



UNIVERSITÀ DEGLI STUDI
DI PERUGIA

ISSN 2385-2275

**Working papers of the
Department of Economics
University of Perugia (IT)**

**Sulla stima degli Undertaking
Specific Parameters e la verifica
delle ipotesi**

Massimo De Felice

Franco Moriconi

Working paper No. 9

April 2015

Sulla stima degli Undertaking Specific Parameters e la verifica delle ipotesi

Massimo De Felice – Sapienza, Università di Roma

Franco Moriconi – Università di Perugia

Indice

I	I modelli teorici e la verifica delle ipotesi sottostanti	3
1.	I modelli teorici alla base dei metodi standardizzati per gli USP	3
1.a.	Il modello di riferimento per il Method 1 (Modello M1) . . .	3
1.b.	Il modello di riferimento per il Method 2 (Modello M2) . . .	7
2.	La verifica delle ipotesi dei modelli teorici	9
2.a.	Verifica delle ipotesi del modello M1	9
2.a.1.	Ipotesi sulla media	10
2.a.2.	Ipotesi sulla varianza	11
2.a.3.	Un esempio di verifica “marketwide” dell’ipotesi M1V	13
2.a.4.	Ipotesi sulla distribuzione	14
2.a.5.	Un confronto per simulazione delle metodologie di test	18
2.a.6.	Appropriatezza del metodo di massima verosimiglianza	23
2.b.	Verifica delle ipotesi del modello M2	23
2.b.1.	Ipotesi sulla media e la varianza condizionate	27
2.b.2.	Ipotesi di indipendenza. Test sui residui “time series”	28
2.b.3.	Ipotesi di indipendenza. Test sui residui di Pearson . .	31
II	Applicazione ai dati d’impresa	33
3.	Premium Risk (modello M1)	33
3.a.	Definizione dei dati in input	33
3.b.	Applicazione del metodo	34
3.c.	Sulla tecnica di minimizzazione	35

4. Reserve Risk – modello M1	35
4.a. Definizione dei dati in input	35
4.b. Applicazione del metodo	37
4.c. Sulla tecnica di minimizzazione	37
5. Reserve Risk – modello M2	37
5.a. Definizione dei dati in input	37
5.b. Applicazione del metodo	38

Premessa – Questo documento illustra il processo di calcolo per la stima degli *Undertaking-Specific Parameters* (USP) definiti nel regime Solvency II, in collegamento coi sottostanti fondamenti teorici. Gli USP considerati sono le deviazioni standard unitarie relative ai sottomoduli “premium risk” e “reserve risk” dell’assicurazione contro i danni; non sono compresi nell’analisi i fattori di aggiustamento per la riassicurazione non proporzionale.

Per ciascun metodo di calcolo delle deviazioni standard unitarie è ripresa l’impostazione formale proposta dalla Commissione Europea; sono richiamati i principi teorici, sono descritte le metodologie adeguate per la verifica delle ipotesi e i test sulla bontà di adattamento ai dati, sono definiti i dati necessari al calcolo e commentati gli aspetti computazionali rilevanti.

Questo documento ha finalità pratica; coordina “documentazione utile” su metodologie, criteri, algoritmi, linee di analisi per “la determinazione dei parametri specifici” (come richiesto dall’IVASS nel punto g di [14]). Per renderne più immediata l’utilizzazione sono occasionalmente richiamate in modo sintetico nozioni – per esempio sui test delle ipotesi – considerate standard nella prassi operativa.

Parte I

I modelli teorici e la verifica delle ipotesi sottostanti

1 I modelli teorici alla base dei metodi standardizzati per gli USP

Per il calcolo delle deviazioni standard unitarie “undertaking specific” gli Atti Delegati prescrivono due “standardised method”. Il “Method 1” può essere utilizzato sia per il sottomodulo di premium risk che per quello di reserve risk; il “Method 2” è un approccio alternativo applicabile al solo sottomodulo di reserve risk. Ciascuno dei due metodi è basato su uno specifico modello stocastico di riferimento, di cui si propone una descrizione sintetica.

1.a Il modello di riferimento per il Method 1 (Modello M1)

Il modello di riferimento per la definizione delle deviazioni standard unitarie di tipo USP secondo il Method 1 è stato definito dal “Joint Working Group on Non-Life and Health NSLT Calibration” (JWG) in [12], all’interno dell’attività di sperimentazione che ha condotto alla quantificazione numerica (“calibration”) delle deviazioni standard unitarie “marketwide” per il premium e il reserve risk prescritte nella Standard Formula.

Il modello teorico sottostante al Method 1 (qui “modello M1”) è uno di quattro modelli alternativi considerati e testati dal JWG nell’attività di calibrazione sui dati di mercato europei. Relativamente a uno specifico “segment” dell’attività non-life, questi modelli considerano una variabile aleatoria Y di cui si vuole determinare – e stimare – la varianza, a partire dalle sue relazioni teoriche con una variabile esplicativa X , che agisce come misura di volume. Nelle applicazioni al premium risk la variabile dipendente Y sarà identificata nel costo aggregato dei sinistri di una fissata generazione, e la variabile indipendente X sarà il corrispondente livello dei premi di competenza. Nelle applicazioni al reserve risk le variabili X e Y saranno il valore degli impegni (a “costo ultimo”) stimati all’inizio e alla fine, rispettivamente, di un fissato anno di esercizio per sinistri accaduti negli esercizi precedenti.

Il modello M1, scelto da EIOPA per il calcolo degli USP successivamente allo studio di calibrazione, è quello caratterizzato dal JWG come classe dei “Lognormal Models, Second Variance Parametrisation”. È basato sulle seguenti ipotesi.

M1M - *Ipotesi sulla media:*

$$\mathbf{E}(Y) = \beta X .$$

M1V - *Ipotesi sulla varianza:*

$$\mathbf{Var}(Y) = \beta^2 \sigma^2 [(1 - \delta)\bar{X}X + \delta X^2] ,$$

con:

$$\bar{X} = \frac{1}{T} \sum_{t=1}^T X_t ,$$

(media aritmetica su una serie storica annuale X_1, X_2, \dots, X_T di valori osservati di X).

M1D - *Ipotesi sulla distribuzione:*

$$\ln Y \sim \text{Normale}(\mu, \omega) ,$$

con:

$$\omega = \ln \{1 + \sigma^2 [(1 - \delta)\bar{X}X + \delta X^2]\} , \quad \mu = \ln(\beta X) - \frac{\omega}{2} .$$

Per la stima del modello è richiesta la stima dei parametri β, σ e δ . In particolare:

- δ è un *mixing parameter*, compreso fra 0 e 1; se $\delta = 1$ la varianza di Y ha una relazione di tipo quadratico con X , mentre se $\delta = 0$ la varianza di Y risulta proporzionale a X .

- σ approssima il (in pratica, coincide col) coefficiente di variazione di Y , $\mathbf{Cv}(Y) = \mathbf{Std}(Y)/\mathbf{E}(Y)$; una sua stima, quindi, quantifica il valore della deviazione standard unitaria specifica dell'impresa per il premium o per il reserve risk (a seconda dell'interpretazione attribuita alle variabili aleatorie X e Y).

Osservazione. L'ipotesi M1V, che specifica la varianza di Y come funzione quadratica di X , è giustificata dal JWG come una estensione “realistica” del classico modello Poisson Composto utilizzato per descrivere l'underwriting risk nella tradizione attuariale (si veda per es. [20], cap. 3). Come è noto, il modello Poisson Composto, che ha parametri costanti nel tempo, prevede per il costo dei sinistri aggregato una media e una varianza lineari come funzioni del volume del portafoglio. Se si passa da un'ipotesi di parametri costanti a un'ipotesi di parametri variabili secondo un processo stocastico (stazionario), si arriva a un'espressione ancora lineare per la media, e a un'espressione del tipo $\mathbf{Var}(Y) = \sigma_1^2 \bar{X}X + \sigma_2^2 X^2$ per la varianza. Se si aggiunge l'ipotesi che $\mathbf{Var}(Y)$ sia proporzionale a β^2 (è questa la “second variance parametrisation”), si giunge, con alcune manipolazioni, alla forma M1V. Questa forma implica che il coefficiente di variazione di Y è indipendente da β . Consente anche di ottenere stime di massima verosimiglianza dei parametri senza ricorrere a procedure di ottimizzazione troppo complesse o infattibili. Questo è probabilmente il motivo per cui il modello M1 è stato scelto da EIOPA per la stima degli USP da parte delle imprese. ■

Osservazione. Va precisato che nell'esperienza di calibrazione effettuata dal JWG la costante \bar{X} è definita come la media aritmetica delle osservazioni X presa su tutte le imprese del mercato. Indicando cioè con $X_{t,i}$ la t -esima osservazione della i -esima impresa operante sul mercato di riferimento (nel segmento considerato), in [12] si è posto:

$$\bar{X} = \frac{\sum_{i=1}^N \sum_{t=1}^{T_i} X_{t,i}}{\sum_{i=1}^N T_i},$$

dove T_i è il numero di osservazioni disponibili per la compagnia i . L'introduzione di questo fattore nell'espressione della varianza è stata effettuata per rendere il coefficiente $\beta^2 \sigma^2 (1 - \delta)$ indipendente dalla scala monetaria. Nella trasposizione del modello del JWG da una logica di mercato a una logica di compagnia, la quantità \bar{X} è stata ridefinita, negli Atti Delegati, come media individuale. Questa scelta può aumentare ulteriormente l'instabilità del modello per serie storiche con poche osservazioni; ha inoltre conseguenze esplicite nei test sulla verifica dell'ipotesi M1V (si veda il paragrafo 2.a.2). ■

Può essere utile richiamare la struttura del modello M1 con qualche dettaglio, dato che il documento del JWG si limita a fornire una presentazione

unificata dell'intero pacchetto dei modelli sperimentati; inoltre, la forma in cui il modello è proposto negli Atti Delegati contiene, rispetto al documento originario, una riparametrizzazione della funzione di stima.

La stima dei parametri del modello M1 è ottenuta col metodo della massima verosimiglianza, a partire da una serie storica di osservazioni:

$$(\mathbf{X}, \mathbf{Y}) = \{(X_t, Y_t); t = 1, 2, \dots, T\},$$

specifica dell'impresa; queste osservazioni debbono essere considerate come realizzazioni indipendenti della variabile aleatoria doppia (X, Y) . Si indichi con $\pi = \omega^{-1}$ la *precisione* (reciproco della varianza) di $\ln Y$. Se si osserva che la v.a.

$$u = \ln(Y/X) + 1/(2\pi) - \ln \beta,$$

ha distribuzione normale con media zero e varianza $1/\pi$, si verifica facilmente che massimizzare la verosimiglianza di Y equivale a minimizzare rispetto a β, σ e δ , dati (\mathbf{X}, \mathbf{Y}) , la "funzione di perdita" (*loss function, criterion function*):

$$\ell(\beta, \sigma, \delta) = \sum_{t=1}^T \pi_t u_t^2 - \sum_{t=1}^T \ln \pi_t, \quad (1.1)$$

dove, per $t = 1, 2, \dots, T$, è $u_t = \ln(Y_t/X_t) + 1/(2\pi_t) - \ln \beta$ e:

$$\pi_t = \frac{1}{\ln \{1 + \sigma^2 [(1 - \delta)\bar{X}/X_t + \delta]\}}.$$

Minimizzando la (1.1) rispetto a $\ln \beta$ si ottiene:

$$\ln \beta = \frac{T/2 + \sum_{t=1}^T \pi_t \ln(Y_t/X_t)}{\sum_{t=1}^T \pi_t}.$$

Utilizzando questa espressione, la minimizzazione della criterion function può essere riparametrizzata in modo da ridursi a un problema in due variabili. Definendo il parametro $\gamma = \ln(\sigma/\beta)$ si ha:

$$\pi_t(\delta, \gamma) = \frac{1}{\ln \{1 + [(1 - \delta)\bar{X}/X_t + \delta] e^{2\gamma}\}},$$

e, dalla (1.1):

$$\sigma(\delta, \gamma) = e^\gamma - \beta = \exp \left[\gamma + \frac{T/2 + \sum_{t=1}^T \pi_t(\delta, \gamma) \ln(Y_t/X_t)}{\sum_{t=1}^T \pi_t(\delta, \gamma)} \right]. \quad (1.2)$$

Quindi la criterion function diventa:

$$\ell(\delta, \gamma) = \sum_{t=1}^T \pi_t(\delta, \gamma) \left\{ \ln \frac{Y_t}{X_t} + \frac{1}{2\pi_t(\delta, \gamma)} + \gamma - \ln [\sigma(\delta, \gamma)] \right\}^2 - \sum_{t=1}^T \ln [\pi_t(\delta, \gamma)].$$

Questa funzione va minimizzata nell'intervallo $D = \{\delta \in [0, 1], \gamma \in \mathbb{R}\}$ utilizzando una adeguata procedura di ottimizzazione numerica. I valori $\hat{\delta}$ e $\hat{\gamma}$ così ottenuti sono le stime di parametro che, sostituite nella (1.2), forniscono la stima di massima verosimiglianza $\hat{\sigma}(\hat{\delta}, \hat{\gamma})$ richiesta per la deviazione standard unitaria USP relativa al segmento considerato.

Osservazione. Tra i diversi approcci alla stima, il metodo della massima verosimiglianza è quello dotato delle migliori proprietà teoriche e delle migliori caratteristiche di coerenza (almeno da un punto di vista bayesiano). Per una valida applicazione pratica è però necessario che il punto di massimo della likelihood (minimo della loss function) sia individuabile in modo efficiente e univoco. Nel modello M1 ciò equivale a richiedere che la procedura di minimizzazione della funzione $\ell(\delta, \gamma)$ – la cui forma, ovviamente, dipende dai dati (\mathbf{X}, \mathbf{Y}) – abbia adeguate proprietà di convergenza. ■

Nelle prescrizioni degli Atti Delegati, una volta effettuata la minimizzazione della loss function, la stima $\hat{\sigma}(\hat{\delta}, \hat{\gamma})$ viene poi moltiplicata per il “fattore di aggiustamento” $\sqrt{(T+1)/(T-1)}$; la stima aggiustata viene infine miscelata col parametro marketwide previsto dalla Standard Formula, tramite il fattore di credibilità c prescritto da EIOPA, dipendente dalla lunghezza T della serie storica utilizzata.

1.b Il modello di riferimento per il Method 2 (Modello M2)

Il modello di riferimento per la definizione delle deviazioni standard unitarie di tipo USP con il secondo metodo standardizzato (Method 2) è un modello di loss reserving ampiamente attestato in letteratura, noto come “Merz-Wüthrich model” [19]. Anche questo modello è stato sperimentato dal JWG nella sua attività di calibrazione dei parametri marketwide per il reserve risk.

A parte una inessenziale modifica nelle ipotesi tecniche¹, il modello di Merz e Wüthrich (qui denominato anche “Modello M2”) coincide col classico modello Distribution-Free Chain Ladder (DFCL) proposto da Mack nel 1993 [17]. Il modello viene però utilizzato con una logica differente da quella originaria; nel modello M2 il *mean square error of prediction* (MSEP), invece che essere considerato relativamente al “full run-off” secondo l'impostazione tradizionale, è calcolato in una logica “one-year”, relativamente al risultato tecnico (*Claims Development Result*, CDR) dell'esercizio corrente. Il passaggio da una “long-term view” a una logica di orizzonte annuale è necessario per rendere la misurazione dell'incertezza coerente col punto di vista adottato in Solvency II.

Osservazione. La logica dell'orizzonte annuale come punto di vista coerente con gli standard di solvibilità era già stata introdotta nel 2003 in [4],

¹Invece dell'ipotesi di markovianità (si veda la successiva ipotesi M2CM), nel DFCL si fanno ipotesi solo su media e varianza.

sotto una diversa denominazione – Year-End Expectation (YEE), invece che CDR – e con riferimento al modello Over-Dispersed Poisson (ODP), invece che al DFCL. Le formule esplicite per il MSEF del DFCL in versione YEE sono state ricavate nel 2006 anche in [5]. La logica YEE è stata applicata in un “field study” condotto dall’ISVAP nel 2006 [9] sui dati del mercato RC Auto italiano, con riferimento sia all’ODP che al DFCL. ■

In riferimento a un fissato “segment” dell’attività non-life, il modello M2 considera i pagamenti per rimborso sinistri X organizzati in un “run-off triangle (trapezoid)” indicizzato per *accident year* $i = 0, 1, \dots, I$ e *development year* $j = 0, 1, \dots, J$, con $I \geq J$. Quindi $X_{i,j}$ rappresenta i pagamenti aggregati “incrementali” per sinistri della generazione i effettuati nell’anno di sviluppo j . I corrispondenti valori dei pagamenti aggregati “cumulati” sono dati da:

$$C_{i,j} = \sum_{k=0}^j X_{i,k}.$$

Le ipotesi alla base del modello M2 sono le seguenti.

M2I - *Ipotesi di indipendenza*. I pagati cumulati $C_{i,j}$ di accident year diversi sono stocasticamente indipendenti.

M2CM - *Ipotesi di markovianità*. Per $i = 0, 1, \dots, I$, il processo $(C_{i,j})_{j \geq 0}$ è una *catena di Markov*:

$$\mathbf{P}(C_{i,j} \leq x | C_{i,0}, C_{i,1}, \dots, C_{i,j-1}) = \mathbf{P}(C_{i,j} \leq x | C_{i,j-1}).$$

M2M - *Ipotesi sulla media condizionata*. Esistono delle costanti $f_j > 0$ tali che per $1 \leq j \leq J$ e per $0 \leq i \leq I$:

$$\mathbf{E}(C_{i,j} | C_{i,j-1}) = f_{j-1} C_{i,j-1}.$$

M2V - *Ipotesi sulla varianza condizionata*. Esistono delle costanti $\sigma_j > 0$ tali che per $1 \leq j \leq J$ e per $0 \leq i \leq I$:

$$\mathbf{Var}(C_{i,j} | C_{i,j-1}) = \sigma_{j-1}^2 C_{i,j-1}.$$

Sotto queste ipotesi si ricava che gli stimatori chain ladder:

$$\hat{f}_j = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{S_j}, \quad \text{con} \quad S_j = \sum_{i=0}^{I-j-1} C_{i,j}, \quad (1.3)$$

sono stimatori non distorti di f_j , per $j = 0, 1, \dots, J-1$. Inoltre gli stimatori:

$$\hat{\sigma}_j^2 = \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j \right)^2, \quad (1.4)$$

sono stimatori non distorti di σ_j^2 , per $j = 0, 1, \dots, J - 2$. Se $I > J$ l'espressione dello stimatore vale anche per $j = J - 1$; altrimenti σ_{J-1}^2 è stimato, per estrapolazione, come:

$$\hat{\sigma}_{J-1}^2 = \min \left\{ \hat{\sigma}_{J-2}^2, \hat{\sigma}_{J-3}^2, \frac{\hat{\sigma}_{J-2}^4}{\hat{\sigma}_{J-3}^2} \right\}. \quad (1.5)$$

La stima di costo ultimo degli accident year aperti è ottenuta proiettando il pagato cumulato dell'“ultima diagonale” tramite le stime dei fattori chain ladder:

$$\hat{C}_{i,J} = C_{i,I-i} \prod_{j=I-i}^{J-1} \hat{f}_j, \quad i = I - J + 1, I - J + 2, \dots, I.$$

Utilizzando questi stimatori si giunge a un'espressione in forma chiusa per il MSEP stimato del risultato tecnico di fine esercizio relativo alla totalità degli accident year aperti. Questa espressione, ben nota, non è qui riportata per brevità.

Osservazione. Il MSEP contiene sia una componente di *process variance*, caratteristica dell'incertezza del processo di sviluppo dei costi, sia una componente di *estimation error*, legata all'incertezza connaturata alla stima dei parametri incogniti del modello. Nonostante l'ipotesi di indipendenza, questa seconda componente di incertezza include un effetto di covarianza che riduce la diversificazione tra accident year. Di questo effetto nella formula per il MSEP totale si tiene adeguatamente conto. ■

Sulla base del modello M2, la stima della deviazione standard unitaria per il reserve risk specifica dell'impresa (per il segmento considerato) è fornita dal rapporto:

$$\widehat{\mathbf{Cv}}_{res} = \frac{\sqrt{\widehat{\text{MSEP}}}}{\widehat{R}}, \quad (1.6)$$

dove $\widehat{R} = \sum_{i=I-J+1}^I (\widehat{C}_{i,J} - C_{i,I-i})$ è la stima delle *Outstanding Loss Liabilities* (stima di riserva, non scontata) fornita dal modello. Come per l'USP fornito dal Method 1, anche questa stima andrà miscelata col parametro marketwide previsto dalla Standard Formula, tramite il fattore di credibilità c prescritto da EIOPA.

2 La verifica delle ipotesi dei modelli teorici

2.a Verifica delle ipotesi del modello M1

Come richiesto negli Atti Delegati, per corroborare l'adattamento del modello di riferimento ai dati vanno verificate le tre ipotesi alla base del modello M1 introdotte nel paragrafo 1.a:

M1M – *Ipotesi sulla media*: relazione lineare tra $\mathbf{E}(Y)$ e X ,
M1V – *Ipotesi sulla varianza*: forma quadratica della varianza $\mathbf{Var}(Y)$ come
funzione di X ,
M1D – *Ipotesi sulla distribuzione*: lognormalità di Y ,
oltre alla verifica della proprietà:
ML – *Appropriatezza del metodo di massima verosimiglianza* utilizzato per
la stima.

2.a.1 Ipotesi sulla media

Per la verifica dell'ipotesi M1M di relazione lineare tra la media di Y e X è sufficiente effettuare una classica analisi di regressione lineare tra $\mathbf{E}(Y)$ e X , con o senza intercetta. Se si accetta che le osservazioni Y_t possano essere interpretate come stime non distorte di $\mathbf{E}(Y)$, l'analisi può essere condotta direttamente sulla serie storica delle osservazioni:

$$(\mathbf{X}, \mathbf{Y}) = \{(X_t, Y_t); t = 1, 2, \dots, T\},$$

specifiche della compagnia, secondo il modello (eventualmente senza intercetta):

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t, \quad t = 1, 2, \dots, T,$$

dove le ε_t sono termini di errore indipendenti a media nulla e varianza costante σ_ε^2 .

In aggiunta, si può effettuare una analisi di regressione lineare di tipo “marketwide”, utilizzando dati pubblicamente disponibili relativi a un campione di N compagnie considerate analoghe alla compagnia che effettua la stima. In questo caso i dati sono:

$$\{(\mathbf{X}, \mathbf{Y})_i; i = 1, 2, \dots, N\} = \{(X_{t,i}, Y_{t,i}); i = 1, 2, \dots, N, t = 1, 2, \dots, T_i\},$$

e si considera il modello (eventualmente senza intercetta):

$$\bar{Y}_i = \beta_0 + \beta_1 \bar{X}_i + \varepsilon_i, \quad i = 1, 2, \dots, N,$$

dove:

$$\bar{X}_i = \frac{1}{T_i} \sum_{t=1}^{T_i} X_{t,i}, \quad \bar{Y}_i = \frac{1}{T_i} \sum_{t=1}^{T_i} Y_{t,i},$$

sono le medie campionarie di X e Y della compagnia i (solito significato per i termini di errore).

Per la verifica dell'ipotesi M1M si tratterà di analizzare la significatività complessiva del modello, controllando il valore della statistica F (relativa all'ipotesi che tutti i parametri, tranne l'intercetta, siano nulli) e il relativo p -value, e considerare il livello della “varianza spiegata” dalla regressione calcolando il valore del coefficiente R^2 (coefficiente di determinazione). Con

riferimento alla stima dei singoli parametri, si tratta inoltre di verificare che il termine β_1 sia significativamente diverso da zero e, nel caso con intercetta, anche che il termine β_0 non sia significativamente diverso da zero. Come d'uso, si assume il fatto che il parametro sia uguale a zero come *ipotesi nulla*, e si adottano i classici test delle ipotesi disponibili in questi casi. L'approccio standard per la significatività dei parametri è quello di un test a due code basato sulla statistica t di Student, relativamente a un fissato livello di significatività α (per esempio $\alpha = 10\%$). Per l'eventuale rifiuto dell'ipotesi nulla si considererà il p -value, associato alla statistica test (in questo caso, la probabilità che la variabile aleatoria t sia, in valore assoluto, maggiore del valore osservato).

Osservazione. In applicazioni di “model selection” è buona regola effettuare confronti tra modelli utilizzando una varietà di indici di goodness-of-fit normalizzati (per il numero di osservazioni e per il numero di parametri); si considera, per esempio, l'SSE (Sum of Squared Errors) aggiustato col criterio “Squared Degree of Freedom” (SDF), l'SSE aggiustato con l'“Akaike Information Criterion” (AIC) e l'SSE aggiustato con il “Bayesian Information Criterion” (BIC). Dato che i test di validazione qui considerati non richiedono un confronto tra modelli alternativi, l'uso di questi indici non è necessario e ci si può limitare, eventualmente, al calcolo di una sola di queste misure di fitting (per esempio l'SSE-SDF). ■

2.a.2 Ipotesi sulla varianza

Per la verifica dell'ipotesi M1V, che la varianza $\mathbf{Var}(Y)$ sia una funzione quadratica di X , è conveniente ricorrere direttamente a una analisi di tipo “marketwide”, dato che di solito non è possibile ricavare in modo affidabile dai soli dati (\mathbf{X}, \mathbf{Y}) specifici dell'impresa osservazioni ripetute e indipendenti della varianza $\mathbf{Var}(Y)$.

Operativamente, si consideri un campione di osservazioni di mercato:

$$\{(\mathbf{X}, \mathbf{Y})_i; i = 1, 2, \dots, N\} = \{(X_{t,i}, Y_{t,i}); i = 1, 2, \dots, N, t = 1, 2, \dots, T_i\},$$

relative a un insieme di N compagnie considerate analoghe alla compagnia che effettua la stima. Per testare l'ipotesi sulla varianza, si può stimare su questi dati il modello:

$$\widehat{\mathbf{Var}}_i(Y) = \beta_0 + \beta_1 \bar{X}_i + \beta_2 \bar{X}_i^2 + \varepsilon_i, \quad i = 1, 2, \dots, N, \quad (2.7)$$

dove:

- $\widehat{\mathbf{Var}}_i(Y)$ è una stima della varianza $\mathbf{Var}_i(Y)$ della compagnia i ,
- \bar{X}_i è, ancora, la media campionaria $\sum_{t=1}^{T_i} X_{t,i}/T_i$.

Data la struttura dell'ipotesi M1V, i parametri nella (2.7) debbono in effetti avere la forma:

$$\beta_0 = 0, \quad \beta_{1,i} = \beta^2 \sigma_i^2 (1 - \delta_i) \bar{X}_i, \quad \beta_{2,i} = \beta_i^2 \sigma_i^2 \delta_i,$$

dove l'indice i sta a indicare la dipendenza dalla singola impresa. In effetti, nelle ipotesi del modello M1 sia δ_i che $\sigma_i = \beta \exp \gamma_i$ sono "entity specific". Inoltre, come osservato nel paragrafo 1.a, la scelta di EIOPA è stata quella di modificare la definizione di \bar{X} passando da una media di mercato (costante per tutte le imprese) a una media individuale di impresa. In questo modo un fattore che entra nel coefficiente $\beta_{1,i}$ viene a coincidere col regressore del modello; per cui, sostituendo le espressioni di $\beta_0, \beta_{1,i}$ e $\beta_{2,i}$ nella (2.7), si ottiene il modello²:

$$\widehat{\mathbf{Var}}_i(Y) = \beta^2 \sigma_i^2 \bar{X}_i^2 + \varepsilon_i, \quad i = 1, 2, \dots, N. \quad (2.8)$$

Se si indica con $\bar{\sigma}^2 = \sum_{i=1}^N \sigma_i^2 / N$ la media aritmetica delle σ_i^2 sull'intero campione di imprese, il modello (2.8) può essere approssimato con:

$$\widehat{\mathbf{Var}}_i(Y) \approx \beta_2 \bar{X}_i^2 + \tilde{\varepsilon}_i, \quad i = 1, 2, \dots, N, \quad (2.9)$$

con $\beta_2 := \beta^2 \bar{\sigma}^2$; l'approssimazione consiste nell'assumere che la variabilità di σ_i^2 tra compagnie (dispersione del parametro) può essere ben rappresentata considerandola incorporata nella varianza σ_ε^2 dei termini di errore $\tilde{\varepsilon}_i$. Si tratta quindi, in definitiva, di verificare l'ipotesi, tramite la regressione lineare (2.9), che la funzione varianza stimata abbia una forma quadratica "pura" (senza termine costante nè termine lineare).

Per la stima della varianza si può considerare un approccio corrispondente agli "Standardised Method 1" proposti nel QIS5 per le USP di premium e di reserve risk, adottando lo stimatore:

$$\widehat{\mathbf{Var}}_i(Y) = \bar{X}_i \frac{1}{T_i - 1} \sum_{t=1}^{T_i} X_{t,i} (Q_{t,i} - \bar{Q}_i)^2,$$

con:

$$Q_{t,i} := \frac{Y_{t,i}}{X_{t,i}} \quad \text{e} \quad \bar{Q}_i := \frac{\sum_{t=1}^{T_i} Y_{t,i}}{\sum_{t=1}^{T_i} X_{t,i}}.$$

Tuttavia la scelta più coerente, qualora possibile, sarebbe quella di usare per la stima di varianza il modello di cui si stanno effettivamente testando

²Questo risultato non cambierebbe sostanzialmente se si scegliesse come variabile indipendente una misura di volume diversa da \bar{X}_i . Se per esempio si scegliesse come regressore X_{T_i} (il più recente valore osservato di X), si avrebbe comunque forte correlazione positiva tra X_{T_i} e il fattore \bar{X}_i che entra nel coefficiente $\beta_{1,i}$, il che suggerirebbe comunque di ridefinire il modello adottando come regressore una misura di volume quadratica.

le ipotesi³. Ciò equivale a porre:

$$\widehat{\mathbf{Var}}_i(Y) = \hat{\sigma}_i^2(\hat{\delta}_i, \hat{\gamma}_i) \cdot \bar{Y}_i^2, \quad (2.10)$$

dove $\hat{\sigma}_i^2$ è la stima della deviazione standard unitaria specifica dell'impresa i fornita dal modello M1, che dovrà essere ottenuta ricavando i parametri $\hat{\delta}_i$ e $\hat{\gamma}_i$ per minimizzazione della corrispondente criterion function.

Qualunque sia lo stimatore della varianza adottato, per la stima – e la validazione – del modello (2.9) è naturale utilizzare le classiche tecniche di regressione lineare, come discusso per il test dell'ipotesi sulla media. Non si può escludere tuttavia che la dispersione del parametro presente nel modello teorico possa generare qualche difficoltà di identificazione. Potrà anche essere opportuno eliminare eventuali outlier nel campione delle varianze stimate.

2.a.3 Un esempio di verifica “marketwide” dell'ipotesi M1V

A puro titolo di esempio, è stato effettuato un esercizio di verifica dell'ipotesi M1V sulla varianza relativa al premium risk nel segmento RC Auto (MTL), basandosi sui dati del mercato italiano⁴. Le informazioni utilizzate sono i dati *public domain* reperibili sul sito ANIA www.infobila.it (“dati Infobila”). Sono state considerate le serie storiche dal 1999 al 2013 dei premi di competenza (variabile X) e della corrispondente stima di costo ultimo al primo anno di sviluppo (variabile Y), relativi a un campione selezionato di $N = 50$ imprese tra quelle operanti nel ramo⁵. Per ciascuna impresa i sono stati individuati i valori $\hat{\delta}_i$ e $\hat{\gamma}_i$ che minimizzano la funzione $\sigma_i(\delta_i, \gamma_i)$, e è stata ricavata la corrispondente stima di varianza (2.10) secondo il modello M1.

In un primo “run” il modello (2.9) è stato stimato sul campione di 50 imprese includendo un termine di intercetta. Nell'analisi dei risultati ha primaria importanza la significatività e la capacità di fitting del modello. Risulta che la statistica F è nettamente significativa, il che indica che il modello spiega una porzione significativa della variabilità dei dati; questa indicazione è confermata dall'elevato valore dell' R^2 e dell' R^2 aggiustato per i gradi di libertà.

Questi risultati dovrebbero essere già di per sé sufficienti per considerare confermata l'accettabilità dell'ipotesi M1V sui dati considerati. Effettuando

³Questa è la strada seguita dal JWG per analizzare l'adeguatezza ai dati di mercato dei vari modelli sperimentati per la calibratura (si veda [12], in particolare par. 91, nota 24).

⁴Va precisato che una verifica dell'ipotesi M1V estesa sull'intero mercato non è necessariamente più espressiva di una analoga analisi effettuata su un segmento del mercato. Se, per esempio, il mercato fosse separato in due segmenti descritti dallo stesso modello, ma con parametri diversi, la verifica dell'ipotesi darà migliore esito se effettuata sul solo segmento al quale appartiene l'impresa interessata.

⁵Sono state escluse le compagnie con meno di 5 osservazioni e quelle con $\hat{\sigma} \geq 1$.

Statistica F	p -value	R^2	R^2 agg.to
737.89	< 0.0001	0.9389	0.9377

anche l'analisi di significatività dei singoli parametri, risulta che, coerentemente con le ipotesi del modello, l'intercetta non è significativamente diversa da zero, mentre il coefficiente di \bar{X}^2 è diverso da zero a un elevato livello di significatività.

Param.	Val. stimato	Errore Std.	Statistica t	p -value
β_0	-78.11673	238.06326	-0.33	0.7442
β_2	0.00467	0.00017	27.16	< 0.0001

In una logica di “model selection”, è stato stimato, sempre sulle 50 compagnie, il modello ridotto senza intercetta, ottenendo i risultati sotto elencati.

Statistica F	p -value	R^2	R^2 agg.to
841.59	< 0.0001	0.9450	0.9439

Param.	Val. stimato	Errore Std.	Statistica t	p -value
β_2	0.00465	0.00016040	29.01	< 0.0001

Come si vede, la significatività generale e la capacità di fitting del modello migliorano ulteriormente (aumento della statistica F e degli R^2), e il coefficiente di “slope” β_2 è confermato diverso da zero a un elevato livello di significatività. Risulta quindi ampiamente e significativamente confermata l'adeguatezza dell'ipotesi M1V sul campione considerato.

2.a.4 Ipotesi sulla distribuzione

Si tratta di verificare l'ipotesi M1D, che i logaritmi $\{\ln Y_t; t = 1, 2, \dots, T\}$ delle osservazioni Y_t costituiscono un campione estratto da una distribuzione normale. Va subito osservato che, data la bassa numerosità del campione disponibile in queste applicazioni, i relativi test di normalità possono presentare criticità, potendo risultare scarsamente significativi o scarsamente “potenti” (nel senso della capacità di limitare errori di II tipo: normalità accettata, benchè assente). Le difficoltà a effettuare test sul tipo di distribuzione sono state segnalate anche dal JWG nella sua attività di calibrazione⁶.

Per affrontare il problema della verifica dell'ipotesi M1D possono essere adottati sia test di tipo “algoritmico”, sia test di tipo grafico.

⁶In [12] al par. 102 si dice: *The empirical findings on this issue [cioè: discriminating between the normal and lognormal distribution] – for example, with regard to the various goodness-of-fit diagnostics and PP-plots – were also inconclusive.*

Test di tipo algoritmico

I test di tipo algoritmico assumono la normalità dei dati come ipotesi nulla (H0), e definiscono una apposita “statistica test” per discriminare rispetto all’ipotesi alternativa - non-normalità (H1). In questo contesto, un basso valore del p -value⁷ è per definizione associato a un basso livello di fiducia che i dati siano effettivamente normali. Secondo una pratica diffusa, valori del p -value più bassi dell’1% forniscono un forte sostegno a H1, cioè alla non-normalità, valori superiori al 10% indicano che i dati non forniscono sostegno a H1, mentre valori compresi tra l’1 e il 10% indicano una situazione incerta. È quindi opportuno precisare come osservazione preliminare che, indipendentemente dall’ampiezza del campione usato, i test di normalità basati sul p -value, pur potendo fornire indicazioni decisive nel senso del rifiuto di H0, possono risultare non decisivi nel senso dell’accettazione di H0 (limitandosi semplicemente a non fornire evidenze contrarie). Il problema è ben illustrato in [16]. Le citate difficoltà incontrate dal JWG possono essere messe in relazione anche con questa situazione.

Tra gli innumerevoli test di normalità di tipo algoritmico attestati in letteratura, i seguenti sono tra i più frequentemente utilizzati.

- **Test di Kolmogorov-Smirnov.** È un test non parametrico, basato sulla funzione di ripartizione empirica (*Empirical Distribution Function*, EDF). Dato un campione $\{X_1, X_2, \dots, X_n\}$ di n osservazioni indipendenti e identicamente distribuite (i.i.d.) della variabile aleatoria X , l’EDF di X è definita come:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{I}_{\{X_i \leq x\}}, \quad x \in \mathbb{R}.$$

Data una funzione di ripartizione teorica $F(x)$ continua ipotizzata come vera (in questo caso la normale), la goodness-of-fit del campione rispetto a $F(x)$ è definita sulla base di una *misura di distanza* tra la funzione di ripartizione empirica $F_n(x)$ e la funzione di ripartizione teorica.

Nel test di Kolmogorov-Smirnov (KS) [18], la misura di distanza è definita come l’estremo superiore D della differenza, in valore assoluto, tra $F_n(x)$ e $F(x)$:

$$D = \sup_{x \in \mathbb{R}} |F_n(x) - F(x)|.$$

Naturalmente, valori tanto più bassi di D sono tanto più favorevoli all’ipotesi nulla. Numericamente, il test consiste nel confrontare $\sqrt{n}D$ con il corrispondente valore critico K_α di Kolmogorov, dove K_α è tale che

⁷Intuitivamente, il p -value esprime la probabilità che i dati possano presentarsi nella forma effettivamente osservata (e, quindi, che la statistica test assuma il valore effettivamente calcolato, o valori più estremi), nel caso sia vera l’ipotesi nulla H0. Quindi valori bassi del p -value spingono a escludere l’ipotesi H0, ma valori alti non escludono che possa essere vera anche l’ipotesi alternativa H1.

$\mathbf{P}(K \leq K_\alpha) = 1 - \alpha$, essendo K la variabile aleatoria di Kolmogorov e α il livello di significatività desiderato. Come per tutti i test basati su misure di distanza, il p -value è la probabilità che D sia *maggiore* del valore osservato.

In pratica, la statistica KS richiede un numero relativamente grande di osservazioni per rifiutare appropriatamente l'ipotesi nulla.

- **Test di Cramer-VonMises.** Il test di Cramer-VonMises (CvM) è anch'esso basato sulla EDF, ma appartiene alla classe dei test "QEDF"; questi test usano una misura di distanza di tipo quadratico, definita come:

$$D^2 = n\omega^2, \quad \text{con} \quad \omega^2 = \int_{-\infty}^{\infty} (F_n(x) - F(x))^2 w(x) dF(x),$$

dove $w(x)$ è una funzione peso fissata. Rispetto al test KS, i test di tipo QEDF tengono maggiormente in considerazione l'intero insieme dei dati, nel senso della somma degli scostamenti; il test KS è invece più sensibile a scostamenti "aberranti".

Il test CvM [2] [25] usa D^2 con $w(x) \equiv 1$:

$$T^2 = n \int_{-\infty}^{\infty} (F_n(x) - F(x))^2 dF(x),$$

e consiste nel confrontare T^2 con il corrispondente valore tabulato, in corrispondenza del desiderato livello di significatività α . Nei test di normalità, CvM dovrebbe avere elevata "potenza", essendo ritenuto uno dei test di tipo EDF più efficaci nel rivelare gli scostamenti dall'ipotesi nulla (basso tasso di errori di II tipo). Se ne consiglia di solito l'uso anche per campioni con $n < 25$ (mentre può fallire su campioni molto grandi).

- **Test di Anderson-Darling.** Il test di Anderson-Darling (AD) [1] è anch'esso della classe QEDF, con la funzione peso:

$$w(x) = 1/[F(x)(1 - F(x))];$$

si ha così la statistica test:

$$A^2 = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{F(x)(1 - F(x))} dF(x).$$

Le proprietà sono analoghe a CvM, con la differenza che la statistica A^2 pone più peso sulle code.

- **Test di Shapiro-Wilk.** Il test di normalità di Shapiro-Wilk (SW) [21] è costruito sul confronto tra uno stimatore della varianza basato sulla combinazione lineare ottimale della statistica d'ordine di una variabile normale, e la tradizionale varianza campionaria. La statistica test W è

il rapporto tra questi due stimatori, e può assumere valori compresi tra 0 e 1. L'ipotesi di normalità sarà rifiutata per valori *bassi* di W e non-rifiutata per valori prossimi a 1. Quindi il p -value è la probabilità che W sia *minore* del valore osservato. Va però precisato che la distribuzione di W è fortemente asimmetrica, al punto che valori di W vicini a 0.9 possono essere considerati bassi ai fini dell'analisi di normalità.

A fini interpretativi, è interessante il fatto che la statistica W può essere interpretata come il quadrato del coefficiente di correlazione in un QQ-plot.

Il test SW è spesso presentato come uno dei test più potenti per la verifica della normalità in piccoli campioni. Può essere poco affidabile in presenza di molti valori ripetuti nei dati.

- **Test di Jarque-Bera.** Il test appartiene alla classe “omnibus moments”, in quanto punta a stabilire contemporaneamente se due momenti, la skewness e la curtosi, stimati sui dati sono coerenti con l'ipotesi di normalità. La statistica test di Jarque-Bera (JB) [15] ha la forma:

$$JB = \frac{T}{24} (4b + (k - 3)^2) ,$$

dove \sqrt{b} e k sono, rispettivamente, la skewness e la curtosi campionarie. Se i dati sono normali, la statistica JB ha distribuzione asintotica chi-quadro non centrata con due gradi di libertà.

Nel test JB, H_0 è un'ipotesi congiunta che sia asimmetria che curtosi abbiano valore nullo. Tale ipotesi viene rigettata per valori di JB troppo grandi. Quindi il p -value è la probabilità che JB sia *maggiore* del valore osservato.

Il test JB è stato utilizzato dal JWG nella procedura di individuazione degli outlier nei valori stimati di deviazione standard su campioni di imprese relativamente ampi. Il test non è però appropriato per piccoli campioni. Infatti in questi casi l'approssimazione chi-quadro è eccessivamente instabile⁸; inoltre la distribuzione del p -value tende a diventare fortemente sbilanciata su valori bassi (right-skewed unimodal). Questi comportamenti tendono a generare alti tassi di errori di I tipo (rifiuto inappropriato dell'ipotesi nulla).

Per tutti questi motivi non appare consigliabile l'applicazione del test JB al problema in questione.

Test di tipo grafico

⁸Il problema è segnalato anche dal JWG. In [12] al paragrafo 9.3 si dice: *Care should be exercised with this test statistic as the asymptotic distribution only holds for fairly large ($n \gg 100$) numbers of observations n .*

- **Istogramma.** È il classico grafico a barre verticali che illustra la frequenza relativa delle osservazioni che ricadono nel k -esimo intervallo di una “griglia” opportunamente scelta sull’asse delle ascisse. Data l’esiguità dei campioni disponibili nei calcoli di USP, questo approccio è quasi sempre di scarsa utilità ai fini di decidere sulla normalità dei dati.
- **PP-plot.** Dato un campione $\{X_1, X_2, \dots, X_n\}$ di n osservazioni indipendenti e identicamente distribuite della variabile aleatoria X , si costruisca il campione ordinato $\{X_{n,n} \leq X_{n-1,n} \leq \dots \leq X_{1,n}\}$. Dato che $X_{k,n} \leq x$ se e solo se $\sum_{i=1}^n \mathbf{I}_{\{X_i > x\}} < k$, sui valori osservati del campione ordinato l’EDF assume i valori:

$$F_n(X_{k,n}) = \frac{n - k + 1}{n}, \quad k = 1, 2, \dots, n.$$

Il *probability plot* (PP-plot) è il grafico bidimensionale:

$$\left\{ \left(F(X_{k,n}), \frac{n - k + 1}{n + 1} \right), k = 1, 2, \dots, n \right\},$$

costruito sul campione ordinato $\{X_{n,n} \leq X_{n-1,n} \leq \dots \leq X_{1,n}\}$ delle n osservazioni (i.i.d.) di X . Per il teorema di Glivenko-Cantelli, se X ha funzione di ripartizione F il grafico dovrebbe avere andamento approssimativamente lineare.

- **QQ-plot.** Il *quantile plot* (QQ-plot) è l’analogo grafico riferito ai quantili:

$$\left\{ \left(X_{k,n}, F^{(-1)} \left(\frac{n - k + b_k}{n + a_k} \right) \right), k = 1, 2, \dots, n \right\},$$

con a_k e b_k scelti in modo da tener conto della discontinuità “empirica” della distribuzione (si veda a es. [13]). Scelte tipiche sono $a_k = b_k \equiv 1$ o $a_k \equiv 0$ e $b_k \equiv 0.5$. Anche in questo caso nell’ipotesi $X \sim F$ il grafico dovrebbe essere approssimativamente lineare.

2.a.5 Un confronto per simulazione delle metodologie di test

È stata effettuata una analisi comparativa per simulazione allo scopo di confrontare la capacità discriminante dei test di normalità considerati, con particolare attenzione ai piccoli campioni.

Logica dell’esercizio di simulazione

I test KS, CvM, AD, SW e JB sono stati applicati a 1000 campioni di T osservazioni (con $T = 6, 10, 15, 100$), estratti per simulazione da una distribuzione normale, lognormale, di Weibull con parametro di forma $\tau > 1$ e di Pareto di tipo II. Per tutte le distribuzioni si è posta una media $m = 100$ e per la normale si è fissato un coefficiente di variazione $\kappa = 0.1$ (valore tipico

delle deviazioni standard unitarie prescritte nella Standard Formula). Ne consegue che il quantile al livello $p = 99.5\%$ è $Q_p = 125.758$. I parametri delle altre tre distribuzioni sono stati scelti in modo da avere stesso valore del quantile Q_p (quindi stesso valore dell'“unexpected loss”)⁹. Ne consegue che la Weibull ha parametro di forma $\tau = 9.4315$ e parametro di scala $\theta = 105.3799$, che implicano una deviazione standard $\sigma = 12.71$; si ha quindi una dispersione maggiore di quella della normale (che ha $\sigma = 10$), coerentemente col fatto che per $\tau > 1$ la Weibull è più “light-tailed”. La dispersione della lognormale e della Pareto, invece, è minore di quella della normale, dato che ambedue queste distribuzioni, e in particolar modo la Pareto, sono più “heavy-tailed”. In particolare, per la lognormale si hanno parametri di media e di deviazione standard (di $\ln Y$) $\mu = 4.60$ e $\omega = 0.09058$, il che implica una deviazione standard $\sigma = 9.07$; per la Pareto, che ha curtosi molto più elevata, si ha parametro di forma $\alpha = 1.0065$ e parametro di scala $\theta = 0.6542$, valori per i quali la varianza non è definita.

In definitiva, quindi, oltre al caso di dati generati da distribuzione normale (che corrisponde all'ipotesi nulla), si sono considerate le tre ipotesi alternative corrispondenti a una distribuzione con curtosi minore (Weibull) e a due distribuzioni (lognormale e Pareto) con curtosi maggiore, una delle quali (la Pareto) a “comportamento estremo”. La coincidenza delle unexpected loss delle quattro distribuzioni è imposta per rendere equivalenti le alternative dal punto di vista dell'SCR implicato secondo Solvency II.

Risultati delle simulazioni

Test di tipo algoritmico. I 1000 valori di ciascuna statistica test e del corrispondente p -value, ogni volta calcolati sui campioni con numerosità $T = 6, 10, 15, 100$, sono stati salvati e confrontati tra loro. Una analisi esaustiva dei risultati si ha confrontando le distribuzioni empiriche così ottenute. Viene qui riportato l'esito di una analisi ridotta, basata solo sull'osservazio-

⁹La funzione di ripartizione di Weibull ha la forma:

$$F(x) = 1 - e^{-(x/\theta)^\tau}, \quad x > 0,$$

con $\theta, \tau > 0$. La media e il quantile a livello p sono:

$$\mu = \theta \Gamma(1 + 1/\tau), \quad Q_p = \theta [-\ln(1 - p)]^{1/\tau}.$$

Per la Pareto di tipo II (nota anche come distribuzione di Lomax), la funzione di ripartizione è:

$$F(x) = 1 - \left(\frac{\theta}{\theta + x} \right)^\alpha, \quad x > 0,$$

con $\alpha, \theta > 0$. La media e il quantile a livello p sono:

$$\mu = \frac{\theta}{\alpha - 1}, \quad Q_p = \theta [(1 - p)^{-1/\alpha} - 1].$$

Per le caratteristiche della distribuzione di Weibull e di Pareto tipo II si può vedere [16].

ne dei valori medi, dei valori modali, dei valori mediani e del “numero di rifiuti” dell’ipotesi nulla.

Nelle tavole 1a e 1b sono riportati, per tutte le dimensioni campionarie considerate, i risultati delle simulazioni relativi ai cinque test di normalità considerati; la tavola 1a si riferisce ai test di Kolmogorov-Smirnov, di Cramer-von Mises e di Anderson-Darling, basati su misure di distanza; la tavola 1b riguarda il test di Shapiro-Wilk e quello di Jarque-Bera. Nelle tavole sono riportati i seguenti dati: il valor medio della statistica test, il valor medio del p -value, il valore modale¹⁰ del p -value, il valor mediano del p -value e il tasso di rifiuto di H_0 al livello α , cioè il numero percentuale r_α di casi, sui 1000 simulati, in cui il p -value è risultato minore del livello di significatività α ; sono stati considerati i valori $\alpha = 1\%, 5\%, 10\%$. Nel caso di dati normali, il tasso di rifiuto r_α individua il tasso di errori di I tipo (rifiuto di H_0 quando è vera); ovviamente si richiede che il valore di r_α sia il più basso possibile. Nei tre casi di dati non normali, invece, r_α dovrebbe essere il più elevato possibile, in quanto misura la “potenza” del test (capacità di escludere l’ipotesi H_0 se è falsa); si richiede di solito che il valore di r_α sui dati non-normali sia non inferiore all’80%. Il complemento a 1 di r_α per i dati non-normali individua il tasso di errori di II tipo (mancato rifiuto di H_0 quando è falsa).

I valori delle tavole evidenziano che il tasso di errori di I tipo è adeguatamente basso per tutti e cinque i test e per tutti i valori della numerosità campionaria T . Nel dettaglio, il valore di r_α per i dati normali è più elevato per i test SW e JB che per i tre test basati su misure di distanza, tra i quali KS tende a sua volta a produrre i valori più bassi.

Se si osservano però i tassi di rifiuto di H_0 per i dati non-normali, per tutti i test risultano valori di r_α inadeguatamente bassi sia sui dati lognormali che sui dati Weibull, per tutti e tre i valori di α . Valori abbastanza alti del tasso di rifiuto si hanno solo per i dati con distribuzione di Pareto, e anche in questo caso si osservano livelli di r_α superiori all’80% solo per i valori più alti di T e di α .

Tutte le statistiche test hanno l’andamento teoricamente previsto: all’allontanarsi dall’ipotesi di normalità, andamento decrescente per il test SW (coerentemente con l’interpretazione della statistica W come quadrato del coefficiente di correlazione del QQ-plot) e andamento crescente per gli altri test. Tuttavia, anche per il campione con $T = 100$ tutti i test falliscono quasi sistematicamente nel “riconoscere” la non-normalità per distribuzioni che non si discostino pesantemente (come la Pareto) da una forma “bell-shaped”. Nel complesso, considerando anche media, moda e mediana dei p -value, si può forse affermare che i test SW e AD dimostrano una potenza

¹⁰Il valore modale è stato ricavato definendo un arrotondamento del valore simulato del p -value alla terza cifra decimale. Nel caso di moda multipla, è stato preso il valore minimo della moda.

distr	T	Kolmogorov-Smirnov										Cramer-von Mises										Anderson-Darling									
		stat			p-value			tasso di rifiuto				stat			p-value			tasso di rifiuto				stat			p-value			tasso di rifiuto			
		test	media	moda	mediana	$\alpha=1\%$	$\alpha=5\%$	$\alpha=10\%$	test	media	moda	mediana	$\alpha=1\%$	$\alpha=5\%$	$\alpha=10\%$	test	media	moda	mediana	$\alpha=1\%$	$\alpha=5\%$	$\alpha=10\%$	test	media	moda	mediana	$\alpha=1\%$	$\alpha=5\%$	$\alpha=10\%$		
Nor	6	0.226	0.139	0.150	0.150	0.00	4.9%	9.1%	0.056	0.222	0.250	0.250	0.6%	4.1%	8.7%	0.338	0.219	0.250	0.250	0.6%	4.4%	9.6%	0.338	0.219	0.250	0.250	0.6%	4.4%	9.6%		
Nor	10	0.183	0.139	0.150	0.150	0.00	5.1%	10.2%	0.058	0.221	0.250	0.250	0.9%	4.7%	8.8%	0.359	0.220	0.250	0.250	0.7%	4.9%	8.7%	0.359	0.220	0.250	0.250	0.7%	4.9%	8.7%		
Nor	15	0.149	0.141	0.150	0.150	0.00	3.3%	8.4%	0.056	0.223	0.250	0.250	0.5%	3.5%	8.9%	0.354	0.222	0.250	0.250	0.7%	3.8%	9.4%	0.354	0.222	0.250	0.250	0.7%	3.8%	9.4%		
Nor	100	0.062	0.140	0.150	0.150	0.00	4.7%	8.9%	0.058	0.223	0.250	0.250	1.1%	3.7%	7.9%	0.376	0.222	0.250	0.250	1.0%	4.0%	8.4%	0.376	0.222	0.250	0.250	1.0%	4.0%	8.4%		
Log	6	0.226	0.139	0.150	0.150	0.00	5.2%	9.5%	0.056	0.221	0.250	0.250	0.9%	4.4%	9.1%	0.341	0.218	0.250	0.250	1.0%	4.7%	10.0%	0.341	0.218	0.250	0.250	1.0%	4.7%	10.0%		
Log	10	0.184	0.138	0.150	0.150	0.00	5.2%	11.3%	0.059	0.218	0.250	0.250	1.3%	5.3%	9.7%	0.365	0.217	0.250	0.250	1.2%	5.2%	9.9%	0.365	0.217	0.250	0.250	1.2%	5.2%	9.9%		
Log	15	0.151	0.141	0.150	0.150	0.00	4.0%	8.5%	0.058	0.220	0.250	0.250	0.8%	4.5%	9.7%	0.365	0.219	0.250	0.250	0.9%	5.0%	10.5%	0.365	0.219	0.250	0.250	0.9%	5.0%	10.5%		
Log	100	0.066	0.132	0.150	0.150	0.00	8.7%	17.1%	0.069	0.203	0.250	0.250	2.4%	10.3%	16.0%	0.446	0.199	0.250	0.250	3.0%	11.4%	18.4%	0.446	0.199	0.250	0.250	3.0%	11.4%	18.4%		
Wei	6	0.227	0.137	0.150	0.150	0.00	5.9%	11.1%	0.057	0.216	0.250	0.250	1.2%	6.1%	11.6%	0.346	0.212	0.250	0.250	1.4%	6.3%	13.0%	0.346	0.212	0.250	0.250	1.4%	6.3%	13.0%		
Wei	10	0.190	0.135	0.150	0.150	0.00	8.0%	13.4%	0.064	0.209	0.250	0.250	1.6%	7.5%	13.5%	0.394	0.207	0.250	0.250	2.0%	7.9%	14.6%	0.394	0.207	0.250	0.250	2.0%	7.9%	14.6%		
Wei	15	0.162	0.132	0.150	0.150	0.00	8.9%	15.8%	0.069	0.201	0.250	0.250	2.3%	10.8%	16.4%	0.426	0.199	0.250	0.250	2.9%	10.7%	17.7%	0.426	0.199	0.250	0.250	2.9%	10.7%	17.7%		
Wei	100	0.083	0.089	0.150	0.097	0.00	36.7%	51.0%	0.135	0.107	0.250	0.069	23.9%	44.2%	57.5%	0.854	0.093	0.005	0.048	28.8%	51.3%	64.2%	0.854	0.093	0.005	0.048	28.8%	51.3%	64.2%		
Par	6	0.323	0.074	0.150	0.052	0.00	49.6%	60.3%	0.143	0.092	0.005	0.038	36.1%	54.6%	64.2%	0.773	0.086	0.005	0.031	37.4%	56.4%	66.0%	0.773	0.086	0.005	0.031	37.4%	56.4%	66.0%		
Par	10	0.332	0.040	0.100	0.010	0.00	76.0%	82.8%	0.290	0.035	0.005	0.005	70.4%	83.1%	88.5%	1.537	0.031	0.005	0.005	72.1%	85.1%	90.5%	1.537	0.031	0.005	0.005	72.1%	85.1%	90.5%		
Par	15	0.337	0.021	0.010	0.010	0.00	90.9%	94.0%	0.496	0.013	0.005	0.005	89.6%	95.7%	97.0%	2.586	0.012	0.005	0.005	90.6%	96.1%	97.5%	2.586	0.012	0.005	0.005	90.6%	96.1%	97.5%		
Par	100	0.380	0.010	0.010	0.010	0.00	100.0%	100.0%	5.046	0.005	0.005	0.005	100.0%	100.0%	100.0%	24.574	0.005	0.005	0.005	100.0%	100.0%	100.0%	24.574	0.005	0.005	0.005	100.0%	100.0%	100.0%		

Tavola 1a. Risultati dei test Kolmogorov-Smirnov, Cramer-von Mises e Anderson-Darling sui dati simulati

		Shapiro-Wilk							Jarque-Bera						
distr	T	stat test	p-value			tasso di rifiuto			stat test	p-value			tasso di rifiuto		
			media	moda	mediana	$\alpha=1\%$	$\alpha=5\%$	$\alpha=10\%$		media	moda	mediana	$\alpha=1\%$	$\alpha=5\%$	$\alpha=10\%$
Nor	6	0.906	0.491	0.447	0.489	1.0%	4.4%	9.1%	1.460	0.597	0.658	0.645	1.2%	3.5%	5.7%
Nor	10	0.926	0.494	0.651	0.495	1.0%	4.0%	9.2%	1.488	0.616	0.000	0.661	1.5%	3.9%	6.2%
Nor	15	0.944	0.527	0.170	0.536	1.0%	3.3%	7.9%	1.527	0.622	0.000	0.663	1.7%	4.1%	5.6%
Nor	100	0.987	0.504	0.478	0.497	1.4%	4.0%	8.5%	1.902	0.543	0.000	0.570	2.0%	4.0%	6.3%
Log	6	0.905	0.486	0.541	0.482	1.2%	4.5%	9.5%	1.481	0.594	0.647	0.631	1.3%	4.2%	6.3%
Log	10	0.924	0.484	0.062	0.475	1.1%	5.4%	10.5%	1.583	0.605	0.000	0.650	2.1%	3.9%	7.2%
Log	15	0.942	0.510	0.369	0.508	1.1%	4.6%	10.1%	1.761	0.605	0.000	0.655	3.0%	5.5%	7.9%
Log	100	0.984	0.402	0.008	0.351	4.7%	14.4%	21.9%	3.324	0.435	0.000	0.421	7.4%	13.4%	17.9%
Wei	6	0.903	0.490	0.135	0.488	1.9%	6.3%	12.4%	1.671	0.584	0.019	0.636	2.1%	6.2%	9.4%
Wei	10	0.917	0.443	0.277	0.419	2.1%	7.8%	14.9%	2.006	0.572	0.000	0.622	4.5%	7.3%	10.0%
Wei	15	0.930	0.434	0.004	0.397	3.5%	11.9%	19.4%	2.931	0.533	0.000	0.580	7.5%	11.6%	13.9%
Wei	100	0.967	0.110	0.000	0.021	39.8%	62.7%	73.3%	11.223	0.131	0.000	0.044	35.8%	51.8%	62.4%
Par	6	0.748	0.128	0.000	0.024	40.1%	57.3%	67.0%	6.595	0.243	0.001	0.088	36.8%	46.6%	51.1%
Par	10	0.658	0.031	0.000	0.000	74.2%	87.6%	91.7%	24.361	0.104	0.000	0.000	60.9%	68.8%	73.9%
Par	15	0.589	0.007	0.000	0.000	91.5%	96.6%	98.2%	72.505	0.032	0.000	0.000	83.0%	87.9%	90.6%
Par	100	0.304	0.000	0.000	0.000	100.0%	100.0%	100.0%	15733.114	0.000	0.000	0.000	100.0%	100.0%	100.0%

Tavola 1b. Risultati dei test Shapiro-Wilk e Jarque-Bera sui dati simulati

leggermente superiore; tuttavia la probabilità di commettere errori di II tipo è elevata per tutti i metodi considerati. Tra i cinque test considerati il test JB evidenzia le performance peggiori, verosimilmente a causa della scarsa numerosità dei campioni considerati. Si conferma quindi l'indicazione di non utilizzabilità dell'approccio per questo tipo di applicazioni.

Test di tipo grafico. Per ottenere anche un confronto di efficacia per test di normalità di tipo grafico, per ciascuno dei livelli della numerosità campionaria T considerati è stato scelto, tra i 1000 campioni simulati, quello che ha fornito, con una fissata statistica test, un p -value più vicino al valore modale. La statistica utilizzata è quella del test di Shapiro-Wilk. Per ciascuno di questi campioni è stato effettuato un PP-plot e un QQ-plot. I risultati sono riportati nelle tavole 2 e 3. Anche da questi test grafici risulta evidente la difficoltà a individuare correttamente i dati normali per i campioni di bassa numerosità, con l'eccezione del caso estremo di dati paretiani. Su questi dati risulta arduo distinguere tramite PP-plot tra l'ipotesi normale e quella lognormale anche sul campione con 100 osservazioni.

2.a.6 Appropriately del metodo di massima verosimiglianza

Per quanto riguarda la proprietà ML, di appropriatezza del metodo di massima verosimiglianza utilizzato per la stima (paragrafo 2.a), questa risulta attestata dalle proprietà di convergenza della procedura di minimizzazione, alla quale si richiede l'individuazione univoca di un minimo della criterion function nell'intervallo D di ottimizzazione. L'univocità del minimo prodotto dalla procedura può essere testata studiando empiricamente la criterion function $\ell(\delta, \gamma)$ su una griglia abbastanza ampia di valori δ e γ .

Nella costruzione della griglia è necessario scegliere i valori γ_{min} e γ_{max} . Ricordando la definizione $\gamma = \ln(\sigma/\beta)$, e dato che $\beta = \mathbf{E}(Y/X)$ (loss ratio atteso, run-off ratio atteso), si può assumere $\beta \approx 1$ e $\sigma \in [0.005, 1]$, da cui:

$$\gamma_{min} = \ln(0.005) = -5.30, \quad \gamma_{max} = \ln(1) = 0.$$

Si restringe quindi il dominio D al dominio:

$$D^* = \{0 \leq \delta \leq 1, \gamma_{min} \leq \gamma \leq \gamma_{max}\},$$

considerando privi di interesse i valori della criterion function situati al di fuori. Un grafico tridimensionale su D^* dovrebbe evidenziare la regolarità della funzione e l'esistenza di un minimo globale (eventualmente sulla frontiera di D^*), ben individuabile dalla procedura di minimizzazione. Nella figura 1 è illustrata, a titolo di esempio, la forma di una tipica "superficie di volatilità" $\sigma(\delta, \gamma)$.

2.b Verifica delle ipotesi del modello M2

Anche per il modello M2 è richiesto negli Atti Delegati che sia verificato l'adattamento dei dati alle ipotesi sottostanti. Precisamente è richiesto che

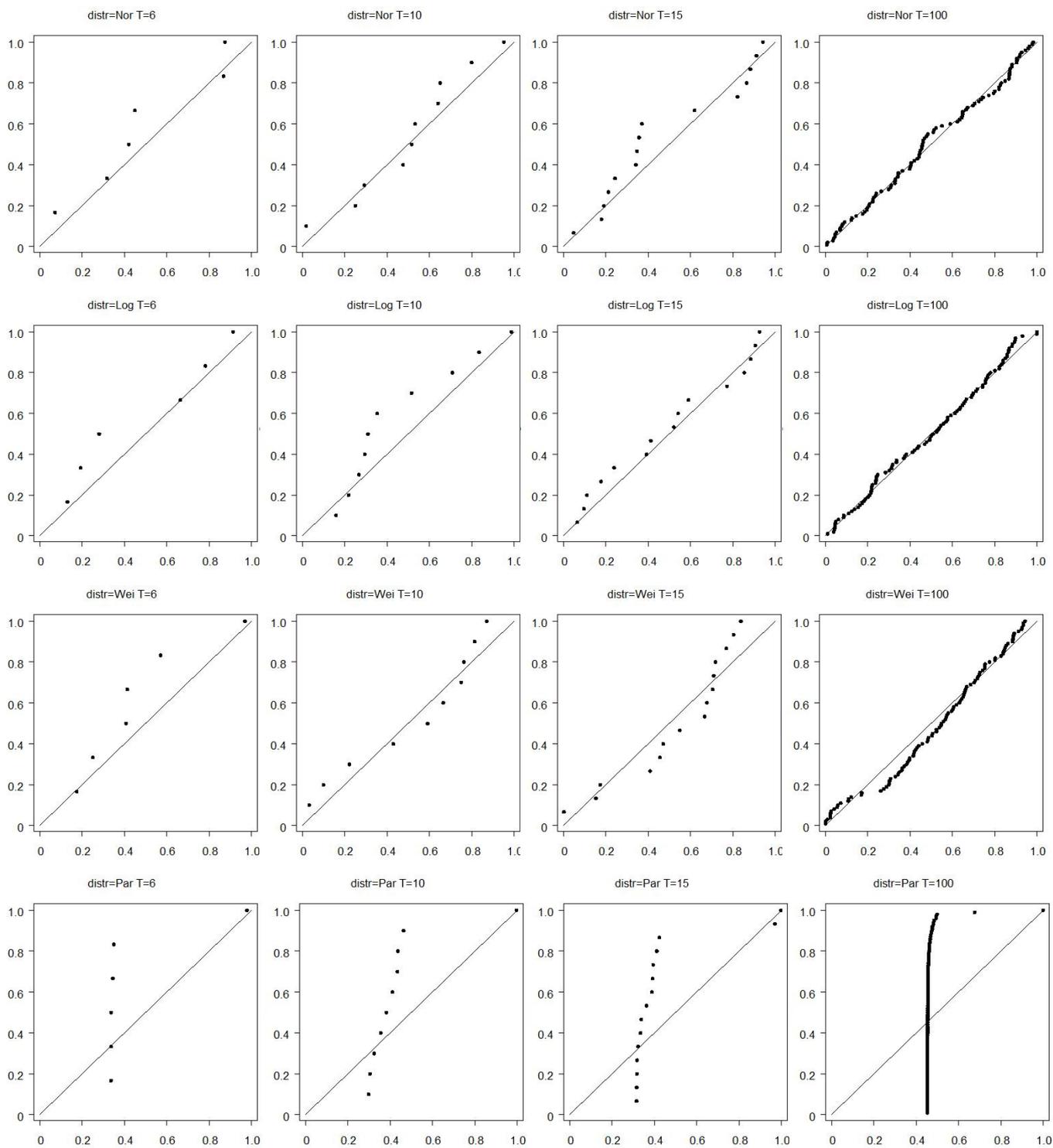


Tavola 2. PP-plot sui campioni con p -value modale (secondo Shapiro-Wilk)

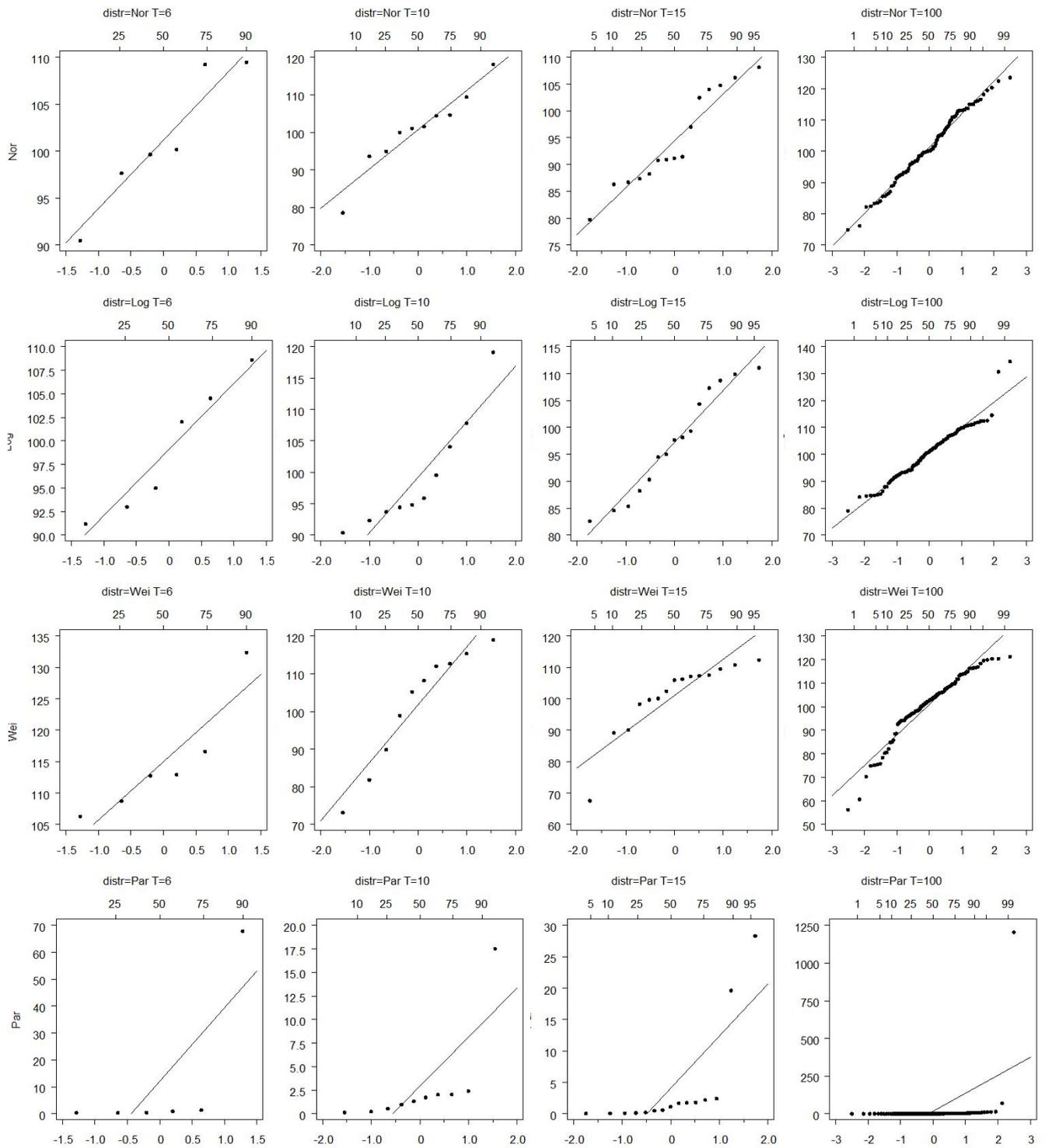
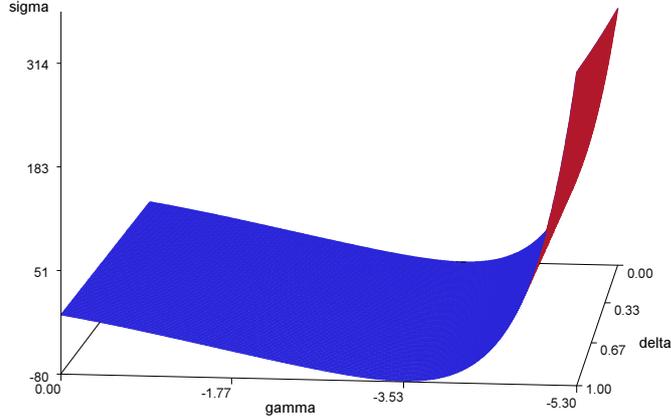


Tavola 3. QQ-plot sui campioni con p -value modale (secondo Shapiro-Wilk)

Figura 1. La superficie $\sigma(\delta, \gamma)$ nel dominio D^*



siano verificate le ipotesi alla base del modello M2 introdotte nel paragrafo 1.b:

M2I – *Ipotesi di indipendenza*: Indipendenza tra i pagati cumulati (e incrementali) di accident year (AY) diversi,

M2M – *Ipotesi sulla media condizionata*: per ogni AY e in ogni anno di sviluppo (DY) di un qualsiasi AY fissato, proporzionalità del valore atteso del pagato cumulato del prossimo DY rispetto al pagato cumulato del DY corrente,

M2V – *Ipotesi sulla varianza condizionata*: per ogni AY e in ogni DY di un qualsiasi AY fissato, proporzionalità della varianza del pagato cumulato del prossimo DY rispetto al pagato cumulato del DY corrente.

È utile riformulare le ipotesi M2M e M2V in forma unificata. Si indichi con:

$$\mathcal{B}_0 := \{C_{0,0}, C_{0,1}, \dots, C_{0,I}, \},$$

l'insieme di tutti i pagati nel primo anno di sviluppo. Allora le ipotesi M2M e M2V si possono unificare nella:

M2MV - *Ipotesi time series*. Esistono delle costanti $f_j > 0$ e $\sigma_j > 0$ e delle variabili aleatorie $\varepsilon_{i,j}$ tali che per $1 \leq j \leq J$ e per $0 \leq i \leq I$:

$$C_{i,j} = f_{j-1} C_{i,j-1} + \sigma_{j-1} \sqrt{C_{i,j-1}} \varepsilon_{i,j}, \quad (2.11)$$

dove le $\varepsilon_{i,j}$ sono termini di errore identicamente distribuiti e condizionatamente indipendenti, dato \mathcal{B}_0 , con media $\mathbf{E}(\varepsilon_{i,j}|\mathcal{B}_0) = 0$ e varianza $\mathbf{Var}(\varepsilon_{i,j}|\mathcal{B}_0) = 1$.

Questa formalizzazione è stata proposta nel 2006 da Buchwalder, Bühlmann, Merz e Wüthrich [3] come estensione “distribuzionale” del modello DFCL di Mack, e caratterizza il cosiddetto Time Series Chain Ladder (TSCL). La (2.11) consente, tra l’altro, di trattare il modello per simulazione¹¹.

2.b.1 Ipotesi sulla media e la varianza condizionate

Per ogni fissato $j = 0, 1, \dots, J - 1$, la (2.11) definisce un modello di regressione lineare relativo alle osservazioni di una coppia di development year consecutivi. Si tratta di J regressioni lineari *pesate* del tipo:

$$y_i = \beta x_i + \frac{\sigma}{\sqrt{w_i}} \varepsilon_i, \quad i = 0, 1, \dots, I,$$

dove $x_i = C_{i,j-1}$, $y_i = C_{i,j}$ e $w_i = 1/x_i = 1/C_{i,j-1}$. Come è noto, il coefficiente β di questa regressione sarà stimato col metodo dei minimi quadrati pesati come:

$$\hat{\beta} = \frac{\sum_{i=1}^n w_i x_i y_i}{\sum_{i=1}^n w_i x_i^2};$$

come immediatamente si verifica, questa espressione coincide con la (1.3), che fornisce lo stimatore chain ladder f_j . Inoltre la varianza dei termini di errore è stimata come:

$$\hat{\sigma}^2 = \frac{\text{SSE}}{n - 1},$$

dove l’SSE ha la forma:

$$\text{SSE} := \sum_{i=1}^n w_i (y_i - x_i \hat{\beta})^2;$$

questa espressione coincide a sua volta con lo stimatore $\hat{\sigma}_j^2$ del DFCL dato dalla (1.5).

Il fatto qui rilevante è che i tradizionali test di ipotesi e di bontà di adattamento possono essere utilizzati per stabilire la significatività di queste stime e l’aderenza del modello ai dati, fornendo quindi un test dell’ipotesi M2MV, cioè sia dell’ipotesi M2M sulla media condizionata, sia dell’ipotesi M2V sulla varianza condizionata. Basterà quindi effettuare (con le opportune modifiche per tener conto della omoschedasticità) il test sulla statistica F

¹¹Dal punto di vista strettamente teorico, la relazione ricorsiva definita dall’ipotesi time series potrebbe generare valori negativi dei pagati cumulati $C_{i,j-1}$. Questo “negativity problem”, già ampiamente discusso nei “Comments” all’articolo originale, potrebbe essere aggirato riformulando le proprietà dei termini di errore $\varepsilon_{i,j}$ condizionatamente al valore assunto da $C_{i,j-1}$. Ciò condurrebbe a un modello con una struttura di dipendenza molto più complessa. Dato che il problema dei negativi è di solito irrilevante nelle applicazioni pratiche, nel TSCL si assume la posizione pragmatica di ignorare questa incoerenza teorica.

e/o sulla statistica t , corredata del relativo p -value, e ricavare misure di “fitting” (SSE) e di varianza spiegata (R^2). Se si esegue una analisi preliminare considerando un modello con intercetta, questa dovrebbe risultare significativamente diversa da zero. A corredo, si può aggiungere un’illustrazione grafica.

Osservazione. Tutte queste misure di significatività e goodness-of-fit rientrano tra i test standard eseguiti dalla procedura Explorer[©], finalizzata all’analisi esplorativa delle qualità di adattamento ai dati di vari modelli stocastici di loss reserving. Uno dei primi lavori dedicati alle metodologie di goodness-of-fit applicate ai modelli di loss reserving è stato proposto nel 1998 da Venter [22]; per sviluppi si veda [6]. ■

Di solito non si avranno dati sufficienti per effettuare tutte le J regressioni teoricamente richieste. Infatti il numero di osservazioni (accident year) disponibili per la stima della regressione j è dato da $I - j$, e decresce quindi al crescere di j . Se, per esempio, si accetta di considerare significativa una analisi di regressione con almeno 5 osservazioni, l’analisi relativa all’ultimo anno di sviluppo potrà essere effettuata solo se si ha a disposizione un trapezoide di paid losses con $I \geq J + 4$; nel caso usuale di “triangolo” ($I = J$) potranno essere considerate solo le prime $I + 1 - 5$ regressioni.

2.b.2 Ipotesi di indipendenza. Test sui residui “time series”

Un possibile metodo per verificare l’indipendenza tra anni di accadimento differenti consiste nel testare l’indipendenza dei residui ricavati dall’equazione time series (2.11). L’idea, proposta anche da Merz e Wüthrich [26], è quella di verificare, tramite una analisi di regressione lineare, l’assenza di trend nei residui considerati in funzione dell’accident year.

Considerando i fattori di sviluppo individuali $F_{i,j}$, la (2.11) si può scrivere:

$$F_{i,j} := \frac{C_{i,j}}{C_{i,j-1}} = f_{j-1} + \frac{\sigma_{j-1}}{\sqrt{C_{i,j-1}}} \varepsilon_{i,j}, \quad (2.12)$$

dove le $\varepsilon_{i,j}$ sono, per ipotesi, somiglianti e condizionatamente indipendenti, dato \mathcal{B}_0 , con media condizionata nulla e varianza condizionata unitaria. Quindi, se vale anche l’ipotesi M21 di indipendenza tra accident year le variabili aleatorie:

$$\epsilon_{i,j} := \frac{F_{i,j-1} - f_{j-1}}{\sqrt{\sigma_{j-1}^2 C_{i,j-1}^{-1}}},$$

sono indipendenti. Per verificare allora l’ipotesi M21 si può testare l’indipendenza delle $\epsilon_{i,j}$ sul trapezoide osservato, cioè l’indipendenza dei *residui*:

$$\epsilon_{i,j} = \frac{F_{i,j-1} - f_{j-1}}{\sqrt{\sigma_{j-1}^2 C_{i,j-1}^{-1}}}, \quad (2.13)$$

$$j = 1, 2, \dots, J, \quad i = 0, 1, \dots, I - j.$$

La numerosità di questi residui è $n^{TS} = J(I - J) + J(J + 1)/2$.

Tuttavia i residui espressi dalla (2.13) non sono in effetti osservabili, dato che i parametri f_j e σ_j non sono noti. Sostituendo nella (2.13) i valori dei parametri stimati con le (1.3) e (1.5) si ottengono allora gli n^{TS} residui osservabili:

$$\hat{\epsilon}_{i,j}^{TS} = \frac{F_{i,j-1} - \hat{f}_{j-1}}{\sqrt{\hat{\sigma}_{j-1}^2 C_{i,j-1}^{-1}}}, \quad (2.14)$$

$$j = 1, 2, \dots, J, \quad i = 0, 1, \dots, I - j,$$

sui quali è effettivamente possibile effettuare test di indipendenza.

L'indipendenza tra i residui di accident year diversi relativi a uno stesso development year risulta già implicitamente testata come conseguenza dell'analisi di regressione svolta per la verifica dell'ipotesi M2MV. Qui si richiede di più: la verifica esplicita dell'indipendenza tra i residui relativi a accident year diversi e a *qualsiasi* development year. Questo controllo di indipendenza può essere effettuato tramite analisi grafica. Se vale l'ipotesi di indipendenza, l'analisi grafica non dovrebbe evidenziare alcun trend nei residui al variare dell'anno di accadimento. L'assenza di trend può anche essere verificata con una analisi di regressione, effettuata per development year¹² oppure, più semplicemente, sull'intero campione dei residui.

Osservazione. La (2.14) può anche essere espressa nella forma:

$$\hat{\epsilon}_{i,j}^{TS} = \frac{C_{i,j} - \hat{C}_{i,j-1}^{TS}}{\hat{\sigma}_{j-1} \sqrt{C_{i,j-1}}}, \quad (2.15)$$

$$j = 1, 2, \dots, J, \quad i = 0, 1, \dots, I - j,$$

dove:

$$\hat{C}_{i,j-1}^{TS} := \hat{f}_{j-1} C_{i,j-1},$$

è interpretabile come il *fitted value* del modello TSCL. ■

Problemi di dipendenza spuria

È importante segnalare che i risultati dei test di indipendenza sui residui possono essere distorti da fenomeni di dipendenza spuria, indotti dall'uso degli stimatori chain ladder \hat{f}_j . In particolare, se si considerano le combinazioni lineari "di colonna" $\sum_{i=0}^{I-j} \sqrt{C_{i,j-1}} \hat{\epsilon}_{i,j}^{TS}$, si ricava che valgono le relazioni:

$$\sum_{i=0}^{I-j} \sqrt{C_{i,j-1}} \hat{\epsilon}_{i,j}^{TS} = 0, \quad j = 1, 2, \dots, J. \quad (2.16)$$

¹²In questo caso valgono anche qui le considerazioni sul numero minimo significativo di osservazioni fatte per la verifica dell'ipotesi M2MV: l'analisi dei residui andrà effettuata solo per gli anni di sviluppo con un numero di osservazioni considerato sufficiente.

Infatti, esprimendo i residui time series secondo la (2.15), si ha, per $j = 1, 2, \dots, J$:

$$\begin{aligned} \sum_{i=0}^{I-j} \sqrt{C_{i,j-1}} \hat{\epsilon}_{i,j}^{TS} &= \sum_{i=0}^{I-j} \sqrt{C_{i,j-1}} \left(\frac{C_{i,j} - \hat{f}_{j-1} C_{i,j-1}}{\hat{\sigma}_{j-1} \sqrt{C_{i,j-1}}} \right) \\ &= \frac{1}{\hat{\sigma}_{j-1}} \sum_{i=0}^{I-j} (C_{i,j} - \hat{f}_{j-1} C_{i,j-1}) \\ &= \frac{1}{\hat{\sigma}_{j-1}} \left(\sum_{i=0}^{I-j} C_{i,j} - \hat{f}_{j-1} \sum_{i=0}^{I-j} C_{i,j-1} \right) = 0, \end{aligned}$$

dove l'ultima uguaglianza è conseguenza della (1.3).

Le (2.16) evidenziano che tra i residui time serie osservabili (cioè calcolati con i parametri stimati) esistono correlazioni negative. In particolare, per $j = I - 1$ risulta che $\hat{\epsilon}_{0,I-1}^{TS}$ e $\hat{\epsilon}_{1,I-1}^{TS}$ sono perfettamente correlati negativamente.

Un'altro fatto collegato, ma di minore importanza, è che le (2.16) implicano la proprietà:

$$\sum_{j=1}^J \sum_{i=0}^{I-j} \sqrt{C_{i,j-1}} \hat{\epsilon}_{i,j}^{TS} = 0, \quad (2.17)$$

che è incompatibile con la proprietà:

$$\frac{1}{n^{TS}} \sum_{j=1}^J \sum_{i=0}^{I-j} \hat{\epsilon}_{i,j}^{TS} = 0. \quad (2.18)$$

Quindi la distribuzione dei residui osservati non è a media nulla.

Osservazione. Le proprietà (2.16) sono già state ricavate da Merz e Wüthrich in [26], par. 7.4, insieme alle proprietà sulla varianza:

$$\mathbf{Var}(\hat{\epsilon}_{i,j}^{TS} | \mathcal{B}_{j-1}) = 1 - \frac{C_{i,j-1}}{\sum_{i=0}^{I-j} C_{i,j-1}} < 1, \quad (2.19)$$

(essendo $\mathcal{B}_k := \{C_{i,j}; i + j \leq I, 0 \leq j \leq k\}$), dalle quali si evince che la distribuzione empirica dei residui ha varianza minore del valore unitario teorico. Tutte queste proprietà dei residui empirici sono state utilizzate dagli Autori in riferimento alla simulazione tramite bootstrap parametrico del modello TSCL. ■

2.b.3 Ipotesi di indipendenza. Test sui residui di Pearson

Una alternativa all'uso dei residui time series si ha considerando i *residui di Pearson* (non aggiustati)¹³:

$$\epsilon_{i,j}^P = \frac{X_{i,j} - X_{i,j}^{fit}}{\sqrt{X_{i,j}^{fit}}}, \quad (2.20)$$

$$j = 0, 1, \dots, J, \quad i = 0, 1, \dots, I - j,$$

dove $X_{i,j}^{fit}$ è il pagato incrementale *fitted* ottenuto con metodo “retrogrado” (*backcasting*) a partire dall'ultima diagonale osservata. Precisamente, si definiscono i pagati cumulati *fitted* tramite la procedure ricorsiva retrograda:

$$C_{i,j}^{fit} = \frac{C_{i,I-j}}{f_j f_{j+1} \cdots f_{I-j-1}}, \quad (2.21)$$

e da questi si ricavano come al solito, per differenza, i corrispondenti pagati incrementali $X_{i,j}^{fit}$. I residui di Pearson $\epsilon_{i,j}^P$ sono largamente utilizzati nella teoria dei *modelli lineari generalizzati* (GLM) e per questo motivo vengono di solito scelti come “generatori di rumore” nella simulazione con bootstrap del chain ladder stocastico quando questo è definito come un modello Over Dispersed Poisson (ODP); è infatti immediato verificare che l'ODP può essere riformulato come un modello GLM (si veda, per esempio, [8], [7], [4]). In questo quadro teorico, i residui di Pearson $\epsilon_{i,j}^P$ hanno media nulla e varianza costante (coincidente col parametro di sovradisersione ϕ dell'ODP).

Come osservato da Verral e England in [23] e [24], i valori fitted retrogradi definiti dalla (2.21) sono quelli più adeguati per costruire i residui di un modello ricorsivo come il chain ladder, e hanno qualità teorica superiore ai valori fitted del tipo $C_{i,j}^{TS} = f_{j-1} C_{i,j-1}$ usati per i residui time series.

La numerosità dei residui di Pearson è superiore a quella dei residui time series, avendosi $n^P = J(I - J) + (J + 1)(J + 2)/2$. Inoltre, mentre i residui time series sono variabili adimensionali (sono puri numeri), i residui di Pearson hanno dimensione euro^{1/2} (il loro quadrato ha dimensioni monetarie) e assumono quindi valori numerici su una scala diversa.

Naturalmente, anche i residui (2.20) non sono osservabili e la loro versione “osservabile” è fornita dalle stime:

$$\hat{\epsilon}_{i,j}^P = \frac{X_{i,j} - \hat{X}_{i,j}^{fit}}{\sqrt{\hat{X}_{i,j}^{fit}}}, \quad (2.22)$$

$$j = 0, 1, \dots, J, \quad i = 0, 1, \dots, I - j,$$

¹³I residui di Pearson “aggiustati” (per il numero di gradi di libertà) sono ottenuti moltiplicando i residui non aggiustati per il fattore $\sqrt{n^{TS}/(n^{TS} - p)}$, dove p è il numero di parametri. L'eventuale inclusione di questa riscalatura è ininfluente ai fini del test di indipendenza.

dove il pagato incrementale fitted stimato $\widehat{X}_{i,j}^{fit}$ è ottenuto dalla (2.21) sostituendo ai fattori di sviluppo incogniti f_j i corrispondenti stimatori chain ladder \hat{f}_j . Si può quindi supporre che anche i residui di Pearson osservati $\hat{\epsilon}_{i,j}^P$ contengano correlazione spuria indotta dall'utilizzazione di questi stimatori, anche se l'utilizzazione di una produttoria, invece che di un singolo stimatore, dovrebbe indurre legami di correlazione più blandi. Nel caso dei residui di Pearson, comunque, lo studio per via teorica di questi effetti è meno agevole che per i residui time series, e per effettuare un confronto tra le prestazioni fornite dai due metodi può essere conveniente ricorrere ad approcci empirici. Un utile confronto si può ottenere per simulazione, generando un campione di "pseudotrapezi" di pagati indipendenti e analizzando i due tipi di residui stimati su ciascun pseudotrapezio.

Parte II

Applicazione ai dati d'impresa

3 Premium Risk (modello M1)

3.a Definizione dei dati in input

Nel sottomodulo *Premium Risk*, fissato un *segmento* (“segment”: un raggruppamento di “lines of business” come descritto in [11]), per la stima della deviazione standard unitaria specifica dell'impresa è utilizzabile il solo modello M1, richiamato nel paragrafo 1.a. In questa applicazione del modello i dati utilizzati sono:

- Y_t : il costo dei sinistri (*aggregated losses*), ottenuto come la somma del pagato e del riservato per i sinistri accaduti e denunciati nell'esercizio t , con $t = 1, \dots, T$ e $T \geq 5$;
- X_t : i premi di competenza (*earned premiums*) dell'esercizio t , con $t = 1, \dots, T$ e $T \geq 5$.

Questi dati possono essere considerati secondo diverse modalità alternative di produzione (lavoro “diretto”, “indiretto”, “diretto+indiretto”) e al netto oppure al lordo dei recuperi.

Se l'opzione adottata per gli USP di premium risk è quella di applicare alle stime di “standard deviation for gross premium risk” il fattore di aggiustamento per la riassicurazione non proporzionale specificato in [10], all'articolo 117(3), tutti i dati considerati saranno al lordo della riassicurazione.

Si richiede che i dati ben rappresentino il “rischio di tariffazione” a cui la compagnia sarà esposta nell'esercizio successivo a quello di valutazione ($t = T + 1$).

Il costo dei sinistri Y_t è dato da:

$$Y_t = P_t + R_t - (P_t^r + R_t^r) - \Delta_t^r,$$

dove (con l'indicazione tra parentesi della voce del modulo di vigilanza IVASS n.17):

- P_t : importi pagati per sinistri accaduti nell'esercizio t (v10);
- R_t : riserva sinistri alla chiusura dell'esercizio per sinistri accaduti nell'esercizio t (v13);
- P_t^r : somme recuperate nell'esercizio, a titolo di franchigia, surrogazione e riscatto del sinistro, da assicurati e da terzi per sinistri accaduti nell'esercizio t (v14);
- R_t^r : somme da recuperare, a titolo di franchigia, surrogazione e riscatto del sinistro, da assicurati e da terzi alla chiusura dell'esercizio per sinistri accaduti nell'esercizio t (v15);
- Δ_t^r : saldo dei movimenti di portafoglio per sinistri accaduti nell'esercizio t (v17).

Osservazione. Nella modulistica di vigilanza è presente il dato aggregato delle somme recuperate e da recuperare, ossia $P_t^r + R_t^r$; il dato disaggregato è presente solo per il lavoro diretto lordo. Il pagato P_t può essere comprensivo dell'ammontare del contributo al F.G.V.S. (v301). ■

Il costo sinistri può essere aggiustato eliminando i costi relativi ai sinistri catastrofali nella misura in cui il rischio di tali sinistri sia già trattato nei sottomoduli dedicati.

I premi di competenza X_t sono definiti dalla:

$$X_t = EP_t = R_{t-1}^p + WP_t - R_t^p + \Delta_t^p + \Delta_t^{cp},$$

con:

- R_{t-1}^p : riserva premi alla chiusura dell'esercizio precedente $t - 1$ (v01);
- WP_t : premi contabilizzati nell'esercizio t (v03);
- R_t^p : riserva premi alla chiusura dell'esercizio t (v04);
- Δ_t^p : saldo dei movimenti di portafoglio relativo ai premi in entrata nell'esercizio t (v05);
- Δ_t^{cp} : saldo delle variazioni per differenza cambi derivante dall'allineamento delle riserve in valuta nell'esercizio t (v02).

3.b Applicazione del metodo

La undertaking-specific standard deviation relativa al segmento s , secondo il Method 1, è data da:

$$\sigma_{(prem,s,USP)} = c \cdot \hat{\sigma}(\hat{\delta}, \hat{\gamma}) \cdot \sqrt{\frac{T+1}{T-1}} + (1-c) \cdot \sigma_{(prem,s)},$$

dove:

- T è l'ampiezza in anni della serie storica (annuale);

- c è il fattore di “credibilità”;
- $\sigma_{(prem,s)}$ è il livello standard (market-wide) della deviazione standard unitaria prescritto da EIOPA;
- $\hat{\sigma}(\hat{\delta}, \hat{\gamma})$ è la stima della deviazione standard unitaria specifica dell’impresa fornita dal modello M1, ottenuta minimizzando la criterion function $\ell(\hat{\delta}, \hat{\gamma})$ (riportata nel paragrafo 1.a) nell’intervallo $D = \{\hat{\delta} \in [0, 1], \hat{\gamma} \in \mathbb{R}\}$.

3.c Sulla tecnica di minimizzazione

Per l’identificazione di $\hat{\delta}$ e $\hat{\gamma}$ si può utilizzare una routine di minimizzazione (per esempio la routine E04JYF o la E04WDF della NAG Fortran Library), oppure un metodo “a griglia”, oppure una combinazione dei due approcci (applicazione della routine di minimizzazione inizializzata nel punto di minimo della griglia). Il calcolo su griglia è comunque utile per analizzare le proprietà di regolarità della criterion function.

4 Reserve Risk – modello M1

4.a Definizione dei dati in input

Nel *Reserve Risk method 1* sono utilizzati, per la stima della deviazione standard unitaria specifica dell’impresa, fissata una Lob e il tipo di lavoro (produzione, riassicurazione):

- Y_t gli impegni stimati a fine esercizio, ottenuti come la somma del pagato e del riservato dei sinistri accaduti negli esercizi precedenti rispetto all’esercizio t , con $t = 1, \dots, T$ e $T \geq 5$ (*year-end obligations*);
- X_t gli impegni stimati all’inizio dell’esercizio (o riserva iniziale), relativi ai sinistri accaduti negli esercizi precedenti rispetto all’esercizio t , con $t = 1, \dots, T$ e $T \geq 5$ (*initial outstanding*).

Si richiede che i dati ben rappresentino il “rischio di riservazione” a cui la compagnia sarà esposta nell’esercizio successivo a quello di valutazione ($t = T + 1$).

Dato che nel caso del reserve risk non è previsto un coefficiente di riduzione “Gross-to-Net” standard, questi dati dovrebbero essere al netto della riassicurazione. La richiesta, tuttavia, non è solitamente facile da soddisfare, dato che i dati “su triangolo” sono usualmente disponibili al lordo della riassicurazione, e la trasformazione da pagati lordi a pagati netti è in genere complessa e delicata. Negli Atti Delegati, al punto D(2)(f) dell’Annesso XVII, si richiede che dai pagati cumulati vengano detratte le somme recuperate dai contratti riassicurativi che siano “coerenti” (*consistent*) con i contratti riassicurativi *that are in place to provide cover for the following*

twelve months. In molti casi, però, queste procedure richiedono interventi non banali di interpretazione e ricostruzione, che potrebbero portare a distorsioni significative della variabilità intrinseca nei dati (che è proprio quello che si dovrebbe stimare). Per questi motivi – anche in ossequio alle prescrizioni sulla “data quality” –, potrebbe essere preferibile adottare un atteggiamento prudentiale, effettuando la stima della deviazione standard unitaria per il reserve risk direttamente sui dati lordi, rinunciando alla potenziale riduzione del requisito che si avrebbe considerando dati netti correttamente ricostruiti.

Gli impegni Y_t stimati a fine esercizio sono dati da:

$$Y_t = P_t + R_t - (P_t^r + R_t^r),$$

dove (con l’indicazione tra parentesi della voce del modulo di vigilanza IVASS n.17):

- P_t importi pagati per sinistri accaduti negli esercizi precedenti rispetto all’esercizio t (v26);
- R_t riserva sinistri alla chiusura dell’esercizio t per sinistri accaduti negli esercizi precedenti rispetto all’esercizio t (v29);
- P_t^r somme recuperate nell’esercizio, a titolo di franchigia, surrogazione e riscatto del sinistro, da assicurati e da terzi per sinistri accaduti negli esercizi precedenti rispetto all’esercizio t (v32);
- R_t^r somme da recuperare, a titolo di franchigia, surrogazione e riscatto del sinistro, da assicurati e da terzi alla chiusura dell’esercizio per sinistri accaduti negli esercizi precedenti rispetto all’esercizio t (v33).

Osservazione. Nella modulistica di vigilanza è presente il dato aggregato delle somme recuperate e da recuperare, ossia $P_t^r + R_t^r$, il dato disaggregato è presente solo per il lavoro diretto lordo. ■

Gli impegni X_t stimati all’inizio dell’esercizio sono così definiti:

$$X_t = R_{t-1} - R_{t-1}^r + \Delta_t^r + \Delta_t^{cr},$$

dove (con l’indicazione tra parentesi della voce del modulo di vigilanza IVASS n.17):

- R_{t-1} riserva sinistri alla chiusura dell’esercizio precedente $t-1$ (v21);
- R_{t-1}^r somme da recuperare da assicurati e da terzi alla chiusura dell’esercizio precedente $t-1$ (v31);
- Δ_t^r saldo dei movimenti di portafoglio per sinistri accaduti negli esercizi precedenti rispetto all’esercizio t (v30);
- Δ_t^{cr} saldo delle variazioni per differenza cambi derivante dall’allineamento delle riserve in valuta per sinistri accaduti negli esercizi precedenti rispetto all’esercizio t (v22).

Osservazione. Nella modulistica di vigilanza è presente il dato relativo al risultato economico delle somme recuperate e da recuperare, ossia $R_{t-1}^r - (P_t^r + R_t^r)$, il dato disaggregato è presente solo per il lavoro diretto lordo. ■

4.b Applicazione del metodo

La undertaking-specific standard deviation relativa al segmento s , secondo il Method 1, è data da:

$$\sigma_{(res,s,USP)} = c \cdot \hat{\sigma}(\hat{\delta}, \hat{\gamma}) \cdot \sqrt{\frac{T+1}{T-1}} + (1-c) \cdot \sigma_{(res,s)},$$

dove:

- T è l'ampiezza in anni della serie storica (annuale);
- c è il fattore di "credibilità";
- $\sigma_{(res,s)}$ è il livello standard (market-wide) della deviazione standard unitaria prescritto da EIOPA;
- $\hat{\sigma}(\hat{\delta}, \hat{\gamma})$ è la stima della deviazione standard unitaria specifica dell'impresa fornita dal modello M1, ottenuta minimizzando la criterion function $\ell(\hat{\delta}, \hat{\gamma})$ (riportata nel paragrafo 1.a) nell'intervallo $D = \{\hat{\delta} \in [0, 1], \hat{\gamma} \in \mathbb{R}\}$.

4.c Sulla tecnica di minimizzazione

Valgono le stesse considerazioni svolte nel paragrafo 3.c.

5 Reserve Risk – modello M2

5.a Definizione dei dati in input

Nel *Reserve Risk – method 2* i dati utilizzati per la stima della deviazione standard unitaria specifica dell'impresa, fissata una Lob, il tipo di lavoro (produzione, riassicurazione), sono i seguenti:

- $X_{i,j}$ importi pagati per i sinistri accaduti nell'anno di accadimento i , con $i = 1, \dots, I$ e $I \geq 5$ e pagati con j anni di differimento, con $j = 1, \dots, J$ e $J \leq I$ (*paid trapezoid*);

Gli importi pagati $X_{i,j}$ sono così definiti:

$$X_{i,j} = X_{i,j}^{lr} - X_{i,j}^r,$$

con:

- $X_{i,j}^{lr}$ importi pagati per i sinistri accaduti nell'anno di accadimento i e pagati j anni di differimento al lordo delle relative somme recuperate;
- $X_{i,j}^r$ Somme recuperate, a titolo di franchigia, surrogazione e riscatto del sinistro, da assicurati e da terzi per sinistri accaduti nell'anno di accadimento i e recuperate con j anni di differimento.

5.b Applicazione del metodo

La undertaking-specific standard deviation relativa al segmento s , secondo il Method 2, è data da:

$$\sigma_{(res,s,USP)} = c \cdot \widehat{\mathbf{Cv}}_{(res,s)} + (1 - c) \cdot \sigma_{(res,s)},$$

dove:

- $\widehat{\mathbf{Cv}}_{(res,s)}$ è la stima del coefficiente di variazione delle *Outstanding Loss Liabilities* (stima della deviazione standard unitaria specifica dell'impresa) espressa dalla (1.6),
- c è il fattore di “credibilità”;
- $\sigma_{(res,s)}$ è il livello standard (market-wide) della deviazione standard unitaria prescritto da EIOPA;

Il valore di $\widehat{\text{MSEP}}$ e di \widehat{R} , rispettivamente numeratore e denominatore di $\widehat{\mathbf{Cv}}_{(res,s)}$, sono ricavati con le espressioni in forma chiusa fornite dal modello M2 (modello di Merz e Wüthrich).

Riferimenti bibliografici

- [1] Anderson, T. W., Darling, D. A., *A test of goodness of fit*, Journal of the American statistical Association, 765-769, 1954.
- [2] Cramér, H., *On the composition of elementary errors*, Scandinavian Actuarial Journal, 1928.
- [3] Buchwalder, M., Bühlmann, H., Merz, M., Wüthrich, M.V., *The mean square error of prediction in the chain ladder reserving method (Mack and Murphy revisited)*, ASTIN Bulletin **36(2)**, 521-542, 2006.
- [4] De Felice, M., Moriconi, F., *Risk based capital in P&C loss reserving or stressing the triangle*, Research Group on “Insurance Companies and Pension Funds”, Working Paper n.1, Rome, December 2003.
- [5] De Felice, M., Moriconi, F., *Process error and estimation error of year-end reserve estimation in the distribution free chain-ladder model*, Alef Working Paper, Rome, November 2006.
- [6] De Felice, M., Moriconi, F., *Measuring Goodness of Fit of Loss Reserving Models*, Rome, Alef Working Paper, July 2008.
- [7] England, P. , *Addendum to “Analytic and bootstrap estimates of prediction errors in claims reserving”*, Insurance: Mathematics and Economics, vol. 31, pp. 461-466, 2002.
- [8] England, P., Verral, R., *Analytic and bootstrap estimates of prediction errors in claims reserving*, Insurance: Mathematics and Economics, vol. 25 pp. 281-293, 1999.
- [9] De Felice, M., Moriconi, F., Cavastracci, S., Matarazzo, L., Pasqualini, S., *Reserve requirements and capital requirements in non-life insurance. An analysis of the Italian MTPL insurance market by stochastic claims reserving models*, ISVAP, Rome, October 2006.
- [10] European Commission, *Text of the delegated Act on Solvency II*, 10 October 2014.
(<http://ec.europa.eu/finance/insurance/solvency/solvency2/index.en.htm>)
- [11] European Commission, *Annex 1 to 21 of the delegated Act on Solvency II*, 10 October 2014.
(<http://ec.europa.eu/finance/insurance/solvency/solvency2/index.en.htm>)
- [12] EIOPA, *Calibration of the Premium and Reserve Risk factors in Standard Formula of Solvency II*, Report of the Joint Working Group on Non-Life and Health NSLT Calibration, 12 December 2011.
- [13] Embrechts, P., Klüppelberg, C., Mikosch, T., *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag 1997.
- [14] IVASS, *Solvency II – utilizzo dei parametri specifici dell’impresa nei moduli del rischio di sottoscrizione per l’assicurazione vita, per l’assicurazione non vita e per l’assicurazione malattia*, Lettera al mercato, 4 novembre 2014.

- [15] Jarque, C.M., Bera, A.K., *Efficient tests for normality, homoscedasticity and serial independence of regression residuals*, Economics Letters, **6**(3), 255-259, 1980.
- [16] Klugman, S.A., Panjer, H.H., Willmot, E., *Loss Models. From data to decisions*, Wiley 2012.
- [17] Mack, T. , *Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates* ASTIN Bulletin **23**(2), 213-225, 1993.
- [18] Massey, Jr., Frank J., *The Kolmogorov-Smirnov test for goodness of fit*, Journal of the American Statistical Association, 68-78, 1951.
- [19] Merz, M., Wüthrich, M. V., *Modelling Claims Development Result for Solvency Purposes*, Casualty Actuarial Society, E-Forum, Fall 2008.
- [20] Mikosch, T., *Non-Life Insurance Mathematics. An introduction with the Poisson process*, 2nd ed., Universitext, Springer, Berlin, 2009.
- [21] Shapiro, S.S, Wilk, M.B., *An analysis of variance test for normality (complete samples)*, Biometrika **52**(3-4), 591-611, 1965
- [22] Venter, G., *Testing the assumptions of age to age factors*, PCAS 1998.
- [23] Verral, R.J. *An investigation into stochastic claims reserving models and the chain-ladder technique*, Insurance: Mathematics and Economics, vol. 26, pp. 91-99, 2000.
- [24] Verrall, R.J., England, P.D., *Comments on "A Comparison of Stochastic Models that Reproduce Chain Ladder Reserve Estimates"*, by Mack and Venter, Insurance: Mathematics and Economics, vol. 26, pp. 109-111, 2000.
- [25] von Mises, R.E., *Wahrscheinlichkeit, Statistik und Wahrheit*, Julius Springer, 1928.
- [26] Wüthrich, M. V., Merz, M., *Stochastic Claims Reserving Methods in Insurance*, Wiley Finance, 2008.