

**UNIVERSITÀ DEGLI STUDI DI ROMA “LA SAPIENZA”  
FACOLTÀ DI SCIENZE STATISTICHE  
DIPARTIMENTO DI SCIENZE ATTUARIALI E FINANZIARIE**

**DOTTORATO DI RICERCA IN SCIENZE ATTUARIALI - XIX CICLO  
SETTORE SCIENTIFICO DISCIPLINARE SECS-S/06**

**TESI DI DOTTORATO**

**Il filtro di Kalman e la valutazione stocastica  
della riserva sinistri**

Coordinatore:  
**Prof. Fabio Grasso**

Candidata:  
**Chiara Parrini**

# INDICE

<b>RIASSUNTO</b>	<b>1</b>
<b>INTRODUZIONE</b>	<b>2</b>
<b>CAPITOLO 1</b>	<b>4</b>
<b>LO STATO ATTUALE DELL'ARTE</b>	<b>4</b>
1.1 Aspetti giuridici.....	4
1.2 Il rischio, la solvibilità e la stabilità di un'impresa assicurativa danni.....	6
1.3 La valutazione della riserva sinistri in Italia.....	8
1.4 Una piccola rivoluzione: i principi IAS e il progetto Solvency II, il primo passo verso un'ampia revisione della vigilanza prudenziale.....	9
1.5 Lo sviluppo dei principi contabili per il settore assicurativo.....	16
1.6 I principi IAS e la riserva sinistri .....	18
<b>CAPITOLO 2</b>	<b>20</b>
<b>LE RISERVE TECNICHE NEI RAMI DANNI</b>	<b>20</b>
2.1 Premesse.....	20
2.2 La riserva premi.....	20
2.3 L'evoluzione dei sinistri e la riserva sinistri .....	24
2.4 Scomposizione della riserva sinistri .....	27
2.5 La rappresentazione dei dati: le coorti di sinistri.....	28
2.6 La base dati: l'informazione triangolare .....	30
<b>CAPITOLO 3</b>	<b>33</b>
<b>MODELLI DETERMINISTICI PER LA STIMA         DELLA RISERVA SINISTRI</b>	<b>33</b>
3.1 Premesse.....	33
3.2 Il metodo della catena.....	35
3.3 Il metodo della separazione.....	39
3.4 Il metodo dei minimi quadrati.....	42

3.5 Il metodo di Bornhuetter-Ferguson .....	44
3.6 Il metodo di Fisher-Lange .....	45
3.7 Il metodo di Tarbell per la stima della riserva IBNR.....	48
<b>CAPITOLO 4</b>	<b>50</b>
<b>MODELLI STOCASTICI PER LA STIMA DELLA RISERVA SINISTRI</b>	<b>50</b>
4.1 Premesse.....	50
4.2 Modelli Chain-Ladder.....	53
4.2.1 Introduzione.....	53
4.2.2 Modelli lineari e tecniche chain-ladder.....	55
4.2.3 I modelli lineari generalizzati.....	58
4.2.4 La curva di Hoerl.....	60
4.3 Stime baynesiane e tecniche di credibilità.....	62
4.3.1 Introduzione.....	62
4.3.2 Stime baynesiane e tecniche chain-ladder.....	63
4.4 Tecniche dinamiche di stima.....	65
4.4.1 Modelli state-space.....	65
4.4.2 Rappresentazione di modelli chain-ladder con modelli state-space.....	67
4.4.3 Stime dinamiche dei parametri riga.....	69
4.4.4 Stime dinamiche dei parametri colonna.....	71
4.4.5 Alcune considerazioni.....	72
4.5 Altri metodi di stima.....	74
4.5.1 Introduzione.....	74
4.5.2 Il modello di Mack.....	76
4.5.3 Il modello di Wright.....	79
4.6 Modelli di smoothing .....	83
4.6.1 Premesse.....	83
4.6.2 Dai modelli di smoothers ai modelli additivi.....	85
4.6.3 I modelli additivi ed i modelli additivi generalizzati.....	86
4.6.4 I modelli di smoothers applicati alla stima della riserva sinistri.....	88
<b>CAPITOLO 5</b>	<b>92</b>
<b>IL FILTRO DI KALMAN</b>	<b>92</b>
5.1 Premesse.....	92
5.1.1 Il prediction error.....	93

5.2 Il filtro di Kalman.....	94
5.2.1 L'algoritmo di calcolo.....	96
5.2.2 Il funzionamento dell'algoritmo di calcolo.....	97
5.2.3 I parametri del filtro.....	98
5.3 La stima della riserva sinistri mediante il filtro di Kalman: il modello.....	99
5.4 Risultati e confronti con altri metodi.....	101
5.5 L'impostazione del problema secondo il progetto Solvency II:	
il contesto per la definizione e il calcolo dei requisiti di riserva.....	111
5.5.1 Risk margin e requisiti di riserva.....	113
5.5.2 Risk margin: risultati.....	114
5.5.3 Un confronto coi requisiti di capitale nei QIS2.....	117
<b>CONCLUSIONI</b> .....	122
<b>BIBLIOGRAFIA</b> .....	127

## RIASSUNTO

Il presente lavoro ha l'obiettivo di partire dall'analisi dei più tradizionali metodi di valutazione della riserva sinistri per giungere allo studio di metodologie di calcolo più complesse ed accurate, quali i modelli stocastici. Inizialmente vengono presentate le principali metodologie presenti in letteratura prima di verificarne l'effettiva capacità di adattamento ai dati tramite alcuni metodi specifici tra cui, in particolare, il filtro di Kalman. Tale filtro sembra adattarsi bene alla stima della riserva sinistri: la best estimate si pone sufficientemente in linea con i risultati prodotti mediante l'utilizzo del metodo deterministico del chain-ladder e, conformemente a quanto avviene per altri modelli stocastici, il prediction error assume un valore più contenuto di quello ottenuto con il modello di Mack. L'utilizzo di diversi scenari, sia relativi a casi di incrementi di variabilità negativi sia a casi di incrementi di variabilità positivi, ha permesso di verificare l'effetto del cambiamento della variabilità di processo sul valore del prediction error. Le analisi hanno confermato quanto originariamente presupposto, ovvero quanto i valori inizialmente assegnati alle matrici di covarianza possono impattare sulle stime ottenute con il modello. Ovviamente la maggiore o minore variabilità presente nei dati ha il suo impatto anche sul calcolo del risk margin. Sfruttando l'ipotesi di lognormalità per la distribuzione di probabilità della riserva sinistri si è potuto constatare che il filtro di Kalman è in grado di cogliere anche questo aspetto. In ultima analisi è stato effettuato un confronto con il requisito di capitale secondo le regole previste dal QIS2 e utilizzando la best estimate ed il risk margin ottenuto attraverso l'applicazione del filtro di Kalman.

## INTRODUZIONE

Una rilevante fonte di rischio per le compagnie di assicurazione contro i danni è costituita dalla possibile sottostima degli impegni di rimborso in essere, da cui la conseguente insufficienza della riserva sinistri. Le tecniche attuariali sono da tempo utilizzate al fine di effettuare le valutazioni della riserva sinistri la cui entità risulta, per sua natura, un importo stimato. Le stime basate sui metodi di inventario (vale a dire, le valutazioni sinistro per sinistro) ed utilizzate dalle compagnie al fine di rappresentare il costo ultimo dei sinistri sembrano in generale essere inadeguate, ad esempio per quei rami caratterizzati da un più lungo processo di liquidazione dei sinistri. L'introduzione del Basilea II relativo al settore bancario, i nuovi Principi Contabili Internazionali (IAS) e, infine, il progetto Solvency II apporteranno inoltre sostanziali modifiche al modo in cui gli operatori del mercato finanziario ed assicurativo gestiranno alcuni aspetti tradizionali (quali, ad esempio, la stima di alcune voci di bilancio e la valutazione del margine di solvibilità). Questi progetti hanno lo scopo di modificare l'intero settore finanziario-assicurativo cercando una maggiore trasparenza e rispondenza con il mercato.

Uno degli argomenti più dibattuti negli ultimi anni, nella letteratura attuariale riferita alle assicurazioni contro i danni, riguarda l'impiego di metodologie di calcolo di tipo deterministico o stocastico per la valutazione della riserva sinistri. Le prime hanno il vantaggio di essere facilmente e rapidamente implementate sotto il profilo computazionale. In particolare, esse necessitano di minori vincoli sui dati di base e, soprattutto, risultano di immediata comprensione a livello manageriale di impresa. Per tale motivo gli attuari, in Italia ed all'estero, hanno in genere preferito i metodi deterministici rispetto a quelli stocastici. Tuttavia, a seguito del progetto Solvency II, l'impiego di metodologie stocastiche dovrà diffondersi sempre di più. Tra i vari aspetti, infatti, il CEIOPS (Committee of European Insurance and Occupational Pension Supervisors), su richiesta della Commissione Europea, ha posto come obiettivo il raggiungimento di un benchmark armonizzato in chiave probabilistica quale requisito regolamentare per la riserva sinistri.

Questo obiettivo può essere raggiunto solamente mediante l'utilizzo di un modello di valutazione di tipo stocastico alla serie storica dei sinistri. Soltanto i modelli stocastici, infatti, nonostante una maggiore complessità applicativa ed ipotesi più impegnative sui dati di base, permettono di conseguire, oltre alla stima puntuale della riserva (best estimate), intervalli di variazione della stessa secondo un prefissato livello di probabilità. Le tecniche deterministiche, invece, restituiscono soltanto un valore puntuale della previsione, trascurando ogni possibile valutazione sulla probabilità del suo realizzarsi.

Il presente lavoro ha pertanto l'obiettivo di partire dall'analisi dei più tradizionali metodi di valutazione della riserva sinistri per giungere allo studio di metodologie di calcolo più accurate, quali i modelli stocastici. La parte fondamentale della tesi riguarda la presentazione di alcune tra le principali metodologie stocastiche di stima e la conseguente verifica dell'effettiva loro capacità di adattamento ai dati empirici. In particolare la ricerca si propone di esaminare un particolare modello stocastico, il filtro di Kalman. Sviluppato negli anni Cinquanta nel campo ingegneristico, il filtro di Kalman ha avuto un utilizzo sempre maggiore nel tempo grazie allo sviluppo crescente di adeguate capacità e tempi di calcolo degli elaboratori. L'interesse verso l'impiego di tale metodologia, che negli ultimi anni è comparsa anche in alcune valutazioni di tipo finanziario (dalla stima della volatilità implicita alla previsione dei depositi bancari alla stima del value at risk), trova giustificazione nella sua elevata efficienza: infatti, il metodo consiste in un insieme di equazioni matematiche in grado di fornire una soluzione efficiente del problema dei minimi quadrati. In ambito assicurativo il suo utilizzo è tuttavia ancora modesto, probabilmente a causa delle difficoltà di implementazione, e la letteratura attuariale non dispone, in particolare, di molti studi relativi all'applicazione di tale metodologia alla stima della riserva sinistri. Nel presente lavoro è pertanto esaminata la bontà di stima che questo algoritmo di calcolo può apportare sulla valutazione della riserva sinistri. In funzione dei risultati ottenuti viene infine effettuato uno studio comparato con altre note metodologie di calcolo, diffusamente presenti in letteratura e frequentemente in uso nella pratica assicurativa.

## CAPITOLO 1

### LO STATO ATTUALE DELL'ARTE

#### 1.1 Aspetti giuridici

Ogni impresa di assicurazione è tenuta, per legge, a presentare una valutazione comprensiva dei risultati di tutti i movimenti di capitale, sia in entrata che in uscita, verificatisi nel corso di ogni anno di esercizio. Tale valutazione viene effettuata annualmente ad una prefissata data che coincide con il 31 dicembre ed il documento che riassume i vari spostamenti di denaro è il bilancio di esercizio. Le regole in base alle quali si ha la redazione dei bilanci sono contenute nel D.lg. n.173 che ha dato attuazione alla Direttiva n. 91/674 sui conti annuali e consolidati delle imprese di assicurazione. Le stime riassunte nel bilancio di esercizio vengono esaminate dalle autorità di controllo e sono affiancate dalla relazione di una società di revisione, per legge iscritta all'albo, tenuto presso la Commissione nazionale per le società e la borsa (Consob), e tra i cui amministratori figurano almeno un attuario. Quest'ultimo ha il compito di certificare la stabilità della compagnia e il rispetto delle norme giuridiche di riferimento; successivamente il documento viene reso pubblico. In realtà di solito si hanno redazioni di bilanci semestrali, trimestrali o addirittura mensili e l'approvazione deve essere effettuata entro il 30 aprile dell'anno successivo a quello cui si riferisce il bilancio stesso, termine questo che può essere prorogato fino al 30 giugno qualora particolari esigenze lo richiedano (ad esempio se l'impresa esercita anche attività riassicurative in misura rilevante).<sup>1</sup> Congiuntamente al documento, le imprese devono depositare il prospetto dimostrativo della situazione del margine di solvibilità e il prospetto delle attività a copertura delle riserve tecniche. Le compagnie di assicurazione sono infatti soggette al rischio di insolvenza come del resto le imprese di tutti i settori; tuttavia in questo ambito, poiché il business consiste nell'assunzione di rischi, è evidente l'importanza ancora più rilevante che assume il tema della solvibilità.

---

<sup>1</sup> Art 11 d.lgs. 26 maggio 1997, n.173, comma 1-3.



A tale proposito, le compagnie utilizzeranno i premi, incassati anticipatamente e in parte accantonati in fondi di riserva (le riserve tecniche), per far fronte agli impegni assunti con gli assicurati. A tutela degli assicurati, oltre alle riserve tecniche, le compagnie devono disporre anche del cosiddetto “margine di solvibilità”, fondo di capitale libero da impegni che ha la funzione di garanzia in quanto consiste nel mantenere una eccedenza di attività rispetto alle passività. Un ulteriore garanzia degli assicurati è attuata mediante il controllo da parte dello Stato, il quale ripone nella figura delle autorità di vigilanza, in Italia l’ISVAP, la salvaguardia delle gestioni assicurative. Nel 1973 fu emanata la prima direttiva della Comunità economica europea per il margine di solvibilità dei rami danni. Questa direttiva prevedeva un margine minimo pari a una percentuale o dei premi o della media dei sinistri degli ultimi tre o sette anni. Da qualche anno si è avvertita sempre di più la necessità di un’informazione finanziaria comparabile e trasparente, in quanto ormai quasi tutte le società si trovano ad operare su scala internazionale (basta pensare all’Unione Europea). In proposito, anche a seguito dell’entrata in vigore della moneta unica europea, si è accentuata l’esigenza di realizzare un mercato dei capitali europeo di tipo integrato, tale da garantire competitività all’interno dei diversi settori ed essere quindi maggiormente attraente nei confronti dei possibili acquirenti di prodotti, soprattutto di tipo finanziario-assicurativo. L’obiettivo è quello di realizzare un quadro di principi contabili internazionali comparabili a livello internazionale con il fine ultimo di fronteggiare l’accelerata globalizzazione dei commerci e l’internazionalizzazione dei mercati finanziari. Non è poi da tralasciare la possibilità di garantire una maggiore trasparenza a seguito di norme di gestione più rigide in grado di ridurre la probabilità di scandali aziendali, che ultimamente stanno compromettendo molto i mercati e che quindi potrebbero far sì che l’investitore privato torni ad avere nelle imprese un grado di fiducia abbastanza elevato. Per raggiungere questo obiettivo la Commissione Europea ha emanato il Regolamento CE n. 1606/2002 che ha imposto l’obbligo di adottare, a partire dal 2005, i principi contabili internazionali per la redazione del bilancio alle società i cui titoli siano negoziati in mercati regolamentati, lasciando però ai singoli stati membri la facoltà o l’obbligo di adottare questi principi anche per la redazione del bilancio d’esercizio.

Al di là dell'Unione Europea, molti altri paesi hanno dichiarato la volontà di aderire ai nuovi principi o almeno di armonizzare ad essi le norme già esistenti, per garantire una completa compatibilità di queste ultime. Si registra quindi una tendenza verso una convergenza che può definirsi globale. Il Canada e l'Australia, ad esempio, hanno dichiarato di voler sostituire i principi contabili con quelli internazionali, mentre gli standard setters giapponesi (cioè i soggetti che definiscono i principi a livello nazionale) cercano di armonizzare nel lungo periodo i principi nazionali al fine di avere norme comparabili. La Swiss Exchange (Borsa svizzera) chiederà alle società quotate di conformarsi agli Standards, mentre il Financial Accounting Standards Board degli USA ha deciso di armonizzare, prima possibile, i principi USA con le norme contabili internazionali.

## **1.2 Il rischio, la solvibilità e la stabilità di un'impresa assicurativa danni**

Una caratteristica peculiare dell'attività assicurativa consiste nell'assunzione dei rischi. È evidente quindi l'importanza fondamentale di una corretta gestione del rischio stesso, aspetto imprescindibile per il raggiungimento della stabilità d'impresa. Uno dei principali obiettivi del management, degli azionisti e degli assicurati rientra nella sopravvivenza dell'impresa a cui si aggiunge la previsione di possibili crisi aziendali. Tale obiettivo è perseguito anche dalle autorità di vigilanza come tutela degli interessi degli assicurati e dei possibili danneggiati. La possibilità di sopravvivenza, anche nota come probabilità di rovina, di un'impresa di assicurazione viene esaminata sotto due aspetti. Il primo riguarda l'esposizione totale dell'impresa alle varie fonti di rischio, che siano esse di natura tecnica, finanziaria, o più semplicemente situazioni generiche di rischio. L'altro aspetto cruciale consiste nella disponibilità di un adeguato capitale di sicurezza sufficiente a fronteggiare perdite inattese e superare pertanto situazioni di crisi. Occorre osservare come l'inversione del ciclo produttivo, per cui i ricavi precedono i costi, caratteristica tipica dell'impresa di assicurazione, comporta maggiori elementi di rischiosità rispetto alle altre imprese industriali. In proposito le compagnie di assicurazione utilizzano opportune metodologie statistiche ed attuariali e la massa dei premi introitati deve essere accantonata, costituendo le cosiddette riserve tecniche, al fine di tutelarsi dai

futuri esborsi, che siano essi per sinistri o per costi di gestione. Come diretta conseguenza, i principali fattori di rischio risiedono soprattutto nelle passività di un'impresa di assicurazione. Le basi statistiche e finanziarie sottostanti le stime dei premi possono infatti non riflettere le previsioni ed essere pertanto smentite dalla realtà, così come le riserve tecniche accantonate possono risultare inadeguate. La classificazione dei rischi presenta margini di discrezionalità, tuttavia i rischi possono essere definiti come le conseguenze derivanti dal verificarsi di eventi aleatori. Queste conseguenze possono essere favorevoli (upside risk) o sfavorevoli (downside risk), ossia possono produrre benefici o svantaggi. Nell'uso comune si è soliti collegare la nozione di rischio a quella di danno economico, vale a dire al downside risk. Una buona definizione di rischio pubblicata nel 1995 dagli Standards Australia e Standards New Zealand è la seguente:

*“Il Rischio è la possibilità che accada qualcosa che avrà un impatto sugli obiettivi da raggiungere. E' misurato in termini di conseguenze e probabilità”.*

L'analisi delle situazioni di rischio che possono influire sulla sopravvivenza di un'impresa di assicurazione abbiamo visto essere strettamente collegata anche al problema della solvibilità e della stabilità finanziaria. Negli ultimi decenni questi aspetti hanno assunto un ruolo sempre più predominante e sono diventati oggetto di studi intensivi. Ciò si è reso possibile anche grazie al miglioramento delle conoscenze teoriche ed empiriche e, non ultimo, per i progressi informatici.

La solvibilità può essere definita come la capacità di far fronte agli impegni assunti. A tal fine i legislatori hanno stabilito leggi ed istituito organi di controllo, mentre l'impresa, oltre alle varie misure preventive, ricorre all'utilizzo di capitale proprio ed alla riassicurazione. Al fine di garantire una maggiore tutela agli assicurati circa l'adempimento degli impegni assunti dalla compagnia è stato inoltre introdotto il margine di solvibilità, una sorta di cuscinetto di sicurezza. Tuttavia va comunque sottolineato come la prima misura per salvaguardare la solvibilità delle imprese di assicurazione non è rappresentato dal margine, bensì dalle riserve tecniche: effettuare un costante controllo e una valutazione corretta delle riserve tecniche rimane certamente un aspetto primario.

Considerare semplicemente il margine minimo di solvibilità, che tuttavia rappresenta un mezzo importante per la stabilità di un'impresa, trascurando una corretta valutazione delle riserve tecniche non risulterebbe adeguatamente sufficiente.

### **1.3 La valutazione della riserva sinistri in Italia**

L'articolo 33 del D.lgs. 173/97 comma 1 prevede che le imprese di assicurazione debbono costituire la riserva sinistri alla fine di ogni anno di esercizio. A tale proposito deve essere riportato nel bilancio l'ammontare complessivo delle somme che, da una prudente ed accurata valutazione effettuata in base ad elementi oggettivi, risultino necessarie per far fronte al pagamento dei sinistri avvenuti nell'esercizio stesso o in quelli precedenti e non ancora pagati. In aggiunta, secondo l'articolo 33 del D.lgs. 173/97 comma 4, le imprese di assicurazione hanno l'obbligo di costituire una riserva anche per sinistri avvenuti ma non ancora denunciati alla data di chiusura dell'esercizio (riserva per sinistri IBNR). Lo stesso articolo del D.lgs. 173/97 (comma 2) suggerisce la metodologia di calcolo della riserva sinistri, che deve essere valutata in misura pari al costo ultimo in modo tale da tener conto di tutti i futuri oneri prevedibili, sulla base dei dati storici e prospettici e comunque delle specifiche caratteristiche dell'impresa. In linea di principio, la riserva sinistri (art. 33 del D.lgs. 173/97 comma 3) deve essere valutata separatamente per ciascun sinistro. In proposito, poiché (circolare ISVAP n.360/99) il sistema dell'inventario, per buona parte dei rami danni ed in particolare per quelli caratterizzati da lunghi processi liquidativi, non consente di pervenire con ragionevole attendibilità al costo ultimo finale, sarà compito delle strutture direzionali dell'impresa l'utilizzo delle risultanze degli smontamenti nel tempo nonché di metodi statistico-attuariali adeguati per il completamento delle stime di inventario. Un concetto basilare ai fini della valutazione delle riserve tecniche è quello di "sufficienza". L'importo delle riserve tecniche deve essere sempre sufficiente a consentire all'impresa di far fronte, per quanto ragionevolmente prevedibile, agli impegni derivanti dai contratti di assicurazione.

È compito dell'attuario (articolo 2 del Provvedimento ISVAP n. 845/1998) esprimere il proprio giudizio in merito alla sufficienza delle riserve tecniche iscritte nel passivo dello stato patrimoniale del bilancio di esercizio. Tale giudizio deve essere formulato in conformità alle vigenti disposizioni di legge e deve rispettare corrette tecniche attuariali che riguardano, in particolare, l'impiego di adeguate basi tecniche, l'impiego di ipotesi evolutive prudenziali, e l'impiego di adeguate metodologie di calcolo. Gli attuari non-life hanno proposto vari modelli, più o meno sofisticati, con il fine ultimo di pervenire a stime ragionevolmente adeguate per le riserve tecniche. L'impiego di un metodo è notevolmente influenzato dal tipo e dall'entità dell'informazione disponibile. In particolare la scelta tra i vari metodi è influenzata dalle particolarità del portafoglio assicurativo, come ad esempio il tipo di rischio assicurato, le dimensioni del portafoglio, la base dati a disposizione, etc. Tuttavia una caratteristica comune alla maggior parte dei metodi utilizzati risiede nel risultato finale, il quale consiste in un valore puntuale di stima. Per tale motivo tali metodi possono risultare inadeguati rispetto all'esigenza di ottenere valori sulla base di elementi oggettivi che non ne limitino la "prudenzialità". Risulta difficile infatti dare limiti al range di tutte le possibili ipotesi plausibili per la previsione dei possibili scenari futuri realizzabili. È evidente come le caratteristiche delle supposizioni siano fortemente influenzate da un accentuata soggettività tipica di chi le effettua. Come meglio vedremo in seguito l'esigenza di una maggiore oggettività diviene ancora più sentita anche a seguito dell'introduzione dei nuovi principi contabili internazionali.

#### **1.4 Una piccola rivoluzione: i principi IAS e il progetto Solvency II, il primo passo verso un'ampia revisione della vigilanza prudenziale**

I principi IAS sono stati formulati al fine di creare una contabilità internazionale uniforme per ciascuna attività economica. In conseguenza, il risultato delle imprese potrà essere confrontato a livello internazionale nel rispetto di un logico e naturale sviluppo della prospettiva di "globalizzazione" dei mercati economici. Con i nuovi principi contabili internazionali (IAS) si tende a privilegiare il concetto della competenza anziché il principio della prudenza che, per anni, ha rappresentato il cardine dei principi contabili nazionali così come stabilito dal Codice Civile.

Secondo tale principio nella redazione di bilancio non vi sarà più asimmetria: per quanto riguarda costi e ricavi si contabilizzeranno anche i proventi non realizzati, per quanto riguarda attività e passività si privilegerà il criterio del valore corrente (fair value) e non quello del costo storico. Tale metodo consiste nel valutare le attività e le passività al valore di mercato, metodologia di stima a lungo discussa dagli organi superiori europei di vigilanza e dalla Banca Centrale Europea. L'impatto dell'adozione dei nuovi principi contabili internazionali va analizzato tenendo in considerazione i notevoli cambiamenti del mercato assicurativo europeo negli ultimi anni: basta pensare alla liberalizzazione tariffaria del 1994 e le nuove norme in materia di concorrenza. Gli ultimi venti anni sono stati caratterizzati da risultati tecnici relativi ai rami danni spesso in perdita: dal 1986 al 2001 il combined ratio per l'intero mercato assicurativo non-life ha sempre superato il 100%. A tale proposito si è reso sempre più necessario un cambiamento radicale del sistema di solvibilità. All'inizio del 2004 è entrato in vigore il cosiddetto Solvency I, un primo aggiustamento in vista della riforma totale che sarà introdotta con il Solvency II. Il progetto Solvency II ha come obiettivo rivedere il quadro europeo di vigilanza prudenziale delle imprese di assicurazione. Si tratta di una riforma a 360° del sistema attuale, quindi interesserà non solo i requisiti minimi di capitale, ma anche la vigilanza e l'informativa di mercato. Lo scopo è l'introduzione di un regime di solvibilità che non si limiti all'analisi dei soli "rischi tecnici" ma che li consideri tutti, che sia flessibile, non eccessivamente complesso e compatibile con i metodi di controllo delle autorità di vigilanza. L'aspetto più importante è l'introduzione di requisiti minimi più aderenti ai rischi sopportati dalle compagnie da calcolarsi o attraverso un sistema standard o attraverso modelli interni di gestione del rischio. È evidente come la necessità di creare una disciplina comune sul margine di solvibilità si scontra con le regole di calcolo e di stima delle attività e delle passività, ed in particolare delle riserve tecniche. Ecco allora che il progetto Solvency II non può prescindere dal guardare con attenzione all'introduzione dei nuovi principi contabili internazionali IAS, i quali, tra l'altro, diminuiranno il livello di prudenza in molte nazioni, aspetto che accentua ancora di più l'importanza di un rafforzamento del sistema di solvibilità e di un'analisi sempre più attenta dei vari rischi a cui l'impresa è esposta.

Non ultimo occorre ricordare l'elevato numero di insolvenze verificatesi negli ultimi anni in UK ed i numerosi limiti della direttiva del 1973 in materia di solvibilità nei rami danni, primo fra tutti la non sensibilità del minimo margine all'effettiva rischiosità dell'impresa di assicurazione. In proposito esistono sistemi molto più evoluti rispetto a quello utilizzato oggi in Europa, tra i quali il Risk Based Capital (RBC) americano e il sistema australiano, possibili punti di partenza per individuare un progetto di solvibilità più adeguato ad una corretta gestione dei rischi. Il 5 Marzo 2002 sono state approvate 2 nuove direttive, la n. 2002/12/CE (vita, sostituita poi con la n.2002/83/EC), e la n. 2002/13/CE (danni) in materia di margine di solvibilità. Questa nuova disciplina UE conclude la prima fase di revisione relativa al settore assicurativo ossia il cosiddetto Solvency I. In Italia le modifiche sono state introdotte con il decreto legislativo n. 307 del 3 novembre 2003 e sono entrate in vigore il 1 Gennaio 2004. Il progetto Solvency II è invece suddiviso in due fasi distinte: la prima, avviata nel 2001 e conclusasi due anni dopo, è sostanzialmente una fase di studio delle principali problematiche ed ha lo scopo di determinare la forma generale del sistema; la seconda fase, avviata nel 2003 e tuttora in corso, è dedicata a definire i dettagli del sistema ed ha come obiettivo la redazione di una bozza della nuova direttiva. Nel corso della prima fase del progetto Solvency II il sottocomitato ha condotto numerose discussioni al suo interno e si è avvalso di molti studi e relazioni elaborati da vari organismi. In concomitanza con queste discussioni, il sottocomitato ha costituito due gruppi di lavoro, composti da esperti di diversi Stati membri. Lo scopo di questi gruppi consiste nello studio di questioni più tecniche, l'uno in materia di assicurazione non vita, l'altro in materia di assicurazione vita. Il gruppo di lavoro non vita ha dedicato i suoi lavori a due tipi di riserve tecniche: le riserve sinistri e le riserve di perequazione. Il fine ultimo è stato quello di ottenere un buon livello di armonizzazione di queste riserve in Europa. Attraverso lo studio di indicatori sintetici il gruppo ha potuto constatare che esisteva una grande diversità nelle modalità di effettuare gli accantonamenti, non soltanto tra i diversi mercati europei, ma talvolta anche all'interno di uno stesso mercato nazionale. Il gruppo ha pertanto messo in luce l'esigenza di far convergere le diverse pratiche di calcolo per pervenire ad un livello comune di prudenza nella costituzione delle riserve.

Tale fine potrebbe essere conseguito sia mediante la definizione di sani principi sia mediante l'adozione di una base comune per il controllo effettivo delle riserve (in particolare relativamente ai dati statistici utilizzati). Il gruppo ha anche sottolineato con interesse il metodo australiano, il quale consiste nel fissare un valore di riferimento numerico per il livello minimo di prudenza nella costituzione delle riserve, sebbene l'attuazione pratica di un tale metodo presenti alcune problematiche ancora da esaminare. Sostanzialmente il progetto Solvency II dovrebbe costituire l'occasione per rafforzare l'armonizzazione delle norme e delle pratiche prudenziali, migliorando così l'integrazione del mercato unico europeo, così come taluni organismi che rappresentano le compagnie di assicurazione hanno sottolineato, soprattutto al fine di evitare le distorsioni della concorrenza. Tali organismi hanno anche identificato due settori fondamentali per i quali nuove regole appaiono fondamentali: le riserve tecniche e le norme quantitative sugli investimenti. Al fine di attuare questi obiettivi, anche e soprattutto per tenere conto maggiormente del profilo di rischio specifico di ogni impresa di assicurazione, è necessario lasciare una certa flessibilità nelle valutazioni. Ad esempio potrebbero essere utilizzate norme quantitative adeguate per le diverse imprese mediante il riconoscimento di modelli interni o valutare i rischi anche da un punto di vista qualitativo, oltre che quantitativo. In ogni caso questa nuova flessibilità deve essere accompagnata da garanzie destinate alla parità di trattamento delle imprese. Non bisogna poi dimenticare che l'intero progetto si inserisce in un contesto di radicale evoluzione delle norme contabili applicabili all'assicurazione, ossia i principi contabili internazionali IAS, dai quali esso non può prescindere. A livello internazionale sono in corso numerosi progetti relativi alle norme prudenziali. In particolare è interessante citare i lavori presso l'Associazione Internazionale delle Autorità di Vigilanza Assicurativa (AICA) e quelli dell'Associazione Internazionale degli Attuari (AAI). Fra gli argomenti di studio più avanzati vi sono le riserve tecniche, il ruolo degli attuari nel sistema prudenziale ed i livelli di controllo della solvibilità eventualmente affiancati da test di stress. La questione di come ripartire le scelte prudenziali tra le riserve tecniche ed i requisiti patrimoniali minimi può risultare secondaria.



Tuttavia la costruzione di un sistema prudenziale pratico impone la ricerca di un livello comune di prudenza nelle riserve che tenga conto della volatilità presente nelle stime dei sinistri dell'assicuratore in questione. È infatti sulla base di questo livello di riserve che continuerà ad essere definita l'entità di quel capitale minimo di "protezione" sulla base del quale l'autorità di controllo può decidere di effettuare interventi di salvaguardia. Al fine di tutelare gli assicurati sarà inoltre necessaria una politica di investimento per tali riserve adeguatamente prudente. Ecco allora che la valutazione della solvibilità non può ridursi alla semplice fissazione di un requisito patrimoniale minimo: il progetto ha una portata ben più ampia e coinvolge molte più voci. Occorrerà mettere in atto tutta una serie di pratiche, dai nuovi metodi di stima delle riserve tecniche ai nuovi e più adeguati processi di controllo, con il fine ultimo di ottenere una migliore gestione del rischio.

A conclusione del paragrafo, esaminiamo la struttura del progetto Solvency II con particolare enfasi sull'impatto che esso ha sulle riserve tecniche. Sull'esempio del sistema bancario, il Solvency II prevedrà una architettura del nuovo sistema di solvibilità a 3 pilastri. Il primo pilastro conterrà i requisiti finanziari quantitativi e i principi generali circa gli investimenti effettuati dall'assicuratore, le disposizioni sulle riserve tecniche e sui mezzi patrimoniali propri. Il secondo pilastro riguarderà tutto ciò che concerne il processo di vigilanza, che, in materia di assicurazione, dovrà includere anche disposizioni sulla valutazione delle riserve e la gestione degli investimenti. Il terzo pilastro tutelerà la disciplina di mercato, la trasparenza, la comparabilità delle informazioni finanziarie delle compagnie e la corretta informazione agli assicurati. In linea di massima, i principali fattori che favoriscono la disciplina di mercato in materia di assicurazione sono i mercati finanziari e le agenzie di rating (per le società quotate), nonché una tendenza generale ad una maggiore trasparenza e ad un'armonizzazione delle norme contabili. Ovviamente ognuno di questi tre pilastri risulta fortemente influenzato dagli altri ed è indispensabile affinché il nuovo modello di solvibilità sia efficiente. Abbiamo visto come il primo pilastro contiene i requisiti quantitativi del sistema e dovrà pertanto fornire protezione alla compagnia a fronte di tutti quei rischi quantificabili tra i quali rientrano l'adeguatezza degli accantonamenti tecnici.

Il primo principio di solvibilità enunciato dall'IAIS (International Association of Insurance Supervisors) riguarda le riserve tecniche che devono essere “adeguate, affidabili, oggettive”. Esse devono inoltre “permettere il raffronto tra gli assicuratori”. Sulla base di questi criteri uno dei principali limiti del sistema europeo, dovuto alla varie pratiche di valutazione utilizzate in Europa con livelli di prudenza differenti, è certamente di non permettere in misura sufficiente il raffronto tra assicuratori. Non è possibile effettuare un confronto tra le compagnie tramite le riserve sinistri. Seppure non sembra che si debba mettere in discussione il principio della prudenza nelle valutazioni effettuate, tuttavia tale principio comprende oggi pratiche troppo diverse che conducono a livelli di prudenza eterogenei tra loro, sia in Europa sia, a volte, anche in uno stesso mercato nazionale, così come sottolineato dalla relazione del gruppo di lavoro delle riserve tecniche non vita. I servizi della Commissione stanno pertanto riesaminando le due possibilità proposte dal gruppo di lavoro sopra menzionato. La prima consiste nel rafforzare il principio di prudenza delle direttive operando sulla gestione dei sinistri e sulle riserve tecniche mediante l'adozione di statistiche comuni (o almeno comparabili), le quali permetterebbero anche un migliore processo di controllo dei livelli di riserva nell'Unione Europea (misure di tipo “secondo pilastro”). La seconda soluzione consisterebbe nel fissare un riferimento quantitativo per il livello di prudenza delle riserve sinistri, ispirandosi pertanto alla soluzione adottata dall'autorità di controllo australiana, APRA. Questa seconda soluzione comporterebbe alcune difficoltà metodologiche, tuttavia avrebbe il vantaggio di inviare un chiaro segnale al mercato, efficace per ottenere l'armonizzazione ricercata. Non è neppure da escludere che essa potrebbe agevolare il passaggio alle nuove norme contabili.

In materia di riserve tecniche, l'altro argomento di particolare interesse risulta infine essere legato alle norme di copertura. L'articolo 20 della terza direttiva (vita e non vita) stabilisce che “gli attivi a copertura delle riserve tecniche devono tener conto del tipo di operazioni effettuate dall'impresa di assicurazione in modo tale da assicurare la sicurezza, il rendimento e la liquidità degli investimenti dell'impresa, la quale provvederà all'adeguata diversificazione e dispersione di tali investimenti”. Il principio base di tale articolo è quello di una gestione finanziaria prudente (principio della “persona prudente”).

In conseguenza la necessità di un'appropriata asset liability management è dunque già presente nelle direttive, tuttavia dovrà essere maggiormente chiarita. Gli investimenti dovranno essere anche sicuri, liquidi, proficui e sufficientemente differenziati e dispersi (teoria ampiamente sviluppata negli articoli 21 e 22 della terza direttiva). L'articolo 21 elenca anche gli investimenti ammissibili come copertura delle riserve tecniche ed i principi per l'interpretazione di tale elenco, mentre l'articolo 22 stabilisce i limiti quantitativi massimi per le varie categorie di attivi, nonché i principi qualitativi che permettono di interpretare più precisamente le nozioni di una dispersione e diversificazione adeguata.

Per quanto riguarda il secondo pilastro, il principio generale che regola la valutazione delle riserve stabilisce che le imprese devono utilizzare metodi appropriati di stima e devono disporre di meccanismi correttivi tali da modificare le proprie valutazioni qualora ciò sia richiesto in seguito al presentarsi di un elemento nuovo. Questo principio può inoltre essere maggiormente specificato mediante l'enunciazione di ulteriori principi in materia di gestione dei sinistri o dei contratti.

In particolare il gruppo di lavoro non vita ha suggerito i seguenti aspetti che dovrebbero riguardare la gestione e la costituzione di riserve sinistri:

- l'obbligo di distinguere le attività di gestione dei sinistri da quelle di sottoscrizione;
- l'utilizzo di norme che garantiscano una registrazione dei sinistri rapida, esatta e aggiornata regolarmente;
- un'idonea classificazione dei dati, l'utilizzo di personale qualificato per la valutazione dei sinistri e la verifica a posteriori dei metodi di stima utilizzati;
- procedure documentate al fine di attuare un controllo interno di tali procedure e, non ultimo, un sistema di comunicazioni adeguato.

Questo ultimo aspetto ha il suo impatto anche in un'ottica di trasparenza e di divulgazione delle informazioni, pertanto si ritrova in materia di terzo pilastro.

## **1.5 Lo sviluppo dei principi contabili per il settore assicurativo**

Nel 1973 è stato costituito un organo del settore privato composto da organismi di contabilità professionali e società private provenienti da vari paesi<sup>2</sup> e specializzati in contabilità e analisi finanziaria: l'International Accounting Standards Committee (IASC). Lo IASC ha avviato il suo lavoro sulla contabilità delle imprese relative al settore assicurativo nel 1997 quando è stato costituito un apposito comitato per lo studio della questione: l'Insurance Steering Committee. Inizialmente il comitato escludeva alcuni tipi di assicurazioni, come quella del credito e i piani previdenziali aziendali per i dipendenti. Per tale motivo molti gruppi finanziari internazionali e organizzazioni di attività assicurative si opposero finché, nel 2000, il comitato fu sottoposto a profonde modifiche che portarono a dar vita, nel 2001, all'International Accounting Standards Board (IASB) ed un nuovo comitato, l'Insurance Advisory Committee. Il 2001 è anche l'anno in cui vengono adottati tutti i principi contabili internazionali (IAS) emanati dallo IASC e ne vengono creati di nuovi dal nuovo organo appena formatosi, lo IASB, i cosiddetti International Financial Reporting Standards (IFRS), aventi lo scopo principale di pervenire a quella armonizzazione contabile globale tanto ricercata dai vari organismi. Nel novembre del 2001 il Board internazionale presenta il DSOP (Draft Statement of Principles of Insurance Contracts) con lo scopo di costituire una base per i principi IFRS appena creati in materia di contratti assicurativi. Nel 2002 lo IASB comincia a pensare che l'originaria scadenza, fissata per il 2005, sia un po' troppo ravvicinata, anche a causa della crescente pressione esercitata dalle imprese operanti nel settore assicurativo e dalle organizzazioni pubbliche. Pertanto l'intero progetto di contabilità viene scisso in due fasi: la prima, transitoria, copre un periodo intermedio a partire dal 1 gennaio 2005 e rientra quindi nei tempi previsti dall'UE, la seconda, in epoca successiva, con la quale viene effettuata la stesura dei bilanci mediante l'utilizzo completo dei principi assicurativi formulati.

---

<sup>2</sup> Tra di essi il maggior supporto fu dato da Australia, Canada, Francia, Germania, Giappone, Messico, Paesi Bassi, Regno Unito, Irlanda e Stati Uniti.

Nel 2003 vi è la pubblicazione da parte dello IASB dell' "Exposure Draft 5 – Contratti assicurativi" (ED 5), una sorta di orientamento per il passaggio agli IFRS durante la prima fase del progetto di contabilità. Tale documento ha permesso di raccogliere commenti e raccomandazioni da parte delle organizzazioni pubbliche e private: base informativa che si è poi rivelata notevolmente utile per la messa a punto dell'IFRS 4 sui contratti assicurativi. Quest'ultimo, pubblicato dallo IASB nel marzo del 2004, rappresenta il passaggio dalla prima alla seconda fase del progetto di contabilità internazionale. Il principio, entrato in vigore il 1 gennaio 2005, oltre a contenere la definizione di contratto assicurativo e i principi base per la contabilizzazione di quest'ultimo, offre anche la base per la valutazione dell'impatto dei nuovi principi contabili a livello di stato patrimoniale e conto economico. L'obiettivo principale è quello di "realizzare miglioramenti in materia di prassi contabile" e di "illustrare quali sono i principali fattori di rischio e di sensibilità dei contratti assicurativi". L'IFRS 4 definisce un contratto assicurativo come "un contratto che espone l'assicuratore a rischi identificati di perdita derivante da eventi o circostanze che avvengono o sono scoperte entro uno specifico periodo, includenti la morte (o la sopravvivenza, in caso di annualità), la malattia, la disabilità, il danneggiamento di proprietà, l'ingiuria verso altri o l'interruzione degli affari". L'altra caratteristica del rischio consiste nella preesistenza al contratto e nella sua natura che deve essere di carattere non finanziario, dove per rischio finanziario si intende la possibilità di una futura variazione di uno dei seguenti fattori: tasso di interesse, prezzo degli strumenti finanziari, prezzo delle merci, tasso di cambio, indice dei prezzi dei tassi e altre variabili ad esse connesse. Il principio prevede inoltre che, ad ogni data di bilancio, l'assicuratore verifichi che le passività accantonate siano adeguate a far fronte agli impegni futuri (cioè ai futuri flussi di cassa stimati per i contratti in corso). Nel caso in cui la passività risulti inadeguata la carenza deve essere imputata al conto economico mentre non è consentita la compensazione tra passività assicurative e le correlate attività connesse ai contratti di assicurazione. L'IFRS 4 consente infine di continuare ad applicare le precedenti prassi contabili purché siano rispettate le prescrizioni minime sopra descritte.

Per i contratti che invece non soddisfano i requisiti dei contratti assicurativi ai sensi dell'IFRS 4 la questione è molto dibattuta. La definizione di contratto assicurativo data dall'IFRS 4 si concentra infatti sulle sole cause scatenanti problemi legati ai contratti di assicurazione, detti rischi assicurativi. Questa definizione evita pertanto quella di rischio finanziario. Tuttavia esistono alcuni contratti che, pur avendo una forma legale che corrisponde a quella di un contratto assicurativo, non trasferiscono all'assicuratore un significativo rischio assicurativo, così come esistono dei contratti i quali presentano una significativa componente di deposito. Per far fronte a questi problemi è stata proposta una soluzione detta unbundling (letteralmente "spacchettatura"), con la quale si è proposto di separare (unbundle) la componente di deposito di un contratto da quella assicurativa. In questo modo, mentre per la seconda componente, prettamente assicurativa, l'assicuratore dovrà riferirsi all'IFRS 4, la prima componente sarà valutata con altri principi, primo fra tutti lo IAS 39 ("Strumenti finanziari: rilevazione e valutazione"), principale riferimento in materia di classificazione e valutazione dei vari strumenti finanziari.

### **1.6 I principi IAS e la riserva sinistri**

Con riferimento alle riserve sinistri dei rami danni i principi IAS richiedono dei metodi per la valutazione tali da coprire gli impegni assunti con un dato livello di "fiducia" e con l'obiettivo di ottenere una maggiore armonizzazione delle pratiche nazionali. L'entrata in vigore dei nuovi principi contabili è tuttavia molto più che una semplice questione di uniformazione a livello internazionale. Gli assicuratori prevedono infatti che i nuovi standards produrranno un impatto molto più ampio e rilevante sul settore assicurativo, impatto che riguarderà l'intera gestione del business, la misurazione delle performance e la comunicazione con gli operatori del mercato. Per quanto riguarda la separazione dei contratti in una componente assicurativa e in una di investimento (il cosiddetto unbundling) non si avranno grandi ripercussioni sugli assicuratori che si occupano dei rami danni. Al contrario del ramo vita, i cui contratti presentano spesso una non trascurabile componente di deposito, nel ramo danni si hanno generalmente contratti di tipo uniperiodale (per lo più annuali), i quali pertanto si basano sui premi di rischio.

Di fondamentale importanza per l'assicuratore danni risulta invece il problema relativo alla gestione dei sinistri e all'accantonamento a riserva. La corretta stima dei futuri impegni, quelli cioè che costituiscono le riserve tecniche, ha un rilievo notevole sulle voci di bilancio. Gli assicuratori dovranno pertanto procedere ad una revisione delle metodologie di accantonamento, sia espandendo i metodi già utilizzati sia sostituendo quelli già in uso con altri più avanzati, primi fra tutti quelli di tipo stocastico. I metodi deterministici di stima risultano poco adeguati anche a seguito delle frequenti revisioni effettuate sulle ipotesi utilizzate per compiere le valutazioni. Ne consegue una maggiore volatilità non catturabile con un valore puntuale di stima. Ben più appropriata risulta essere una valutazione basata sul concetto di "best estimate" (valore attuale atteso) a cui aggiungere un margine di sicurezza, il cosiddetto "safety margin", per poter considerare l'incertezza presente nelle stime.

## CAPITOLO 2

### LE RISERVE TECNICHE NEI RAMI DANNI

#### 2.1 Premesse

Le riserve tecniche, con il cui termine vengono indicate sia la riserva premi che la riserva sinistri, indicano gli accantonamenti che detiene l'assicuratore. Questi importi monetari costituiscono un debito nei confronti dell'assicurato e si contraddistinguono dalle cosiddette riserve patrimoniali, fondi liberi da impegni gestionali e pertanto quantità a disposizione dell'impresa. Questi ultimi fondi si formano mediante accantonamento di utili e sono utilizzati per fronteggiare altri obblighi di gestione e non quelli tecnici, così come avviene con le riserva premi e la riserva sinistri<sup>3</sup>. Come specificato nel precedente capitolo, nonché come diretta conseguenza della stessa definizione di riserva tecnica, queste entità monetarie vengono contabilizzate, alla fine di ogni anno di esercizio, al passivo delle voci di bilancio. Per tale motivo è facile comprendere l'importanza di una loro accurata valutazione. I valori ottenuti sono quantità stimate poiché all'epoca di bilancio non è possibile conoscere l'esatto ammontare degli impegni futuri che gravano sull'impresa di assicurazione. Queste prime considerazioni evidenziano come il premio incassato dall'assicuratore per ogni singolo contratto stipulato risulta essere la base di partenza per ogni valutazione riguardante le riserve tecniche stesse. In particolare si dovrà focalizzare l'attenzione sulla gestione del premio durante il periodo contrattuale, ossia l'intervallo di tempo coperto dalla polizza assicurativa.

#### 2.2 La riserva premi

La prima questione da affrontare riguarda la corretta gestione del premio assicurativo al fine di garantire un'adeguata copertura degli eventuali esborsi derivanti dai sinistri che colpiscono il contratto durante il periodo di valenza.

---

<sup>3</sup> Non dimentichiamo che sulla produzione dell'impresa assicuratrice gravano, come in ogni altra attività industriale, imposte e tasse.



In particolare è opportuno considerare che il premio risulta essere di competenza anche per anni di esercizio successivi a quello relativo al momento in cui l'importo viene versato. Come ipotesi di partenza si consideri un portafoglio di contratti sufficientemente omogenei in base a quanto è possibile affermare all'epoca di stipula dei contratti. In particolare:

- tutti i contratti sono stipulati ad una stessa data, che per convenzione possiamo fissare con il 1 gennaio di un certo anno  $x$
- la gestione di impresa ha avvio proprio in questo anno (sia  $t=0$  l'inizio di esercizio e  $t=1$  la chiusura dello stesso)
- tutti i contratti hanno una durata annuale (ipotesi non molto restrittiva per le assicurazioni dei rami danni)
- sia  $W$  l'importo del fondo iniziale, riserva di sicurezza di cui l'assicuratore dispone<sup>4</sup>.

Ci si ponga adesso nell'ottica di un assicuratore che voglia seguire durante questo primo anno di esercizio la gestione dell'impresa. Nelle date ipotesi egli incassa, al tempo 0, un premio  $C$ , somma dei premi di tariffa complessivi e comprensivo dei vari caricamenti. In particolare sia:

- $\varepsilon_1 C$ , il caricamento per spese di acquisizione (provvigioni d'acquisto che vengono corrisposte agli agenti e spese di produzione, assunte quindi entrambe proporzionali all'importo dei premi);
- $\varepsilon_2 C$ , il caricamento per spese di gestione (anche esse proporzionali all'importo totale dei premi introitati).

La terza componente dei caricamenti per spese, quella relativa alla liquidazione dei sinistri, si considera invece compresa nell'entità dei risarcimenti,  $X$ .

Con l'ulteriore condizione di spese di gestione uniformemente distribuite nel corso dell'intervallo annuale  $[0,1]$  e tralasciando per il momento l'importo  $W$ , al generico tempo  $t$  resta quindi a disposizione dell'impresa l'importo

---

<sup>4</sup> Fondo costituito per far fronte ad andamenti della sinistrosità diversi da quello atteso, che costituisce la base di calcolo del premio.

$$d(t) = C(1 - \varepsilon_1 - t\varepsilon_2) - X(t)$$

con il quale l'assicuratore dovrà far fronte ai sinistri che potranno colpire i rischi del portafoglio durante il periodo assicurato (che nel caso specifico coincide con l'anno di esercizio in questione).

Occorre osservare come questa quantità non rappresenti il guadagno dell'assicuratore al tempo  $t$ . Tra  $t$  ed 1, infatti, la compagnia di assicurazione dovrà sostenere ulteriori uscite, sia per spese di gestione sia per i relativi risarcimenti, mentre le entrate saranno pari a 0. In conseguenza, all'epoca  $t$ , ci sono ancora impegni futuri che l'assicuratore detiene nei confronti degli assicurati e che devono essere calcolati sulla base della suddetta disponibilità. Tali impegni costituiscono la riserva premi<sup>5</sup>,  $R_p(t)$ , la quale, con l'ulteriore ipotesi di uniforme distribuzione del verificarsi dei sinistri nel corso del tempo, risulta essere determinata da

$$R_p(t) = (1-t)(1-\varepsilon_1)C.$$

Fino adesso si è trattato il problema in ipotesi semplificatrici della realtà le quali non riflettono pienamente il problema di valutazione cui un assicuratore deve far fronte. L'ipotesi iniziale di acquisizione di tutti i contratti del portafoglio nello stesso istante di tempo (coincidente con l'inizio di un esercizio: nello specifico il primo), ad esempio, non appare conforme alla realtà assicurativa. L'acquisizione dei contratti avviene infatti pressoché in ogni istante di tempo. Si può pertanto pensare di analizzare il caso più generico di acquisizione di contratti annuali in un qualunque giorno dell'anno, analisi possibile mediante la discretizzazione dell'intervallo di tempo di riferimento. Proprio in questo contesto assume la sua importanza la contabilizzazione delle poste a riserva all'interno delle voci del bilancio, cosa che altrimenti non avrebbe senso. Mantenendo l'ipotesi che tutti i contratti abbiano una durata annuale ci si trova nella situazione in cui l'assicuratore non incassa più un unico premio  $C$ , bensì diversi premi  $C_1, C_2, \dots, C_m$ , relativi alle varie epoche di acquisizione dei contratti.

---

<sup>5</sup> Tale riserva è anche detta "riserva per rischi in corso o per danni non ancora denunciati" oppure "riporto delle aliquote di premio per rischi ancora scoperti alla chiusura dell'esercizio".

Estendendo quindi l'analisi al caso più generale nel quale i premi sono destinati a coprire i rischi nel periodo di competenza  $[t_i, t_i + 1]$  e mantenendo l'ipotesi di uniforme distribuzione degli oneri nel corso del tempo<sup>6</sup>, la disponibilità al tempo  $t$  diventa

$$d(t) = \sum_{t_j \leq t} C_j (1 - \varepsilon_1 - \varepsilon_2(t - t_j)) - X(t),$$

mentre la riserva premi diviene

$$R_p(t) = \sum_{t_j \leq t} (t_j + 1 - t)(1 - \varepsilon_1) C_j.$$

All'epoca di bilancio,  $t=I$ , si ha

$$R_p(I) = \sum_{t_j \leq I} C_j (1 - \varepsilon_1) t_j,$$

quantità che fornisce la stima della riserva premi alla chiusura di bilancio.

La riserva premi può essere valutata in ogni istante di vita del contratto e la valutazione si basa sulla durata media residua dei contratti assicurativi.

In ipotesi di uniforme distribuzione di spese, di accadimento dei sinistri e di risarcimenti, si può utilizzare il metodo del pro rata temporis, vale a dire contratto per contratto. Se si utilizza questo metodo vengono considerate tutte le durate residue relative ad ogni singolo contratto stipulato dall'assicuratore. In altre parole si attribuisce ad ogni polizza la parte di premio commisurata alla porzione di durata residua senza alcuna aggregazione contrattuale. Nello specifico viene individuata un'unità di misura di tempo rispetto alla quale calcolare la durata residua relativamente ad ognuno dei contratti stipulati. L'unità di misura, cioè l'intervallo in cui è suddiviso un esercizio, ha generalmente durata mensile, ma si possono fare valutazioni anche considerando il mezzo mese o il giorno, e si parla rispettivamente di metodo dei 12-esimi, dei 24-esimi, o dei 365-esimi.

---

<sup>6</sup> E' stata sottintesa un'ipotesi per la quale i caricamenti sono previsti secondo gli stessi coefficienti per ogni contratto del portafoglio.

In alternativa, all'interno di qualche ramo assicurativo, può essere autorizzata dall'autorità di controllo una valutazione di tipo forfetario che consiste nell'individuare un'unica durata residua media a partire dalle durate residue relative ai singoli contratti. Per farlo si sintetizzano tutti i valori effettuandone una sorta di media. Ipotizzando ad esempio uniforme distribuzione nell'incasso premi si può supporre che gli introiti siano tutti concentrati in  $t = \frac{1}{2}$  e pertanto

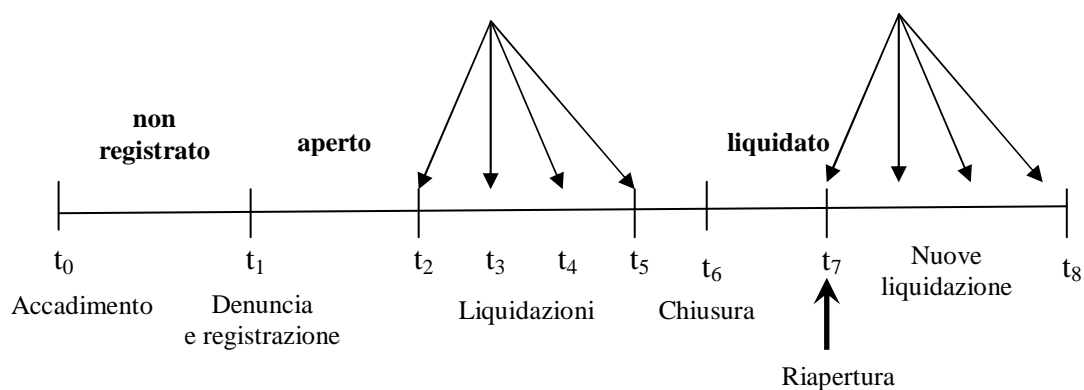
$$R_p(1) = \frac{1}{2}(1 - \varepsilon_1)C_j,$$

dove le diverse durate residue sono state sintetizzate in un'unica durata media residua posta pari a metà anno. Questo metodo impone numerosi vincoli: ad esempio la riserva premi dei rischi R.C.A. calcolata con il metodo forfetario non deve essere inferiore al 40 % dei premi di tariffa d'esercizio, mentre per i rischi che hanno durata più breve, come quello del trasporto merci, il limite scende al 15 %.

### **2.3 L'evoluzione dei sinistri e la riserva sinistri**

Lo studio dell'evolversi di un sinistro ed in particolare del risarcimento ad esso associato, evidenzia la presenza di vari istanti temporali, che vanno da quello iniziale in cui avviene il sinistro stesso, a quello finale nel quale il sinistro risulta definitivamente liquidato. Molto spesso il processo di liquidazione di un sinistro non è piatto e costante nel tempo: sono di comune consuetudine sia ritardi nella denuncia dello stesso evento, sia ritardi nel pagamento dell'indennizzo da parte dell'assicuratore, così come liquidazioni di tipo parziale fino al momento in cui non si conosca l'esatto ammontare finale dell'importo dovuto all'assicurato.

La storia di un sinistro può quindi essere riprodotta con il seguente schema che rappresenta i vari stati di un sinistro nei diversi intervalli di tempo individuati dalle varie epoche:



**Figura 1: La storia di un sinistro**

Gli istanti fondamentali per la descrizione di un sinistro sono fondamentalmente tre:

- l'istante  $t_0$  di accadimento del sinistro;
- l'istante  $t_1$  di denuncia dello stesso e relativa registrazione da parte della compagnia;
- l'istante  $t_6$  di precisazione dell'entità definitiva del risarcimento relativo al sinistro stesso.

Come diretta conseguenza si hanno i seguenti stati di definizione di un sinistro, così come evidenziato in figura:

- sinistro accaduto ma non ancora denunciato nell'intervallo temporale  $[t_0, t_1)$ , sinistro **non registrato**;
- sinistro **aperto**, nell'intervallo temporale  $[t_1, t_6)$ ;
- sinistro **liquidato (o chiuso)** a partire dall'istante  $t_6$ .

Occorre osservare che lo schema proposto risulta comunque una semplificazione della realtà nella quale si hanno spesso situazioni solo parzialmente descritte in figura 1. Innanzitutto non sempre l'istante di denuncia e di registrazione coincidono. Possono inoltre verificarsi una serie di risarcimenti parziali relativi ad un sinistro aperto, mentre il pagamento definitivo avverrà con quasi assoluta certezza dopo la fine del periodo contrattuale stesso (ad esempio nel caso di un sinistro denunciato in prossimità della fine del periodo contrattuale). Ci sono poi situazioni in cui si verifica una successiva riapertura di un sinistro anche a seguito della chiusura dello stesso.

Risulta evidente, pertanto, la necessità da parte dell'assicuratore di porre a riserva importi monetari destinati a coprire quei sinistri che sono accaduti, ma non sono stati ancora denunciati, o quelli il cui processo di liquidazione ha preso avvio, ma non è ancora terminato. L'ammontare dell'accantonamento cui l'assicuratore deve disporre al fine di corrispondere gli importi dovuti per le situazioni sopra descritte prende il nome di riserva sinistri. La valutazione di tale riserva può essere fatta sia seguendo nel tempo la storia di ogni singolo sinistro, dall'epoca in cui avviene al momento in cui è definitivamente liquidato, secondo un'ottica di monitoraggio ed ottenendo così una valutazione dinamica del costo finale dello stesso (metodo dell'inventario, approccio individuale o analitico di stima). In alternativa è possibile analizzare un portafoglio sufficientemente ampio di sinistri e quanto più omogeneo in termini di risarcimenti (approccio collettivo o sintetico che porta all'utilizzo di metodi di tipo statistico-probabilistici). La scelta di uno di questi due metodi dipende dalla tipologia dell'insieme di sinistri con cui si ha a che fare. In particolare se si ha un portafoglio con un numero sufficientemente grande ed omogeneo di rischi e caratterizzato da sinistri di basso importo si tende ad optare per il metodo collettivo, viceversa avendo a che fare con pochi sinistri, ma di importo elevato è preferibile l'approccio individuale.

Per dare buoni risultati l'approccio collettivo dovrebbe essere caratterizzato anche da stabilità delle condizioni contrattuali nel tempo, ossia dovrebbero restare costanti le modalità di liquidazione dei sinistri, così come i relativi costi. Spesso nella pratica attuariale si adotta un approccio di tipo misto: si utilizza un metodo di tipo collettivo sui sinistri di medio-basso importo e si stima la coda destra della distribuzione (vale a dire i sinistri di ammontare più elevato) attraverso un metodo di tipo individuale. Qualunque sia il procedimento utilizzato la stima della riserva sinistri risulta più complessa rispetto alla valutazione della riserva premi e comporta l'impiego di molte più risorse, prima fra tutte l'informazione disponibile, fondamentale base di partenza per la formulazione di future ipotesi sull'evoluzione di grandezze di tipo economico (quali l'inflazione, la redditività degli investimenti a copertura delle riserve o i mutamenti di tipo legislativo), nonché per la stima dei futuri oneri a carico dell'assicuratore.

In proposito occorre osservare come l'errato utilizzo dell'informazione disponibile può precludere la correttezza delle stime mettendo così a repentaglio le gestioni future della compagnia (in caso di sottostima) o la posizione degli azionisti della stessa (in caso di una sovrastima). In merito alle considerazioni fatte fino a questo momento è possibile pensare alla riserva sinistri come l'importo atteso del costo dei risarcimenti del sinistro o dell'insieme dei sinistri. Questa definizione mette in evidenza che la riserva sinistri non è il valore attuale atteso dell'importo dei futuri risarcimenti, bensì l'importo stimato del valore finale.

#### **2.4 Scomposizione della riserva sinistri**

Le considerazioni svolte nel precedente paragrafo portano a una naturale scomposizione della riserva sinistri basata sui motivi per cui tale quantità viene predisposta.

- La riserva **IBNER** (Incurred But Not Enough Reserved) generata dal fatto che alla chiusura dell'esercizio in questione alcuni sinistri sono stati denunciati e registrati dall'impresa, ma non sono stati ancora completamente pagati. In particolare, riprendendo lo schema precedentemente esposto, essa comprende i sinistri aperti e i liquidati.
- La riserva **IBNR** (Incurred But Not Reported) generata dal fatto che alla chiusura dell'esercizio ci sono sinistri avvenuti, ma non ancora denunciati. In particolare essa comprende le tardive denunce (o, per riprendere lo schema del paragrafo precedente, i sinistri non registrati).

La riserva IBNR risulta essere per l'assicuratore di importo abbastanza modesto rispetto a quella IBNER in quanto le tardive denunce si hanno, di solito, solo per quei sinistri avvenuti in prossimità della chiusura d'esercizio (in genere gli ultimi 2-3 mesi dell'anno). Una situazione ben diversa si ha invece per i riassicuratori, i quali possono trovarsi a coprire risarcimenti che, con il passare del tempo, hanno superato la priorità e che invece non sarebbero stati di loro competenza se la denuncia fosse

avvenuta contestualmente al verificarsi dell'evento stesso<sup>7</sup>. Per la valutazione della riserva IBNR si ricorre all'esperienza passata, valutando l'influenza quantitativa dei risarcimenti a fronte di tardive denunce. Le stime sono in genere effettuate in termini di numero medio di sinistri e costo medio di risarcimento relativamente ad ogni anno di esercizio. Vista la maggiore importanza nel seguito tratteremo i soli metodi per la stima della riserva IBNER, accennando soltanto a quelli che possono essere alcuni metodi di calcolo della riserva IBNR.

## **2.5 La rappresentazione dei dati: le coorti di sinistri**

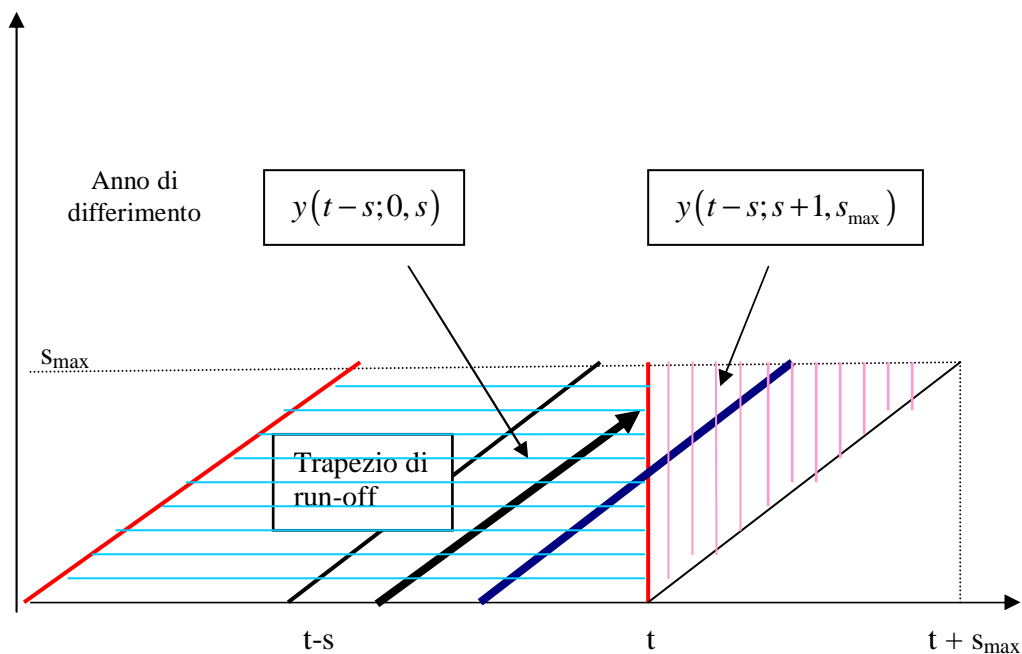
Per ottenere una stima della riserva sinistri risulta opportuno raggruppare i dati per anno di accadimento, costruendo cioè delle coorti di sinistri che abbiano in comune l'anno in cui è avvenuto il sinistro stesso, tecnicamente detto anno di generazione o di origine. Come diretta conseguenza dell'evolversi di un sinistro è poi necessario specificare ulteriori ipotesi sull'andamento dei risarcimenti nel tempo. Se si ipotizza che l'evento sinistro provochi un risarcimento ad una certa epoca  $t$ , esso verrà denunciato presso la compagnia in un istante di tempo successivo o pari a  $t$ , mentre verrà liquidato totalmente nell'anno  $t + s$  ( $s \geq 0$ ) con la possibilità di avere risarcimenti parziali nell'intervallo di tempo  $[t, t + s]$ .

È prassi comune stabilire un ritardo massimo di risarcimento dei sinistri ( $s_{\max}$ ), oltre il quale la compagnia non deve più corrispondere alcuna somma per quei danni provocati da sinistri di una stessa generazione. Più in generale, l'ammontare di denaro che una compagnia deve corrispondere a titolo di risarcimento per i sinistri appartenenti ad una stessa generazione decresce nel tempo fino a diventare nullo dopo un certo periodo di ritardo. Per questo motivo, occorre effettuare l'analisi dei dati sia in funzione della coorte di appartenenza sia in funzione del ritardo nel risarcimento. La situazione appena descritta può essere rappresentata mediante un diagramma di tipo Lexis che mette a fianco più coorti consecutive, così come mostra la seguente figura.

---

<sup>7</sup> In certi casi può essere stabilita la cosiddetta "clausola di stabilità", secondo la quale la soglia riassicurativa viene modificata nel tempo sulla base di cambiamenti dovuti a fattori esterni.





**Figura 2: Il processo di risarcimento mediante un diagramma di tipo Lexis**

La quantità  $y(t-s; 0, s)$  rappresenta l'ammontare cumulato fino all'anno di differimento  $s$  dei risarcimenti relativi ai sinistri totalmente o parzialmente liquidati durante l'arco temporale  $[t-s, s]$  più la riserva per i sinistri ancora aperti alla fine dell'anno  $s$  stesso.

L'anno di valutazione corrente è indicato con  $t$  e la coorte originata nell'anno di accadimento  $t-s$  è rappresentata dalla diagonale ascendente evidenziata in figura. Il grafico evidenzia che la base dati cui la compagnia si trova a disporre per la stima della riserva sinistri è costituita dal cosiddetto trapezio run-off, reticolato azzurro in figura, mentre i risarcimenti che devono essere ancora stimati al fine del calcolo della riserva appartengono a coorti che risultano ancora aperte all'epoca di valutazione e che quindi non hanno ancora raggiunto il ritardo massimo di pagamento. Questi valori sono riprodotti da un triangolo e sono indicati in figura dal reticolato in rosa a destra di  $t$ .

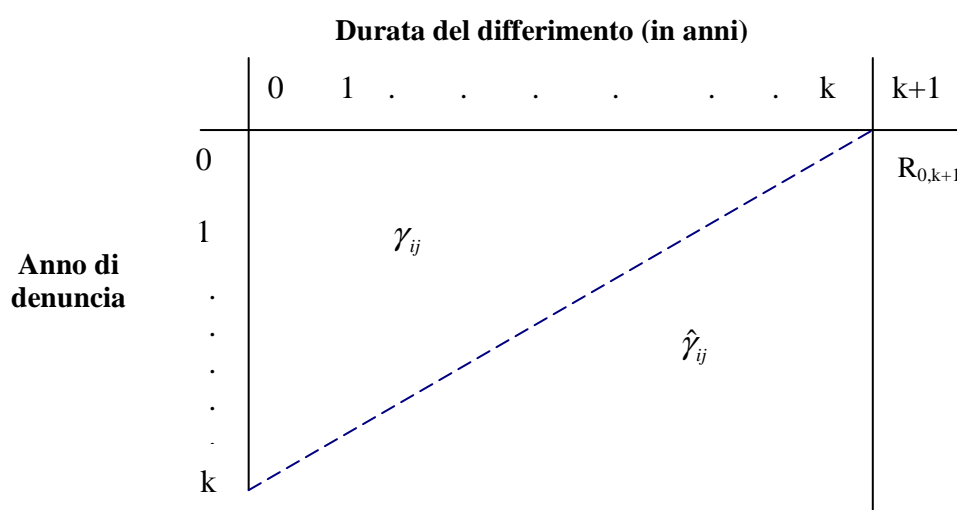
## 2.6 La base dati: l'informazione triangolare

Quanto detto nel precedente paragrafo costituisce l'informazione base, punto di partenza per effettuare la stima della riserva sinistri. L'approccio collettivo necessita di specifiche ipotesi per l'impostazione del problema in condizioni di tipo generali. Innanzitutto ipotizziamo che il differimento del risarcimento abbia un limite superiore e che questo sia indicato con  $k$ . Tale limite viene misurato in anni ed è specifico del ramo che si prende in considerazione: ad esempio per i rischi dei rami infortuni o malattie  $k \leq 5$ , mentre per i rischi dell'R.cAuto  $k$  può assumere anche valori fino a 8 o 9 anni. Si supponga inoltre che i sinistri vengano raggruppati per anno di denuncia e quindi, trovandosi in  $k+1$ , l'assicuratore si può sentire "al sicuro" in quanto ha già pagato gli indennizzi che doveva agli assicurati per i sinistri avvenuti fino a  $k$  anni prima. Ponendosi alla data di bilancio, che corrisponde al 31 dicembre di un certo anno  $t$ , l'assicuratore conosce le liquidazioni per i sinistri relativi a questo anno e a quelle per i sinistri appartenenti alle generazioni degli anni precedenti,  $t-1, t-2, \dots, t-k, \dots$ . Ovviamente il numero di queste generazioni risulta essere non inferiore alla massima durata di differimento. Di solito l'informazione relativa ai risarcimenti con più di  $k$  anni di differimento viene compattata raggruppando nella generazione  $t - k$  anche i sinistri di generazioni precedenti non ancora totalmente liquidati dopo  $k$  anni di ritardo. Si ipotizza di liquidare questi esborsi con un risarcimento uguale all'importo posto a riserva. Sotto queste ipotesi è possibile ottenere una rappresentazione dell'informazione di base e della situazione generale del problema molto utile ai fini applicativi, sia per quanto riguarda la facile lettura ed interpretazione di questa sia per ciò che concerne l'utilizzo di particolari metodi di stima della riserva sinistri stessa. Lo schema può essere raffigurato da una tabella che classifica i risarcimenti secondo un criterio che mostri sia la coorte di appartenenza, che il ritardo nel pagamento del risarcimento. Si ottiene così una matrice quadrata nella quale gli elementi del triangolo superiore (triangolo di run-off) corrispondono agli elementi noti, poiché corrispondono a informazione sul passato e sul presente, mentre quelli del triangolo inferiore sono gli elementi da stimare per ottenere il valore della riserva.

La matrice così definita può essere studiata sotto diversi punti di vista.

In particolare:

- sulle diagonali si ha l'evoluzione dei sinistri per i diversi anni di calendario
- su ogni riga si ha l'evoluzione di una generazione di sinistri per i diversi anni di sviluppo
- sulle colonne si legge il trend evolutivo dei sinistri, in un dato anno di sviluppo, per differenti generazioni



**Figura 3: Rappresentazione della base dati**

Le informazioni che vengono registrate in questa tabella possono riguardare quantità diverse: numero di sinistri, pagamenti totali, pagamenti medi, etc..., mentre i valori possono apparire sia in forma cumulata che non. Si deve anche notare che in alcuni casi, anziché di un triangolo di dati, si dispone di un insieme che può assumere la forma di un trapezio o di un parallelogramma, in relazione al fatto che il numero di generazioni non è necessariamente uguale a quello degli anni di differimento. L'elemento aggiuntivo,  $R_{0,k+1}$ , indica l'importo collocato a riserva per i sinistri della generazione 0 e superstiti di generazioni precedenti. In altre parole tale valore riguarda quei sinistri il cui risarcimento non è ancora stato completato dopo k anni dal loro accadimento.

Tramite opportuni metodi di stima, dagli elementi noti presenti in questa tabella si ricavano i  $\hat{\gamma}_{i,j}$ ,  $i=1\dots k$ ,  $j=k-i+1\dots k$ , del triangolo inferiore della matrice e successivamente i termini aggiuntivi  $\hat{R}_{i,k+1}$ <sup>8</sup>,  $i=1\dots k$ , che indicano le riserve relative a pagamenti con più di k anni di differimento. A partire da queste stime è poi immediata la costruzione della previsione delle future liquidazioni e quindi alla stima della riserva sinistri del portafoglio.

---

<sup>8</sup> In alcuni modelli si pone questo importo uguale a 0 in quanto i sinistri non ancora definitivamente liquidati dopo k anni risultano essere in numero limitato.

## CAPITOLO 3

### MODELLI DETERMINISTICI PER LA STIMA DELLA RISERVA SINISTRI

#### 3.1 Premesse

In questo capitolo verrà esaminata la stima della riserva sinistri secondo un approccio di tipo collettivo. Come già sottolineato nel capitolo precedente il metodo collettivo richiede da parte della compagnia la disponibilità di alcuni elementi essenziali al fine di pervenire alla valutazione finale. In particolare l'assicuratore avrà bisogno di una base statistica appropriata formata da una adeguata informazione circa il numero di sinistri (aperti e liquidati), dei risarcimenti (cumulati e non) per i sinistri liquidati e dei premi di competenza. Innanzitutto risulta infatti importante poter disporre di una base informativa sufficientemente ampia ed attendibile che si compone di precedenti esperienze di liquidazione di sinistri il più omogenei possibile rispetto a quelli che costituiscono il portafoglio in questione. A partire da questi elementi sarà poi necessario un opportuno modello di proiezione in grado di fornire, con quanta più accuratezza possibile, la stima dei futuri esborsi. I modelli di proiezione che usualmente vengono impiegati possono essere classificati in modelli di tipo deterministico e modelli di tipo stocastico. I primi forniscono una stima puntuale della riserva senza però riuscire ad identificare la volatilità della stessa. In questa prima tipologia rientrano tutti quei modelli che vengono di norma utilizzati nei principali mercati assicurativi e che godono sicuramente di una maggiore semplicità rispetto ai secondi. I modelli stocastici permettono invece di considerare l'aleatorietà del portafoglio analizzato, fornendo una stima intervallare, che consiste in un intervallo di confidenza entro il quale la riserva si collocherà, anziché una stima puntuale. In questa seconda tipologia rientrano dunque tutti quei modelli che riconoscono la natura stocastica del processo di arrivo dei sinistri e sul quale verranno formulate opportune ipotesi di tipo probabilistico.

Sono pertanto evidenti i vantaggi dei metodi stocastici in quanto capaci di dare atto della variabilità della stima con il conseguente livello di fiducia ad essa attribuita.

Per quanto riguarda i metodi deterministici, essi possono basarsi su due tecniche di proiezione distinte: il “grossing up” ed il “link ratio”. La prima calcola i fattori di sviluppo,  $r_0, r_1, \dots, r_k$ , che rappresentano la legge di “smontamento” dei sinistri, (ritardo nel pagamento del risarcimento relativo ai sinistri), basandosi sulla proporzione di risarcimenti effettuati nel corso degli anni di differimento in rapporto al totale

$$r_0 = \left( \frac{Y_{i0}}{\sum_j Y_{ij}} \right); \quad r_1 = \left( \frac{Y_{i1}}{\sum_j Y_{ij}} \right); \quad \dots; \quad r_h = \left( \frac{Y_{ih}}{\sum_j Y_{ij}} \right); \quad \dots;$$

mentre la seconda considera la proporzione dei risarcimenti effettuata nel corso degli anni di differimento in rapporto a due anni consecutivi

$$r_0 = \left( \frac{Y_{i1}}{Y_{i0}} \right); \quad r_1 = \left( \frac{Y_{i2}}{Y_{i1}} \right); \quad \dots; \quad r_h = \left( \frac{Y_{ih+1}}{Y_{ih}} \right); \quad \dots$$

La logica è poi quella di moltiplicare gli elementi che si trovano sulla diagonale (valori più recenti a disposizione) per tali coefficienti di sviluppo individuati,  $r_0, r_1, \dots, r_k$ , a secondo della tecnica utilizzata. In seguito, una volta individuate le stime relative agli elementi del triangolo inferiore, l'entità della riserva sinistri è determinata semplicemente sommando tutti questi valori ed aggiungendo i termini aggiuntivi (anche questi stimati),  $\hat{R}_{i,k+1}$ , ( $i = 1, 2, \dots, k$ ), relativi alle riserve per i pagamenti effettuati con più di k anni di differimento.

Pertanto, in caso di risarcimenti non cumulati,  $\hat{Y}_{ij}$ , la stima della riserva sinistri risulta essere semplicemente

$$R_{(k)}^{IBNER} = \sum_{i=1}^k \left( \sum_{j=k-i+1}^k \hat{Y}_{ij} + \hat{R}_{i,k+1} \right) + R_{0,k+1}$$

ed indicando con  $\hat{X}_{ij}$  i risarcimenti cumulati tale espressione può essere riscritta

$$R_{(k)}^{IBNER} = \sum_{i=1}^k \left[ (\hat{X}_{ik} - X_{ik-i}) + \hat{R}_{i,k+1} \right] + R_{0,k+1}.$$

Per la scelta di un modello di proiezione rispetto ad un altro risulta di fondamentale importanza il ruolo dell'informazione che l'assicuratore si trova a disposizione e delle caratteristiche dei rischi studiati: è sufficiente ricordare come tali peculiarità influenzino largamente la scelta tra un approccio di tipo individuale e uno di tipo collettivo. Per quanto riguarda l'accuratezza della stima, infine, risulta essenziale l'analisi sistematica dei fattori che possono agire sui risarcimenti e modificarne l'evoluzione (sia agendo sul numero di sinistri che sull'importo associato a ogni singolo risarcimento). A questo proposito si può distinguere tra influenze di tipo **endogeno** (condizioni di polizza, gestione dei sinistri) sulle quali l'impresa di assicurazione può esercitare un adeguato controllo e influenze di tipo **esogeno** (quali inflazione, fattori economici, fattori climatici, modifiche della legislazione, etc.), sulle quali l'impresa non ha invece alcun potere.

Nel seguito del capitolo verranno presentati alcuni metodi di stima della riserva sinistri di tipo deterministico, largamente utilizzati fino ad oggi, ma che nel tempo sono stati modificati per arrivare a metodi di tipo stocastico: un esempio rilevante è il metodo della catena, nato come deterministico e nel seguito raffinato fino a crearne uno di tipo stocastico.

### 3.2 Il metodo della catena

Il metodo della catena, anche noto con il nome "Chain Ladder Method", è nato intorno al 1970 nel Regno Unito ed è tuttora uno dei più utilizzati. Il modello si basa sugli importi cumulati,  $X_{ij}$ , e la stima dei rapporti tra risarcimenti relativi a due anni consecutivi di differimento si ricava con la tecnica del "link ratio", tecnica di proiezione dei dati precedentemente esposta. L'ipotesi fondamentale che sta alla base del modello è che il rapporto  $X_{ij}/X_{ij-1}$  non dipenda dall'anno di origine, ma solo dal numero di anni di differimento  $j$ .

In altre parole esiste una “legge di smontamento” dei sinistri, che rappresenta la progressione dei risarcimenti cumulati nel corso degli anni di differimento, costante per ogni anno di generazione, a meno di variazioni aleatorie.

In sostanza:

$$\frac{X_{01}}{X_{00}} = \frac{X_{11}}{X_{10}} = \frac{X_{21}}{X_{20}} = \dots,$$

la progressione dei risarcimenti cumulati è la stessa per ogni generazione.

Sotto tali ipotesi la stima dei rapporti  $m_j = \frac{X_{ij}}{X_{ij-1}}$ , relativa agli importi cumulati di

due colonne adiacenti, viene effettuata nel seguente modo:

$$m_j = \frac{\sum_{i=0}^{k-j} X_{ij}}{\sum_{i=0}^{k-j} X_{ij-1}}, \quad i = 1, 2, \dots, k.$$

Osserviamo che, in questo caso, si considerano tutti i valori a disposizione relativamente a due anni di differimento consecutivi. Tuttavia è possibile considerare ulteriori versioni del metodo, le cui stime si basano solo su alcuni elementi, relativi agli anni di differimento, presenti nella matrice dei dati. Ad esempio si può pensare di utilizzare i soli valori relativi agli ultimi 3 o 5 anni di generazione oppure si può decidere di assegnare dei “pesi” diversi alle quantità in questione con il fine ultimo di attribuire loro un'importanza diversa in base al loro riferimento temporale: ad esempio è plausibile ritenere più affidabili i dati più recenti.

In ogni caso, una volta determinati tutti i coefficienti  $m_i$ ,  $i = 1, 2, \dots, k$ , è possibile ottenere delle stime per i futuri esborsi e quindi valutare la riserva.

I rapporti  $m_i$ ,  $i = 1, 2, \dots, k$ , individuano infatti dei parametri che ci consentono di passare da una colonna a quella successiva, ovvero:

$$\hat{X}_{ij} = X_{ik-i} \prod_{h=k-i+1}^j m_h, \quad i = 1, 2, \dots, k; \quad j = k - h + 1, \dots, k.$$

Osserviamo che il termine  $X_{i,k-i}$ , essendo sulla diagonale della matrice di run-off, risulta essere noto e quindi non è oggetto di stima.



In modo del tutto analogo si stimano i termini aggiuntivi (importi cumulati finali):

$$\hat{R}_{i,k+1} = \hat{X}_{ik} m_{k+1} \quad \text{dove} \quad m_{k+1} = \frac{X_{0k} + R_{0,k+1}}{X_{0k}} = 1 + \frac{R_{0,k+1}}{X_{0k}}.$$

La riserva globale è infine data da:

$$R_{(k)}^{IBNER} = \sum_{i=1}^k \left[ (\hat{X}_{ik} - X_{ik-i}) + \hat{R}_{i,k+1} \right] + R_{0,k+1}.$$

Il metodo che è stato presentato risulta essere molto appetibile per l'esiguità dell'informazione che viene richiesta, tuttavia, per l'applicazione, devono essere soddisfatte vincolanti condizioni che assicurino la basilare ipotesi di costanza precedentemente introdotta. Si osserva che sulla base di tale ipotesi non sarebbe necessario escludere o dare meno affidabilità ai valori più lontani (negli anni di differimento) presenti nella tabella di run-off, mentre invece accade spesso il contrario. A causa di fattori endogeni o esogeni, quali ad esempio cambiamenti nell'amministrazione aziendale, con la conseguente modifica nella liquidazione dei sinistri ed in particolare nei tempi di indennizzo, mutamenti della legislazione, inflazione,...etc, la suddetta condizione può non sussistere. Proprio per attenuare almeno alcuni degli effetti generati dalla presenza di queste cause perturbatrici e rimediare così a questa carenza del metodo della catena, si possono apportare alcune modifiche all'impianto principale. A questo proposito si introduce adesso il metodo della catena aggiustato per l'inflazione. Questo nuovo modello opera esattamente come il precedente, ma utilizza risarcimenti dai quali l'effetto di smontamento dei sinistri è stato rimosso e convertito in valore corrente della moneta. Per farlo si procede ad un aggiustamento dei dati introducendo dei tassi inflattivi<sup>9</sup> sulla base delle osservazioni passate e si ottiene quindi una nuova stima della riserva seguendo i passi elencati di seguito:

- I. si costruisce la tabella dei risarcimenti non cumulati  $Y_{ij} = X_{ij} - X_{ij-1}$ ;

---

<sup>9</sup> Tali tassi non corrispondono a quelli ufficiali dell'inflazione economica in quanto considerano anche gli effetti, diretti ed indiretti, sulla gestione di impresa in materia di costi per risarcimenti.

II. si inflazionano gli  $Y_{ij}$  assumendo come anno di riferimento quello della generazione più recente (nel caso trattato la generazione  $k$ ) ottenendo in questo modo gli importi  $Y_{ij}^*$  espressi in valore della moneta dell'anno cui è riferita la valutazione della riserva;

III. si ricostruiscono i risarcimenti cumulati  $X_{ij}^* = \sum_{h=0}^j Y_{ih}^*$ .

Si applica adesso il classico metodo della catena a questi nuovi valori (a questo punto espressi in valore della moneta corrente), ossia:

IV. si ricavano le stime degli elementi del triangolo inferiore della matrice di runoff ( $\hat{X}_{ij}^*$ ,  $i = 1 \dots k$ ,  $j = k - i + 1 \dots k$ ) e di  $\hat{R}_{i,k+1}^*$  ( $i = 1 \dots k$ ).

Osserviamo tuttavia che la stima non è ancora completa. Occorre infatti considerare anche l'inflazione presente negli anni futuri la valutazione. Il procedimento di calcolo è analogo a quello utilizzato per gli anni passati con l'ulteriore ipotesi di un tasso di inflazione costante nel tempo e calcolato secondo una qualche media dei tassi riscontrati negli anni precedenti la valutazione stessa. I passi da seguire sono pertanto:

V. si costruisce la tabella dei risarcimenti non cumulati,  $\hat{Y}_{ij}^* = \hat{X}_{ij}^* - \hat{X}_{ij-1}^*$ , previsti per ciascuna generazione nei futuri anni di differimento;

VI. si inflazionano gli  $\hat{Y}_{ij}^*$  in base all'ipotesi sulla futura inflazione a tasso costante così come precedentemente formulata e si ottengono così le stime  $\hat{Y}_{ij}^*$  dei risarcimenti non cumulati e le stime  $\hat{R}_{i,k+1}^*$  degli accantonamenti a riserva espressi al valore corrente di moneta;

VII. si calcola infine la riserva (che pertanto considera sia l'inflazione relativa agli anni passati sia l'inflazione relativa agli anni futuri) utilizzando l'espressione riferita ai risarcimenti non cumulati (secondo la formula precedentemente esposta) ottenendo:

$$R_{(k)}^{IBNER} = \sum_{i=1}^k \left( \sum_{j=k-i+1}^k \hat{Y}_{ij}^* + \hat{R}_{i,k+1}^* \right) + R_{0,k+1}.$$

### 3.3 Il metodo della separazione

Il metodo della separazione trae le sue origini nel 1977 con Taylor, il quale però ne attribuisce l'essenza a Verbeek nel 1972. Il Dipartimento del Commercio<sup>10</sup> del Regno Unito, primo ente che ha introdotto il metodo, percepiva in esso la possibilità di ovviare ad alcuni limiti connessi al metodo della catena precedentemente esposto. In particolare le applicazioni pratiche del metodo della catena risultavano essere poco affidabili per diversi motivi legati alle ipotesi di base del modello. In proposito occorre osservare che, intorno a quegli anni, l'inflazione risultava alta e del tutto instabile. Come diretta conseguenza utilizzare il metodo della catena aggiustato per l'inflazione non sembrava essere adeguato in quanto esso richiedeva l'adozione di un indice di inflazione per il risarcimento che sarebbe stato estremamente controverso tra i diversi assicuratori. Anche per tale motivo risultava sempre più evidente la necessità di costruire un modello che considerasse soltanto i dati disponibili, lasciando che questi ultimi “parlassero da soli”<sup>11</sup> a proposito dell'inflazione, così come sosteneva il Dipartimento del Commercio del Regno Unito. Fu in questa ottica che venne introdotto il metodo della separazione, con il fine ultimo di stimare l'inflazione in modo endogeno a partire da ogni set di dati a disposizione. Per quanto appena detto, la struttura base del modello è sostanzialmente quella del metodo della catena aggiustato per l'inflazione. I valori considerati sono i singoli risarcimenti  $Y_{ij}$ , con l'ulteriore base informativa costituita dal numero complessivo dei sinistri relativi alla generazione  $i$  e indicato con  $N_i$ , oltre agli accantonamenti a riserva  $R_{0,k+1}$ . L'ipotesi fondamentale del modello riprende quella su cui si basa il metodo della catena: vi è costanza nello smontamento dei sinistri nel tempo. In altre parole, la stima della riserva IBNER si basa sull'assunzione che la progressione dei risarcimenti non cumulati avvenga secondo una legge costante nel tempo (la stessa per ogni generazione), disturbata però da fattori esogeni (ad esempio l'inflazione).

---

<sup>10</sup> Department of Trade.

<sup>11</sup> “Let the data speak for themselves”.

Come dice il nome il metodo si pone l'intento di separare la legge di base (costante nel tempo) dai disturbi che la accompagnano e associabili a variazioni di carattere economico o generazionale<sup>12</sup>. Per quanto appena detto si può pensare che il generico risarcimento  $Y_{ij}$ , relativo quindi a sinistri della generazione  $i$  liquidati con  $j$  anni di differimento, possa essere fattorizzato nel modo seguente:

$$Y_{ij} = N_i r_j \lambda_{i+j}$$

con

-  $r_j$  parametro adimensionale che esprime la legge di smontamento dei sinistri e

tale che  $\sum_{j=0}^k r_j = 1$ <sup>13</sup>;

-  $\lambda_{i+j}$  importo che esprime il fattore di disturbo dipendente dall'anno  $i+j$  al quale corrisponde il risarcimento del sinistro (epoca di valutazione dello stesso).

A partire dai dati che si hanno a disposizione, quindi quelli del triangolo di runoff, attraverso l'utilizzo della tecnica di proiezione del "grossing up" applicata al risarcimento medio relativo ad ogni singolo sinistro,

$$\bar{Y}_{ij} = \frac{Y_{ij}}{N_i} = r_j \lambda_{i+j},$$

si ricavano le stime dei fattori  $r_j$ ,  $j = 0, 1, \dots, k$ , e  $\lambda_{i+j}$ ,  $i + j = 0, 1, \dots, k$ , a meno di una costante moltiplicativa  $\varepsilon$  non nulla.

Si può dimostrare che le stime dei fattori  $r_j$  e  $\lambda_{i+j}$  corrispondono a:

$$(a) \quad \hat{\lambda}_h = \frac{\sum_{i=0}^h \bar{Y}_{i, h-i}}{1 - \sum_{i=0}^{k-h-1} \hat{r}_{k-i}}, \quad h = 0, 1, \dots, k-1;$$

$$(b) \quad \hat{r}_w = \frac{\sum_{i=0}^{k-w} Y_{i, w}}{\sum_{i=w}^k \hat{\lambda}_i}, \quad w = 0, 1, \dots, k-1.$$

<sup>12</sup> Per variazione di carattere economico si può pensare all'inflazione mentre per variazioni di carattere generazionale si può pensare, ad esempio, al volume di affari trattati nei diversi anni.

<sup>13</sup> Condizione di normalizzazione.

I termini aggiuntivi, relativi agli accantonamenti a riserva, si ottengono in modo del tutto analogo e supponendo che un importo non ancora risarcito entro i primi  $k$  anni di differimento sia completamente liquidato nell'anno immediatamente successivo,  $k+1$ .

In altre parole  $R_{i,k+1} = Y_{ik+1}$ , da cui:

$$(c) \quad \hat{R}_{i,k+1} = \frac{N_i}{N_0} R_{0,k+1} \frac{\hat{\lambda}_{k+i+1}}{\hat{\lambda}_{k+1}}.$$

A questo punto vengono formulate opportune ipotesi sugli importi futuri  $\lambda_{i+j}$  ( $i+j = k+1, \dots, 2k+1$ ). Di solito si assume che i  $k+1$  rapporti  $\frac{\lambda_{k+1}}{\lambda_k}, \frac{\lambda_{k+2}}{\lambda_{k+1}}, \dots, \frac{\lambda_{2k+1}}{\lambda_{2k}}$  siano uguali alla media aritmetica dei  $k$  rapporti stimati  $\frac{\hat{\lambda}_1}{\hat{\lambda}_0}, \frac{\hat{\lambda}_2}{\hat{\lambda}_1}, \dots, \frac{\hat{\lambda}_k}{\hat{\lambda}_{k-1}}$ .

Osserviamo adesso che è possibile individuare i valori di  $\lambda_{i+j}$  solo per  $i+j = 0, 1, \dots, k$ . Per le restanti quantità  $\lambda_{i+j}$ , con  $i+j > k$ , ( $i+j = k+1, \dots, 2k+1$ ), occorre effettuare opportune ipotesi di previsioni future.

Di solito si assume che i  $k+1$  rapporti  $\frac{\lambda_{k+1}}{\lambda_k}, \frac{\lambda_{k+2}}{\lambda_{k+1}}, \dots, \frac{\lambda_{2k+1}}{\lambda_{2k}}$  siano costanti ed uguali alla media aritmetica dei  $k$  rapporti precedentemente stimati  $\frac{\hat{\lambda}_1}{\hat{\lambda}_0}, \frac{\hat{\lambda}_2}{\hat{\lambda}_1}, \dots, \frac{\hat{\lambda}_k}{\hat{\lambda}_{k-1}}$ .

Osserviamo l'analogia di questa ipotesi con quella utilizzata nel metodo della catena per stimare i tassi di inflazioni relativi agli anni successivi la valutazione.

A questo punto abbiamo a disposizione tutti i valori necessari per ottenere una stima della riserva IBNER.

In particolare la stima che compete alla generazione  $i$  risulta essere l'importo

$$R_i = \sum_{j=k-i+1}^k \hat{Y}_{ij} + \hat{R}_{i,k+1};$$

da cui deriva la riserva totale

$$R_{(k)}^{IBNER} = \sum_{i=1}^k R_i + R_{0,k+1}.$$

Riprendendo l'espressione riferita ai risarcimenti non cumulati:

$$R_{(k)}^{IBNER} = \sum_{i=1}^k \left( \sum_{j=k-i+1}^k \hat{Y}_{ij} + \hat{R}_{i,k+1} \right) + R_{0,k+1}.$$

Notiamo che sotto la condizione di normalizzazione e l'ipotesi di costanza di  $\lambda_{i+j}$  al variare di  $i$  e di  $j$ , i fattori  $r_j$  assumerebbero il significato di rapporto tra il pagato con  $j$  anni di differimento ed il totale pagato entro i primi  $k$  anni di differimento, indipendentemente dalla generazione. Si realizza così l'ipotesi di invarianza nel tempo della "legge di smontamento" dei sinistri e come diretta conseguenza, se vale la condizione di normalizzazione, i fattori  $\lambda_{i+j}$  risultano denotare l'importo medio di risarcimento che si avrebbe se il costo fosse lo stesso in ogni periodo.

In ultima analisi le considerazioni che si possono fare per il metodo della catena valgono anche per il metodo della separazione. Risulta difficile effettuare una scelta motivata del livello futuro di inflazione a causa dell'assenza di un reale esame del meccanismo del processo di smontamento dei sinistri. Un'ulteriore difficoltà sta nella superficialità dell'interpretazione di  $\lambda_{i+j}$  come indice di inflazione. Esso può essere visto come un riassunto di tutte le influenze esogene operanti nell'anno  $i+j$ , visione un po' semplicistica con cui pertanto ci troviamo ad operare utilizzando questo metodo.

### 3.4 Il metodo dei minimi quadrati

Il metodo dei minimi quadrati fu inizialmente introdotto da De Vylder in un lavoro che risale al 1978 e che si occupava della stima della riserva IBNR (quella relativa alle tardive denunce). Con il passare del tempo questo modello ha trovato largo impiego anche per la valutazione dell'altra componente della riserva sinistri, la riserva IBNER. Come dice il nome il metodo si basa sul criterio dei minimi quadrati. La base informativa è costituita dai singoli risarcimenti,  $Y_{ij}$ , relativi quindi alla generazione  $i$  e liquidati con  $j$  anni di differimento e dall'accantonamento iniziale a riserva  $R_{0,k+1}$ . Questo metodo tiene in considerazione il fatto che non tutti i dati del triangolo superiore della matrice di run-off, quelli cioè che costituiscono la base statistica, siano necessariamente noti o sufficientemente attendibili.

In altre parole si ha a disposizione solo un insieme di informazioni parziali e se si indica con  $\Omega$  il sottoinsieme delle coppie  $(i,j)$  relative ai dati del triangolo di run-off repute affidabili, la stima della riserva IBNER si basa su di un ipotesi di fattorizzazione del risarcimento non cumulato che si esprime nel seguente modo:

$$Y_{ij} = p_i r_j;$$

dove

- $p_i$  è una quantità che dipende dall'anno di generazione  $i$  e ne esprime il risarcimento complessivo;
- $r_j$  è un parametro adimensionale che dipende dall'anno di differimento  $j$  e rappresenta la legge di smontamento dei sinistri.

Le stime ottime dei fattori  $p_i$  ( $i=0,1,\dots,k$ ) e  $r_j$  ( $j=0,1,\dots,k$ ) si ricavano utilizzando la tecnica di proiezione del “grossing up” e sono tali da minimizzare la somma

$$\sum_{(i,j) \in \Omega} (Y_{ij} - p_i r_j)^2$$

dove  $\Omega$  ricordiamo essere l'insieme delle coppie  $(i,j)$  effettivamente disponibili.

La soluzione del problema è individuata a meno di una costante moltiplicativa  $\epsilon$  non nulla ed è univocamente determinata nel caso si assuma la condizione di

normalizzazione  $\sum_{j=0}^k r_j = 1$ .

In particolare la soluzione normalizzata del problema è ottenuta utilizzando un procedimento iterativo applicato alle seguenti quantità

$$(a) \quad p_i = \frac{\sum_j Y_{ij} r_j}{\sum_j r_j^2};$$

$$(b) \quad r_j = \frac{\sum_i Y_{ij} p_i}{\sum_i p_i^2};$$

ed assegnando arbitrariamente valori iniziali ai fattori  $r_0, r_1, \dots, r_k$ .

In questo modo si ottengono le stime

$$\hat{Y}_{ij} = \hat{p}_i \hat{r}_j$$

relative ai risarcimenti non cumulati del triangolo inferiore della matrice e

$$\hat{R}_{i,k+1} = R_{0,k+1} + \frac{\hat{p}_i}{\hat{p}_0}$$

relative agli accantonamenti a riserva, (nonché termini aggiuntivi del modello).

In conseguenza una stima per la riserva IBNER espressa mediante la solita formula relativa ai risarcimenti non cumulati è data da

$$R_{(k)}^{IBNER} = \sum_{i=1}^k \left( \sum_{j=k-i+1}^k \hat{Y}_{ij} + \hat{R}_{i,k+1} \right) + R_{0,k+1}.$$

### 3.5 Il metodo di Bornhuetter-Ferguson

Il metodo di Bornhuetter-Ferguson risale al 1972 e deve la sua popolarità ai limiti del metodo della catena. Tale metodo ha infatti avuto il suo sviluppo nel momento in cui i risultati ottenuti con il metodo della catena sono divenuti insoddisfacenti a causa dell'instabilità presente nella proporzione tra risarcimenti appartenenti a diversi anni di differimento, vale a dire quando l'ipotesi fondamentale alla base di questo modello è venuta a mancare. È nata quindi l'esigenza di tentare di stabilizzare i risultati con qualche meccanismo al fine di rendere le valutazioni più attendibili e veritiere. Una possibile idea è stata quella di utilizzare una stima iniziale esterna relativamente agli esborsi che la compagnia deve effettuare ed a cui poi vengono applicati i classici fattori di sviluppo del metodo della catena o altri simili. In questo modo è possibile disporre di informazioni endogene circa il risarcimento relativo ad ogni sinistro ed in particolare circa le sue modalità di liquidazione. Ricordiamo che utilizzando il metodo della catena la stima dei risarcimenti per sinistri rimasti in sospeso è data da

$$\hat{X}_{ij} = X_{ik-i} \prod_{h=k-i+1}^j m_h, \quad i = 1, 2, \dots, k; \quad j = k - h + 1, \dots, k.$$

In altre parole gli importi cumulati più recenti, ossia quelli che si trovano sulla diagonale della matrice di run-off, vengono moltiplicati per il prodotto dei fattori di sviluppo successivi a quello in esame per ottenere una stima dell'ultimo risarcimento cumulato.



La valutazione della riserva è poi ottenuta come differenza fra questo importo e quello che si trova sulla diagonale, a cui vengono sommati i termini aggiuntivi, ulteriori accantonamenti a riserva relativi ai sinistri non ancora liquidati dopo il termine massimo prefissato inizialmente, che nel caso trattato coincide con l'anno  $k$  ed è un elemento prescelto dalla compagnia nel momento in cui viene costruito il triangolo di run-off. Utilizzando il metodo introdotto da Bourhuetter-Ferguson il procedimento rimane sostanzialmente invariato con l'unica, ma non trascurabile, differenza che riguarda la stima dell'ultimo risarcimento, il quale risulta adesso valutato sulla base di informazioni esogene e giudizi esperti. Queste informazioni esterne si concretizzano nella scelta di un modello di tipo baynesiano.

In particolare l'informazione a priori di cui sopra può essere riassunta utilizzando una distribuzione di probabilità per i parametri di riga opportunamente prescelta ed unico elemento di differenza tra i due metodi in questione. Molto spesso viene utilizzata una distribuzione di tipo Gamma. In questo caso i parametri di riga,  $X_i$ , vengono assunti indipendenti tra loro e distribuiti quindi come una distribuzione di tipo Gamma di parametri  $(\alpha_i, \beta_i)$ . Ne consegue che la valutazione dell'ultimo risarcimento cumulato, che corrisponde al valore atteso dei parametri di riga, è uguale a  $\frac{\alpha_i}{\beta_i}$ . Fondamentalmente possiamo concludere che questa tecnica di stima

della riserva presuppone che vi sia una perfetta informazione a priori circa i parametri di riga. Essa si materializza nella scelta di una appropriata distribuzione di probabilità e pertanto per questa parte di stima non vengono utilizzati interamente i dati a disposizione.

### 3.6 Il metodo di Fisher-Lange

Il metodo di Fisher-Lange rientra fra quelli che utilizzano il risarcimento medio in relazione ai sinistri liquidati e il costo medio dei sinistri in relazione ai sinistri registrati.

Per tale ragione la base informativa di partenza è costituita da:

- i risarcimenti cumulati,  $X_{ij}$  ;
- il termine aggiuntivo o accantonamento iniziale a riserva,  $R_{0,k+1}$  ;

- il numero complessivo dei sinistri relativo alla generazione  $i$ ,  $N_i$ ;
- il numero cumulato dei sinistri relativi alla generazione  $i$  e complessivamente liquidati non oltre i primi  $j$  anni di differimento,  $N_{ij}^C$ .

La stima della riserva si basa sull'ipotesi che sia il risarcimento medio sia la velocità di liquidazione dei sinistri non dipendono dall'anno di generazione  $i$ , ma soltanto dal numero  $j$  di anni di differimento.

È necessario pertanto ottenere le stime dei seguenti due fattori:

- $q_i$ ,  $i = 0, 1, \dots, k$ , che rappresenta l'evoluzione del risarcimento medio relativamente alle varie generazioni;
- $w_j$ ,  $j = 0, 1, \dots, k$ , che rappresenta l'evoluzione della velocità di liquidazione dei sinistri nel corso dei diversi anni di differimento.

Tali quantità vengono stimate a partire dai dati disponibili e operando con un procedimento analogo a quello utilizzato nel metodo della catena. Per quanto concerne le proiezioni, di solito si fa ricorso ai soli dati relativi all'ultimo anno a disposizione (valori che si trovano sulla diagonale della tabella di run-off), nonché anno di esercizio in chiusura, che nel caso trattato coincide con l'anno  $k$ . Tuttavia si può facilmente estendere il meccanismo utilizzando una maggiore quantità di dati a disposizione, opportunamente ponderati sulla base della loro epoca di riferimento e con adeguate medie aritmetiche, semplici o ponderate, attuando in questo modo un miglioramento della stima causato dall'utilizzo di una informazione più raffinata. I risultati sono espressi in termini di risarcimenti medi, i cui valori sono facilmente reperibili a partire dagli importi cumulati.

Essi si ottengono facendo le differenze tra i risarcimenti cumulati e dividendo le quantità così ottenute per il numero di sinistri complessivamente registrato per una certa generazione  $i$ . I costi medi sono gli elementi su cui fare riferimento per conseguire le stime dei due fattori,  $q_i$  e  $w_j$ .

Occorre infine osservare che per ottenere le proiezioni dei futuri esborsi è inoltre necessario formulare opportune ipotesi circa l'andamento dell'inflazione nel corso degli anni futuri la valutazione,  $i+j$ , ( $i+j=k+1, \dots, 2k+1$ ), così come accadeva per il metodo della catena.

Attraverso l'utilizzo di questa procedura è possibile conseguire le seguenti stime:

- le stime  $\hat{Y}_{ij}$  dei risarcimenti non cumulati relativi al triangolo inferiore della matrice di run-off;
- le stime  $\hat{N}_{ij}$  del numero cumulato di sinistri relativi al triangolo inferiore della matrice di run-off;
- le stime  $\hat{N}_{i,k+1}$  del numero cumulato di sinistri relativi alla generazione  $i$  non ancora liquidati dopo  $k$  anni di differimento;
- le stime  $\hat{R}_{i,k+1}$  degli accantonamenti a riserva.

A partire da queste stime è possibile ricavare anche quelle relative ai risarcimenti cumulati  $\hat{X}_{ij}$  del triangolo inferiore della matrice di run-off: in particolare basta utilizzare il procedimento inverso a quello precedentemente descritto per individuare i singoli risarcimenti a partire dagli importi cumulati. Da questi valori risulta poi semplice ottenere una stima della riserva IBNER tramite la formula relativa ai risarcimenti cumulati

$$R_{(k)}^{IBNER} = \sum_{i=1}^k [(\hat{X}_{ik} - X_{ik-i}) + \hat{R}_{i,k+1}] + R_{0,k+1}.$$

Per concludere è interessante osservare che tale metodo non è stato inizialmente creato come criterio di valutazione della riserva sinistri, bensì come strumento di controllo. Sostanzialmente è una specie di spia utilizzata per verificare la bontà del modello effettivamente impiegato ai fini della valutazione della riserva sinistri stessa.

Se il risultato ottenuto con Fisher-Lange è simile a quello ottenuto con il metodo di stima prescelto dall'impresa di assicurazione, allora il criterio adottato risulta essere valido, viceversa l'assicuratore può disporre di un segnale di allarme circa l'attendibilità delle stime effettuate. A tale proposito, però, occorre anche osservare che il metodo può dare risultati distorti nel caso in cui i calcoli si basino soltanto sui dati più recenti.

### 3.7 Il metodo di Tarbell per la stima della riserva IBNR

Come descritto nel precedente paragrafo la riserva per sinistri IBNR viene creata al fine di tutelarsi dalle tardive denunce. Le stime si affidano all'esperienza passata valutando l'influenza quantitativa dei risarcimenti a fronte di ritardi nelle denunce dei sinistri. In particolare le valutazioni sono di solito effettuate in termini di costi medi relativamente ad ogni anno di esercizio. Nella pratica la componente IBNR viene aggiunta alla IBNER per ottenere così la stima complessiva della riserva sinistri. I metodi di valutazione possono essere gli stessi utilizzati anche per la riserva IBNER basta sostituire:

- con l'anno di generazione l'anno di accadimento;
- con l'anno di differimento l'anno di denuncia.

In proposito si segnala in particolare il metodo dei minimi quadrati che ha il vantaggio di operare anche in presenza di una base informativa carente, caratteristica spesso diffusa in presenza di dati relativi alle tardive denunce. Nonostante ciò, anche per la riserva sinistri IBNR sono stati ideati vari metodi statistici, che tuttavia risultano essere sempre in netta minoranza rispetto alla numerosità di quelli introdotti per la valutazione della componente IBNER. Nel seguito riportiamo il metodo di Tarbell, sul quale viene spesso rivolta particolare attenzione e che si occupa dei ritardi nelle denunce.

Gli elementi su cui si basa sono i seguenti:

- i risarcimenti,  $Y_{1,\dots,12}^{(k-1)}$ , che si riferiscono ai sinistri accaduti nell'anno  $k-1$  e tardivamente denunciati nel corso dei dodici mesi dell'anno  $k$  (costo complessivo per le tardive denunce relative dell'anno  $k-1$ );
- i risarcimenti medi,  $y_{10,11,12}^{(k)}$ , che si riferiscono ai sinistri liquidati nell'ultimo trimestre dell'anno  $k$  (costo medio per sinistri denunciati negli ultimi tre mesi dell'anno  $k$ );
- il numero di sinistri,  $n_{10,11,12}^{(k)}$ , denunciati nell'ultimo trimestre dell'anno  $k$  (numero di sinistri denunciati negli ultimi tre mesi dell'anno  $k$ ).

L'ipotesi fondamentale del modello è che ogni sinistro sia denunciato al massimo entro un anno dal suo accadimento. Sotto queste considerazioni la stima per la riserva sinistri IBNR alla fine dell'anno  $k$  è data da:

$$\hat{R}_k^{(IBNR)} = \left( \frac{n_{10,11,12}^{(k)}}{n_{10,11,12}^{(k-1)}} \right) \left( \frac{y_{10,11,12}^{(k)}}{y_{10,11,12}^{(k-1)}} \right) Y_{1,\dots,12}^{(k-1)} .$$

Tale stima risulta essere il prodotto tra l'importo dei risarcimenti,  $Y_{1,\dots,12}^{(k-1)}$ , relativi ai sinistri accaduti nell'anno  $k-1$  e tardivamente denunciati nel corso dei dodici mesi dell'anno  $k$  e un fattore di correzione ex-post che aggiusta la stima di  $Y_{1,\dots,12}^{(k-1)}$ .

In proposito osserviamo che il valore  $Y_{1,\dots,12}^{(k-1)}$  potrebbe già di per se rappresentare una stima della riserva IBNR in quanto esprime l'importo destinato a coprire le tardive denunce relative all'anno precedente la valutazione stessa. Tuttavia è utile sottolineare l'importante ruolo del fattore di correzione: si può pensare che esso rappresenti i possibili cambiamenti verificatisi nel corso del tempo e dovuti a fattori esterni o interni la gestione assicurativa stessa (rispettivamente fattori esogeni o endogeni), primo fra tutti l'effetto dell'inflazione con le conseguenti ripercussioni sui costi medi per risarcimento relativamente ad ogni sinistro. In ultima analisi si può osservare che le stime ottenute attraverso il metodo della catena e quello di separazione sono attribuite all'anno di accadimento e quindi, nel calcolo delle riserve per mezzo di queste tecniche, si includono implicitamente i sinistri IBNR per tutte le generazioni tranne l'ultima<sup>14</sup>.

---

<sup>14</sup> Sempre sotto la condizione che la registrazione di un sinistro avvenga entro dodici mesi dal suo accadimento.

## CAPITOLO 4

### MODELLI STOCASTICI PER LA STIMA DELLA RISERVA SINISTRI

#### 4.1 Premesse

Nel capitolo precedente sono stati presentati modelli di tipo deterministico i quali fanno solamente assunzioni che riguardano il valore atteso degli esborsi futuri. Tuttavia, il quadro normativo che si sta delineando nell'ambito della Comunità Europea in merito ai principi contabili ed ai requisiti di capitale necessari ai fini della solvibilità per le imprese assicuratrici operanti nei rami danni, rende, di fatto, fondamentale basare le valutazioni delle grandezze rilevanti su modelli di tipo stocastico. In particolare, per quanto disposto con i principi contabili, è previsto che attività e passività siano valutate a fair value, ossia in modo consistente con il mercato. In riferimento, ad esempio, alle passività relative a flussi di pagamenti futuri per sinistri già avvenuti il fair value si può ottenere attualizzando con un tasso per attività prive di rischio la somma di certi equivalenti di ciascun esborso. A sua volta il certo equivalente di un esborso si può ottenere sommando alla sua speranza matematica un margine per il rischio che, in modo consistente con il mercato, tenga conto sia del rischio del processo, legato all'aleatorietà intrinseca del pagamento, sia del rischio di stima, legato al procedimento utilizzato per determinare i parametri del modello introdotto per descrivere il flusso aleatorio dei pagamenti.

In sostanza:

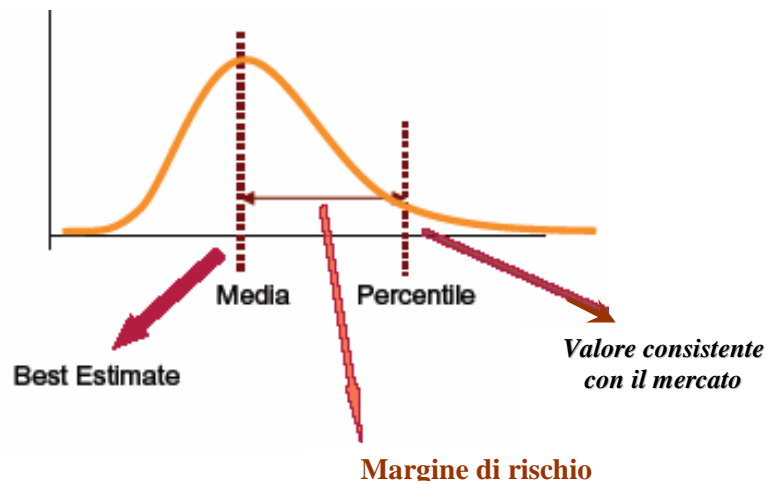
$$\text{Riserve tecniche a valori di mercato} = \text{Riserve tecniche Best Estimate} + \\ \text{Margine per il Rischio.}$$

La prima componente (Riserve tecniche Best Estimate), determinata dalla speranza matematica dei futuri esborsi, serve a catturare implicitamente la volatilità dei rischi cosiddetti diversificabili (hedgeable), ossia tutti quei rischi che possono essere trasferiti, attraverso l'utilizzo di un portafoglio replicante, ad un terzo soggetto (per esempio la maggior parte dei rischi di tipo finanziario).

Il margine per il rischio, invece, è posto a copertura della volatilità dei rischi non diversificabili (non hedgeable), aspetto non catturato nelle riserve best estimate e, quindi, con i modelli deterministici. Attualmente si considerano implementabili due diverse metodologie valutative del margine di rischio: la prima, basata sul costo del capitale, utilizzata principalmente nel settore vita dal Swiss Solvency Test o come elemento aggiuntivo opzionale dal CEIOPS (Committee of European Insurance and Occupational Pension Supervisors) per il calcolo del QIS2 (Quantitative Impact Study); la seconda, tipica delle assicurazioni contro i danni, segue l'approccio dei percentili, metodo richiesto dal CEIOPS per il QIS2.

Quest'ultimo si compone sostanzialmente di tre fasi:

- la definizione di una distribuzione di probabilità delle riserve tecniche per ogni tipologia di rischio non diversificabile;
- il calcolo della media (best estimate) e del percentile della distribuzione (ad esempio il 75°):
- il calcolo del margine per il rischio come differenza tra i due valori ottenuti sopra.



**Figura 4: Il margine di rischio**

È evidente che l'approccio dei percentili richiede la definizione di distribuzioni di probabilità, modelli stocastici, parametri complessi e inusuali nel settore vita per i rischi non diversificabili. Le richieste normative prevedono di determinare, a fronte dei flussi aleatori di pagamenti futuri, importi, detti capitali per il rischio, in grado di assicurare che, con elevata probabilità, l'impresa possa fare fronte ai

propri impegni. La valutazione delle componenti di rischio nel settore assicurativo danni si basa dunque sulla costruzione della distribuzione del costo ultimo attraverso modelli stocastici e sulla scelta dell'indice sintetico di variabilità della riserva attuariale utilizzato per quantificare il rischio di riserazione, sia esso il percentile o la varianza della distribuzione adottata. I temi appena trattati sono oggetto di dibattito assicurativo ormai da molti anni, così come appare nella recente letteratura attuariale.

Ai fini della determinazione sia del fair value sia del capitale di rischio, sono utilizzati diversi modelli stocastici che verranno in parte proposti in questo capitolo. In particolare, i modelli lineari generalizzati sono già stati ampiamente impiegati nella valutazione delle riserve. Si tratta di una classe di modelli statistici che rappresentano una generalizzazione naturale dei modelli lineari classici. La loro peculiarità consiste nella flessibilità insita nel modello. Questi strumenti statistici consentono di scegliere in un'ampia classe le distribuzioni della variabile risposta e di considerare diverse strutture di regressione e funzioni di collegamento nonché di analizzare la bontà dell'accostamento ai dati in modo tale da controllare anche il rischio del modello stesso.

Prima di passare in rassegna alcuni dei principali modelli stocastici attualmente in uso nella pratica assicurativa cerchiamo di delineare ancora una volta le principali differenze tra un modello deterministico ed un modello di tipo stocastico. Innanzitutto, i primi possono sì da un lato produrre una stima "migliore" (best estimate) all'interno del ventaglio dei risultati possibili, ma non sono in grado di fornire una misura della precisione di questa stima. Tale questione si supera con i modelli stocastici i quali fanno assunzioni riguardanti la variabilità dei futuri pagamenti altrochè il valore atteso dei medesimi. Pertanto essi risultano essere in grado di fornire, accanto alla migliore stima dei futuri esborsi, anche una misura di precisione di quest'ultima, trattando il processo che porta alla determinazione della riserva come un esercizio di analisi dei dati su cui basarsi per costruire modelli inseriti in un contesto di lavoro di tipo statistico che consenta controlli in grado di testare, ad esempio, la bontà di adattamento o l'analisi dei residui. È importante sottolineare inoltre come i modelli deterministici siano spesso utilizzati senza una chiara conoscenza delle assunzioni che si stanno facendo;



cosa che non accade con i modelli stocastici in quanto basati su ipotesi del tutto esplicite ed in grado di essere testate con una grande varietà di tecniche. Un'altra importante questione che deriva da quanto appena detto è quella della stima: un modello di tipo deterministico fornisce semplicemente una stima di tipo puntuale dei futuri esborsi. È evidente invece che i pagamenti reali potranno differire da quelli stimati e, mentre i modelli deterministici non danno un'idea di quanto possa essere significativo questo scostamento, i modelli stocastici forniscono invece un intervallo entro i cui limiti ci si può attendere cada l'importo dei futuri esborsi con un determinato livello di confidenza. A sfavore dei modelli stocastici sta invece il fatto che essi modellano una serie di eventi molto ampia attraverso pochi parametri. Si hanno pertanto critiche sul fatto che si facciano assunzioni troppo semplicistiche, a volte irreali, dovute anche alla rigidità di questi modelli che non consentono di incorporare giudizi o fattori estranei. Inoltre, anche l'applicazione pratica e computazionale risulta spesso complessa da attuare e richiede un'abilità statistica e di calcolo non indifferente. Pertanto anche i risultati ottenuti risultano spesso di comparazione non semplice rispetto a quelli ottenuti con metodi deterministici. Nonostante ciò i modelli stocastici risultano essere molto utili, basti pensare al fatto che essi riescono dove gli altri falliscono. Comunque l'utilità che maggiormente ci interessa è quella che deriva dal fatto che questi modelli sono in grado di fornire maggiore informazione rispetto agli altri, cosa che può risultare utile alla compagnia non solo nel processo di costituzione della riserva, ma anche nella sua gestione complessiva.

## **4.2 Modelli chain-ladder**

### **4.2.1 Introduzione**

Il metodo della catena, o chain ladder method, è probabilmente il metodo più diffuso per la stima della riserva sinistri. Il motivo principale del suo largo impiego risiede nella sua semplicità applicativa e nel fatto che non occorre formulare alcuna ipotesi statistica relativamente al modello da utilizzare per rappresentare la distribuzione dei dati di partenza. Tuttavia, è anche noto come tale metodo sia molto sensibile alle variazioni dei dati osservati, specialmente per quanto riguarda la stima della riserva

relativa agli anni più recenti la valutazione. Ricordiamo in proposito che l'ipotesi basilare di questa tecnica di stima consiste nella costanza nel tempo di quel che è definita la legge di smontamento dei sinistri o, in altre parole, la progressione con cui nel tempo vengono liquidati i sinistri avvenuti. È evidente invece come nel corso degli anni si possano riscontrare cambiamenti di vario genere, sia interni che esterni all'impresa assicuratrice stessa, e che influenzano sensibilmente la velocità con cui avvengono i risarcimenti (ad esempio, la politica gestionale prescelta dall'impresa o l'inflazione presente sul mercato).

La tecnica del metodo della catena è sostanzialmente un metodo che consente di ottenere previsioni attraverso l'utilizzo dei fattori che legano ogni colonna a quella successiva, fino ad ottenere la stima dell'ultimo<sup>15</sup> esborso e fornendo in questo modo una stima di tipo puntuale. Anche questo risulta essere un grosso limite del metodo in quanto, come più volte sottolineato nei precedenti capitoli, una valutazione di questo tipo risulta essere carente da un punto di vista statistico. L'insufficienza si riscontra nel fatto che accanto ad un valore atteso non si legge una stima della variabilità, importante per le decisioni della compagnia<sup>16</sup>. Inoltre, nel corso degli anni sono state utilizzate molte altre procedure statistiche per il calcolo della riserva sinistri ed alcune di queste hanno dato risultati di gran lunga diversi da quelli prodotti dal metodo della catena. Pertanto, a maggior ragione risulta fondamentale riuscire a fornire, accanto alla stima puntuale della riserva, anche una stima della variabilità della stessa. A questo proposito, la misura di variabilità che di solito si utilizza è l'errore di predizione, definito come deviazione standard della distribuzione dei possibili risultati della riserva. Tale errore ha lo scopo di catturare l'incertezza presente nei dati ed è in grado di cogliere se vi sono differenze significative tra il metodo della catena e le altre possibili tecniche di stima. Il primo passo che si deve compiere per ottenere l'errore di predizione è quello della formulazione di un modello statistico per la distribuzione dei dati. Se l'intento è quello di fornire un modello stocastico analogo al metodo della catena, allora per prima cosa si deve richiedere che i valori che esso produce devono essere gli stessi della tecnica del chain-ladder. Per ottenere ciò vi sono due metodi: specificare le distribuzioni per i dati, altrimenti specificare i primi due momenti.

---

<sup>15</sup> Ultimo inteso come "maggior ritardo possibile" che è stato osservato.

<sup>16</sup> Ad esempio su come effettuare gli investimenti o sulla costituzione di riserve extra.

Nel corso degli anni sono stati esaminati diversi metodi al fine di determinare l'errore di predizione della riserva. Negli anni '80 Taylor e Ashe (1983), Zehnwirth (1985) e Renshaw (1989) hanno utilizzato la tecnica statistica della regressione lineare applicata ai logaritmi dei pagamenti incrementali e la procedura dei minimi quadrati per ottenere le stime dei parametri del modello. Tale approccio è stato poi ripreso negli anni '90 da Christofides (1990) e Verrall (1990 e 1991); mentre Wright (1990) e Mack (1991) hanno seguito orientamenti del tutto diversi. Il primo si è basato sui modelli lineari generalizzati; mentre il secondo ha utilizzato una distribuzione gamma e stime di massima verosimiglianza per valutare i pagamenti futuri ed in conseguenza l'entità della riserva. Nei seguenti paragrafi verranno presentati i modelli appena descritti.

#### 4.2.2 Modelli lineari e tecniche chain-ladder

Vediamo adesso com'è possibile rappresentare il metodo della catena mediante un modello di tipo lineare. Partiamo dall'ipotesi di disporre i dati in forma triangolare e ricordiamo che il metodo chain ladder si basa sugli importi cumulati,  $X_{ij}$ , e che la stima dei rapporti tra risarcimenti relativi a due consecutivi anni di differimento si ricava con la tecnica del "link ratio", tecnica di proiezione esposta nel capitolo precedente. Ricordiamo, inoltre, che l'ipotesi fondamentale che sta alla base del modello è che il rapporto  $X_{ij}/X_{ij-1}$  non dipende dall'anno di origine, ma solo dal numero di anni di differimento  $j$ . In altre parole, esiste una "legge di smontamento" dei sinistri, che rappresenta la progressione dei risarcimenti cumulati nel corso degli anni di differimento, costante per ogni anno di generazione, a meno di variazioni aleatorie.

In sostanza:

$$\frac{X_{01}}{X_{00}} = \frac{X_{11}}{X_{10}} = \frac{X_{21}}{X_{20}} = \dots$$

Sotto tali ipotesi, la stima dei rapporti  $m_j = \frac{X_{ij}}{X_{ij-1}}$ , relativa agli importi cumulati di

due colonne adiacenti, viene effettuata nel seguente modo:

$$\hat{m}_j = \frac{\sum_{i=0}^{k-j} X_{ij}}{\sum_{i=0}^{k-j} X_{ij-1}}, \quad j = 1, 2, \dots, k.$$

Una volta determinati tutti i coefficienti  $m_i$ ,  $i = 1, 2, \dots, k$ , è possibile ottenere delle stime per i futuri esborsi e, quindi, valutare la riserva aggiungendo al valore finale i termini aggiuntivi (importi cumulati finali) che verranno stimati in modo del tutto analogo.

I rapporti  $m_i$ ,  $i = 1, 2, \dots, k$ , individuano infatti dei parametri che ci consentono di passare da una colonna a quella successiva, ovvero:

$$\hat{X}_{ij} = X_{ik-i} \prod_{h=k-i+1}^j m_h, \quad i = 1, 2, \dots, k; \quad j = k - h + 1, \dots, k.$$

Sostanzialmente il modello che sta alla base del metodo della catena si può riassumere in

$$E(X_{ij} / X_{i1}, X_{i2}, \dots, X_{ij-1}) = m_j \cdot X_{ij-1}, \quad j = 1, 2, \dots, k;$$

dove

$$m_j = \frac{\sum_{i=0}^{k-j} X_{ij}}{\sum_{i=0}^{k-j} X_{ij-1}}, \quad i = 1, 2, \dots, k.$$

Si tratta di un modello molto semplice ed immediato le cui stime relative agli esborsi futuri hanno effetti sia di riga che di colonna. L'effetto colonna è evidente ed è rappresentato dal parametro  $m_j$ ,  $j = 1, 2, \dots, k$ . L'effetto riga è invece espresso dal termine  $X_{ik-i}$ , ultimo esborso disponibile e che si trova dunque sulla diagonale della matrice dati.

Per poter rappresentare il metodo della catena mediante un schema lineare è pertanto necessario considerare un modello a due parametri, in modo tale da catturare sia l'effetto riga che l'effetto colonna. In particolare, viene utilizzato un approccio moltiplicativo applicato ai dati non cumulati,  $Y_{ij}$ , in accordo a:

$$E(Y_{ij}) = C_i \lambda_j,$$

dove:

- $C_i$  individua il parametro di riga,  $i$ ,
- $\lambda_j$  individua il parametro di colonna,  $j$ .

Ovviamente  $\sum_{j=1}^k \lambda_j = 1$  e sono previsti errori di previsione con struttura moltiplicativa.

Sostanzialmente:

- $\lambda_j$  rappresenta la proporzione attesa dei risarcimenti finali effettuati con  $j$  anni di differimento;
- $C_i$  rappresenta il risarcimento atteso complessivo relativo alla generazione  $i$  (esclusi i termini aggiuntivi).

Kremer (1982) ha mostrato che valgono le seguenti relazioni tra i parametri del modello moltiplicativo sopra introdotto ed i valori calcolati con il metodo chain ladder:

$$\lambda_1 = \frac{1}{\prod_{i=2}^k m_i}; \quad \lambda_j = \frac{m_j - 1}{\prod_{i=j}^k m_i}, \quad j \geq 2;$$

$$C_i = E(X_{ik}).$$

In conseguenza le stime per  $C_i$  ( $i = 1, 2, \dots, k$ ) e  $\lambda_j$  ( $j = 1, 2, \dots, k$ ) possono essere ottenute applicando un modello lineare ai logaritmi dei singoli risarcimenti secondo lo schema:

$$E(Z_{ij}) = \mu + \alpha_i + \beta_j,$$

dove è

$$Z_{ij} = \log Y_{ij}$$

e gli errori di previsione assumono adesso una struttura additiva e si ipotizzano a media nulla. In genere tali errori vengono considerati anche tutti ugualmente distribuiti.

Ai fini esemplificativi si ipotizzi che i dati della nostra matrice siano disponibili per un intervallo temporale di 3 anni. In questo caso il modello lineare assume la forma:

$$\begin{bmatrix} z_{11} \\ z_{12} \\ z_{21} \\ z_{13} \\ z_{22} \\ z_{31} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_2 \\ \alpha_3 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{13} \\ e_{22} \\ e_{31} \end{bmatrix}.$$

Invertendo infine la trasformazione logaritmica si può mostrare che

$$C_i = E(X_{ik}) = e^{\alpha_i} e^{\mu} \sum_{j=1}^k e^{\beta_j}.$$

Lo stesso Kremer ha fatto vedere che nel caso in cui la stima di  $C_i$  avvenga nel seguente modo, allora si ottengono risultati molto simili a quelli che si avrebbero utilizzando il metodo della catena. Se poi le medie geometriche coinvolte nelle valutazioni sono sostituite da medie aritmetiche, allora questo modello lineare ed il metodo chain ladder producono esattamente gli stessi risultati.

In ultima analisi si può osservare che il modello lineare consente di ottenere stime migliori da un punto di vista statistico; tuttavia, per quanto appena detto, non ci sono grosse differenze tra utilizzare il metodo della catena o un opportuno modello lineare: l'unica differenza consiste nella tecnica di stima.

#### 4.2.3 I modelli lineari generalizzati

Nel precedente paragrafo abbiamo visto come il metodo della catena possa essere espresso sotto forma di un opportuno modello lineare. In realtà la linearità si ha nei logaritmi dei singoli pagamenti, in ipotesi dunque che i dati originali abbiano una distribuzione lognormale. Una volta ottenute le stime si ha pertanto il problema di invertire i risultati ottenuti per esprimere i valori finali secondo la scala iniziale della matrice dati. Sostanzialmente abbiamo a che fare con i cosiddetti modelli lineari generalizzati.

Con il termine modelli lineari generalizzati si fa riferimento ad una classe di modelli statistici che rappresentano una generalizzazione naturale dei modelli lineari classici.

Gli elementi comuni ai diversi tipi di modelli lineari generalizzati sono:

- le variabili risposta  $Y_i$  sono fra loro indipendenti e seguono una distribuzione esponenziale con valore atteso  $\mu_i$ ;
- le variabili esplicative esercitano il loro ruolo sotto forma di combinazione lineare dei loro effetti:

$$\sum_j \beta_j x_{ij} = \eta_i,$$

dove  $\eta_i$  è detto predittore lineare ed i valori  $x_{ij}$ , considerati fissi, vengono usati per spiegare la variabilità degli  $Y_i$  (a questo fine è necessario stimare i parametri sconosciuti  $\beta_j$ );

- il valore atteso degli  $Y_i$ , indicato con  $\mu_i$ , è legato al predittore lineare  $\eta_i$  attraverso una funzione nota  $g(\cdot)$ , che viene chiamata funzione di link e tale per cui

$$\eta_i = g(\mu_i).$$

Per definire un modello lineare generalizzato è pertanto necessario specificare tre componenti:

- la componente casuale o stocastica (appartenente alla famiglia esponenziale);
- la componente sistematica (lineare nei parametri della regressione);
- la funzione link (che lega la componente casuale a quella sistematica).

In questo modo il valor medio della variabile  $Y$ , che abbiamo indicato con  $\mu$ , si può esprimere come combinazione lineare delle variabili esplicative  $X_1, \dots, X_n$  e di una funzione  $g(\cdot)$  in accordo a:

$$g(\mu) = \eta = \alpha + \beta_1 X_1 + \dots + \beta_n X_n.$$

Una tale impostazione, molto generale, comprende una grande varietà di modelli utilizzati in diversi tipi di analisi e, come caso particolare, comprende anche il modello di regressione lineare. Se infatti si introduce l'ipotesi che la distribuzione delle  $Y$  sia normale, con media  $\mu$ , e che la funzione di collegamento sia la funzione

identità, siamo nel caso del modello di regressione lineare multipla. Riassumendo, la teoria dei modelli lineari generalizzati permette di mettere in evidenza quanto è comune a numerose procedure di regressione, apparentemente diverse fra di loro, ma che in realtà possono essere studiate per mezzo di una teoria unificata. I modelli lineari generalizzati permettono di trattare modelli non lineari (che possono essere espressi sotto forma di modelli lineari) e variabili di risposta non gaussiane (ma comunque appartenenti alla famiglia esponenziale). In ultima analisi anche i modelli lineari generalizzati richiedono che sia nota la funzione di verosimiglianza (basata sulla famiglia esponenziale) ed assumono che le osservazioni siano tra loro indipendenti. I risultati, che possono essere rappresentati sia da dati discreti sia da dati continui, possono essere utilizzati, oltre che per fini descrittivi, anche per effettuare previsioni sull'evoluzione futura del fenomeno. In particolare, una volta individuata la stima dei futuri esborsi a cui la compagnia di assicurazione deve fare fronte, è possibile fornirne anche un intervallo di previsione in modo da avere anche la variabilità delle stime ottenute. Risulta importante infatti disporre di un valore cosiddetto soglia, al di là del quale l'ammontare complessivo degli esborsi si potrà collocare solo con una certa probabilità presumibilmente bassa e comunque nota all'assicuratore. Se  $R$  rappresenta il valore finale della riserva e  $E(R)$  ne è una sua stima, l'intervallo di confidenza al 95% è dato da quel valore  $k$  tale per cui

$$P(R \leq E(R) + k) = 0.95,$$

ossia

$$P(R - E(R) \leq k) = 0.95.$$

Ciò significa che solo nel 5% dei casi il valore finale della riserva si collocherà fuori dai limiti di questo intervallo.

#### **4.2.4 La curva di Hoerl**

A partire dal modello lineare di tipo chain ladder precedentemente esposto è possibile effettuare scelte di tipo alternativo in modo tale da rappresentare meglio i dati in esame. Nel tempo numerose critiche sono state riferite ai modelli di tipo chain-ladder: una di queste riguarda il fatto che questi modelli utilizzino un singolo parametro per ogni anno di sviluppo, anziché cercare di trovare una sintesi di questi.



Un approccio alternativo riguarda la possibilità di utilizzare una curva parametrica per il modello: una di queste è la curva di Hoerl, detta anche curva Gamma. La scelta di questa funzione è dovuta essenzialmente alla sua somiglianza con i dati del triangolo di run-off: la sua forma cresce molto rapidamente fino a raggiungere il punto di massimo, per poi decrescere in maniera esponenziale. L'introduzione di questa funzione nei modelli log-normali o comunque in quelli che utilizzano una funzione link di tipo logaritmico risulta agevole: basta infatti sostituire il predittore lineare di tipo chain-ladder, con il quale si ottiene un parametro per ogni anno di sviluppo, con

$$\eta_{ij} = \nu_i + \delta_i \log j + \gamma_i(j-1).$$

Pertanto

$$z_{ij} = \nu_i + \delta_i \log j + \gamma_i(j-1) + e_{ij},$$

dove

$$z_{ij} = \log y_{ij}.$$

Invertendo poi la funzione logaritmica, applicando l'esponenziale, si ottiene:

$$e^{\eta_{ij}} = e^{\nu_i} j^{\delta_i} e^{\gamma_i(j-1)}.$$

L'anno di sviluppo  $j$  è dunque adesso una covariata di tipo continua e non discreta così come avveniva nel precedente modello. Il fatto di poter utilizzare una covariata continua risulta un vantaggio nel senso che non si è più costretti ad operare all'interno del campo di variazione temporale offerto dagli anni di sviluppo, cosa che risulta particolarmente utile nel momento in cui si vanno a stimare quelli che sono i parametri di coda. Un altro vantaggio che si ottiene operando nel continuo è che si possono ottenere fattori di sviluppo anche al di là del campo di variazione temporale che si ha a disposizione secondo la tecnica statistica detta estrapolazione. Tuttavia utilizzando questa tecnica si può incorrere in vari rischi. Ad esempio, per alcune combinazioni dei parametri può accadere che la curva di Hoerl, dopo aver mostrato un trend di tipo decrescente, inizi invece a salire quando si attua l'estrapolazione, non rispecchiando così i dati empirici. Un caso di particolare interesse è quello in cui si pone  $\delta_i = \delta$  e  $\gamma_i = \gamma$  per ogni  $i$ . In questo modo il modello ottenuto risulta lo stesso per ogni anno di origine e si hanno quindi solamente due parametri da stimare.

Un metodo che risulta molto efficace nel passare l'informazione da un anno di origine ad un altro in modo tale da ottenere le stime dei parametri  $\delta_i$  e  $\gamma_i$  è l'utilizzo del filtro di Kalman, impiegato per la prima volta in ambito attuariale da De Jong e Zehnwirth (1983). Questo argomento verrà comunque ampiamente trattato nel prossimo capitolo e pertanto adesso non ci soffermiamo oltre.

### **4.3 Stime baynesiane e tecniche di credibilità**

#### **4.3.1 Introduzione**

Procedimenti di stima baynesiani applicati ai modelli lineari sono stati per la prima volta analizzati da Lindley e Smith (1972). Per quanto riguarda la letteratura attuariale e finanziaria, Klugman (1989) si è dedicato allo studio di modelli lineari gerarchici in ambito di stima dei rating aziendali. Abbiamo già visto nel precedente paragrafo come uno dei metodi deterministici più comunemente utilizzato per effettuare la stima dei pagamenti futuri, quale risulta essere il metodo della catena, possa essere riesaminato sotto un ottica di un opportuno modello lineare. Analizzeremo adesso tale aspetto da un punto di vista baynesiano. L'analisi ha sostanzialmente due obiettivi che costituiscono, tra l'altro, i vantaggi dell'utilizzo di un approccio baynesiano. Innanzitutto nel momento in cui vengono effettuate le stime è possibile avvalersi di informazioni disponibili circa i dati in esame, ad esempio fonti che derivano da esperienze passate o da altre compagnie assicurative che presentano caratteristiche simili a quella che si sta analizzando. Tali notizie aggiuntive possono essere impiegate per specificare la distribuzione a priori ed i relativi parametri del modello. Secondariamente non bisogna trascurare il fatto che un'analisi di tipo baynesiano consente in modo naturale di pervenire ad una stima che ha una sua interpretazione nell'ambito della teoria della credibilità. Per quanto riguarda il primo aspetto, la prassi è quella di determinare una distribuzione a priori dei dati in esame sulla base delle informazioni disponibili, per poi giungere ad una distribuzione a posteriori degli stessi ed alle conseguenti stime future. In genere i modelli che vengono utilizzati assumono una distribuzione di tipo lognormale; pertanto è sufficiente specificare solo la media e la varianza della stessa in base all'andamento dei dati. Ad esempio, se l'evidenza empirica suggerisce che i parametri di riga siano tutti circa dell'ordine di 0.1, si può

pensare di utilizzare una distribuzione normale (sulle trasformate logaritmiche dei valori) con media 0.1 e varianza bassa. Qualora invece non sia disponibile un'ampia base dati o siamo in presenza di informazione non molto affidabile, si opterà per la scelta di una varianza più elevata intorno al valore medio selezionato.

Un secondo metodo consiste nel fissare come distribuzione a priori la distribuzione empirica che si ottiene dai dati. In questo modo è possibile giungere a formule di credibilità dove l'obiettivo ultimo è quello di stabilire il giusto peso da dare all'esperienza (quello cioè che si osserva nel tempo) ed in conseguenza alla statistica (quello cioè che si stima dai dati disponibili). Sostanzialmente un approccio così strutturato ricalca il problema analizzato dal Bühlmann (1967) relativo al calcolo di un premio di rischio che presupponesse stime di credibilità: basta pensare agli elementi della matrice dl triangolo di run off (che nel nostro caso individuano i pagamenti per risarcimento) come un insieme di rischi presenti in un portafoglio assicurativo. Occorre osservare però che in questo caso le righe del triangolo di run off contengono un numero diverso di elementi (in particolare un numero decrescente all'aumentare dell'anno di generazione) ed è necessario considerare anche l'influenza dei parametri di colonna. Risulta pertanto ancora più sofisticato e laborioso stabilire il giusto peso da dare al cosiddetto fattore di credibilità (fattore che individua il peso da dare all'esperienza). La letteratura attuariale ha i suoi studi in De Vylder (1982) il quale si è occupato specificatamente dello studio di questo problema.

#### **4.3.2 Stime baynesiane e tecniche chain-ladder**

Vedremo adesso com'è possibile ottenere stime baynesiane in un modello deterministico di tipo chain ladder. Sostanzialmente si tratta di stabilire opportune distribuzioni di probabilità a priori che rappresentano i parametri del modello lineare che individua il metodo chain ladder.

Ricordando infatti che

$$E(Z_{ij}) = \mu + \alpha_i + \beta_j,$$

dove

$$Z_{ij} = \log Y_{ij},$$

si tratta di individuare la distribuzione a priori per  $\mu$ ,  $\alpha_i$  e  $\beta_j$  da cui poi giungere alla distribuzione a posteriori ed alle conseguenti stime. Vediamo, ad esempio, come scegliere una distribuzione a priori per i parametri di riga  $\alpha_i$ . Si ipotizza che i parametri di riga siano tra loro indipendenti ed identicamente distribuiti. Ovviamente si presuppone di non conoscere il loro valore: pertanto è necessario fare ipotesi sul loro andamento. Osserviamo che queste ipotesi sono esattamente le stesse effettuate nella teoria classica della credibilità: quando si vuol calcolare un premio utilizzando la tecnica della credibilità si presuppone che i rischi siano tra loro indipendenti ed identicamente distribuiti. L'insieme dei rischi è poi visto come facente parte di un collettivo da cui stimare la componente di premio che rappresenta la fonte statistica da combinare con un opportuno peso (il fattore di credibilità) con la componente che individua l'esperienza relativa ad ognuno dei singoli rischi presenti nel portafoglio. Nel caso in cui si vuol stimare i pagamenti incrementali la situazione è pertanto esattamente la stessa con l'unica differenza che adesso è fondamentale considerare anche l'effetto congiunto dei parametri colonna della matrice dati. Per farlo, sebbene si ricerchi la giusta distribuzione di probabilità a priori relativa ai parametri di riga, si combina le informazioni che derivano dalle righe della matrice dati con le informazioni che si deducono considerando l'intero triangolo di run off. Pertanto non ci basiamo solo sulle fonti che derivano dalle righe della matrice dati, ma si analizzano tutti gli elementi disponibili nel loro complesso.

Tale procedimento ha un'interpretazione di tipo baynesiano per il fatto che la distribuzione a priori viene stimata dai dati. Senza perdersi troppo nei calcoli, è possibile dimostrare che la stima dei parametri riga  $\alpha_i$ ,  $i = 1, 2, \dots, k$ , che si ottiene utilizzando la tecnica statistica della massima verosimiglianza applicata al modello lineare del tipo

$$z/\underline{\alpha} \approx N(X\underline{\alpha}, \sigma^2 I)$$

ha una forma di credibilità. Infatti, essa è costituita da una media ponderata della stima di massima verosimiglianza globale ed una media della valutazione che se ne trae considerando i dati di riga.

La situazione è complicata dal fatto che le stime di  $\mu$ ,  $\beta_1, \beta_2, \dots, \beta_k$ , coinvolgono quelle di  $\alpha_1, \alpha_2, \dots, \alpha_k$  e viceversa, sebbene ciò sia del tutto ovvio e naturale: le stime dei parametri di riga influenzano sensibilmente le stime dei parametri di colonna e la situazione complessiva della matrice dati espressa dal parametro  $\mu$ . Ma ciò è solo una conferma del fatto che la matrice dati viene esaminata nel suo complesso e non separando gli effetti riga dagli effetti colonna: in questo modo è possibile considerare anche le correlazioni tra i due effetti in questione.

#### **4.4 Tecniche dinamiche di stima**

##### **4.4.1 Modelli state-space**

Nel paragrafo precedente abbiamo analizzato le tecniche baynesiane di stima in cui si presuppone che i parametri di riga siano generati da una stessa distribuzione di probabilità. Risulta evidente pertanto la possibilità di mettere in relazione i diversi anni di generazione. Questo vantaggio non si ha applicando il più comune metodo della catena, mediante il quale ogni anno di generazione viene considerato separatamente. Le stime baynesiane consentono di ovviare a tale inconveniente così come l'utilizzo di un approccio state space, che, sotto certi punti di vista, risulta addirittura migliore rispetto alle tecniche baynesiane. I modelli state space, infatti, stabiliscono un legame ricorsivo tra i parametri che individuano le varie righe della matrice dati, mentre un approccio di tipo baynesiano si basa invece su ipotesi statiche secondo cui tutti i parametri di riga sono simili tra loro (in quanto rappresentati da uno stesso tipo di distribuzione spesso con il medesimo valor medio). In ultima analisi, il metodo chain ladder presenta un altro svantaggio che è possibile superare con i modelli state space: se da un lato non è in grado di considerare le relazioni tra i vari anni di generazione, dall'altro, paradossalmente, impone troppa similitudine tra gli stessi. Il metodo della catena infatti presuppone che l'andamento del triangolo di run off (che individua la velocità di liquidazione dei sinistri) sia lo stesso per ogni anno di generazione: in altre parole, si utilizzano gli stessi fattori di sviluppo (ossia gli parametri che individuano gli effetti di colonna) per ogni anno di generazione.

Un modello state space applicato ai parametri di colonna è in grado di superare tale limite. In generale, un modello state space ha la seguente forma:

$$\begin{aligned}\underline{Y}_1 &= F_1 \underline{\theta}_1 + \underline{e}_1 \\ \underline{Y}_2 &= F_2 \underline{\theta}_2 + \underline{e}_2 \\ &\cdot \\ &\cdot \\ &\cdot \\ \underline{Y}_k &= F_k \underline{\theta}_k + \underline{e}_k\end{aligned}$$

dove  $\underline{Y}_k = \log \underline{Z}_k$ .

Questa equazione è nota come l'equazione di osservazione ed una parte fondamentale del sistema di stato a cui ad esempio può essere applicato un filtro di Kalman al fine di giungere alla stima dei parametri. Il vettore  $\underline{\theta}_k$  è detto vettore di stato ed è legato al precedente valore  $\underline{\theta}_{k-1}$  dalla cosiddetta equazione di sistema. In conseguenza l'equazione di osservazione e l'equazione di sistema descrivono completamente un modello state space. In particolare l'equazione di sistema mette in relazione due vettori di stato consecutivi,  $\underline{\theta}_k$  e  $\underline{\theta}_{k-1}$  ad esempio, e definisce pertanto come il vettore di stato si evolve nel tempo. Il vettore di stato stabilisce l'evoluzione del sistema ed il vettore delle osservazioni  $\underline{Y}_1, \underline{Y}_2, \dots, \underline{Y}_k$  (elementi del modello che di cui vogliamo ottenere una stima) è legato al vettore di stato dall'equazione di osservazione.

Vi sono molte possibili scelte per l'equazione di sistema, la più generale è del tipo:

$$\underline{\theta}_{k+1} = G_k \underline{\theta}_k + H_k \underline{u}_k + w_k,$$

dove  $\underline{u}_k$  è un vettore stocastico dato in input e  $w_k$  è un vettore che rappresenta il disturbo (termine di errore).

In genere si assume che  $\underline{u}_k$  e  $w_k$  siano rappresentate da una distribuzione normale del tipo:

$$\begin{aligned}\underline{u}_k &\approx N(\underline{u}_k, \underline{U}_k), \\ w_k &\approx N(0, \underline{W}_k),\end{aligned}$$

e la scelta di  $G_k$  e  $H_k$  individua la dinamica del sistema.

Vediamo adesso com'è possibile rappresentare un modello lineare di tipo chain ladder con un modello state space.

#### 4.4.2 Rappresentazione di modelli chain-ladder con modelli state-space

Per poter rappresentare il modello lineare di tipo chain ladder con un modello di tipo state space è necessario considerare gli elementi presenti nella matrice dati nel seguente modo:

$$Z_{1,1}, \begin{bmatrix} Z_{1,2} \\ Z_{2,1} \end{bmatrix}, \begin{bmatrix} Z_{1,3} \\ Z_{2,3} \\ Z_{3,1} \end{bmatrix}, \dots,$$

e all'epoca k

$$\begin{bmatrix} Z_{1,k} \\ Z_{2,k-1} \\ \dots \\ \dots \\ Z_{k,1} \end{bmatrix}.$$

Osserviamo che l'insieme dei dati, i quali nel loro complesso formano il triangolo di run off, si presentano sotto la forma di serie storiche. Questo ha il vantaggio di operare con elementi che si generano casualmente in modo naturale. La serie storica dei dati si espande in k. In altre parole il numero di elementi presente in ogni vettore,  $\underline{Z}_1, \underline{Z}_2, \dots, \underline{Z}_k$ , cresce al crescere di k: ad esempio, per l'intero insieme dei dati del triangolo di run off,  $\dim(\underline{Z}_k) = k$ .

Nel caso in cui ci basiamo sui dati relativi ad un arco temporale di tre anni il modello lineare di tipo chain ladder assume una forma del tipo:

$$\begin{bmatrix} Z_{11} \\ Z_{12} \\ Z_{21} \\ Z_{13} \\ Z_{22} \\ Z_{31} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_2 \\ \beta_2 \\ \alpha_3 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{13} \\ e_{22} \\ e_{31} \end{bmatrix};$$

con

$$Z_{ij} = \log Y_{ij}.$$

Per esprimere gli stessi dati in forma ricorsiva occorre esprimere il modello espresso sopra nel seguente modo:

$$Z_{1,1} = \mu + e_{1,1}$$

$$\begin{bmatrix} Z_{1,2} \\ Z_{2,1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_2 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} e_{12} \\ e_{21} \end{bmatrix}$$

$$\begin{bmatrix} Z_{1,3} \\ Z_{2,2} \\ Z_{3,1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_2 \\ \beta_2 \\ \alpha_3 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} e_{1,3} \\ e_{2,2} \\ e_{3,1} \end{bmatrix}.$$

Quelle appena descritte rappresentano le equazioni di osservazione dove il vettore di stato è dato da

$$\begin{bmatrix} \mu \\ \alpha_2 \\ \beta_2 \\ \alpha_3 \\ \beta_3 \end{bmatrix}$$

e, più in generale, il vettore di stato all'epoca k è espresso da

$$\underline{\theta}_k = \begin{bmatrix} \mu \\ \alpha_2 \\ \beta_2 \\ \dots \\ \dots \\ \alpha_k \\ \beta_k \end{bmatrix}.$$

Abbiamo visto che il vettore di stato relativo all'epoca k è legato al vettore di stato relativo all'epoca k-1 mediante la cosiddetta equazione di sistema che, nel caso di un modello lineare di tipo chain ladder, assume la forma:

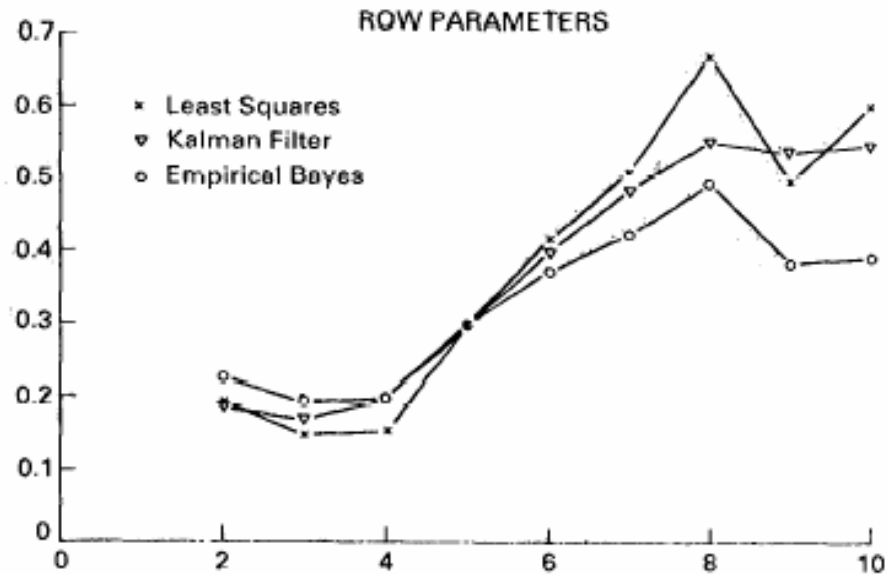




Questa è la relazione che lega tra loro i parametri di riga. Osserviamo che i parametri di colonna non compaiono in questa espressione come se la loro distribuzione di probabilità fosse incerta; tuttavia anche la loro stima varia al variare della stima fornita per i parametri di riga.

Ovviamente le valutazioni sono altamente correlate con la scelta della varianza del processo espressa dal valore  $w_t$ : tanto più alto è il valore assegnato alla varianza e tanto più bassa è la relazione tra i parametri di riga. È inoltre possibile far dipendere il valore della varianza dall'epoca di valutazione,  $k$ , lasciando in questo modo la possibilità di considerare possibili cambiamenti nell'evolversi delle modalità di liquidazione dei sinistri dovute a cambiamenti interni l'impresa, quali ad esempio la politica gestionale aziendale, o cambiamenti esterni all'impresa, primo fra tutti l'inflazione. Un altro approccio è poi quello di stimare la varianza del processo direttamente dai dati, utilizzando così un procedimento di stima empirico (ad esempio, mediante la tecnica statistica della massima verosimiglianza). Le tecniche di stima della varianza possono pertanto essere molteplici. Il valore finale del termine  $w_t$  è comunque basilare sia ai fini della valutazione dei parametri di riga sia sull'impatto che questo valore ha sui parametri colonna, seppure non compaiono esplicitamente nella relazione. In ultima analisi facciamo alcune osservazioni sull'errore standard. In particolare, è possibile affermare che seguendo un approccio baynesiano o basandosi su di un modello state space si ottengono errori standard più bassi e stabili rispetto a procedimenti in cui non si sfruttino alcune informazioni iniziali per stimare i parametri. Il fatto di considerare le informazioni che derivano dai dati a disposizione consente di avere maggiore stabilità nei risultati, cosa che si manifesta con errori standard più bassi.

Il seguente grafico mostra la stima dei parametri riga ottenuta mediante l'utilizzo di un approccio in cui non viene sfruttata l'informazione iniziale; un approccio di tipo baynesiano ed un approccio di tipo state space con stime effettuate mediante un filtro di Kalman.



**Figura 5: Stima dei parametri di riga**

Osserviamo come l'approccio di tipo baynesiano ed il modello state space generino stime il cui andamento non ha inaspettati picchi o oscillazioni troppo marcate, così come avviene nel caso di un procedimento in cui non vengono utilizzate le informazioni iniziali: è infatti sufficiente guardare la parte finale del grafico per trovare conferma a questa affermazione. Ne consegue una maggiore stabilità nelle stime ed errori standard più bassi proprio grazie al fatto che vengono considerate tutte le informazioni possibili che si possono dedurre dai dati a disposizione.

#### **4.4.4 Stime dinamiche dei parametri colonna**

Uno dei grossi limiti del metodo chain ladder è quello di considerare gli stessi fattori di sviluppo per ogni anno di generazione; in altre parole, tale metodo assume che la curva dei run off sia la stessa per ogni anno di generazione. È evidente come tale aspetto si riscontri difficilmente nella pratica assicurativa: cambiamenti interni ed esterni all'impresa apportano modifiche a volte sostanziali nel processo di liquidazione dei sinistri. Al contrario è invece possibile stabilire un legame tra l'evoluzione del processo di liquidazione dei sinistri relativo a due anni di generazione consecutivi ed è pertanto opportuno formulare un modello di tipo state space in modo tale da considerare tale aspetto senza dover imporre lo stesso andamento nel tempo della politica di liquidazione dei risarcimenti.

Ricordiamo che il modello base lineare di tipo chain ladder assume la forma

$$E(Z_{ij}) = \mu + \alpha_i + \beta_j,$$

dove

$$Z_{ij} = \log Y_{ij}.$$

Lasciando che i fattori di sviluppo siano diversi nei vari anni di generazione si ottiene

$$E(Z_{ij}) = \mu + \alpha_i + \beta_{ij}.$$

Imponendo invece che gli stessi fattori di sviluppo siano legati l'un l'altro mediante una relazione che lega ogni fattore di sviluppo relativo ad un anno di generazione a quello successivo si ha

$$\beta_{i+1,j} = \beta_{ij} + \text{termine di disturbo stocastico.}$$

Per la varianza del termine di disturbo è possibile fare le stesse osservazioni effettuate per la stima dei parametri di riga così come abbiamo visto nel paragrafo precedente. Possiamo invece osservare che in questo modo è possibile far variare l'evoluzione dei differimenti nei pagamenti relativi ai diversi anni di generazione in base alla scelta della varianza del termine di disturbo stocastico. Se ad esempio assumiamo un valore pari a zero si ricade nel caso del metodo della catena secondo cui i fattori di sviluppo sono gli stessi per ogni anno di generazione. Possiamo inoltre far dipendere la varianza dall'epoca di valutazione  $k$  ed assumere un valore particolarmente elevato nel momento in cui si prevedano cambiamenti significativi sull'impatto delle politiche liquidative, come ad esempio periodi di crisi economica con la conseguente inflazione.

#### **4.4.5 Alcune considerazioni**

Nei precedenti paragrafi abbiamo visto diverse tecniche di stima per la riserva sinistri le quali introducono una componente stocastica nelle valutazioni a partire da una struttura deterministica di tipo chain ladder. La prima osservazione che se ne può trarre è la maggiore stabilità nella stima dei parametri. Ciò si spiega con le maggiori informazioni che vengono sfruttate nel caso in cui si applicano approcci di tipo baynesiano anziché procedimenti in cui non viene considerata alcuna informazione a priori estrapolata dai dati a disposizione.

È evidente infatti come l'impiego di tutti i dati disponibili e di tutte le informazioni aggiuntive che si possono dedurre da essi possa apportare un miglioramento nelle valutazioni, cosa che si esprime in errori standard più bassi ed una conseguente maggiore stabilità nei valori dei parametri stimati con il modello. Abbiamo anche visto procedure dinamiche di stima, in particolare modelli di tipo state space in cui le valutazioni vengono effettuate mediante un filtro di Kalman. I primi studi in ambito attuariale relativi alla stima delle riserve sinistri mediante l'utilizzo di un approccio ricorsivo realizzato con un filtro di Kalman risalgono al 1983 e sono opera di De Jong e Zehnwirth. I procedimenti ricorsivi permettono, come dice il nome, di legare i parametri in modo ricorsivo tra loro. Inoltre, essi utilizzano metodi di stima dinamici per effettuare le valutazioni. È evidente come questo permetta di collegare ogni parametro con i valori di quelli precedentemente stimati: ad ogni passo le valutazioni dipendono fortemente da tutte quelle effettuate fino a quel momento. Affinché il modello disponga sempre di informazioni aggiuntive occorre introdurre nella struttura statica di base un vettore stocastico di input che, ad ogni ulteriore passo, individui la distribuzione a priori dei nuovi parametri che devono essere stimati dal modello, distribuzione di probabilità che viene di volta in volta aggiornata in base alle nuove conoscenze che si dispongono, siano esse di carattere interno od esterno all'impresa assicuratrice. Abbiamo visto come il vantaggio di utilizzare un approccio di tipo baynesiano di stima o un modello di tipo state space consiste nella possibilità di considerare tutti i dati e le informazioni ricavabili dai dati a disposizione e presenti nella matrice di run off. Mentre infatti il più comune metodo deterministico di stima del chain ladder stima i fattori di differimento sulla base dei soli dati relativi a due anni (di differimento) consecutivo, i procedimenti sopra descritti si basano sull'intero set di dati a disposizione. Tuttavia esistono delle differenze anche tra un approccio di tipo baynesiano ed un modello dinamico di stima di tipo state space. In particolare, mentre il primo stima la varianza del processo direttamente dai dati, quindi mediante una tecnica empirica di valutazione, i modelli di tipo state space necessitano di una stima a priori della varianza del processo. La relazione che lega ogni parametro al precedente coinvolge infatti un termine di disturbo, precedentemente espresso con  $w_t$ , e di cui si è parlato a lungo nei precedenti paragrafi, che individua la varianza del

processo e che deve essere specificato prima di iniziare ad effettuare le valutazioni. Se da un lato i modelli state space hanno il vantaggio di legare ricorsivamente tutti i parametri stimati; dall'altro un approccio di tipo baynesiano non necessita di alcun input aggiuntivo a priori: può stimare interamente tutti i valori di cui necessita dai dati seguendo un approccio cosiddetto di tipo empirico, senza dover specificare pertanto nessuna quantità in input. Prima di concludere è opportuno osservare anche l'importanza del metodo deterministico di stima di tipo chain ladder, punto di partenza fondamentale per lo studio di nuove e più sofisticate tecniche stocastiche di valutazione come ad esempio quelle precedentemente proposte. In ultima analisi ci soffermiamo sull'utilizzo dei risultati ottenuti con i vari metodi di stima. Se da un lato un metodo che non sfrutta informazioni a priori aggiuntive porta necessariamente a valutazioni meno stabili e conseguenti errori standard più elevati come più volte sottolineato; dall'altro occorre considerare quali sono gli effettivi valori che interessano ad una compagnia assicuratrice. Molto spesso, infatti, non è necessario ricavare un opportuno livello di confidenza per ogni riga della matrice dati, bensì è sufficiente ricavare un adeguato livello di sicurezza nel valore complessivo che si ha per la stima della riserva relativo all'intero set di dati, considerando pertanto tutte le righe della matrice contemporaneamente. Questo è importante perché, in generale, l'errore standard che si ha per le stime relative ad ogni singola riga è significativamente più elevato rispetto all'errore standard riferito al totale. L'errore standard sul totale dei dati è spesso molto più contenuto ed in conseguenza anche le stime sono più stabili. La pratica assicurativa ha inoltre spesso necessità di conoscere i valori della riserva complessiva, piuttosto che i valori specifici relativi ad ogni anno di generazione: anche le dimensioni degli errori di stima possono pertanto essere più limitati.

## **4.5 Altri metodi di stima**

### **4.5.1 Introduzione**

Vediamo adesso alcuni metodi diversi rispetto a quelli presentati fino a questo momento per stimare l'errore standard relativo alla stima della riserva sinistri effettuata col metodo del chain ladder.

Com'è ben noto l'errore standard risulta particolarmente utile al fine di verificare l'incertezza presente nei dati e, quindi, l'attendibilità delle stime. Tale errore è pertanto un indice di bontà del modello utilizzato per effettuare le valutazioni e può risultare utile per confrontare più modelli di stima. In particolare risulta interessante verificare se vi sono differenze significative tra i valori stimati con due o più approcci differenti. Mack è stato uno dei primi autori a proporre, nel suo lavoro del 1993, modelli di tipo stocastico che riproducessero le stesse stime del metodo della catena, specificando semplicemente i primi due momenti della distribuzione dei risarcimenti cumulati e considerando pertanto il modello privo di distribuzione. In realtà un primo decisivo passo verso un traguardo simile era già stato effettuato dallo studioso Schnieper nel 1991. Egli utilizzò una mistura della tecnica di Bornhuetter-Ferguson e del metodo chain ladder per ottenere una approssimazione dell'errore standard relativo alla stima del tasso di premio per un contratto di assicurazione che prevedeva una copertura riassicurativa di tipo excess of loss. Mack sfrutta l'idea di Schnieper e la applica alla stima della riserva sinistri ottenendo ulteriori risultati.

In particolare:

1. intanto il modello è specificato solo per il caso chain ladder (e non come mistura di più modelli, in linea con Schnieper);
2. oltre alla stima dell'errore standard relativo alla riserva sinistri, viene fornita anche una valutazione della varianza del processo, che in questo caso risulta decisamente opportuna vista la natura stocastica della riserva sinistri, nonché oggetto di indagine;
3. Mack riesce a dimostrare che i fattori di sviluppo non sono correlati tra loro e ciò comporta una stima non distorta della riserva sinistri;
4. infine, oltre all'errore standard relativo ad ogni anno di generazione, Mack fornisce anche una stima dell'errore standard complessivo per la riserva globale, somma delle riserve relative ai vari anni di generazione. Tale stima ha l'importante proprietà di tenere in considerazione le correlazioni tra le stime individuali relative alle singole riserve riguardanti i vari anni di generazione.

#### 4.5.2 Il modello di Mack

Ricordiamo che  $X_{ij}$  rappresenta il risarcimento cumulato relativo all'anno di generazione  $i$ ,  $i = 1, 2, \dots, k$ , pagato con un differimento di  $j$  anni,  $j = 1, 2, \dots, k$ . Si ipotizza che  $X_{ij}$  sia una variabile aleatoria di cui disponiamo di alcuni suoi valori nel caso  $i + j \leq k + 1$  (dati del triangolo run off). L'obiettivo resta la stima del risarcimento cumulato relativo all'ultimo anno della matrice dati, ( $X_{ik}$  con  $i = 1, 2, \dots, k$ ), da cui si ha in conseguenza la stima della riserva sinistri

$$R_i = X_{ik} - X_{i, k+1-i},$$

per ogni anno di generazione  $i = 2, \dots, k$ .

L'ipotesi che sta alla base del metodo della catena si può riassumere in

$$E(X_{i, j+1} / X_{i1}, X_{i2}, \dots, X_{ij}) = m_j \cdot X_{ij}, \quad i = 1, 2, \dots, k \quad \text{e} \quad j = 1, 2, \dots, k-1;$$

dove  $m_1, m_2, \dots, m_{k-1} > 0$  rappresentano i fattori di sviluppo, così stimati:

$$\hat{m}_j = \frac{\sum_{i=1}^{k-j} X_{i, j+1}}{\sum_{i=1}^{k-j} X_{ij}}, \quad j = 1, 2, \dots, k-1.$$

Si ha poi l'ultimo risarcimento cumulato

$$\hat{X}_{ik} = X_{i, k+1-i} \cdot \hat{m}_{k+1-i} \cdot \dots \cdot \hat{m}_{k-1},$$

con i conseguenti valori della riserva sinistri

$$\hat{R}_i = X_{i, k+1-i} \cdot (\hat{m}_{k+1-i} \cdot \dots \cdot \hat{m}_{k-1} - 1) = \hat{X}_{ik} - X_{i, k+1-i}$$

per ogni anno di generazione  $i = 2, \dots, k$ .

Osserviamo come la legge fondamentale che regola il metodo della catena è quella di costanza nel tempo della progressione con cui vengono liquidati i sinistri, legge nota come "progressione di smontamento dei sinistri". Questo significa escludere ogni possibile dipendenza tra gli anni di generazione, vale a dire ipotizzare che le variabili aleatorie  $X_{ij}$  sono indipendenti nei diversi anni di generazione  $i$ . Quindi:

$$\{X_{i1}, X_{i2}, \dots, X_{ik}\} \text{ e } \{X_{j1}, X_{j2}, \dots, X_{jk}\} \text{ con } i \neq j \text{ sono indipendenti.}$$



Ovviamente questa è un'ipotesi un po' forte in quanto, col tempo, possono verificarsi fattori esterni od interni all'impresa che modificano le attuali condizioni di pagamento dei sinistri. Ad esempio, possono mutare le modalità con cui vengono trattati i sinistri oppure può venire aggiornato il processo di liquidazione senza pensare alla sempre presente inflazione che pregiudica in modo considerevole le somme di denaro in gioco.

Ciò premesso, è possibile esprimere l'errore standard della stima  $\hat{X}_{ik}$  di  $X_{ik}$  come:

$$mse(\hat{X}_{ik}) = E\left(\left(\hat{X}_{ik} - X_{ik}\right)^2 / D\right),$$

dove  $D = \{X_{ij} / i + j \leq k + 1\}$ , ossia sono i valori della matrice dati osservati.

Osserviamo che, al fine di considerare le deviazioni medie tra  $\hat{X}_{ik}$  e  $X_{ik}$  dovute alla sola aleatorietà futura, si è considerata la quantità  $E\left(\left(\hat{X}_{ik} - X_{ik}\right)^2 / D\right)$  anziché

$$E\left(\left(\hat{X}_{ik} - X_{ik}\right)^2\right) = E\left(E\left(\left(\hat{X}_{ik} - X_{ik}\right)^2 / D\right)\right).$$

Occorre osservare inoltre che

$$mse(\hat{R}_i) = E\left(\left(\hat{R}_i - R_i\right)^2 / D\right) = E\left(\left(\hat{X}_{ik} - X_{ik}\right)^2 / D\right) = mse(\hat{X}_{ik}).$$

Infine dalla relazione

$$E(X - a)^2 = Var(x) + (E(X) - a)^2$$

segue

$$mse(\hat{X}_{ik}) = Var(X_{ik} / D) + \left(E(X_{ik} / D) - \hat{X}_{ik}\right)^2,$$

la quale mostra come l'errore quadratico medio sia scomponibile nella somma di due termini: il primo termine rappresenta la varianza del processo, mentre il secondo termine è l'effettivo errore di stima.

È evidente pertanto che per ottenere una stima dell'errore quadratico medio di  $\hat{X}_{ik}$  occorre specificare una formula per la varianza di  $X_{ik}$ .

Poiché i fattori di sviluppo  $\hat{m}_j$  sono delle medie ponderate dei fattori  $\frac{X_{i,j+1}}{X_{ij}}$ , ciò suggerisce che la  $Var(X_{i,j+1} / X_{ij} | X_{i1}, \dots, X_{ij})$  deve essere inversamente proporzionale a  $X_{ij}$ , ossia:

$$Var(X_{i,j+1} | X_{i1}, \dots, X_{ij}) = X_{ij} \sigma_j^2 \quad \text{con } i = 1, 2, \dots, k \quad \text{e } j = 1, 2, \dots, k-1$$

e dove  $\sigma_j^2$ ,  $j = 1, 2, \dots, k-1$ , è un parametro incognito da stimare dai dati.

Una possibile stima corretta per tale parametro è:

$$\hat{\sigma}_j^2 = \frac{1}{k-j-1} \sum_{i=1}^{k-j} X_{ij} \left( \frac{X_{i,j+1}}{X_{ij}} - \hat{m}_j \right)^2 \quad \text{con } j = 1, 2, \dots, k-2.$$

Questa espressione non è altro che una media ponderata dei residui; il denominatore rappresenta il numero dei residui meno uno, il che fornisce una stima non distorta di  $\sigma_j^2$ . Osserviamo invece che non è presente nessuna correzione che tenga conto della distorsione dovuta al numero dei parametri e che, inoltre, per completare la stima di  $\sigma_j^2$  per ogni  $j = 1, 2, \dots, k-1$  occorre assegnare un valore a  $\sigma_{k-1}^2$ . Nel caso in cui  $\hat{m}_{k-1} = 1$ , ossia si ipotizza che tutti i risarcimenti vengono effettuati al più in  $k$  anni, si può porre  $\sigma_{k-1}^2 = 0$ ; altrimenti un possibile metodo è quello di attribuire a  $\sigma_{k-1}^2$  la quantità:

$$\hat{\sigma}_{k-1}^2 = \min(\hat{\sigma}_{k-2}^4 / \hat{\sigma}_{k-3}^2, \min(\hat{\sigma}_{k-3}^2, \hat{\sigma}_{k-2}^2)).$$

Sotto tutte queste ipotesi è possibile dimostrare che l'errore quadratico medio di  $\hat{R}_i$ ,  $i = 2, \dots, k$ , è dato da:

$$mse(\hat{R}_i) = \hat{X}_{ik}^2 \sum_{j=k+1-i}^{k-1} \frac{\hat{\sigma}_j^2}{\hat{m}_j^2} \left( \frac{1}{\hat{X}_{ij}} + \frac{1}{\sum_{z=1}^{k-j} X_{zj}} \right),$$

dove  $\hat{X}_{ij} = X_{i,k+1-i} \cdot \hat{m}_{k+1-i} \cdot \dots \cdot \hat{m}_{k-1}$  con  $j > k+1-i$  sono le stime future.

Spesso risulta di particolare interesse la stima dell'errore quadratico medio complessivo, quello relativo alla riserva globale, nonché somma di tutte le riserve nei vari anni di generazione,  $\hat{R} = \hat{R}_2 + \dots + \hat{R}_k$ .

Tale quantità non può essere data dalla semplice somma dei singoli errori quadratici medi poiché è opportuno considerare anche la correlazione presente tra i fattori di sviluppo  $\hat{m}_j^2$  e tra le quantità  $\hat{\sigma}_j^2$ .

Un'espressione in grado di effettuare una correzione che tenga conto delle covarianze è data da:

$$mse(\hat{R}) = \sum_{i=2}^k \left\{ (s.e(\hat{R}_i))^2 + \hat{X}_{ik} \left( \sum_{z=i+1}^k \hat{X}_{zk} \right) \sum_{j=k+1-i}^{k-1} \frac{2\hat{\sigma}_j^2 / \hat{m}_j^2}{\sum_{n=1}^{k-j} X_{nj}} \right\}$$

dove  $\hat{R} = \sum_{i=2}^k \hat{R}_i$ .

In ultima analisi è opportuno concludere con un'osservazione sulle caratteristiche che l'errore standard appena introdotto riesce a cogliere. Tale errore infatti riflette inesattezze della stima dovuta all'aleatorietà delle variabili coinvolte nelle valutazioni; tuttavia non è in grado di cogliere il cosiddetto errore di specificazione, ossia l'errore che si commette nello specificare il modello utilizzato per effettuare l'analisi. Secondo tale considerazione non si può escludere che il modello adottato sia sbagliato. Infine, l'errore standard si basa sull'ipotesi che il futuro ricalchi il passato: in altre parole prevede che non vi siano discordanze da ciò che si è verificato nel passato in termini di processo di liquidazione dei sinistri (i fattori di sviluppo futuri sono in accordo con quelli registrati nel passato dei dati a disposizione per le stime).

#### 4.5.3 Il modello di Wright

Il modello di Wright si differenzia da quelli presentati fino a questo momento per il fatto di basarsi su alcuni concetti fondamentali relativi alla teoria del rischio in ambito attuariale. È noto infatti come, sotto tale teoria, il rischio assicurativo sia scomposto in due distinte variabili aleatorie che lo rappresentano: il numero di sinistri  $N$  e l'ammontare dei risarcimenti  $Y$ . In particolare, l'obiettivo è quello di individuare la distribuzione di probabilità del risarcimento complessivo, anche detta base tecnica del rischio.

A sua volta il risarcimento complessivo o danno aggregato è dato dalla quantità

$$X = \sum_{i=0}^N Y_i$$

dove  $N$  è il numero complessivo di sinistri che colpiscono il contratto ed  $Y_i$  è il risarcimento relativo all' $i$ -esimo sinistro in ordine cronologico con l'usuale convenzione  $Y_0 = 0$ .

Nella maggior parte dei casi sono poi adottate le ipotesi standard della teoria classica del rischio:

- indipendenza tra ammontare dei risarcimenti e numero di sinistri;
- indipendenza ed identica distribuzione dei risarcimenti per ogni prefissato numero di sinistri verificatisi.

Si tratta dunque di specificare un'appropriata distribuzione di probabilità per la variabile aleatoria numero dei sinistri e la variabile aleatoria ammontare dei risarcimenti, da cui in conseguenza se ne ricava la distribuzione relativa al danno aggregato. L'approccio più classico è quello di modellare la variabile aleatoria numero dei sinistri con una distribuzione di tipo Poisson e la variabile aleatoria ammontare dei risarcimenti con una distribuzione di tipo Gamma. Il risultato relativo al risarcimento complessivo è una distribuzione composta del tipo

$$F_X(x) = \sum_{k=0}^{+\infty} P(N = k) F_Y^{k*}(x) = \frac{e^{-\lambda} \lambda^k}{k!} F_Y^{k*}(x)$$

dove  $\lambda$  è il parametro della distribuzione di Poisson relativo alla variabile aleatoria numero di sinistri, mentre

$$F_Y^{k*}(x) = \Pr\left(\sum_{i=0}^k Y_i \leq x\right)$$

con  $F_Y^{0*}(x) = 1$  per ogni  $x \geq 0$ .

Tale quantità è nota in statistica come la convoluzione  $k$ -esima della variabile aleatoria  $Y$ . In altre parole, essa rappresenta la probabilità che la somma di  $k$  importi aleatori non negativi indipendenti ed identicamente distribuiti sia inferiore o uguale a  $x$  (pertanto funzione di ripartizione della somma di  $k$  importi aleatori non negativi indipendenti ed identicamente distribuiti).

Il lavoro presentato da Wright nel 1990 segue un approccio di tale tipo in quanto la componente sistematica e quella random del modello da lui stesso suggerito sono basate sui concetti della teoria del rischio per quanto concerne il processo di generazione dei risarcimenti. In particolare, l'Autore ha proposto di considerare il singolo risarcimento incrementale  $y_{ij}$  come somma di  $N_{ij}$  pagamenti indipendenti di ammontare  $Y_{ij}$  ed il numero di sinistri  $N_{ij}$  è stato modellato come una variabile casuale di Poisson. Nello specifico,  $N_{ij}$  è rappresentata da una distribuzione di tipo Poisson con valore atteso e varianza pari a:

$$E[N_{ij}] = c_i a_j k_i j^{A_i} e^{-b_i j},$$

$$Var[N_{ij}] = E[N_{ij}],$$

dove  $k$ ,  $A$  e  $b$  sono i parametri (costanti incognite) da stimare,  $c$  è una misura dell'esposizione al rischio che dipende dall'anno di generazione  $i$ , ed  $a$  è un termine noto di aggiustamento, che, nel suo lavoro, Wright specifica per ogni anno di differimento  $j$ <sup>17</sup>.

Gli importi dei singoli risarcimenti seguono invece una distribuzione con valore atteso e varianza dati rispettivamente da:

$$E[Y_{ij}] = e^{\delta} k \cdot j^{\omega}$$

$$Var[Y_{ij}] = v(E[Y_{ij}])^2$$

con  $k$  e  $\omega$  costanti incognite.

Esaminando l'espressione relativa al valore atteso si può notare la presenza del termine  $e^{\delta}$  incluso per considerare la presenza dell'inflazione. In esso infatti  $t = i + j$  rappresenta l'anno di calendario, mentre  $\delta$  è una costante che esprime la forza di inflazione dei pagamenti. Originariamente Wright suggerì una distribuzione di tipo Gamma per rappresentare i singoli risarcimenti. In realtà, osservando le espressioni dei momenti, egli non ipotizzò una distribuzione di tipo Gamma, ma solamente una distribuzione di probabilità in cui la varianza esiste finita ed è proporzionale al

---

<sup>17</sup> L'Autore inserisce degli aggiustamenti agli anni di sviluppo  $j$  che vengono qui ignorati per semplicità in quanto esulano da uno scopo puramente descrittivo che si intende fornire.

quadrato della media. Il modello fin qui proposto risulta molto interessante in quanto ricalca molto da vicino quella che è la teoria di pricing delle polizze assicurative dei rami danni, nella quale si tiene conto sia della distribuzione del numero di sinistri che di quella del costo del singolo risarcimento, per passare poi ad un modello di tipo mistura che incorpori le due componenti. Nel caso trattato in questo paragrafo, passando quindi ad una mistura dei modelli e rifacendosi alle ipotesi della teoria classica del rischio precedentemente esposte si ottiene:

$$E[y_{ij}] = m_{ij} = c_i a_j k_i j^{A_i} e^{-b_i j} e^{\delta} k \cdot j^\omega,$$

$$Var[y_{ij}] = c_i a_j k_i j^{A_i} e^{-b_i j} E[Y_{ij}^2].$$

Si può dimostrare che ciò rappresenta un modello di tipo lineare generalizzato di cui Renshaw ne ha ricavato una semplice derivazione. In particolare, egli dimostrò che il modello per i dati incrementali può essere riscritto nel seguente modo:

$$E[y_{ij}] = e^{u_{ij} + v_i + \delta_i \log j + \gamma_i (j-1) + \delta},$$

$$Var[y_{ij}] = \phi_{ij} E[y_{ij}] = \phi_{ij} m_{ij}.$$

Con questa formulazione si ottiene un modello di tipo lineare generalizzato nel quale la variabile risposta  $y_{ij}$  è modellata attraverso una funzione link di tipo logaritmica e la varianza è proporzionale alla media. I  $\phi_{ij}$  sono parametri di scala incogniti da stimare tramite il modello ad esempio mediante il metodo del “joint modelling”. Tale metodo consiste nell’assegnare inizialmente valori positivi ai  $\phi_{ij}$  e fittare poi un modello lineare generalizzato ponderato con pesi inversamente proporzionali alla varianza. Tali pesi sono poi utilizzati come variabile risposta in un secondo modello ed i loro valori sono utilizzati per aggiornare quelli dei  $\phi_{ij}$  prima di riadattare nuovamente il primo modello al fine di rivedere le stime dei  $\phi_{ij}$  stessi.

In ultima analisi si può notare che l’esponente di  $E[y_{ij}]$ , a meno del termine  $u_{ij}$  e di quello relativo all’inflazione  $\delta$ , risulta essere l’espressione della curva di Hoerl precedentemente presentata. I termini  $u_{ij}$  sono noti e rappresentano degli aggiustamenti del modello. La differenza sostanziale che si può riscontrare tra la curva di Hoerl ed il modello di Wright appena presentato risiede nel fatto che mentre

il primo utilizza una variabile risposta su scala logaritmica, il secondo utilizza i pagamenti incrementali stessi e non una loro trasformazione. Inoltre non bisogna trascurare il fatto che Wright introduce nel suo modello una componente aggiuntiva,  $e^{\alpha}$ , in grado di considerare anche la presenza dell'inflazione.

## 4.6 Modelli di smoothing

### 4.6.1 Premesse

I modelli di smoothing nascono dall'esigenza di trovare una soluzione alla troppa rigidità delle curve di tipo parametrico. I primi autori ad utilizzare questi modelli di tipo non parametrico sono stati England e Verral (2001). Uno smoother consiste in una tecnica in grado di rappresentare i dati in maniera meno variabile, cioè più facilmente leggibili ed interpretabili. Le medie mobili sono un semplice esempio di smoother, mentre una retta di regressione non può essere considerata uno smoothing per via della forma funzionale ad essa associata.

Gli smoothers hanno due utilizzi principali:

- *descrittivo* nel caso di uno smoother relativo ad una sola variabile predittiva rappresentato da un grafico, detto *scatterplot smoother*, che può essere utilizzato al fine di facilitare la visione del trend che lega la variabile dipendente  $y$  alla variabile esplicativa  $x$ ;
- *di studio delle dipendenze* per analizzare la dipendenza della media della variabile risposta  $y$  dalle variabili predittive  $x$ , con il fine ultimo di specificare il modello più opportuno per la variabile oggetto di indagine  $y$ .

Molti smoothers utilizzano un procedimento cosiddetto di “media locale”, ossia effettuano una media dei valori delle variabili dipendenti attorno ad alcuni specifici valori delle variabili esplicative. Nell'ipotesi, ad esempio, in cui si abbia una sola variabile esplicativa,  $x$ , ed una sola variabile dipendente,  $y$ , lo smoother andrà a calcolare una media locale dei valori relativi alla variabile  $y$  in un intorno  $I_{x_0}$  di  $x_0$ . In conseguenza l'utilizzo dello smoother è in genere legato, oltre alla scelta del tipo di smoother da utilizzare, alla scelta del valore  $x_0$  ed alla dimensioni dell'intorno  $I_{x_0}$ , le cui entità sono generalmente espresse in termini di parametro di smoothing.

Scendendo nei particolari, uno smoother è definito come una funzione di  $x$  e di  $y$  (nell'ipotesi di una sola variabile esplicativa e di una sola variabile dipendente), il cui risultato è una funzione  $s(\cdot)$  il cui dominio coincide con quello della variabile  $x$ , ossia si considera la relazione:

$$s(\cdot) = S(y/x).$$

Un classico esempio di smoothing è la media mobile, modello di smoothing principalmente utilizzato per esaminare il trend di serie temporali. Nell'ipotesi che il punto di riferimento sia il valore  $x_0$  e che in corrispondenza di esso si abbiano valori di  $y$  replicati, ossia per un determinato valore della variabile esplicativa si ha che la variabile dipendente assume più valori, allora si può semplicemente effettuare la media dei valori della  $y$  in corrispondenza del valore  $x_0$  come stima di  $s(x_0)$ . Se invece il fenomeno non presenta valori replicanti per la variabile  $y$  in corrispondenza di  $x_0$ , si può considerare un opportuno intorno di  $x_0$ ,  $I_{x_0}$ , centrato su tale punto e che abbia  $q/2$  elementi alla destra e  $q/2$  elementi alla sinistra di  $x_0$  stesso. In questo caso la stima di  $s(x_0)$  può essere effettuata mediante l'utilizzo di una media, semplice o ponderata che sia, che tenga conto della distanza di ogni punto considerato dal valore centrale  $x_0$  per tale motivo anche detto *symmetric nearest neighbourhood*.

Questi semplici modelli di smoothing rappresentano esempi di medie mobili. Tali procedimenti hanno la caratteristica di essere facilmente applicabili ad ogni fenomeno per la loro semplicità, tuttavia essi risultano essere poco stabili in corrispondenza dei valori estremi di una serie. In particolare in tali punti potrebbero risultare significativamente distorti: pertanto una semplice generalizzazione delle medie mobili è il cosiddetto *running line smoother*. Tale metodo contribuisce ad eliminare il problema della distorsione e viene calcolato considerando una stima dei minimi quadrati nell'intorno  $I_{x_0}$  secondo la relazione

$$s(x_0) = \bar{\beta}(x_0),$$

che ne individua la retta di regressione nel punto  $x_0$ . L'approccio deve essere tale da apportare valide ed utili indicazioni dall'inclinazione della retta di regressione in corrispondenza dei punti estremi della serie.



#### 4.6.2 Dai modelli di smoothers ai modelli additivi

Fino ad ora sono stati presentati smoothers che effettuano un adattamento locale. Tale procedura risulta estremamente più flessibile rispetto a quella che può essere eseguita utilizzando delle funzioni di tipo parametrico, lineari o polinomiali che siano. Tuttavia a volte può rivelarsi utile un adattamento locale di funzioni polinomiali, ed in questo caso si possono applicare trasformazioni cosiddette di tipo spline.

La regressione polinomiale possiede proprietà di interpolazione su tutto il dominio della variabile dipendente, ma se considerata in sottoinsiemi di tale dominio può essere considerata alla pari di uno smoother, permettendo così di combinare la flessibilità dei polinomi con l'adattamento locale. Si consideri pertanto di suddividere il dominio della variabile indipendente in sottointervalli delimitati da nodi. Per adattare le splines si costruisce una funzione polinomiale di grado  $p$ , che assume valori nei tratti delimitati dai nodi, derivabile  $p$  volte negli intervalli tra due nodi e  $p-1$  volte nei nodi stessi. Sebbene sia possibile considerare vari tipi di funzioni polinomiali, la più utilizzata è la funzione cubica, sotto ipotesi di continuità nei punti che costituiscono i nodi, ed avente in tali punti derivata prima e derivata seconda continua. La principale difficoltà che si può incontrare quando si considerano tali tipi di trasformazioni è legata alla scelta del numero e delle posizioni degli intervalli. Un approccio molto semplice è quello che considera solo il numero di essi, mentre la loro posizione viene scelta secondo una funzione uniforme sul campo di variazione della variabile indipendente. Un altro approccio consiste nel porre gli intervalli in corrispondenza dei punti che individuano i quartili sempre in riferimento alla variabile indipendente. Le splines vengono adattate effettuando una regressione multipla in corrispondenza dei valori osservati delle  $x$ . La singola funzione spline viene poi costruita come combinazione lineare delle splines calcolate nei sottointervalli in cui è stato suddiviso il dominio della variabile  $x$  e la funzione risultante è, per costruzione, vincolata ad essere continua. Questo tipo di smoother è molto utile perché è statisticamente semplice e flessibile; inoltre non presenta elevati problemi computazionali in sede di adattamento. Nel caso più generale in cui si abbiano  $n$  variabili esplicative  $X_1, \dots, X_n$ , si pone il problema di interpolare una superficie  $n$ -dimensionale per ottenere la stima della variabile dipendente  $Y$ .

In tale ambito è possibile generalizzare i concetti di medie mobili e di rette mobili, considerando un intorno di punti centrato in  $x_1^0, \dots, x_n^0$ . In questo caso, però, vi è un'ulteriore difficoltà legata alla metrica che definisce la forma dell'intorno stesso nello spazio avente  $n$  dimensioni e un altro problema importante nel caso multidimensionale consiste nel fatto che gli intorni, all'aumentare di  $n$ , perdono rapidamente il carattere locale, diventando sempre più ampi. Inoltre, sempre all'aumentare di  $n$ , le osservazioni diventano sempre più rare e distanziate tra di loro e quindi, nel caso si volessero considerare intorni contenenti un numero ampio di unità, questi dovrebbero avere dimensioni sempre più grandi. Per questo motivo e per i problemi computazionali legati alle stime, spesso non trascurabili, vengono di solito utilizzati smoothers che fanno riferimento a modelli con un'unica variabile esplicativa, oppure vengono considerati i modelli cosiddetti additivi ed i modelli additivi generalizzati, i quali costituiscono una particolare classe di modelli di tipo regressivo. Sotto tale punto di vista gli smoothers unidimensionali vengono semplicemente utilizzati come punto di partenza al fine di applicare algoritmi di stima non parametrica più efficaci.

#### **4.6.3 I modelli additivi ed i modelli additivi generalizzati**

Pur godendo di proprietà vantaggiose i modelli di tipo regressivo possono talvolta non essere appropriati al fenomeno oggetto di indagine. In molte situazioni, infatti, tali modelli vincolano l'analisi ad ipotesi molto restrittive sulle variabili ed inoltre possono causare problemi di rappresentazione grafica nel caso in cui le variabili esplicative siano più di una. Si può pertanto ricorrere a modelli più flessibili in grado di cogliere gli aspetti fondamentali dei fenomeni analizzati e di essere maggiormente adattabili ad ogni tipo di esigenza. Il modo più efficace per fare questo consiste nel considerare i modelli additivi, i quali possono essere definiti dalla seguente relazione:

$$Y = \sum_{i=1}^n f_i(X_i, \theta_i) + \varepsilon .$$

Sostanzialmente all'unica funzione che componeva il modello lineare si è sostituita una somma di funzioni, introducendo così l'ipotesi di additività degli effetti delle variabili esplicative sulla variabile risposta  $Y$ .

In altre parole la variabile risposta  $Y$  è completamente scomponibile nella somma di  $n$  funzioni univariate più il termine di errore  $\varepsilon$ . Apparentemente una tale ipotesi può apparire restrittiva in quanto non tiene conto dei termini di interazioni tra le variabili, tuttavia la maggior parte degli algoritmi di stima prevede la possibilità di inserire nel modello termini di interazione. Si può inoltre osservare che il modello additivo conserva delle importanti caratteristiche riguardo l'interpretazione dei coefficienti, similmente al modello lineare: ognuna delle funzioni  $f_i$  adattate non dipende dal valore assunto dalle altre funzioni  $f_j$ , per ogni  $i$  diverso da  $j$ . Per questo motivo l'adattamento di ognuna delle funzioni  $f_i$  indica in quale modo la variabile  $X_i$  agisce sulla variabile risposta  $Y$ . Il metodo più generale per stimare le funzioni  $f_i$  di un modello additivo consiste nel porre ognuna di esse uguale ad uno smoother arbitrario, come, ad esempio, quello che può essere ottenuto mediante delle funzioni di tipo spline. In questo modo, però, il numero di equazioni che andrebbe risolto sarebbe decisamente elevato (nell'ordine di  $(np)^2$  per  $n$  variabili e  $p$  osservazioni), per cui si può fare ricorso ad un particolare algoritmo di adattamento molto flessibile, il cosiddetto backfitting, il quale consiste in una procedura iterativa che assume le seguenti condizioni:

$$\begin{aligned} E(\varepsilon) &= 0, \\ \text{var}(\varepsilon) &= \sigma^2, \\ \text{cov}(\varepsilon, X) &= 0, \end{aligned}$$

dalle quali si può derivare

$$E\left(Y - \sum_{i=1}^n f_i(X_i)\right) = f_i(X_i).$$

Quest'ultima relazione suggerisce l'algoritmo iterativo che può essere utilizzato per il calcolo delle singole  $f_i$ . Il metodo del backfitting parte infatti da funzioni iniziali stabilite a priori e procede ad ulteriori stime delle stesse funzioni sulla base del seguente metodo:

$$f_j(X_j) = S_j\left(Y - \sum_{i \neq j} f_i / X_j\right),$$

dove  $S_j(\cdot)$  individua uno smoother applicato alla variabile residuo,  $Y - \sum_{i \neq j} f_i / X_j$ .

Tale algoritmo nello stimare iterativamente le funzioni  $f_j$  rimuove gli effetti che tutte le altre funzioni diverse dalla  $j$ -esima hanno nei confronti della variabile risposta  $Y$ . L'algoritmo va iterato fino a quando le singole funzioni non subiscono più variazioni tra una iterazione e l'altra e la convergenza di esso è stata dimostrata da Hastie Tibshirani (1990). Un esempio di stima del valore iniziale da cui far partire l'algoritmo può essere il valore della stima dei minimi quadrati ottenuti mediante il modello di regressione lineare applicato ad ognuna delle variabili esplicative. Il vantaggio dell'algoritmo di backfitting è quello di specificare modelli con componenti molto articolate, in quanto si possono inserire nell'analisi funzioni lineari e non. Inoltre, in ultima osservazione, si possono considerare variabili qualitative o fattori che rappresentano l'interazione tra due o più variabili.

Infine, così come i modelli additivi costituiscono una generalizzazione dei modelli lineari, i modelli additivi generalizzati costituiscono una generalizzazione dei modelli lineari generalizzati. In particolare, i modelli additivi generalizzati differiscono dai modelli lineari generalizzati per il fatto che i primi considerano l'additività degli effetti delle variabili indipendenti nei confronti della variabile dipendente per mezzo dell'introduzione di una procedura di smoothing.

#### **4.6.4 I modelli di smoothers applicati alla stima della riserva sinistri**

In questo paragrafo verranno presentati risultati che tengono conto del fatto che l'obiettivo dell'indagine resta la stima dei risarcimenti incrementali da cui si giunge, additivamente, alla valutazione della riserva sinistri. Per questo motivo gli smoothing sono costruiti nel momento stesso in cui il modello viene interpolato. In altre parole non vengono assunte forme a priori per l'andamento del modello di run-off, bensì la procedura di smoothing permette al modello di adattarsi al trend presente nei dati. I modelli di smoothing sono implementati utilizzando i modelli additivi di tipo generalizzato (*Generalized Additive Models, GAMs*).

Nei modelli additivi generalizzati il predittore lineare presente nei più classici modelli lineari generalizzati è sostituito con uno smoother di tipo non parametrico così come esprime la seguente espressione:

$$\eta_{ij} = \sum_{v=1}^p s_v(x_{ij}).$$

Ovviamente la funzione  $s(x)$  definisce lo smoother di tipo non parametrico in  $x$  e la scelta di questa funzione può avvenire all'interno di una vasta gamma di smoothers. Nel nostro caso si sceglie uno smoother appartenente alla famiglia degli splines cubici. Ricordiamo che le splines sono sostanzialmente funzioni di tipo polinomiale a tratti. Quando si costruisce una funzione interpolante, nasce l'esigenza di controllare il comportamento di tipo oscillatorio del polinomio utilizzato. Un metodo che consente di raggiungere tale scopo è quello di mantenere il grado del polinomio fisso e diminuire l'intervallo considerato nell'interpolazione.

Formalizziamo matematicamente il problema:

sia  $[a, b]$  l'intervallo inizialmente considerato e sia  $\Delta = \{a = x_0 < x_1 < \dots < x_n = b\}$  una suddivisione di esso. Una funzione di tipo splines di grado  $p$  con nodi nei punti  $x_i$   $\forall i = 0, \dots, n$  è una funzione  $x \rightarrow s_p(x)$ ,  $x \in [a, b]$  con le seguenti proprietà:

1. su ogni intervallo  $[x_i, x_{i+1}]$ ,  $i = 0, \dots, n-1$  la funzione  $s_p(x)$  è un polinomio di grado  $p$ ;
2. la funzione  $s_p(x)$  e le sue  $p-1$  derivate sono continue su  $[a, b]$ , ossia  $s_p(x) \in C^{p-1}[a, b]$ ;
3.  $s_p(x)$  interpola  $f(x)$  sui nodi di  $\Delta$  se  $s_p(x_i) = f_i$ .

Questa funzione ha un certo numero di gradi di libertà individuati dal numero dei coefficienti del polinomio diminuito del numero dei vincoli, ossia  $(p+1)n - p(n-1) = n + p$ . Lo spazio generato dalle funzioni splines ha dimensione pari ai gradi di libertà. Data la definizione generale, le splines cubiche si ottengono per  $p = 3$ . Sono quindi funzioni dello spazio  $C^2[a, b]$  e sono polinomi di terzo grado su ogni sottointervallo della suddivisione dell'intervallo iniziale stesso  $[a, b]$ .

I gradi di libertà sono dati da  $4n - 3(n - 1) = n + 3$  e pertanto se si prendono le coppie  $(x_i, f_i)$ ,  $i = 0, \dots, n$  tali che  $a = x_0 < x_1 < \dots < x_n = b$  e si impongono le condizioni di interpolazione  $s_3(x_j) = y_j$ , con  $x_j \in \Delta$ , rimangono 2 gradi di libertà. Il polinomio interpolante deve però essere unico e per questo motivo è necessario imporre altre due condizioni. Le più comuni sono:

- a) *spline naturale*:  $s_3''(a) = s_3''(b) = 0$ ;
- b) *spline periodica*:  $s_3'(a) = s_3'(b)$ ;  $s_3''(a) = s_3''(b)$ ;  $f_n = f_0$ ;
- c) *spline vincolata*:  $s_3'(a) = f_0'$ ;  $s_3'(b) = f_n'$ .

La spline cubica naturale è quindi la soluzione del problema di minimo:

$$\min_f \int_a^b |f''(x)|^2 dx.$$

Nel caso in cui i dati siano distribuiti normalmente, lo spline cubico di tipo univariato  $\hat{y} = s(x)$  si trova minimizzando la somma dei quadrati dei residui, ossia:

$$\sum_i (y_i - s(x))^2 + \theta \int_a^b (s''(t))^2 dt;$$

dove il secondo addendo definisce una quantità di penalizzazione di smoothing dovuta alla curvatura della funzione di spline  $s(x)$ .

In particolare, il parametro  $\theta > 0$  controlla il livello di smoothing: quando questo tende a 0 la penalità dovuta allo smoothing è anch'essa nulla ed il modello si adatta perfettamente ai dati; quando, al contrario, esso tende all'infinito, i valori predetti cadono lungo una linea dritta forzando perciò la relazione di linearità in  $x$ . È pertanto evidente che il parametro  $\theta$  viene scelto in modo tale da avere un giusto compromesso tra il livello di smoothness che si desidererebbe avere e la bontà di adattamento.

Generalizziamo adesso l'espressione relativa al predittore su scala logaritmica introdotto con la curva di Hoerl, al caso in cui si utilizzano funzioni di smoothing:

$$\eta_{ij} = s_{\theta_i}(i) + s_{\theta_j}(\log(j)) + s_{\theta_j}(j).$$

La funzione  $s(i)$  rappresenta un “lisciamiento” dell’anno di origine  $i$ , ottenuto attraverso uno spline con parametro di smoothing  $\theta_i$ ; mentre, in modo analogo, le funzioni  $s(j)$  e  $s(\log(j))$  rappresentano le splines che specificano la forma del modello di run-off con parametro di smoothing  $\theta_j$ , cioè lo stesso per entrambe le funzioni. Si può osservare come l’uso del predittore appena introdotto implichi in modo implicito l’assunzione dello stesso modello di run-off per ogni anno di origine,  $i$ . In ultima analisi risulta interessante, anche per fini di tipo comparativo, effettuare alcune osservazioni sugli estremi dei parametri di smothing. Si è già sottolineato che quando  $\theta_i$  è 0 non vi è smoothing: in questo modo, il modello è vincolato a passare attraverso ogni valore di  $i$ , il che significa trattare ogni anno di origine come se fosse un fattore. Lo stesso discorso vale per  $\theta_j$ : anche in questo caso, infatti, al tendere di questo parametro a 0, gli anni di sviluppo vengono trattati come fattori. La conseguenza di ciò è che il modello collassa nel metodo della catena e restituisce gli stessi risultati del più comune metodo deterministico. Quando, infine,  $\theta_j$  tende all’infinito, la parte del modello che si riferisce agli anni di sviluppo risulta lineare sia in  $j$  che nel  $\log(j)$ : il modello collassa pertanto nella curva di Hoerl.

## CAPITOLO 5

### IL FILTRO DI KALMAN

#### 5.1 Premesse

Il filtro di Kalman si colloca nell'ambito dei modelli dinamici di stima e dei filtri adattivi. I filtri adattivi introducono l'approccio baynesiano nella stima degli stati di un sistema governato da equazioni differenziali stocastiche. In particolare le stime a priori degli stati iniziali del sistema ad un dato istante vengono aggiornate sulla base delle osservazioni e le nuove stime a posteriori rappresentano le stime a priori per il successivo step. È infatti evidente che quanto più ampio è il periodo campionario di riferimento, tanto più alta risulta la possibilità che alcuni parametri del modello non possono essere tenuti costanti poiché su di essi possono avere la loro influenza una serie di fattori di cambiamento esterni od interni al processo in esame. Tra gli strumenti analitici più frequentemente utilizzati per la stima dei modelli a struttura variabile vi è appunto il filtro di Kalman e le sue possibili estensioni. Occorre subito mettere in evidenza come l'utilizzo di questi strumenti presenta molti problemi di tipo applicativo relativi alla definizione delle condizioni iniziali di stima, della struttura dinamica del modello e dell'algoritmo di stima della matrice di transizione e delle matrici di covarianza. Tutto ciò rende tale strumentazione di difficile adozione per un attuario applicato e ne limita fortemente l'utilizzo. Da un punto di vista letterario, tale filtro è stato introdotto da Kalman (1960) nel campo dell'ingegneria, successivamente Harrison e Stevensen (1976) lo hanno esteso in ambito statistico fino ad arrivare a De Jong e Zehnwirth (1983) i quali lo hanno applicato alle tecniche attuariali. Prima di descriverne il suo funzionamento, appare opportuno soffermarsi sulla definizione delle misure di variabilità che interessano l'attuario nell'ambito della stima della riserva sinistri. In particolare l'attenzione è focalizzata sulla radice quadrata dell'errore quadratico medio di predizione, meglio noto con il termine di prediction error.



Nel seguito verrà pertanto fornita la definizione di prediction error evidenziandone la differenza che intercorre tra quest'ultimo e il più classico standard error.

### 5.1.1 Il prediction error

Sul calcolo dell'errore di previsione sono stati effettuati numerosi approfondimenti, sia a carattere teorico che computazionale. L'obiettivo principale dei metodi stocastici non è infatti solo quello di fornire una semplice previsione riguardo il possibile valore da accantonare, quanto piuttosto ottenere anche e soprattutto delle misure di variabilità della stima della riserva sinistri. A tale proposito evidenziamo fin da adesso che spesso si ricercano modelli in grado di replicare esattamente gli stessi valori prodotti con un metodo non stocastico di stima, quale ad esempio il chain ladder. Un modello stocastico è un processo predittivo in quanto le stime dei risarcimenti futuri sono ricavate sulla base dei dati osservati. Per tale motivo la variabilità totale del processo, che è rappresentata dal prediction error, deve includere sia la variabilità insita nei dati stimati (varianza del processo), in quanto anch'essi frutto di previsione, sia l'incertezza nella stima dei parametri (varianza della stima). In proposito si assuma che la variabile casuale  $X$  rappresenti la riserva totale. Mediante l'adozione di un opportuno modello stocastico per  $X$  sui dati del triangolo passato si ottiene un suo valore previsto,  $\hat{X}$ , il quale rappresenta una stima del valor medio  $E(X)$ . Siamo interessati alla variabilità totale del processo ovvero all'errore quadratico medio di previsione (RMSEP, dall'inglese Root Mean Squared Error of Prediction), meglio noto come prediction error e definito come

$$E\left[(X - \hat{X})^2\right].$$

Sotto la tipica assunzione di indipendenza tra le osservazioni del passato e quelle future ed ipotizzando la correttezza  $E(X) = E(\hat{X})$  si ottiene la seguente scomposizione:

$$RMSEP = E\left[(X - \hat{X})^2\right] = E\left[(X - E(X) + E(X) - \hat{X})^2\right] =$$

$$\begin{aligned}
&= E\left[\left(X - E(X) - (\hat{X} - E(X))\right)^2\right] = \\
&= E\left[(X - E(X))^2\right] + E\left[(\hat{X} - E(X))^2\right].
\end{aligned}$$

Il primo termine è solitamente riferito alla varianza del processo e rappresenta la variabilità presente nei dati iniziali da cui si ricavano le valutazioni future, mentre il secondo termine è la varianza della stima, la cui radice quadrata è lo standard error e, com'è intuitivo attendersi, rappresenta la variabilità dovuta alla stima. Il prediction error è infine la radice quadrata della somma di queste due quantità, che complessivamente individuano la varianza di previsione.

## 5.2 Il filtro di Kalman

Il filtro di Kalman si è originariamente sviluppato nella letteratura ingegneristica e rappresenta un sistema di equazioni che permette ad uno stimatore di essere aggiornato ogni volta che una nuova osservazione diviene disponibile. L'algoritmo di calcolo si può pensare come un processo distinto in due fasi contraddistinte da:

- un'equazione di previsione, che riguarda la formazione del previsore ottimo dalla prima osservazione, date tutte le informazioni disponibili fino al momento corrente;
- un'equazione di aggiornamento, che riguarda la raccolta della nuova osservazione e quindi della ulteriore informazione a posteriori, da incorporare nello stimatore.

Vi è pertanto una combinazione di informazioni a priori con informazioni a posteriori e ciò avviene secondo un meccanismo di correzione basato, in particolare, sul contributo di due grandezze. La prima è la cosiddetta sequenza innovativa, ossia un insieme di innovazioni verificatesi nell'arco temporale su cui si osserva la serie storica espressa come sequenza degli errori di previsione un passo avanti (one step ahead) direttamente ricavabili dalla differenza tra l'osservazione effettiva in un certo istante e la previsione della stima ottenuta all'istante precedente. La seconda grandezza è il cosiddetto guadagno del filtro di Kalman, una matrice di pesi solitamente indicata con il simbolo  $K$ , variabile nel tempo in quanto rappresenta la quantità di informazioni contenute nella sequenza innovativa circa lo stato del

periodo di cui si vuole effettuare la stima. È necessario precisare che l'utilizzo del filtro di Kalman richiede la specificazione di un modello sottostante su cui essere applicato; questo è lo state space model, o modello a spazio degli stati, descritto nel precedente capitolo. Sostanzialmente tale modello prevede due assunzioni fondamentali incentrate su due gruppi di equazioni: l'equazione di transizione (o di sistema) e l'equazione di osservazione (o di misurazione). L'equazione di transizione descrive il processo stocastico attraverso un sistema lineare dinamico caratterizzato da un errore con distribuzione di tipo normale con media zero e varianza nota. In conseguenza dal sistema deriva un processo casuale dato dalla soluzione di un'equazione di transizione basata sulla proprietà di markovianità, ovvero di indipendenza dello stato futuro del processo dal proprio passato, dato lo stato presente. L'equazione di osservazione, così com'è intuitivo pensare, descrive invece l'osservazione, intesa questa ultima come somma del segnale ricercato e rappresentato dal vettore di stato e di una componente erratica. In altre parole tale equazione individua la relazione tra lo stato ed un vettore di misurazione ed è data nella nota forma della regressione, dove il vettore dei coefficienti di regressione è lo stato, ovvero un vettore di variabili stocastiche, che viene moltiplicato per una matrice di regressione e, quindi, il loro prodotto sommato ad un vettore di errori casuali di osservazione.

Esprimendo in formule il sistema di equazioni appena descritto:

$$\begin{cases} \beta_t = H_t \beta_{t-1} + G_t w_t \\ Y_t = X_t \beta_t + v_t \end{cases}$$

dove

$\beta = k \times 1$  è il vettore di stato;

$H = k \times k$  è la matrice di transizione;

$w = g \times 1$  è il vettore dei disturbi del sistema;

$G = k \times g$  è la matrice di transizione o di regolazione dei disturbi del sistema;

$Y_t = n \times 1$  è il vettore delle osservazioni;

$X_t = n \times k$  è la matrice dei regressori;

$v = n \times 1$  è il vettore degli errori di osservazione.

Le variabili aleatorie  $w_t$  e  $v_t$  rappresentano rispettivamente l'errore di processo e l'errore di misura. Si assume che esse siano indipendenti l'una dall'altra e con distribuzione di probabilità normale del tipo

$$p(w) \approx N(0, Q) \quad \text{e} \quad p(v) \approx N(0, R).$$

Si assume inoltre che la matrice di covarianza  $Q$  dell'errore di processo, la matrice di covarianza  $R$  dell'errore di misura e le matrici di transizione  $H, G$  e  $X$  siano costanti nel tempo.

### 5.2.1 L'algoritmo di calcolo

Definiamo con  $\bar{\beta}_k \in \mathfrak{R}^n$  la stima a priori dello stato al passo  $k$ , data la conoscenza del processo precedente l'istante  $k$  stesso, e con  $\hat{\beta}_k \in \mathfrak{R}^n$  la stima dello stato a posteriori al passo  $k$  data la misura  $y_k$ . Gli errori sulla stima a priori e a posteriori possono essere definiti come

$$\bar{e}_k = \beta_k - \bar{\beta}_k \quad \text{e} \quad \hat{e}_k = \beta_k - \hat{\beta}_k.$$

La covarianza dell'errore sulla stima a priori si definisce come

$$\bar{P}_k = E[\bar{e}_k \bar{e}_k^T],$$

mentre la covarianza dell'errore sulla stima a posteriori è

$$\hat{P}_k = E[\hat{e}_k \hat{e}_k^T].$$

Per ottenere le equazioni del filtro di Kalman, si procede con l'obiettivo di trovare un'equazione che calcoli una stima a posteriori  $\hat{\beta}_k$  dello stato come combinazione lineare di una stima a priori  $\bar{\beta}_k$  e di una differenza pesata tra l'attuale misura  $y_k$  e la predizione della misura  $X\bar{\beta}_k$  come mostra la seguente formula

$$\hat{\beta}_k = \bar{\beta}_k + K(y_k - X\bar{\beta}_k).$$

La differenza  $(y_k - X\bar{\beta}_k)$  è detta innovazione sulla misura o residuale. Tale differenza rappresenta la discrepanza tra la misura predetta  $X\bar{\beta}_k$  e la misura attuale  $y_k$ . Un residuale pari a zero significa che le due misure sono in completo accordo.

La matrice  $K$ , di dimensione  $n \times m$ , è il fattore di guadagno del filtro ed è scelta in modo tale da minimizzare la covarianza dell'errore a posteriori  $\hat{P}_k$ . Un possibile risultato per  $K$  è dato da:

$$K_k = \frac{\bar{P}_k X^T}{X \bar{P}_k X^T + R},$$

dove  $R$  ricordiamo essere la matrice di covarianza dell'errore di misura e  $X$  è la matrice dei regressori. Si può osservare che, se la covarianza dell'errore di misura  $R$  tende a zero, il guadagno  $K$  pesa maggiormente il residuale. Infatti:

$$\lim_{R \rightarrow 0} K_k = X^{-1}.$$

Viceversa, quando è la covarianza dell'errore sulla stima a priori  $\bar{P}_k$  ad avvicinarsi a zero, allora il guadagno  $K$  pesa il residuale in maniera inferiore. Formalmente:

$$\lim_{\bar{P}_k \rightarrow 0} K_k = 0.$$

Queste osservazioni consentono di dare una nuova interpretazione alla funzione di  $K$ , ovvero quando la covarianza dell'errore di misura  $R$  tende a zero l'attuale misura  $y_k$  diventa sempre più affidabile, mentre la misura predetta  $X\bar{\beta}_k$  lo diventa sempre meno. È pertanto evidente come  $K$  possa essere interpretato come un fattore di credibilità che assegna più o meno importanza a ciò che si osserva nel tempo  $y_k$  o a ciò che si prevede di osservare  $X\bar{\beta}_k$ .

### 5.2.2 Il funzionamento dell' algoritmo di calcolo

Il filtro di Kalman stima lo stato di un processo utilizzando previsioni e correzioni delle stesse. In particolare lo stato viene stimato ad un certo istante e successivamente, in seguito alle misurazioni, viene effettuata una opportuna correzione. In conseguenza, le equazioni del filtro di Kalman si dividono in due gruppi: equazioni di previsione (o di "time update") ed equazioni di correzione (o di "measurements update"). Le equazioni di previsione servono a proiettare lo stato in avanti nel tempo e consentono pertanto di ottenere una stima a priori dello stato dell'oggetto all'istante successivo.

Le equazioni di correzione servono ad ottenere una migliore stima a posteriori. L'algoritmo è pertanto una continua alternanza tra le fasi di previsione e correzione.

Le equazioni specifiche per le previsioni sono le seguenti:

$$\begin{aligned}\bar{\beta}_k &= H\hat{\beta}_{k-1} + Gw_k; \\ \bar{P}_k &= H\hat{P}_{k-1}H^T + Q.\end{aligned}$$

Le equazioni per la fase di correzione sono invece:

$$\begin{aligned}K_k &= \bar{P}_k X^T (X\bar{P}_k X^T + R)^{-1}; \\ \hat{\beta}_k &= \bar{\beta}_k + K_k (y_k - X\bar{\beta}_k); \\ \hat{P}_k &= (I - K_k X)\bar{P}_k.\end{aligned}$$

Più precisamente, l'algoritmo si articola nei seguenti passi. Innanzitutto, durante la fase di correzione, bisogna calcolare il guadagno del filtro,  $K_k$ . Successivamente si misura il processo per ottenere  $y_k$ , quindi è possibile ottenere una stima dello stato a posteriori, basata sulla previsione precedente e sulla misura attuale. Il passo finale è la stima della covarianza dell'errore a posteriori  $\hat{P}_k$ . Dopo ogni fase di previsione e correzione, il processo viene reiterato utilizzando la precedente stima a posteriori per calcolare una nuova stima a priori. La natura ricorsiva è una delle caratteristiche più interessanti del filtro.

### 5.2.3 I parametri del filtro

Per essere implementato il filtro di Kalman necessita del settaggio della covarianza dell'errore di misura  $R$  e della covarianza dell'errore di processo  $Q$ . Una misura concreta di  $R$  è tipicamente possibile in quanto spesso, prima di avviare il processo da analizzare, si cerca di ottenere alcuni campioni di misura da cui determinarne la varianza. La determinazione della covarianza dell'errore di processo  $Q$  è invece generalmente più complessa. Spesso non esiste la possibilità di osservare direttamente il processo che si dovrà stimare. In alcuni casi è pertanto opportuno immettere sufficiente incertezza nel processo stesso tramite il settaggio di  $Q$ . Osserviamo infatti come una buona parte dell'affidabilità del filtro si gioca sulla regolazione di tale parametro  $Q$ .

In un caso o nell'altro, sia che si disponga o meno di una base razionale per settare i parametri, osserviamo come la performance del filtro (statisticamente parlando) dipenda dalla scelta di  $Q$  ed  $R$ . In particolare, se ad esempio  $Q$  ed  $R$  sono di fatto costanti, sia la covarianza dell'errore sulla stima  $P_k$  che il guadagno di Kalman si stabilizzano velocemente e rimangono costanti. In questo caso pertanto gli aggiornamenti delle stime si stabilizzano dopo un breve ciclo di iterazioni.

### 5.3 La stima della riserva sinistri mediante il filtro di Kalman: il modello

Ci proponiamo adesso di effettuare la valutazione della riserva sinistri mediante il filtro di Kalman al fine di verificarne alcune considerazioni circa la capacità di stima. Il modello si basa sugli algoritmi di calcolo proposti da Zehnwirth e De Jong nel loro articolo del 1983 "Claims Reserving, State-Space Models and the Kalman Filter". In particolare si assume che i risarcimenti effettuati nell'anno di calendario  $t = w + d$ , quindi relativi alla generazione  $w$  e liquidati con  $d$  anni di differimento, siano espressi, per ogni anno di generazione  $w$ , dalla relazione:

$$x(w, d) = \sum_{j=1}^p \phi_j(d) b_j(w),$$

dove  $\phi_j(d)$ ,  $j = 1, 2, \dots, p$  sono  $p$  funzioni di  $d$  note; mentre  $b_j(w)$ ,  $j = 1, 2, \dots, p$  sono parametri incogniti che dipendono dall'anno di generazione  $w$ .

Nel presente lavoro si è ipotizzato  $p = 1$ , ovvero che vi sia un'unica funzione che rappresenta il ritardo nei pagamenti  $d$  secondo la relazione

$$\phi(d) = (d + 1) \exp(-d),$$

la quale implica che i sinistri, relativamente ad ogni anno di origine, decrescono in modo per lo più esponenziale, al muoversi lungo ogni riga della matrice di run-off.

In conseguenza:

$$x(w, d) = b(w) \phi(d) = b(w) (d + 1) \exp(-d),$$

ed il vettore dei coefficienti  $b(w)$  è composto da un'unica componente incognita che individua la velocità con cui l'esponenziale tende a decrescere relativamente ad ogni anno di generazione  $w$ . Occorre osservare che il filtro di Kalman effettua le stime basandosi sui dati presi lungo le diagonali del triangolo di run-off.

Pertanto, a partire dalla matrice dati, il primo passo consiste nel disporre i valori in vettori, di dimensione diversa, ed i cui elementi siano i risarcimenti presi lungo le diagonali della matrice dati (in particolare per i risarcimenti relativi alla generazione 1 avremo un vettore di dimensione 1; per i risarcimenti relativi alla generazione 2 avremo un vettore di dimensione 2; e così via per le altre). In conseguenza la prima osservazione  $y(1)$ , relativa alla prima generazione, avrà un risarcimento atteso che coinvolge solo la quantità  $x(1,0)$  ed, in particolare il solo parametro  $b(1)$ ; la seconda osservazione  $y(2)$  coinvolgerà le quantità  $x(2,0)$  e  $x(1,1)$  e quindi i parametri  $b(1)$  e  $b(2)$ ; e così via fino al caso generale relativo alla  $t$ -esima osservazione che coinvolgerà le quantità  $x(t,0), x(t-1,1), \dots, x(1,t-1)$ , quindi i parametri  $b(t), b(t-1), \dots, b(1)$ . Questi ultimi valori  $b(1), b(2), \dots, b(t)$  individuano il vettore di stato; mentre la matrice dei regressori  $X$  è una matrice diagonale dove l'elemento in  $j$ -esima posizione è dato dalla quantità  $\exp(-j+1) \cdot j$ . Per collegare i vettori di stato tra loro occorre specificare come si evolvono i risarcimenti  $x(w,d)$  al variare dell'anno di origine  $w$  per un prefissato ritardo nei pagamenti  $d$ . In particolare, per ogni  $d$ :

$$x(w,d) = E[x(w,d) / x(w-1,d), x(w-2,d), \dots] + \eta(w,d),$$

dove  $\eta(w,d)$  è il termine di errore a media 0.

Questa espressione, per ogni prefissato ritardo  $d$ , esprime il legame, a meno del termine di errore  $\eta(w,d)$ , tra il risarcimento della generazione  $w$  e quello delle generazioni precedenti. Il modello assume che la media condizionata sia anche essa un polinomio di grado  $q-1$  che dipende dai valori osservati nelle passate generazioni, ovvero:

$$x(w,d) = \sum_{j=1}^q a(j)x(w-j,d) + \eta(w,d),$$

con

$$a(j) = \binom{q}{j} (-1)^{j-1}, \quad j = 1, 2, \dots, q.$$



Riprendendo l'espressione precedente  $x(w,d) = b(w)\phi(d)$  e sostituendola in questa si ottiene

$$\phi(d)b(w) = \sum_{j=1}^q a(j)\phi(d)b(w-j) + \eta(w,d),$$

da cui invertendo la funzione  $\phi$  si ha

$$b(w) = \sum_{j=1}^q a(j)b(w-j) + \Psi(d)\eta(w,d),$$

dove  $\Psi(d) = \phi^{-1}(d)$ .

La funzione  $\Psi(d)$  individua il vettore di regolazione dei disturbi del sistema, mentre dalle quantità  $\sum_{j=1}^q a(j)$  si ricava la matrice di transizione  $H$ . In particolare, nel modello esaminato si assume un polinomio di grado  $q = 1$ ,  $a(1) = 1$  per cui il legame tra i  $b(\cdot)$  è dato da:

$$b(w) = b(w-1) + \Psi(d)\eta(w,d),$$

per ogni  $d$ .

Ovvero

$$b(w) = b(w-1) + v(w),$$

dove  $v(w)$  è il termine di errore a media 0, e rappresenta il vettore dei disturbi del sistema. In altri termini, queste ipotesi implicano che la media condizionata di  $x(w,d)$  dato  $x(w-1,d), x(w-2,d), \dots$  è pari a  $x(w-1,d)$  per ogni  $d$ .

#### 5.4 Risultati e confronti con altri metodi

Utilizziamo la seguente matrice dati:

	1	2	3	4	5
2001	40.409.721	30.694.295	15.880.667	8.201.559	5.214.168
2002	46.294.758	44.353.590	18.304.246	8.833.488	
2003	51.590.220	50.605.820	17.603.713		
2004	56.599.000	53.743.000			
2005	62.461.276				

**Tabella 1: Matrice dati**

Il triangolo considerato è stato costruito sulla base dei dati medi di mercato osservato sulle imprese e distinto per fasce di mercato tratto dalle circolari Isvap e dalla documentazione ufficiale presente sul sito dell'Ania. Tale triangolo è stato costruito ipotizzando una compagnia di medie dimensioni che eserciti il ramo RcAuto. Si è considerato unicamente i valori relativi agli ultimi 5 anni di generazione, caratterizzati da un andamento relativamente omogeneo nei pagamenti dei sinistri. E' opportuno sottolineare che nelle valutazioni non sono state formulate ipotesi in merito all'eventuale effetto della componente inflazionistica sul costo del pagato ed è stato ipotizzato che i sinistri denunciati precedentemente l'anno 2001 siano stati completamente estinti, trascurando pertanto l'effetto del tail factor. Allo scopo di stimare l'ammontare dell'eventuale riserva sinistri che l'impresa dovrebbe accantonare a bilancio a fronte dei sinistri denunciati e non ancora liquidati, i dati considerati sono stati inizialmente analizzati attraverso l'utilizzo del metodo chain-ladder, avvalendosi di diverse varianti per la stima dei fattori di sviluppo. Applicando il classico metodo deterministico del chain-ladder si ottengono i seguenti risultati:

<b>Metodo chain-ladder classico</b>	
2001	-
2002	6.452.155,60
2003	17.548.560,48
2004	40.984.413,47
2005	102.049.986,16
<b>Riserva</b>	<b>167.035.115,70</b>

**Tabella 2: Chain-ladder classico**

Si osserva una riserva di circa 167 milioni di Euro con i seguenti fattori di sviluppo:

$\lambda_j$	1,92048489	1,19620739	1,08694133	1,05477859
-------------	------------	------------	------------	------------

**Tabella 3: Fattori di sviluppo del chain-ladder classico**

Utilizzando adesso una variante del metodo chain-ladder, basato sull'utilizzo dei fattori di sviluppo incrementali pesati secondo la seguente formula

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} w_{i,j} \cdot f_{i,j}}{\sum_{i=1}^{n-j+1} w_{i,j}}, \text{ dove } w_{i,j} = X_{i,j} \text{ e } f_{i,j} = \frac{X_{i,j}}{X_{i,j-1}},$$

si ottengono i seguenti fattori pesati

$\hat{\lambda}_j$	1,9279982	1,1984079	1,0874369	1,0547786
-------------------	-----------	-----------	-----------	-----------

**Tabella 4: Fattori di sviluppo del chain-ladder pesati**

ed i risultati sono i seguenti:

<b>Metodo chain-ladder con lambda pesati</b>	
2001	-
2002	6.452.155,60
2003	17.611.179,46
2004	41.331.905,56
2005	103.072.830,65
<b>Riserva</b>	<b>168.468.071,27</b>

**Tabella 5: Chain-ladder con lambda pesati**

Applichiamo adesso un primo metodo che costituisce una versione stocastica del tradizionale metodo chain-ladder, nel senso che il valor medio della distribuzione (Best Estimate non scontata) coincide con le stime prodotte da questo metodo deterministico. In particolare, analizziamo il metodo di Mack (“Distribution-Free Chain-Ladder”, DFCL), il quale ricordiamo non prevede ipotesi specifiche sulla distribuzione dei pagamenti futuri, limitandosi a modellarne la media e la varianza. Tale metodo consente di ottenere, relativamente alle diverse generazioni, la stima dell’errore di previsione (prediction error), dato dalla somma dell’errore di processo e dell’errore di stima. Utilizzando poi la formula di aggregazione, che non è semplicemente la somma dei singoli prediction error, ma utilizza una correzione che tiene conto della correlazione tra i dati, si ottiene una stima del prediction error globale. I risultati sono espressi nella seguente tabella:

<b>Modello di Mack</b>	<b>Pred. Error</b>	<b>PE/BE</b>
2001	-	-
2002	6.452.155,60	558.999,02
2003	17.548.560,48	1.426.071,56
2004	40.984.413,47	2.285.479,92
2005	102.049.986,16	3.854.514,61
<b>Riserva</b>	<b>167.035.115,70</b>	<b>6.006.439,47</b>

**Tabella 6: Prediction error con il modello di Mack**

Il prediction error globale è pari a circa 6 milioni. Osserviamo che tale valore è superiore a quello che si otterrebbe come somma delle singole prediction variance delle diverse generazioni pari a circa 4,7 milioni. La differenza, come anticipato, è dovuta alla correlazione insita nei dati tra le diverse generazioni, che, come detto, è considerata mediante l'introduzione del fattore di correzione. Si osserva inoltre un prediction error decrescente per le generazioni più recenti. L'effetto è dovuto, sia all'andamento regolare nella liquidazione dei pagamenti, ovvero risarcimenti maggiori nei primi anni di sviluppo per poi diminuire sempre di più nel corso del tempo, sia ai minori anni di differimento (per cui meno dati a disposizione). Infine si evidenzia un rapporto tra il prediction error e la best estimate, il cui valore è ovviamente coincidente con quello ottenuto con il metodo deterministico del chain-ladder, pari al 3,60%. Passiamo adesso ad analizzare il filtro di Kalman. In proposito ricordiamo che tale filtro è stato implementato ipotizzando che il parametro beta possa essere stimato sulla base di una funzione esponenziale del ritardo di pagamento. Abbiamo anche già sottolineato come uno dei punti più critici dell' algoritmo di calcolo, come si vedrà meglio in seguito nel dettaglio delle analisi di sensitività, è rappresentato dalla stima della matrice di covarianza dell'errore di misura, precedentemente indicata con  $R$  e che nel seguito indicheremo con  $U(t)$ . Tale matrice è stata stimata ipotizzando incorrelazione tra i diversi anni di generazione e stimando le varianze utilizzando semplicemente le varianze dei costi sinistri relativi ai diversi anni di sviluppo. La varianza relativa all'ultimo anno di differimento, non essendo possibile stimarla direttamente dai dati, è stata ottenuta individuando una funzione di interpolazione che meglio interpolasse le varianze precedenti. In particolare è stata individuata una curva esponenziale decrescente con la seguente equazione

$$y = 10^{14} \cdot x^{-4,6462},$$

da cui si ottiene la seguente matrice iniziale  $U(t)$ :

<b>U(t)</b>	1	2	3	4	5
1	$5,92 \cdot 10^{13}$	-	-	-	-
2	-	$7,82 \cdot 10^{13}$	-	-	-
3	-	-	$1,04 \cdot 10^{12}$	-	-
4	-	-	-	$9,98 \cdot 10^{10}$	-
5	-	-	-	-	$1,52 \cdot 10^{11}$

**Tabella 7: Matrice di covarianza dell'errore di misura**

Ai fini di un confronto tra il chain-ladder classico ed il filtro di Kalman, riportiamo nella seguente tabella i valori delle riserve ottenute con le due diverse metodologie, sia per ogni singola generazione, sia a livello globale:

<b>Generazione</b>	<b>Chain-ladder</b>	<b>Kalman</b>
2001	-	-
2002	6.452.155,60	10.770.363,08
2003	17.548.560,48	20.684.280,76
2004	40.984.413,47	46.043.936,19
2005	102.049.986,16	92.000.392,81
<b>Riserva</b>	<b>167.035.115,70</b>	<b>169.498.972,85</b>

**Tabella 8: Valori della riserva, chain-ladder e filtro di Kalman**

I primi risultati derivanti dall'applicazione del filtro di Kalman sembrano sovrastimare l'ammontare dei costi sinistri relativi alle generazioni passate e sottostimare i sinistri delle generazioni più recenti. In ogni caso, si osserva un valore della riserva globale molto simile a quello ottenuto mediante l'applicazione del chain-ladder classico. Tale risultato è ancora più evidente nel caso si consideri il chain-ladder basato sull'utilizzo dei fattori di sviluppo incrementali pesati, come mostra la seguente tabella:

	<b>Chain-ladder con lambda pesati</b>	<b>Kalman</b>
<b>Riserva globale</b>	168.468.071,27	169.498.972,85

**Tabella 9: Valori della riserva, chain-ladder con lambda pesati e filtro di Kalman**

Ovviamente la stima ottenuta con il filtro di Kalman sarà maggiormente vicina a quella ottenuta attraverso l'applicazione del chain-ladder, quanto più la funzione scelta per stimare il parametro beta sia effettivamente in grado di ben adattarsi ai dati osservati.

Tuttavia, conformemente a quanto avviene per altri modelli stocastici, come ad esempio l'Over Dispersion Poisson o il Bootstrapping, anche il filtro di Kalman restituisce un prediction error più contenuto di quello ottenuto con il modello di Mack.

	<b>Mack</b>	<b>Kalman</b>
Prediction Error	6.006.439,47	3.904.327,46
<b>PE/BE</b>	3,60%	2,30%

**Tabella 10: Prediction error, modello di Mack e filtro di Kalman**

In realtà, come precedentemente anticipato, il filtro di Kalman dipende notevolmente dai valori della matrice  $U(t)$  inizialmente ipotizzati. A tale proposito, utilizziamo ad esempio una nuova matrice  $U(t)$  in cui viene modificato unicamente il valore relativo al primo anno di sviluppo  $U(1)$ , stimato non più direttamente dai dati, bensì attraverso la funzione di interpolazione esponenziale precedentemente introdotta. In questo modo si ottiene un nuovo valore  $U(1)$  che crea un andamento decrescente della varianza al crescere del ritardo dei pagamenti:

<b>U(t)</b>	1	2	3	4	5
1	$2 \cdot 10^{14}$	-	-	-	-
2	-	$7,82 \cdot 10^{13}$	-	-	-
3	-	-	$1,04 \cdot 10^{12}$	-	-
4	-	-	-	$9,98 \cdot 10^{10}$	-
5	-	-	-	-	$1,52 \cdot 10^{11}$

**Tabella 11: Matrice di covarianza dell'errore di misura**

La nuova matrice  $U(t)$  produce un identico valore di riserva, ma una nuova stima del prediction error, in questo caso, molto vicino a quello ottenuto con il modello di Mack.

	<b>Mack</b>	<b>Kalman</b>
Prediction Error	6.006.439,47	5.475.403
<b>PE/BE</b>	3,60%	3,23%

**Tabella 12: Prediction error, modello di Mack e filtro di Kalman**

Continuiamo l'analisi di sensitività con lo scopo di testare il comportamento del modello al variare delle caratteristiche dei dati esaminati. A tal fine è stato creato un secondo triangolo di run-off, costruito modificando opportunamente i sinistri pagati nei primi due anni di sviluppo, in modo da ottenere un identico valore della riserva,

stimato con il metodo del chain-ladder, ma rendendo più irregolare l'ammontare dei pagamenti sostenuti dall'impresa.

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>2001</b>	40.129.785	20.694.295	15.880.667	8.201.559	5.214.168
<b>2002</b>	46.014.822	44.353.590	18.304.246	8.833.488	
<b>2003</b>	71.228.970	60.605.820	17.603.713		
<b>2004</b>	60.381.969	53.743.000			
<b>2005</b>	62.461.276				

**Tabella 13: Nuova matrice dati**

L'applicazione del modello di Mack, basato sul chain-ladder classico, ha permesso di ottenere nuovamente una stima della riserva sinistri e del prediction error.

	<b>Modello di Mack</b>	<b>Pred. Error</b>	<b>PE/BE</b>
2001	-	-	-
2002	7.216.152,01	795.828,18	11,03%
2003	23.752.950,85	3.051.094,46	12,85%
2004	42.341.904,98	5.200.957,00	12,28%
2005	93.724.107,87	7.610.563,26	8,12%
<b>Riserva</b>	<b>167.035.115,70</b>	<b>12.598.379,77</b>	<b>7,54%</b>

**Tabella 14: Prediction error con il modello di Mack**

L'ammontare della riserva sinistri risulta ovviamente coincidente con il caso precedente, in quanto il triangolo è stato costruito in modo da ottenere una identica best estimate della riserva. Il comportamento maggiormente variabile nei dati, si vedano ad esempio i pagamenti molto elevati effettuati nel terzo e nel quarto anno di generazione rispetto al caso precedente, si riflette invece sul prediction error. Infatti, a fronte di un incremento del 10% circa del coefficiente di variabilità relativo al costo dei sinistri (si passa da un valore pari a circa lo 0,54 nel caso precedente ad un valore di circa 0,62 in questo caso), si ottiene un rapporto tra PE/BE pari al 7,54%, ovvero raddoppiato rispetto all'esempio effettuato sul triangolo originale. Ciò trova la sua giustificazione proprio nell'effetto dell'incremento della variabilità dei dati.

A questo punto, utilizzando il nuovo triangolo di run-off, sono stati nuovamente inizializzati i valori necessari per implementare il filtro di Kalman, fra cui la matrice di covarianza  $U(t)$ .

<b>U(t)</b>	1	2	3	4	5
1	$1,29 \cdot 10^{14}$	-	-	-	-
2	-	$2,28 \cdot 10^{14}$	-	-	-
3	-	-	$1,04 \cdot 10^{12}$	-	-
4	-	-	-	$9,98 \cdot 10^{10}$	-
5	-	-	-	-	$1,52 \cdot 10^{11}$

**Tabella 15: Matrice di covarianza dell'errore di misura**

I risultati sono i seguenti:

<b>Generazione</b>	<b>Chain-ladder</b>	<b>Kalman</b>
2001	-	-
2002	6.452.155,60	10.770.363,08
2003	17.548.560,48	20.684.280,76
2004	40.984.413,47	46.043.936,19
2005	102.049.986,16	92.000.392,81
<b>Riserva</b>	167.035.115,70	169.498.972,85

**Tabella 16: Valori della riserva, chain-ladder e filtro di Kalman**

Il filtro di Kalman permette di ottenere valori identici al caso precedente: non sono state variate le ipotesi relative al parametro beta, alla funzione esponenziale utilizzata per stimarlo ed inoltre i valori del costo sinistri pagati nell'ultimo anno di calendario (diagonale della matrice) sono gli stessi di quelli relativi al caso precedente. Tuttavia le nuove inizializzazioni sulle matrici di covarianza (di misura e di processo) restituiscono la stima di un nuovo prediction error.

	<b>Mack</b>	<b>Kalman</b>
Prediction Error	12.598.379,77	8.104.770
<b>PE/BE</b>	7,54%	4,78%

**Tabella 17: Prediction error, modello di Mack e filtro di Kalman**

Il filtro di Kalman sembra cogliere in misura meno marcata rispetto al modello di Mack la variazione tra i prediction error. Tale risultato si conferma in linea con quelli ottenuti da England e Verrall, secondo i quali i principali metodi relativi ai modelli lineari generalizzati (GLM) o al metodo del Bootstrapping restituiscano valori del prediction error in alcuni casi molto più contenuti di quelli ottenuti col modello di Mack.



Anche considerando una nuova matrice  $U(t)$ , corretta mediante una funzione di interpolazione allo scopo di garantire una maggiore regolarità nell'andamento della varianza, si ottiene un valore del prediction error più vicino a quello ottenuto con il modello di Mack, ma comunque inferiore.

	<b>Mack</b>	<b>Kalman</b>
Prediction Error	12.598.379,77	8.104.770
<b>PE/BE</b>	7,54%	4,78%

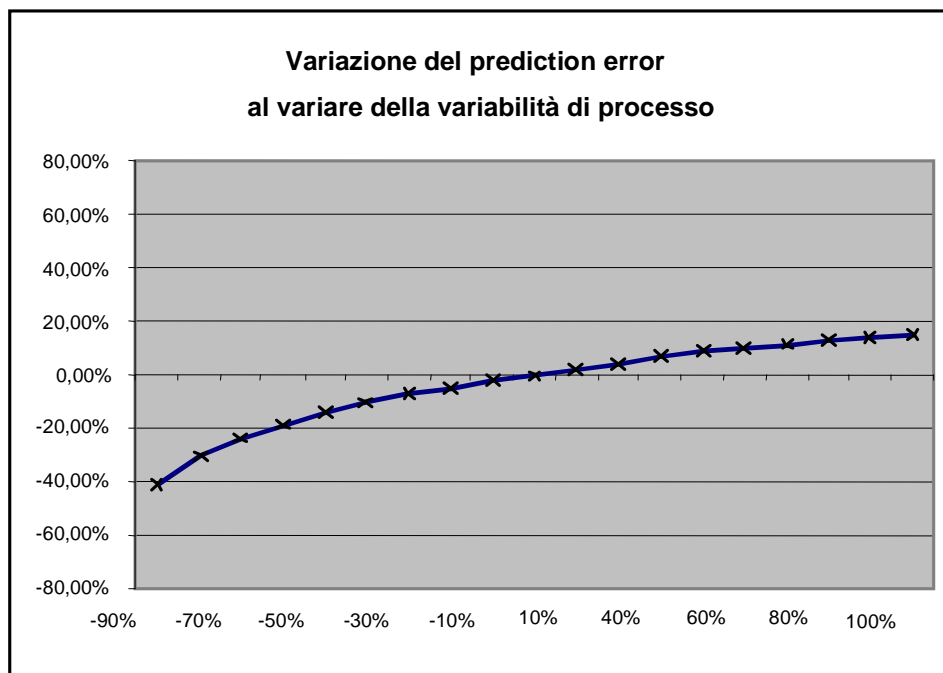
**Tabella 18: Prediction error, modello di Mack e filtro di Kalman**

Per concludere l'analisi sembra interessante verificare l'andamento del prediction error al variare dell'inizializzazione della matrice di covarianza di processo, precedentemente indicata con  $Q$  e che nel seguito indicheremo con  $V(t)$ . Di seguito sono rappresentati alcuni risultati:

<b>Incremento della variabilità del processo</b>	<b>Prediction error</b>	<b>Incremento del prediction error</b>
-90%	2.310.018,26	-40,83%
-80%	2.720.383,04	-30,32%
-70%	2.988.139,55	-23,47%
-60%	3.190.781,41	-18,28%
-50%	3.355.170,21	-14,07%
-40%	3.494.085,05	-10,51%
-30%	3.614.689,28	-7,42%
-20%	3.721.432,93	-4,68%
-10%	3.817.283,84	-2,23%
<b>Standard</b>	<b>3.904.327,46</b>	
10%	3.984.089,12	2,04%
20%	4.057.720,50	3,93%
40%	4.189.975,84	7,32%
50%	4.249.875,56	8,85%
60%	4.306.279,19	10,30%
70%	4.359.573,40	11,66%
80%	4.410.082,75	12,95%
90%	4.458.082,47	14,18%
100%	4.503.808,11	15,35%

**Tabella 19: Prediction error al variare della variabilità del processo**

La tabella mostra il caso standard relativo al triangolo dei dati originali confrontato con diversi scenari di variabilità del processo. In particolare sono stati considerati sia casi di incrementi di variabilità negativi che casi di incrementi di variabilità positivi. Si osserva che la variazione del prediction error aumenta all'aumentare della variabilità del processo, sebbene in quantità nettamente minore che proporzionale. Analogamente un effetto del tutto opposto si osserva in caso di riduzione di variabilità di processo. In questo caso la riduzione influisce in misura più marcata: ad esempio a fronte di una diminuzione della variabilità di processo pari al 90% si registra una riduzione dell'errore di previsione pari al 40% circa, valore quindi non trascurabile. Ciò conferma quanto possano influire i valori inizialmente assegnati alle matrici di covarianza sulle stime ottenute con il modello. Al fine di visualizzare meglio gli incrementi di errore di misura ottenuti sono stati riprodotti i valori espressi nella precedente tabella (tabella 19) nel grafico che segue.



**Figura 6: Variazione del prediction error al variare della variabilità di processo**

### **5.5 L'impostazione del problema secondo il progetto Solvency II: il contesto per la definizione e il calcolo dei requisiti di riserva.**

Alla luce dello stato attuale delle cose, non è possibile concludere il lavoro senza considerare l'impatto che il progetto Solvency II produrrà sulla valutazione delle riserve e dei requisiti minimi di capitale. Negli ultimi anni si sono infatti sviluppati diversi sistemi di solvibilità e tuttora si stanno sviluppando. Tutti presentano l'esigenza di fornire una struttura di capitale più rappresentativa dell'effettiva rischiosità di ogni singola impresa, caratteristica che dovrebbe manifestarsi sia mediante una più opportuna valutazione dei rischi tecnici, sia mediante l'introduzione di nuove categorie di rischio. Per quel che attiene al presente lavoro, molti studi hanno sottolineato la necessità di correlare i requisiti minimi di capitale, oltre che ai principi contabili, anche a ciò che verrà stabilito in termini di valutazione delle riserve tecniche. In particolare la Commissione Europea ha richiesto al CEIOPS di dare una forma quantitativa al principio di prudenza espresso nelle direttive al fine di promuovere in Europa, attraverso standard comuni, un approccio più armonizzato per quel che riguarda il livello di prudenza da considerare nelle riserve tecniche. Il nuovo sistema di calcolo delle riserve dovrebbe basarsi sulla somma di due distinte componenti:

- la best estimate, quale valore medio atteso dei futuri impegni nei confronti degli assicurati;
- il risk margin, quale margine addizionale predisposto proprio per formalizzare quantitativamente l'incertezza degli impegni futuri stessi, generata da fattori di rischio esogeni.

È evidente pertanto come l'impiego di metodologie stocastiche si diffonderà sempre di più, visto che, in un simile contesto, una definizione di best estimate e risk margin è possibile solo mediante l'adozione di un modello di valutazione stocastico alle serie storiche dei sinistri. Mediante le procedure stocastiche è infatti sempre possibile ottenere una stima dei momenti di primo e secondo ordine, in particolare della media e della varianza. Nel caso in cui le ipotesi del modello specificano anche la forma della distribuzione della riserva, sarà ovviamente possibile ricavare anche la distribuzione completa della stessa.

Per i modelli cosiddetti “distribution free” (vedi Mack), invece, la distribuzione completa della riserva dovrà essere ricavata dai momenti di primo e secondo ordine, specificando un’opportuna ipotesi aggiuntiva. I calcoli effettivi potranno poi essere ricavati analiticamente, oppure tramite tecniche di simulazione. La best estimate sarà poi ricavata dalla media della distribuzione della stima degli impegni futuri in essere, mentre il risk margin sarà considerato come un opportuno caricamento in grado di considerare la variabilità del processo di liquidazione futura dei sinistri. Tuttavia, mentre la definizione della best estimate come valore atteso della distribuzione dei futuri impegni nei confronti degli assicurati è condivisa da molti, vi è in corso un ampio dibattito per quel che riguarda la definizione quantitativa del risk margin, anche se è prassi comune che esso debba rappresentare un margine prudenziale destinato a compensare l’incertezza attinente gli importi futuri delle prestazioni secondo una impostazione coerente con una valutazione di mercato. In questa ottica il risk margin rappresenterebbe il premio per il rischio che un altro assicuratore richiederebbe per partecipare agli impegni in essere del portafoglio considerato in una libera trattativa di mercato.

Per quanto riguarda i requisiti di capitale sembra che debba essere costituito un Minimum Capital Requirement (MCR), calcolato con una formula semplice e con il ruolo di rappresentare il livello di sicurezza, ed un Solvency Capital Requirement (SCR) destinato ad individuare il livello di capitale che un assicuratore dovrebbe disporre, in un approccio di tipo going-concern, prefissato un certo livello di confidenza ed un certo orizzonte temporale. Al fine di facilitare l’intervento dell’Autorità di Vigilanza si sta valutando anche un ulteriore indicatore, basato su un orizzonte temporale più lungo, con il ruolo di fornire un aggiuntivo segnale di rischio. La principale caratteristica richiesta al requisito MCR è la semplicità, in modo tale da essere facilmente comparabile. Alcuni studi hanno messo in evidenza la possibilità di considerare nel calcolo anche un margine di rischio legato alle riserve tecniche. Per il Solvency Capital Requirement, invece, la direzione è quella di una formula di calcolo basata sulla combinazione di due scenari: un Factor-Based-Approach e una Scenario-Based-Approach. Riassumendo, rispetto a quanto avvenuto fino ad oggi, la valutazione dei requisiti minimi di capitale sembra dovere includere più fonti di rischio, eventualmente variabili da impresa ad impresa, ed un approccio

che consideri diversi scenari, siano essi statici o dinamici, con lo scopo di verificare l'impatto che questi possono avere sul patrimonio dell'impresa. Tuttavia ci sono ancora molte questioni aperte: i fattori di rischio da utilizzare, la probabilità di rovina da prefissare (probabilmente pari allo 0,5%), l'orizzonte temporale su cui effettuare le valutazioni (probabilmente 1 anno) ed inoltre la misura di rischio con cui fare le stime (Var o TailVar). Come ultima osservazione non è da escludere che possano essere previsti modelli interni alle imprese (Internal Risk Models) al fine di conseguire una più accurata stima del rischio rispetto ad una formula standard. Sicuramente tali modelli presentano una maggiore flessibilità ed adattabilità alle specifiche caratteristiche di ogni singola impresa, seppure più complessi e costosi da implementare rispetto ad una formula standard di calcolo, la quale, tra l'altro, non potrebbe essere rivista così frequentemente dall'Autorità di Vigilanza.

### **5.5.1 Risk margin e requisiti di riserva**

Abbiamo visto come il risk margin debba essere valutato secondo un approccio basato sui valori di mercato. Sono evidenti le difficoltà che si presentano in uno schema di riferimento di questo tipo. È infatti noto che in generale non esistono mercati talmente ampi ed efficienti da comprendere le passività danni, né i tipi di contratto ad esse assimilabili e sui cui prezzi potere, quindi, calibrare un modello consistente con il mercato e che incorpori il premio per il rischio. Pertanto in questi casi è necessario costruire modelli che trattino coerentemente le aspettative di mercato ed i premi per il rischio mediante distribuzioni di probabilità naturali, nel caso rappresentate dalle metodologie stocastiche di stima della riserva sinistri. In altre parole l'entità della riserva  $R$  dovrà essere pari a

$$R = v(t, L),$$

dove  $v(t, X)$  rappresenta il prezzo di mercato al tempo  $t$  dell'importo aleatorio degli impegni futuri  $L$ , ovvero il valore di mercato degli impegni di rimborso in essere. Secondo questa definizione il risk margin risulterebbe individuato da un caricamento implicito (risk loading),  $\lambda$ , da aggiungere all'aspettativa degli obblighi futuri,  $E(L)$ , come ulteriore premio richiesto nel mercato da investitori avversi al rischio per diventare titolari della passività  $L$ .

In questa ottica il caricamento aggiuntivo  $\lambda$  può essere appropriatamente definito market value margin. Nell'ambito del progetto Solvency II sono stati presi in esame due metodi alternativi per il calcolo del risk margin:

- il quantile approach, secondo il quale il risk margin è definito come la differenza tra il valore degli obblighi futuri scelto ad un prefissato livello di probabilità, tipicamente uguale al 75-esimo o 90-esimo quantile, e la best estimate;
- il cost-of-capital approach, secondo il quale il risk margin è dato dal costo del patrimonio di vigilanza (reserve risk capital) richiesto dalla normativa per coprire il rischio di insolvenza fino all'estinzione completa della riserva sinistri.

Nel secondo caso, per calcolare il risk margin come costo del capitale, è ovviamente necessario definire quale capitale prendere a riferimento per il calcolo. Innanzitutto tale capitale (reserve risk capital) viene richiesto all'assicuratore come una disponibilità di vigilanza aggiuntiva al risk margin che può risultare utile a fine anno nel caso in cui la riserva sinistri, pur includendo il risk margin, non sia sufficiente a coprire una inattesa perdita sui pagamenti. Ai fini dei calcoli, si prende a riferimento un orizzonte temporale di un anno e si seleziona un valore particolarmente pessimistico del fabbisogno di riserva (Worst case), determinato sulla base di un prefissato ed elevato livello di probabilità. Il reserve risk capital è poi dato dalla differenza tra il Worst case ed il valore della riserva richiesto (best estimate più risk margin). Nel presente lavoro verrà analizzato il calcolo del risk margin secondo l'approccio del quantile e sarà effettuato un confronto con il Solvency Capital Requirement (SCR) precedentemente introdotto per il reserve risk calcolato così come viene descritto nel secondo studio di impatto quantitativo (QIS2) effettuato dal CEIOPS.

### **5.5.2 Risk margin: risultati**

Volendo calcolare il risk margin per il modello di Mack precedentemente analizzato è necessario specificare ulteriori ipotesi circa la distribuzione di probabilità della riserva.

Essendo infatti il modello di Mack distribution free, a differenza di altri metodi stocastici, garantisce la stima unicamente del prediction error, senza fare alcuna assunzione sulla distribuzione sottostante i sinistri pagati. Riprendendo alcuni autori, tra cui England e Verrall, è stata ipotizzata l'ipotesi di una distribuzione normale e lognormale, con media pari alla best estimate e standard deviation pari al prediction error stimato dal modello di Mack stesso, allo scopo di fornire una prima quantificazione indicativa del risk margin.

I risultati sono stati i seguenti:

<b>Risk Margin (Modello di Mack)</b>		
	Ipotesi Normale	Ipotesi Lognormale
75° percentile	171.086.397,56	171.082.560,32
90° percentile	174.732.677,61	174.797.758,78
75°/B.E.	2,425%	2,423%
90°/B.E.	4,608%	4,647%

**Tabella 20: Risk margin, modello di Mack**

Dall'analisi dei dati si osserva che le ipotesi della distribuzione sottostante non producono grosse differenze in termini di risk margin. Il comportamento maggiormente asimmetrico dell'ipotesi di lognormalità si riflette solo sui percentili più elevati (nel nostro caso il 90-esimo). Sul 75-esimo percentile, che rappresenta il livello attualmente individuato dall'Apra e preso come livello standard per il calcolo del risk margin secondo il "quantile approach" dallo stesso Ceiops, si osservano invece valori molto simili. In realtà le differenze contenute sono dovute anche al basso grado di variabilità e di asimmetria presente nei dati: la riserva totale assume infatti un coefficiente di variabilità pari a 0,036 e un indice di asimmetria prossimo allo zero (pari a 0,000046). Confrontiamo adesso il risk margin ottenuto con il modello di Mack con il risk margin che si ottiene effettuando le stime con il filtro di Kalman, sfruttando in entrambi i casi l'ulteriore ipotesi di lognormalità della distribuzione della riserva. In particolare, il risk margin calcolato mediante il filtro di Kalman è stato valutato sia utilizzando la prima matrice di covarianza  $U(t)$  stimata dai dati, sia utilizzando la matrice corretta con interpolazione lineare per garantire maggiore regolarità nell'andamento della varianza (in tabella indicato con modello di Kalman bis).

<b>Risk Margin</b>			
Lognormale	Modello di Mack	Modello di Kalman	Modello di Kalman (bis)
BE	167.035.115,70	169.498.972,85	169.498.972,85
75° percentile	171.082.560,32	172.144.336,52	173.239.253,30
90° percentile	174.797.758,78	174.529.853,81	176.615.630,47
75°/B.E.	2,42%	1,56%	2,21%
90°/B.E.	4,65%	2,97%	4,20%

**Tabella 21: Risk margin, confronto tra modelli**

L'incidenza del risk margin rispetto alla best estimate dipende ovviamente dalla variabilità della distribuzione della riserva stessa. In questo caso il modello di Mack fornisce stime più elevate rispetto al primo modello di Kalman, specialmente sul 90-esimo percentile. Tuttavia tali stime tendono ad allinearsi se si considera il modello di Kalman bis. In particolare il modello di Kalman bis si dimostra più prudentiale rispetto al modello di Kalman originario, nel senso che fornisce, a parità di best estimate, intervalli di confidenza più ampi e, quindi, riserve sinistri (best estimate + risk margin) più elevate. A conferma del grande impatto che la variabilità presente nei dati apporta al calcolo del risk margin, si effettuano le stesse valutazioni utilizzando la seconda matrice dati precedentemente introdotta (tabella 13) e relativa ad una sinistrosità volutamente più variabile.

<b>Risk Margin (Modello di Mack)</b>		
Lognormale	Modello di Mack	Modello di Kalman
BE	167.035.115,70	169.498.972,85
75° percentile	175.532.593,72	174.930.964,12
90° percentile	183.180.589,02	179.998.816,19
75°/B.E.	5,09%	3,20%
90°/B.E.	9,67%	6,19%

**Tabella 22: Risk margin, confronto tra modelli**

In questo caso si osservano differenze significative già al 75-esimo percentile. Tali differenze aumentano maggiormente sui percentili più elevati anche per effetto dell'ipotesi di lognormalità utilizzata. La lognormale a due parametri ha infatti un indice di asimmetria crescente al crescere del coefficiente di variabilità, che adesso è circa 0,62 a differenza di un valore di circa 0,54 nel caso precedente.



### 5.5.3 Un confronto coi requisiti di capitale nei QIS2

Può essere interessante calcolare il requisito di capitale secondo le regole previste dal QIS2 utilizzando la best estimate ed il risk margin ottenuto attraverso l'applicazione del filtro di Kalman. Il secondo studio di impatto quantitativo (QIS2) ha avuto inizio nel marzo 2006 e fa seguito ad un precedente studio (QIS1) effettuato negli ultimi mesi del 2005 con lo scopo di verificare i livelli di prudenza delle riserve tecniche al variare di alcune ipotesi di scenario. Nelle specifiche tecniche associate al QIS2, il requisito di capitale previsto dovrà essere ottenuto dalla combinazione di più requisiti di capitale relativi a diverse tipologie di rischio stabilite dalla Commissione Europea e dal CEIOPS. I vari rischi dovranno inoltre essere aggregati considerando diversi scenari possibili: di correlazione, di indipendenza e senza gli effetti della diversificazione. Per quel che attiene agli obiettivi del presente lavoro ci si soffermerà unicamente sul metodo previsto per la valutazione dell'underwriting risk nei rami danni che, come è noto, rappresenta il rischio complessivo derivante dalla stipula dei contratti assicurativi, somma di tre sotto rischi: reserve risk, premium risk e cat risk. In particolare verrà esaminato il primo dei tre sotto rischi, quello relativo alla riservazione. Ai fini dei calcoli del Solvency Capital Requirement relativo al reserve risk  $K^{(QIS)}$  si eseguono i seguenti passi:

- 1) vengono definiti 11 rami o linee di business (Line of Business, LoB) e per ogni ramo  $k$  viene individuato un valore "market-wide" relativo al fattore di volatilità  $f_k$ ;
- 2) la volatilità di ramo per ogni impresa è calcolata moltiplicando il fattore di volatilità  $f_k$  sopra introdotto per un cosiddetto size factor  $s_k$  rappresentativo della dimensione dell'impresa.

Nello specifico:

$$\sigma_k = f_k s_k,$$

dove il size factor  $s_k$  è una funzione non crescente della riserva sinistri (PCO, Provision for Claims Outstanding) al lordo della riassicurazione.

In particolare tale fattore è circa 2,236 per valori di riserva lorda minori di 20 milioni di euro, ha un andamento decrescente all'aumentare del valore della riserva fino a 100 milioni ed assume un valore costantemente pari ad 1 se la riserva è maggiore di 100 milioni.

Ossia:

$$s_k = \begin{cases} 1 & \text{se } PCO \geq 100m \\ \frac{10}{\sqrt{PCO \cdot 10^{-6}}} & \text{se } 100m > PCO \geq 20m \\ \frac{10}{\sqrt{20}} & \text{altrimenti} \end{cases}$$

- 3) la volatilità per unità di riserva  $\sigma$  relativa all'intero comparto danni si ricava aggregando le singole volatilità di ramo secondo la seguente formula:

$$\sigma^2 = \sum_{k=1}^{11} \sum_{j=1}^{11} w_k w_j \sigma_k \sigma_j c_{kj},$$

dove

-  $\{c_{kj}\}$  è una matrice di correlazione specificata dal CEIOPS;

- i coefficienti di ponderazione  $w_k$  esprimono l'incidenza percentuale della riserva sinistri di ramo (al netto della riassicurazione)  $R_k$ , rispetto alla riserva complessiva di tutti i rami  $R_{tot}$ , ovvero  $w_k = \frac{R_k}{R_{tot}}$ ;

- 4) il Basic Solvency Capital Requirement (BSCR) relativo al rischio di riservazione è ottenuto come:

$$BSCR = \rho(\sigma) R_{tot},$$

dove  $\rho(\cdot)$  è una funzione espressa da

$$\rho(x) = \frac{0.99 - N\left(N^{-1}(0.99) - \sqrt{\log(x^2 + 1)}\right)}{0.01},$$

e  $N(\cdot)$  è la funzione di ripartizione della normale standard.

5) il SCR è infine ottenuto da:

$$K^{(QIS)} = BSCR - PL,$$

dove  $PL$  individua il profitto o la perdita attesa relativa al prossimo anno di esercizio. In particolare tale quantità è data dalla somma di due componenti:

- una relativa ai premi  $PL_{prem}$  ed espressa con il valore atteso della differenza tra i premi di competenza ed i corrispondenti costi (per rimborso dei sinistri e spese) attinenti al prossimo anno di esercizio;
- una relativa alle riserve  $PL_{res}$  il cui valore è essenzialmente dato dalla parte di risk margin predisposta per il prossimo anno di esercizio (in proporzione ai livelli di riserva).

È stato dunque effettuato il calcolo del requisito di capitale, secondo le regole previste dal QIS2, utilizzando la best estimate ed il risk margin ottenuti mediante l'applicazione del filtro di Kalman ai dati analizzati e relativi al ramo RcAuto (vedi tabella, valori in migliaia di euro).

<b>2001</b>	40.409.721	30.694.295	15.880.667	8.201.559	5.214.168
<b>2002</b>	46.294.758	44.353.590	18.304.246	8.833.488	
<b>2003</b>	51.590.220	50.605.820	17.603.713		
<b>2004</b>	56.599.000	53.743.000			
<b>2005</b>	62.461.276				

**Tabella 23: Matrice dati**

Per semplicità, è stato ipotizzato che l'impresa non faccia alcun ricorso a coperture riassicurative ottenendo pertanto una piena coincidenza tra riserve al netto e riserve al lordo della riassicurazione. In linea con i valori osservati nella matrice dati per le quantità del primo anno di sviluppo in relazione alle generazioni precedenti, si è inoltre ipotizzato che la proporzione  $\alpha$  di sinistri posti a riserva e pagati nell'anno successivo sia pari al 40%, così come mostra la seguente tabella.

<b>Total value Best Estimate (BE)</b>	Gross	Net
Claims outstanding	167.035	167.035
<b>Total value 75th percentile</b>	Gross	Net
Claims outstanding	173.042	173.042
Expected prop. of claims paid in next year $\alpha$	40%	

**Tabella 24: Best Estimate, 75-esimo percentile e valore di  $\alpha$ .**

Nel Qis2, il ramo RcAuto corrisponde alla linea di business  $k = 2$  ed il relativo fattore di volatilità “market wide” è  $f_2 = 0,15$ . Poiché l’analisi è stata effettuata considerando solamente il ramo RcAuto, anche il calcolo del Solvency Capital Requirement è stato realizzato a livello di singolo ramo, ponendo  $R_k = 0$  per  $k \neq 2$  e considerando il valore di SCR ottenuto come un requisito di capitale “stand-alone”. Anche i profitti/perdite attesi  $PL$  sono ovviamente stati calcolati per il singolo ramo. Come previsto nelle indicazioni del QIS2, la PCO si è fatta corrispondere con il valore della riserva al 75-esimo percentile. Tale valore ha permesso la stima del size factor  $s_k = 1$  (in quanto la riserva al lordo della riassicurazione è superiore a 100 milioni di Euro) e quindi della volatilità di ramo in relazione alla dimensione dell’impresa (“market wide standard deviation”)  $\sigma_k = 0,15$ . La trasformazione basata sul 99-esimo percentile permette poi di giungere alla stima del Basic Solvency Capital Requirement (BSCR) o Risk Charge, moltiplicando il Volatility Scr ottenuto dall’applicazione di tale trasformazione per l’ammontare delle riserve totali al netto della riassicurazione. Infine, per quanto riguarda il calcolo dei profitti/perdite attesi, (in tabella indicati con Expected Surplus), la componente di riserva  $PL_{res}$ , così come previsto dal QIS2, è stata calcolata come frazione di risk margin predisposta per il prossimo anno di esercizio,

$$PL_{res} = risk\ margin \times \alpha ;$$

mentre è stata trascurata la componente di profitti/perdite relativa ai premi  $PL_{prem}$ .

I risultati sono i seguenti:

Size factor ( $s_k$ )	1
Volatility factor ( $f_2$ ), RcAuto	15%
Market wide standard deviation ( $\sigma_k$ )	15%
Volatility Scr	47,3%
Risk Charge (BSCR)	81.909,47
Expected Surplus	2.402,58
<b>SCR</b>	<b>79.506,89</b>
<b>SCR/BE</b>	<b>47,60%</b>

**Tabella 25: Solvency Capital Requirement.**

Il requisito di capitale, riguardante unicamente il reserve risk, risulta pari a circa il 47% della best estimate. In particolare ipotizzando ad esempio che la riserva sinistri sia esattamente pari al 100% dei premi di tariffa<sup>18</sup> si ottiene un requisito SCR pari al 47% dei premi. Tale requisito, pur rappresentando solo una minima parte dell'eventuale requisito complessivo, risulta dunque già notevolmente superiore all'attuale Margine di Solvibilità previsto dalle regole del Solvency I. Occorre però ricordare che l'esempio considera solamente un'ipotetica compagnia con un'unica linea di business (lob). Tale ipotesi porta ad una semplificazione molto significativa, in quanto trascura sia gli effetti della correlazione tra i rami sia gli effetti della diversificazione tra i rischi. Alcune analisi hanno mostrato come la gestione di più rami assicurativi, anche correlati positivamente, porta ad una riduzione del rapporto SCR/Premi ottenuto con il Qis2 per effetto appunto della diversificazione. Tale risultato è chiaramente in linea con un principio base della Teoria del Rischio che prevede che il Capitale a Rischio sia una funzione crescente e con concavità rivolta verso il basso al crescere del portafoglio dell'impresa.

---

<sup>18</sup> Si sottolinea che generalmente per il ramo RcAuto la riserva sinistri assume valori decisamente superiori al 100% dei premi

## CONCLUSIONI

Il presente lavoro si inserisce nell'ampio filone di ricerca riguardante la valutazione della riserva sinistri per le imprese di assicurazioni contro i danni. La possibilità di stabilire modelli stocastici di stima da imporre alle imprese assicurative e, quindi, agli attuari, al fine di valutare le loro passività appare sempre più plausibile anche alla luce di alcune recenti proposte in discussione presso il CEIOPS.

Il presente studio ha pertanto inizialmente proposto una panoramica di alcune delle principali metodologie statistiche di calcolo utilizzate nella letteratura attuariale per valutare la riserva sinistri stessa. I risultati ottenuti dall'analisi empirica dei dati ci portano innanzitutto ad effettuare alcune considerazioni sulle ipotesi di partenza alla base dei metodi di stima prescelti. Uno dei punti più critici nelle applicazioni dei modelli statistici risulta, infatti, la scarsità di dati, la stabilità delle condizioni di stima e la scelta delle ipotesi da utilizzare circa la distribuzione di probabilità per rappresentare, nel nostro caso, i risarcimenti per sinistro. Questi aspetti dovranno essere regolamentati dagli Organi di Vigilanza con lo scopo di favorire la trasparenza dei modelli e limitare o, quantomeno, controllare la soggettività presente nelle valutazioni di bilancio, comunque non del tutto eliminabile. La soggettività ha il suo peso sia sulla scelta del modello che sulla scelta delle ipotesi da utilizzare per implementarlo.

L'utilizzo del filtro di Kalman permette di superare, in parte, tale inconveniente. Anche in questo modello, infatti, sono presenti alcuni elementi soggettivi in input, tuttavia tale metodologia di calcolo ha il grosso vantaggio di potere essere aggiornata nel tempo sulla base di nuove informazioni disponibili, provenienti dall'esterno. È pertanto evidente la grande flessibilità che contraddistingue il filtro e che si contrappone ai modelli statici che impongono ipotesi rigide e fisse. Con il filtro di Kalman è possibile considerare le variazioni che possono verificarsi nel corso del tempo sul processo di liquidazione dei sinistri a seguito di fattori esterni o interni all'impresa di assicurazione stessa.

A tal fine, ricordiamo come tale filtro sia individuato, ad ogni ciclo di iterazione, da due equazioni fondamentali: la prima lega l'osservazione corrente al parametro incognito, la seconda stabilisce il legame tra il parametro incognito dell'epoca attuale ed il parametro incognito relativo allo step precedente.

Questo è uno dei più grandi meriti del filtro ed è ciò che permette di aggiornare le stime. Non è da escludere poi che tali relazioni possano essere modificate nel corso del tempo, stabilendo nuovi legami tra le quantità oggetto di studio. Si può pensare, infatti, di far variare le relazioni introducendo, ad esempio, ulteriori parametri incogniti o cambiare totalmente la forma funzionale che collega i parametri tra due step consecutivi. In questo modo, sfruttando tecniche dinamiche di stima, è possibile incorporare in tempo reale (step dopo step) le informazioni aggiuntive che possono influenzare il fenomeno analizzato. Quanto all'analisi dei dati, abbiamo visto che il filtro di Kalman sembra adattarsi bene alla stima della riserva sinistri. Come altri modelli stocastici, esso consente di ottenere la best estimate ed il relativo errore di previsione sia separatamente per ogni anno di generazione sia per il valore complessivo della riserva. In particolare, sui dati analizzati, la stima della best estimate si pone sufficientemente in linea con i risultati prodotti mediante l'utilizzo del metodo deterministico del chain-ladder. Tale risultato è ancora più evidente nel caso si consideri il chain-ladder basato sull'utilizzo di fattori di sviluppo incrementali pesati. Ovviamente la stima ottenuta con il filtro di Kalman sarà maggiormente vicina a quella ottenuta attraverso l'applicazione del chain-ladder, quanto più la funzione scelta per stimare il parametro incognito sia effettivamente in grado di ben adattarsi ai dati osservati. Conformemente a quanto avviene per altri modelli stocastici, come ad esempio l'Over Dispersion Poisson o il Bootstrapping, anche il filtro di Kalman restituisce un prediction error più contenuto di quello ottenuto con il modello di Mack. In realtà, il filtro dipende notevolmente dai valori di inizializzazione dell'algoritmo di calcolo tra cui la matrice di covarianza dell'errore di misura  $U(t)$  e la matrice di covarianza dell'errore di processo  $V(t)$ . Le analisi di sensitività hanno permesso di constatare che modificando gli elementi della matrice  $U(t)$  si ottengono valori del prediction error diversi a parità di best estimate.

Utilizzando, ad esempio, una matrice  $U(t)$  modificata unicamente nel valore relativo al primo anno di sviluppo  $U(1)$ , opportunamente stimato in modo da creare un andamento decrescente della varianza al crescere del ritardo dei pagamenti, si ottiene un identico valore di riserva, ma una nuova stima del prediction error, in questo caso, molto vicino a quello ottenuto con il modello di Mack.

Ulteriori analisi effettuate allo scopo di verificare l'adattamento del filtro al variare dei dati di partenza, questi ultimi modificati per rendere più irregolare l'ammontare dei pagamenti sostenuti dall'impresa, hanno mostrato che il filtro di Kalman sembra cogliere in misura meno marcata rispetto al modello di Mack la variazione tra i prediction error. Anche considerando una ulteriore matrice  $U(t)$ , corretta mediante una funzione di interpolazione allo scopo di garantire una maggiore regolarità nell'andamento della varianza, si è osservato un valore del prediction error più vicino a quello ottenuto con il modello di Mack, ma comunque inferiore.

L'effetto del cambiamento della variabilità di processo sul valore del prediction error è stato verificato ipotizzando diversi scenari, sia relativi a casi di incrementi di variabilità negativi sia a casi di incrementi di variabilità positivi. In questo modo si è potuto constatare che la variazione del prediction error aumenta all'aumentare della variabilità del processo, sebbene in quantità nettamente meno che proporzionale. Analogamente, un esito del tutto opposto si è osservato in caso di diminuzione di variabilità: tuttavia, in questo caso la riduzione influisce in misura più marcata. Le analisi hanno pertanto potuto confermare quanto inizialmente presupposto, ovvero quanto i valori inizialmente assegnati alle matrici di covarianza possono impattare sulle stime ottenute con il modello. Tuttavia, la scelta delle inizializzazioni può anche essere vista come un punto di forza del modello stesso: infatti, periodi caratterizzati da incertezza o da aspettative di repentini cambiamenti nel processo di liquidazione dei sinistri possono essere rappresentati da un legame meno stretto tra i parametri incogniti del modello, ovvero dalla scelta di una matrice di covarianza dell'errore di processo  $V(t)$  ad alta variabilità.



Analogamente, andamenti irregolari della variabile claims experience possono essere rappresentati mediante una scelta della matrice  $U(t)$  (matrice di covarianza dell'errore di misura) più instabile. Tutti questi aspetti giocano ovviamente a favore del filtro di Kalman quale possibile strumento di calcolo utilizzabile ai fini della valutazione della riserva sinistri stessa. La maggiore o minore variabilità presente nei dati ha ovviamente il suo impatto anche sul calcolo del risk margin. Sfruttando l'ipotesi di lognormalità per la distribuzione di probabilità della riserva sinistri, in linea con altri autori tra i quali England e Verral, si è potuto constatare che il filtro di Kalman è in grado di cogliere anche questo aspetto. Aumentando la variabilità nei dati del triangolo di run off si sono riscontrate differenze significative già al 75-esimo percentile. Tali differenze sono risultate nettamente più evidenti sui percentili più elevati anche per l'effetto dell'ipotesi di lognormalità utilizzata. La lognormale a due parametri ha infatti un indice di asimmetria crescente al crescere del coefficiente di variabilità. In ultima analisi è stato inoltre effettuato un confronto con il requisito di capitale, o Solvency Capital Requirement (SCR), secondo le regole previste dal QIS2 e utilizzando la best estimate ed il risk margin ottenuto attraverso l'applicazione del filtro di Kalman. Tale requisito, pur rappresentando solo una minima parte dell'eventuale requisito complessivo, in quanto calcolato sotto alcune ipotesi semplicistiche visti gli scopi perseguiti nel presente lavoro, è comunque risultato già notevolmente superiore all'attuale Margine di Solvibilità previsto dalle regole del Solvency I. L'esempio ha tuttavia considerato solamente un'ipotetica compagnia con un'unica linea di business (lob). Tale ipotesi trascura pertanto sia gli effetti della correlazione tra i rami sia gli effetti della diversificazione tra i rischi, entrambe caratteristiche che apporterebbero una riduzione del rapporto SCR/Premi per effetto della diversificazione, così come è stato dimostrato in alcuni studi ed in linea con un principio base della Teoria del Rischio il quale prevede che il Capitale a Rischio sia una funzione crescente e con concavità rivolta verso il basso al crescere del portafoglio dell'impresa.

Appare interessante concludere la tesi con alcune considerazioni che danno enfasi alle numerose possibilità di ampliare il progetto. Oltre all'eventualità di effettuare ulteriori confronti tra i modelli attualmente disponibili nella letteratura attuariale, si potrebbe pensare di estendere il modello sotto vari punti di vista. Ad esempio, si potrebbero scegliere numerose altre funzioni per rappresentare il parametro incognito oppure si potrebbero inserire parametri nuovi o nuove forme di riparametrizzazione oppure si potrebbero stabilire legami funzionali diversi da quelli utilizzati nel presente modello sia per collegare tra loro i parametri incogniti sia per collegare il vettore di stato con la misura. Questi sono ovviamente solo alcuni dei possibili esempi di estensione del modello, il quale, vista l'enorme flessibilità che lo caratterizza, lascia ovviamente spazio a molteplici altri sviluppi.

## BIBLIOGRAFIA

- “A global Framework for Insurer Solvency Assessment”, 2003 *I.A.A.*
- APRA, “Prudential Standard GPS 210. Liability Valuation for General Insurers”, *Australian Prudential Regulation Authority, luglio 2002.*
- Barnett, G., and B. Zehnwirth (1997), “Standard errors and diagnostics for link ratio techniques” *Proc. 28th Int. ASTIN Colloquium Cairns, pag. 385–412.*
- Barnett, G. and Zehnwirth, B. (2000), “Best estimates for reserves”, *CAS, Volume LXXXVII, pag.245-321.*
- Blum, K. A., and Otto, D. J. (1998), “Best estimate loss reserving: An actuarial perspective”, *CAS Forum Fall 1998, pag. 55–101.*
- Brown, R. L. (1994), “Variability of loss reserves”, *CAS Forum Spring, pag. 279–298.*
- CAS Committee on Reserve (2001), “Survey of loss reserving actuaries report”, *CAS Forum Fall, pag. 393–428.*
- CEIOPS, “Technical provision in Non-Life Insurance”, *Call for advice n.8, 2 maggio 2005.*
- CEIOPS, “Quantitative Impact Study 2. Technical Specification”, *CEIOPS-PI-08/06.*
- “Current and future solvency work in the IAIS and within the actuarial profession from a Solvency 2 point of view”, *settembre 2002.*
- Daboni, L. (1989), “Lezioni di tecnica attuariale delle assicurazioni contro i danni”, *Edizioni Lint, Trieste.*
- Daykin, C.D., Penttinen, T., Pesonen, M. (1994), “Practical Risk Theory for Actuaries”, *Chapman & Hall, London.*
- De Jong, P. and Zehnwirth, B. (1983), “Claims reserving state space models and the Kalman filter”, *Journal of the Institute of Actuaries, vol. 110, pag. 157-181.*
- De Jong, P. and Zehnwirth, B., (1983), “Credibility Theory and Kalman Filter”, *Insurance: Mathematics and Economics, vol. 2, pag. 281-286.*
- Directive 2002/13 EC for non-life insurance undertakings.

- England, P. D., and Verrall, R. J. (2001), “A flexible framework for stochastic claims reserving”, *Proc. CAS 88*, pag. 1–38.
- England, P. D., and Verrall, R. J. (2000), “Comments on: A comparison of stochastic models that reproduce chain–ladder reserve Estimates”, *Insurance Math. Econom.*, vol. 26, pag. 109–111.
- England, P. D., and Verrall, R. J. (1998), “Standard errors of prediction in claims reserving: A comparison of methods”, *General Insurance Convention & Astin Colloquium*, vol. 1, pag. 459–478.
- England, P.D. and Verrall, R.J. (2002), “Stochastic claims reserving in general insurance”, *British Actuarial Journal*, vol. 8, pag. 443-518.
- “Executive summary of the report of the working group on non-life technical provisions” settembre 2002.
- Gibson, L. (2000), “Reserving – Best estimates and ranges”, *Proceedings of the General Insurance Conerntion*, vol. 2, pag. 185–187.
- Harrison, P.J. and Stevens, C.F. (1976), “Bayesian forecasting” *Journal of the Royal Statistical Society, Series B*, vol. 38, pag. 205-247.
- Hayne, R. M. (2003), “Measurement of reserve variability”, *CAS Forum Fall*, pag. 141–171.
- Hess, K. T., and Schmidt, K. D. (2002), “A comparison of models for the chain–ladder method” *Insurance Math. Econom.* 31, pag.351–364.
- Hossack, I. B., Pollard, J. H., and Zehnwirth, B. (1999), “Introductory Statistics with Applications in General Insurance”, *Second Edition*, London New York: Cambridge University Press.
- Institute of Actuaries, “Guidance Notes 20 and 33”, *Manual of Actuarial Practice*.
- ISVAP, “Disposizioni in materia di sistema dei controlli interni e gestione dei rischi”, *Circolare n. 577/D*, 30 dicembre 2005.
- Kalman, R.E. (1960), “A new approach to linear filtering and prediction problems”, *Journal of Basic Engineering*, vol.82, pag. 340-345.
- Kreps, R. E. (2002), “Management’s best estimates of loss reserves”, *CAS Forum Fall*, pag. 247–257.
- Mack, T. (1993), “Distribution-Free Calculation of the Standard Error of Chain-Ladder Reserve Estimates”, *ASTIN Bulletin*, vol.23, pag. 213-225.

- Mack, T. (1994), “Measuring the Variability of Chain-Ladder Reserve Estimates”, *Casualty Actuarial Society, Spring Forum*.
- Mack, T. (1999), “The standard error of chain–ladder reserve estimates: Recursive calculation and inclusion of a tail factor” *ASTIN Bull.*, vol. 29, pag. 361–366.
- Mack, T. (1994), “Which Stochastic Model is Underlying the Chain-Ladder Model?” *Insurance: Mathematics and Economics*, vol.15, pag. 133-138.
- Mack, T., Venter, G. (2000), “A Comparison of Stochastic Models that Reproduce Chain-Ladder Reserve Estimates”, *Insurance: Mathematics and Economics*, vol.26, pag. 101-107.
- McCullagh P. and Nelder J.A. (1989), “Generalised Linear Models”, *Chapman and Hall, New York*.
- Moore, B., Drab, S., Christie, J., and Shah, S. (2004), “International accounting standards applied to property and casualty insurance: Overview of reserving issues”, *CAS Forum Fall*, pag. 155–205.
- Nelder, J.A. and Verrall, R.J. (1997), “Credibility theory and generalised linear Models”, *Astin Bulletin*, vol. 27, pag. 370-384.
- Neuhaus, W. (1989), “Dynamic linear models and the discrete Kalman filter”, *Statistical Memoir No. 1/89 from the University of Oslo*.
- Pollard, J.H. (1983), “Outstanding claims provisions: a distribution-free statistical approach”, *Journal of the Institute of Actuaries*, vol. 109, pag. 417-433.
- Reid, D.H. (1978), “Claim reserves in general insurance” *Journal of the Institute of Actuaries*, vol. 105, pag. 211-296.
- Reid, D.H. (1986), “Discussion of methods of claim reserving in non–life insurance” *Insurance: Mathematics and Economics*, vol. 5, pag. 45–56.
- Renshaw, A.E., Verral, R.J. (1998), “A stochastic model underlying the chain ladder technique” *British Actuarial Journal*, vol. 4, pag. 903-923.
- Sawkins, R.W. (1979), “Methods of analysing claim payments in general insurance”, *Transactions of the Institute of Actuaries of Australia*, pag. 435-519.
- Schiegl, M. (2002), “On the safety loading of chain–ladder estimates”, *ASTIN Bull.*, vol. 32, pag. 107–128.

- Schmidt, K. D., and Schnaus, A. (1996), “An extension of Mack’s model for the chain–ladder method”, *ASTIN Bull.*, vol. 26, pag. 247–262.
- “Solvency II Challenging Issues for Insurance Industries”, *Convegno 23 Novembre 2005*.
- “Solvency 2 – Presentation of the proposed work” *marzo 2001*.
- “Solvency II Working Party” *Sito Internet*.
- Straub, E. (1988), “Non–Life Insurance Mathematics”, *Berlin – Heidelberg – New York: Springer*.
- “Study into the methodologies to assess the overall financial position of an insurance undertaking from the perspective of prudential supervision”, *Kpmg Maggio 2002*.
- Taylor, G.C. (1986), “Claim Reserving in Non-life Insurance”, *North-Holland, Amsterdam*.
- Taylor, G.C. (2000), “Loss Reserving : An Actuarial Perspective”, *Kluwer Academic Publishers, London*.
- Taylor, G.C. and Ashe, F.R. (1983), “Second moments of estimates of outstanding claims” *Journal of Econometrics*, vol. 23, pag. 37-61.
- Taylor, G. C., and McGuire, G. (2004), “Loss Reserving with GLMs – A Case Study”, *CAS Discussion Paper Program*, pag. 327–391.
- “The review of the overall financial position of an insurance undertaking – Solvency 2 Review” *dicembre 1999*.
- Verrall, R.J. (2000), “An Investigation into Stochastic Claims Reserving Models and the Chain-Ladder Technique”, *Insurance: Mathematics and Economics*, vol.26, pag. 91-99.
- Verrall, R.J. (1989), “A State Space Representation of the Chain Ladder Linear Model”, *Journal of the Institute of Actuaries*, vol. 116, pag. 589-609.
- Verrall, R. J. (1994), “Statistical methods for the chain–ladder technique” *CAS Forum Spring*, pag. 393–446.
- Welch, G. and Bishop, G. (2001), “An introduction to the kalman filter”, *Technical report, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3175*.

- Wright, T.S. (1990), “A Stochastic Method for Claims Reserving in General Insurance”, *Journal of the Institute of Actuaries*, vol.117, pag. 677-732.
- Zehnwirth, B. (1989), “The chain–ladder technique: A stochastic model”, *Claims Reserving Manual*, vol. 2, London: Institute of Actuaries.
- Zehnwirth, B. (1994), “Probabilistic development factor models with applications to loss reserve variability. Prediction intervals and risk based capital”, *CAS Forum Spring*, vol. 2, pag. 447–606.