Università
Ca'Foscari
Venezia

Ca' Foscari
Dorsoduro 3246
30123 Venezia

Corso di Laurea Magistrale
in Economia e Finanza

Tesi di Laurea

La determinazione dei premi assicurativi nel ramo danni: dalla teoria della credibilità ai sistemi Bonus-Malus

Relatore
Ch.ma Prof.ssa Antonella Basso

Laureando
Giorgia Perissinotto
Matricola 827232

Anno Accademico
2013/2014

A Mamma e Papà,

la mia forza.

A Lorenzo,

il mio amore.

"Devo pur sopportare qualche bruco

se voglio conoscere le farfalle,

sembra che siano così belle"

Il piccolo principe

SOMMARIO

INTRODUZIONE

Spinto da un innato bisogno di sicurezza, l'uomo comune ha da sempre cercato strumenti che fossero in grado di proteggerlo dall'aleatorietà degli eventi. Non a caso le prime forme assicurative vengono fatte risalire addirittura gli Egizi, nel 2700 a.C., presso i quali si costituiva un fondo per le spese funebri di coloro che avevano il compito di tagliare le pietre[1]. Tale forma di tutela non rimase legata a quegli anni, e soprattutto a quella tipologia di rischio, ma vide il suo maggior sviluppo con l'insediamento di nuove culture, le quali incentravano la loro vita sulle attività di mare.

È chiaro quindi che lo sviluppo di nuove forme assicurative è semplicemente stato dettato dalle necessità via via riscontrata nell'affrontare nuove situazioni, dalla necessità di tutelarsi verso eventi che potrebbero mettere a rischio, o comunque compromettere, la persona o beni di proprietà della stessa. Proprio per questo motivo, ai giorni nostri, si possono trovare soluzioni per ogni tipo di esigenza, in grado quindi di andare ed offrire protezione a tutti i soggetti: basterà pagare un prezzo, che poi definiremo premio, per trasferire i nostri rischi ad un ente esterno, ovvero la Compagnia di Assicurazione.

Ma quale sarà il giusto prezzo da pagare?

La tesi si propone quindi di andare a rispondere a tale domanda. Di conseguenza si cercherà, non solo di chiarire quali siano i diversi elementi da considerare per la formulazione di un premio, ma anche di analizzare quali siano i differenti studi da dover effettuare perché una compagnia Assicurativa sia in grado di offrire diverse coperture assicurative avendo un profitto.

[1] Il mestiere del tagliapietre è paragonabile all'attuale muratore, uno dei mestieri maggiormente praticati all'epoca.

Il capitolo 1 farà una panoramica generale sui contratti di assicurazione, andando a delineare brevemente le diverse tipologie di premi: dal semplice premio equo di equilibrio, al premio di tariffa. L'attenzione si sposterà poi sulle diverse funzioni di risarcimento, elemento essenziale nel contratto di assicurazione, che andrà di volta in volta definito in base al rischio dal quale il soggetto intende proteggersi.
Nel capitolo 2 si entra nella sostanza della trattazione, iniziando a delineare gli strumenti base per il calcolo dei tre premi principali, ovvero: il premio equo, il premio puro, ed il premio di tariffa. Per tale analisi si farà spesso riferimento alle forme distributive dei dati, quindi, per facilitarne la comprensione nei capitoli seguenti, si è preferito introdurre anche alcuni modelli probabilistici adeguati per studiare il numero di sinistri ed il danno provocato dal sinistro.
Il capitolo 3 è dedicato alla tariffazione a priori, e quindi il primo approccio che una compagnia assicurativa ha con i dati rilevati sul portafoglio di assicurati. Grazie anche all'utilizzo di esempi al fine di chiarire il concetto espresso, si cerca di far comprendere al lettore non solo l'utilità, ma anche la natura dei dati raccolti, facendo spesso riferimento all'assicurazione Responsabilità Civile Autoveicoli (in breve RCA).
Nel capitolo 4, invece, la teoria della credibilità ha un ruolo di notevole importanza in quando elemento di congiunzione tra un premio privo di personalizzazione, ottenuto dalla tariffazione a priori, ed un premio individuale che va a riflettere le caratteristiche dominanti dei diversi assicurati.
Il concetto di credibilità non è così univoco, infatti negli anni esso ha portato ad un proliferare di sistemi di trattazione dei dati rilevati a posteriori, sistemi che possiamo ricondurre a due principali approcci: *limited fluctuation credibility* e *greatest accuracy credibility*. Nonostante questi abbiano avuto un enorme successo negli anni, attualmente ad essi vengono preferiti i sistemi Bonus-Malus, i quali in parte sono riconducili ai più datati approcci di credibilità, ma allo stesso tempo presentano regole evolutive in

grado di rispondere meglio alle diverse esigenze, così come presentato nel capitolo 5. Sempre in quest'ultimo capitolo si dà l'opportunità al lettore di osservare l'utilità delle formule finora esposte, semplicemente attraverso l'introduzione delle tabelle attualmente utilizzate da diverse compagnie per effettuare un preventivo RCA. Il capitolo si conclude poi con un confronto fra il sistema Bonus-Malus attualmente in vigore nello Stato Italiano ed i sistemi presenti in altri paesi, utilizzando come metodo di paragone alcuni indicatori standard.

Nel capitolo 6 troviamo la conclusione della tesi, ovvero l'analisi del processo di tariffazione normalmente utilizzato per stimare il numero di sinistri[2]. Grazie a due modelli differenti, scelti appositamente per la diversità dei dati trattati, si cercherà di chiarire l'utilità della tariffazione in generale, per poi spostare l'attenzione sull'importanza del numero di variabili tariffarie considerate.

---

[2] Colgo qui l'occasione per ringraziare l'Agenzia Generali di Jesolo per la collaborazione dimostrata.

1. I CONTRATTI ASSICURATIVI

1.1. DEFINIZIONI

La definizione di contratto di assicurazione viene fornita direttamente dall'articolo 1882 del codice civile, ovvero "il contratto col quale l'assicuratore, verso pagamento di un premio, si obbliga a rivalere l'assicurato, entro i limiti convenuti, del danno ad esso prodotto da un sinistro (assicurazione contro i danni), ovvero a pagare un capitale o una rendita al verificarsi di un evento attinente alla vita umana (assicurazione sulla vita)".

Una prima distinzione, doverosa da effettuare, sono gli ambiti in cui operano le assicurazioni: un soggetto ha la facoltà, e talvolta l'obbligo[3], di proteggere il valore di un bene piuttosto che ridurre al minimo la propria esposizione monetaria al rischio, come garantirsi un risarcimento qualora in futuro si verifichi un determinato evento; il nostro studio va a focalizzarsi comunque sulle assicurazioni che trattano le prime caratteristiche citate, ovvero quelle che compongono il ramo danni.

I rischi che queste polizze coprono, e quindi contro cui è possibile proteggersi, sono numerosi e, come elenca il decreto legislativo 17 marzo 1995 n. 175, vanno a trattare veicoli e responsabilità civile piuttosto che perdite pecuniarie ed infortuni, tutti accomunati comunque dal fatto che alla certezza dell'importo pagato dal soggetto intenzionato ad assicurarsi, si va a contrapporre l'aleatorietà della prestazione da parte dell'assicuratore (PARRINI, 2012).

Il contratto, denominato polizza, prevede che a fronte di un pagamento certo, ovvero il *premio*, da parte del contraente, vi sia l'impegno aleatorio

[3] Si fa riferimento alle polizze Rc auto e alle polizze vita legate ad esempio alla stipula di un mutuo.

dell'assicuratore a pagare l'intero risarcimento per tutti i danni prodotti dai sinistri avvenuti durante il periodo di copertura della polizza. Il pagamento di tale polizza può avvenire tramite frazionamento annuale o semestrale, prevedere una clausola di tacito rinnovo oppure no, ma quel che è certo è che l'assicuratore sarà obbligato a rispondere ogni qualvolta il rischio assicurato sarà colpito da sinistro (GIGANTE, PICECH, & SIGALOTTI, 2010).

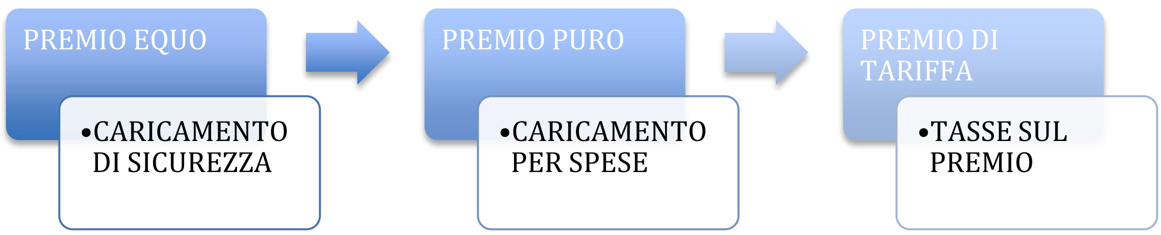
Parlare semplicemente di premio però è troppo generale, nonostante da un punto di vista commerciale spesso si faccia riferimento solamente a quanto viene fatto pagare al futuro assicurato; dietro tale definizione ci sono dei concetti e dei calcoli molto più dettagliati. Il primo premio di cui sarebbe necessario parlare, infatti, è quello che va a costituire la base per rendere possibile l'esistenza stessa di un'assicurazione, ovvero, il premio equo (P_e): definito "equo" poiché rappresenta il valore attuale di quanto l'assicuratore si impegna a pagare al verificarsi dell'evento dannoso futuro ed incerto. Proprio per "l'equità" che lo contraddistingue, tale premio non permette alla compagnia di trarne un minimo vantaggio, quindi per far fronte al totale dei risarcimenti e mantenere un equilibrio tecnico della gestione si rende necessario sommare a tale cifra un importo definito caricamento di sicurezza, che potrà essere calcolato attraverso criteri come quello dell'utilità attesa o della probabilità di rovina: questo tipo di premio viene definito netto o puro[4] (DABONI, 1989).

Il premio di tariffa, ovvero il premio realmente pagato dall'assicurato, include inoltre ulteriori spese di acquisizione e gestione del contratto (oltre alle eventuali spese per la gestione dei sinistri futuri), che vengono raggruppate sotto il nome "caricamento per spese" e permettono di raggiungere alla compagnia assicurativa l'equilibrio anche dal punto di vista economico. Inoltre a quest'ultimo premio, che è l'unico del quale il

[4] Il caricamento di sicurezza inoltre permettere la formazione di riserve patrimoniale utili a garantire il pagamento degli eventi assicurati.

cliente effettivamente conosce il valore, spesso vengono aggiunte anche le tasse.[5]

Figura 1.1: Il passaggio dal premio equo al premio di tariffa



Il risarcimento totale viene normalmente indicato con il simbolo X, e sta a rappresentare il rischio che ricade sulla compagnia assicurativa di dover rimborsare un determinato ammontare in funzione del numero N di sinistri che verranno a verificarsi. Non è ovviamente possibile stabilire a priori il numero di incidenti che si verificheranno durante il periodo di copertura, e quindi questo può assumere valori che vanno teoricamente da 0 ad infinito.

Al verificarsi di un sinistro, questo provocherà un danno di importo Z_i (con

[5] Per poter valutare se una compagnia è particolarmente costosa piuttosto che un'altra bisogna confrontare il caricamento per spese perché è l'unica parte in cui c'è l'effettiva possibilità di creare un margine di guadagno. Tra le diverse assicurazioni possono presentarsi differenze anche sul caricamento di sicurezza ma questo non è imputabile ad un tentativo di profitto ma semplici differenze nella gestione piuttosto che nella composizione del portafoglio clienti.

$i = 1,2, \dots, n$), danno che sensatamente sarà pari a 0 nel momento in cui non si verifichi alcun sinistro. L'indennizzo Y_i che andrà a versare l'assicuratore sarà direttamente fun zione del danno provocato: $Y_i = \rho(Z_i)$[6], dove a sua volta il risarcimento totale sarà dato dal totale degli importi aleatori versati per ogni singolo sinistro avvenuto

$$X = \sum_{i=1}^{N} Y_i$$

1.2. LE FUNZIONI DI RISARCIMENTO

I contratti assicurativi prevedono, attraverso numerose clausole, la possibilità di scegliere la funzione di risarcimento ρ che più risponde alle esigenze del futuro assicurato, ed in questo modo è possibile individuare le più frequenti forme assicurative:

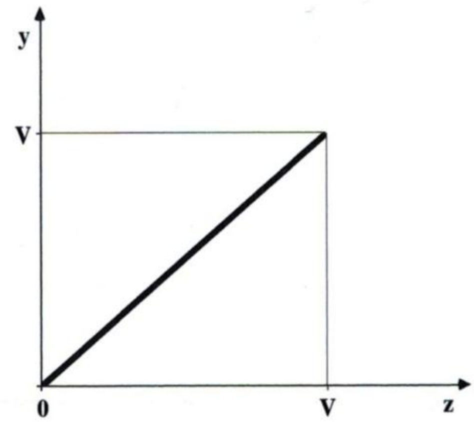
- <u>L'assicurazione a valore intero</u> è caratterizzata da una funzione di risarcimento $Y = Z$, tipicamente utilizzata quando si vuole garantire il bene contro eventi come il furto e l'incendio, assume invece il nome di assicurazione a garanzia illimitata nei casi in cui si tratti di responsabilità civile. Per poter sottoscrivere questa tipologia di assicurazione è importante che il futuro assicurato sia in grado di rilevare a priori il valore del bene V da assicurare e quindi renderlo noto alla compagnia.

---

[6] In tale funzione si dovrà sempre rispettare la regola $Y \leq Z$ per il principio di non arricchimento: lo scopo di un'assicurazione è riportare il soggetto alla condizione in cui si trovava prima che si verificasse il sinistro, quindi garantire che il fatto avvenuto non vada a gravare sulle sue spese; di conseguenza non ci si può trovare nella condizione per cui il risarcimento superi il danno effettivo, altrimenti il sottoscrittore trarrebbe vantaggio ogni qualvolta fosse vittima di un incidente.

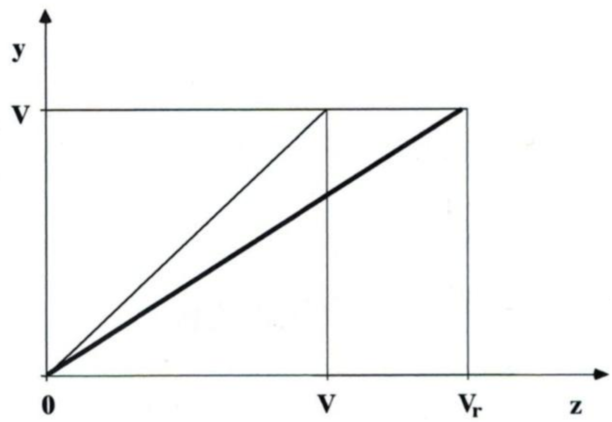
Nel caso in cui si verifichi il sinistro ed il valore del bene rilevato V_r risulti superiore a quello dichiarato, si avrà un caso di sottoassicurazione, ed il risarcimento sarà $Y = \frac{V}{V_r} Z$

Figura 1.2: Assicurazione a valore intero



Fonte: GIGANTE, PICECH, & SIGALOTTI (2010)

Figura 1.3: Assicurazione a valore intero nel caso in cui il valore rilevato superi quello dichiarato



Fonte: GIGANTE, PICECH, & SIGALOTTI (2010)

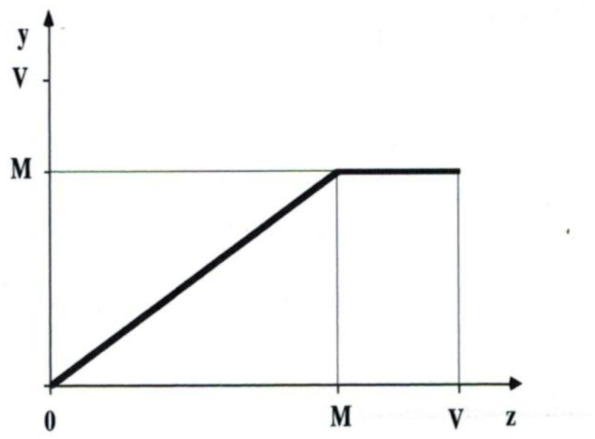
- Anche <u>l'assicurazione a primo rischio relativo</u> si fonda sul concetto del valore V del bene assicurato, ma viene introdotto anche un secondo

valore M che indica il massimale di copertura[7]. L'effettivo risarcimento in ogni caso non supererà mai il massimale, ma al momento della stipula del contratto dovranno comunque essere indicati entrambi i valori.
In questo caso la funzione di risarcimento assume la forma

$$Y = \begin{cases} Z, & Z \leq M \\ M, & Z > M \end{cases} = \min(Z, M)$$

Come nel caso precedente, se il danno effettivamente periziato supera la soglia V indicata nella polizza, si è in presenza di sottoassicurazione; il risarcimento dovrà quindi seguire la regola proporzionale $Y = \min\left(\frac{V}{V_r} Z, M\right)$.

Figura 1.4: Assicurazione a primo rischio relativo

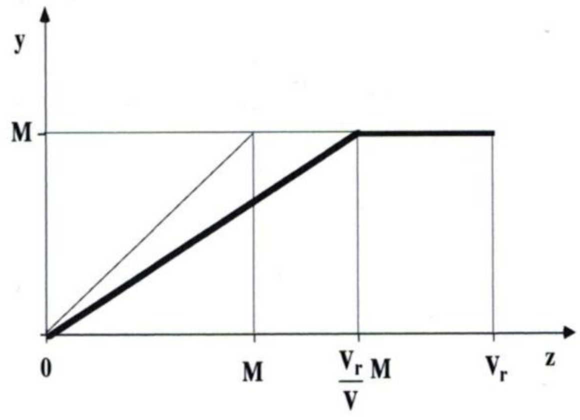


Fonte: GIGANTE, PICECH, & SIGALOTTI (2010)

[7] In realtà il valore M può in questo caso essere rappresentato dal massimo danno probabile (indicato con MPL, dall'inglese *maximum probable loss*), ovvero l'estremo superiore degli importi ai quali viene attribuita una probabilità positiva di verificarsi

$$MPL = sup\{z \in \Re: \Pr(Z = z) > 0\}$$

Figura 1.5: Assicurazione a primo rischio relativo nel caso di sottoassicurazione



Fonte: GIGANTE, PICECH, & SIGALOTTI (2010)

- Per <u>l'assicurazione a primo rischio assoluto</u> non è più corretto parlare di valore del bene assicurato, ma di valore che il soggetto intende assicurare; tramite questa clausola, infatti, l'indennizzo che dovrà corrispondere l'assicurazione sarà al massimo il valore M stabilito contrattualmente. È evidente che il sottoscrittore ha un ruolo chiave poiché dovrà stimare l'ammontare di denaro che sarà autonomamente in grado di coprire in caso di sinistro, con il vantaggio comunque di non trovarsi mai nella situazione di sottoassicurazione o di ottenere un risarcimento che segua delle regole proporzionali[8].
  Ciò equivale ad un risarcimento
  $$Y = \begin{cases} Z, & Z \leq M \\ M, & Z > M \end{cases} = \min(Z, M)$$

- Quando si parla di <u>assicurazioni con franchigia</u>[9] bisogna distinguere la franchigia relativa da quella assoluta: nel primo caso il risarcimento sarà totale se il danno arrecato supera una determinata soglia d

---

[8] Per ulteriori approfondimenti si consiglia di consultare il sito www.amicoassicuratore.it.

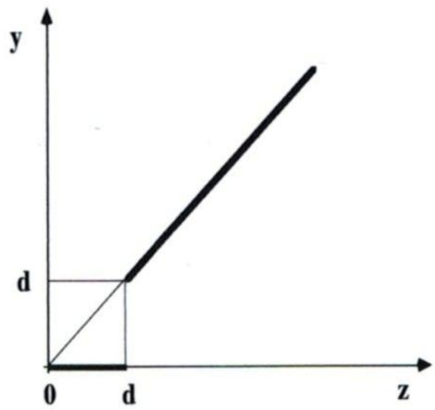
[9] La franchigia viene introdotta per un duplice scopo: da un lato rende l'assicurato più responsabile, evitare quindi il *moral hazard* a favore di una guida più sicura; dall'altro lato

$$Y = \begin{cases} 0, & se\ Z \leq d \\ Z, & se\ \ Z > d \end{cases}$$

in presenza di franchigia assoluta, invece, l'ammontare di danno che non eccede la soglia rimarrà sempre a carico dell'assicurato, mentre la compagnia andrà a risarcire esclusivamente l'eccedenza

$$Y = \begin{cases} 0, & se\ Z \leq d \\ Z - d, & se\ Z > d \end{cases} = \max{(0, Z - d)}$$

Figura 1.6: Assicurazione con franchigia relativa



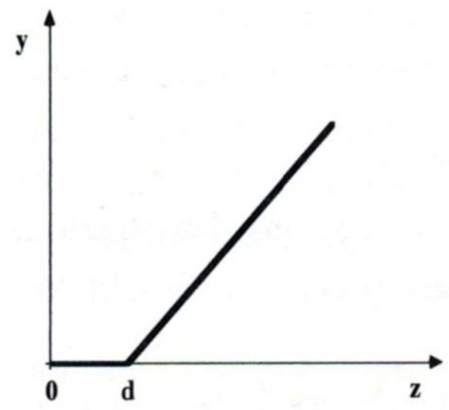
Fonte: GIGANTE, PICECH, & SIGALOTTI (2010)

---

consente di pagare un premio molto più contenuto, dato che l'assicurazione non dovrà più intervenire al di sotto della soglia stabilita (spesso se si tratta di franchigia assoluta viene espressa in % sul danno, mentre nei casi di franchigia relativa viene espressa come quantitativo in denaro).

Oltre alla franchigia di valore, sopra citata, è doveroso citare anche un'altra tipologia di franchigia, definita di valore; questa prevede una clausola per cui per un determinato arco temporale, a partire dalla data di sottoscrizione del contratto fino ad una data successiva indicata, la polizza non risulta attiva e di conseguenza non risponderà qualora si verifichino dei sinistri. Questa può essere vista come una tutela da parte della compagnia contro coloro che vanno stipulare una polizza solo dopo aver subito il danno, cercando così di limitare l'esborso monetario (PARRINI, 2012).

Figura 1.7: Assicurazione con franchigia assoluta

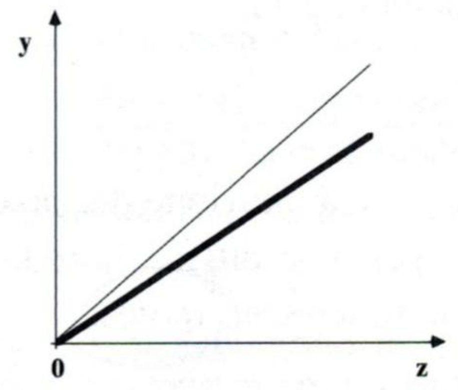


Fonte: GIGANTE, PICECH, & SIGALOTTI (2010)

- L'ultima tipologia trattata riguarda le assicurazioni con scoperto: normalmente è utilizzata nelle coperture contro il rischio d'insolvenza dei crediti commerciali o dove si teme che il sinistro avvenga quasi per volontà dell'assicurato stesso che potrebbe trarne dei vantaggi. Questa ruota attorno al valore di un'aliquota ξ (con $0 < \xi < 1$), che corrisponde a quella parte di risarcimento che rimane a carico solo ed esclusivamente dell'assicurato, mentre la parte risarcita dalla compagnia sarà $Y = (1 - \xi)Z$. [10]

[10] Unica tipologia di assicurazione dove è corretto parlare di scoperto "consapevole".

Figura 1.8: Assicurazioni con scoperto



Fonte: GIGANTE, PICECH, & SIGALOTTI (2010)

In realtà nulla vieta alle diverse compagnie di proporre una funzione di risarcimento che risulti dalla combinazione di più tipologie sopra indicate, ad esempio proponendo di stipulare una polizza in cui è presente un massimale oltre ad una franchigia (PITACCO, 2000).

Nonostante risulti evidente che le possibilità nel campo assicurativo siano svariate, ciò che viene offerto con maggior frequenza ruota attorno a tre forme: primo rischio assoluto, primo rischio relativo e valore intero. Come accennato in precedenza, la peculiarità della prima di queste tre forme è quella di non sfociare mai in un caso di sottoassicurazione, scenario possibile invece per le altre due forme. Questo viene spiegato dalla regola proporzionale enunciata dall'articolo 1907 del codice civile: "*se l'assicurazione copre solo una parte del valore che la cosa assicurata aveva nel tempo del sinistro, l'assicuratore risponde dei danni in proporzione della parte suddetta, a meno che sia diversamente convenuto*[11]*";* il legislatore ha quindi stabilito che nelle due forme che prevedono di indicare nel contratto il valore del bene assicurato (e non solamente quanto di esso vorrò andare ad assicurare), se al momento del sinistro si rileva un valore diverso da

[11] Con quest'ultima espressione si fa riferimento alle possibili clausole contenute nel contratto, quindi eventuali deroghe al principio di proporzionalità enunciato.

quest'ultimo, la compagnia assicurativa avrà la facoltà di indennizzare i soggetti tenendo conto di questo: quindi, se al momento della stipula il valore indicato è in realtà 20% in meno dell'effettivo valore, anche l'assicuratore potrà predisporre un pagamento pari al 20% in meno dell'effettivo danno subito[12].

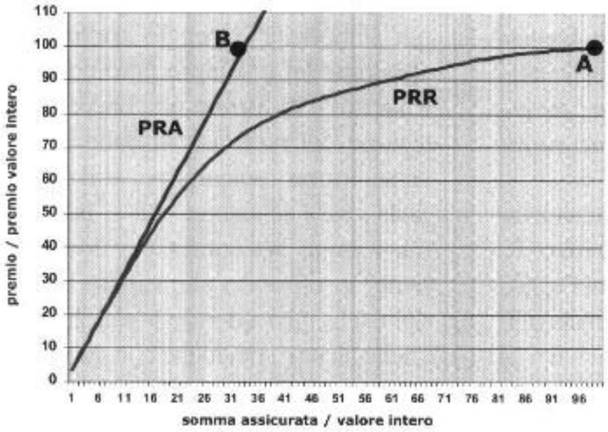
Fermo restando che l'obiettivo di ogni soggetto è quello di ottimizzare la sua copertura assicurativa, ovvero ottenere la massima copertura ad un prezzo minore possibile, la forma assicurativa favorita è quella a primo rischio assoluto. Il vantaggio che essa offre è quello di poter decidere la somma da assicurare senza dover mai comunicare alla compagnia qual è il reale ammontare del bene, quindi una maggior flessibilità rispetto alla forma a valore intero che obbliga ad assicurare l'intero valore, compreso il rischio improbabile[13]; d'altra parte questo vantaggio ha pur sempre un costo per la compagnia, perché non essendo a conoscenza di tale informazione va a tutelarsi imponendo un premio che è pari a circa tre volte quello pagato per una polizza a valore intero. Come si evince dal grafico seguente, sarà conveniente riporre la scelta in questa forma (dove la dicitura PRA sta per primo rischio assoluto) solo quando la parte assicurata sarà massimo pari al 33% dell'intero valore, poiché superata tale soglia si avrà una situazione antieconomica tale da favorire l'assicurazione a valore intero (indicato col punto B).

---

[12] In questo caso risulta chiara la volontà di andare a "punire" i soggetti che per pagare un premio inferiore vanno ad indicare dei valori che non rispecchiano la realtà.

[13] L'utilizzo della parola improbabile rappresenta perfettamente lo svantaggio che risulta da questa forma: nel caso di una polizza furto, sebbene una persona abbia all'interno della propria abitazione beni per un ammontare di 40000€, difficilmente vuole assicurarli tutti, perché è possibile ma sicuramente poco probabile che gli vengano rubati tutti i beni, compresi quelli che materialmente fanno parte dell'arredamento.

Figura 1.9: Confronto fra le tre forme assicurative principali



Fonte: LOMAZZI (2005)

Sempre lo stesso grafico suggerisce di utilizzare una seconda forma quando il valore che si vuole assicurare superi 1/3 del valore intero del bene: la forma a primo rischio relativo (indicata con PRR).

In questo caso si ripresenta la necessità di verificare il valore totale del bene oggetto del contratto, costringendo quindi il soggetto ad ottenere un risarcimento che sarà massimo pari a quanto esso ha stabilito, ma con un minimo che potrà seguire la regola proporzionale, e di conseguenza non definibile a priori; talvolta la certezza non ha prezzo, ed è proprio per questo motivo che difficilmente chi sceglie di trasferire il proprio rischio ad una compagnia, se ne accolli comunque una parte scegliendo la forma a primo rischio assoluto. (LOMAZZI, 2005)

2. LA TARIFFAZIONE NEL RAMO DANNI

2.1. LA PRATICA DI TARIFFAZIONE

Il premio richiesto ad un assicurato deriva proprio dalla tariffazione, pratica che consiste nell'assegnare la giusta distribuzione di probabilità al numero di sinistri ed al singolo danno causato da ognuno di questi, così da ottenere la base tecnica utile ai fini del calcolo del premio.

2.1.1. IL CALCOLO DEL PREMIO EQUO

Dopo aver definito N il numero di sinistri, ed indicato con Y_i, dato $Y_i = \varphi(Z_i)$, l'entità dell' i-esimo sinistro, si può passare a valutare il risarcimento X

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

in un'ottica probabilistica.

Avendo a disposizione tali dati, ovvero quelli che appunto vanno a costituire la base tecnica, è possibile costruire un modello che permetta di fornire una distribuzione di probabilità anche per il risarcimento totale X . Perché ciò sia comunque possibile è fondamentale imporre due ipotesi:

a) Stazionarietà: tutti i risarcimenti Y_i considerati sono identicamente distribuiti[14];
b) Indipendenza (condizionata): le variabili casuali Y_i (con $i = 1,2, \dots, N$) sono indipendenti (BUHLMAN & GISLER, 2005).

[14] L'assunzione fatta permette di stabilire una relazione tra il passato ed il futuro, e nonostante sia un'assunzione forte è comunque raggiungibile aggiustando fattori come l'inflazione (BUHLMAN & GISLER, 2005).

Indicando con F_Z la funzione di ripartizione del danno per sinistro e con F_Y la funzione di ripartizione del risarcimento per il singolo sinistro, e tenendo in considerazione le due ipotesi sopra citate, è possibile interpretare la distribuzione di probabilità di $X = \sum_{i=1}^{N} Y_i$ come una distribuzione composta. Per l'ipotesi di stazionarietà si suppone che i risarcimenti Y_i abbiano tutti la stessa distribuzione, quindi è corretto scrivere $E(Y_i) = E(Y) \;\; per\ i = 1,2, \dots, N$; indicando poi con p_n la probabilità che N sia uguale ad n, si dimostra che:

$$E(X) = \sum_{n=0}^{+\infty} p_n\, E(S|N=n) = \sum_{n=1}^{+\infty} p_n E(S|N=n) = \sum_{n=1}^{+\infty} p_n E(Y_1 + Y_2 + \ldots + Y_N|N=n)$$
$$= \sum_{n=1}^{+\infty} p_n \sum_{j=1}^{+\infty} E(Y_j|N=n)$$

Per l'ipotesi d'indipendenza il valor medio dei risarcimenti non sarà influenzato dal numero di sinistri, quindi $E\left(Y_j \middle| N=n\right) = E(Y_i)$; tenendo in considerazione questo, oltre all'ipotesi di stazionarietà per cui $E(Y_i) = E(Y)$, si può risolvere l'equazione in questo modo

$$\sum_{n=1}^{+\infty} p_n \sum_{j=1}^{n} E(Y_j|N=n) = \sum_{n=1}^{+\infty} p_n n E(Y) = E(Y) \sum_{n=1}^{+\infty} p_n n = E(Y)E(N)$$

Compattando le formule otteniamo quindi il concetto elementare da cui deriva tutto ciò che riguarda il *pricing* assicurativo: il premio equo, che è pari al risarcimento totale atteso, è dato dal prodotto del numero atteso di sinistri per il risarcimento di ogni singolo sinistro, ovviamente nell'ipotesi che il sinistro si verifichi (PITACCO, 2000).

$$E(X) = E(N)E(Y)$$

2.1.2. IL CALCOLO DEL PREMIO PURO

Il premio equo però non corrisponde a quanto effettivamente viene fatto pagare ai soggetti assicurati; affinché si possa avere un equilibrio finanziario tra la compagnia e i diversi soggetti è necessario sommare dei caricamenti di sicurezza, detti anche caricamenti impliciti, che permettono a colui che andrà ad accollarsi il rischio di limitare la sua perdita ed al tempo stesso di avere un guadagno.

Per procedere al calcolo del premio netto ci si affida ai principi di calcolo più conosciuti, ma per poter meglio orientare la scelta dei diversi studiosi sono state individuate delle proprietà che permettono di preferire un metodo piuttosto che un altro, e queste sono:

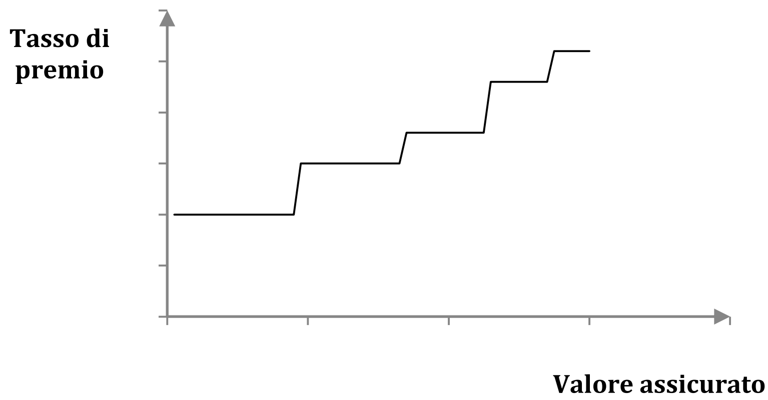
1) Additività. Indicato con $\Pi(X)$ il premio corrispondente per assicurarsi contro il rischio X, e con $\Pi(Y)$ il costo che si dovrà andare a sostenere per coprirsi dal richio Y, questa proprietà potrebbe essere riassunta attraverso l'espressione $\Pi(X+Y) \leq \Pi(X) + \Pi(Y)$; ovvero si prevede che sia vantaggioso sottoscrivere un'unica polizza che al suo interno abbia più garanzie, e che questo sia conveniente al punto da pagare un prezzo minore o al massimo uguale alla somma che pagherei se per avere tutte le garanzie fosse necessario sottoscrivere più polizze. Proprietà che peraltro impedisce di fare profitti a rischio zero attraverso operazioni di arbitraggio all'interno del mercato[15] (BORCH K. H., 1990).
2) Omogeneità d'importo. Assumendo che ci si voglia assicurare contro i rischi X ed Y, ci si attende che all'aumentare delle coperture presenti in una stessa polizza, ci sia un aumento anche del premio che si dovrà andare a pagare. Quindi, sulla base di questa proprietà, il premio da far

[15] Normalmente questo tipo di operazione viene riassunta attraverso l'espressione "Ther ain't no free lunch", semplicemente non si può pretendere di guadagnare su arbitraggi privi di rischio come non si può mangiare senza pagare.

pagare ai futuri assicurati dovrà crescere proporzionalmente con il massimo valore risarcibile; proporzionalità che metterebbe in difficoltà l'assicuratore qualora ritenesse necessario applicare un caricamento maggiore nel caso si trovasse a garantire rischi molto probabili per importi elevati.

Per ovviare a questo problema, nell'ambito assicurativo si sceglie di adottare un tasso di premio che viene chiamato "costante a tratti", e che si distingue per il che fatto rimane costante sui diversi intervalli, presentando così un'omogeneità localizzata esclusivamente in determinati punti (PITACCO, 2000).

Figura 2.1: Tasso di premio su intervalli



Fonte: PITACCO (2000)

Sul fronte opposto si posizionano invece le compagnie che non vedono questa proprietà come un ostacolo, anzi la sfruttano per cercare di proporre dei prodotti alternativi, creando delle polizze che siano in grado di coprire diverse problematiche (pensiamo alle polizze denominate usualmente "casa" che includono non solo furto e incendio, ma anche responsabilità civile), ad un prezzo così contenuto da risultare nettamente inferiore rispetto a quelle delle compagnie concorrenti, il tutto con lo scopo di provocare una migrazione verso il proprio portafoglio clienti.

3) Positività del caricamento di sicurezza. Affinchè vi sia un equilibrio economico, e quindi la possibilità di guadagno per la compagnia assicurativa, il premio puro dovrà necessariamente essere superiore a quello che è il premio equo; quindi rispettivamente $\Pi(X) \geq E(X)$.

   Questa proprietà merita però un'ulteriore riflessione; molte volte una compagnia può avere la convenienza a far pagare un prezzo che sia quasi svantaggioso per la compagnia stessa solo per questioni di marketing o per la necessità di avere nel portafoglio una tipologia di clienti che in altro modo non riuscirebbe ad ottenere[16]. È proprio il mercato stesso che quindi pone dei limiti, degli ostacoli o puri e semplici artifici che permettono di violare la norma comune per poter essere efficienti sul mercato e quindi non badare a quelle che sarebbero le regole da rispettare, al fine di permettere alla compagnia assicurativa di stare a galla ed arricchire il portafoglio di nuovi sottoscrittori (SUNDT, 1984).
4) L'ultima proprietà, che agli occhi del futuro assicurato è praticamente ovvia, prevede che il premio pagato sia necessariamente inferiore a quello che è il valore del bene assicurato, altrimenti non ci sarebbe nessuna convenienza ad assicurarsi (PITACCO, 2000).

2.1.2.1. IL PRINCIPIO DEL VALORE ATTESO O DELLA SPERANZA MATEMATICA

Se gli attuari preferissero la semplicità nei calcoli, sicuramente preferirebbero questo metodo poiché prevede l'utilizzo di un unico parametro della distribuzione del rischio: il valore atteso; il valore calcolato, in sostanza, assume il ruolo di caricamento di sicurezza e quindi diventa parte dell'espressione seguente:

$$\Pi = (1 + \alpha)E(X) \quad con\, \alpha > 0$$

[16] Si fa riferimento a quella che è la diversificazione di portafoglio.

Nonostante l'immediatezza dei calcoli, non potrà mai risultare un principio di calcolo affidabile, perché farebbe pagare lo stesso premio a tutti i soggetti ai quali corrispondono rischi che presentano lo stesso valore atteso. È facile comprendere che sono molto più pericolosi i rischi che hanno un'elevata varianza, rispetto a quelli che a parità di valore atteso hanno invece una minor varianza.
A prescindere da quanto esposto, non si può sostenere che sia uno dei peggiori principi di calcolo, infatti, riflette quasi tutte le proprietà richieste: l'unica nota negativa è quella di non riuscire a far sempre pagare un premio inferiore rispetto al valore del bene assicurato (SUNDT, 1984).

2.1.2.2. IL PRINCIPIO DELLA VARIANZA

Proprio per colmare le lacune create dal principio di calcolo precedente, si è introdotto questo metodo, che prevede di aggiungere al premio equo un caricamento di sicurezza che sia in grado di riflettere la rischiosità, misurata per l'appunto con la varianza del risarcimento totale $var(X)$; in questo modo nella formula non comparirà solo il valore atteso, ma anche la rischiosità espressa, appunto, dalla dispersione intorno alla media:

$$\Pi = E(X) + \lambda\, var(X) \qquad con\, \lambda > 0$$

Nonostante questo principio dimostri una maggior flessibilità e adeguatezza rispetto al principio esposto precedentemente, l'unica proprietà che è in grado di rispettare è quella della positività del caricamento di sicurezza[17] (SUNDT, 1984). Solo nel caso in cui si vogliano assicurare due differenti rischi X ed Y non correlati, o meglio ancora stocasticamente indipendenti, si potrà attribuire anche la proprietà additiva (DABONI, 1989).

[17] In merito alle proprietà presenti in questo principio di calcolo le opinioni sono discordanti fra i diversi autori. Si veda per il confronto anche PITACCO (2000), DABONI (1989), e BORCH (1990).

2.1.2.3. IL PRINCIPIO DELLO SCARTO QUADRATICO MEDIO

Come suggerisce il nome stesso, si differenzia del principio precedente per l'utilizzo nella formula dello scarto quadratico medio, quindi $\sigma = \sqrt{var(x)}$, come caricamento di sicurezza nella formula

$$\Pi = E(X) + \beta\sigma(X) \quad con\, \beta > 0$$

Nonostante la distinzione dal principio della varianza sia limitata a tale elemento, utilizzando questo metodo il premio netto potrà possedere le proprietà che in precedenza abbiamo indicato come additività del premio e positività del caricamento di sicurezza (SUNDT, 1984).

2.1.2.4. IL PRINCIPIO DELL'UTILITA' ATTESA

Questo tipo di approccio si distingue dai precedenti, infatti solo in questo caso per ottenere il premio è necessario risolvere l'equazione $E[u(\Pi - X)] = 0$, ovvero si cercherà, data una funzione u di utilità, un premio che soddisfi tale eguaglianza.

Provando a sviluppare l'equazione nel modo seguente, si dimostra (CASTELLANI, 2011) che:

$$E[u(\Pi - X)] - u(0) = \Pi - \mathrm{E(X)} - \frac{a}{2}E[(\Pi - \mathrm{X})^2] = \Pi - \frac{a}{2}[\Pi^2 - 2\Pi\mathrm{E(X)} +$$

$$\mathrm{EX2} = -a2\Pi 2 + \Pi\alpha\mathrm{EX} + 1 - \mathrm{EX} - a2varX + E2X = 0,$$

quindi esplicitando il premio netto si ottiene:

$$\Pi = \mathrm{E(X)} + \frac{1}{a}\left(1 - \sqrt{1 - a^2 var(X)}\right) \cong E(X) + \frac{a}{2}var(X)$$

Basta osservare il secondo elemento dell'equazione per capire chiaramente che nonostante il calcolo si strutturi in maniera differente, in modo approssimato il valore ottenuto è proporzionale a quanto si è visto per il metodo di calcolo che prevede l'utilizzo della varianza; proprio per questo motivo esso si caratterizza per la medesima proprietà: è possibile ottenere un premio che non sia equivalente al premio equo, ovvero che includa un caricamento di sicurezza (DABONI, 1989).

Prima di concludere l'argomento è importante osservare che, analizzati tutti e 4 i maggiori principi di calcolo utilizzati, nessuno di essi è in grado di garantire la quarta proprietà, ovvero che l'importo del premio sia sempre inferiore al valore del rischio stesso. Nonostante ciò appaia strano, non crea grossi problemi, così come pure il fatto di non rispettare le altre proprietà è poi così vincolante: questo perché talvolta sono gli stessi assicuratori a scegliere di non offrire più garanzie all'interno dello stesso pacchetto (probabilmente per strategie di marketing) ma nulla vieta che i premi così calcolati vengano offerti al pubblico sottoscrittore.

2.1.3. IL CALCOLO DEL PREMIO DI TARIFFA

La compagnia assicurativa, giunta a questo punto, difficilmente propone i premi calcolati attraverso i principi sopra esposti al pubblico di futuri assicurati, perché si riserva la facoltà di arrotondare la somma con degli ultimi caricamenti (espliciti) per la copertura di tutte le spese che essa dovrà comunque sostenere durante il periodo di copertura.
Normalmente esse possono essere ricondotte a tre tipologie, ed in particolare si parla di:

- Spese per acquisire la polizza, quindi eventuali commissione all'agente, stime di valori o comunque tutto ciò che è necessario per stipulare il contratto. Data l'elevate numerosità di contratti assicurati in carico a ciascuna agenzia, non si andrà realmente a conteggiare le diverse spese sostenute, ma si opererà sulla base di un'aliquota α che sia in grado di approssimare tale valore.
- Spese sostenute al momento dell'incasso dei premi, quindi si ripresenta la provvigione da dare all'agente, oltre alle eventuali spese bancarie o

postali. Anche in questo caso si parlerà di aliquota da applicare, denominata β.

- Spese per la gestione della pratica, in cui vanno a confluire le restanti voci. Questi costi, incorporati nel premio attraverso l'aliquota γ, sono quelli che in realtà hanno un maggior peso perché riguardano affitto dei locali, stipendi del personale, cartoleria, e tutto ciò che è necessario per mantenere operativa una compagnia assicurativa.

Indicando con Π^{α}, Π^{β} e Π^{γ} i tre caricamenti espliciti appena esposti, il premio di tariffa $\overline{\Pi}$ sarà dato dall'equazione:

$$\overline{\Pi} = \Pi + \Pi^{\alpha} + \Pi^{\beta} + \Pi^{\gamma}$$

Quindi per comprendere quale sarà la somma che pagherà il soggetto assicurato, basterà sommare al premio puro Π, attraverso l'aliquota adeguatamente valutata, le spese derivanti da acquisizione, gestione ed incasso del premio (PORZIO, PREVIATI, COCOZZA, MIANI, & PISANI, 2011).

2.2. I MODELLI PROBABILISTICI

In realtà, per poter procedere correttamente al calcolo di un premio che possa corrispondere adeguatamente ai criteri sopra descritti, è importante valutare singolarmente il numero di sinistri N e l'ammontare del danno Z per ciascun sinistro. La strumentazione statistica che meglio permette di prevedere i risarcimenti attesi per ogni polizza sono i modelli probabilistici, in particolare ci si va a focalizzare su quelli che nel modo migliore permettono di studiare i due differenti elementi della base tecnica.

2.2.1. I MODELLI PER IL NUMERO DI SINISTRI

Vengono presentati i quattro modelli più frequentemente utilizzati per catturare le informazioni sulla variabile N numero di sinistri, non con lo scopo di doverli implementare tutti per ottenere i risultati che a noi interessano, semplicemente verrà adottato solo quello che in modo ottimale provvede a perequare i dati osservati.

2.2.1.1. IL MODELLO DI POISSON

Il modello maggiormente noto e a cui si è fatto riferimento per molti anni è quello che si avvale della distribuzione di Poisson; tale distribuzione, detta anche legge degli eventi rari, va ad indicare la probabilità che in un determinato arco di tempo si verifichi l'evento; in realtà siamo in grado di effettuare questo tipo di studio solo se si è già a conoscenza di almeno una caratteristica di tale distribuzione, quindi ad esempio essere a conoscenza che in un mese sono stati registrati 10 sinistri mi permette di studiare se in un anno è possibile che i sinistri che verranno registrati siano superiori a 100.

Tale modello si distingue per avere:

- Funzione di probabilità $p_k = e^{-\lambda} \frac{\lambda^k}{k!}$
- Media $E(N) = \lambda$
- Varianza $var(N) = \lambda$

Per meglio comprendere in che modo utilizzare il modello introdotto, è opportuno introdurre un esempio; nel caso che segue l'obiettivo è quello di valutare il comportamento di un campione di automobilisti in base al numero di sinistri effettuati nel periodo di copertura (FERRARA, 2012).

Tabella 2.1: Frequenza sinistri di un campione di automobilisti

Numero di sinistri k	Numero di conducenti n_k
0	20592
1	2651
2	297

3	41
4	7
5	0
6	1
7	0
Totale	23589

Fonte: FERRARA (2012)

Il valore atteso, e quindi la media della nostra distribuzione sarà:

$$E(N) = \lambda = 0{,}1442$$

A partire da questo dato è possibile calcolare la probabilità che il numero di sinistri sia pari ad un ipotetico numero k semplicemente applicando la formula precedente ed ottenere il risultato illustrato nella tabella che segue:

Tabella 2.2: Calcolo delle probabilità secondo il modello di Poisson

Numero di sinistri k	Numero di conducenti n_k	Probabilità
0	20592	0,86570
1	2651	0,12485
2	297	0,00900
3	41	0,00043
4	7	0,00002
5	0	0,00000
6	1	0,00000
7	0	0,00000
Totale	23589	

Fonte: FERRARA (2012)

Nonostante i risultati ottenuti attraverso questo modello, bisogna sottolineare che esso per poter esser utilizzato necessita di determinati elementi, ed in particolar modo si fa riferimento a:

- L'arco temporale necessario per effettuare l'analisi deve essere scelto in modo che non siano presenti al suo interno i picchi di sinistrosità, e quindi il valore massimo possibile di sinistri, sennò potrebbe influenzare i calcoli; in altri termini si sta chiedendo che il periodo di tempo considerato deve fare in modo che la probabilità che si verifichi un sinistro rimanga sempre piccola e costante.
- Il fatto che si verifichi un sinistro non potrà mai influenzare, positivamente o negativamente, la probabilità che ne avvenga oppure no un secondo (CORAIN, 2011).

Si può infine notare che possedere media e varianza con il medesimo valore, peculiarità del modello di Poisson, può essere visto come un vantaggio di questo modello; non solo perché permette di facilitare i calcoli, ma anche perché permette di studiare non solo un numero $N_1, N_2, \dots N_n$ di variabili indipendenti di Poisson ma anche la loro somma $N = N_1 + N_2 + \dots + N_n$, poiché si caratterizza anch'essa per avere la stessa distribuzione.

Allo stesso tempo però questa importante caratteristica può porre dei limiti al modello stesso, infatti evidenze empiriche portano ad evidenziare che normalmente si è in presenza di una varianza con valore maggiore del valore medio (PARRINI, 2012).

2.2.1.2. LA BINOMIALE NEGATIVA

Attorno agli anni '70 inizia a diffondersi il concetto di propensione al rischio dei diversi soggetti, un elemento nuovo che per poter essere assimilato in modo appropriato con le caratteristiche già allora note, necessitava l'utilizzo di una forma distributiva che non presentasse i difetti e vincoli descritti precedentemente del modello poissoniano; per accogliere al meglio la richiesta di maggiore flessibilità si suggerì l'utilizzo di modelli dati dalla mistura tra la funzione di Poisson e la funzione Gamma: ovvero una delle possibili formulazione della binomiale negativa.

La capacità di studiare situazioni in cui vi siano bassi valori medi e campioni di dimensioni non generose, ha portato a considerare questo modello come quello migliore per descrivere le distribuzioni ottenute, in particolar modo per quanto riguarda l'ambito RCauto.

Data p la probabilità di successo, ed indicato con x e λ rispettivamente il numero di incidenti e la propensione al rischio, la probabilità che descrive il numero di fallimenti che si devono verificare prima di ottenere il k-esimo successo è definita dall'espressione (IODICE D'ENZA, 2009):

$$P_k(X = x) = \binom{x-1}{k-1} p^k (1-p)^{x-k} \quad \forall_x = k, k+1 \ldots$$

In altri parole si va a studiare $p(x, \lambda, t)$, ovvero la probabilità che si verifichino x sinistri, con propensione al rischio λ in un arco temporale che va da 0 a t (FERRARA, 2012).

Dato il valore medio è $E(X) = k\frac{1}{p}$ e la varianza è $var(X) = k\frac{1-p}{p^2}$, si fa riferimento all'esempio numerico trattato in precedenza per calcolare la probabilità ottenuta mediante l'utilizzo del modello con la distribuzione binomiale negativa:

Tabella 2.3: Esempio numerico con distribuzione binomiale negativa

Numero di sinistri k	Numero di conducenti n_k	Probabilità
0	20592	0,87353
1	2651	0,11088
2	297	0,01368
3	41	0,00167
4	7	0,00020
5	0	0,00002

6	1	0,00000
7	0	0,00000
Totale	23589	

Fonte: FERRARA (2012)

2.2.1.3. LA DISTRIBUZIONE BINOMIALE

La distribuzione binomiale, detta anche distribuzione di Bernoulli, si caratterizza per la possibilità di studiare un numero n di osservazioni indipendenti e ripetute, dove gli unici esiti possibili possono essere successo o insuccesso.

Il modello in questione per poter essere applicato necessità comunque che il valore della varianza sia inferiore a quello della media, quindi sarà il risultato ottenuto dagli stessi dati che potrà indicare se la binomiale sia la soluzione adeguata per studiare tale fenomeno.

Definita p la probabilità di successo e $q = 1 - p$ la probabilità di insuccesso, la funzione che indica la probabilità di ottenere k risultati dopo aver effettuato n prove indipendenti, è la seguente:

$$P_k = \binom{n}{k} p^k q^{n-k} \qquad 0 \leq k \leq n$$

Il tipico esempio a cui si fa riferimento per dimostrare l'utilizzo di questo modello è il lancio di una moneta; si veda comunque l'esempio che segue per osservarne l'applicazione anche nel contesto assicurativo.

Tabella 2.4: Numero di sinistri per polizza con copertura un anno, in un campione di 15160 polizze

Numero di sinistri k	Numero di polizze n_k
0	5367
1	5893
2	2870

3	842
4	163
5	23
6	1
7	1
8+	0

Fonte: FERRARA (2012)

Dato il valore medio $E(N) = np = 0.9854$ e la varianza $var(N) = npq = 0.8904$, ipotizzando un numero n pari a 10 prove, la probabilità che si ottengano k sinistri per ogni contratto attraverso la distribuzione binomiale sarà:

Tabella 2.5: Risultato ottenuto attraverso la distribuzione binomiale

Numero di sinistri k	Numero di polizze n_k	Probabilità
0	5367	0,35440
1	5893	0,38740
2	2870	0,19060
3	842	0,05550
4	163	0,01060
5	23	0,00140
6	1	0,00010
7	1	0,00000
8	0	
9	0	
10	0	
Totale	15160	

Fonte: FERRARA (2012)

2.2.2. I MODELLI PER IL DANNO PROVOCATO DAL SINISTRO

I modelli che vengono utilizzati per studiare questo tipo di variabile si distinguono da quelli introdotti precedentemente perché si basano su dati di natura diversa, ovvero fenomeni che non possono essere conteggiati ma solamente misurati e per questo assumono il nome di variabili continue. In questi casi è corretto parlare solo di funzioni di densità di probabilità e quindi la probabilità che vengano assunti valori compresi all'interno di un intervallo, dato che la probabilità di assumere un particolare valore è pari a zero.

La variabile aleatoria Z sarà quindi rappresentata da una distribuzione di probabilità continua con densità $f_Z(x)$ continua per tutti i valori di $x > 0$. (PARRINI, 2012)

2.2.2.1. LA DISTRIBUZIONE LOGNORMALE

Questo tipo di funzione viene generalmente definita come «la distribuzione di probabilità di una variabile aleatoria x il cui logaritmo $logx$ segue una distribuzione normale» (CORAIN, 2011).

In base a questa distribuzione la funzione di densità della variabile x presa in esame, è data da

$$f_z(x) = \frac{1}{\sigma\sqrt{2\pi}}\frac{1}{x}e^{-\frac{1}{2\sigma^2}(lnx-\mu)^2}$$

con media $E(x) = e^{(\mu+\frac{\sigma^2}{2})}$ e varianza $Var(x) = e^{(2\mu+\sigma^2)}e^{(\sigma^2-1)}$.

Spesso nella assicurazione la scelta di utilizzare questo strumento è data dal fatto che essendo caratterizzata da una coda "pesante" o anche detta "heavy tailed" permette di assegnare una probabilità anche ad valori molto alti, così rende possibile studiare anche quelli che vengono definiti casi estremi; la seconda motivazione invece deriva dalla notevole semplificazione dei calcoli, dettata dall'equivalenza:

$$Z \sim LNorm(\mu, \sigma^2) \equiv \ln(Z) \sim Norm(\mu, \sigma^2)$$

La scelta di adottare questo modello comunque non può prescindere dal vincolo imposto dal teorema centrale del limite, ovvero che la distribuzione lognormale possa rappresentare il danno solo nell'ipotesi in cui esso sia generato da un numero sufficientemente ampio di caratteristiche del rischio che siano indipendenti, identicamente distribuite e operanti in senso moltiplicativo; ipotesi, per altro, decisamente poco realistica (PARRINI, 2012).

2.2.2.2. LA DISTRIBUZIONE DI PARETO

Questo tipo di distribuzione nasce in realtà in un contesto ben diverso da quello assicurativo, infatti, si prefigge di descrivere in che modo venga ripartito il reddito all'interno della popolazione; col passare degli anni però si è osservato che tale struttura ben si collocava anche nell'ambito delle assicurazioni, in particolare si è dimostrata utile a studiare l'andamento del costo dei sinistri per quanto riguarda i casi definiti estremi.

Nonostante siano state individuate due versioni di questa distribuzione, quella di maggior interesse, definita europea, viene definita attraverso un termine c che sta ad indicare il valore (in termini economici) del sinistro di minori dimensioni che è stato rilevato, ed un termine a che riassume il comportamento della coda della distribuzione; l'eterogeneità dell'importo dei sinistri all'interno di un portafoglio si distribuisce quindi come la distribuzione di Pareto secondo la formula:

$$F(x) = 1 - \left(\frac{c}{x}\right)^a \qquad x > c \qquad a, c > 0,$$

oppure con espressione equivalente (traslata verso destra):

$$F(x) = 1 - \left(\frac{l + c}{l + x}\right)^a \qquad x > c > 0$$

Da qui è semplice intuire che se il termine l arrivasse ad assumere valore 0, le due formule corrisponderebbero.

Per meglio comprendere l'espressione appena esposta è doveroso chiarire che all'aumentare del valore del termine a ci si aspetta che la coda di destra abbia un peso minore, e di conseguenze un comportamento opposto si verificherà qualora il valore di a diminuisca.

Quest'ultimo termine comunque non è noto a priori, quindi si suggeriscono diversi approcci che permettono di stimare tale valore:

- Il primo metodo preso in considerazione prevede di stimare a attraverso lo stimatore di massima verosimiglianza

$$\hat{a} = \frac{n}{\sum_{i=1}^{n} ln\left(\frac{x_i}{c}\right)} \ ;$$

  metodo che non rientra tra i favoriti, considerato che, dall'analisi della varianza emerge chiaramente che non si tratta di uno stimatore corretto.
- Il secondo approccio si basa sul metodo dei momenti, un metodo che non possiamo definire di poca utilità, ma presenta il difetto di poter essere applicato solo nei casi in cui si è in grado di conoscere non solo il singolo sinistro ma addirittura l'ammontare complessivo dei sinistri che eccedano la soglia c citata in precedenza.
- L'ultima soluzione, alla quale si fa più spesso riferimento in ambito assicurativo, permette di superare i difetti dei due metodi appena elencati; tale metodologia prevedere semplicemente di ponderare il risultato ottenuto per il valore presente nel mercato (MORALDI, 2010).

2.2.2.3. DISTRIBUZIONE GAMMA

Un'ultima funzione che è doveroso introdurre per la frequenza con cui viene utilizzata all'interno dei modelli attuariali è la funzione Gamma. Spesso utilizzata in combinazione con la distribuzione di Poisson, la funzione gamma è la seguente

$$\Gamma(a) = \int_0^{\infty} x^{a-1} e^{-x} dx \ ,$$

con $a > 0$.

Le proprietà di questa funzione possono essere riassunte attraverso l'espressione (MINKOVA, 2010):

$$\Gamma(a) = (a-1)\Gamma(a-1),$$

$$\Gamma(n) = (n-1)!\ , n \geq 1$$

e

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}\,.$$

Se una variabile casuale X si distribuisce secondo la distribuzione Gamma di parametro $a > 0$ e $\beta > 0$, ovvero $X \sim \Gamma(a, \beta)$, la sua funzione di densità sarà data da

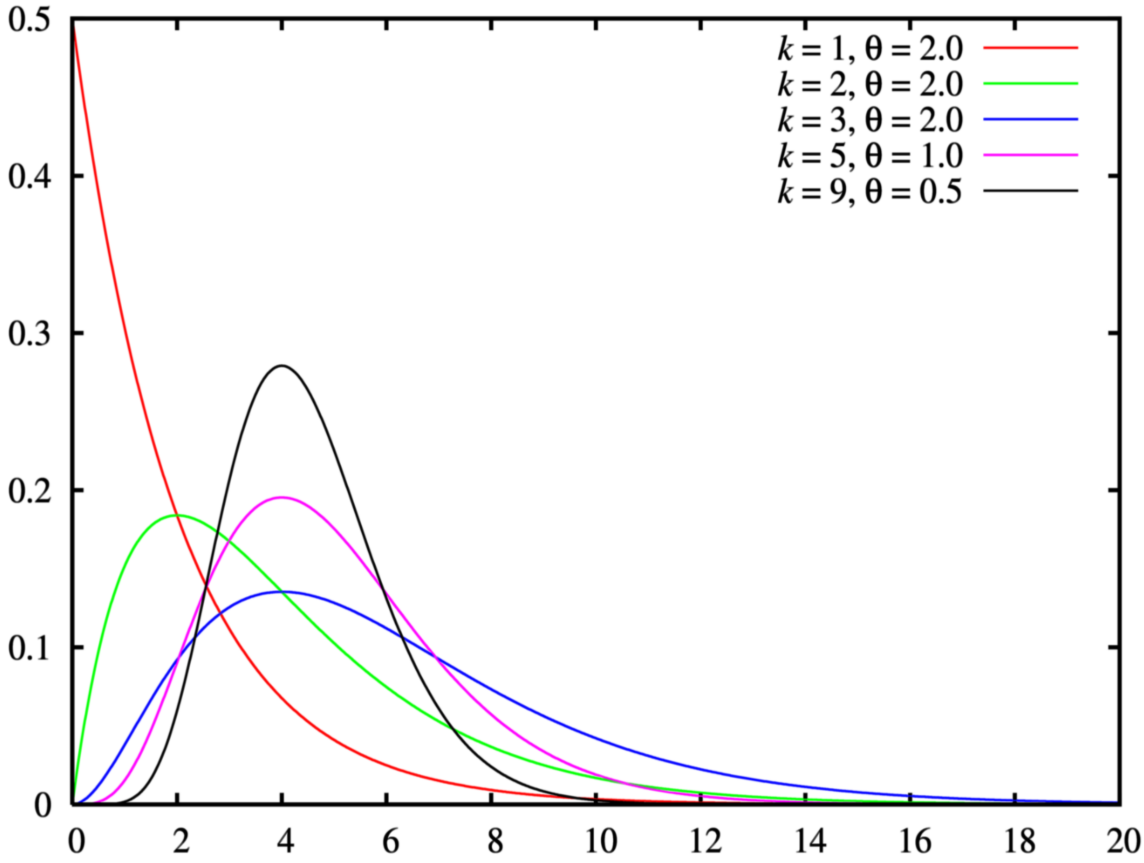
$$f(x) = \frac{\beta^a}{\Gamma(a)} x^{a-1} e^{-\beta x}, con\ x > 0\,,$$

in cui media e varianza saranno rispettivamente:

$$E(X) = \frac{a}{\beta} \qquad Var(X) = \frac{a}{\beta^2}\,.$$

Si osserva, infine, che il parametro di forma per questo tipo di distribuzione è a, quindi al variare di tale valore la distribuzione Gamma potrà assumere diverse forme, così come viene illustrato nella figura 2.2 attraverso il parametro k.

Figura 2.2: Distribuzione Gamma al variare del parametro di forma



Fonte: "Gamma distribution pdf". Con licenza CC BY-SA 3.0 tramite Wikimedia Commons - http://commons.wikimedia.org/wiki/File:Gamma_distribution_pdf.png#mediaviewer/File:Gamma_distribution_pdf.png

3. DALLA RACCOLTA DI INFORMAZIONI AL PREMIO DI TARIFFA

3.1. ANALISI DEI FATTORI DI RISCHIO: LA PERSONALIZZAZIONE A PRIORI

Il primo passo da fare, ed anche il più elementare, è quello della raccolta ed analisi dei dati relativi ai soggetti che entreranno a far parte del portafoglio assicurativo, e quindi esattamente i dati che costituiscono la base tecnica per il calcolo del premio puro.

In base alla numerosità delle informazioni raccolte, varierà l'accuratezza con la quale si andrà a studiare il valore della frequenza sinistri, e l'ammontare stimato del danno; sebbene la preferenza ricada per un'informazione dettagliata sul singolo soggetto, l'impossibilità di ottenere ciò non preclude alla compagnia la possibilità di studiare in modo approfondito la collettività. La soluzione consiste nell'analizzare delle sottoclassi costituite da soggetti che presentano rischi omogenei, denominate anche classi di rischio, in cui le variabili oggetto di studio saranno caratteristiche della vettura o del soggetto assicurato.

L'unica caratteristica che merita una trattazione separata è il *moral hazard*, ovvero, ciò che spinge un individuo a comportarsi in un determinato modo; un'informazione implicita nel comportamento stesso di chi la compie, al punto tale da non riuscire ad essere mai correttamente percepita.

L'approccio più immediato per utilizzare le informazioni relative alla frequenza sinistri ed il danno è quello di calcolare la quota danni attraverso l'approccio seguente:

$$Q = \frac{c_1 + c_2 + \cdots + c_n}{r} ;$$

questo valore, detto anche stima grezza del premio equo, si ottiene dal rapporto fra il costo globale ed il numero di esposizioni.

La quota danni può anche essere espressa come prodotto tra l'indice di sinistrosità (numero medio di sinistri) ed il costo medio di sinistro (DABONI, 1989):

$$Q = \frac{n}{r} \cdot \bar{c} \, .$$

Per chiarire in che modo può essere calcolata, è utile introdurre un esempio numerico basato su 100 contratti, suddivisi tra uomini e donne, come illustra la tabella 3.1;

Tabella 3.1: Esempio numerico considerando un'unica variabile

	UOMINI	DONNE
Numero contratti	60	40
Esborso 0€	8	0
Esborso 150€	40	20
Esborso 300€	12	18
Esborso 500€	0	2

L'esborso complessivo sarà dato da:

$$0€ \times 8 \ + 150€ \times 60 \ + 300€ \times 30 \ + 500€ \times 2 \ = 19000€ \, ;$$

ed in particolare, ammontare relativo agli uomini è:

$$0€ \times 8 + 150€ \times 40 + 300€ \times 12 + 500€ \times 0 = 9600€ \, ,$$

mentre quello relativo alle donne sarà:

$$0€ \times 0 + 150€ \times 20 + 300€ \times 18 + 500€ \times 2 = 9400€ \, .$$

Sulla base dei risultati appena conseguiti, è quindi possibile calcolare il costo medio del danno:

$$costo\ medio\ per\ gli\ uomini = \frac{(esborso\ per\ gli\ uomini)}{contratti\ sinistrati\ uomini)} = \frac{9600}{52} = 184{,}61€ \, ,$$

$$costo\ medio\ per\ le\ donne = \frac{esborso\ per\ le\ donne}{contratti\ sinistrati\ donne} = \frac{9400}{40} = 235€ \, .$$

Il secondo elemento della base tecnica che è necessario calcolare è la frequenza sinistri, ottenibile attraverso l'equazione seguente:

$$frequenza\ sinistri\ uomini = \frac{contratti\ sinistrati\ uomini}{contratti\ stipulati\ uomini} = \frac{52}{60} = 0{,}87\ ,$$

$$frequenza\ sinistri\ donne = \frac{contratti\ sinistrati\ donne}{contratti\ stipulati\ donne} = \frac{40}{40} = 1\ .$$

Infine, sarà sufficiente moltiplicare il costo medio e la frequenza sinistri per ricavare il valore della quota danni per entrambi i sessi:

$$quota\ danni\ uomini = 0{,}87 \times 184{,}61 = 160{,}61€\ ,$$

$$quota\ danni\ donne\ \ = 1 \times 235 = 235€\ .$$

Qualora una compagnia assicurativa ottenesse dei risultati simili, le soluzioni plausibili sarebbero sostanzialmente due. Ipotizzando che il numero e l'entità dei sinistri rimangano stabili, si potrebbe stabilire un premio medio pari a 190€ da far pagare ad ambo i sessi; la seconda soluzione prevede, invece, di far pagare un premio differenziato in base alla classe di rischio di appartenenza, che sarà in questo caso 235€ per le donne, e 160,61€ per gli uomini.
In realtà, però, nessuna delle due soluzioni è ottimale; la prima farebbe pagare di più alla categoria maschile, rischiando di disincentivare tale tipologia di soggetti dall'assicurarsi con la compagnia in questione, la seconda va a penalizzare chi fa parte della categoria donne, ma alla guida ha un atteggiamento decisamente meno rischioso rispetto ad altre (PORZIO, PREVIATI, COCOZZA, MIANI, & PISANI, 2011).
Proprio per ovviare a questi annosi problemi, si è preferito adottare un sistema di personalizzazione del premio, un premio che vada quindi ad analizzare il singolo e non semplicemente la classe di rischio a cui

appartiene. La personalizzazione può essere effettuata a priori, e quindi sulla base delle caratteristiche che si vengono a conoscere al momento della stipula del contratto, oppure a posteriori, e quindi sulla base di regole di adeguamento del premio che saranno modellate sui dati rilevati durante un orizzonte temporale più ampio (quello di copertura del contratto).

Premesso ciò, si passa all'individuazione di quali siano le variabili che vengono considerate influenti da un punto di vista della rischiosità; attraverso opportune metodologie statiche, come ad esempio i modelli lineari generalizzati, si possono creare dei gruppi omogenei di assicurati che presentano le stesse caratteristiche, ed infine attribuire a ciascuna di essi il premio ritenuto più opportuno.[18]

Le variabili che possono essere prese in considerazione sono numerose, ma la scelta ricade solo su quelle che possono essere realmente significative, e quindi suggerire un possibile profilo di rischiosità; tra queste vi sono quelle che riguardano dati soggettivi, e quindi coinvolgono direttamente colui sul quale viene stipulata la polizza, come:

- Età. È stato dimostrato che le persone al di sotto di una certa soglia di età (stabilita intorno ai 25/26 anni) sono più soggetti a commettere sinistri, e quindi avranno valori maggiori di frequenza sinistri.
- Anzianità patente. Gli anni trascorsi dal conseguimento della patente vengono considerati un elemento molto importante, si dimostra infatti che, l'esperienza alla guida può giovare sulla prudenza.

[18] Nonostante vi siano delle analogie tra i soggetti che vengono collocati nella stessa classe tariffaria, rimane comunque una notevole eterogeneità nei comportamenti, eterogeneità che si cerca di eliminare con una personalizzazione a posteriori del premio. Ovviamente è un processo di aggiustamento continuo, e grazie alla storia del soggetto esaminato si cerca nel lungo termine di far pagare un premio il più fedele possibile alle caratteristiche riscontrate per poi ottenere un progressivo passaggio dal premio collettivo a quello che è un premio individuale.

- Numero dei conducenti. Il fatto che un mezzo possa essere guidato da soggetti diversi da colui al quale viene intestata la polizza, pone maggiori rischi; per questo viene assegnato uno sconto a coloro che dichiarano di essere gli unici conducenti del mezzo, mentre nel caso contrario verrà fatto pagare un prezzo maggiore.
- Modalità di utilizzo. Degli studi hanno dimostrato che chi utilizza la macchina per lavoro è più esposto a determinati rischi, quindi dovrà pagare una cifra maggiore rispetto ad un soggetto che la utilizza esclusivamente per brevi tratti di strada.
- Classe Bonus-Malus. Questa variabile, che sarà introdotta nei capitoli seguenti, nonostante si basi su quella che è l'esperienza, e quindi è ottenibile esclusivamente a posteriori, entra comunque a far parte di un sistema di catalogazione dei diversi soggetti.
- Sesso. Anche se in realtà, attualmente, non è più possibile distinguere i soggetti in base a tale variabile; questo è dovuto al fatto che la Corte di Giustizia Europea ha stabilito (con la sentenza del 1 Marzo 2011) che a partire dalla data 21 Dicembre 2012 non verrà più fatta una distinzione del premio sulla base del sesso, ma si dovranno emettere esclusivamente polizze definite "unisex" onde evitare discriminazioni.

A queste variabili vanno poi aggiunte quelle che riguardano prettamente il veicolo, ovvero:

- Alimentazione. Talvolta la scelta di acquistare una macchina alimentata a gas o diesel è dettata dal fatto che vi sarà una maggior percorrenza di chilometri, fattore che la compagnia associa ad una maggior frequenza nei sinistri.
- Potenza del veicolo. Questo dato risulta particolarmente importante perché, a differenza degli altri, può suggerire anche l'entità del danno che potrebbe essere causato in caso di sinistro; in particolar modo si osserva un proporzionale aumento del premio da pagare, all'aumentare della potenza del veicolo.

- Zona di circolazione.[19] Non in tutte le città, e soprattutto non in tutte le regioni, si verifica lo stesso numero di sinistri, quindi si cerca di adeguare il premio anche alla zona in cui andrà prevalentemente a circolare la macchina.
- Presenza di particolari caratteristiche. Rientrano sotto questa voce quelli che possono essere considerati equipaggiamenti di sicurezza, quindi la presenza di ABS (solo passeggero o anche conducente) o scatole nere. Queste ultime riducono notevolmente il prezzo che dovrà essere pagato dall'assicurato perché permettono alla compagnia di verificare costantemente la posizione del veicolo, e quindi qualora si verifichi un sinistro sarà inequivocabile la determinazione di chi sia la responsabilità, e di conseguenza anche l'entità del danno che dovrà rimborsare la compagnia[20].

Ogni compagnia poi potrà valutare se richiedere delle ulteriori informazioni che permettano una miglior valutazione del soggetto.

Lo studio attento sul campione si focalizza in particolar modo sulle variabili di cui ci interessa la distribuzione, come la quota danni e la frequenza dei sinistri, ma normalmente si ritiene necessario anche fare un'analisi dei costi sostenuti dalla compagnia per i guidatori che hanno causato un numero superiore a 4 incidenti, e confrontarli con chi ne ha provocati un numero inferiore, per capire di quanto si discostano i due valori.

Nel momento in cui si ha il numero sufficiente di dati, si deve creare un collegamento logico tra le variabili esplicative (qualitative e quantitative) e le variabili risposta (valutazioni probabilistiche), operazione che è possibile fare anche solo grazie all'utilizzo di tabelle che esplicitino tale relazione, come illustra la tabella 3.2.

[19] Talvolta si preferisce utilizzare come dato la provincia di immatricolazione del veicolo.

[20] Per maggiori approfondimenti si consiglia di consultare il sito www.ania.it.

Tabella 3.2: Relazione tra le classi di determinazione della potenza fiscale e le loro variabili risposta

POTENZA FISCALE	NUMERO POLIZZE	NUMERO SINISTRI	AUTOV. ANNO	FREQ. SINISTRI	DANNO MEDIO	QUOTA DANNI
≤8 HP	2.059	100	1838,93	5,44	1858	101,45
9-10 HP	2.354	178	2083,14	8,55	2319	198,158
11-12 HP	8.432	681	7406,61	9,19	2681	246,514
13-14 HP	7.279	615	6376,92	9,64	2147	207,043
15-16 HP	6.919	636	6014,67	10,57	2165	228,882
17-18 HP	5.662	608	4886,25	12,44	2427	301,970
19-20 HP	3.644	404	3123,78	12,93	2480	320,778
>20 HP	1.548	241	1350,11	17,85	2706	483,062
globale	37.987	3463	33080,41	10,47	2383	249,426

Fonte: SIGALOTTI & PICECH (1994)

Tabella 3.3: Relazione tra le classi di determinazione dell' anzianità di patente e le loro variabili risposta

ANZIAN. PAT.	NUM. POLIZZE	NUM. SIN.	AUTOV. ANNO	FREQ. SIN.	DANNO MEDIO	QUOTA DANNI
0-2 anni	753	130	638,59	20,36	2782	566,404
3-5 anni	2217	257	182,972	14,05	2955	414,988

6-10 anni	4286	403	3733,41	10,79	2290	247,240
>10 anni	25855	2079	22579,48	9,21	2304	212,167
globale	33021	2869	28781,20	9,97	2382	237,470

Fonte: SIGALOTTI & PICECH (1994)

Grazie a tale schematizzazione emergono immediatamente i fattori che maggiormente impattano sulla sinistrosità, così da facilitare la costruzione di un modello tariffario che permetta la formazione di classi di rischio che rispondano a determinati requisiti: omogeneità interna, dimensioni adeguate, incentivo ad una guida più prudente, e che permetta di mettere in relazione più variabili (OUTREVILLE, 1998) .
Fermo restando quanto appena esposto, la scelta di adottare metodi di classificazione che prevedano una o più voci sopra esposte non è dettata dalla conoscenza personale dei soggetti preposti a tale ruolo, ma è suggerita dall'utilizzo di specifici strumenti parametrici; lo stesso vale anche per il numero ottimale di classi tariffarie, ottenuto dall'utilizzo di strumenti come la cluster analysis.

Dopo aver introdotto le possibili variabili, è opportuno esemplificare attraverso un esempio numerico in che modo è possibile formulare i calcoli in termini di quota danni se le variabili considerate diventano due: l'anzianità della patente e la zona di residenza del conducente.
In questo caso si potrà parlare di binomio (i, j) che si presenterà nelle modalità n_i e n_j dove il primo parametro avrà $i = 1{,}2$ ed allo stesso modo per il secondo $j = 1{,}2$.

Tabella 3.4: Esempio numerico per illustrare il concetto di quota danni

Anzianità patente	Zona residenza	Numero Polizze	Numero Sinistri	Costo Medio
≤ 5 anni (1)	Rischiosa(1)	1957	354	1863

≤ 5 anni (1)	Non rischiosa(2)	2632	302	1702
> 5 anni (2)	Rischiosa(1)	5735	1073	1697
> 5 anni (2)	Non rischiosa(2)	6800	566	1654
TOTALE		17124	2295	1729

Fonte: CERCHIARA (2014)

Per l'intera collettività è possibile calcolare la quota danni semplicemente risolvendo l'equazione

$$Q = \frac{n.sinistri}{n.polizze} \times costo\ medio = \frac{2295}{17124} \times 1729 = 231{,}72€\,.$$

Volendo analizzare le singole classi di rischio invece otterremo:

$$Q(1{,}1) = \frac{354}{1957} \times 1863 = 337$$

$$Q(1{,}2) = \frac{302}{2632} \times 1702 = 195{,}29$$

$$Q(2{,}1) = \frac{1073}{5735} \times 1697 = 317{,}50$$

$$Q(1{,}1) = \frac{566}{6800} \times 1654 = 137{,}67$$

L'informazione ottenuta attraverso questo approccio risulta significativa ai fini di calcoli superficiali, ma i rischi nei quali si può incorrere sono diversi, come quello di avere delle classi di rischio vicine che pagano premi eccessivamente differenti. Quindi, per evitare simili inconvenienti, oltre a quello che su un premio si possano andare a ripercuotere semplici errori di stima, o addirittura periodici picchi di sinistrosità, bisogna provvedere con metodologie più precise: un modello tariffario (CERCHIARA, 2014).

Per procedere bisogna quindi operare in un'ottica diversa, dove acquistano maggiore importanza le diverse modalità in cui ogni dato può presentarsi, ovvero i valori che ogni variabile può assumere (ad esempio per la variabile zona di residenza, le informazioni da considerare saranno il numero di persone che abitano per ciascuna città), perché solo su queste basi è possibile formulare un premio specifico τ_{ij} [21] per ogni classe di rischio considerata.

L'ampia conoscenza di strumenti statistici permette non solo di catalogare facilmente questi dati, ma anche di capire il legame che tra essi intercorre; le informazioni che ne derivano possono riguardare la correlazione tra una determinata zona residenziale e la maggiore frequenza di sinistri rilevata, piuttosto che una relazione tra la potenza del veicolo e l'entità del danno causato.

Sempre avvalendosi di questi strumenti è, poi, possibile comprendere in che modo le variabili rilevate vadano ad influenzare il premio; in particolare si osserva che qualora esse abbiano un'influenza di tipo additivo, si adotterà un modello additivo

$$\tau_{ij} = \tau + \alpha_i + \beta_j \,,$$

mentre se l'influenza è di tipo moltiplicativo, si adotterà un modello moltiplicativo

$$\tau_{ij} = \tau \times \alpha_i \times \beta_j \,.$$

I coefficienti α e β che compaiono nei due modelli vengono chiamati relatività, e vanno valutati in modo tale che il risultato ottenuto si avvicini a quello calcolato con la quota danni (PORZIO, PREVIATI, COCOZZA, MIANI, & PISANI, 2011). Ponendo il parametro τ pari alla quota danni Q, si possono individuare diversi criteri di stima per tali relatività, e tra questi andiamo a segnalare:

- <u>Il metodo delle relatività intuitive</u> che suggerisce di calcolare le relatività attraverso le seguenti formule

$$\widehat{\alpha_i} = \frac{Q_{i\cdot}}{Q} \qquad \widehat{\beta_j} = \frac{Q_{\cdot j}}{Q} \,,$$

---

[21] Solo qualora si avesse a disposizione una base di dati sufficientemente ampia sarebbe possibile associare fin da subito ad ogni classe un premio "grezzo" τ'_{ij}.

ovvero ponendo a numeratore, rispettivamente, la quota danni delle classi di rischio che presenta la prima variabile tariffaria con modalità i, e la quota danni delle classi che invece hanno come seconda variabile la modalità j.

Il metodo delle relatività intuitive prevede quindi che, qualora uno dei due rapporti dia un risultato superiore ad uno, la relatività verrà considerata in senso negativo perché andrà a "pesare" sul premio come sovrapprezzo, viceversa un risultato inferiore ad 1 potrà alleggerire il premio finale grazie ad uno sconto.

Possiamo quindi riscrivere il modello moltiplicativo anche come

$$P_{ij} = Q\hat{\alpha}_i\hat{\beta}_j = \frac{Q_{i\cdot}Q_{\cdot j}}{Q}.$$

Tenendo sempre in considerazione l'esempio introdotto in precedenza, possiamo applicare quanto appena esposto per calcolare la quota danni relativa alla variabile anzianità patente:

$$Q_{i\cdot} = \frac{risarcimento\ totale\ per\ rischi\ con\ anzianità\ patente\ i}{numero\ totale\ polizze\ per\ rischi\ con\ anzianità\ patente\ i}$$

$$Q_{1\cdot} = \frac{1173506}{4589} = 255{,}72 \quad Q_{2\cdot} = \frac{2757045}{12535} = 219{,}95\,,$$

mentre per la variabile zona di residenza:

$$Q_{i\cdot} = \frac{risarcimento\ totale\ per\ rischi\ con\ zona\ residenza\ j}{numero\ totale\ polizze\ per\ rischi\ con\ zona\ residenza\ j}$$

$$Q_{\cdot 1} = \frac{2480383}{7692} = 322{,}46 \quad Q_{\cdot 2} = \frac{1450168}{9432} = 153{,}75\,.$$

I risultati ottenuti sono riassunti nella tabella 3.5.

Tabella 3.5: Premi ricavati tramite il metodo di stima delle relatività intuitive

Anzianità patente	Zona residenza	Premio
1	1	82459,47
1	2	39316,95
2	1	70925,08
2	2	33817,31

Fonte: GIGANTE, PICECH, & SIGALOTTI (2010)

- Il metodo dei minimi quadrati propone invece di calcolare le due relatività in modo che sia minima la funzione

$$\sum_{i=1}^{I}\sum_{j=1}^{J}(Q_{ij} - f(\alpha_i, \beta_j))^2,$$

  quindi rendendo minimo il valore dato dalla somma dei quadrati degli scostamenti, tra quanto stimato e quanto realmente osservato (GIGANTE, PICECH, & SIGALOTTI, 2010).

L'equilibrio tecnico finanziario promosso dall'ISVAP presuppone, poi, l'aggiunta a tale premio specifico delle spese di acquisizione e gestione che sosterrà la compagnia durante il periodo di copertura, ovvero il caricamento di sicurezza: parte integrante del premio di tariffa proposto a chi intende sottoscrivere una polizza.

Se questo elemento aggiuntivo viene calcolato in modo tale da rispettare l'eguaglianza

$$\sum_{i=1}^{n_i}\sum_{j=1}^{n_j}\gamma_{ij} P_{ij} = \sum_{i=1}^{n_i}\sum_{j=1}^{n_j}\gamma_{ij} P^{(0)}\alpha_i \times \beta_j ,$$

allora si potrà indicare il premio finale come

$$P_{ij} = P^{(0)}\alpha_i \times \beta_j ,$$

dove

$$P^{(0)} = \frac{\gamma\Pi}{\sum_{i=1}^{n_i}\sum_{j=1}^{n_j}\gamma_{ij}\, P^{(0)}\alpha_i \times \beta_j}.$$

quindi, il premio specifico sarà modellato in base alle relatività su quello che è il premio di tariffa, e non semplicemente sul premio equo (PORZIO, PREVIATI, COCOZZA, MIANI, & PISANI, 2011).

3.2. PERSONALIZZAZIONE A POSTERIORI: LA VARIABILE PROFILO DI RISCHIO

A prescindere dall'accuratezza adottata nella selezione e nella catalogazione delle informazioni riguardanti i nuovi assicurati, ogni classe di rischio sarà comunque caratterizzata da elementi di eterogeneità, che possono derivare da fattori esterni e talvolta non prevedibili. A partire dagli anni '50 si è cercato di risolvere questo problema attraverso l'introduzione di un parametro di rischio θ, che incorporasse al suo interno le informazioni in grado di diminuire, se non eliminare, il vincolo di disomogeneità (DE FERRA, 1995).

Il vero concetto di premio, quello che viene generalmente chiamato "pricing", richiede infatti l'introduzione di questo elemento aggiuntivo così che ogni assicurato si vedrà attribuita una diversa propensione al rischio a prescindere dalla classe di rischio di appartenenza; il meccanismo prevede che, ad esempio, se per un soggetto venga stabilito un valore di θ pari a 50%, ci si aspetti che durante il periodo di copertura assicurativa ad esso saranno attribuiti, con molta probabilità, un numero di incidenti inferiore rispetto a quelli che appartengono alla sua stessa classe tariffaria, e per la precisione esattamente la metà; al contrario, attribuendo un valore di θ pari a 300%, l'aspettativa è quella di un comportamento molto più rischioso, tanto da rendere possibile un numero di sinistri pari a tre volte quello attribuito generalmente alla sua classe tariffaria.

Per chiarire, quando si parla di parametro θ, ci si riferisce in realtà ad una delle possibili realizzazioni della variabile casuale Θ, ovvero la variabile in grado di riassumere tutti i possibili valori assumibili dai diversi profili di rischio ancora non completamenti noti; quindi, dalla stessa definizione se ne ricava che in presenza di notevole eterogeneità aumenteranno i possibili valori assumibili dal parametro di rischio [22] (DENUIT, MARECHAL, PITREBOIS, & WALHIN, 2007).

Sebbene il profilo di rischio sia già stato citato in precedenza, ed in particolare nell'elenco delle possibili variabili tariffarie, tale forma di catalogazione potrà essere applicata solo in un secondo momento, ovvero quando sono state raccolte informazioni sufficienti per poter parlare di esperienza. Il momento in questione non potrà essere precedente a quello di rinnovo della prima copertura assicurativa, ed è proprio in tale data che, dopo aver stipulato per l'anno 0 il premio iniziale τ_0, la compagnia assicurativa ha il dovere di stipulare per l'anno $t+1$ un premio che incorpori tutte le informazioni raccolte, ovvero "experience rating"[23], e quindi ricreare un nuovo premio $\tau_{t+1} = f(\tau_0, n_1, n_2, \dots, n_t)$ che sia funzione non solo del numero di sinistri avvenuti durante il periodo di copertura, ma anche del premio corrisposto l'anno precedente (PITACCO, 2000).

Fermo restando che il fine ultimo di un attuario è quello di ricavare un premio individuale basato sul profilo di rischio

$$P^{ind} = E[\tau_{n+1}|\theta] = \mu(\theta)$$

(in cui sia θ che $\mu(\theta)$ sono valori non noti e che dovranno essere stimati), sulla base degli elementi finora analizzati, si ha la possibilità di calcolare il premio collettivo nel seguente modo:

---

[22] Proprio perché derivano tutti dallo stesso spazio campionario si giustifica il fatto che i soggetti con tali caratteristiche facciano parte della medesima classe tariffaria.

[23] Il concetto di experience rating verrà esposto in modo approfondito nel capitolo seguente.

$$P^{coll} = \int_{\Theta} \mu(\theta)\, dU(\theta) = \mu_0 = E(X_{n+1});$$

quest'ultimo corrisponderà al premio, in media, adeguato per la classe tariffaria presa in esame, dove all'aumentare dei dati disponibili ci si attende una stima più accurata.

Nonostante l'attenzione degli attuari si concentri sul premio individuale, in quanto maggiormente competitivo ai fini commerciali[24], il premio collettivo ha la peculiarità di poter essere stimato con una maggiore accuratezza, tenendo sempre presente che si tratta di un premio dinamico, ed in quanto tale subirà un continuo processo di aggiustamento ad ogni scadenza prestabilita (il rinnovo della polizza).

Giunta a questo livello di classificazione, la compagnia si riserva la facoltà di scegliere quale sia l'approccio più adeguato per ottenere tale livello di informazione, affidandosi a strumenti che richiamano la teoria della credibilità, e quindi un'analisi su base collettiva, oppure attraverso formule di adeguamento su base individuale, come l'approccio bayesiano o i sistemi Bonus-Malus (BUHLMAN & GISLER, 2005).

[24] Saper distinguere i premi in base a quella che è l'effettiva rischiosità del soggetto assicurato è di enorme importanza: avere un premio uguale per tutti porterebbe a far pagare la stessa quota sia ad un "buon" guidatore, che ad uno che invece non lo è. Dato che il costo da sostenere per il secondo soggetto è nettamente superiore, perché sarà predisposto ad effettuare un numero maggiore di incidenti, la strategia consiste nell'avere nel portafoglio assicurativo un numero sufficiente di assicurati con bassa rischiosità, in maniera tale da bilanciare i due effetti. Tale strategia agisce proprio su quello che è il premio individuale: solo proponendo dei premi veramente bassi a questi soggetti, quasi al limite del guadagno, si potrà avere un portafoglio ben diversificato.

4. LA TEORIA DELLA CREDIBILITA'

4.1. INTRODUZIONE

Spiegare a cosa si riferisce la teoria della credibilità è più semplice da un punto di vista pratico che teorico, infatti, sebbene non sia uno strumento in sé, va a risolvere tutte le criticità che normalmente si riscontrano quando si è in procinto di assimilare ed elaborare i dati provenienti da quello che è il singolo individuo o la collettività, dati che appartengono al passato piuttosto che quelli che sono appena stati catalogati; a supportare tale teoria vi sono comunque numerosi strumenti matematici, e tra questi fa da padrone la statistica bayesiana.

L'utilizzo del termine credibilità è stato inizialmente introdotto nelle scienze attuariali per andare a distinguere le situazioni nelle quali il numero di dati a disposizione è sufficientemente ampio da permetterne l'utilizzo per valutare il premio più consono, da quelle situazioni in cui invece si presentano un numero relativamente basso di informazioni, tale da indurre gli studiosi a non tenerle in considerazione ai fini del calcolo del premio. Quindi, in poche parole, era il termine che permetteva di creare una netta distinzione tra i dati credibili, in quanto in numero sufficiente a supportare quanto riscontrato, e quelli che invece risultavano privi di credibilità[25] poichè erano relativamente insufficienti per procedere con il loro studio.

Questa metodologia, che era poco nota in Italia ma già ben presente negli Stati Uniti nella prima metà degli anni '50, si poneva spesso in contrapposizione con quella che è l'approccio statistico. Difatti la sua

---

[25] È importante sottolineare che quando si fa riferimento alla credibilità dei dati osservati, non si sta sostenendo che essi in qualsiasi contesto vengano utilizzati non saranno comunque mai in grado di andare a supportare quanto affermano; la credibilità va sempre valutata rispetto ad un determinato contesto, una determinata situazione, quindi ogni qualvolta si è in presenza di nuovi dati è possibile che il valore attribuito ad essi cambi.

introduzione derivava proprio dalla necessità di poter studiare ambiti particolari in cui le normali formule non potevano essere applicate, oltre che per sopperire a quella che si potrebbe definire una carenza dell'impianto statistico nel dimostrare elementi facenti parte di un mondo diverso da quello matematico (DE FINETTI, 1964).

Forse la differenza fondamentale tra i due approcci appena esposti sta proprio nel modo di valutare le informazioni che si hanno a disposizione: perché se uno statistico difficilmente riesce ad effettuare i suoi studi dando un grosso peso a quelle che sono le informazioni note a priori, per la teoria della credibilità esse sono invece proprio il punto di partenza, ovvero la base che permette poi di inglobare le nuove informazioni raccolte.

Un ulteriore elemento di distinzione sta poi nel fatto di valutare singolarmente tutti i soggetti, come se ognuno di essi avesse delle caratteristiche proprie, che solo ad esso appartengono; ed è proprio questo elemento che ha portato gli studiosi negli anni a porre continuamente sotto esame tutte le nuove informazioni raccolte, come se ognuna di esse fosse in grado di suggerire qualcosa di nuovo in merito al soggetto assicurato.

Il risultato di quanto appena esposto è stato quindi un progressivo mutamento dal mondo statico, tipico dell'approccio statistico, a quello sempre più dinamico introdotto dalla teoria della credibilità (LONGLEY-COOK, 1962).

4.1.1. EXPERIENCE RATING E CREDIBILITY THEORY

Prima di focalizzarsi sulla teoria della credibilità e sul modo in cui si sono evolute le varie correnti di pensiero in merito, è utile richiamare il concetto, espresso nel capitolo precedente, di experience rating, ovvero la valutazione in base all'esperienza, in modo da chiarire il ruolo svolto nell'applicazione di questa teoria.

Si è a lungo parlato di variabili di classificazione, informazioni note a priori ed in che modo calcolare un premio su base collettiva, ma tutto ciò non rappresenta realmente il percorso di vita di un assicurato, si limita a descriverne solo la prima parte, o meglio, i primi anni; il passo successivo, ed anche il più importante ai fini assicurativi, può avvenire solo grazie ad un sistema di catalogazione dei diversi individui in base all'esperienza individuale, e perché ciò sia possibile ci si avvale dell'utilizzo generale del sistema di classificazione in base all'esperienza in cui vengono applicati tutti i metodi derivanti dalla teoria della credibilità.
La scelta di raccogliere informazioni sulla storia dell'assicurato ha lo scopo di far comprendere a quale profilo di rischio egli appartenga, così da avere costantemente la certezza di aver fatto pagare un premio adeguato.
Per capire l'importanza che assume questo metodo di valutazione, si può assumere di osservare un portafoglio composto da 10 assicurati con le medesime caratteristiche osservate a priori, a cui per il primo anno si è scelto di far pagare un premio collettivo $\Pi = 0{,}25$.
Dopo il primo anno si ipotizza una situazione pari a quella illustrata nella tabella seguente[26], dove solo due assicurati hanno avuto un incidente.

Tabella 4.1: Esperienza degli assicurati dopo un anno

ANNI	ASSICURATI									
	1	2	3	4	5	6	7	8	9	10
1					1			1		

Fonte: NORBERG (1979)

[26] Nella tabella 4.1 e 4.2 vengono indicati solo il numero sinistri avvenuti a carico di ciascun soggetto; l'assenza di un numero nella casella sta ad indicare che quell'anno l'assicurato non ha fatto incidenti.

Il costo medio osservato dei sinistri sarà pari a $\overline{\mathrm{X}} = \frac{numero\ sinistri\ osservati}{numero\ totale\ osservazioni} = \frac{2}{10} = 0.2$, quindi il premio assegnato in precedenza risulta corretto ed in grado di coprire le spese che dovrà sostenere la compagnia.

Tabella 4.2: Esperienza degli assicurati dopo due anni

ANNI	ASSICURATI									
	1	2	3	4	5	6	7	8	9	10
1					1			1		
2					1					

Fonte: NORBERG (1979)

In realtà nemmeno osservando l'anno successivo non si può andare a criticare il premio collettivo, perché la media dei sinistri riporta un valore $\overline{\mathrm{X}} = \frac{3}{20} = 0.15$, ancora sufficientemente basso per non mettere in difficoltà l'assicuratore; la situazione invece cambia se la visione diventa più ampia, e quindi come nell'esempio si vadano a studiare dieci anni.

Tabella 4.3: Esperienza degli assicurati dopo dieci anni

ANNI	ASSICURATI									
	1	2	3	4	5	6	7	8	9	10
1					1			1		
2					1					
3	1			1					1	1
4		1				1				1
5		1						1	1	1
6	1	1						1		1
7	1					1		1		1
8		1				1		1		1

9		1								
10				1			1			
$\overline{X}_i$	0,3	0,5	0,0	0,2	0,2	0,3	0,1	0,5	0,2	0,6
$\overline{X}$	0,29									

Fonte: NORBERG (1979)

Infatti, si nota subito che l'ammontare medio del costo dei sinistri $\overline{X} = \frac{29}{100} = 0{,}29$ non risulta più così equo, ed allo stesso tempo metterebbe in difficoltà la compagnia assicurativa. Confrontando i diversi valori di $\overline{X}_i$, e quindi l'ammontare riferito a ciascun individuo, si nota inoltre che non è corretto imporre a tutti lo stesso premio; nonostante secondo una prima classificazione sembrava potessero tutti stare sullo stesso piano, andando ad osservare un periodo più ampio si è osservata una situazione differente.
La motivazione di tutto ciò è riconducibile alla presenza di eterogeneità all'interno delle classi di rischio individuate; da ciò deriva che la compagnia potrebbe decidere di adottare un numero elevato di variabili di classificazione, od essere addirittura portata a creare un numero consistente di sottoclassi, ma a prescindere da ciò rimarrà sempre l'elemento di disomogeneità. Quindi si deduce che per poter individuare il premio universalmente corretto per ciascun assicurato è indispensabile ricorrere ad una valutazione in base all'esperienza (GOULET, 1998).

4.2. I DIVERSI APPROCCI ALLA TEORIA DELLA CREDIBILITA'

Per inglobare le informazioni raccolte sull'esperienza dei diversi assicurati, negli anni si sono sviluppate diverse metodologie: dal noto sistema bonus-malus, al sistema merito-demerito, per poi passare a sistemi alternativi che prevedono ad esempio delle commissioni in riassicurazione. Si osserva comunque che nonostante nei vari paesi si possano riscontrare diversità nell'utilizzo delle informazioni raccolte, la maggior parte dei modelli adottati si basa sulla teoria della credibilità.

Il merito per l'introduzione nel mondo attuariale di questa teoria va attribuito a Mowbray che attraverso il suo paper "How Expensive a Payroll Exposure is Necessary to Give a Dependable Pure Premium", scritto nel 1914, va a chiarire quali siano le caratteristiche che devono possedere le informazioni rilevate per essere definite credibili, e quindi poter essere utilizzate per calcolare il premio assicurativo.

In realtà il suo intervento si riferisce alla politica salariale dei lavoratori nei confronti delle assicurazioni sanitarie obbligatorie[27], quindi un mondo ben distante da quello in cui attualmente trova spazio tale concetto, ma fin dall'inizio ciò non ha impedito di vederne la sua utilità.

Mowbray analizza per ogni singolo lavoratore la sua storia passata: se l'esperienza dimostrata varia in maniera moderata in ogni periodo, allora si è in presenza di dati che rispondono al requisito di piena credibilità.

Andando ad analizzare t periodi, si va ad osservare un numero N_t di sinistri, dove ognuno comporterà un valore Y_{jt} di indennizzo, per un totale di risarcimenti per periodo pari a X_t; quindi si deduce che $X_t = Y_{1t} + Y_{2t} + \cdots + Y_{Nt}$, dove i singoli risarcimenti sono per ipotesi tutti indipendenti ed identicamente distribuiti (i.i.d.).

Ricordando[28], inoltre, che le espressioni di media e varianza del risarcimento totale sono:

$$E(X_t) = E(N_t)E(Y_{jt})$$

$$Var(X_t) = E(N_t)Var(Y_{jt}) + Var(N_t)E(Y_{jt})^2,$$

[27] Nonostante il nostro Paese non preveda questo tipo di assicurazione, moltissimi paesi nel mondo prevedono che per poter usufruire delle prestazioni sanitarie un soggetto debba essere in possesso di questa forma di assicurativa.

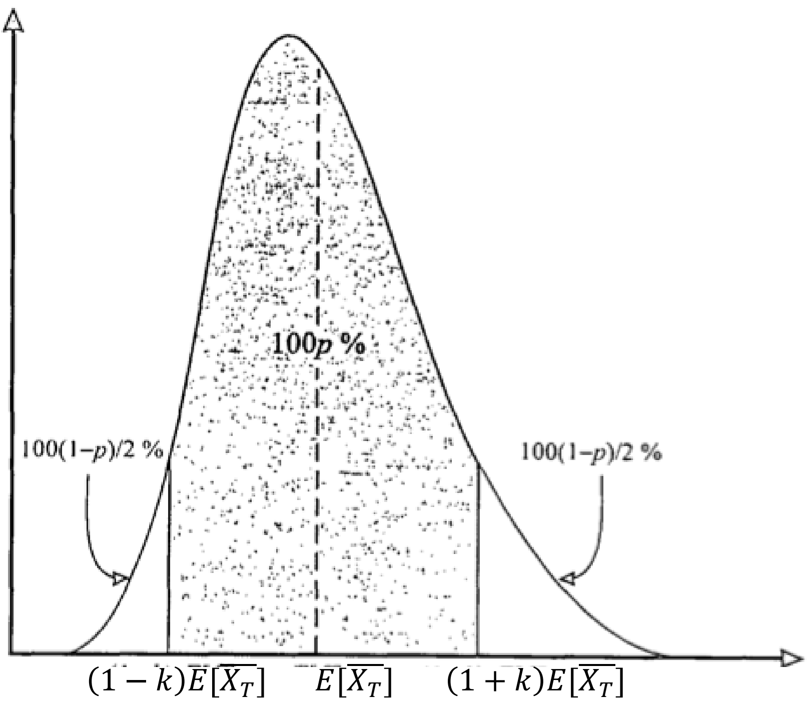
[28] Come già esposto nel capitolo precedente.

la soglia proposta da Mowbray può essere formalizzata attraverso l'espressione

$$P\left[(1-k)E[\overline{X_T}] \leq \overline{X_T} \leq (1+k)E[\overline{X_T}]\right] \geq p^{29}$$

In pratica, qualora andando ad analizzare la singola posizione di un assicurato si rilevi che il risarcimento totale ha un valore che con probabilità p, si distribuisce per un valore percentuale pari a k attorno al suo valore medio, allora la sua esperienza avrà piena credibilità. Il vantaggio che deriva dall'essere definito un assicurato con credibilità di ordine (k, p) sta nel fatto che il premio puro su cui si basa il premio di tariffa che dovrà pagare sarà calcolato esclusivamente sulla base del comportamento dimostrato, senza alcun riferimento alla classe di appartenenza (e quindi al premio collettivo) (MOWBRAY, 1914).

Figura 4.1: Soglia indicata da Mowbray



[29] In questa espressione si utilizza il risarcimento medio, e quindi $\overline{X_T} = \frac{(X_1+X_2+\cdots+X_T)}{T}$. La scelta di utilizzare il risarcimento per ogni periodo non è vincolante, spesso infatti si preferisce studiare il valore atteso del numero di sinistri.

Non essendoci un criterio fisso per assegnare il corretto valore di k e p, la soglia può essere più o meno ampia e come conseguenza anche i valori ottenuti potranno oscillare, come dimostra l'esempio seguente.

Tabella 4.4: Esempio numerico riguardante il numero di sinistri al variare dei valori assegnati di k e p

Numero N di sinistri		Probabilità attribuita a p		
		99%	95%	90%
Massimo scostamento dal valore medio (k)	2,50%	10,623	6,147	4,326
	5%	2,656	1,537	1,082
	7,50%	1,18	683	481
	10%	664	384	271

Fonte: LONGLEY-COOK (1962)

In precedenza abbiamo definito il termine X_t come somma di variabili casuali indipendentemente e identicamente distribuite, quindi, sfruttando il teorema del limite centrale[30], possiamo approssimare la sua distribuzione di probabilità con quella della variabile casuale Normale Standardizzata:

$$\frac{(\overline{X_T}-E(\overline{X_T}))}{\sqrt{Var(\overline{X_T})}} \sim N(0,1).$$

A questo punto è possibile reinterpretare il vincolo esposto da Mowbray da un punto di vista statistico attraverso l'espressione

$$Pr\left[\left|\frac{(\overline{X_T}-E(\overline{X_T}))}{\sqrt{Var(\overline{X_T})}}\right| \leq \frac{kE(\overline{X_T}))}{\sqrt{Var(\overline{X_T})}}\right] \approx 2\phi\left(\frac{kE(\overline{X_T}))}{\sqrt{Var(\overline{X_T})}}\right) - 1 \geq p.$$

[30] Il teorema del limite centrale sostiene che la somma, o la media, di variabili casuali indipendenti e identicamente distribuite possieda una distribuzione di probabilità che converga sempre ad una distribuzione centrata sulla media, e quindi alla variabile casuale Normale (PICCOLO, 2004).

Da ciò si deduce che per parlare di piena credibilità e quindi poter utilizzare i dati raccolti sull'esperienza individuale è necessario che il quadrato della media dei risarcimenti sia rispettivamente

$$[E(\overline{X_T})]^2 \geq \left(\frac{\alpha_{1-\varepsilon/2}}{k}\right)^2 \frac{Var(\overline{X_T})}{T},$$

dove il termine ε è il complemento a 1 della probabilità p, mentre α_i è l'i-esimo percentile della distribuzione normale (GOULET, 1998).

Nonostante attraverso questo tipo di distribuzione sia semplice interpretare il fenomeno, questo tipo di analisi può comunque essere svolta avvalendosi di altre forme distributive, come ad esempio la distribuzione Bernoulliana in cui la probabilità che si verifichi l'evento $X_t = 1$ è pari a θ, mentre l'evento contrario ($X_t = 0$) vedrà associata la probabilità $1 - \theta$.

Attraverso questo tipo di distribuzione, di media e varianza rispettivamente pari a $E(\overline{X_T}) = \theta$ e $E(\overline{X_T}) = \frac{\theta(1-\theta)}{T}$, la soglia sarà espressa in termini di T, ovvero il numero di eventi che si devono andare ad osservare, così come segue

$$T \geq \left(\frac{\alpha_{1-\varepsilon/2}}{k}\right)^2 \frac{1-\theta}{\theta}.$$

Qualora invece attraverso l'analisi emerga che ogni assicurato può incorrere in un numero n di sinistri all'anno, dove ognuno si verifica in modo indipendente rispetto agli altri con probabilità θ, la distribuzione di X_t può essere meglio rappresentata attraverso la forma distributiva binomiale caratterizzata dai parametri (n, θ). Quindi si deriva che il vincolo sopra esposto sarà determinato dal numero di sinistri e dai periodi di osservazione, attraverso l'espressione

$$nT \geq \left(\frac{\beta_{1-\frac{\varepsilon}{2}}}{k}\right)^2 \frac{1-\theta}{\theta}.$$

Un'ultima forma distributiva che è doveroso citare dato il suo ampio utilizzo in ambito assicurativo, è la distribuzione di Poisson. Perché ciò sia possibile è necessario conoscere anche il numero annuale λ di sinistri,

mentre la media e la varianza per il risarcimento totale saranno rispettivamente $E(X_t) = \lambda E(Y_{jt})$ e $Var(X_t) = \lambda E(Y_{jt}{}^2)$.
Il vincolo riformulato attraverso l'utilizzo di questa distribuzione assume la forma (GOULET, 1998)

$$nT \geq \left(\frac{\beta_{1-\frac{\varepsilon}{2}}}{k}\right)^2 \left[1 + \frac{Var(Y_{jt})}{E(Y_{jt})^2}\right] .$$

Utilizzando quanto appena esposto come punto di partenza, negli anni si sono andate poi a sviluppare varie correnti che hanno permesso differenti interpretazione del concetto di credibilità. Principalmente si fa riferimento a due diversi approcci: il primo che si è sviluppato, definito limited fluctuaction credibility, si focalizza sul concetto di esperienza stabile mentre il secondo è l'approccio greatest accuracy credibility, che preferisce una visione più completa del portafoglio di assicurati, spostando maggiormente l'attenzione sul concetto di omogeneità.
Nonostante si tratti di aspetti diversi attraverso l'utilizzo dello stesso concetto, questo non ha impedito negli anni lo svilupparsi delle diverse correnti, ragion per cui tuttora analizzando i diversi paesi si nota che non vi è omogeneità nell'utilizzo delle formule legate alla credibilità (AA.VV, 2008).

4.2.1. LA LIMITED FLUCTUATION CREDIBILITY

Il primo dei due approcci citati, indicato anche col termine *American Credibility* grazie al suo sviluppo e all'utilizzo prevalentemente nel territorio americano, si sviluppa poco dopo l'uscita del paper di Mowbray come critica al concetto di credibilità che egli stesso aveva proposto.
La prima critica mossa derivava dall'eccessiva rigidità del vincolo proposto: è chiaro che solamente i soggetti che dimostravano un'esperienza limitatamente stabile potevano godere di un premio che riflettesse le informazioni raccolte sulla loro esperienza, beneficio di cui non potevano

godere i soggetti che anche solamente di poco non rispettavano il vincolo imposto; questa situazione portava a notevoli malcontenti, poiché soggetti con esperienza pressoché simile si ritrovavano a pagare premi ben diversi tra loro.

Per evitare che si creasse questo disequilibrio all'interno dei portafogli assicurativi, si cercò di reinterpretare il concetto di credibilità come una sorta di collegamento tra due concetti che nel mondo assicurativo vivono pari passo: rischio ed esperienza (GOULET, 1998).

Consapevoli quindi che la piena credibilità è un concetto troppo limitativo, per non dire utopico, si preferisce parlare in termini di credibilità parziale, permettendo quindi di andare ad utilizzare anche i dati che fino a prima non venivano considerati perché mancavano di stabilità; sicuramente fu un passo in avanti, ma che portò ad un ulteriore problema: in che modo combinare le due tipologie di informazioni.

La prima soluzione proposta prevedeva di assemblare la probabilità di sinistro p_0 calcolata sui dati già in possesso e la probabilità p_1 valutata sulla base dei nuovi dati in base all'espressione

$$p = \frac{I_0 + I_1}{E_0 + E_1},$$

in cui p_0 e p_1 sono state calcolate come rapporto tra sinistri denunciati e polizze stipulate, quindi secondo la formula $p_0 = \frac{I_0}{E_0}$ e $p_1 = \frac{I_1}{E_1}$.

Per quanto apparentemente corretta, questa soluzione non ha riscosso grande successo in quanto poneva ulteriori problemi: non solo l'impossibilità talvolta riscontrata a reperire i dati necessari, ma anche per il fatto che in caso di piena credibilità dei nuovi dati raccolti, la formulazione impediva che $p = p_1$ (LONGLEY-COOK, 1962).

La miglior interpretazione di credibilità parziale, tuttora adottata da coloro che seguono questo tipo di approccio, viene quindi offerta da Whitney (1918); egli infatti, attraverso una profonda analisi del problema, suggerisce in che modo utilizzare sia le informazioni relative al rischio che venivano rilevate

attraverso le variabili tariffarie (età, residenza,…) che quelle rappresentate dall'esperienza del singolo individuo.
La complessità nel dover considerare contemporaneamente i due rischi citati deriva dal fatto che essi, per dare una rappresentazione coerente del soggetto, dovranno assumere pesi diversi; da ciò deriva che per valutare quanto possa influire l'esperienza del singolo piuttosto che le informazioni emerse al momento della sottoscrizione della polizza, è opportuno andare ad osservare quali sono i quattro elementi principali che fanno spostare l'ago della bilancia: l'esposizione, il moral hazard, l'omogeneità della classe di rischio ed il premio collettivo.
Quando un soggetto decide di assicurare un rischio particolarmente ingente, l'attenzione si sposterà dalla classe a cui appartiene a quella della sua esperienza passata: se con un rischio relativamente basso il premio sarà comunque in linea con quello della classe tariffaria, all'aumentare dell'esposizione vi sarà un progressivo aumento del peso attribuito al profilo di rischio dimostrato dal soggetto. Per quanto riguarda invece la concentrazione dei rischi, e quindi l'omogeneità all'interno della classe, si osserva il comportamento contrario: finché non si notano grossi scostamenti rispetto alla media non ha senso penalizzare il singolo, e così allo stesso modo avverrà anche quando il premio collettivo già ingloba al suo interno un numero sufficiente di dati.
La trattazione della propensione al rischio invece non ha una regola generale da seguire, ma è molto discrezionale; ogni caso va valutato separatamente (WHITNEY, 1918).
Sulla base delle premesse sopra enunciate, nel paper viene formulato il premio di credibilità come una media ponderata del premio collettivo per la classe di rischio P_0 e del premio basato sull'esperienza individuale P_1[31] rilevato in un secondo momento, strutturandolo nel modo seguente:

[31] La scelta di inserire all'interno della formula, e quindi attribuire due pesi, al profilo emerso dalla collettività e quello emerso dal singolo individuo non è vincolante; si poteva

$$P = ZP_1 + (1 - Z)P_0.$$

La novità introdotta attraverso questa espressione sta proprio nel termine Z, ovvero il fattore di credibilità, che permette di attribuire un peso diverso ai due rischi rilevati; Whitney propone inoltre di calcolare tale fattore attraverso la formula

$$Z = \frac{n}{(n+K)},$$

in cui n corrisponderà al numero di nuovi dati raccolti, mentre il termine K varierà in funzione del giudizio degli attuari.

Semplicemente osservando la formula è facile intuire che il valore di Z può oscillare da 0 a 1, ed in particolare nel caso in cui Z assuma valore pari a 1, ci si troverà nel caso in cui i nuovi dati raccolti (ovvero l'esperienza) avranno piena credibilità e di conseguenza $P = P_1$; intuitivamente si comprende che qualora Z risulti pari a 0 si preferirà procedere con lo studio del premio individuale solo sulla base dei dati già precedentemente noti grazie alle variabili di classificazione, e quindi si avrà $P = P_0$ (DE FINETTI, 1964).

Negli anni sono state elaborate anche altre formulazioni del fattore di credibilità, coinvolgendo anche il numero N di dati necessari per avere piena credibilità, come nelle espressioni seguenti

$$Z = (1 + k)\frac{n}{(n+kN)} \quad \text{o} \qquad Z = \sqrt{\frac{n}{N}},$$

ma la formulazione proposta da Whitney continua a conservare un ruolo chiave per la Limited Fluctuation Credibility.

Nell'esempio che segue si utilizza esattamente quest'ultima definizione di Z per dimostrare in che modo si possa distribuire il fattore di credibilità, dato un valore K convenzionale pari a 500:

---

infatti confrontare anche semplicemente le informazioni relative al passato con quelle considerate "nuove".

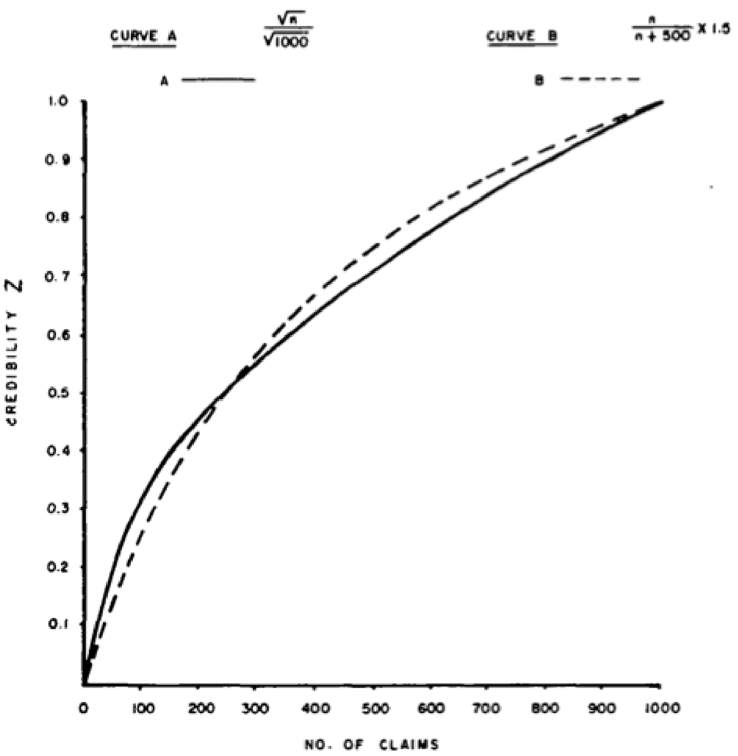
Tabella 4.5: Esempio di calcolo del fattore di credibilità

Numero di sinistri (n)	$Z = \frac{n}{n + 500}$
1000	1,00
900	0,96
800	0,92
700	0,87
600	0,82
500	0,75
400	0,67
300	0,56
200	0,43
100	0,25
0	0

Fonte: LONGLEY-COOK (1962)

Se si fosse scelto di utilizzare un'altra formulazione per il fattore di credibilità Z, si sarebbero sicuramente ottenuti altri risultati, ma a prescindere da ciò, attraverso tali dati si sarebbe comunque giunti alla stessa forma grafica, ovvero una curva (come illustra la figura 4.2).

Figura 4.2: Rappresentazione grafica della curva di credibilità



Fonte: LONGLEY-COOK (1962)

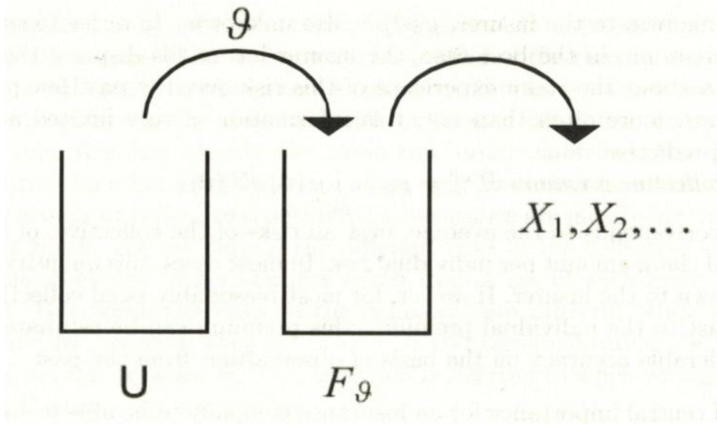
Come già accennato in precedenza, nonostante la semplicità dei calcoli prevista da questo tipo di credibilità, le obiezioni mosse verso questo approccio ne hanno comportato una diffusione limitata, rimanendo quindi in territorio americano. Le obiezioni a cui si fa riferimento riguardano innanzitutto il termine K proposto da Whitney: un termine soggettivo che si basa sulla discrezione degli attuari. Bisogna poi sempre ricordare che questa corrente nasce con lo scopo di valutare quali dei dati relativi all'esperienza individuale possano essere così significativi da venire inclusi nel calcolo del premio; perché si possa parlare di premio realmente equo, e quindi utilizzare i dati per far sì che ogni individuo paghi il premio giusto, bisogna fare un passo ulteriore (GOULET, 1998).

4.2.2. LA GREATEST ACCURACY CREDIBILITY

A differenza della limited fluctuaction credibility, all'interno di questa corrente non si colloca un solo tipo di approccio a quella che è la materia assicurativa, ma nel tempo si sono sviluppate diverse formulazioni, tutte con lo stesso fine ultimo di analizzare l'eterogeneità presente all'interno del portafoglio assicurativo, così da creare un premio il più aderente possibile al profilo di rischio evidenziato dal soggetto durante il periodo di copertura[32].
Già nel capitolo precedente abbiamo introdotto il concetto di parametro di rischio, indicandolo come il fattore che permette di andare ad attribuire ad ogni singolo soggetto un valore che sia in grado di riflettere la sua rischiosità. Sulla base di ciò dobbiamo, quindi, immaginare l'intera popolazione rappresentata attraverso la distribuzione definita cumulativa o strutturale $U(\theta)$, rappresentata spesso come un'urna contenente tutti i possibili valori di rischio, dove al suo interno vadano a distinguersi tutti i singoli profili di rischio θ, caratterizzati ognuno dalla sua funzione di distribuzione $F(\theta)$. Inoltre, come già enunciato, il singolo parametro θ è solo una delle possibili realizzazioni della variabile casuale $\Theta = \theta, X_1, X_2, \ldots$ (BUHLMAN & GISLER, 2005).

[32] Per ulteriori approfondimenti sulle varie formulazioni a cui si è accennato, si rimanda a Goulet (1998).

Figura 4.3: Modello a due urne



Fonte: BUHLMAN & GISLER (2005)

Sulla base di questi elementi si vanno a distinguere quattro tipologie di premio puro (o netto): il premio per il rischio, il premio collettivo, il premio Bayesiano ed il premio di Credibilità o di Bühlmann.

Il premio collettivo è semplicemente il premio a cui si è fatto riferimento fin dall'inizio come l'ammontare che viene fatto pagare agli assicurati quando non si è ancora in possesso di un numero sufficiente di dati; indicato attraverso l'espressione

$$m = E(X) = E[\mu(\Theta)],$$

si nota chiaramente che viene calcolato come media dei possibili premi per il rischio.

Questo, che potremmo definire il primo approccio che si ha verso un soggetto neoassicurato, è un premio comunemente adottato da entrambi gli approcci, quindi sulla base di questo elemento non possono essere riscontrate differenze. La vera differenza fra i due approcci emerge però al momento in cui si introduce il premio individuale; la limited fluctuaction credibility, proprio in base a quelle che sono le sue caratteristiche di base, non sempre ammette l'esistenza di questa tipologia di premio: perché esso vi sia, infatti, è necessario che il soggetto abbia dimostrato, nel periodo di osservazione, un'esperienza stabile; in caso contrario sarà ammesso solamente il premio collettivo m (GOULET, 1998).

La greatest accuracy credibility parte invece da un punto di vista ben diverso: perché sia possibile parlare di premio individuale, basterà conoscere l'effettivo livello di rischio attribuibile ai diversi soggetti; noto quest'ultimo elemento e l'ammontare totale di risarcimenti per il periodo di tempo considerato, basterà calcolarne il valore medio attraverso la formula del valore atteso condizionato:

$$\mu(\theta) = E(X|\Theta = \theta).$$

Nonostante sia possibile riassumere il concetto attraverso la semplice formula appena esposta, non è altrettanto semplice andare a calcolare l'effettivo valore che essa può assumere. La difficoltà deriva dal fatto che il parametro Θ, a differenza degli altri dati, non è direttamente osservabile; questo limite ha portato negli anni a sviluppare diverse metodologie di stima, con l'obiettivo di calcolare uno stimatore per $\mu(\theta)$ che portasse ad un valore più vicino possibile al quello che è il valore reale. I metodi che hanno acquisito maggior valore col tempo, e che tuttora occupano una posizione di spicco, sono quello riferito alla teoria Bayesiana, ed il modello proposto da Bühlmann; si utilizza preferibilmente il primo quando si tratta di analizzare danni riferiti al ramo vita, mentre il secondo viene considerato lo strumento più adatto per studiare i dati riferiti al ramo danni.

Nel seguito si andranno a confrontare i due diversi approcci; pur partendo da un punto di vista ben diverso, e quindi con diversi metodi di calcolo, si dimostrerà che in determinate condizioni le due metodologie andranno a divergere.

4.2.2.1. L'APPROCCIO BAYESIANO

Considerato da molti come l'approccio alla credibilità per eccellenza, l'approccio Bayesiano consente di combinare informazioni relative al presente ed al passato basandosi esclusivamente sulle probabilità individuali e sul noto Teorema di Bayes[33] dal quale prende il nome. L'analisi sviluppata sulla base di questi elementi prevede un largo uso delle distribuzioni di probabilità, ed in particolare si parlerà di probabilità inverse: si fa riferimento a questa tipologia perché effettivamente grazie al teorema di Bayes non si cerca mai la conseguenza del problema, ma si vuole risalire alla causa; le domande a cui si cerca di rispondere, infatti, saranno l'ammontare di danni subiti dalla compagnia sulla base degli esiti ottenuti, piuttosto che il numero di incidenti che sono stati causati (HERZOG, 2008).

Prima di addentrarci nelle caratteristiche di questa tipologia di approccio, è utile far luce su alcuni elementi chiave, o meglio peculiarità, proprie di questo modo di procedere (MAHLER & DEAN, 2001).

- Si parla in termini di distribuzione condizionata, e quindi l'attenzione è posta non tanto sul singolo evento, ma sulla probabilità che esso si realizzi in concomitanza con altri elementi chiave; si fa quindi riferimento a due eventi A e B, per cui la probabilità condizionata sarà
$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$
- Sempre rimanendo in tema di distribuzioni, un altro elemento chiave è la distribuzione $U(\theta)$, ovvero la forma distributiva della variabile casuale incognita Θ; la definizione non è sempre stata così univoca, infatti, nei primi anni si assumeva generalmente che tale distribuzione fosse il risultato dell'espressione

$$U(\theta) = \frac{numero\ \ polizze\ presenti\ nel\ portafoglio\ con\ parametro\ di\ rischio < \theta}{numero\ totale\ delle\ polizze\ presenti\ in\ portafoglio},$$

[33] Il teorema di Bayes tratta la possibilità che, dati due eventi A e B, l'evento A si realizzi nell'ipotesi in cui si sia verificato l'evento B; relazione normalmente espressa attraverso la seguente formulazione:
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}.$$

dalla cui formulazione si comprende chiaramente che non si potevano ricavare molte informazioni, anzi, non andava nemmeno a considerare il numero di nuove polizze che entravano a far parte del portafoglio assicurativo.

Proprio a partire da questa carenza di informazioni, negli anni, è stata proposta una diversa interpretazione, tuttora utilizzata, di linea guida per ogni singolo profilo di rischio θ; si suppone, infatti, che ogni singolo profilo sia indipendentemente e identicamente distribuito rispetto agli altri profili di rischio, secondo la variabile casuale Θ con distribuzione U (SUNDT, 1984).

- Un altro concetto chiave riguarda il Teorema della Probabilità Totale, il quale sostiene che dati gli eventi $E_1, E_2, \dots$ partizione dello spazio ψ, per ogni altro evento A, in cui A è sottoinsieme di ψ, sarà valida la seguente espressione:

$$P(A) = \sum P(A|E_i)P(E_i) \, .$$

Quindi la probabilità di ogni singolo evento A sarà data dalla media aritmetica delle probabilità condizionate $P(A|E_i)$, ponderate per i pesi $P(E_i)$.

Dopo aver esposto gli elementi chiave che vanno a caratterizzare l'approccio Bayesiano, si andrà a chiarire in che modo utilizzare le formule proposte attraverso un esempio pratico, che non solo renderà più chiaro in che modo possono essere utilizzate le formule, ma andando a riproporre lo stesso esempio anche per il modello di credibilità lineare proposto da Bühlmann, renderà possibile un confronto fra i due doversi approcci (HERZOG, 1989).

Immaginiamo di avere a disposizione due dadi con la stessa probabilità di essere pescati: il primo, al quale sarà associato l'evento A_1 ogni volta che sarà lanciato, sarà caratterizzato da un faccia segnata e cinque no; con la dicitura A_2, invece, si indicherà il lancio del secondo dado, caratterizzato da tre facce segnate e tre prive di ogni simbolo.

Figura 4.4: I due dadi con rispettivamente una e tre facce segnate

In base al risultato ottenuto attraverso il lancio dei dati si può inoltre fare un'ulteriore distinzione; sarà indicato con S_i, dove $i = 1,2, \dots$, l' i-esimo lancio del dado che darà come risultato una faccia segnata, mentre verrà indicato con N_i il lancio in cui l'esito sarà una faccia non segnata.

Come si evince già graficamente, la probabilità di trovare attraverso il lancio del primo dato una faccia senza simboli sarà $P(N_i|A_1) = \frac{5}{6}$, mentre dal lancio del secondo dado sarà $P(N_i|A_2) = \frac{3}{6}$.

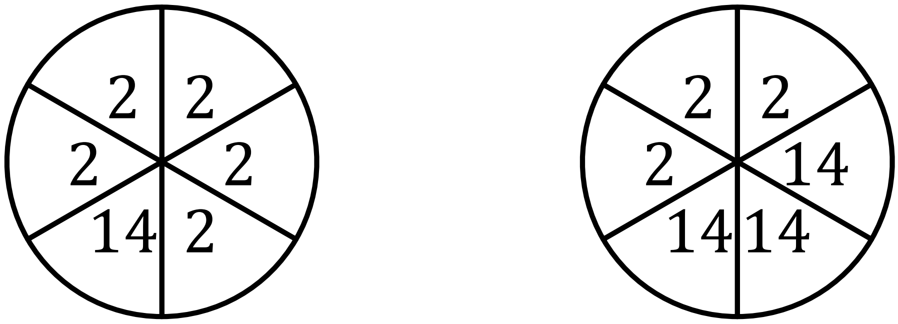
Perché attraverso l'utilizzo di esempi simili si possa andare a simulare il processo di catalogazione e aggiustamento dei premi in base alle caratteristiche emerse dai diversi soggetti, è opportuno inserire un ulteriore elemento; attraverso il lancio dei dadi è possibile distinguere un evento positivo o negativo, mentre per stabilire l'ammontare del danno in caso di esito positivo si ricorre ad una ruota numerata.

L'esempio a cui si fa riferimento prevede la presenza di due ruote, entrambe divise in sei settori, e dotate di una freccia che vada ad indicare il valore su di esse:

- quando sarà fatta girare la prima ruota, si farà riferimento all'evento B_1; i possibili esiti sono il numero "2" per cinque settori, mentre in solamente un settore è presente il numero "14";

- quando si andrà a far girare la seconda ruota, si farà riferimento all'evento B_2, ed i possibili esiti sono sempre "2" o "14", ma entrambi i numeri occupano tre settori.

Figura 4.5: Le due ruote dell'esempio



Sarà poi indicato con G_i, in cui $i = 1,2,\dots$, ogni i-esimo giro che verrà fatto fare alla ruota.

Combinando assieme i due esempi appena introdotti, si può andare a simulare l'operazione che normalmente viene svolta in ambito assicurativo, attraverso due passaggi:

1) Inizialmente un soggetto assicurato non si vede attribuito una maggiore o minore probabilità di compiere un sinistro rispetto a coloro che fanno parte della sua stessa classe di rischio, situazione che viene associata ai dadi del primo esempio, che posseggono uguale probabilità di essere scelti. Il lancio del dado, o meglio l'esito che deriva dal lancio è il punto di partenza per l'analisi: se la faccia è segnata questo corrisponderà ad un sinistro, se la faccia è priva di segni, allora al soggetto non verrà riconosciuto nessun sinistro. Solo qualora si verifichi la prima delle due situazioni, si procederà al giro della ruota per stabilire l'ammontare del danno causato.

   Il lancio dei dati è in questo modo paragonabile al periodo di osservazione degli assicurati: solo dopo ripetuti lanci, e quindi solo dopo l'i-esima osservazione, si potrà risalire all'esborso totale X_i; ripetute osservazioni per poter calcolare il premio netto iniziale, indicato con $E(X_1)$.

2) Il passaggio successivo prevede di aggiustare il premio appena ottenuto grazie agli esiti ottenuti nel secondo periodo di osservazione, e quindi andare a calcolare $E(X_1|X_2)$.

Procediamo quindi col primo passaggio andando a calcolare il premio netto iniziale attraverso i possibili esiti ottenibili dal giro della ruota.

Dato che gli unici esiti sono 0 (nel caso in cui non il dado abbia la faccia non segnata, non si procederà nemmeno al giro della ruota), 2 e 14 il suddetto premio sarà calcolato nel modo seguente

$$E(X_1) = 0P[X_1 = 0] + 2P[X_1 = 2] + 14P[X_1 = 14].$$

Come detto in precedenza, perché si abbia la condizione $X_1 = 0$, è necessario che il dado lanciato mostri un lato non segnato, quindi si deriva che

$$P[X_1 = 0] = P(N_1) = P(N_1|A_1)P(A_1) + P(N_1|A_2)P(A_2) = \left(\frac{5}{6} * \frac{1}{2}\right) + \left(\frac{3}{6} * \frac{1}{2}\right) = \frac{2}{3};$$

Per gli altri esiti si procede in modo analogo, tenendo però in considerazione che perché essi si realizzino è necessario che con il lancio del dado si ottenga una faccia segnata, e di conseguenza si dovrà tenere in considerazione anche l'esito della ruota.

$$\begin{aligned} P[X_1 = 2] &= P(S_1\ e\ G_1 = 2\ ) = P(S_1)\ P(G_1 = 2\ ) \\ &= \{P(S_1|A_1)P(A_1) + P(S_1|A_2)P(A_2)\} \\ &\quad * \{P(S_1 = 2|B_1)P(B_1) + P(S_1 = 2|B_2)P(B_2)\} \\ &= \left[\left(\frac{1}{6} * \frac{1}{2}\right) + \left(\frac{3}{6} * \frac{1}{2}\right)\right] * \left[\left(\frac{5}{6} * \frac{1}{2}\right) + \left(\frac{3}{6} * \frac{1}{2}\right)\right] = \frac{2}{9}\ , \end{aligned}$$

e

$$\begin{aligned} P[X_1 = 14] &= P(S_1\ e\ G_1 = 14\ ) = P(S_1)\ P(G_1 = 14\ ) \\ &= \{P(S_1|A_1)P(A_1) + P(S_1|A_2)P(A_2)\} \\ &\quad * \{P(S_1 = 14|B_1)P(B_1) + P(S_1 = 14|B_2)P(B_2)\} \\ &= \left[\left(\frac{1}{6} * \frac{1}{2}\right) + \left(\frac{3}{6} * \frac{1}{2}\right)\right] * \left[\left(\frac{1}{6} * \frac{1}{2}\right) + \left(\frac{3}{6} * \frac{1}{2}\right)\right] = \frac{1}{9}. \end{aligned}$$

Quindi, richiamando il concetto di teorema della probabilità totale, per ottenere il premio netto iniziale basta utilizzare la formula

$$E(X_1) = \sum_{i=1}^{2}\sum_{j=1}^{2} E[X_1 | A_i \cap B_j] P[A_i \cap B_j],$$

ed in questo modo ottenere i risultati riassunti attraverso la figura 4.6.

Figura 4.6: Tabella con gli esiti derivanti dal primo passaggio

POSSIBILI ESITI	PROBABILITA' INIZIALE	PRODOTTO DEL POSSIBILE ESITO CON LA SUA PROBABILITA'
0	2/3	0
2	2/9	4/9
14	1/9	14/9
TOTALE	1	2

Fonte: HERZOG (1989)

I diversi scenari possibili, riassunti attraverso l'espressione A_i e B_j, sono ottenuti combinando la frequenza di sinistri grazie all'esito ottenuto dal lancio dei dadi, e l'ammontare del sinistro attraverso l'esito emerso grazie alla ruota; il risultato che si ottiene è quindi un premio netto per ognuno degli scenari possibili.

Figura 4.7: Premio netto a priori

COMBINAZIONE DEI POSSIBILI EVENTI	FREQUENZA	SINGOLO RISARCIMENTO	PREMIO NETTO
A_1 e B_1	1/6	4	2/3
A_1 e B_2	1/6	8	4/3
A_2 e B_1	1/2	4	2
A_2 e B_2	1/2	8	4

Il passaggio successivo, nonché quello di maggiore interesse, permette di andare a studiare e poi combinare tra loro le nuove informazioni raccolte, che nell'ambito delle assicurazioni vengono riassunte normalmente attraverso l'espressione esperienza. Per mantenere la spiegazione sempre sul piano numerico, si vanno a riprendere gli esempi precedenti, e si cercherà di dimostrare come calcolare

$$E[X_2|X_1 = k] \qquad \text{con } k = 0, 2, 14 \ ,$$

ovvero, il premio a posteriori.

Dato che ogni esito ottenuto dal lancio dei dadi o dal giro della ruota è indipendente rispetto agli altri possibili esiti, si deduce che il valore di X_2 sarà funzione esclusivamente degli esiti di A_i e B_j e quindi della loro probabilità

$$P[A_i \cap B_j|X_1 = k].$$

Per raggiungere l'obiettivo prefissato, e quindi stimare il premio a posteriori sapendo che nel primo periodo è stato rilevato l'ammontare $X_1 = k$, sarà necessario risolvere l'equazione

$$E[X_2|X_1 = k] = 0P[X_2 = 0|X_1 = k] + 2P[X_2 = 2|X_1 = k] + 14P[X_2 = 14|X_1 = k]$$

Chiaramente il modo di procedere non è dei più semplici, quindi richiamando ancora la Teoria della Probabilità Totale, questa espressione può essere risolta attraverso la formula

$$P[X_2 = m|X_1 = k] = \sum_{i=1}^{2} \sum_{j=1}^{2} P[X_2 = m|\ A_i \ e \ B_j] P[A_i \ e \ B_j|X_1 = k] \ ,$$

dove $m = 0, 2, 14$.

Scomponendo in due parti l'equazione, si riassumono i risultati ottenuti attraverso la tabella 4.6.

Tabella 4.6: Esito dell'espressione $P[X_2 = m|A_i \; e \; B_j]$

COMBINAZIONE DEI POSSIBILI EVENTI	VALORE ASSUNTO DAL TERMINE m (IN TRENTASEIESIMI)			VALORE TOTALE ESPRESSO IN TRENTASEIESIMI	VALORE TOTALE ESPRESSO IN DECIMALI
	0	**2**	**14**		
$A_1 \; e \; B_1$	30	5	1	36	1
$A_1 \; e \; B_2$	30	3	3	36	1
$A_2 \; e \; B_1$	18	15	3	36	1
$A_2 \; e \; B_2$	18	9	9	36	1

Per risolvere la seconda parte dell'equazione, e quindi $P[A_i \; e \; B_j|X_1 = k]$, si fa invece riferimento alla classica formula del Teorema di Bayes:

$$P[A_i \; e \; B_j|X_1 = k] = \frac{P[X_1 = k|A_i \; e \; B_j]P[A_i \; e \; B_j]}{P[X_1 = k]} \; .$$

Tabella 4.7: Soluzione dell'equazione $P[A_i \; e \; B_j|X_1 = k]$ al variare del valore k

COMBINAZIONE DEI POSSIBILI EVENTI	VALORE ASSUNTO DAL TERMINE k		
	0	2	14
$A_1 \; e \; B_1$	0,3125	0,1562	0,0625
$A_1 \; e \; B_2$	0,3125	0,0937	0,1875
$A_2 \; e \; B_1$	0,1875	0,4687	0,1875
$A_2 \; e \; B_2$	0,1875	0,2812	0,5625

Grazie ai risultati appena esposti abbiamo la possibilità di andare a calcolare con quale probabilità è possibile che venga lanciato uno dei due dati, che venga girata una delle due ruote, e che da quest'ultima si ottenga un risultato ad esempio pari a 2. Ricorrendo, infatti, al Teorema di Bayes ricaviamo

$$P[A_2\ e\ B_1|X_1 = 2] = \frac{P[X_1 = 2|A_2\ e\ B_1]P[A_2\ e\ B_1]}{P[X_1 = 2]} = 0{,}4687\ \ ,$$

ovvero ci sarà una probabilità del 46,87% che la prima delle due ruote indichi il valore 2 grazie al lancio del dado con tre lati segnati.

Ma lo scopo di questo approccio è quello di andare a combinare tra loro dati di due periodi di osservazioni diverse, e quindi saper prevedere in che modo si distribuirà la variabile X_2 dato il valore di X_1.

Tabella 4.8: Soluzione dell'equazione $P[X_2 = m|X_1 = k]$ al variare di m e k

VALORE ASSUNTO DAL TERMINE m	VALORE ASSUNTO DAL TERMINE k		
	0	2	14
0	0,7083	0,5833	0,5833
2	0,1944	0,2951	0,243
14	0,0972	0,1215	0,1736
TOTALE	1	1	1

Grazie ai dati raccolti finora, ed in particolar modo quelli che riguardano la tabella 4.8, possiamo andare a fare le opportune sostituzioni all'interno della formula

$E[X_2|X_1 = k] = 0P[X_2 = 0|X_1 = k] + 2P[X_2 = 2|X_1 = k] + 14P[X_2 = 14|X_1 = k]$,

così da ottenere il valore atteso condizionale, che in ambito assicurativo corrisponde al premio netto corretto per il profilo di rischio individuale.

E quindi rispettivamente:

- $k = 0;\ E[X_2|X_1 = 0] = 0P[X_2 = 0|X_1 = 0] + 2P[X_2 = 2|X_1 = 0] + 14P[X_2 = 14|X_1 = 0] = 2 * 0{,}1944 + 14 * 0{,}0972 = 1{,}7496$;
- $k = 2;\ E[X_2|X_1 = 2] = 0P[X_2 = 0|X_1 = 2] + 2P[X_2 = 2|X_1 = 2] + 14P[X_2 = 14|X_1 = 2] = 2 * 0{,}2951 + 14 * 0{,}1215 = 2{,}2761$;
- $k = 14;\ E[X_2|X_1 = 14] = 0P[X_2 = 0|X_1 = 14] + 2P[X_2 = 2|X_1 = 14] + 14P[X_2 = 14|X_1 = 14] = 2 * 0{,}243 + 14 * 0{,}1736 = 2{,}9164$.

Il risultato, o meglio, il contributo che ha offerto questo modo di operare sui dati lo si può notare direttamente dal confronto col primo valore atteso stimato; da un valore di $E[X_1] = 2$, grazie all'influenza dei nuovi dati raccolti si osserva uno spostamento dei valori verso destra, a dimostrazione del fatto che se il premio netto iniziale sembrava adeguato per quella classe di rischio, in realtà non lo è; per fare in modo che si mantenga un giusto equilibrio all'interno della compagnia, i premi dovranno essere aggiustati di conseguenza (HERZOG, 1989).

Come già anticipato, i dati che si vanno a raccogliere e di conseguenza a combinare tra loro, possono avere natura diversa e di conseguenza sarà diversa anche la distribuzione di probabilità a seconda che si tratti di variabili numeriche o di classificazione.

Gli esempi fatti finora andavano a trattare variabili del primo tipo, quindi una categoria nella quale possono rientrare l'ammontare del risarcimento o ad esempio il numero di sinistri; spesso però ci si trova a valutare dati come la zona di residenza piuttosto che l'alimentazione della macchina, e nasce da qui la necessità di andare brevemente ad analizzare non solo le distribuzioni di probabilità ma anche le funzioni di densità di probabilità.

Definita $f_\theta(\theta)$ la funzione di densità di probabilità a priori di θ, qualora attraverso l'analisi su dei campioni casuali si rilevino dei dati diversi da inglobare all'interno della distribuzione, ad esempio $X = x$, la funzione assumerà la forma:

$$f_\theta(\theta|X = x) = \frac{f_x(x|\theta) f_\theta(\theta)}{f_x(x)}.$$

Correggendo quindi la stima del profilo θ attraverso i nuovi dati rilevati, è possibile passare dalla distribuzione a priori alla distribuzione a posteriori, ma per chiarire in modo opportuno ciò a cui si sta facendo riferimento, si ricorre nuovamente ad un esempio numerico.

Se per il periodo di copertura, fissato convenzionalmente pari ad un anno, si prevede che con probabilità pari a θ un assicurato farà un sinistro, la

funzione di densità di probabilità relativa a questa variabile sarà data da $f_\theta(\theta) = 1$, ed il valore di θ sarà compreso tra 0 e 1.
Se all'interno del portafoglio di assicurati viene estratto il Sig. Rossi, il quale ha compiuto esattamente un sinistro, sarà necessario includere l'informazione $X = 1$ all'interno della forma distributiva; per effettuare questo tipo di aggiornamento si dovrà procedere nel modo seguente (MAHLER & DEAN, 2001):

$$f_\theta(\theta|X = x) = \frac{f_x(X=1|\theta)f_\theta(\theta)}{f_x(1)} = \frac{(\theta)(1)}{\frac{1}{2}} = 2\theta.$$

4.2.2.2. IL MODELLO DI BUHLMANN

Un altro approccio alla credibilità, altrettanto noto, è quello che viene indicato come Credibilità dei minimi quadrati o Credibilità di Bühlmann.
A differenza di quanto esposto nel capitolo precedente, il modo di operare proposto non è così nuovo, anzi, egli ripropone l'espressione lineare di credibilità già esposta anni prima da Whitney,

$$P = zP_1 + (1 - z)P_0.$$

Il contributo che era stato offerto dallo scrittore, infatti, non era stato accantonato per l'incapacità di saper pesare le informazioni relative al presente o al passato, ma per la difficoltà riscontrata nel calcolo del giusto peso z da attribuire a tale espressione. La prima soluzione proposta per il calcolo del peso della credibilità si otteneva risolvendo l'equazione

$$z = \frac{n}{n + K},$$

dove n era il numero di osservazioni, mentre K variava in base al giudizio degli attuari. Nonostante nel territorio americano tuttora venga adottato questo metodo di calcolo per il fattore di credibilità, nel territorio europeo tale definizione ebbe meno fortuna: la scelta di inserire un elemento di soggettività all'interno di una formula così frequentemente utilizzata venne criticato al punto tale da mettere in cattiva luce anche la relazione di credibilità proposta.

Il primo a fornire un'adeguata soluzione al problema, proponendo un metodo di calcolo alternativo, fu Arthur Bailey (1945,1950); egli, infatti, non solo suggerì di calcolare K attraverso strumenti algebrici, ma al tempo stesso avanzò una pesante critica nei confronti dell'approccio Bayesiano. Quest'ultimo, infatti, non solo prevedeva il ricorso a numerosi calcoli matematici di non poca difficoltà, ma sviluppava l'insieme di formule proprie dell'approccio Bayesiano sulle distribuzioni a priori, distribuzioni che in realtà non potevano essere note. Non a caso la scelta di adottare senza remore questa idea, ed il fatto di assumere che tutte le possibilità siano equiprobabili [34], è stata per molti anni battezzata col nome di *Assunzione di Distribuzione di Equa Ignoranza.* La soluzione che egli propone è quindi quella di utilizzare esclusivamente la versione del teorema rielaborata da Laplace[35], definita talvolta come la versione generalizzata di Bayes, creando quindi l'opportunità di applicare tale approccio anche a dati che non presentano la stessa distribuzione di probabilità (BAILEY A. L., 1950).

Nonostante la critica mossa da Bailey avesse un fondamento, il pubblico non ancora pronto intellettualmente a recepire tale novità fece in modo che tale scoperta andasse a perdersi nel tempo, in particolare fino a quando non fu lo stesso Bühlmann a farla tornare alla luce.

L'autore, infatti, diede due grossi contributi alla teoria della credibilità: prima riuscì a dare un'adeguata espressione generale per la costante K, e poi per semplificare le procedure, propose uno strumento di calcolo che

[34] Peculiarità dell'approccio Bayesiano.

[35] Nota anche come regola di successione, permette di calcolare la probabilità che si realizzi un evento futuro dopo aver osservato un numero n di eventi simili. Dato H_i l'insieme dei possibili esiti, l'evento futuro E sarà dato da

$$P(E|n) = \sum_i P(H_i|n)P(E|H_i).$$

facesse un minor uso di capacità e competenze computazionali, e che al tempo stesso non andasse a richiedere l'utilizzo di distribuzioni a priori; il lavoro da egli prodotto viene normalmente individuato come il miglior stimatore nella classe degli stimatori lineari (HERZOG, 2008).
Lo studio di quello che ora viene definito " Parametro K della Credibilità di Bühlmann" inizia con un'analisi mai affrontata prima; abbandonando quindi la necessità di raccogliere molte informazioni, egli propone la formulazione seguente (HERZOG, 1989):

$$K = \frac{valore\ atteso\ del\ processo\ di\ varianza}{varianza\ della\ media\ ipotetica}.$$

L'espressione presente a numeratore, indicata anche con EPV (dall'inglese *Expected value of the Process Variance*[36]), si differenzia dal normale concetto di varianza perché si va ad osservare il valore della varianza condizionato ad una determinata combinazione di rischi (MAHLER & DEAN, 2001).
Ancora una volta, per rendere la teoria più comprensibile, si richiama l'esempio del lancio dei due dadi introdotto nel paragrafo precedente (HERZOG, 1989); considerato che il primo dei due dadi era caratterizzato da un lato segnato e cinque lati no, la varianza di un singolo lancio viene calcolata nel modo seguente:

$$n°\ di\ lanci * lati\ segnati * lati\ non\ segnati = 1 * \frac{1}{6} * \frac{5}{6} = \frac{5}{36}.$$

Il secondo dado, invece, era caratterizzato da tre lati segnati e tre privi di segni; sempre attraverso la stessa formula si ricava:

$$1 * \frac{1}{2} * \frac{1}{2} = \frac{1}{4}\ .$$

Poiché il processo di varianza prevede di pesare la varianza sopra calcolata in base alla tipologia di rischio considerata, sarà necessario conoscere la

[36] La varianza riferita al rischio viene distinta attraverso l'espressione "processo di varianza".

probabilità che ogni rischio si realizzi; avendo due dadi con la medesima probabilità pari a $\frac{1}{2}$ di essere pescati, il risultato sarà

$$EPV = \frac{1}{2} * \frac{5}{36} + \frac{1}{2} * \frac{1}{4} = \frac{7}{36}.$$

La varianza della media ipotetica, indicata anche con il termine VHM (dall'inglese *Variance of the Hypothetical Means*), potrebbe essere identificata come la varianza del valore atteso oltre il possibile scenario di diverse combinazioni di rischio. Quindi, sempre facendo riferimento allo stesso esempio, possiamo individuare due diverse combinazioni di rischi: l'ammontare medio di sinistri che corrisponde al primo dado è $\frac{1}{6}$, dado che presenta una sola faccia segnata; mentre per il secondo dado sarà $\frac{1}{2}$ in base al medesimo ragionamento. Poiché, come detto in precedenza, i due dadi hanno egual probabilità p, si deduce che il valore medio derivante dalla loro somma è pari a

$$\mu_{VHM} = \frac{1}{2} * \frac{1}{2} + \frac{1}{2} * \frac{1}{6} = \frac{1}{3}.$$

La varianza della media ipotetica quindi viene calcolata Raccolti i dati, per arrivare alla varianza della media ipotetica basterà risolvere l'equazione:

$$VHM = p * [varX_1 - (\mu_{VHM})^2] + p * [varX_2 - (\mu_{VHM})^2]$$
$$= \frac{1}{2}\left[\frac{1}{6} - \left(\frac{1}{3}\right)^2\right] + \frac{1}{2}\left[\frac{1}{2} - \left(\frac{1}{3}\right)^2\right] = \frac{1}{36}.$$

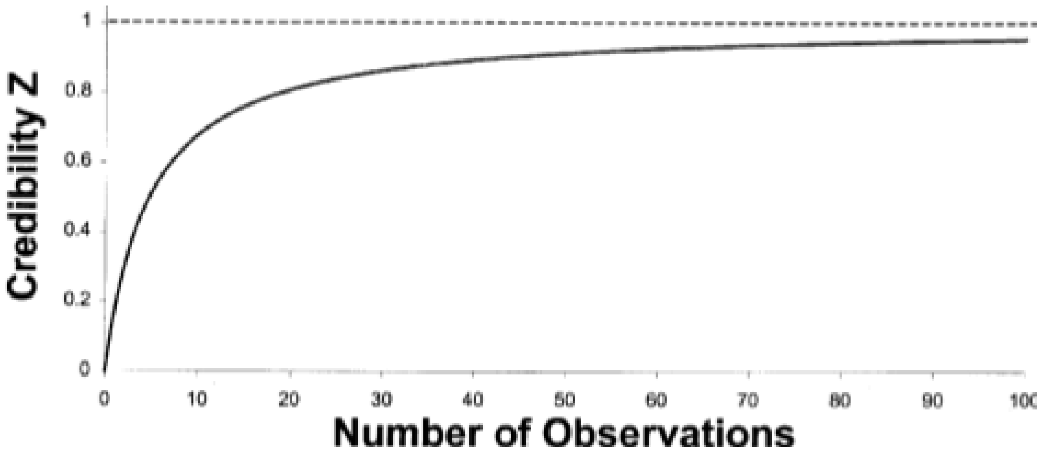
Avendo a disposizione il valore a numeratore ed a denominatore, per ottenere il valore della costante K basterà sostituire i valori all'interno dell'espressione, e quindi

$$K = \frac{EPV}{VHM} = \frac{\frac{7}{36}}{\frac{1}{36}} = 7.$$

Grazie al contributo offerto da Bühlmann, non solo si è andati a ripristinare il valore attribuito all'espressione di credibilità lineare inizialmente introdotto da Whitney, ma si è potuto dare maggior valore al parametro di credibilità che si caratterizza per essere (HERZOG, 1989):

- Funzione crescente rispetto al numero di osservazioni: all'aumentare del numero di dati raccolti il peso attribuito ai nuovi dati aumenta progressivamente fino a raggiungere il valore 1 di piena credibilità.

Figura 4.8: Z funzione crescente rispetto al numero di osservazioni



Fonte: MAHLER & DEAN (2001)

- Z è funzione decrescente rispetto al termine K, e di conseguenza anche rispetto al termine EPV; questo comporta che qualora la varianza oscilli fortemente a causa della combinazione di rischi presi in considerazione, verrà ridotto il peso attribuito ai nuovi dati, in favore delle osservazioni meno recenti.
- Z è funzione crescente rispetto al termine VHM, quindi se aumenta l'oscillazione del valore atteso dato dai diversi rischi, anch'esso aumenterà proporzionalmente.

L'intervento effettuato da Bühlmann non si è limitato a definire la questione tanto annosa dell'adeguato metodo di valutazione del parametro sopra citato, ma ha fornito anche un approccio di calcolo del premio a posteriori che tuttora viene definito come il premio di Credibilità di Bühlmann.

L'idea nasce dalla volontà di proporre un metodo che, a differenza di tutti quelli proposti fino ad allora, potesse provare la relazione di credibilità senza dover necessariamente fare riferimento alle funzioni di distribuzione che governano i rischi individuali, mantenendo pur sempre la dipendenza con i dati rilevati a priori; ma soprattutto uno strumento che non implicasse necessariamente l'ipotesi di indipendenza e identica distribuzione[37] delle variabili casuali, ma che si distinguesse per una maggior flessibilità così da poter essere applicato universalmente, a prescindere dalla forma assunta dai dati.

Egli dimostra quindi che la miglior approssimazione lineare dell'espressione $E[\mu(\theta)|X_1, X_2, \dots X_n]$, stimatore del premio a posteriori $\mu(\theta)$ dopo aver osservato un numero di variabili casuali $X_i$omogenee rispetto al tempo, si ottiene minimizzando l'espressione

$$E[a + z\bar{X} - \mu(\theta)]^2 \,,$$

rispetto alle variabili a e z.

Quindi, riscrivendo l'equazione nel modo seguente

$$E[z(\bar{X} - \mu(\theta))]^2 + E[a - (1 - z)\mu(\theta)]^2 \,,$$

si ricava che, per minimizzare l'espressione, i termini a e z devono essere rispettivamente pari a:

$$a = (1 - z)E[\mu(\theta)]$$

e

$$z^2 = E(\bar{X} - \mu(\theta))^2 + (1 - z)^2 Var[\mu(\theta)] \,,$$

da cui si ricava:

$$z = \frac{Var[\mu(\theta)]}{Var[\mu(\theta)] + E(\bar{X} - \mu(\theta))^2} \,.$$

---

[37] Bühlmann riassume le caratteristiche di indipendenza e identica distribuzione con il termine omogeneità, omogeneità che però lui individua in due forme diverse (BUHLMANN, 1967):

- omogeneità nella massa dei rischi, riferita quindi al fatto che le variabili si distribuiscano omogeneamente rispetto ad uno stesso tipo di rischio;
- omogeneità rispetto al tempo, dove invece si valuta se le diverse variabili sono distribuite omogeneamente rispetto alla variabile tempo.

Quindi, assumendo l'indipendenza e l'identica distribuzione delle variabili casuali X_i, si ricava che

$$E\left(\bar{X}-\mu(\theta)\right)^2=\frac{1}{n}E\left(X_1-\mu(\theta)\right)^2=\frac{1}{n}E[\sigma^2(\theta)]\,,$$

da cui si ottiene la seguente relazione di credibilità:

$$(1-z)E[\mu(\theta)]+z\bar{X}\,,$$

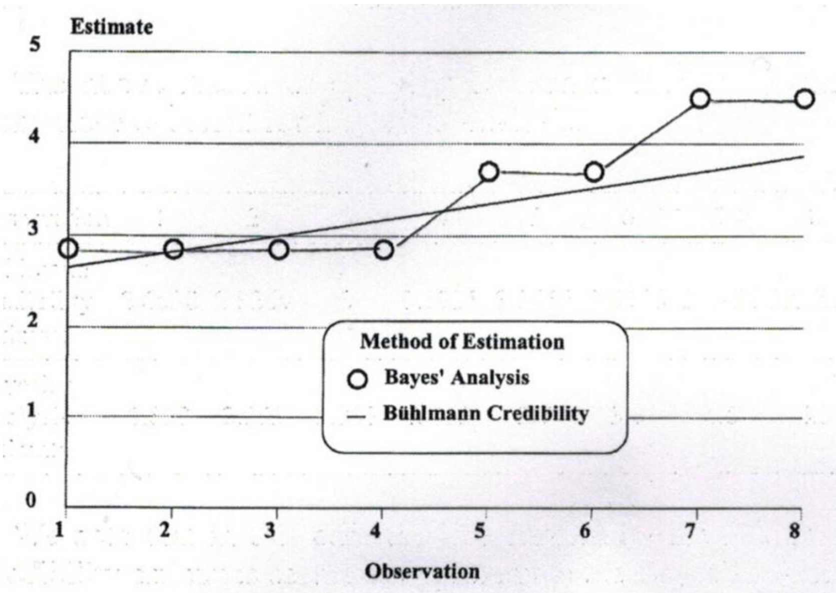
in cui z corrisponde al parametro o peso di credibilità e si ricava dall'ormai nota relazione $z=\frac{n}{n+K}$

4.2.2.3. LA CONVERGENZA TRA LO STIMATORE BAYESIANO E LA CREDIBILITA' LINEARE

Nonostante la formula di credibilità ottenuta attraverso l'applicazione dei due precedenti approcci sia apparentemente molto diversa, in realtà, grazie alla combinazione di particolari distribuzioni, si dimostra che i due premi valutati attraverso questa metodologia, in alcuni casi coincidono. La mistura di distribuzioni a cui si fa riferimento può riguardare ad esempio il modello composto da: due distribuzioni normali, la bernoulliana e la distribuzione Beta, ma forse la più utilizzata in ambito assicurativo è la mistura Poisson-Gamma.

Quest'ultimo modello, infatti, si compone di due elementi: la distribuzione di Poisson, che descrive il numero di sinistri data la frequenza sinistri, e la distribuzione Gamma, che spiega come si distribuisca quest'ultimo elemento all'interno della popolazione; grazie a questa impostazione si potrà andare a studiare il numero di sinistri per un assicurato scelto a caso all'interno della popolazione (Si rimanda per la dimostrazione a (MAHLER & DEAN, 2001)).

Figura 4.9: Confronto fra i risultati ottenuti attraverso l'approccio Bayesiano e le formule di credibilità proposte da Bühlmann, a fronte del medesimo numero di osservazioni



Fonte: MAHLER & DEAN (2001)

5. I SISTEMI BONUS-MALUS

5.1. CARATTERISTICHE PRINCIPALI

L'idea di creare un sistema di personalizzazione a posteriori come il sistema Bonus-Malus deriva dalla necessità di colmare il divario di conoscenze che separa il soggetto assicurato dalla compagnia di assicurazione; questo avviene perché, nonostante le variabili utilizzate per una classificazione a priori siano numerose e tra loro ben diversificate, nel ramo RCauto in particolare, vi sono elementi come la prontezza di riflessi, l'aggressività alla guida o la stessa conoscenza del codice della strada che non possono essere individuate tramite dei semplici questionari, e di conseguenza creano un elemento di eterogeneità all'interno delle classi tariffarie, tato da rendere necessaria un'ulteriore valutazione al fine di individuare il premio adeguato da far corrispondere ai diversi assicurati (BROUHNS, GUILLEN, DENUIT, & PINQUET, 2002).

Intorno alla metà degli anni '50 nasce quindi quella che è stata definita "personalizzazione in base all'esperienza", applicata attraverso diversi sistemi, che in base alle regole di transizione "one up/one down" o ad esempio "two up/one down", possono essere individuate all'interno dei modelli di experience rating, merit rating, no-claim discount, ma sicuramente il più diffuso nel territorio europeo è il sistema Bonus-Malus.

Quest'ultimo nasce in Svizzera nel 1963 ed è rappresentato come una scala suddivisa in 22 classi, tutte soggette alla regola per cui in assenza di sinistri l'assicurato aveva diritto di scendere di una classe, mentre nel caso contrario sarebbe salito di tre. La scelta di introdurre questo meccanismo deriva dalla volontà delle compagnie assicurative di utilizzare le formule di credibilità senza dover necessariamente ricorrere ai numerosi ed articolati calcoli da esse previsti, quindi attraverso la scala bonus-malus si trova la risposta commerciale, ed inevitabilmente più semplice, del complesso di formule introdotto nel capitolo precedente (OUTREVILLE, 1998).

In realtà, però, l'introduzione di questi sistemi nasce anche con lo scopo di incentivare i guidatori a mantenere una guida più virtuosa, che porti a ridurre il moral hazard; proprio per questo motivo negli anni si sono sviluppate diverse scale di valutazione, tutte accomunate dal principio di voler premiare attraverso uno sconto sul premio chi rientra nella categoria dei bravi guidatori, e di penalizzare, invece, con una maggiorazione chi dimostra una guida inaffidabile, per arrivare nel lungo termine ad avere un premio sempre più coerente con la sinistrosità dimostrata dal soggetto.

Come per tutte le novità, anche ad esso sono state avanzate delle critiche, ed in particolare c'è chi sostiene che tale sistema di personalizzazione sia andato ad intaccare quelli che sono i principi fondamentali che sottendono a tutto il sistema assicurativo (LEMAIRE, 1995). Si fa riferimento in particolare a:

- Lo stesso concetto di prestazione aleatoria, il quale è venuto a mancare; ora il premio che viene pagato dal soggetto che intende assicurarsi non è più aleatorio sotto tutti i fronti, ma in parte dipende da quella che è la sua storia di sinistrosità.
- Viene a mancare anche il concetto di equità, perché nel momento in cui un assicurato non compie sinistri otterrà uno sconto e quindi pagherà meno, ma per mantenere l'equilibrio finanziario, la compagnia assicurativa farà in modo di ottenere quei soldi da chi ha registrato dei sinistri con responsabilità.
- Per quella che è la legge dei grandi numeri, i singoli risarcimenti avevano un peso quasi nullo, tali da essere praticamente ininfluenti sull'operato della compagnia; invece il sistema Bonus-Malus porta continuamente a dover prendere in considerazione ogni singola denuncia di sinistro per poter progressivamente adeguare il premio.
- Inoltre, vi sono caratteristiche che possono subire modificazioni in base alla regolamentazione del paese in cui vengono applicate, come

ad esempio le regole che impongono che all'interno dello stesso paese vi sia la medesima tariffazione; in questo modo viene a mancare anche la competizione tra le diverse compagnie.

Non sono poi mancate le critiche anche dal punto di vista degli assicurati stessi; si è sollevata, infatti, la questione che questo tipo di approccio vada a penalizzare più del dovuto coloro che a priori già presentavano un profilo rischioso: essi non solo pagheranno un premio notevolmente maggiore rispetto ad altri, ma verranno collocati nella classi più alte della scala, dovendo in questo modo attendere molti anni prima di poter pagare un premio minore. A questo problema sono state offerte comunque delle soluzioni[38], ma eccessivamente difficoltose da mettere in pratica (BROUHNS, GUILLEN, DENUIT, & PINQUET, 2002).

Dall'introduzione del primo sistema BM (Bonus-Malus) a quello attualmente in vigore, vi sono stati diversi approcci e studi finalizzati alla ricerca della miglior formulazione, ed un grosso contributo per la correttezza e stabilità del sistema è da riconoscere a Norberg(1975) ed a Verico(2002). Invece è da attribuire a Lemaire il merito di aver interpretato il movimento degli assicurati all'interno delle classi di merito come una catena di Markov (MAGGINA, 2008); quest'ultimo elemento può essere definito come una successione di eventi casuali la cui probabilità condizionata dipenderà esclusivamente dallo stato precedente, e mai farà riferimento alla storia evolutiva precedente, chiamato per questo motivo un processo privo di memoria.

Il sistema BM viene normalmente descritto come un processo markoviano grazie a quattro caratteristiche che si riscontrano in esso (CERCHIARA, 2014):

[38] Una delle soluzioni, che prevedrebbe l'utilizzo all'interno della tariffazione a priori della variabile classe BM, verrà studiata all'interno del modello.

- La classe di merito che viene assegnata ad un soggetto dipenderà esclusivamente dalla classe occupata in precedenza, e non dal percorso che ha dovuto fare per trovarsi in tale classe;
- Il fatto che il passaggio all'interno del sistema non sia casuale ma segua determinate regole evolutive;
- La presenza di una classe di arrivo, situazione che tutti ambiscono a raggiungere poichè presenta il premio con il maggior sconto;
- La scelta di suddividere la scala così realizzata in un numero finito di classi, così da poter associare ad ognuna un coefficiente e sulla base di questo variare il premio.

Il fatto che questi sistemi possano essere riassunti attraverso un procedimento markoviano ne fa un ulteriore elemento di distinzione rispetto all'approccio Bayesiano introdotto nel capitolo precedente; infatti, non si farà una valutazione basandosi sul comportamento a priori, ma semplicemente si andrà a calcolare un coefficiente π_j che potrebbe essere definito medio per la classe di merito di appartenenza; ovvero, si farà una valutazione basata sul presente, e non sul passato.

Ogni sistema BM prevede, inoltre, la suddivisione in un numero finito di k livelli[39], ognuno dei quali permetterà l'assegnazione di un coefficiente per tutti i soggetti che facciano parte di una determinata classe di merito; un nuovo assicurato quindi avrà la possibilità di partire da un livello prestabilito, ed al termine di ogni anno la sua posizione potrà variare all'interno della scala sulla base del numero di sinistri riportato e delle regole di transizione previste per tale sistema.

Si comprende quindi che gli elementi che vanno a costituire tale modello di catalogazione in base all'esperienza non si esauriscono attorno al numero di livelli, ma richiedono di stabilire anche il livello di partenza per ogni nuovo

[39] La scelta di indicare in generale col termine k il numero di livelli non è casuale, ma dettata dal fatto che il numero massimo di livelli per i sistemi bonus-malus non è predefinito, ma può variare di paese in paese.

assicurato e le regole evolutive che seguirà il modello stesso, elementi che possono variare, come già anticipato, di paese in paese (NORBERG, 1976).

Nonostante sia arduo definire il sistema BM perfetto, e pensare che riesca a cogliere interamente tutti gli elementi di eterogeneità rimasti in seguito alla tariffazione a priori, sono state individuate tre proprietà che permettono di preferire un sistema rispetto ad un altro. Le proprietà stabiliscono che (GIGANTE, PICECH, & SIGALOTTI, 2000):

- Qualora due rischi appartengano a classi tariffarie diverse, rispettivamente x e y, con $\mu_x < \mu_y$, a parità di esperienza rilevata ci si aspetterà un coefficiente di aggiustamento del premio per la classe x maggiore; ovvero un fattore di aggiustamento a posteriori maggiore per la classe con numero di sinistri atteso, a priori, minore.
- Se a priori due premi sono diversi sulla base delle variabili tariffarie, ma negli anni si riscontra sempre la stessa esperienza per entrambi, col passare degli anni la differenza tra i due premi andrà a diminuire fino ad annullarsi.
- Infine, per il principio dell'equilibrio finanziario che vige all'interno delle compagnie assicurative, il premio totale atteso dovrebbe eguagliare la perdita totale attesa.

Si osserva comunque che, nonostante la compresenza di tutte le proprietà rappresenti una situazione ottimale, è ben difficile che esse vengano rispettate.

5.1.1. LA RESPONSABILITA' CIVILE AUTOVEICOLI (RCA)

Considerato che il sistema del quale si sta parlando vede nelle assicurazioni RCA il suo ambito applicativo per eccellenza, prima di andare ad analizzare le peculiarità dei diversi sistemi di personalizzazione, è necessario andare a chiarire a quale tipo di assicurazione essi fanno riferimento.

Il 24 Dicembre 1969, con la legge 990 viene introdotta "l'assicurazione obbligatoria della responsabilità civile derivante dalla circolazione dei veicoli a motore e dei natanti", ovvero una legge che obbligava coloro che utilizzavano mezzi a motore non solo a ritenersi responsabili per tutti i danni che potevano causare alla guida, ma soprattutto ad assicurarsi contro questa evenienza per poter circolare; sempre con la stessa legge però si sancisce anche l'obbligo delle compagnie assicurative a contrarre, e quindi ad essere informate e saper effettuare diverse proposte sulla base delle esigenze di chi vorrà sottoscrivere il contratto.

Parlando precedentemente di tariffe e di variabili tariffarie si è scelto deliberatamente di analizzare il sistema attualmente in vigore, sistema che gode di un regime di liberalizzazione delle tariffe grazie alla direttiva comunitaria 92/49/CEE; solo a partire da tale data, infatti, le assicurazioni hanno avuto la facoltà di scegliere liberamente in che modo operare la personalizzazione a priori ed a posteriori sul portafoglio di assicurati.

Un ulteriore sconvolgimento in ambito assicurativo è stato ad opera del decreto legge n. 223 del 2006, definita poi Legge Bersani; tra le diverse novità introdotte, infatti, quella che ha maggiormente cambiato il modo di operare delle assicurazioni riguarda il sistema Bonus-Malus. Grazie a questa legge, coloro che acquistavano delle nuove vetture potevano godere della stessa classe di merito che avevano in precedenza, beneficio però del quale potevano godere in generale tutti coloro che vivevano nello stesso nucleo familiare: un neo-patentato non doveva più entrare necessariamente in quattordicesima classe come faceva in precedenza, ma poteva slittare direttamente alla classe di merito minore presente nel nucleo familiare.

Il sistema, che aveva lo scopo implicito di agevolare le famiglie favorendo quindi l'acquisto di nuove vetture (e un rilancio dell'economia), ha portato ad un grosso squilibrio da un punto di vista del rischio: le compagnie, in realtà, si trovavano a dover coprire rischi considerevoli avendo in contropartita degli introiti minimi. Per questo motivo, a partire da tale data, si è osservato un progressivo disgregarsi degli elementi fondamentali del sistema Bonus-Malus, del quale rimane intatto solo il nome, ma i principi di fondo stanno cambiando irreparabilmente[40] (PORZIO, PREVIATI, COCOZZA, MIANI, & PISANI, 2011).

5.2. LE REGOLE EVOLUTIVE

Quando si parla di regole evolutive, si fa riferimento al percorso che dovrà compiere ogni soggetto sulla base del numero di sinistri con responsabilità che egli ha compiuto e segnalato; sarà proprio in conformità a questi, infatti, che egli avrà uno sconto o un aggravio sul premio di polizza che andrà a pagare nella nuova classe di merito occupata.

Attraverso queste regole riusciamo quindi a riassumere quali possano essere i possibili esiti al termine dell'anno di copertura, e lo strumento più adatto per rappresentare questo tipo di percorso è la matrice di transizione T_k (DENUIT, MARECHAL, PITREBOIS, & WALHIN, 2007):

$$T(k) = \begin{vmatrix} t_{00}(k) & t_{01}(k) & \cdots & t_{0s}(k) \\ t_{10}(k) & t_{11}(k) & \cdots & t_{1s}(k) \\ \vdots & \vdots & \ddots & \vdots \\ t_{s0}(k) & t_{s1}(k) & \cdots & t_{ss}(k) \end{vmatrix}$$

Supponiamo ad esempio che un assicurato si trovi attualmente in classe i; la classe che andrà ad occupare l'anno successivo, che indicheremo con j, sarà funzione del numero di sinistri ad esso imputabili. In particolare, i singoli elementi della matrice seguiranno la regola:

$$t_{ij}(k) = \begin{cases} 1, & se\ T_k(i) = j \\ 0, & seT_k(i) \neq j \end{cases},$$

[40] Risulta chiaro, effettivamente, che un sistema che nasce con lo scopo di penalizzare coloro che dimostrano comportamenti maggiormente rischiosi (come coloro che hanno appena conseguito la patente) vede nella Legge Bersani la fine di tutti i principi cardine in quanto non vi sarà più un'equa distribuzione dei premi.

ovvero, assumeranno valore 1 se, sulla base del numero k di sinistri registrati, ci sarà l'effettivo spostamento dalla classe i alla classe j, mentre nel caso contrario assumeranno valore 0.

Da tale relazione si generano un numero k di matrici in cui gli unici elementi presenti saranno 0 e 1 in base alla regola appena esposta, ed in particolare in base al sistema italiano vigente[41] le matrici assumeranno le seguenti forme:

$$T(0) = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{vmatrix}$$

$$T(1) = \begin{vmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

$$T(2) = \begin{vmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

$$T(k) = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix} \quad \text{per tutti i valori di } k \geq 3.$$

[41] Il sistema Bonus-Malus Italiano prevede che qualora non vi siano sinistri, il soggetto scenderà di una classe, mentre nel caso questi si verifichino salirà di due classi. Al termine del capitolo l'argomento verrà trattato più approfonditamente.

Quanto appena illustrato però costituisce solamente la base, mentre per fornire una spiegazione più utile, ed allo stesso tempo più completa, bisogna ricorrere alle probabilità. I sistemi BM, infatti, presentano al loro interno delle regole molto più complesse, che non si limitano a definire quali siano i casi possibili, ma per determinare il coefficiente da associare ad ogni classe di merito è necessario prevedere con quale probabilità vi sarà il passaggio dall'ipotetica classe i alla classe di merito j al termine dell'anno di copertura.

Quindi, assumendo che la frequenza di sinistro di un assicurato sia pari a ϑ, si dovrà calcolare la probabilità $p_{ij}(\vartheta)$ che sulla base di quel numero di sinistri passi dalla classe i alla classe j attraverso la probabilità condizionata

$$p_{ij}(\vartheta) = P[L_{k+1}(\vartheta) = j | L_k(\vartheta) = i] \,,$$

in cui $\{L_1(\vartheta), L_2(\vartheta), \dots\}$ indica semplicemente la traiettoria che seguirà col passare degli anni ogni soggetto. Ma poiché le singole traiettorie sono indipendenti l'una dall'altra e dipenderanno esclusivamente dal numero di sinistri $N_1, N_2, \dots,$[42] possiamo riscrivere l'equazione come (DENUIT, MARECHAL, PITREBOIS, & WALHIN, 2007):

$$p_{ij}(\vartheta) = \sum_{n=0}^{+\infty} P[L_{k+1}(\vartheta) = j | N_{k+1}(\vartheta) = n, L_k(\vartheta) = i] P[N_{k+1}(\vartheta) = n | L_k(\vartheta) = i]$$

$$= \sum_{n=0}^{\infty} \frac{\vartheta^n}{n!} \exp(-\vartheta)\, t_{ij}(n)$$ [43].

[42] La serie così definita è costituita, per ipotesi, dal numero di sinistri tra loro tutti indipendenti. Tale ipotesi permette di procedere coi calcoli senza problemi, ma da un punto di vista logico è priva di senso. L'idea di fare una correzione a posteriori nasce proprio dal fatto di penalizzare i soggetti maggiormente propensi ad effettuare sinistri, perché implicitamente si sta sostenendo che chi ha fatto molti sinistri in futuro sarà più propenso ad effettuarne degli altri; l'idea su cui quindi poggia questo modo di pensare è quella di una certa dipendenza fra il numero di sinistri, e quindi andrebbe a sfatare l'ipotesi di indipendenza sostenuta in tutte le dimostrazioni che coinvolgono questa variabile (BROUHNS, GUILLE'N, DENUIT, & PINQUET, 2002).

Le probabilità di transizione hanno un ruolo essenziale per gli attuari, perché permettono di fare delle valutazioni sulle possibili traiettorie; ruolo reso ancora più facile dal fatto che la scala BM può essere definita come un processo markoviano: sarà sufficiente conoscere la classe di merito precedente per fare tutte le valutazioni del caso, e non si dovrà risalire al percorso fatto in precedenza.
Sulla base di queste considerazioni si deriva, quindi, la matrice markoviana di transizione

$$P(\vartheta) = \begin{vmatrix} P_{00}(\vartheta) & P_{01}(\vartheta) & \cdots & P_{0s}(\vartheta) \\ P_{10}(\vartheta) & P_{11}(\vartheta) & \cdots & P_{1s}(\vartheta) \\ \vdots & \vdots & \ddots & \vdots \\ P_{s0}(\vartheta) & P_{s1}(\vartheta) & \cdots & P_{ss}(\vartheta) \end{vmatrix},$$

in cui $P(\vartheta) = \sum_{k=0}^{\infty} \frac{\vartheta^k}{k!} \exp(-\vartheta)\, T(k)$.
Facendo nuovamente riferimento al sistema BM adottato in Italia, la matrice può essere riscritta nella seguente forma:

---

[43] Si vuole sottolineare che, il calcolo delle probabilità di transizione poggia sul fatto che gli N_{k+1} sinistri, come le $L_k(\vartheta)$ traiettorie, sono indipendenti; di conseguenza la probabilità che si verifichi un sinistro, nota la traiettoria, si potrà calcolare nel seguente modo:

$$P[N_{k+1} = n | L_k(\vartheta) = l_1] = P[N_{k+1} = n] = \frac{\vartheta^n}{n!} \exp(-\vartheta).$$

$$P(\vartheta) = \left| \begin{matrix} exp(-\vartheta) & 0 & \vartheta exp(-\vartheta) & 0 & \frac{\vartheta^2}{2} exp(-\vartheta) & 1 - \sum_1 \\ exp(-\vartheta) & 0 & 0 & \vartheta exp(-\vartheta) & 0 & 1 - \sum_2 \\ 0 & exp(-\vartheta) & 0 & 0 & \vartheta exp(-\vartheta) & 1 - \sum_3 \\ 0 & 0 & exp(-\vartheta) & 0 & 0 & 1 - exp(-\vartheta) \\ 0 & 0 & 0 & exp(-\vartheta) & 0 & 1 - exp(-\vartheta) \\ 0 & 0 & 0 & 0 & exp(-\vartheta) & 1 - exp(-\vartheta) \end{matrix} \right|$$

dove:

$\sum_i, con\ i = 1,2,3$, corrisponde alla somma di tutti gli elementi presenti in ogni riga, e quindi si avrà:

$$\sum_1 = \exp(-\vartheta)\left(1 + \vartheta + \frac{\vartheta^2}{2}\right)$$

$$\sum_2 = \sum_3 = \exp(-\vartheta)\,(1 + \vartheta)\,.$$

La matrice appena illustrata viene indicata anche come matrice ad un solo passaggio, per distinguerla da quella che permette di prevedere la traiettoria di un singolo individuo in un numero di passaggi superiore ad uno (convenzionalmente n).

Quindi si potrebbe vedere nella matrice multi-passaggio un'espressione più evoluta, dove si evidenziano le probabilità che il soggetto in n anni passi dalla classe i alla classe j attraverso la probabilità condizionata

$$P_{ij}^{(n)}(\vartheta) = \mathrm{P}[L_{k+n}(\vartheta) = j | L_k(\vartheta) = i],$$

che si traduce in forma matriciale come segue:

$$P^n(\vartheta) = \begin{pmatrix} P_{00}^{(n)}(\vartheta) & P_{01}^{(n)}(\vartheta) & \dots & P_{0s}^{(n)}(\vartheta) \\ P_{10}^{(n)}(\vartheta) & P_{11}^{(n)}(\vartheta) & \cdots & P_{1s}^{(n)}(\vartheta) \\ \vdots & \vdots & \ddots & \vdots \\ P_{s0}^{(n)}(\vartheta) & P_{s1}^{(n)}(\vartheta) & \dots & P_{ss}^{(n)}(\vartheta) \end{pmatrix}.$$

5.3. IL PREMIO DI RIFERIMENTO PER IL SISTEMA BONUS-MALUS

Abbiamo a lungo trattato le variabili tariffarie, i sistemi di personalizzazione a priori ed a posteriori, ma con il solo impiego di questi elementi, sarebbe difficoltoso ottenere il premio assicurativo più idoneo; il punto di partenza, ovvero il valore a cui verranno associati i coefficienti più consoni, è il premio di riferimento.

Quando si va parlare di premio assicurativo, implicitamente si fa riferimento al modello moltiplicativo (sempre nel caso specifico dell'RCA):

$$P\gamma_i\pi_j\ ,$$

in cui al premio di riferimento viene associata: una relatività γ, legata alla classe tariffaria i determinata sulla base della valutazione a priori; ed un coefficiente di premio π che invece varierà in funzione della j-esima classe bonus-malus occupata.

Mentre per il calcolo dei due coefficienti si richiamano modelli specifici, facenti spesso riferimento a strumenti statistico-econometrici, per il premio base la valutazione è innanzitutto soggettiva: ogni compagnia assicurativa fa le proprie previsioni in merito.

La scelta può, infatti, ricadere verso un ammontare unico per tutti gli assicurati presenti nel portafoglio, oppure fare una differenziazione sulla base di quanto è emerso dalla valutazione a priori. Per effettuare questa scelta potrebbe essere utile ricondursi alle proprietà enunciate nel paragrafo n° 5.1., ovvero le proprietà espresse da Gigante, Picech & Sigalotti (2000), che permettono di distinguere un buon sistema Bonus-Malus; soluzione non definitiva, comunque, e la ragione è dovuta all'impossibilità di ottenere un premio in grado di soddisfare le proprietà suddette.

Si osserva, infatti, che ponendo ad esempio un premio di equilibrio P^e, calcolato in modo che soddisfi l'equazione:

$$\sum_{i=1}^{s}\sum_{j=1}^{J} P^e \gamma_i \pi_j P(I=i, Y=j) = E(N_t)\,,$$

si riuscirebbe a mantenere l'equilibrio finanziario all'interno della compagnia, poichè l'ammontare che verrà incassato sarà necessariamente uguale all'ammontare del numero di sinistri; emerge chiaramente che la soluzione per l'equazione è unica, quindi un premio di riferimento universale, valido per tutti i soggetti.

Ricordiamo che un buon sistema BM può essere definito tale anche se ha la capacità di adeguare progressivamente il premio a posteriori, al punto tale da imporre premi quasi identici per soggetti che nel lungo periodo dimostrano un'esperienza simile; perché ciò sia possibile, però, è necessario che i premi di riferimento varino in base alla classe tariffaria considerata.

Si comprende chiaramente che, qualsiasi sia la scelta, non si riuscirà a rispettare entrambe le proprietà; proprio per questo motivo si osserva che nella pratica le assicurazioni preferiscono adottare un unico premio di partenza, non perché reputino l'equilibrio finanziario l'unica proprietà importante da rispettare, ma perché è evidente la complessità che si genererebbe nel caso contrario.

5.4. DALLA STIMA DEI COEFFICIENTI AL CALCOLO DEL PREMIO: UN ESEMPIO

Ogni compagnia assicurativa utilizza le formule finora introdotte per calcolare i coefficienti da associare alle diverse caratteristiche degli assicurati. Tali coefficienti vengono raccolti all'interno di tabelle facenti parte delle norme tariffarie, tabelle che col tempo sono state eliminate e sostituite da complessi programmi in grado di effettuare i preventivi sulle polizze senza ricorrere a strumenti cartacei, ma che per semplicità vengono presentate qui di seguito come miglior dimostrazione dell'utilità di quanto esposto finora.

Come già accennato, ogni compagnia ha la facoltà di favorire o sfavorire soggetti con determinate caratteristiche all'interno del proprio portafoglio assicurati, e quindi di compagnia in compagnia, e non solamente di paese in paese, si vedranno dei coefficienti diversi. Gli strumenti utilizzati in questo capitolo fanno riferimento alla compagnia assicurativa UBI Assicurazioni (UBI-Assicurazioni, 2014), che all'interno delle norme tariffarie indica in che modo andare a calcolare il premio RCAuto in base alle caratteristiche rilevate sull'assicurato; in particolare essa stabilisce un premio di riferimento iniziale di € 708,33 uguale per tutti i soggetti, e sulla base dei coefficienti indicati nelle tabelle di seguito presentate si arriverà al calcolo del premio finale[44].

Il primo coefficiente per cui moltiplicare tale premio sarà dato dalla classe Bonus-Malus occupata, valore che andrà giustamente a penalizzare coloro che si troveranno nelle classi più alte, ed a premiare coloro che in base al loro profilo di rischio hanno dimostrato di essere dei buoni guidatori.

Tabella 5.1: Coefficienti assegnati in base alla classe Bonus-Malus di appartenenza

CLASSE BM	COEFFICIENTE DI ASSUNZIONE
1	0,48
2	0,50
3	0,53
4	0,56
5	0,62
6	0,66
7	0,70
8	0,74

[44] Il premio così calcolato sarà al netto della tassazione e delle eventuali garanzie accessorie (cristalli, assistenza stradale, furto e incendio), le quali potranno essere aggiunte in secondo momento a discrezione dell'assicurato.

9	0,78
10	0,82
11	0,88
12	0,94
13	1,00
14	1,20
15	1,40
16	1,80
17	2,50
18	3,00

Fonte: UBI Assicurazioni (2014)

La distinzione fra i diversi assicurati avviene anche sulla base della zona di residenza del proprietario del veicolo, variabile spesso preferita alla zona di circolazione del mezzo.

Figura 5.2: Tabella con i coefficienti relativi alla zona di residenza

PROVINCIA DI RESIDENZA DEL PROPRIETARIO DEL VEICOLO	SIGLA	COEF.	PROVINCIA DI RESIDENZA DEL PROPRIETARIO DEL VEICOLO	SIGLA	COEF.
Agrigento	AG	0,6379	Nuoro	NU	0,7940
Alessandria	AL	0,4205	Ogliastra	OG	0,5101
Ancona	AN	0,7435	Oristano	OR	0,6001
Aosta	AO	0,5544	Olbia Tempio	OT	0,4103
Ascoli Piceno	AP	0,6988	Palermo	PA	0,7028
L'Aquila	AQ	0,5544	Piacenza	PC	0,5037
Arezzo	AR	0,8175	Padova	PD	0,6270
Asti	AT	0,4773	Pescara	PE	0,9890

Avellino	AV	1,2332	Perugia	PG	0,5028
Bari	BA	0,9314	Pisa	PI	1,0468
Bergamo	BG	0,4713	Pordenone	PN	0,5994
Biella	BI	0,4180	Prato	PO	1,0072
Belluno	BL	0,5880	Parma	PR	0,6307
Benevento	BN	0,9669	Pistoia	PT	1,2344
Bologna	BO	0,6880	Pesaro Urbino	PU	0,7207
Brindisi	BR	1,0537	Pavia	PV	0,4853
Brescia	BS	0,4768	Potenza	PZ	0,5797
Barletta Andria Trani	BT	1,0036	Ravenna	RA	0,8196
Bolzano	BZ	0,5766	Reggio Calabria	RC	1,5463
Cagliari	CA	0,9719	Reggio Emilia	RE	0,5961
Campobasso	CB	0,5978	Ragusa	RG	0,7234
Caserta	CE	2,3837	Rieti	RI	0,5563
Chieti	CH	0,7749	Roma	RM	0,7834
Carbonia Iglesias	CI	0,6007	Rimini	RN	1,0019
Caltanissetta	CL	0,7554	Rovigo	RO	0,7248
Cuneo	CN	0,4088	Salerno	SA	1,2033
Como	CO	0,6058	Siena	SI	0,7138
Cremona	CR	0,4108	Sondrio	SO	0,5198
Cosenza	CS	0,6994	La Spezia	SP	0,8780
Catania	CT	0,7928	Siracusa	SR	0,6045
Catanzaro	CZ	0,9270	Sassari	SS	0,8447
Enna	EN	0,6524	Savona	SV	0,6152
Forlì Cesena	FC	0,6600	Taranto	TA	1,4341
Ferrara	FE	0,5264	Teramo	TE	0,7914
Foggia	FG	0,9899	Trento	TN	0,5499
Firenze	FI	0,9319	Torino	TO	0,5985
Fermo	FM	0,5965	Trapani	TP	0,7620

Frosinone	FR	0,8356	Terni	TR	0,7633
Genova	GE	0,6590	Trieste	TS	0,8032
Gorizia	GO	0,5015	Treviso	TV	0,5611
Grosseto	GR	0,7635	Udine	UD	0,5696
Imperia	IM	0,5585	Varese	VA	0,4838
Isernia	IS	0,8177	Verbania	VB	0,5799
Crotone	KR	1,5392	Vercelli	VC	0,5037
Lecco	LC	0,4272	Venezia	VE	0,7172
Lecce	LE	0,8292	Vicenza	VI	0,5388
Livorno	LI	0,6967	Verona	VR	0,4095
Lodi	LO	0,4567	Medio Campidano	VS	0,6453
Latina	LT	1,1496	Viterbo	VT	0,4781
Lucca	LU	1,0367	Vibo Valenzia	VV	1,3178
Monza Brianza	MB	0,4857	S. O. Malta	EE	0,7827
Macerata	MC	0,6738	Corpo Diplomatico	SMOM	0,7827
Messina	ME	0,6943	Escursionisti Esteri	CD	0,7827
Milano	MI	0,5586	AFI	AFI	0,7827
Mantova	MN	0,4269	FTASE	FTASE	0,7827
Modena	MO	0,5867	Targhe Estere	E	0,7827
Massa Carrara	MS	0,8180	Croce Rossa Italiana	CRI	0,7827
Matera	MT	0,6402	Corpo Consolare	CC	0,7827
Napoli	NA	2,2371	Citta del Vaticano	SCV	0,7827
Novara	NO	0,4157	Repubblica San Marino	RSM	1,0552

Fonte: UBI Assicurazioni (2014)

La successiva caratteristica osservata riguarda l'alimentazione e la cilindrata del veicolo, sulla base delle quali verrà attribuito un coefficiente che sarà tendenzialmente più elevato all'aumentare della cilindrata del mezzo assicurato.

Figura 5.3: Coefficienti relativi alla potenza del veicolo assicurato

ALIMENTAZIONE (ESCLUSO ELETTRICHE)	POTENZA FISCALE	CILIDRATA	KW	COEFFICIENTE
Qualsiasi alimentazione	fino a 8	fino a 569,5	da 0 a 9999,00	1,0276
Qualsiasi alimentazione	da 9 a 10	da569,6 a 774,0	da 0 a 9999,00	1,7343
Qualsiasi alimentazione	da 11 a 12	da 774,1 a 999,2	da 0 a 40	1,8085
Qualsiasi alimentazione	da 11 a 12	da 774,1 a 999,2	da 40,01 a 9999,00	1,9037
Qualsiasi alimentazione	da 13 a 14	da 999,3 a 1243,6	da 0 a 43	2,0680
Qualsiasi alimentazione	da 13 a 14	da 999,3 a 1243,6	da 43,01 a 9999,0	2,1738
Benzina	da 15 a 16	da 1234,7 a 1505,9	da 0 a 55,00	2,2261
Benzina	da 15 a 16	da 1234,7 a 1505,9	da 55,01 a 9999,00	2,3433
Diesel	da 15 a 16	da 1234,7 a 1505,9	da 0 a 53,00	2,3167
Diesel	da 15 a 16	da 1234,7 a 1505,9	da 53,01 a 9999,00	2,4386
Benzina	da 17 a 18	da 1506,0 a 1784,9	da 0 a 75,00	2,3511
Benzina	da 17 a 18	da 1506,0 a 1784,9	da 75,01 a 9999,00	2,4491
Diesel	da 17 a 18	da 1506,0 a 1784,9	da 0 a 65,00	2,4405
Diesel	da 17 a 18	da 1506,0 a 1784,9	da 65,01 a 9999,00	2,5422
Benzina	da 19 a 20	da 1785,0 a 2080,1	da 0 a 9999,00	2,6616
Diesel	da 19 a 20	da 1785,0 a 2080,1	da 0 a 100,00	2,8153
Diesel	da 19 a 20	da 1785,0 a 2080,1	da 100,01 a 9999,00	2,9024
Benzina	oltre 20	oltre 2080,1	da 0 a 9999,00	2,9380
Diesel	oltre 20	oltre 2080,1	da 0 a 100,00	3,1953
Diesel	oltre 20	oltre 2080,1	da 100,01 a 9999,00	3,2625

Fonte: UBI Assicurazioni (2014)

Tabella 5.4: Coefficienti per le sole autovetture elettriche

ALIMENTAZIONE ELETTRICA	COEFFICIENTE
Alimentazione elettrica fino a kw 60,00	1,8085
Alimentazione elettrica da kw 60,01 fino a kw 80,00	2,3433
Alimentazione elettrica da kw 80,01 fino a kw 100,00	2,4491
Alimentazione elettrica da kw 100,01 fino a kw 120,00	2,6616
Alimentazione elettrica oltre kw 120,00	2,9380

Fonte: UBI Assicurazioni (2014)

Importante è anche l'anzianità del veicolo, valutata in base alla data di immatricolazione; nella compagnia presa in considerazione si osserva che vengono agevolati i veicoli più datati, a sfavore di quelli che risultano immatricolati da meno anni.

Tabella 5.5: Coefficienti associati alla vetustà della macchina

PERSONALIZZAZIONE VETUSTA' DEL VEICOLO	COEFFICIENTE
Immatricolazione fino al 2° anno completo	1,0100
Immatricolazione dal 3° anno fino al 5° anno completo	1,0100
Immatricolazione dal 6° anno fino all'8° anno completo	1,0080
Immatricolazione dal 9° anno fino all'11° anno completo	1,0080
Immatricolazione dal 12° anno fino al 14° anno completo	0,9879
Immatricolazione dal 15° anno fino al 20° anno completo	0,9600
Immatricolazione dal 21° anno e oltre	0,9479

Fonte: UBI Assicurazioni (2014)

La marca del veicolo può essere considerata come una variabile aggiuntiva, un elemento che non sempre permette di far variare notevolmente il premio finale.

Tabella 5.6: Fattore di personalizzazione del premio in base alla marca dell'autoveicolo

PERSONALIZZAZIONE MARCA DEL VEICOLO	COEFFICIENTE
ALFA ROMEO	1,0286
AUDI	1,0228
BMW	1,0130
CITROEN	0,9752
DAEWOO	0,9814
FIAT	0,9937
FORD	0,9853
HONDA	1,0403
HYUNDAI	1,0204
LANCIA	1,0286
MERCEDES	0,9502
NISSAN	0,9823
OPEL	0,9649
PEUGEOT	0,9753
RENAULT	1,0172
SEAT	0,9861
SMART	1,0238
SUZUKI	0,9818
TOYOTA	1,0124
VOLKSWAGEN	0,9779
Altre marche	1,0443

Fonte: UBI Assicurazioni (2014)

Si osserva, poi, che generalmente vi è una forma di disincentivo verso gli assicurati più giovani, ai quali verrà associato un coefficiente nettamente superiore.

Tabella 5.7: Coefficiente adeguato all'età del sottoscrittore della polizza assicurativa

PERSONALIZZAZIONE ETA' DEL PROPRIETARIO DEL VEICOLO	COEFFICIENTE
Fino a 20	2,4702
Da 21 a 23	1,5896
Da 24 a 26	1,0829
Da 27 a 29	0,9438
Da 30 a 32	0,8402
Da 33 a 35	0,8618
Da 36 a 38	0,8816
Da 39 a 41	0,9101
Da 42 a 44	1,0316
Da 45 a 47	1,1095
Da 48 a 50	1,1370
Da 51 a 53	1,0804
Da 54 a 56	0,9998
Da 57 a 59	0,9482
Da 60 a 62	0,9318
Da 63 a 65	0,8752
Da 66 a 70	0,8901
Da 71 a 75	0,9739
Da 76 a 80	1,1741
Oltre 80	1,2211
Persona giuridica	0,9360

Fonte: UBI Assicurazioni (2014)

Infine, sul premio pagato incide considerevolmente anche la storia di sinistrosità di ogni assicurato, informazione che si ricava dall'attestato di rischio; quest'ultimo documento viene consegnato dalla compagnia assicurativa con cui si era assicurati precedentemente e permette di vedere se negli ultimi cinque anni vi sono stati dei sinistri con responsabilità e se essi sono già stati liquidati. Da tale documento emerge anche il numero di anni in cui il mezzo è stato assicurato: un soggetto che non è assicurato da almeno cinque anni riceverà una penalizzazione data dal fatto che, non possedendo una storia di sinistrosità completa, non si avranno sufficienti informazioni per assegnare il premio corretto.

Quindi sulla base del numero di anni assicurati, e in base al numero di sinistri rilevati, si procederà con l'assegnazione dei coefficienti presenti nella tabella 5.8.

Tabella 5.8: Attestato di rischio con relativo coefficiente

	V	IV	III	II	I	ANNO IN CORSO	COEF.
Assicurato da 5 anni	0	0	0	0	0	0	1,000
Assicurato da 4 anni	NA	0	0	0	0	0	1,060
Assicurato da 3 anni	NA	NA	0	0	0	0	1,060
Assicurato da 2 anni	NA	NA	NA	0	0	0	1,115
Assicurato da 1 anno	NA	NA	NA	NA	0	0	1,115
Nuovo veicolo da ass.	NA	NA	NA	NA	NA	NA	1,300

Fonte: UBI Assicurazioni (2014)

Tabella 5.9: Coefficiente associato al numero di sinistri con colpa

NUMERO DI SINISTRI PRESENTI	COEFFICIENTE
Somma dei sinistri = 0	1,0000
Somma dei sinistri = 1	1,1792

Somma dei sinistri = 2 o più	1,2500

Fonte: UBI Assicurazioni (2014)

Come già annunciato in precedenza, vi sono numerosi elementi sulla base dei quali è possibile personalizzare il premio, elementi che qui non vengono considerati come: impiego del soggetto assicurato, presenza di caratteristiche accessorie nel veicolo, luogo in cui viene abitualmente parcheggiata l'auto.
Quindi si deduce che, sulla base dei coefficienti sopra descritti, se un soggetto di 40 anni residente nella provincia di Venezia e possessore di una Volkswagen Golf immatricolata nel 2013 scegliesse di assicurarsi con la UBI Assicurazioni, verrebbe classificato nel seguente modo:

Tabella 5.10: Esempio

Premio di riferimento €	708,33
CLASSE BM=1	0,4800
GOLF BENZINA 77 KW	2,3433
IMMATRICOLATO 2013	1,0100
VOLKSWAGEN	0,9779
VENEZIA	0,7172
40 ANNI	0,9101
ASSICURATO DA 2 ANNI	1,115
0 SINISTRI	1,0000
PREMIO FINALE	**572,69**

5.5. I DIVERSI SISTEMI BONUS-MALUS NEL MONDO

Come già si capisce dal titolo del capitolo, non esiste un unico sistema Bonus-Malus valido per tutti i Paesi del mondo, ma ve ne sono diversi, ognuno con le caratteristiche proprie della regolamentazione vigente nello Stato in cui esso viene applicato.

Proprio per le diversità che essi presentano, non si può effettuare un semplice confronto sulla base dei premi pagati dagli assicurati, ma bisogna introdurre degli indicatori standard che permettano un equo paragone (LEMAIRE & ZI, 1994).

Il primo indicatore è la stazionarietà media relativa di livello (RSAL) :

$$RSAL = \frac{livello\ medio\ stazionario - livello\ minimo}{livello\ massimo - livello\ minimo},$$

in cui il livello medio stazionario rappresenta la classe mediamente più occupata, mentre i livelli di massimo e minimo dipenderanno rispettivamente dalla classe massima e minima presente nella scala Bonus-Malus considerata. Il risultato, espresso in termini percentuali, indicherà la posizione media degli assicurati all'interno delle diverse scale BM; per valori bassi si dovrà dedurre che mediamente la popolazione occupa le classi più basse della scala, mentre per valori elevati ci sarà una maggior distribuzione nei livelli.

Tabella 5.11: Stazionarietà media relativa di tutti i sistemi analizzati

CLASSIFICA	STATO	RSAL
1	Kenya	28,79%
2	Spagna	25,67%
3	Malesia	21,17%
4	Finlandia (nuovo) [45]	16,04%

[45] L'analisi sui diversi Paesi non si è limitata allo studio dei sistemi Bonus-Malus attualmente utilizzati, ma si sono presi in considerazione anche i sistemi più "vecchi", ovvero quelli che dopo l'anno 1988 hanno subito dei mutamenti; l'indicazione vecchio o nuovo serve a distinguere a quale dei due sistemi si sta facendo riferimento.

5	Svezia	14,20%
6	Olanda	11,78%
7	Regno Unito (protetto)	11,37%
8	Taiwan	9,55%
9	Finlandia (vecchio)	8,46%
10	Hong Kong	8,35%
11	Tailandia	8,03%
12	Regno Unito (non protetto)	7,07%
13	Portogallo	6,75%
14	Norvegia (vecchio)	6,61%
15	Svizzera (nuovo)	6,47%
16	Germania (nuovo)	5,85%
17	Giappone (nuovo)	4,63%
18	Belgio (nuovo)	4,05%
19	Danimarca	3,78%
20	Svizzera (vecchio)	2,90%
21	Francia	2,12%
22	Norvegia (nuovo)	2,11%
23	Brasile	1,85%
24	Corea	1,37%
25	Lussemburgo (nuovo)	1,36%
26	Italia (nuovo)	1,30%
27	Lussemburgo (vecchio)	1,01%
28	Giappone (vecchio)	0,88%
29	Belgio (vecchio)	0,74%
30	Italia (vecchio)	0,01%

Fonte: LEMAIRE & ZI (1994)

Facendo una panoramica generale, si può constatare che la diversità fra le regolamentazioni che governano i vari paesi incide anche su quello che è il sistema adottato per valutare i nuovi assicurati, categoria già penalizzata di per sé, ma che con determinati sistemi viene ulteriormente penalizzata a causa dell'elevato sovrapprezzo applicato, come si vede nella tabella 5.12.

Tabella 5.12: Sovrapprezzo implicito per i nuovi assicurati

CLASSIFICA	STATO	SOVRAPPREZZO
1	Germania (nuovo)	212,97%
2	Norvegia (nuovo)	195,80%
3	Danimarca	189,50%
4	Norvegia (vecchio)	159,13%
5	Svezia	158,89%
6	Olanda	146,29%
7	Giappone (vecchio)	144,12%
8	Finlandia (vecchio)	143,39%
9	Finlandia (nuovo)	142,57%
10	Corea	135,51%
11	Hong Kong	122,04%
12	Giappone (nuovo)	121,76%
13	Italia (nuovo)	121,38%
14	Lussemburgo (nuovo)	100,89%
15	Regno Unito (non protetto)	98,75%
16	Svizzera (nuovo)	94,10%
17	Lussemburgo (vecchio)	92,25%
18	Regno Unito (protetto)	84,65%
19	Francia	77,55%
20	Malesia	76,65%
21	Kenya	74,60%
22	Taiwan	68,20%
23	Svizzera (vecchio)	67,88%

24	Italia (vecchio)	64,26%
25	Brasile	52,33%
26	Tailandia	50,55%
27	Belgio (nuovo)	41,87%
28	Belgio (vecchio)	39,26%
29	Spagna	28,70%
30	Portogallo	26,05%

Fonte: LEMAIRE & ZI (1994)

Il secondo strumento di confronto è il coefficiente di variazione dei premi: ovvero di quanto variano i premi assicurativi sulla base dell'esperienza raccolta. Calcolato semplicemente come rapporto tra la deviazione standard e la media, contribuisce a chiarire in che modo i diversi sistemi analizzati possano influire nel calcolo del premio; nella tabella 5.13 si valuta il coefficiente di variazione per un assicurato con frequenza sinistri pari al 10%.

Tabella 5.13: Coefficiente di variazione del premio

CLASSIFICA	STATO	COEFFICIENTE DI VARIAZIONE
1	Svizzera (nuovo)	0,4595
2	Norvegia (vecchio)	0,3900
3	Kenya	0,3835
4	Finlandia (nuovo)	0,3834
5	Svezia	0,3769
6	Olanda	0,3523
7	Giappone (nuovo)	0,3283
8	Taiwan	0,3162
9	Malesia	0,3075

10	Danimarca	0,3017
11	Svizzera (vecchio)	0,2700
12	Finlandia (vecchio)	0,2570
13	Germania (nuovo)	0,2536
14	Hong Kong	0,2518
15	Regno Unito (non protetto)	0,2419
16	Lussemburgo (nuovo)	0,2147
17	Belgio (nuovo)	0,2128
18	Francia	0,2049
19	Norvegia (nuovo)	0,2049
20	Portogallo	0,1956
21	Tailandia	0,1925
22	Spagna	0,1533
23	Corea	0,1271
24	Giappone (vecchio)	0,1261
25	Regno Unito (protetto)	0,1260
26	Lussemburgo (vecchio)	0,1075
27	Italia (nuovo)	0,0934
28	Belgio (vecchio)	0,0586
29	Brasile	0,0304
30	Italia (vecchio)	0,0046

Fonte: LEMAIRE & ZI (1994)

Particolarmente importante come elemento di confronto è anche l'efficienza di un sistema Bonus-Malus; considerando che esso non si basa su quella che è l'entità del danno causato dal sinistro, ma semplicemente sul numero di sinistri con responsabilità che vengono fatti da ciascun assicurato[46], il

[46] Da qui deriva l'indipendenza del numero di sinistri dall'ammontare del danno provocato.

sistema si potrà definire efficiente se sarà in grado di cogliere tutte le oscillazioni della frequenza sinistri per poi adattarsi di conseguenza .

Tale caratteristica viene normalmente espressa in termini di elasticità, ed indicando con λ la frequenza sinistri, e con $P(\lambda)$ il premio medio di stazionarietà, si potrà parlare di sistema perfettamente elastico solo quando sarà valida la seguente equazione

$$\frac{\frac{\mathrm{d}\lambda}{\lambda}}{\frac{\mathrm{dP}(\lambda)}{\mathrm{P}(\lambda)}} = 1 \, .$$

Nonostante l'equazione precedente mostri una situazione ideale che nella realtà difficilmente si presenta, si può comunque approssimare l'elasticità di ogni sistema attraverso la formula

$$\eta(\lambda) = \frac{dP(\lambda)/P(\lambda)}{d\lambda/\lambda} \, ,$$

e confrontare i sistemi che meglio rispondono ai cambiamenti in termini di frequenza sinistri (LEMAIRE, 1995).

Tabella 5.14: Classifica dei diversi sistemi sulla base dell'efficienza

CLASSIFICA	STATO	EFFICIENZA
1	Svizzera (nuovo)	0,449
2	Finlandia (nuovo)	0,403
3	Svezia	0,298
4	Olanda	0,275
5	Norvegia (vecchio)	0,263
6	Germania (nuovo)	0,257
7	Kenya	0,237
8	Giappone (nuovo)	0,232
9	Svizzera (vecchio)	0,208
10	Francia	0,200
11	Belgio (nuovo)	0,195

12	Finlandia (vecchio)	0,194
13	Lussemburgo (nuovo)	0,183
14	Malesia	0,165
15	Danimarca	0,165
16	Taiwan	0,136
17	Hong Kong	0,133
18	Regno Unito (non protetto)	0,129
19	Norvegia (nuovo)	0,127
20	Portogallo	0,111
21	Tailandia	0,081
22	Spagna	0,079
23	Corea	0,078
24	Italia (nuovo)	0,063
25	Lussemburgo (vecchio)	0,058
26	Giappone (vecchio)	0,052
27	Regno Unito (protetto)	0,051
28	Belgio (vecchio)	0,024
29	Brasile	0,011
30	Italia (vecchio)	0,001

Fonte: LEMAIRE & ZI (1994)

Uno dei rovesci della medaglia dei sistemi BM viene riassunto con l'espressione "hunger for bonus", ovvero il fenomeno per cui gli assicurati preferiscono non comunicare alla loro compagnia di aver fatto un sinistro solo per il timore di andare in Malus e di conseguenza pagare un premio maggiore l'anno successivo (LEMAIRE, 1995).

Un buon sistema di personalizzazione a posteriori dovrebbe essere in grado di evitare questo fenomeno, e quindi far sì che anche per incidenti di valori minimi i soggetti avvisino senza timore la loro assicurazione. Abitualmente si definisce quindi una soglia, calcolata attraverso complessi algoritmi, che

permette di identificare fino a che valore in termini monetari un soggetto ha la convenienza a non comunicare il fatto accaduto; questa soglia prende il nome di ritenzione media ottimale.
Sulla base degli esiti ottenuti attraverso l'analisi dei vari indicatori sopra esposti, è stato possibile quindi stilare una classifica generale in cui l'elemento discriminante è la rigidità del sistema.

Tabella 5.15: La classifica degli Stati con i sistemi BM più rigidi

CLASSIFICA	STATO
1	Svizzera (nuovo)
2	Finlandia (nuovo)
3	Kenya
4	Svezia
5	Taiwan
6	Norvegia (vecchio)
7	Olanda
8	Malesia
9	Germania (nuovo)
10	Finlandia (vecchio)
11	Giappone (nuovo)
12	Danimarca
13	Svizzera (vecchio)
14	Hong Kong
15	Regno Unito (non protetto)
16	Spagna
17	Portogallo
18	Belgio (nuovo)
19	Lussemburgo (nuovo)
20	Francia

21	Norvegia (nuovo)
22	Tailandia
23	Regno Unito (protetto)
24	Corea
25	Lussemburgo (vecchio)
26	Italia (nuovo)
27	Giappone (vecchio)
28	Belgio (vecchio)
29	Brasile
30	Italia (vecchio)

Fonte: LEMAIRE & ZI (1994)

A prescindere comunque dalla presenza di indicatori e classifiche, si osserva che rispetto ai sistemi generalmente adottati nei vari paesi europei, di notevole complessità e severità appare l'approccio adottato nel Nord-America, poiché la valutazione dell'esperienza non considera solo il numero di incidenti con responsabilità, ma fa una valutazione anche sulla guida stessa dell'assicurato. Scelta dettata dalla volontà dello stato, nasce non solo per incoraggiare una guida più sicura, ma anche evitare le violazioni del codice della strada.

Il sistema del Nord-America si presenta come una scala formata da 35 livelli, in cui un nuovo assicurato sarà collocato nella quindicesima classe; l'anno successivo potrà scendere di una classe se dimostrerà una guida virtuosa, potrà salire di 2 o 5 livelli se infrange il codice della strada, mentre se farà un incidente potrà salire di 3 o 4 livelli in base alla gravità riscontrata (BROUHNS, GUILLEN, DENUIT, & PINQUET, 2002).

6. ANALISI COMPARATA DI DUE MODELLI: LA TARIFFAZIONE NEL RAMO RCA ATTRAVERSO I MODELLI LINEARI GENERALIZZATI

Semplicemente sfogliando le principali riviste dedicate al mondo attuariale, si osserva che il tema "sistemi Bonus-Malus" è molto dibattuto, ed è in continua evoluzione. Con l'obiettivo comune di studiare quale sia il sistema di tariffazione ottimale, negli anni sono stati proposti diversi modelli, ognuno con l'intenzione di cogliere, e progressivamente ridurre, tutti gli elementi di eterogeneità che potrebbero presentarsi al momento della sottoscrizione della polizza assicurativa. Seppur numerosi, tali modelli possono essere ricondotti a due macro categorie, nelle quali l'elemento discriminatorio è la tipologia d'informazione utilizzata; se da un lato c'è chi sostiene che la tariffazione sia un processo da effettuare sulla base delle sole informazioni rilevate a priori, ad essi si contrappone la corrente a favore dell'inclusione della classe Bonus-Malus all'interno del modello (FRANGOS & VRONTOS, 2001).

Utilizzando alcuni dati relativi ad una compagnia assicurativa, opportunamente ponderati attraverso indagini sul settore (S.I.F.A., 2014) (ACI, 2013) (ANIA, 2014), è stato possibile ricostruire due modelli di tariffazione che andassero in qualche modo a riflettere entrambe le correnti di pensiero; avvalendosi, in particolar modo, dei modelli lineari generalizzati (general linear model-GLM) si è potuta sfruttare l'informazione fornita per stimare un modello che fosse in grado di descrivere il numero di sinistri attesi.

6.1. I MODELLI LINEARI GENERALIZZATI

La scelta di utilizzare questa tipologia di modelli, e non affidarsi semplicemente ai più noti modelli di regressione lineare, è stata dettata dai vincoli posti da quest'ultimo strumento; vincoli che potremmo riassumere nella necessità che le variabili endogene si distribuiscano in modo normale, una condizione che le variabili che si andranno ad utilizzare non possiedono. Adottando i GLM, invece, alle variabili risposta è possibile assegnare una delle distribuzioni appartenenti alla famiglia esponenziale lineare, e quindi : oltre alla distribuzione Normale, anche le distribuzioni Gamma, Gaussiana Inversa, di Poisson e Binomiale Negativa, a seconda della natura della variabile oggetto di studio.
Avendo a disposizione un insieme di informazioni (y_i, x_i), date da un insieme di unità statistiche $i = 1, \dots, n$, il valore che assumeranno le variabili esplicative, e quindi quelle in grado di descrivere il fenomeno che si va a studiare, verrà riassunto all'interno del vettore x_i. Attraverso la regressione delle variabili esplicative si ottiene, poi, il vettore $\hat{y} = (y_1, \dots, y_n)'$, elemento essenziale per formulare un'ipotesi sulla distribuzione della nostra variabile d'interesse Y; ipotesi che sarà realizzata, appunto, attraverso l'utilizzo dei vettori delle grandezze assunte dai regressori.
Le ipotesi sulle quali poggia questa tipologia di strumento riguardano, innanzitutto, le variabili risposta $Y_1, \dots, Y_n$, le quali devono necessariamente avere una distribuzione facente parte della famiglia esponenziale lineare, oltre ad essere stocasticamente indipendenti[47]. La seconda ipotesi riguarda, invece, il legame tra il valore atteso μ_i della variabile dipendente Y_i, ed il

[47] L'indipendenza stocastica fa in modo che, in presenza di due eventi denominati A e B, il realizzarsi di A non influenzi la realizzazione di B; ovvero $P(A|B) = A$ e $P(B|A) = B$, condizione normalmente riassunta attraverso l'espressione $P(A \cap B) = P(A) * P(B)$ (PICCOLO, 2004).

vettore x_i; tale legame viene normalmente riassunto attraverso la funzione di collegamento g:

$$g(\mu_i) = x_i'\beta \,.$$

Dato che la funzione di collegamento è invertibile, l'espressione precedente può anche essere espressa nel seguente modo:

$$E(Y_i) = \mu_i = g^{-1}(x_i'\beta) \,,$$

in cui β è un vettore di parametri[48].

A seconda della funzione di collegamento si possono, poi, distinguere:

- Modelli tariffari additivi, in cui g è funzione identica:

$$\mu_i = x_i'\beta = \sum_{j=0}^{m} x_{ij}\beta_j \;;$$

- Modelli tariffari moltiplicativi, in cui g è funzione logaritmo:

$$\mu_i = e^{x_i'\beta} = \prod_{j=0}^{m} e^{x_{ij}\beta_j} \,.$$

Nonostante l'enorme diffusione di entrambe le tipologie di modelli, quelli maggiormente utilizzati in ambito assicurativo sono del secondo tipo, così come quelli che verranno utilizzati nel paragrafo seguente (GIGANTE, PICECH, & SIGALOTTI, 2010).

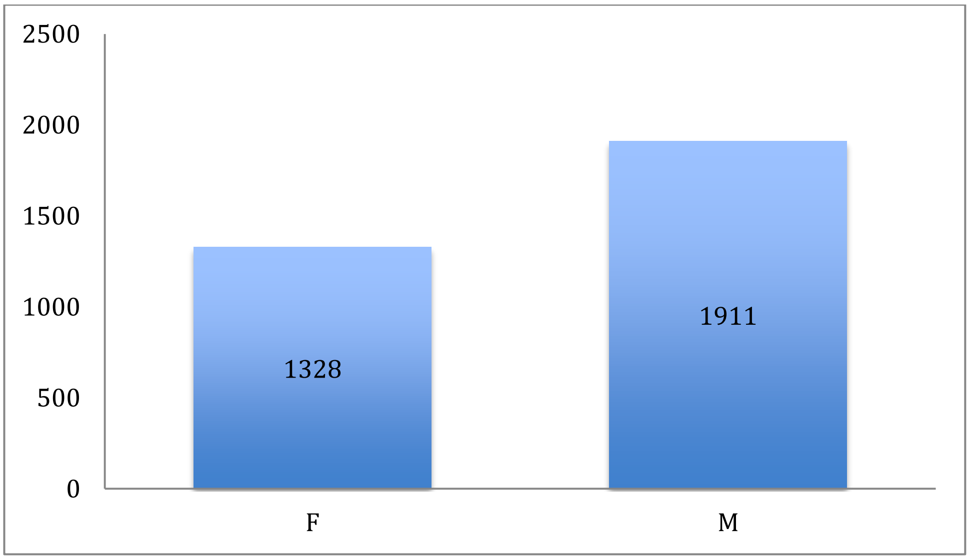
6.2. IL NUMERO ATTESO DI SINISTRI: DUE MODELLI A CONFRONTO

Per le successive analisi verrà utilizzato un campione di 3239 assicurati, con dati relativi agli ultimi 6 anni (2009-2014), frazionati in semestri; si sottolinea, inoltre, che il campione è suddiviso in 1328 femmine e 1911 maschi.

---

[48] Per ulteriori approfondimenti sul tema si rimanda a Gigante, Picech & Sigalotti (2010).

Figura 6.1: Istogramma rappresentante il campione di dati suddiviso in base al sesso



La variabile esplicativa che verrà utilizzata per analizzare il primo modello, ovvero il modello che utilizza solo dati conoscibili a priori, sarà l'età dei diversi assicurati.

Per esigenze di calcolo le diverse età sono state raggruppate in 6 classi secondo il seguente criterio[49]:

$$X_1 = \begin{cases} 1 \; se \; 18 \leq età \leq 20 \\ 0 \quad\quad altrimenti \end{cases}, \qquad X_2 = \begin{cases} 1 \; se \; 21 \leq età \leq 35 \\ 0 \quad\quad altrimenti \end{cases},$$

$$X_3 = \begin{cases} 1 \; se \; 36 \leq età \leq 45 \\ 0 \quad\quad altrimenti \end{cases}, \qquad X_4 = \begin{cases} 1 \; se \; 46 \leq età \leq 60 \\ 0 \quad\quad altrimenti \end{cases},$$

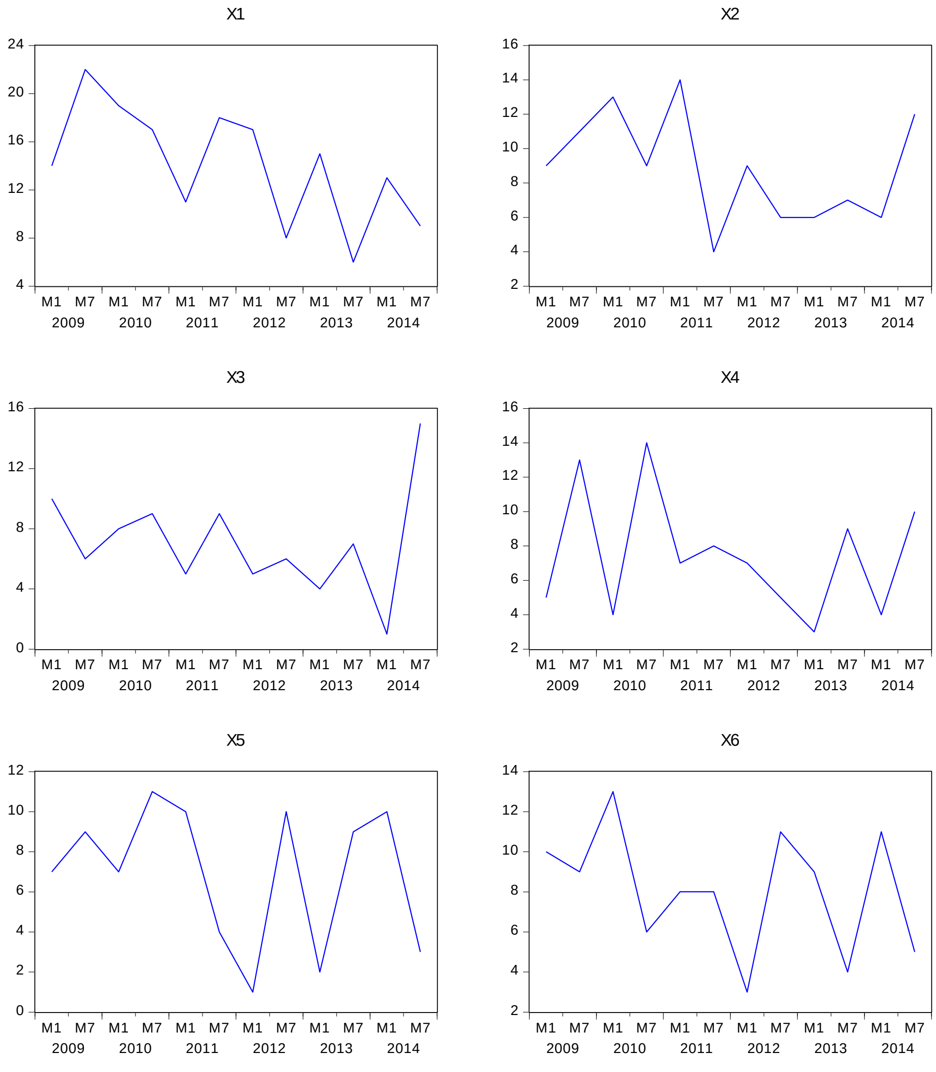
$$X_5 = \begin{cases} 1 \; se \; 61 \leq età \leq 75 \\ 0 \quad\quad altrimenti \end{cases}, \qquad X_6 = \begin{cases} 1 \quad se \; età \geq 76 \\ 0 \quad\quad altrimenti \end{cases}.$$

Dalla semplice analisi delle diverse classi d'età, come illustra la figura 6.2, emerge chiaramente che il numero di sinistri rilevato non segue un andamento costante, ma fortemente discontinuo; solo nei soggetti con età

[49] Per poter studiare le singole età, e quindi non ricorrere ai raggruppamenti, sarebbe necessaria una numerosità campionaria nettamente superiore a quella disponibile; la scelta di occupare le fasce d'età, e poi, per il modello successivo, anche le classi BM, è stata l'unica soluzione possibile per permettere l'analisi dei modelli.

inferiore ai vent'anni si osserva una sinistrosità decrescente, giustificata probabilmente dalle normative sempre più severe verso i soggetti neopatentati.

Figura 6.2: Andamento del numero di sinistri dal primo semestre del 2009 al secondo semestre del 2014, rispetto alle diverse classi d'età



Utilizzando Eviews si è quindi cercato di studiare l'andamento del numero di sinistri esclusivamente sulla base della variabile appena introdotta, andando quindi a focalizzare l'attenzione sulla prima categoria di modelli a cui si è accennato nell'introduzione al capitolo.

Come già anticipato nei capitoli precedenti, la distribuzione del numero di sinistri viene normalmente approssimata attraverso la distribuzione di Poisson, alla quale corrisponde la funzione canonica di collegamento logaritmo. Per questo motivo, per andare a calcolare $E(N|I = i)$, ovvero il numero di sinistri dato il valore assunto dalla variabile tariffaria età, si andrà ad utilizzare un modello tariffario del tipo

$$\mu_i = E(N_i) = \exp(x_i'\beta) = e^{(x_i'\beta)}.$$

Mantenendo come variabile risposta il numero annuo di sinistri N_i, per l'i-esimo assicurato, e utilizzando come regressori le determinazioni delle diverse classi di età sopra citate, si ottiene il risultato evidenziato dalla figura 6.3.

Figura 6.3: Risultato ottenuto dalla regressione, in cui compare solo la variabile di tariffazione a priori

Dependent Variable: N_SINISTRI
Method: Generalized Linear Model (Quadratic Hill Climbing)
Date: 01/18/15 Time: 11:57
Sample: 2009S1 2014S2
Included observations: 12
Family: Poisson
Link: Log
Dispersion fixed at 1
Coefficient covariance computed using observed Hessian
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	z-Statistic	Prob.
X1	0.019023	0.027337	0.695862	0.4865
X2	0.018604	0.022436	0.829205	0.4070
X3	0.022582	0.031307	0.721318	0.4707
X4	0.014598	0.052860	0.276163	0.7824
X5	0.021743	0.045363	0.479307	0.6317
X6	0.016475	0.056983	0.289117	0.7725
C	2.958352	0.365673	8.090159	0.0000

Mean dependent var	52.41667	S.D. dependent var	10.17536
Sum squared resid	3.088005	Log likelihood	-34.73197
Akaike info criterion	6.955329	Schwarz criterion	7.238191

Hannan-Quinn criter.	6.850603	Deviance	0.063247
Deviance statistic	0.012649	Restr. deviance	21.44697
LR statistic	21.38372	Prob(LR statistic)	0.001565
Pearson SSR	0.063275	Pearson statistic	0.012655
Dispersion	1.000000		

Per dare un significato ai risultati ottenuti, e quindi calcolare in concreto quale valore di sinistri potremmo attenderci per ogni classe, andiamo a risolvere l'equazione associata al modello tariffario moltiplicativo; ad esempio, ci attenderemo che il numero atteso di sinistri per la classe d'età 21-35 sia:

$$E\left(N_{X_2}\right) = e^{(2{,}958352+0{,}018604)} = 19{,}6273 \cong 20\,.$$

Tabella 6.1: Numero annuale di sinistri stimato per le diverse classi di età[50]

Età	NUMERO DI SINISTRI STIMATO
18-20	19,6362
21-35	19,6280
36-45	19,7062
46-60	19,5495
61-75	19,6897
76+	19,5862

Seppur apparentemente anomalo, il risultato ottenuto attraverso la regressione non mostra grosse differenze fra le diverse età analizzate, e questo è causato principalmente da due fattori. La prima motivazione riguarda i dati stessi analizzati, i quali evidenziano per le due fasce d'età

[50] Nella tabella 6.1, come nell'esempio precedente, si è volutamente deciso di lasciare anche i numeri decimali, con la consapevolezza che il numero di sinistri può solo essere un numero intero; la scelta è stata dettata dalla volontà di evidenziare che, pur con minime differenze, il risultato ottenuto non è uguale per le diverse classi di età.

estreme, X_1 e X_6, un'attuale contrazione del numero di sinistri rispetto al passato; situazione differente, invece, per le altre classi, che nei vari anni hanno dimostrato numerosi picchi di sinistrosità, andando inevitabilmente a intaccare la bontà delle stime effettuate.
La seconda motivazione è dovuta alla scelta, dettata da necessità di calcolo, di andare a raggruppare soggetti con diverse età all'interno della stessa classe; in questo modo si è persa parte dell'informazione relativa al profilo dei singoli assicurati.

Nella seconda tipologia di modelli si cerca di andare a definire meglio il numero di sinistri, aggiungendo alla variabile età anche la classe Bonus-Malus occupata. Sempre per esigenze di calcolo, si è scelto di aggregare le diverse classi nel modo illustrato nella tabella 6.2.

Tabella 6.2: Aggregazione delle diverse classi Bonus-Malus

CLASSI BM	C1	C2	C3	C4	C5
1	1	0	0	0	0
2	1	0	0	0	0
3	1	0	0	0	0
4	0	1	0	0	0
5	0	1	0	0	0
6	0	1	0	0	0
7	0	0	1	0	0
8	0	0	1	0	0
9	0	0	1	0	0
10	0	0	0	1	0
11	0	0	0	1	0
12	0	0	0	1	0
13	0	0	0	1	0
14	0	0	0	0	1
15	0	0	0	0	1
16	0	0	0	0	1
17	0	0	0	0	1
18	0	0	0	0	1

Tabella 6.3: Risultato ottenuto dalla regressione, con l'utilizzo simultaneo di variabili rilevate a priori ed a posteriori

Dependent Variable: N_SINISTRI
Method: Generalized Linear Model (Quadratic Hill Climbing)
Date: 01/18/15 Time: 16:14
Sample: 2009S1 2014S2
Included observations: 12
Family: Poisson
Link: Log
Dispersion fixed at 1
Coefficient covariance computed using observed Hessian
Convergence achieved after 9 iterations
No d.f. adjustment for standard errors & covariance

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C1	-0.029443	0.350152	-0.084087	0.9330
C2	-0.036841	0.415275	-0.088714	0.9293
C3	-0.047491	0.741424	-0.064054	0.9489
C4	-0.034393	0.430726	-0.079849	0.9364
C5	-0.031169	0.409617	-0.076093	0.9393
X1	0.055582	0.461597	0.120413	0.9042
X2	0.055547	0.479159	0.115926	0.9077
X3	0.057597	0.448915	0.128304	0.8979
X4	0.043044	0.357154	0.120519	0.9041
X5	0.058738	0.463941	0.126607	0.8993
X6	0.041440	0.315749	0.131242	0.8956
C	3.052595	2.030357	1.503477	0.1327

Mean dependent var	52.41667	S.D. dependent var	10.17536
Sum squared resid	1.84E-20	Log likelihood	-34.70035
Akaike info criterion	7.783392	Schwarz criterion	8.268298
Hannan-Quinn criter.	7.603862	Deviance	3.46E-14
Restr. deviance	21.44697	LR statistic	21.44697
Prob(LR statistic)	0.029025	Pearson SSR	4.26E-22
Dispersion	1.000000		

Al fine di evidenziare i risultati ottenuti, e quindi il numero di sinistri stimato per l’anno successivo sulla base delle due variabili, i risultati vengono riassunti nella tabella 6.4.

Tabella 6.4: Numero di sinistri stimato in base all'età ed alla classe Bonus-Malus di appartenenza[51]

Età	CLASSE BM	NUMERO DI SINISTRI STIMATO	NUMERO DI SINISTRI APPROSSIMATO
18-20	1-2-3	21,7309	22
18-20	11-12-13	21,6236	22
18-20	14-15-16-17-18	21,6934	22
21-35	1-2-3	21,7301	22
21-35	4-5-6-7	21,5699	22
21-35	8-9-10	21,3414	21
21-35	11-12-13	21,6228	22
21-35	14-15-16-17-18	21,6926	22
36-45	1-2-3	21,7747	22
36-45	4-5-6-7	21,6142	22
36-45	8-9-10	21,3852	21
46-60	1-2-3	21,4601	21
46-60	4-5-6-7	21,3019	21
61-75	1-2-3	21,7996	22
76+	1-2-3	21,4257	21

Si osserva, ad esempio, che per la classe tariffaria 21-35, in prima classe di merito, ci si attenderà un numero di sinistri per l'anno successivo pari a:

$$E(N|I=i,Y=j)=e^{(3{,}052595+0{,}055547-0{,}029443)}=21{,}7301\cong 22\,.$$

Il risultato ottenuto non si discosta molto da quanto ottenuto sulla base della sola classe d'età, ma presenta pur sempre una differenza.

[51] Nella tabella 6.4 non vengono analizzate tutte le possibili combinazioni tra fasce d'età e classi BM, ma esclusivamente quelle occupate da un numero considerevole di assicurati, stabilito per semplicità intorno all'1% del totale di assicurati presenti nel portafoglio analizzato. Le possibili combinazioni rimangono numerose, comunque, grazie all'introduzione della Legge Bersani; come introdotto in precedenza, infatti, tale legge permette ai nuovi assicurati di adottare la classe minore presente nel nucleo familiare (che non sempre corrisponde alla classe più bassa del sistema BM), ed in questo modo violare le normali regole evolutive.

Quanto evidenziato dal secondo modello, da un punto di vista assicurativo, non può ritenersi ancora soddisfacente; nonostante il risultato sia leggermente variato rispetto a prima, non si è ancora nella condizione di poter definire dei coefficienti in grado di rispecchiare in maniera opportuna la vera rischiosità dei soggetti assicurati.

Per osservare in che modo possano variare i risultati al variare del campione considerato, si procede con l'analisi di un campione più ristretto, e per la precisione si decide di non considerare la fascia d'età X_6, occupata dai soggetti con età superiore ai 75 anni[52].

Per poter operare un confronto fra i diversi risultati, si utilizza nuovamente lo Eviews.

Tabella 6.5: Risultato ottenuto dalla regressione della sola variabile età, tenendo in considerazione il modello ridotto

Dependent Variable: N_SINISTRI
Method: Generalized Linear Model (Quadratic Hill Climbing)
Date: 02/07/15 Time: 15:13
Sample: 2009S1 2014S2
Included observations: 12
Family: Poisson
Link: Log
Dispersion fixed at 1
Coefficient covariance computed using observed Hessian
Convergence achieved after 5 iterations

Variable	Coefficient	Std. Error	z-Statistic	Prob.
X1	0.004564	0.030580	0.149248	0.8814
X5	0.025468	0.060560	0.420536	0.6741
X4	0.033283	0.062849	0.529576	0.5964
X3	0.013667	0.047085	0.290260	0.7716
X2	0.030855	0.065114	0.473864	0.6356
C	2.864996	0.261634	10.95038	0.0000

Mean dependent var	53.41667	S.D. dependent var	17.43799
Sum squared resid	33.28514	Log likelihood	-34.90555
Akaike info criterion	6.817592	Schwarz criterion	7.060045
Hannan-Quinn criter.	6.727827	Deviance	0.640178
Deviance statistic	0.106696	Restr. deviance	65.44032
LR statistic	64.80014	Prob(LR statistic)	0.000000
Pearson SSR	0.632914	Pearson statistic	0.105486

[52] Nell'operare tale scelta si è tenuto in considerazione il fatto che i soggetti appartenenti a tale classe vanno ad occupare esclusivamente le prime classi di merito. Di conseguenza non ci si aspetta una perdita di informazioni rilevanti, ma una stima più accurata del modello.

Dispersion	1.000000

Tabella 6.6: Numero annuale stimato di sinistri per il campione ridotto

Età	NUMERO DI SINISTRI STIMATO
18-20	17,6293
21-35	18,0017
36-45	18,1429
46-60	17,7905
61-75	18,0989

Dai risultati presenti nella tabella 6.6 emerge chiaramente che, rispetto a quanto evidenziava la tabella 6.1, c'è una riduzione del numero atteso di sinistri. In particolare, analizzando la classe d'età 21-35, si osserva che il numero atteso di sinistri è:

$$E(N_{25}) = e^{(2,864996+0,030855)} = 18,0017 \cong 18\,.$$

Dinnanzi a tale risultato, seppur diverso da quello ottenuto in precedenza, è difficoltoso trarre delle conclusioni soddisfacenti, poiché si osserva che le diverse classi di età sono tutte orientate verso il medesimo numero di sinistri.

Per avere una visione più completa dell'effetto provocato dall'eliminazione di alcuni dati, si continua l'analisi del numero di sinistri attraverso il secondo modello considerato, ovvero il modello che tiene in considerazione anche la classe Bonus-Malus di appartenenza.

Tabella 6.7: Risultato ottenuto dalla regressione del campione ridotto, con utilizzo simultaneo di variabili rilevate a priori ed a posteriori

Dependent Variable: N_SINISTRI
Method: Generalized Linear Model (Quadratic Hill Climbing)
Date: 02/07/15 Time: 15:30
Sample: 2009S1 2014S2
Included observations: 12
Family: Poisson
Link: Log
Dispersion fixed at 1
Coefficient covariance computed using observed Hessian

Convergence achieved after 7 iterations

Variable	Coefficient	Std. Error	z-Statistic	Prob.
X4	-0.002418	0.457530	-0.005284	0.9958
X3	-0.040371	0.387558	-0.104167	0.9170
X2	0.004184	0.472758	0.008850	0.9929
X1	-0.047860	0.312001	-0.153396	0.8781
X5	0.003725	0.273226	0.013634	0.9891
C5	0.039750	0.302339	0.131476	0.8954
C4	0.051873	0.384671	0.134849	0.8927
C3	0.043503	0.394985	0.110140	0.9123
C2	0.033782	0.294251	0.114805	0.9086
C1	0.030669	0.311678	0.098401	0.9216
C	2.812503	0.823099	3.416970	0.0006

Mean dependent var	53.41667	S.D. dependent var	17.43799
Sum squared resid	8.263076	Log likelihood	-34.68367
Akaike info criterion	7.613945	Schwarz criterion	8.058443
Hannan-Quinn criter.	7.449376	Deviance	0.196411
Deviance statistic	0.196411	Restr. deviance	65.44032
LR statistic	65.24391	Prob(LR statistic)	0.000000
Pearson SSR	0.197292	Pearson statistic	0.197292
Dispersion	1.000000		

Tabella 6.8: Numero stimato di sinistri nel campione ridotto, in base all'età ed alla classe Bonus-Malus di appartenenza,

Età	CLASSE BM	NUMERO DI SINISTRI STIMATO	NUMERO DI SINISTRI APPROSSIMATO
18-20	1-2-3	16,3677	16
18-20	11-12-13	16,7185	17
18-20	14-15-16-17-18	16,5170	17
21-35	1-2-3	17,2421	17
21-35	4-5-6-7	17,2959	17
21-35	8-9-10	17,4648	17
21-35	11-12-13	17,6116	18
21-35	14-15-16-17-18	17,3994	17
36-45	1-2-3	16,4908	16
36-45	4-5-6-7	16,5422	17
36-45	8-9-10	16,7038	17
46-60	1-2-3	17,1287	17
46-60	4-5-6-7	17,1821	17
61-75	1-2-3	17,2342	17

Attraverso i risultati presenti nella tabella 6.8, si osserva che la scelta di ridurre il campione d'interesse ha portato ad un'ulteriore contrazione del numero stimato di sinistri; infatti, andando nuovamente ad analizzare i soggetti con età compresa tra i 21 e 35 anni, ed appartenenti alla prima classe di merito, si osserva che il numero atteso di sinistri è pari a:

$$E(N|I = i, Y = j) = e^{(2{,}812503+0{,}004184+0{,}030669)} = 17{,}2421 \cong 17\,.$$

Nonostante grazie all'introduzione della variabile di tariffazione Bonus-Malus si ottenga un risultato più coerente con i livelli di sinistrosità rilevati negli anni[53], è evidente che non si possa ritenere soddisfacente quanto ottenuto dallo studio dei due modelli.

Si conclude quindi che per ottenere un modello tariffario in grado di prevedere il corretto numero di sinistri, non è sufficiente una numerosità campionaria adeguata, ma anche la presenza di un maggior numero di variabili di tariffazione, variabili delle quali si è già parlato nei capitoli precedenti.

[53] Si fa riferimento alle statistiche presenti nel sito www.ania.it.

CONCLUSIONE

Il tema centrale, attorno a cui verte tutta la trattazione della tesi, sono i premi assicurativi del ramo danni, con particolare attenzione per il ramo Responsabilità Civile Autoveicoli.
Quest'ultima forma assicurativa, divenuta obbligatoria con la legge n. 990 del 24 Dicembre 1969, prevede che l'attribuzione del premio sia legata ad una dettagliata analisi delle caratteristiche distintive dei diversi soggetti assicurati, caratteristiche che vengono poi implementate per il calcolo del premio attraverso appurati modelli tariffari.
Ripercorrendo l'intero processo di tariffazione adottato dalle diverse compagnie assicurative, si è cercato di chiarire non solo l'effettiva utilità delle diverse informazioni raccolte, ma soprattutto l'effettivo contributo delle stesse alla creazione del premio assicurativo finale.
Quello che però ha permesso di mettere in luce maggiormente il processo di determinazione dei premi assicurativi è stato il capitolo conclusivo, in cui si cerca di effettuare una simulazione. Ispirandosi ai modelli tariffari proposti da Gigante, Picech & Sigalotti (2010), si è cercato di prevedere il numero di sinistri per l'anno successivo, sulla base della variabile età in primis, per poi passare ad un risultato migliore grazie all'informazione relativa alla classe Bonus-Malus occupata dal campione di assicurati.
Nel primo modello si è scelto di utilizzare, appunto, un'informazione ottenibile dalla semplice tariffazzione a priori, variabile che racchiude al suo interno un'informazione non di poco conto sul soggetto assicurato, ma che si dimostra eccessivamente scarna da un punto di vista della bontà della stima ottenuta. Si osserva, infatti, che suddividendo il portafoglio di assicurati in fasce d'età, non si potrebbe giungere ad un risultato corretto in base a quest'unico elemento, poiché l'insufficienza di dati disponibili porterebbe a prevedere un egual numero di sinsitri per qualsiasi fascia d'età considerata.

La scelta di includere come seconda variabile l'informazione relativa all'esperienza degli assicurati, racchiusa nella classe Bonus-Malus di appartenenza, deriva sempre da un modello tariffario proposto da Gigante, Picech & Sigalotti (2004); la valutazione che ne deriva non si discosta molto dal risultato del modello precedente, ma rappresenta pur sempre un miglior approccio alla materia, permettendo di distinguere le fasce d'età con maggior propensione ad effettuare sinistri.

La letteratura si è comunque dimostrata negli anni molto interessata all'argomento, ed il numero di modelli tariffari proposti ne è la testimonianza; ma nonostante ciò la scelta di quali modelli tariffari adottare sarà comunque a discrezione delle compagnie assicurative, poiché nonostante i diversi risultati evidenziati, non esiste ancora un modello di tariffazione che possa essere considerato preferibile agli altri.

BIBLIOGRAFIA

AA.VV. (2008). Credibility Practice Note. UNited States of America: American Accademy of Actuaries.

ACI. (2013). *Incidentalità.* Tratto il giorno Gennaio 9, 2015 da http://www.aci.it/laci/studi-e-ricerche/dati-e-statistiche/incidentalita.html

ANIA. (2014). *L'assicurazione italiana 2013-2014.* Tratto il giorno Gennaio 2, 2015 da http://www.ania.it/export/sites/default/it/pubblicazioni/rapporti-annuali/2014/LAssicurazione-italiana-2013-2014.pdf

BAILEY, A. L. (1945). A Generalized Theory of Credibility. *Proceedings of the Casualty Actuarial Society , 32*, 13-20.

BAILEY, A. L. (1950). Credibility Procedures, Laplace's generalization of Bayes' Rule and the Combination of Collateral Knowledge with Observed data. *Proceedings of Casualty Actuarial Society , 37*, 7-23.

BAILEY, R. A., & SIMON, L. J. (1960). Two studies in automobile insurance ratemaking. *Astin Bullettin , 1* (IV), 192-217.

BERMUDEZ, L., DENUIT, M., & DHAENE, J. (2001). Exponential Bonus-Malus Systems integrating a priori risk classification. *Journal of Actuarial practice , 9.*

BORCH, K. H. (1990). *Economics of insurance.* Netherlands: Elsevier science publishers B.V.

BORCH, K. (1974). *The Mathematical Theory of Insurance.* United States of America: Lexington Books.

BROUHNS, N., GUILLEN, M., DENUIT, M., & PINQUET, J. (2002). *Optimal bonus-malus scales in segmented tariffs.* Unversitè catholique de louvain, Institute de statistique.

BUHLMAN, H., & GISLER, A. (2005). *A Course in Credibility Theory and its Applications.* The Netherlands: Springer.

BUHLMANN, H. (1967). Experience Rating and Credibility. *Astin Bulletin , 4*, 199-207.

CASTELLANI, G. (2011). Corso di Teoria del Rischio. *Calcolo del premio.* Università La Sapienza , Roma.

CERCHIARA, R. R. (2014). Corso di metodi matematici dell'economia e delle scienze attuariali e finanziarie. *Classi di Rischio e Personalizzazione del premio.* Università della Calabria.

CERCHIARA, R. R. (2014). Corso di metodi matematici dell'economia e delle scienze attuariali e finanziarie. *Il modello markoviano per la rappresentazione del Sistema Bonus-Malus.* Università della Calabria, Calabria.

CERCHIARA, R. R. (2014). Corso di metodi matematici dell'economia e delle scienze attuariali e finanziarie. *La tariffazione RCA in Italia.* Università della Calabria.

CORAIN, L. (2011). Corso di Ingegneria Civile. *Metodi probabilistici.* Facoltà di Ingegneria, Padova.

DABONI, L. (1989). *Lezioni di tecnica attuariale delle assicurazioni contro i danni.* Trieste: Lint.

DE FERRA, C. (1995). *L'assicurazione: nozioni, concetti, basi matematiche.* Milano: La Musicografica Lombarda-Brugherio.

DE FINETTI, B. (1964). Sulla teoria della credibilità. *Giornale dell'Istituto italiano degli attuari , 27*, 219-231.

DENUIT, M., MARECHAL, X., PITREBOIS, S., & WALHIN, J.-F. (2007). *Actuarial Modelling of claim counts.* England: John Wiley & Sons Ltd.

FELLER, W. (1966). *An Introduction to Probability Theory and its Application* (Vol. 2). New York: John Wiley & Sons.

FERRARA, G. (2012). *Modelli matematici per i processi di sinistri.* Tratto il giorno Gennaio 4, 2015 da http://www.sifa-attuari.it/materiale/GFCorso_251012.pdf

FRANGOS, N. E., & VRONTOS, S. D. (2001). Design of optimal Bonus-Malus Systems with a frequency and a severity component on an individual basis in automobile insurance. *Astin Bulletin , 31* (1), 1-22.

GERBER, H. U. (1995). A Theacher's Remark on Exact Credibility. *Astin Bullettin , 25* (2), 189-192.

GERBER, H. U. (1979). *An Introduction to matematical Risk Theory.* United States of America: S.S. Huebner Foundation for insurance education.

GIGANTE, P., PICECH, L., & SIGALOTTI, L. (2000). Bonus-malus experience rating and rating factors. *Proceedings of the XXXI International ASTIN Colloquium*, (p. 399-412). Porto Cervo.

GIGANTE, P., PICECH, L., & SIGALOTTI, L. (2004). La classe bonus-malus come variabile esplicativa nella tariffazione R.C.A. *Atti del IX convegno di Teoria del Rischio, 26 Giugno 2002*, (p. 117-139). Campobasso.

GIGANTE, P., PICECH, L., & SIGALOTTI, L. (2010). *La tariffazione nei rami danni con modelli lineari generalizzati.* Trieste: EUT.

GIGANTE, P., PICECH, L., & SIGALOTTI, L. (2003). *Revisione dei premi con sistemi bonus-malus per portafogli ripartiti in classi tariffarie.* Dip. di Mat. App. alle Sc. Ec. Stat. e Att."B. De Finetti". Trieste: Università di Trieste.

GOOVAERTS, M. J., VYLDER, F., & HAEZENDONCK, J. (1991). *Insurance premiums.* Netherlands: Elsevier Science Publishers B. V.

GOULET, V. (1997). On Approximations in Limited Fluctuation Credibility. *Proceedings of the Casualty Actuarial Society , 84*, 533-552.

GOULET, V. (1998). Principles and Application of Credibility Theory. *Journal of Actuarial Practice , 6*, 5-62.

HERZOG, T. N. (1989). The Bayesian Model versus Buhlmann's model. *Transactions of Society of Actuaries , 41*, 43-88.

HERZOG, T. N. (2008). The Rev. Thomas Bayes and Credibility Theory. *Contingencies* , 40-42.

IODICE D'ENZA, A. (2009). Corso di Statistica. *Esercitazione.* Università degli Sudi di Cassino, Cassino.

LEMAIRE, J. (1995). *Bonus-Malus Systems in Automobile insurance.* United States of America: Kluwer Academic Publishers.

LEMAIRE, J., & ZI, H. (1994). A comparative analysis of 30 Bonus-Malus Systems. *Astin Bulletin , 24* (2), 287-309.

LOMAZZI, M. (2005). *La riscoperta della forma "a primo rischio relativo".* Tratto il giorno Novembre 6, 2014 da TUA assicurazioni: http://www.tuaassicurazioni.it/public/file/2004_12_LA%20RISCOPERTA%20DELLA%20FORMA.....ASSINEWS.pdf

LONGLEY-COOK, L. H. (1962). An Introduction to Credibility Theory. *Proceedings of the Casualty Actuarial Society , 49*, 194-221.

LORETI, M. (1998). *Teoria degli errori e fondamenti di statistica. Introduzione alla fisica sperimentale.* Bologna: Zanichelli.

MAGGINA, F. (2008). Scomposizione delle misure di efficienza dei sistemi bonus-malus. *Università La Sapienza.* Dipartimento di Scienze Attuariali e Finanziarie, Roma.

MAHLER, H. C., & DEAN, C. G. (2001). Credibility: Chapter 8. *Foundation of Casualty Actuarial Science* .

MINKOVA, L. (2010). *Insurance Risk Theory.* Tratto il giorno Novembre 6, 2014 da https://www.fmi.uni-sofia.bg/sms/fam/insurance-risk-theory-lectures/InsuranceRisk.pdf

MOWBRAY, A. H. (1914). How Extensive a Payroll Exposure is Necessary to Give a Dependable Pure Premium? *Proceedings of the Casualty Actuarial Society , 1*, 24-30.

NORBERG, R. (1976). A Credibility Theory for Automobile Bonus System. *Scandivian Actuarial Journal* , 92-107.

NORBERG, R. (1975). Credibility premium plans which make allowance for bonus hunger. *Scandinavian Actuarial Journal* , 73-86.

NORBERG, R. (2004). Credibility Theory. *Encyclopedia of Actuarial Science*.

NORBERG, R. (1979). The Credibility Approach to Experience Rating. *Scandinavian Actuarial Journal* , 181-221.

OUTREVILLE, J. F. (1998). *Theory and Practice of Insurance.* United States of America: Kluwer Academic Publishers.

PARRINI, C. (2012). Corso di Laurea in Scienze Statistiche ed Economiche. *La gestione tecnico attuariale delle imprese di assicurazione.* Università degli Studi Milano Bicocca.

PICCOLO, D. (2004). *Statistica per le decisioni.* Bologna: Il Mulino.

PITACCO, E. (2000). *Elementi di Matematica delle assicurazioni.* Trieste: Lint.

PORZIO, C., PREVIATI, D., COCOZZA, R., MIANI, S., & PISANI, R. (2011). *Economia delle imprese assicurative.* Milano: McGraw-Hill.

ROLSKI, T., SCHMIDLI, H., SCHMIDT, V., & TEUGELS, J. (1998). *Stochastic Processes for Insurance and Finance.* Chichester: John Wiley & Sons.

S.I.F.A., s. (2014). *La tariffazione nel ramo RCA: operatività della compagnia e modalità di verifica dell'AIRCA, in base alla normativa vigente.* Tratto il giorno Gennaio 9, 2015 da http://www.sifa-attuari.it/materiale/Esempio_fabbisogno_Sammartini.pdf

SCHNIEPER, R. (1995). On the Estimation of the Credibility Factor: a Bayesian Approach. *Astin Bulletin , 25* (2), 137-151.

SIGALOTTI, L., & PICECH, L. (1994). La scelta delle variabili tariffarie e la personalizzazione. . *Le tariffe R.C.Auto nella fase di passaggio da tariffe amministrative a tariffe libere.* Istituto Italiano degli Attuari, Roma.

SUNDT, B. (1984). *An Introduction to Non-Life Insurance Mathematics.* Karlsruhe: Verlag Versicherungswirtschaft e.V.

UBI-Assicurazioni. (2014). *Norme tariffarie-Tariffa 2014 (2) - Ed. 01.07.2014.* Tratto il giorno Novembre 15, 2014 da http://www.ubiassicurazioni.it/media/WQQ_-_Norme_tariffarie_-_Ed._01_07_2014.pdf

VENTER, G. G. (2003). Credibility Theory for Dummies. *Casualty Actuarial Society* , 621-627.

VERICO, P. (2002). Bonus-malus systems:"lack of transparency" and Adequacy Measure. *Astin Bulletin , 32* (2), 315-318.

VYLDER, F., GOOVAERTS, M., & HAEZENDONCK, J. (1984). *Premium Calculation in Insurance.* Netherlands: D. reidel Publishing Company.

WHITNEY, A. W. (1918). The Theory of Experience Rating. *Proceedings of the Casualty Actuarial Society , 4*, 274-292.

SITOGRAFIA

www.aci.it
www.amicoassicuratore.it
www.ania.it
www.istat.it

RINGRAZIAMENTI

Concluso il mio percorso universitario, voglio qui ringraziare tutti coloro che, in modi diversi, hanno contribuito al raggiungimento di questo risultato.
Ringrazio innanzitutto la Professoressa Antonella Basso, che in questi mesi mi ha seguito con estrema professionalità e disponibilità, mettendo costantemente a disposizione non solo il suo tempo ma anche la sua esperienza.
Ringrazio la mia Mamma, sicuramente la migliore mamma al mondo, che con il suo infinito amore è sempre riuscita a farmi tornare il sorriso anche quando sembrava impossibile. Ringrazio mio Papà, che ha sempre fatto il possibile per non farmi mai mancare niente, e nonostante cerchi sempre di nascondere i suoi sentimenti, so quanto è orgoglioso di me.
Lorenzo, ringrazio anche te, che mi hai dato tutto l'amore e l'appoggio che potessi desiderare; sappi che sono fiera di te, così come tu lo sei di me.
Ringrazio di cuore tutti gli zii e i cugini che mi hanno aiutato, che si sono preoccupati per me e per i miei studi, che mi hanno fatto sorridere ed hanno sempre creduto in me. Ve ne sarò sempre riconoscente.
Ringrazio i nonni Anna e Bruno, che con orgoglio, raccontano sempre di avere una Dottoressa in famiglia; e mi sento di ringraziare anche i nonni che non ci sono più, che purtroppo non hanno potuto condividere questa esperienza con me, ma da lassù mi hanno sempre protetta.

Grazie di cuore a tutti voi.