

# *ALEA*

## *Tech Reports*

---

Un modello per l'incorporazione  
del rischio specifico nel VaR

Marco Bee

Tech Report Nr. 13  
Gennaio 2002

---

**Alea - Centro di ricerca sui rischi finanziari**  
**Dipartimento di informatica e studi aziendali**  
**Università di Trento - Via Inama 5 - 38100 - Trento**  
**<http://www.aleaweb.org/>**

*ALEA, Centro di ricerca sui rischi finanziari è un centro di ricerca indipendente costituito presso il Dipartimento di informatica e studi aziendali dell'Università di Trento. Il centro si propone di elaborare conoscenze innovative in materia di risk management, e di favorirne la diffusione mediante pubblicazioni e programmi di formazione. Il centro è diretto dal prof. Luca Erzegovesi. Nella collana ALEA Tech Reports sono raccolti lavori di compendio e rassegna della letteratura e della prassi operativa su tematiche fondamentali di risk management.*

# UN MODELLO PER L'INCORPORAZIONE DEL RISCHIO SPECIFICO NEL VAR

Marco Bee

*IntesaBci*

*Risk Management - Development & Research (\*)*

## RIASSUNTO

Questo articolo tratta il problema di incorporare il rischio specifico nel calcolo del VaR. Il problema viene affrontato ipotizzando che il processo generatore dei dati sia una mistura di tre distribuzioni normali: la distribuzione relativa ai periodi "normali" genera la maggior parte delle osservazioni, le altre due distribuzioni sono caratterizzate da una media spostata rispettivamente verso il basso e verso l'alto. L'entità di tali salti permette di incorporare il rischio di evento, che si riflette in una distribuzione a code pesanti. La metodologia utilizzata per stimare i parametri, basata sull'algoritmo EM, permette di stimare tutti i parametri a partire dai dati, senza introdurre alcuna ipotesi a priori. Gli esempi presentati confermano, anche tramite un confronto con altre metodologie, l'appropriatezza del modello per la stima del VaR.

---

(\*) Desidero ringraziare Giuseppe Espa e Demetrio Maffei per utili suggerimenti e discussioni. La responsabilità per eventuali errori è interamente mia.

# 1 Introduzione

Il Valore a Rischio (VaR) è da alcuni anni divenuto il principale strumento utilizzato dai dipartimenti di Risk Management delle grandi banche. Date la diffusione e l'importanza di questa misura nella determinazione del capitale allocato, si è fatto ricorso ad una metodologia di misurazione standardizzata e computazionalmente stabile, che è stata codificata da RiskMetrics [RiskMetrics (1995)] ed è basata sull'ipotesi di normalità dei rendimenti.

L'attività di ricerca sull'argomento negli ultimi anni ha d'altro canto verificato che la distribuzione dei rendimenti si allontana significativamente da questa ipotesi, per via di una curtosi più alta di quella della distribuzione normale (leptocurtosi). Ciò implica che le code della distribuzione sono più pesanti di quanto ipotizzato dalla normale e dunque che il VaR, essendo un quantile della distribuzione, è sottostimato, in particolare per alti livelli di confidenza: questo è uno dei motivi per cui nello standard RiskMetrics il VaR è calcolato al 95%. Tuttavia è sempre più evidente che le istituzioni finanziarie, essendo generalmente attrezzate a sostenere perdite che si verificano una volta ogni 20 giorni (questa è l'informazione fornita dal VaR al 95%), sono interessate a quantificare la probabilità di perdite più rare ma più consistenti e quindi potenzialmente più pericolose. In questo caso il VaR normale è senza dubbio inadeguato ed è necessario utilizzare tecniche che permettano di stimare in modo più preciso la coda della distribuzione.

Quanto detto finora vale per il rischio di mercato, ovvero per il rischio derivante da fluttuazioni di prezzi e tassi al quale sono soggette tutte le attività finanziarie. E' tuttavia evidente che nella maggior parte dei casi (vanno escluse solo le obbligazioni governative dei Paesi con il rating più alto, comunemente considerate *risk-free*) esiste un'altra dimensione di rischio, legata alla specifica attività finanziaria. E' ciò che si definisce comunemente *rischio specifico* o *rischio di evento*, legato per esempio a notizie di *default* o profitti imprevisti. I modelli comunemente usati per misurare il rischio di mercato sono certamente inadatti a cogliere questo aspetto del rischio di un'attività finanziaria ed è quindi necessario sviluppare modelli alternativi.

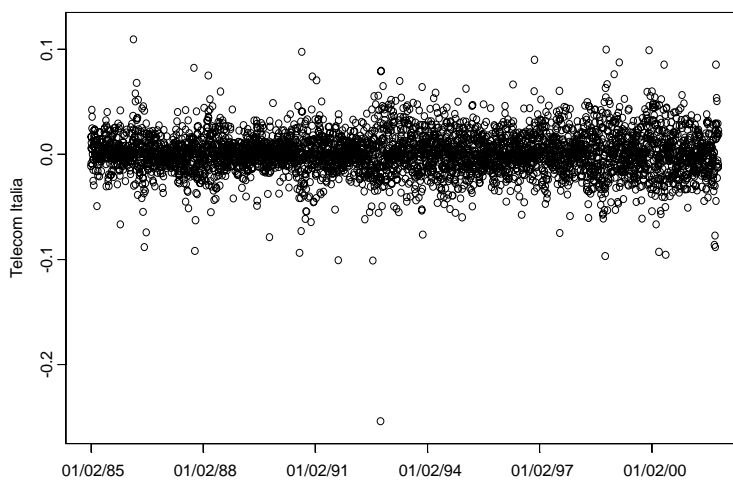
L'articolo è organizzato come segue. La sezione 2 costituisce una breve rassegna critica delle metodologie proposte per il calcolo del VaR. Nella sezione 3 si presentano il modello *Jump-VaR* e la sua interpretazione in ambito finanziario. La sezione 4 è dedicata alle applicazioni del modello al calcolo del VaR di vari titoli azionari. La sezione 5 conclude. In appendice vengono forniti dettagli tecnici sugli aspetti statistici della metodologia di stima.

## 2 Modelli per il calcolo del VaR

Le metodologie per il calcolo del VaR sono classificabili in vari modi. Alcune metodologie (GARCH, normalità alla RiskMetrics) sono condizionali, cioè la misura di VaR dipende dalle condizioni di mercato dei giorni precedenti, altre (Teoria del Valore Estremo, simulazione storica) sono non condizionali. E' inoltre possibile riferirsi al fatto che si stimi tutta la distribuzione (normalità, GARCH, simulazione storica) oppure solo la coda (Teoria del Valore Estremo).

L'uso di modelli condizionali o non condizionali è tuttora oggetto di dibattito. La metodologia standard proposta da RiskMetrics fornisce una misura di VaR condizionale, e ciò ha contribuito a far sì che nella pratica si tendano a considerare più appropriati i metodi condizionali. Tuttavia questo punto di vista non sembra completamente condivisibile. Infatti il motivo principale per cui i modelli della famiglia GARCH hanno avuto successo nel prevedere la volatilità di serie finanziarie sta nel fatto che tali serie presentano evidenti fenomeni di *volatility clustering*: le osservazioni caratterizzate da alta volatilità tendono ad essere adiacenti, e dunque ha certamente senso costruire la previsione della volatilità per il tempo  $t + 1$  facendola dipendere dalla volatilità al tempo  $t$ .

Fig. 1: Rendimenti giornalieri del titolo Telecom Italia



Nel calcolo del VaR siamo invece interessati a cogliere correttamente gli eventi di coda. Tali eventi, per definizione, si presentano raramente; inoltre non sembrano essere dipendenti dal livello di volatilità né presentare dipendenza temporale. Come dimostrano Danielsson - de Vries (1997, sezione 2), tali osservazioni possono essere considerate iid. Senza scendere in dettagli, si consideri in proposito la figura 1, che riporta i rendimenti giornalieri del titolo Telecom Italia su un periodo di circa 15 anni: le osservazioni estreme appaiono distribuite in modo pressoché uniforme su tutto il periodo.

L'utilizzo di metodi non condizionali ha l'importante conseguenza che (a differenza di ciò che accade con i metodi condizionali) il VaR può essere calcolato con tutte le osservazioni disponibili, il che è fondamentale per stimare correttamente la coda della distribuzione, dove le realizzazioni della variabile casuale che genera i dati sono rare.

Il modello più usato per il calcolo del VaR è basato sulla distribuzione normale. In particolare, RiskMetrics si basa sull'ipotesi di normalità condizionale: la distribuzione della serie dei rendimenti al tempo  $t$  date le osservazioni precedenti è data da

$$r_t \sim N(0, \sigma_t^2).$$

Si tratta quindi di una misura di VaR condizionale. Per stimare la varianza sono utilizzati due metodi:

- (i) GARCH(1, 1) [Engle (1982), Bollerslev (1986)]:

$$\sigma_t^2 = \text{var}(r_t|r_{t-1}) = \alpha\sigma_{t-1}^2 + \beta r_{t-1}^2, \quad \alpha + \beta < 1. \quad (1)$$

- (ii) *Exponentially Weighted Moving Average* (EWMA): è lo stimatore più utilizzato e costituisce un caso particolare del GARCH; la stima della varianza è infatti data da:

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1 - \lambda)r_{t-1}^2, \quad \lambda < 1.$$

E' chiaro che lo stimatore si ottiene dalla (1) ponendo  $\beta = 1 - \alpha$ . Questo modello, denominato IGARCH (Integrated GARCH), è poco adeguato per almeno due motivi. In primo luogo, essendo un processo non stazionario, la varianza non condizionale è infinita; ciò tuttavia non influenza significativamente i risultati in quanto gli orizzonti temporali di previsione sono di solito brevi. Un problema più grave è il fatto che, nel caso

multivariato, esiste una probabilità positiva che la stima della matrice di covarianza così ottenuta sia definita negativa.

Parallelamente alla metodologia RiskMetrics si sono dunque sviluppati molti modelli che tengono conto, nella stima del VaR, della pesantezza delle code della distribuzione. Si vedano per esempio Longin (2001) e Danielsson - de Vries (1997) per una rassegna critica.

La simulazione storica consiste nel ricampionare dalla distribuzione dei rendimenti più recenti (di solito un anno); è semplice da implementare ma ha il difetto di essere implicitamente basata sull'ipotesi che il processo generatore dei rendimenti resti invariato nel tempo. Essendo una metodologia non parametrica, non può fornire alcuna informazione sulla probabilità di osservazioni al di fuori del *range* campionario. Poiché il periodo considerato è di solito pari ad un anno, è molto probabile che non sia disponibile un numero sufficientemente grande di osservazioni estreme e che di conseguenza le code non vengano stimate correttamente.

La Teoria del Valore Estremo (EVT) è una metodologia che si propone di stimare solo la coda della distribuzione, cioè quella parte della distribuzione che ha generato le osservazioni al di sopra (al di sotto) di una soglia  $u$  sufficientemente grande (piccola). Rimandando a Embrechts *et al.* (1997) per i dettagli, gli eccessi  $y_i = x_i - u$  hanno approssimativamente, per un valore di  $u$  sufficientemente grande, distribuzione di Pareto generalizzata.

La metodologia è molto potente ma dipende dalla soglia oltre la quale si decide di stimare la distribuzione. La scelta di tale soglia non è sempre agevole: i metodi grafici presentati, per esempio, in Embrechts *et al.* (1997) non forniscono risposte univoche e l'unica metodologia analitica disponibile [Danielsson *et al.* (2001)] è di difficile implementazione. Inoltre, solo le osservazioni al di sopra della soglia vengono utilizzate per la stima e quindi è necessaria una serie molto lunga per avere a disposizione un numero sufficiente di dati: se la soglia è troppo alta, gli stimatori avranno una varianza molto alta, se è troppo bassa l'approssimazione tramite la distribuzione di Pareto generalizzata non è più appropriata.

Nel seguito si tratterà della metodologia *Jump-VaR*. Da un punto di vista statistico, il suo impiego è giustificato in quanto può approssimare arbitrariamente bene qualsiasi distribuzione di probabilità; da un punto di vista finanziario rappresenta un modello che spiega il rischio specifico in termini di un salto nel rendimento.

### 3 Il modello Jump-VaR

Consideriamo la serie storica di un'attività finanziaria. Se la serie è sufficientemente lunga, si osservano generalmente alcuni (rari) salti, cioè variazioni di rendimento molto più pronunciate di quelle che si osservano in tempi "normali". Tipicamente, questi salti sono collegati ad eventi specifici all'emittente dell'attività in questione: nel caso di un salto negativo (positivo) nel prezzo di un'obbligazione *corporate* o delle obbligazioni governative di un Paese emergente si può pensare ad un *downgrade* (*upgrade*) dell'emittente. In ambito azionario, i salti derivanti dal rischio specifico saranno spesso coincidenti con le relazioni trimestrali.

Il processo generatore dei dati, descritto in Gibson (2001), è dunque dato da una distribuzione che genera le osservazioni relative ai periodi "normali", vale a dire privi di salti, e da una distribuzione che genera la componente di salto:

$$X = X_0 + X_J,$$

dove  $X_0$  può avere qualsiasi distribuzione e la componente di salto ha una distribuzione trinomiale:

$$X_J = \begin{cases} 0 & \text{con probabilità } (1 - p_D - p_U); \\ D & \text{con probabilità } p_D; \\ U & \text{con probabilità } p_U, \end{cases}$$

dove in generale sarà  $D \in \mathbf{R}^-$  e  $U \in \mathbf{R}^+$ . Ne segue che la distribuzione di  $X$  è una mistura delle distribuzioni di tre variabili casuali  $X_D$ ,  $X_0$  e  $X_U$ , avente densità:

$$f_X(x) = p_D f_D(x) + (1 - p_D - p_U) f_0(x) + p_U f_U(x), \quad x \in \mathbf{R}, \quad (2)$$

dove  $X_D$  e  $X_U$  hanno la stessa distribuzione di  $X$  con la media spostata rispettivamente di  $D$  ed  $U$ <sup>(1)</sup>. Nell'ipotesi di normalità, che si adotterà nel prosieguo del lavoro, si ha che  $f_0(x) \sim N(\mu_2, \sigma^2)$ ,  $f_D(x) \sim N(\mu_2 - D, \sigma^2)$ ,  $f_U(x) \sim N(\mu_2 + U, \sigma^2)$ . La rarità dei salti si rifletterà nel fatto che le probabilità a priori  $p_D$  e  $p_U$  saranno "piccole".

L'idea di calcolare il VaR modellando la distribuzione dei rendimenti tramite una mistura non è nuova: si vedano Hamilton (1991), Venkataraman

---

<sup>(1)</sup> La formalizzazione della mistura data in Gibson (2001), sebbene i ragionamenti ed i calcoli successivi siano corretti, è errata, in quanto rappresenta la somma di tre variabili casuali normali, che ha ancora distribuzione normale.



(1997), Zangari (1996) e la nuova edizione del documento tecnico di RiskMetrics [RiskMetrics (2001), pag. 98], ma la sua diffusione è stata probabilmente limitata dalla convinzione che la stima dei parametri con il metodo della massima verosimiglianza sia problematica, almeno nel caso di varianze diverse nelle popolazioni. In proposito, vale la pena di osservare che:

- (i) la stima di massima verosimiglianza non presenta alcuna difficoltà quando le varianze delle popolazioni siano uguali. Infatti in questo caso la funzione di verosimiglianza è limitata e spesso ha un unico massimo;
- (ii) quando le varianze siano diverse, la funzione di verosimiglianza presenta effettivamente delle singolarità. A titolo di esempio, si consideri la funzione di log-verosimiglianza per una mistura di due normali con medie e varianze diverse:

$$l(\boldsymbol{\theta}; x_1, \dots, x_n) = \sum_{i=1}^n \log \left( \frac{p_1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2} \frac{(x_i - \mu_1)^2}{\sigma_1^2}} + \frac{p_2}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2} \frac{(x_i - \mu_2)^2}{\sigma_2^2}} \right).$$

Se si pone, per esempio,  $x_1 = \mu_1$ , è immediato verificare che, quando  $\sigma_1^2 \rightarrow 0$ ,  $l(\boldsymbol{\theta}; x_1, \dots, x_n) \rightarrow +\infty$ . Questo problema non sussiste nel caso (i), in quanto, sotto la condizione sempre verificata in pratica che il numero di osservazioni disponibili sia maggiore del numero di popolazioni della mistura, la stima della varianza non può essere arbitrariamente piccola e quindi la funzione di verosimiglianza è limitata.

Comunque anche nel caso (ii) l'algoritmo, se inizializzato appropriatamente, converge al massimo globale della funzione di verosimiglianza, e fornisce stimatori che godono delle usuali proprietà degli stimatori di massima verosimiglianza. Dunque in questo caso il problema, puramente numerico, consiste in "keeping the variance away from zero" [Flury (1997, pag. 653)], per far sì che l'algoritmo converga ad un massimo locale. Per assicurarsi che il massimo locale raggiunto sia anche il massimo globale sarà eventualmente utile tentare diverse inizializzazioni.

Come testimoniano numerose applicazioni pratiche in diversi campi, in tutti i casi in cui si abbia a disposizione a priori una stima approssimativa dei parametri con cui inizializzare l'algoritmo non si presenta nessuno dei problemi ora menzionati. Nella nostra esperienza la stima di massima verosimiglianza di misture normali, anche nel caso (ii), non ha mai comportato sostanziali difficoltà ed ha prodotto risultati incoraggianti [Bee (2001)].

Tornando al modello *jump-VaR* esposto nella (2), Gibson (2001) utilizza, per la calibrazione del modello, un metodo in due passi che stima separatamente la distribuzione  $f_0(x)$  ed i parametri di salto. Questi ultimi vengono

stimati decidendo a priori, arbitrariamente, quali osservazioni vengono considerate salti [Gibson (2001), pag. 9]. Questa decisione risulta poi avere un impatto rilevante sul VaR.

In questo lavoro si è invece optato per un'interpretazione del modello che consente di stimare tutti i parametri endogenamente con il metodo della massima verosimiglianza, tramite l'algoritmo EM<sup>(2)</sup>. Oltre ad evitare qualsiasi elemento di arbitrarietà, tale metodo fornisce stime che godono delle consuete proprietà degli stimatori di massima verosimiglianza. Inoltre il modello, grazie al fatto che la varianza delle tre popolazioni è uguale, rientra nel caso (i), dunque la funzione di verosimiglianza è limitata e l'algoritmo, a meno di scelte estremamente anomale nei valori di inizializzazione, che in pratica si possono sempre evitare, converge al massimo globale della funzione di verosimiglianza.

Un altro processo generatore dei dati basato sulle distribuzioni a mistura finita è stato utilizzato in Bee (2001). Si tratta di un modello che prevede che la maggior parte delle osservazioni siano generate da una normale a bassa volatilità, mentre le rimanenti osservazioni provengono da una normale con la stessa media ma volatilità molto maggiore. Trattandosi di rendimenti, è ragionevole ipotizzare che la media sia nulla. Formalmente si ha

$$f_X(x) = p_1 f_1(x) + p_2 f_2(x), \quad x \in \mathbf{R}, \quad (3)$$

dove  $f_i(x)$  è la densità normale di parametri  $\mu$  e  $\sigma_i^2$ ,  $i = 1, 2$ , con  $\sigma_2^2 \gg \sigma_1^2$ . Anche in questo caso il modello è dunque agevolmente interpretabile in termini sia finanziari che statistici e nella prossima sezione la sua performance verrà confrontata con quella del *Jump-VaR*.

## 4 Applicazioni e risultati

Per il calcolo del VaR sono stati utilizzati i rendimenti giornalieri delle seguenti serie: Fiat, Olivetti, Telecom Italia (per queste tre azioni si è considerato il periodo 2 gennaio 1985 - 17 ottobre 2001), Sun Microsystem (13 maggio 1986 - 17 ottobre 2001), IBM (19 ottobre 1970 - 17 ottobre 2001), MIB30 (4 gennaio 1983 - 17 ottobre 2001), Standard & Poor's 500 (1 gennaio 1970 - 17 ottobre 2001)<sup>(3)</sup>.

---

<sup>(2)</sup> Alcuni dettagli sul funzionamento dell'algoritmo sono forniti in appendice.

<sup>(3)</sup> Tutti i dati sono stati ottenuti da Bloomberg<sup>TM</sup>.

Nella tabella 1 sono riportate le stime dei parametri ottenute con la metodologia *Jump-VaR* (J) e con la metodologia esposta nella (3), che chiameremo *Emphasized Volatility* (E); i parametri  $\mu_1$  e  $\mu_3$  nel caso *Jump-VaR* sono da interpretare rispettivamente come  $\mu_2 + D$  e  $\mu_2 + U$ , con  $D \in \mathbf{R}^-$  e  $U \in \mathbf{R}^+$ .

**Tabella 1 - Stime dei parametri<sup>(4)</sup>**

	$p_1$	$p_2$	$p_3$	$\mu_1$	$\mu_2$	$\mu_3$	$\sigma_1$	$\sigma_2$
Sun <sub>J</sub>	0.011	0.978	0.011	-0.123	0.001	0.113	0.003	-
Sun <sub>E</sub>	0.886	0.114	-	0.001	-0.001	-	0.027	0.071
IBM <sub>J</sub>	0.006	0.982	0.012	-0.093	-0.001	0.069	0.016	-
IBM <sub>E</sub>	0.882	0.118	-	0.000	0.002	-	0.014	0.042
Fiat <sub>J</sub>	0.012	0.957	0.031	-0.064	-0.001	0.054	0.016	-
Fiat <sub>E</sub>	0.785	0.215	-	-0.001	0.004	-	0.013	0.034
OL <sub>J</sub>	0.016	0.958	0.026	-0.077	-0.001	0.065	0.02	-
OL <sub>E</sub>	0.696	0.304	-	-0.001	0.001	-	0.015	0.038
TI <sub>J</sub>	0.011	0.968	0.021	-0.07	0.000	0.052	0.017	-
TI <sub>E</sub>	0.78	0.22	-	0.000	0.002	-	0.014	0.034
MIB <sub>J</sub>	0.015	0.974	0.011	-0.045	0.001	0.043	0.013	-
MIB <sub>E</sub>	0.857	0.143	-	0.001	-0.001	-	0.012	0.026
SPX <sub>J</sub>	0.001	0.990	0.009	-0.079	0.000	0.033	0.009	-
SPX <sub>E</sub>	0.884	0.116	-	0.003	-0.000	-	0.007	0.020

Tutte le stime sono state ottenute senza problemi di convergenza, e con diverse inizializzazioni si sono sempre ottenute le stesse stime.

Utilizzando queste stime si è poi calcolato il VaR. Il VaR è dato dal quantile della distribuzione empirica ottenuta simulando 200000 numeri aleatori dalle misture (2) e (3), con i valori dei parametri ottenuti nella tabella 1. La tabella 2 presenta i risultati. In parentesi è riportata la percentuale di eccedenze, cioè di osservazioni che, nella serie disponibile, eccedono la misura di VaR calcolata. A titolo di confronto sono state riportate anche le misure di VaR ottenute tramite l' Extreme Value Theory: per ciascun titolo, i parametri della Generalized Pareto Distribution sono stati stimati tramite massima verosimiglianza provando diversi valori di soglia. In alcuni casi la soglia utilizzata è più piccola del quantile 5%, per cui non è stato possibile calcolare il VaR al 95%.

<sup>(4)</sup> OL = Olivetti, TI = Telecom Italia, MIB = MIB30, SPX = Standard & Poor's 500.

**Tabella 2 - Valore a Rischio<sup>(5)</sup>**

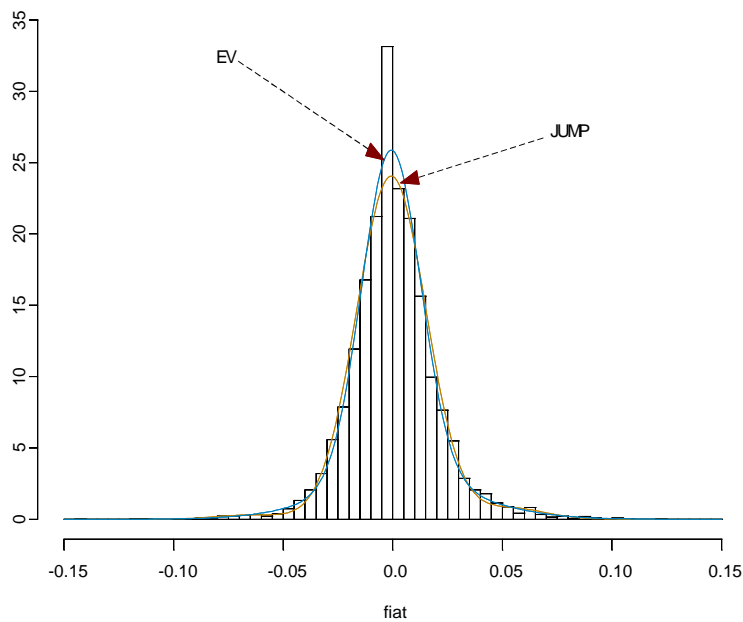
	$VaR_{95\%}$	$VaR_{99\%}$	$VaR_{99,5\%}$
Sun <sub>J</sub>	5.16 (5.06)	9.3 (0.89)	12.62 (0.4)
Sun <sub>E</sub>	5.13 (5.09)	9.52 (0.84)	12.02 (0.5)
Sun <sub>N</sub>	5.66 (3.95)	8.05 (1.62)	8.9 (1.27)
Sun <sub>EVT</sub>	–	9.35 (0.94)	11.6 (0.52)
IBM <sub>J</sub>	2.43 (4.9)	3.62 (1.63)	4.34 (0.85)
IBM <sub>E</sub>	2.37 (5.35)	4.59 (0.72)	5.75 (0.35)
IBM <sub>N</sub>	2.73 (3.56)	3.88 (1.3)	4.29 (0.89)
IBM <sub>EVT</sub>	–	4.26 (0.89)	5.31 (0.4)
Fiat <sub>J</sub>	2.84 (5.51)	5.11 (0.94)	6.73 (0.41)
Fiat <sub>E</sub>	2.88 (5.24)	5.33 (0.89)	6.41 (0.57)
Fiat <sub>N</sub>	3.23 (3.89)	4.58 (1.3)	5.06 (0.96)
Fiat <sub>EVT</sub>	–	5.04 (0.96)	6.18 (0.57)
OL <sub>J</sub>	3.62 (5.55)	7.14 (0.71)	8.66 (0.43)
OL <sub>E</sub>	3.83 (4.86)	6.96 (0.87)	8.1 (0.48)
OL <sub>N</sub>	4.04 (4.43)	5.71 (1.55)	6.31 (1.23)
OL <sub>EVT</sub>	–	6.74 (1)	8.25 (0.48)
TI <sub>J</sub>	2.98 (5.23)	5.25 (1.07)	7.02 (0.39)
TI <sub>E</sub>	2.98 (5.23)	5.51 (0.82)	6.55 (0.43)
TI <sub>N</sub>	3.24 (4.25)	4.61 (1.48)	5.11 (1.16)
TI <sub>EVT</sub>	–	5.29 (1.05)	6.5 (0.43)
MIB <sub>J</sub>	2.29 (5.63)	4.02 (1.05)	4.96 (0.39)
MIB <sub>E</sub>	2.33 (5.37)	4.04 (1.05)	4.87 (0.44)
MIB <sub>N</sub>	2.42 (4.89)	3.45 (1.48)	3.82 (1.09)
MIB <sub>EVT</sub>	–	3.98 (1.05)	4.71 (0.57)
SPC <sub>J</sub>	1.44 (5.28)	2.07 (1.74)	2.32 (1.16)
SPC <sub>E</sub>	1.39 (5.71)	2.84 (0.57)	2.47 (0.25)
SPC <sub>N</sub>	1.57 (4.16)	2.24 (1.4)	2.47 (0.95)
SPC <sub>EVT</sub>	–	2.45 (0.96)	3.03 (0.43)

La performance delle due metodologie basate sulle misture è molto simile e paragonabile alla performance dell'EVT. Vale la pena di notare come il modello  $J$  sia piuttosto impreciso nella stima del VaR per l'indice Standard & Poor's; in termini intuitivi è molto probabile che, trattandosi di un indice che contiene molti titoli, i salti dei vari titoli si compensino dando origine ad una distribuzione priva di salti; la discreta performance del VaR normale in questo caso fa pensare che la distribuzione non sia in effetti troppo distante dalla

<sup>(5)</sup> OL = Olivetti, TI = Telecom Italia, MIB = MIB30.

normale. Questo ragionamento può essere formalizzato tramite il teorema del limite centrale nella versione di Lindeberg-Feller, che richiede solo l'ipotesi di indipendenza <sup>(6)</sup>. Tale ipotesi è giustificata dalla stessa definizione di rischio specifico.

Fig. 2 - Funzione di densità della mistura (Fiat)

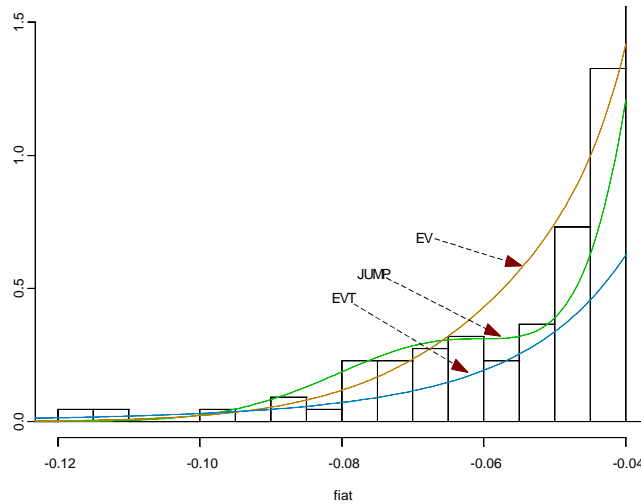


La figura 2 mostra le densità stimate con i due metodi (J ed EV) per il titolo Fiat, assieme all'istogramma dei dati normalizzato in modo che l'area sia uguale ad 1.

---

<sup>(6)</sup> I dettagli, omessi per ragioni di spazio, sono disponibili presso l'autore.

Fig. 3 - Coda della funzione di densità stimata (Fiat)



Più interessante per i nostri scopi è la figura 3, che si focalizza sulla coda delle densità esposte in figura 2, mostrando inoltre la coda della distribuzione stimata tramite EVT, cioè la densità stimata della Generalized Pareto Distribution. Vale la pena di sottolineare che la densità della mistura stimata con il metodo  $J$  presenta una sorta di "gobba" (approssimativamente coincidente con il valore stimato del parametro  $D$ ) che sembra riflettere un salto effettivamente presente nei dati, come evidenziato dall'istogramma.

Sulla base dei risultati esposti è possibile fornire qualche indicazione su quale delle due metodologie basate sulle misture sia preferibile utilizzare. Da un punto di vista computazionale è certamente preferibile, per i motivi già menzionati, l'approccio *Jump-VaR*. Da un punto di vista finanziario tale approccio sembra particolarmente adatto a situazioni in cui i salti verso l'alto e verso il basso siano effettivamente presenti nei dati. Per verificarlo, è certamente raccomandabile rappresentare i dati in un istogramma: gli istogrammi di figura 2 e 3 rappresentano importanti indicazioni a favore dell'uso del modello *Jump-VaR*. Per stimare i parametri di una generica distribuzione a code pesanti è invece preferibile utilizzare il metodo  $E$ .

Tuttavia per rispondere in modo più completo alla questione si è ritenuto opportuno effettuare un esperimento su dati simulati: sono stati simulati

100000 numeri aleatori da una mistura di tipo  $J$  (tre popolazioni normali con varianza uguale) e 100000 numeri aleatori da una mistura di tipo  $E$  (due normali con varianza della seconda popolazione molto maggiore della prima); per generare tali dati si sono utilizzati i valori dei parametri stimati per il titolo Fiat, allo scopo di ottenere dei dati che abbiano le tipiche caratteristiche dei dati finanziari. Infine, allo scopo di mettere alla prova l'algoritmo su una generica distribuzione a code pesanti, sono stati simulati 100000 numeri aleatori da una distribuzione  $t$  di Student con 4 gradi di libertà. Nei tre casi sono stati stimati i due modelli  $J$  ed  $E$ . Con le stime ottenute si sono calcolati il VaR e le eccedenze: i risultati sono esposti in tabella 3.

**Tabella 3 - Valore a Rischio su dati simulati**

Mod. vero	Mod. stimato	95%	99%	99.5%
J	J	2.84 (5.04)	5.09 (1.00)	6.70 (0.49)
J	E	2.9 (4.71)	4.97 (1.05)	6.08 (0.7)
E	J	2.88 (4.97)	5.4 (0.96)	6.4 (0.47)
E	E	2.87 (5.01)	5.23 (1.07)	6.27 (0.52)
$t_4$	J	2.11 (5.20)	3.70 (1.01)	5.03 (0.37)
$t_4$	E	2.09 (5.33)	4.1 (0.73)	5.11 (0.35)

I risultati evidenziano la notevole flessibilità delle distribuzioni a mistura finita, in quanto tutti i VaR calcolati sono estremamente vicini ai loro valori teorici. Per quanto riguarda la scelta del tipo di mistura da utilizzare, si rimanda alle considerazioni del paragrafo precedente.

E' inoltre importante notare che la convergenza dell'algoritmo quando si stimino i parametri di una mistura di tipo EV è sensibilmente più lenta; ciò peraltro può costituire un problema solo per le simulazioni Monte Carlo, mentre nell'applicazione a dati reali l'aumento dei tempi di stima è modesto.

Infine, l'implementazione di test formali per determinare il numero di popolazioni di una mistura è piuttosto problematica [Titterington *et al.* (1985)]. A questa possibilità si preferisce un approccio empirico basato sul fatto che, qualora si tenti di stimare una mistura di  $n$  popolazioni quando i dati sono stati generati da una mistura di  $n - 1$  popolazioni, si ottengono stime identiche per due popolazioni [Aitkin e Tunnicliffe Wilson (1980)].

Sebbene dal punto di vista statistico la metodologia analizzata si possa estendere senza difficoltà al caso multivariato, gli esperimenti effettuati con i

dati a disposizione non sono stati pienamente convincenti. L'interpretazione più plausibile è che, essendo i salti non molto ampi (o, in altre parole, essendo le popolazioni della mistura non ben separate), sia necessaria una numerosità campionaria estremamente grande per stimare correttamente i parametri, il cui numero aumenta considerevolmente nel contesto multivariato. Nel nostro caso sarebbe probabilmente necessario disporre di qualche decennio di dati; tuttavia ciò potrebbe porre il problema, abbastanza comune quando si abbia a che fare con dati finanziari, della non omogeneità temporale della serie: nello specifico, è ragionevole supporre che nella prima parte della serie i salti siano assenti o molto più rari che nell'ultima parte, e dunque la maggior numerosità campionaria potrebbe non essere di grande aiuto per la stima.

Applicare direttamente la metodologia alla serie dei rendimenti di un portafoglio non ha dato risultati convincenti, ed in questo caso ci sono valide ragioni teoriche, le stesse per cui i risultati ottenuti nell'applicazione alla serie dell'indice Standard & Poor's non sono stati brillanti; in proposito si rimanda dunque alla discussione svolta in precedenza. Risultati migliori si ottengono stimando una mistura del tipo  $E$ : si veda Bee (2001) per maggiori dettagli.

## 5 Conclusioni

In questo articolo si è proposta una metodologia per il calcolo del VaR che permette di misurare la componente di rischio specifico, ovvero la componente legata alle caratteristiche dell'emittente dell'attività in questione. Le applicazioni presentate hanno dimostrato che l'approccio proposto riesce a stimare con precisione le code delle distribuzioni dei rendimenti di serie finanziarie, tipicamente piuttosto pesanti. In particolare i risultati sono paragonabili a quelli ottenuti tramite l'impiego dell'Extreme Value Theory, attualmente ritenuta il metodo più affidabile per la stima delle code di distribuzioni leptocurtiche. Oltre ad essere computazionalmente affidabile, la metodologia analizzata stima i parametri in modo completamente endogeno, fatta salva l'inizializzazione dell'algoritmo, che, come si è visto, non presenta difficoltà particolari. Più problematica risulta essere l'estensione ad un portafoglio; su quest'argomento sarà necessaria ulteriore attività di ricerca.

L'applicazione a dati provenienti dai mercati azionari proposta in questo articolo è forse la più immediata ma si ritiene che altri tipi di dati finanziari si prestino ad essere modellati in questo modo. In proposito un'analisi relativa agli *spread* di credito ottenuti dalle curve di rendimento di obbligazioni



*corporate* costituirà oggetto di un prossimo lavoro.

## Appendice: l'algoritmo EM per una mistura normale

L'introduzione dell'algoritmo EM ad opera di Dempster *et al.* (1977) ha risolto in modo soddisfacente il problema della stima di massima verosimiglianza dei parametri di una distribuzione a mistura finita [Titterington *et al.* (1985)]. Si presentano qui le equazioni dell'algoritmo nel caso in cui la mistura sia del tipo  $J$ ; si rinvia a McLachlan e Krishnan (1996) per maggiori dettagli.

Dalla (2) si ha che la funzione di verosimiglianza è data da

$$l(\boldsymbol{\theta}; x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^3 \left( \frac{p_j}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \frac{(x_i - \mu_j)^2}{\sigma^2}} \right),$$

dove  $\mu_1 = \mu - D$ ,  $\mu_2 = \mu$ ,  $\mu_3 = \mu + U$ . Alla base dell'algoritmo sta l'idea che, se fossimo a conoscenza della popolazione di provenienza di ciascuna osservazione, la stima di massima verosimiglianza sarebbe elementare: si potrebbero infatti stimare i parametri di ciascuna popolazione (in questo caso distribuita come una normale) con le sole osservazioni generate da quella popolazione. Quest'idea viene applicata costruendo una funzione di verosimiglianza, detta funzione di verosimiglianza completa, basata, oltre che sui dati effettivamente osservati, anche su dati (ipotetici) che indicano la popolazione di appartenenza di ciascuna osservazione:

$$\begin{aligned} l_c(\boldsymbol{\theta}; \mathbf{z}, \mathbf{x}) &= \sum_{i=1}^n \sum_{j=1}^3 z_{ij} \log\{p_j f_j(x_i)\} \\ &= \sum_{i=1}^n \sum_{j=1}^3 z_{ij} \log p_j + \sum_{i=1}^n \sum_{j=1}^3 z_{ij} \log f_j(x_i), \end{aligned} \quad (4)$$

dove  $z_{ij}$  è uguale ad uno se la  $i$ -esima osservazione appartiene alla  $j$ -esima popolazione e zero altrimenti. Poiché tale funzione non è completamente determinata, in quanto dipende da osservazioni ignote  $(z_1, \dots, z_n)$ , se ne calcola il valore atteso condizionato ai dati osservati, ottenendo così il passo  $E$  dell'algoritmo:

$$\tau_j(x_i; \boldsymbol{\theta}) \equiv \tau_{ij} = \frac{p_j f_j(x_i; \boldsymbol{\theta})}{f(x_i; \boldsymbol{\theta})}. \quad (5)$$

Il passo  $M$  è poi dato dai consueti stimatori di massima verosimiglianza della normale, in cui la variabile ignota  $z_{ij}$  è sostituita da  $\tau_{ij}$ , che ha l'interpretazione

di probabilità a posteriori di appartenenza della  $i$ -esima osservazione alla  $j$ -esima popolazione:

$$\begin{aligned} p_i &= \frac{1}{n} \sum_{j=1}^n \tau_{ij}; \\ \mu_i &= \frac{1}{np_i} \sum_{i=1}^n \tau_{ij} x_i = \frac{\sum_{i=1}^n \tau_{ij} x_i}{\sum_{i=1}^n \tau_{ij}}; \\ \sigma^2 &= \frac{1}{n} \sum_{i=1}^n \tau_{ij} (x_i - \mu_i)(x_i - \mu_i)'. \end{aligned} \tag{6}$$

Le equazioni (5) e (6) sono dunque le equazioni fondamentali per programmare l'algoritmo e vanno iterate fino a convergenza; come si può intuire, l'implementazione risulta piuttosto agevole.

## Bibliografia

- Aitkin, M. e Tunnicliffe Wilson, G.** (1980), "Mixture Models, Outliers, and the EM Algorithm", *Technometrics*, 22, 325-331.
- Bee, M.** (2001), "Mixture Models for VaR and Stress Testing", *Tech Report nr. 12*, Alea, Centro di Ricerca sui Rischi Finanziari, Università di Trento, giugno.
- Bollerslev, T.** (1986), "Generalized Conditional Autoregressive Heteroskedasticity", *Journal of Econometrics*, 31, 307-327.
- Danielsson, J., L. C.G. de Vries** (1997), "Value-at-Risk and Extreme Returns", *working paper*, London School of Economics.
- Danielsson, J., L. de Haan, L. Peng e C.G. de Vries** (2001), "Using a Bootstrap Method to Choose the Sample Fraction in Tail Index Estimation", *Journal of Multivariate Analysis*, 76, 226-248.
- Dempster, A.P., Laird, N.M. e Rubin, D.B.** (1977), "Maximum Likelihood from Incomplete Data via the EM Algorithm (with discussion)", *Journal of the Royal Statistical Society B*, 39, 1-38.
- Embrechts, P., C. Kluppelberg e T. Mikosch** (1997), *Modelling Extremal Events for Insurance and Finance*, New York, Springer.
- Engle, R.F.** (1982), "Auto-regressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, 50, 987-1007.
- Flury, B.** (1997), *A First Course in Multivariate Statistics*, New York, Springer.
- Gibson, M.S.** (2001), "Incorporating Event Risk into Value-at-Risk", *working paper*, Federal Reserve Board.
- Hamilton, J.D.** (1991), "A Quasi-Bayesian Approach to Estimating Parameters for Mixtures of Normal Distributions", *Journal of Business & Economic Statistics*, 9, 27-39.
- Longin, F.M.** (2001), "Beyond the VaR", *working paper*.
- McLachlan, G.J. e Krishnan, T.** (1996), *The EM Algorithm and Extensions*, New York, Wiley.
- Riskmetrics - Technical Document**, 3rd edition, J.P. Morgan, 1995.
- Return to RiskMetrics: The Evolution of a Standard**, J.P. Morgan, 2001.

**Titterington, D.M., Smith, A.F.M. e Makov, U.E.** (1985), *Statistical Analysis of Finite Mixture Distributions*, New York, Wiley.

**Venkataraman, S.** (1997), "Value at Risk for a Mixture of Normal Distributions: the Use of Quasi-Bayesian Estimation Techniques", *Economic Perspectives*, 2-13.

**Zangari, P.** (1996), "An Improved Methodology for measuring VaR", *Risk-Metrics Monitor*, Second Quarter 1996, 7-25.

## Collana ALEA Tech Reports

- Nr.1 F. Sguera, *Valutazione e copertura delle opzioni binarie e a barriera*, Marzo 1999.
- Nr.2 A. Beber, *Introduzione all'analisi tecnica*, Marzo 1999.
- Nr.3 A. Beber, *Il dibattito su dignità ed efficacia dell'analisi tecnica nell'economia finanziaria*, Marzo 1999.
- Nr. 4 L. Erzegovesi, *Capire la volatilità con il modello binomiale*, Luglio 1999.
- Nr. 5 G. Degasperì, *La dinamica delle crisi finanziarie: i modelli di Minsky e Kindleberger*, Agosto 1999
- Nr. 6 L. Erzegovesi, *Rischio e incertezza in finanza: classificazione e logiche di gestione*, Settembre 1999
- Nr. 7 G. Degasperì, L. Erzegovesi, *I mercati finanziari come sistemi complessi: il modello di Vaga*, Settembre 1999.
- Nr.8 A.Beber e L.Erzegovesi, *Distribuzioni di probabilità implicite nei prezzi delle opzioni*, dicembre 1999.
- Nr.9 M.Filagrana, *Le obbligazioni strutturate nel mercato italiano: principali tipologie e problematiche di valutazione e di rischio*, marzo 2000.
- Nr.10 A.Beber, *Determinants of the implied volatility function on the Italian Stock Market*, marzo 2001.
- Nr.11 F.Bazzana, *I modelli interni per la valutazione del rischio di mercato secondo l'approccio del Value at Risk*, giugno 2001.
- Nr.12 M.Bee, *Mixture models for VaR and stress testing*, giugno 2001.
- Nr.13 M.Bee, *Un modello per l'incorporazione del rischio specifico nel VaR*, gennaio 2002.