dots, d\}$.

Le prime due proprietà sono di semplice interpretazione. La prima consegue banalmente dal fatto che C è una funzione di ripartizione congiunta. La seconda è necessaria affinché le sue marginali siano distribuite uniformemente. Il fondamentale ruolo delle copule nella descrizione della dipendenza stocastica di un vettore aleatorio è garantito dal seguente celeberrimo Teorema.

Teorema di Sklar A.4.1.1

Sia F una funzione di ripartizione congiunta d -dimensionale con marginali F_i , $i = 1, \dots, d$. Allora esiste una copula $C : [0, 1]^d \rightarrow [0, 1]$, tale che, per ogni $x_1, \dots, x_d \in \mathbb{R}$, si ha:

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)). \quad (\text{A.5})$$

Se le marginali F_i sono funzioni continue, allora C è unica; altrimenti C è univocamente determinata sul prodotto cartesiano dei range di variazione delle marginali. Vale anche il viceversa: se C è una copula e F_1, \dots, F_d sono funzioni di ripartizione univariate allora, F , definita come in (A.5), è una

funzione di ripartizione congiunta con F_1, \dots, F_d come marginali.

Il Teorema di Sklar mostra dunque che una funzione di ripartizione congiunta può essere espressa tramite la copula e le marginali. Inoltre, nel caso di marginali continue, la nozione di copula è più naturale in quanto è garantita l'unicità. La seguente osservazione fa comprendere l'importanza, nell'ambito della teoria base delle copule, della trasformazione integrale di probabilità.

Considerato un vettore aleatorio $\mathbf{X} = (X_1, \dots, X_d)$ con funzioni di ripartizione marginali F_i continue per $i = 1, \dots, d$. Applicando ad ogni componente la trasformazione integrale di probabilità si ottiene un vettore aleatorio la cui funzione di ripartizione è la copula:

$$(U_1, \dots, U_d) = (F_1(X_1), \dots, F_d(X_d)),$$

ove dunque (U_1, \dots, U_d) è un vettore aleatorio avente le marginali distribuite uniformemente sull'intervallo $(0, 1)$. Disponendo di una metodologia per effettuare simulazioni da (U_1, \dots, U_d) è possibile ricondursi a simulazioni dal vettore aleatorio \mathbf{X} tramite la relazione (metodo dell'inversione):

$$(X_1, \dots, X_d) = (F_1^{-1}(U_1), \dots, F_d^{-1}(U_d)).$$

Questa osservazione fa comprendere il meccanismo tramite il quale si possono generare campioni pseudo-casuali impiegando le marginali e la copula in un ambiente di simulazione. Formalmente, quanto osservato spinge inoltre ad introdurre la definizione di copula di una funzione di ripartizione congiunta F con marginali continue.

Definizione

Considerato un vettore aleatorio \mathbf{X} con funzione di ripartizione congiunta F e marginali F_1, \dots, F_d continue, si dice copula di F la funzione di ripartizione congiunta del vettore aleatorio $(F_1(X_1), \dots, F_d(X_d))$.

Si potrebbero fornire molti interessanti risultati generali concernenti le copule. Per i fini di questa trattazione ne proponiamo uno in quanto sarà utile nel seguito.

Proposizione A.4.1.2

Considerato il vettore aleatorio $\mathbf{X} = (X_1, \dots, X_d)$ con marginali F_i , $i = 1, \dots, d$ continue. Sia $T_i : \mathbb{R} \rightarrow \mathbb{R}$, una funzione reale crescente per ogni $i = 1, \dots, d$. Allora il vettore aleatorio $(T_1(X_1), \dots, T_d(X_d))$ ammette la medesima copula di \mathbf{X} .

La copula gode dunque di una proprietà di invarianza per trasformazioni crescenti.

Procediamo l'esposizione tramite alcuni importanti esempi di copula. Gli esempi che si potrebbero considerare sono moltissimi, ci concentriamo su quelli impiegati nel Capitolo 4. Il primo esempio è la copula di indipendenza che rappresenta un caso patologico.

La copula di indipendenza Π è definita dalla seguente espressione:

$$\Pi(u_1, \dots, u_d) = \prod_{i=1}^d u_i. \quad (\text{A.6})$$

Alla luce del Teorema di Sklar, e della relazione (A.5), appare chiaro che, un vettore aleatorio con marginali continue e stocasticamente indipendenti, è caratterizzato dalla copula di indipendenza (A.6).

Impiegando il Teorema di Sklar è possibile estrarre, a partire da vettori aleatori con distribuzioni congiunte note (ad esempio la normale multivariata), la struttura della copula. Ciò apre la possibilità di introdurre una vastissima classe di copule. Si parla in questo caso di copule implicite in quanto esse non hanno necessariamente un'espressione in forma chiusa e sono definite implicitamente a partire dalla distribuzione multivariata di partenza.

Sia \mathbf{X} un vettore aleatorio distribuito come una normale multivariata d -dimensionale: $\mathbf{X} \sim N_d(\mathbf{m}, \Sigma)$. La copula ottenibile dalla funzione di ripartizione multivariata della normale è detta copula gaussiana. In questo caso, segue dalla proposizione A.4.1.2 che, il vettore aleatorio \mathbf{X} ammette medesima copula di $\mathbf{Y} \sim N_d(\mathbf{0}, \mathcal{P})$, ove \mathcal{P} è la matrice di correlazione di \mathbf{X} . Ciò permette di comprendere un risultato fondamentale concernente le copule. Consideriamo, per semplicità, il caso $d = 2$. Alla luce di quanto appena osservato, il coefficiente di correlazione ρ individua in questo caso univocamente la copula gaussiana (il risultato vale in generale per le distribuzioni ellittiche si veda Embrechts et al. (2002)). Ciò fa intuire che, qualora si operi con distribuzioni normali, si può impiegare in modo naturale il coefficiente di correlazione per descrivere la copula. Il risultato però non è vero in generale. Analizzeremo in dettaglio questo aspetto successivamente. Indicata con $\Phi_{\mathcal{P}}$ la funzione di ripartizione congiunta della normale multivariata con media nulla e \mathcal{P} matrice di correlazione e con Φ la funzione di ripartizione della normale standard, l'espressione della copula in generale è:

$$C_{\mathcal{P}}^{Ga}(\mathbf{u}) = \Phi_{\mathcal{P}}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d)).$$

Si impiega la notazione $C_{\mathcal{P}}^{Ga}$ per enfatizzare la dipendenza dalla matrice di correlazione. La copula gaussiana con $\mathcal{P} = I_d$, ove I_d è la matrice identica d -dimensionale, degenera nella copula di indipendenza. Come si è fatto per la copula gaussiana, si può estrarre univocamente la copula implicita a partire da qualsiasi distribuzione multivariata con marginali continue. Un importante esempio, largamente impiegato nella pratica, è la copula t-student. In questo caso l'espressione della copula è dunque la seguente:

$$C_{\nu, \mathcal{P}}^t(\mathbf{u}) = t_{\nu, \mathcal{P}}(t_{\nu}^{-1}(u_1), \dots, t_{\nu}^{-1}(u_d)),$$

ove t_{ν} è la funzione di ripartizione della distribuzione t-student standard univariata con ν gradi di libertà e $t_{\nu, \mathcal{P}}$ è la funzione di ripartizione congiunta del vettore aleatorio $\mathbf{X} \sim t_{\nu, \mathcal{P}}(\mathbf{0}, \mathcal{P})$. Nel caso della copula t-student, se $\mathcal{P} = I_d$, non si ottiene la copula di indipendenza in quanto t-student incorrelate non

sono indipendenti (McNeil et al. (2015)). Presentiamo infine un ultimo importante esempio: la copula di Gumbel. A differenza della copula gaussiana e quella di t-student, la copula di Gumbel appartiene alla famiglia di distribuzioni che ammette una espressione esplicita per la copula. Assumendo per semplicità $d = 2$, la sua espressione è la seguente:

$$C_{\theta}^{Gu}(u_1, u_2) = \exp(-((- \log u_1)^{\theta} + (- \log u_2)^{\theta})^{\frac{1}{\theta}}), \quad 1 \leq \theta < +\infty.$$

Se $\theta = 1$, C_{θ}^{Gu} degenera nella copula di indipendenza. L'importanza della copula di Gumbel è dovuta alla particolare tipo di dipendenza che essa implica tra le componenti, discuteremo questo aspetto nel seguito. Per comprendere alcune importanti differenze tra le copule introdotte si possono impiegare opportune misure di dipendenza che introduciamo nella successiva sottosezione.

A.4.2 Misure di dipendenza

Considereremo, in questa sottosezione, quattro misure di dipendenza. Con misura di dipendenza si intende una funzione che associa ad una coppia aleatoria (X_1, X_2) un numero reale, che in qualche modo ne descrive la struttura di dipendenza. La metodologia con il quale a partire dalla coppia aleatoria viene associato il numero reale è peculiare di ogni misura.

Correlazione lineare (di Pearson)

La correlazione lineare tra le variabili aleatorie X_1 e X_2 , dotate di varianza finita e positiva, è definita dal rapporto:

$$\rho(X_1, X_2) = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}}.$$

La disuguaglianza di Cauchy-Schwartz garantisce che il coefficiente di correlazione lineare assuma valori nell'intervallo $[-1, 1]$. Si ha che, se X_1 e X_2 sono stocasticamente indipendenti, allora $\rho(X_1, X_2) = 0$; ma non è vero il viceversa (a meno di distribuzioni ellittiche). La condizione $|\rho(X_1, X_2)| = 1$

implica perfetta dipendenza lineare tra le due variabili aleatorie, ovvero X_2 è esprimibile come $X_2 = \alpha + \beta X_1$ con $\beta > 0$ qualora la dipendenza sia positiva e $\beta < 0$ se negativa. L'indice misura pertanto la dipendenza lineare tra le due variabili aleatorie. Si osservi inoltre che la correlazione lineare non si conserva per trasformazioni crescenti. Ovvero, considerata una trasformazione $T : \mathbb{R} \rightarrow \mathbb{R}$ si ha in generale:

$$\rho(X_1, X_2) \neq \rho(T(X_1), T(X_2)).$$

Si rammenti che la proposizione (5.2) mostrava che la copula è invariante per questo tipo di trasformazioni. Pertanto, unendo i due risultati, si comprende che la correlazione lineare non può dipendere esclusivamente dalla copula ma in generale dipende anche dalle marginali. Questa limitazione ha particolare rilevanza quando si è interessati a costruire un modello multivariato a partire dalle marginali e impiegando la correlazione per descriverne la struttura di dipendenza. Alternativamente, per osservare che la correlazione lineare dipende sia dalle marginali che dalla copula, si può considerare la seguente espressione per la covarianza:

$$\text{Cov}(X_1, X_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (F(x_1, x_2) - F_1(x_1)F_2(x_2)) dx_1 dx_2,$$

da cui tramite la (A.5) si arriva ad un'espressione non riducibile che dipende da copula e marginali. Introduciamo un secondo indice di dipendenza che rimedia ad alcune delle problematiche della correlazione lineare.

Indice di correlazione di Spearman

L'indice di correlazione di Spearman prende in considerazione il rango delle variabili aleatorie di riferimento. A partire dalle variabili aleatorie X_1 e X_2 , esso può essere definito come:

$$\rho_S(X_1, X_2) = \rho(F_1(X_1), F_2(X_2)).$$

Si tratta pertanto dell'indice di correlazione lineare applicato ad una particolare trasformazione (la trasformazione integrale di probabilità) di X_1 e X_2 . Anche l'indice di correlazione di Spearman assume valori nell'intervallo $[-1, 1]$. Inoltre esso vale 0 qualora X_1 e X_2 sono stocasticamente indipendenti (non vale il viceversa). Si può dimostrare che i valori 1 e -1 corrispondono al più estremo tipo di dipendenza stocastica rispettivamente positiva e negativa, ovvero la comonotonia e la contromonotonia. La correlazione di Spearman misura pertanto la dipendenza monotona tra le due variabili aleatorie. Per altri dettagli e proprietà si veda Embrechts et al. (2002).

La seguente proposizione fornisce un'importante proprietà di questo indice. Considerata la coppia aleatoria (X_1, X_2) con marginali continue e copula C , vale la seguente rappresentazione alternativa per l'indice di correlazione di Spearman:

$$\rho_S(X_1, X_2) = 12 \int_0^1 \int_0^1 (C(u_1, u_2) - u_1 u_2) du_1 du_2.$$

La proposizione dimostra dunque che, diversamente dall'indice di correlazione lineare, l'indice di correlazione di Spearman può essere rappresentato dalla sola copula. Per questa ragione nel corso della trattazione, negli esempi numerici, si fa sempre ricorso all'indice di Spearman. Si noti inoltre che, sebbene la correlazione di Spearman non dipenda dalle marginali, ciò non comporta che a partire essa e le marginali si individuino univocamente una distribuzione multivariata (a meno che non si lavori con marginali appartenenti alla famiglia delle distribuzioni ellittiche). Un altro vantaggio della correlazione di Spearman è che consente di costruire, assegnate due marginali, una distribuzione congiunta con qualsiasi livello di correlazione $\rho_S \in [-1, 1]$; ciò non è in generale vero per la correlazione lineare.

Coefficienti di dipendenza di coda

I coefficienti di dipendenza di coda misurano la forza della dipendenza tra le code delle marginali di una coppia aleatoria. Come il coefficiente di correlazione di Spearman, anche i coefficienti di dipendenza di coda dipendono

esclusivamente dalla struttura della copula e non dalle marginali; sono dunque invarianti per trasformazioni crescenti (McNeil et al. (2015) e Embrechts et al. (2002)). Si distinguono il coefficiente di dipendenza di coda superiore (λ_u), e quello inferiore (λ_l). Per entrambi gli indici l'idea è di considerare congiuntamente il comportamento sulle code (superiore per λ_u e inferiore per λ_l) della coppia aleatoria. Analiticamente λ_u è definito dall'espressione:

$$\lambda_u(X_1, X_2) = \lim_{q \rightarrow 1^-} P(X_2 > F_2^{-1}(q) \mid X_1 > F_1^{-1}(q)).$$

Il coefficiente di dipendenza di coda superiore λ_u assume valori nell'intervallo chiuso $[0, 1]$, qualora fosse pari a 0 si parla di indipendenza asintotica (di coda superiore). λ_l è definito specularmente dalla relazione:

$$\lambda_l(X_1, X_2) = \lim_{q \rightarrow 0^+} P(X_2 \leq F_2^{-1}(q) \mid X_1 \leq F_1^{-1}(q)).$$

Anche il coefficiente di dipendenza di coda inferiore assume valori nell'intervallo chiuso $[0, 1]$. In entrambi i casi, valori maggiori di 0, misurano la forza della dipendenza di coda. A partire dalle definizioni, si può facilmente dimostrare che, nel caso di marginali continue, i due coefficienti sono rappresentabili dalla sola copula. Ad esempio, per il coefficiente di coda inferiore, si ha:

$$\lambda_l(X_1, X_2) = \lim_{q \rightarrow 0^+} \frac{P(X_2 \leq F_2^{-1}(q), X_1 \leq F_1^{-1}(q))}{P(X_1 \leq F_1^{-1}(q))} = \lim_{q \rightarrow 0^+} \frac{C(q, q)}{q}.$$

In alcuni semplici casi, è possibile ottenere per questi coefficienti espressioni in forma chiusa che permettono di comprendere come al variare dei parametri della distribuzione si comporti la dipendenza di coda della copula. Con particolare riferimento alla coda superiore, esponiamo alcuni esempi in quanto sono utilizzati nella sezione concernente la dipendenza stocastica between del modello sul rischio di credito 4.4.3. Nel caso della copula gaussiana, si può dimostrare che essa è asintoticamente indipendente in entrambe le code ovvero $\lambda_u = \lambda_l = 0$. Anche con una correlazione molto alta, eventi estremi sono indipendenti. Questa proprietà fa comprendere che la copula gaussiana

non è adatta a modellare situazioni nelle quali si crede che ci possa essere una dipendenza tra eventi estremi. Nel caso della copula Gumbel o quella t-student, si possono ottenere espressioni esplicite per i due coefficienti, che mostrano il comportamento di queste al variare dei parametri. In particolare, per λ_u , si hanno le seguenti espressioni:

$$\lambda_u^t = 2t_{\nu+1} \left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}} \right), \quad \lambda_u^{Gu} = 2 - 2^{\frac{1}{\theta}}.$$

La dipendenza di coda superiore per la t-student diminuisce all'aumentare dei gradi di libertà, mentre per la Gumbel aumenta all'aumentare di θ . A titolo di esempio, per visualizzare graficamente il fenomeno, si veda la Figura A.1. Si sono rappresentate 10 000 simulazioni da una coppia aleatoria con medesime marginali e coefficiente di correlazione lineare e diversa copula. Nel caso della copula Gumbel, a differenza della gaussiana, eventi estremi sulla coda superiore tendono ad accadere in contemporanea.

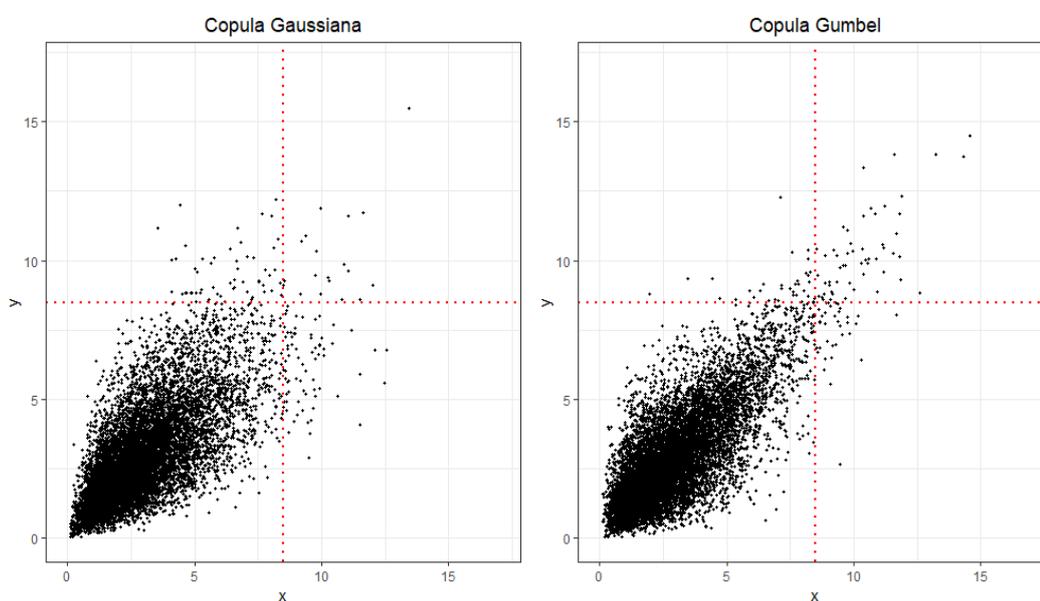


Figura A.1: 10 000 simulazioni da due distribuzioni con identiche marginali distribuite come Gamma(3,1), medesima correlazione lineare $\rho = 0.7$ ma diversa struttura di dipendenza.

Appendice B

Modelli stocastici per la mortalità

Introduciamo in questa seconda Appendice alcuni importanti modelli demografici per la descrizione della mortalità di una popolazione. Nella prima sezione si presenta il famoso modello di Lee-Carter che rappresenta il punto di partenza dell'analisi. A partire da questo modello, sono state introdotte in letteratura numerose estensioni e generalizzazioni. La seconda sezione, seguendo Villegas et al. (2018), espone la struttura della famiglia dei modelli *Age - Period - Cohort generalizzata* (GAPC) che consente di descrivere, sotto un'unica ottica, diversi modelli stocastici presenti in letteratura. L'Appendice si conclude con un'analisi dell'accostamento di cinque modelli di riferimento della famiglia GAPC stimati con dati di popolazione riferiti all'Italia. Il miglior modello, selezionato dall'analisi, è stato impiegato nel terzo capitolo per descrivere la mortalità degli assicurati del portafoglio di rendite vitalizie.

B.1 Il modello di Lee - Carter

Il più famoso approccio stocastico per modellare i tassi di mortalità proiettati è stato introdotto in Lee e Carter (1992). A partire da questo modello si è sviluppata un'intensa letteratura che ha studiato con attenzione le proprietà

del modello e sono state analizzate diverse possibili estensioni alla sua struttura base (Villegas et al. (2018), Pitacco et al. (2009), Giroi e King (2007) e Brouhns et al. (2002)).

Nella formulazione originale di Lee e Carter, è proposta per il logaritmo del tasso centrale di mortalità $m_{x,t}$, con riferimento agli individui di età x nell'anno di calendario t , la seguente formulazione:

$$\log m_{x,t} = a_x + b_x k_t + \varepsilon_{x,t}. \quad (\text{B.1})$$

Il termine di errore $\varepsilon_{x,t}$ è ipotizzato essere distribuito come una normale di media 0 e varianza σ^2 . Infine, è necessaria un'ulteriore ipotesi che imponga alcuni vincoli sui parametri in quanto la struttura (B.1) è sovrapparametrizzata. Lee e Carter hanno proposto i seguenti due vincoli: $\sum_x b_x = 1$, $\sum_t k_t = 0$. Impiegando questi vincoli è possibile tramite un procedimento di stima determinare univocamente tutti i parametri del modello.

Analizzando la (B.1) si può osservare che il logaritmo della tasso centrale di mortalità dipende da un primo termine a_x legato unicamente all'età, da un secondo termine bilineare $b_x k_t$ i cui parametri dipendono rispettivamente da età e anno di calendario e infine da una componente di errore $\varepsilon_{x,t}$. Il termine a_x individua dunque la forma della mortalità attraverso le età e, alla luce dei vincoli, rappresenta la media dei $\log m_{x,t}$ rispetto l'anno di calendario. L'effetto dell'anno di calendario interviene nel modello esclusivamente nel termine k_t ; dunque il termine b_x permette di assegnare a questo effetto un peso che varia con l'età. Un ulteriore aspetto molto importante da tenere in considerazione è che nell'espressione (B.1) tra le variabili indipendenti non ci sono quantità osservabili. Per dettagli sulla procedura di stima a due stadi proposta da Lee e Carter si veda ad esempio Lee - Carter (1992) e Giroi e King (2007). Una volta che si dispone del modello stimato, per ottenere una previsione futura si modella la componente k_t attraverso un processo ARIMA. Molto spesso viene utilizzato, per la proiezione, un semplice processo di passeggiata casuale con drift in quanto l'evidenza empirica ha dimostrato in diversi studi un buon adattamento.

Il modello originale appena descritto ipotizza che gli errori siano distribuiti normalmente e siano inoltre omoschedastici (medesima varianza). Le due ipotesi risultano in generale essere molto forti e poco realistiche, è stata pertanto proposta una struttura alternativa (Brouhns et al. (2002)). Con riferimento all'età x e all'anno di calendario t , indichiamo con $D_{x,t}$ il numero aleatorio di individui deceduti, con $d_{x,t}$ la corrispondente osservazione e con $E_{x,t}^c$ il numero centrale di esposti al rischio. Il rapporto tra $d_{x,t}$ e $E_{x,t}^c$ rappresenta una stima grezza del tasso centrale di mortalità $m_{x,t}$ per individui di età x nell'anno di calendario t .

Assumiamo inoltre che l'intensità istantanea di mortalità $\mu_{x,t}$ sia costante per ogni età e anno di calendario; pertanto essa coincide con il tasso centrale $m_{x,t}$. Per la variabile aleatoria $D_{x,t}$ si formula la seguente ipotesi distributiva:

$$D_{x,t} \sim \text{Poisson}(E_{x,t}^c \mu_{x,t}), \quad (\text{B.2})$$

e inoltre, si assume che, al variare di x e t , condizionatamente a $\{\mu_{x,t}\}$, siano indipendenti. Per il tasso centrale di mortalità (che si rammenti coincide con l'intensità istantanea) è proposta nuovamente la struttura impiegata da Lee-Carter. Si ha dunque:

$$\log \mu_{x,t} = \alpha_x + \beta_x k_t, \quad (\text{B.3})$$

ove si impiegano i vincoli sui parametri già enunciati in precedenza. Per la stima del modello viene sfruttata una classica procedura di massima verosimiglianza e risolta tramite l'algoritmo numerico di Newton-Raphson; la funzione di log-verosimiglianza del problema è siffatta:

$$l = \sum_{x,t} \{D_{x,t}(\alpha_x + \beta_x k_t) - E_{x,t}^c e^{\alpha_x + \beta_x k_t}\} + \text{costante}$$

Per quanto concerne la proiezione, vengono impiegate anche in questo modello le medesime metodologie introdotte in precedenza.

B.2 Estensioni al modello di Lee-Carter

A partire dal modello di Lee-Carter si è sviluppata nel corso degli ultimi anni un'intensa e feconda letteratura che ha introdotto varie estensioni al modello base. Alcuni autori hanno individuato una struttura comune a molti dei nuovi modelli introdotti; presentiamo, in particolare, quella introdotta da Villegas et al. (2018). Essa si chiama famiglia dei modelli Age - Period - Cohort generalizzata (GAPC). In linea con la (famosa) famiglia dei modelli lineari generalizzati (GLM) introdotta da McCullagh and Nelder (1989), un modello appartenente alla famiglia GAPC può essere descritto da quattro elementi: una componente stocastica, una componente deterministica, la funzione di collegamento e un insieme di vincoli sui parametri del modello. Descriviamo brevemente le componenti.

1. Componente stocastica

Consideriamo, come fatto nella sezione precedente, la variabile aleatoria $D_{x,t}$ che rappresenta il numero aleatorio di decessi per individui di età x nell'anno di calendario t . Per questo numero aleatorio sono possibili all'interno della famiglia GAPC due tipologie distinte di specificazione probabilistica:

$$\begin{aligned} D_{x,t} &\sim \text{Poisson}(E_{x,t}^c \mu_{x,t}), \\ D_{x,t} &\sim \text{Binom}(E_{x,t}^0, q_{x,t}), \end{aligned} \tag{B.4}$$

ove, a fronte degli individui di età x nell'anno di calendario t , $E_{x,t}^0$ rappresenta il numero iniziale di esposti al rischio e $q_{x,t}$ la probabilità monoannuale di decesso. Si noti che l'ipotesi Poissoniana coincide con la (B.2) utilizzata in precedenza. Alla luce delle due distinte ipotesi probabilistiche consegue rispettivamente che:

$$\text{E} \left(\frac{D_{x,t}}{E_{x,t}^c} \right) = \mu_{x,t}, \quad \text{E} \left(\frac{D_{x,t}}{E_{x,t}^0} \right) = q_{x,t}.$$

Queste sono le quantità che, tramite la funzione di collegamento, sono modelate dalla componente sistematica. Si ipotizza inoltre che le variabili aleatorie

$D_{x,t}$ siano stocasticamente indipendenti condizionatamente rispettivamente ai parametri $\mu_{x,t}$ o $q_{x,t}$.

2. Componente sistematica

Nei modelli appartenenti alla famiglia GAPC sono individuabili in generale tre effetti. L'effetto dell'età x , l'effetto dell'anno di calendario t e infine l'effetto anno di nascita $t - x$ (coorte). Essi sono espressi all'interno di un previsore siffatto (Hunt and Black (2015)):

$$\eta_{x,t} = \alpha_x + \sum_{i=1}^N \beta_x^{(i)} \kappa_t^{(i)} + \beta_x^{(0)} \gamma_{t-x}. \quad (\text{B.5})$$

Nell'espressione (B.5) si individuano tre componenti. Il primo termine α_x è un termine statico che descrive la forma della mortalità attraverso le età. Il secondo termine $\sum_{i=1}^N \beta_x^{(i)} \kappa_t^{(i)}$, ove N indica il numero di addendi bilineari età - anno di calendario inseriti nella componente sistematica (ad esempio, nel modello di Lee-Carter eteroschedastico analizzato in precedenza, $N = 1$), descrive l'effetto di interazione tra l'età x e l'anno di calendario t . L'ultima componente di (B.5) è ascrivibile all'effetto dell'anno di nascita (effetto di coorte). A seconda del modello considerato, i termini $\beta_x^{(i)}$ possono essere pre-determinati o, come nel caso di Lee-Carter, stimati. Si noti che nei modelli GAPC le componenti κ_t^i , $i = 1, \dots, N$ e γ_{t-x} sono da considerarsi processi stocastici. I tassi di mortalità proiettati si ottengono a partire dalle proiezioni di questi processi.

3. Funzione di collegamento

La componente stocastica e quella sistematica sono tra loro collegate da una funzione g detta, come nel caso dei GLM, funzione di collegamento. La relazione che lega le due componenti è la seguente:

$$g \left(E \left(\frac{D_{x,t}}{E_{x,t}} \right) \right) = \eta_{x,t} \quad (\text{B.6})$$

ove $E_{x,t}$ rappresenta gli esposti al rischio iniziali o centrali. Per la funzione g le scelte ricadono, nel caso della distribuzione di Poisson sul logaritmo e nel caso della distribuzione binomiale sul logit, che sono rispettivamente i legami canonici delle distribuzioni. Per una discussione più approfondita su queste scelte si veda Currie (2016).

4. Insieme di vincoli sui parametri

Molti dei modelli che soddisfano le 3 ipotesi appena descritte non permettono un'identificazione univoca dei parametri senza l'utilizzo di vincoli sui medesimi. Indicato con θ il generico vettore di parametri del modello,

$$\theta = (\alpha_x, \beta_x^{(1)}, \dots, \beta_x^{(N)}, \kappa_t^{(1)}, \dots, \kappa_t^{(N)}, \beta_x^{(0)}, \gamma_{t-x}),$$

sia v una funzione che trasforma questo vettore di parametri in un vettore di parametri alternativo che soddisfa i vincoli del modello senza cambiare $\eta_{x,t}$ ovvero:

$$v(\theta) = \tilde{\theta} = (\tilde{\alpha}_x, \tilde{\beta}_x^{(1)}, \dots, \tilde{\beta}_x^{(N)}, \tilde{\kappa}_t^{(1)}, \dots, \tilde{\kappa}_t^{(N)}, \tilde{\beta}_x^{(0)}, \tilde{\gamma}_{t-x}),$$

$$\eta_{x,t} = \alpha_x + \sum_{i=1}^N \beta_x^{(i)} \kappa_t^{(i)} + \beta_x^{(0)} \gamma_{t-x} = \tilde{\alpha}_x + \sum_{i=1}^N \tilde{\beta}_x^{(i)} \tilde{\kappa}_t^{(i)} + \tilde{\beta}_x^{(0)} \tilde{\gamma}_{t-x}.$$

Stima del modello nella famiglia GAPC

Per la stima del modello, l'approccio generalmente impiegato è quello della massima verosimiglianza. A seconda che per la componente stocastica sia impiegata l'ipotesi di Poisson o quella binomiale, si ottengono rispettivamente le due seguenti funzioni di log-verosimiglianza:

$$\mathcal{L} = \sum_x \sum_t \omega_{x,t} \{d_{x,t} \log \hat{d}_{x,t} - \hat{d}_{x,t} - \log \hat{d}_{x,t}!\}, \quad (\text{B.7})$$

$$\mathcal{L} = \sum_x \sum_t \omega_{x,t} \left\{ d_{x,t} \log \left(\frac{\hat{d}_{x,t}}{E_{x,t}^0} \right) + (E_{x,t}^0 - d_{x,t}) \log \left(\frac{E_{x,t}^0 - \hat{d}_{x,t}}{E_{x,t}^0} \right) + \log \left(\frac{E_{x,t}^0}{d_{x,t}} \right) \right\}, \quad (\text{B.8})$$

ove, in entrambi i casi, $\omega_{x,t}$ è un peso assegnato all'osservazione (x, t) che può essere 1 qualora l'osservazione è inclusa nella verosimiglianza o 0 altrimenti. Il termine $\hat{d}_{x,t}$ è il valore atteso del numero di decessi con riferimento agli individui di età x nell'anno di calendario t , implicato dal modello ed è definito dalla seguente espressione:

$$\hat{d}_{x,t} = E_{x,t} g^{-1} \left(\alpha_x + \sum_{i=1}^N \beta_x^{(i)} \kappa_t^{(i)} + \beta_x^{(0)} \gamma_{t-x} \right).$$

Le funzioni di verosimiglianza introdotte sono massimizzate tramite un approccio di tipo numerico.

B.2.1 Esempi di modelli della famiglia GAPC

Molti dei modelli stocastici per la descrizione della mortalità presenti in letteratura appartengono alla famiglia GAPC. Seguendo Villegas et al. (2018), presentiamo schematicamente cinque dei modelli più significativi. Per tutti e cinque consideriamo l'ipotesi di distribuzione binomiale per la variabile aleatoria $D_{x,t}$ e assumiamo, come funzione di collegamento, il logit. Le specificità di ciascun modello sono pertanto individuate dalla componente sistematica e dall'insieme di vincoli sui parametri. Di particolare interesse per il confronto è la struttura della componente sistematica. In ciascun caso, un diverso insieme di vincoli sui parametri permette di conseguire una stima univoca del modello; questo aspetto, di natura prettamente tecnica, non è di interesse per gli obiettivi di questa Appendice e viene pertanto tralasciato, per dettagli si veda Villegas et al. (2018).

1. Modello di Lee-Carter (LC)

Abbiamo già introdotto in precedenza il modello di Lee-Carter; analizziamo ora come esso si inserisce nel contesto dei modelli GAPC. La struttura della componente sistematica, come già visto, è:

$$\eta_{x,t} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)}.$$

Con riferimento al previsore $\eta_{x,t}$, si ha dunque $N = 1$ e non è presente l'effetto di coorte. Per la proiezione di κ_t si utilizza un processo stocastico ARIMA, in particolare viene spesso impiegata una semplice processo di passeggiata casuale con drift, ovvero:

$$\kappa_t^{(1)} = \delta + \kappa_{t-1}^{(1)} + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_\epsilon^2) \text{ i.i.d.} \quad (\text{B.9})$$

Procediamo con la prima estensione al modello di Lee-Carter nel quale viene introdotto un effetto di coorte.

2. Modello di Renshaw and Haberman (RH)

A partire dal modello di Lee-Carter, si introduce in Renshaw and Haberman (2006) un modello che incorpora anche l'effetto di coorte; si ha quindi la seguente struttura:

$$\eta_{x,t} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)} + \beta_x^{(0)} \gamma_{t-x}.$$

Per la proiezione delle componenti κ_t e γ_{t-x} sono impiegati due processi ARIMA univariati ipotizzando l'indipendenza dei due effetti. In particolare, per κ_t è utilizzato un processo di passeggiata aleatoria con drift ARIMA(0,1,0) e per γ_{t-x} un processo ARIMA(1,1,0). In Haberman and Renshaw (2011), i medesimi autori hanno proposto una struttura alternativa semplificata per il previsore, nella quale $\beta_x^{(0)} = 1$, dimostrando che essa rappresenta un miglioramento in termini di stabilità del modello. Si ha dunque:

$$\eta_{x,t} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)} + \gamma_{t-x}.$$

3. Modello Age - Period - Cohort (APC)

A partire dal modello di Renshaw-Haberman è possibile definire un interessante caso particolare nel quale $\beta_x^{(0)} = \beta_x^{(1)} = 1$. Si ha dunque la seguente struttura per il previsore del modello:

$$\eta_{x,t} = \alpha_x + \kappa_t^{(1)} + \gamma_{t-x}.$$

Il modello APC è stato largamente impiegato in statistica medica e demografia ed è stato introdotto in ambito attuariale in Currie (2006). Anche in questo caso per la proiezioni sono usualmente impiegati due distinti processi ARIMA univariati e indipendenti per κ_t e γ_{t-x} .

4. Modello di Cairns, Blake e Dowd (CBD)

I tre modelli finora considerati sono, seppur diversi, molto simili in termini componente sistematica. In Cairns et al. (2006) viene considerata una struttura molto diversa dalle precedenti per il previsore. In particolare si considera $N = 2$ con i parametri che dipendono dall'età $\beta_x^{(1)}$ e $\beta_x^{(2)}$ prespecificati e pari rispettivamente a 1 e $x - \bar{x}$, ove \bar{x} è l'età media dei dati; non sono considerati nè l'effetto statico dell'età e neppure l'effetto di coorte. Il previsore del modello CBD è dunque:

$$\eta_{x,t} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)}.$$

Un aspetto peculiare del modello CBD è che esso, a differenza di tutti gli altri che si sono considerati, non ha problemi di identificazione, può essere dunque stimato senza la necessità di imporre vincoli sui parametri. Concludiamo questa presentazione con il modello M7 che è un'estensione del modello CBD.

5. Modello CBD quadratico con effetto coorte (M7)

In Cairns et al. (2009) gli autori hanno esteso la struttura del previsore del modello CBD comprendendo anche un termine quadratico per l'età e l'effetto coorte. La struttura del modello M7 è riassumibile nella seguente tabella.

$$\eta_{x,t} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + ((x - \bar{x}) - \hat{\sigma}_x^2)\kappa_t^{(3)} + \gamma_{t-x}.$$

Nel previsore il termine $\hat{\sigma}_x^2$ è la media empirica della quantità $(x - \bar{x})$.

B.3 Un confronto sui dati italiani 1950 - 2014

In questa sezione si propone un confronto tra i cinque modelli stocastici per descrivere la mortalità introdotti in precedenza. Il modello selezionato come migliore alla luce dell'analisi è quello impiegato nel terzo capitolo per la reverse sensitivity analysis del portafoglio di rendite vitalizie. Consideriamo i dati sulla mortalità della popolazione italiana (fonte: Human Mortality Database (2014)), che riguardano l'intervallo temporale 1872 - 2014 e l'intervallo di età 0 - 110. In linea con il terzo capitolo, siamo interessati esclusivamente alla mortalità di individui con un'età maggiore di 65 anni. Si osservi che in generale dati concernenti la mortalità relativi ad età estreme (maggiori di 89 anni) sono da considerarsi poco affidabili soprattutto qualora riguardino epoche molto passate (si pensi ad esempio alla mortalità dei centenari nel 1900); si propende quindi per considerare per la stima del modello l'intervallo di età 60 - 89. Per quanto concerne l'intervallo temporale consideriamo esclusivamente epoche successive al 1950. La Tabella B.1 presenta il confronto tra le cinque componenti sistematiche.

Un semplice approccio per confrontare i diversi modelli stocastici per la descrizione della mortalità della popolazione richiede l'uso di criteri informativi. Sebbene, come è naturale, molto spesso modelli con più parametri garantiscono un accostamento migliore, può accadere che questi siano sovrapparametrizzati. Per un confronto in termini relativi tra i modelli si utilizzano dunque indici che tengano conto, in qualche modo, del numero di parametri

Confronto tra previsori	
LC	$\eta_{x,t} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)}$
RH	$\eta_{x,t} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)} + \gamma_{t-x}$
APC	$\eta_{x,t} = \alpha_x + \kappa_t^{(1)} + \gamma_{t-x}$
CBD	$\eta_{x,t} = \kappa_t^{(1)} + (x - \bar{x}) \kappa_t^{(2)}$
M7	$\eta_{x,t} = \kappa_t^{(1)} + (x - \bar{x}) \kappa_t^{(2)} + ((x - \bar{x}) - \hat{\sigma}_x^2) \kappa_t^{(3)} + \gamma_{t-x}$

Tabella B.1: Confronto tra le componenti sistematiche dei cinque modelli considerati nello studio.

impiegati dal modello. Due dei più famosi criteri informativi sono il Criterio di Informazione di Akaike e il Criterio di Informazione Bayesiano. Essi sono così definiti:

$$AIC = 2k - 2 \log(\hat{\mathcal{L}}), \quad BIC = \log(n)k - 2 \log(\hat{\mathcal{L}}),$$

ove k è il numero di parametri del modello, n la grandezza del campione e $\hat{\mathcal{L}}$ il valore della funzione di verosimiglianza nel suo massimo. Per procedere operativamente con il confronto numerico stimiamo dunque i 5 modelli. Per la stima è impiegata nella massima verosimiglianza (B.8) una matrice di pesi $\omega_{x,t}$. Questa matrice, come si è già menzionato, contiene solo 0 o 1 a seconda che l'osservazione sia rispettivamente esclusa o inclusa nella verosimiglianza del modello. Sono escluse in generale le prime e ultime tre coorti ai due estremi del dataset. Per operare un confronto significativo, utilizziamo per ogni modello per la componente stocastica la distribuzione binomiale e come funzione di collegamento il logit. Si osservi che, in questo caso, è necessario trasformare (tramite un'approssimazione) i numeri centrali di esposti al rischio $E_{x,t}^c$ per i quali si dispone dei dati, in numeri iniziali di esposti al rischio $E_{x,t}^0$. Nella Tabella B.2 è rappresentato il confronto dei due criteri informativi scelti e del numero di parametri dei modelli.

	LC	RH	APC	CBD	M7
AIC	37 499 (5)	24 645 (2)	28 534 (3)	36 905 (4)	23 545 (1)
BIC	38 184 (5)	25 814 (2)	29 537 (3)	37 629 (4)	25 104 (1)
N° Parametri	123	210	180	130	280

Tabella B.2: Confronto in termini di criteri informativi e numero di parametri per i cinque modelli presi in considerazione.

Sebbene il modello M7 abbia più parametri, risulta essere il migliore sia per AIC che per BIC. Il secondo miglior modello è quello di Renshaw - Haberman; come si può osservare i due criteri informativi hanno prodotto il medesimo ranking.

La bontà di adattamento di questi modelli è tipicamente valutata a partire dai residui. Si considerano in particolare i residui di devianza scalati che, se il modello è adatto alla descrizione dei dati, non dovrebbero presentare

particolari andamenti. Essi sono così definiti:

$$r_{x,t} = \text{sign}(d_{x,t} - \hat{d}_{x,t}) \sqrt{\frac{\text{dev}(x,t)}{\hat{\phi}}}, \quad \hat{\phi} = \frac{D(d_{x,t}, \hat{d}_{x,t})}{K - \nu}, \quad (\text{B.10})$$

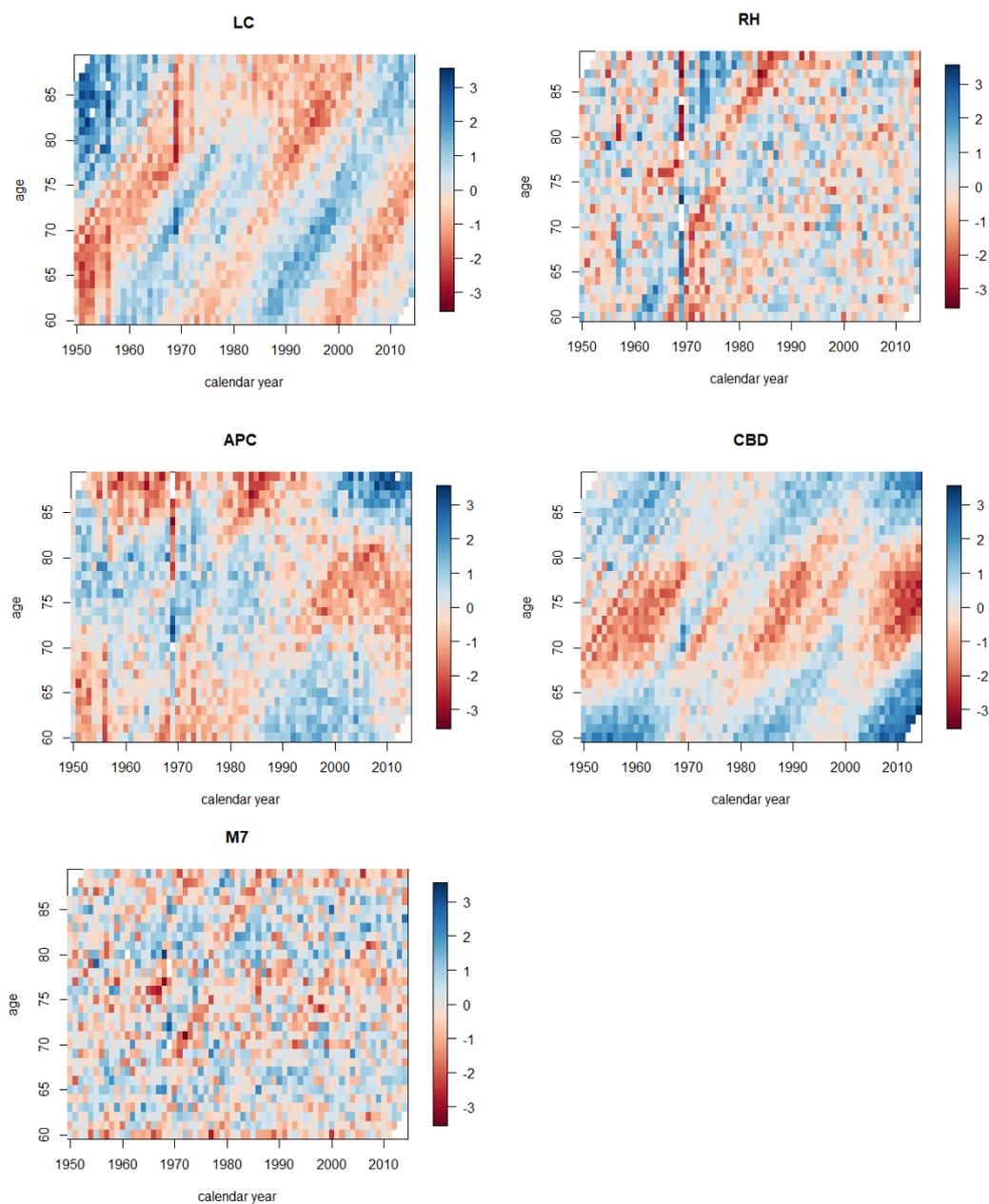


Figura B.1: Confronto tramite mappa di calore dei residui di devianza scalati dei 5 modelli.

ove è posto:

$$\text{dev}(x, t) = 2 \left\{ d_{x,t} \log \frac{d_{x,t}}{\hat{d}_{x,t}} + (E_{x,t}^0 - d_{x,t}) \log \left(\frac{E_{x,t}^0 - d_{x,t}}{E_{x,t}^0 - \hat{d}_{x,t}} \right) \right\},$$

$$D(d_{x,t}, \hat{d}_{x,t}) = \sum_x \sum_t \omega_{x,t} \text{dev}(x, t), \quad K = \sum_x \sum_t \omega_{x,t}.$$

e dove infine ν è il numero di parametri del modello. Procediamo quindi tramite un confronto grafico. Controlliamo inizialmente i residui di devianza scalati dei modelli tramite mappe di calore, si veda a tal fine la Figura B.1. Nei modelli LC, APC e CBD si può osservare un chiaro andamento nei residui; alcune zone presentano residui dello stesso segno ad evidenziare che il modello non si adegua bene ai dati della popolazione italiana. Come si era già osservato dai criteri informativi, i due modelli migliori in termini di

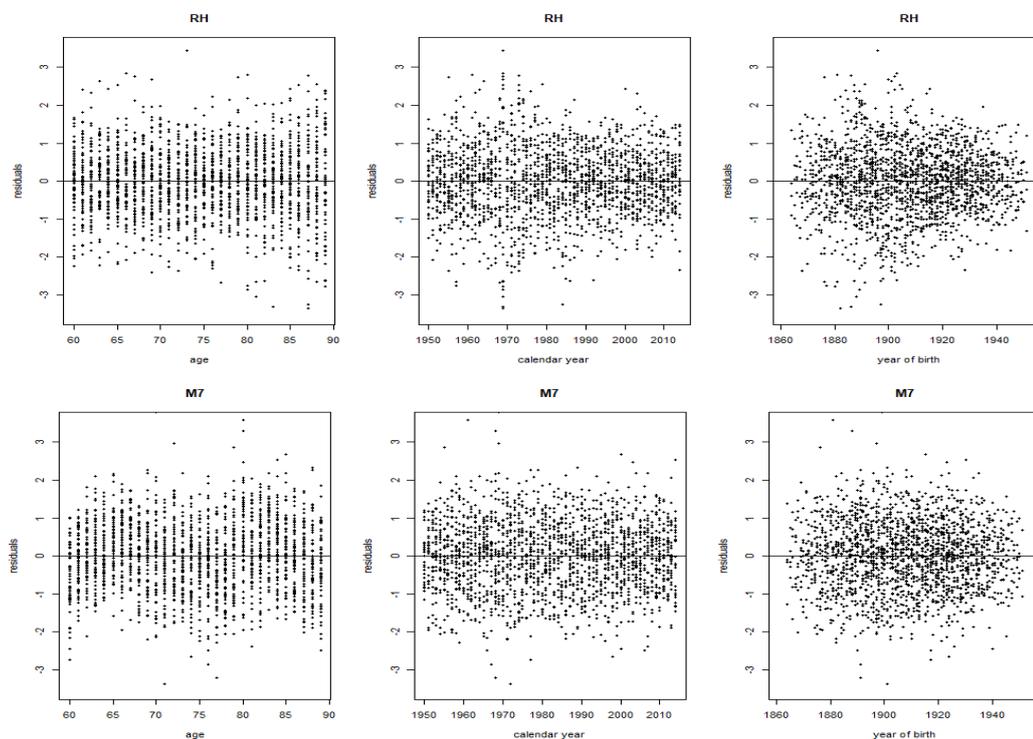


Figura B.2: Confronto tra residui di devianza scalati contro età, anno di calendario e anno di nascita per il modello RH (sopra) e il modello M7 (sotto).

adattamento sono anche in questo caso il modello M7 e quello di Renshaw - Haberman (RH). Tra i due, in particolare, il modello M7 sembra essere il preferibile. Procediamo con un'indagine sui residui più dettagliata andando a confrontare i grafici marginali per età, anno di calendario e anno di nascita, si veda la Figura B.2. Per entrambi i modelli i residui non presentano andamenti particolarmente evidenti. Nel modello RH sembra esserci nel grafico relativo all'anno di calendario qualche problema di adattamento intorno al 1970, mentre il grafico relativo all'anno di nascita presenta una leggera asimmetria. Alla luce di quanto evidenziato si decide infine, in ultima istanza, di scegliere M7 come miglior modello. Pertanto esso è impiegato per la descrizione della mortalità degli assicurati del portafoglio di rendite vitalizie del terzo capitolo.

Bibliografia

- [1] Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking and Finance*, 26 (7), 1505-1518.
- [2] Ali, S. M., & Silvey, S. D. (1966). A general class of coefficients of divergence of one distribution from another. *Journal of the Royal Statistical Society. Series B Methodological*, 131-142.
- [3] Artzner, P., Delbaen, F., Eber, J.-M., Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9 (3), 203-228.
- [4] Ash, R.B. (1972). *Real Analysis and Probability*, Academic Press, New York.
- [5] BCBS (2008). Principles for sound stress testing practices and supervision. Technical Report. Basel Committee on Banking Supervision, Bank for International Settlements.
- [6] BCBS (2010). Basel III: A global regulatory framework for more resilient banks and banking systems. Technical Report. Basel Committee on Banking Supervision, Bank for International Settlements.
- [7] BCBS (2012). Fundamental review of the trading book. Technical Report. Basel Committee on Banking Supervision, Bank for International Settlements.
- [8] BCBS (2013). Fundamental review of the trading book: A revised market risk framework. Technical Report. Basel Committee on Banking Supervision, Bank for International Settlements.

- [9] BCBS (2017). Supervisory and bank stress testing: range of practices. Technical Report. Basel Committee on Banking Supervision, Bank for International Settlements.
- [10] Billingsley, P. (1995). Probability and measure. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition.
- [11] Borgonovo, E., Plischke, E. (2016). Sensitivity analysis: A review of recent advances. *European Journal of Operational Research*, 248 (3), 869-887.
- [12] Brouhns, N., Denuit, M., Vermunt, J. (2002). A Poisson Log-Bilinear Regression Approach to the Construction of Projected Lifetables. *Insurance: Mathematics and Economics*, 31 (3), 373-393.
- [13] Brigo, D., Mercurio, F. (2006). *Interest Rate Models - Theory and Practice*. Springer Finance. Springer-Verlag, Berlin, second edition, 2006.
- [14] Cairns, AJG., Blake, D., Dowd, K. (2006). A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty: Theory and Calibration. *Journal of Risk and Insurance*, 73(4), 687-718.
- [15] Cambou, M. , & Filipovic, D. (2017). Model uncertainty and scenario aggregation. *Mathematical Finance*, 27 (2), 534-567.
- [16] Chang, C., Jimenez-Martin, J.A., Maasoumi, E., McAleer, M., Amaral, T. (2017). Choosing Expected Shortfall over VaR in Basel III Using Stochastic Dominance.
- [17] Cooke, R. M. , & van Noortwijk, J. M. (1999). Local probabilistic sensitivity measures for comparing FORM and Monte Carlo calculations illustrated with dike ring reliability calculations. *Computer Physics Communications*, 117 (1-2), 86-98 .
- [18] Cover, T. M. , & Thomas, J. A. (2012). *Elements of information theory*. John Wiley & Sons.

- [19] Csiszár, I. (1975). I-divergence geometry of probability distributions and minimization problems. *The Annals of Probability*, 3 (1), 146-158.
- [20] Currie, ID. (2006). Smoothing and Forecasting Mortality Rates with P-Splines.
- [21] Currie, ID. (2016). On Fitting Generalized Linear and non-linear Models of Mortality. *Scandinavian Actuarial Journal*, (4), 356-383.
- [22] Dhaene, J., Vanduffel, S., Tang, Q., Goovaerts, M. J., Kaas, R., Vyncke, D. (2006). Risk Measures and Comonotonicity: A Review.
- [23] EIOPA (2009). Directive 2009/138/EC of the european parliament and of the council. Technical Report. European Insurance and Occupational Pensions Authority.
- [24] EIOPA (2010). QIS5 Technical Specifications. European Insurance and Occupational Pensions Authority.
- [25] EIOPA (2018). Insurance Stress Test 2018 Technical specifications. European Insurance and Occupational Pensions Authority.
- [26] Embrechts, P., McNeil, A. and Straumann, D. (2002) Correlation and Dependence in Risk Management: Properties and Pitfalls. *Risk Management: Value at Risk and Beyond*. 1st Edition, Cambridge University Press, Cambridge, 176-223.
- [27] Föllmer, H., Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, 6 (4), 429-447.
- [28] Girosi F, King G. (2007), *Demographic forecasting*. Cambridge: Cambridge University Press.
- [29] Kullback, S., Leibler, R.A. (1951). On information and sufficiency. *Annals of Mathematical Statistics*. 22 (1): 79-86
- [30] Kullback, S. (1959). *Information theory and statistics*. John Wiley and Sons.

- [31] Haberman S, Renshaw A (2011). A Comparative Study of Parametric Mortality Projection Models. *Insurance: Mathematics and Economics*, 48(1), 35-55.
- [32] Haugh M. (2016). An introduction to copulas. *Quantitative Risk Management*. Springer 2016.
- [33] Holzer S. (2007). Quaderni della scuola di dottorato di ricerca in finanza nell'Università degli Studi di Trieste. EUT - Edizioni Università di Trieste.
- [34] Holzer S. (2016). Probabilità elementare. EUT - Edizioni Università di Trieste.
- [35] Hunt A., Blake D. (2015). On the Structure and Classification of Mortality Models Mortality Models. Pension Institute Working Paper.
- [36] IAA (2013). Stress Testing and Scenario Analysis. International Actuarial Association.
- [37] IVASS. (2016). Solvency II The new prudential regulation of the insurance sector: a simplified guide.
- [38] Lee RD., Carter LR. (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association*, 87 (419), 659 - 671.
- [39] McCullagh P., Nelder J (1989). *Generalized Linear Models*. second edition. Chapman & Hal, London.
- [40] MacKay, D.J.C. (2003). *Information Theory, Inference, and Learning Algorithms* (First ed.). Cambridge University Press. p. 34.
- [41] McNeil, A., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, Princeton.
- [42] Olivieri, A., Pitacco, E. (2003). Solvency requirements for pension annuities.

- [43] Pesenti, S. M., Millosovich, P., Tsanakas, A. (2019). Reverse sensitivity testing: What does it take to break the model?
- [44] Pitacco E., M. Denuit, S. Haberman, A. Olivieri (2009). *Modelling Longevity Dynamics for Pensions and Annuity Business*. Oxford University Press, 2009
- [45] Saltelli, A. , Ratto, M. , Andres, T. , Campolongo, F. , Cariboni, J. , Gatelli, D. , et al. (2008). *Global sensitivity analysis: The primer*. John Wiley & Sons.
- [46] Sklar, A. (1959). Fonctions de repartition ‘a n dimensions et leurs marges. *Publications de l’Institut de Statistique de l’Universite de Paris* 8:229-231.
- [47] Vasicek, O.A (1977). An equilibrium characterization of the term structure. *Journal of financial economics*, 5: 177-88.
- [48] Villegas, A., Kaishev, V.K., Millosovich, P. (2018). *StMoMo: An R Package for Stochastic Mortality Modelling*.