

Introduzione al Value-at-Risk

Flavio Angelini, Stefano Herzel

9 giugno 2010

Indice

1	Preliminari	5
1.1	Piano del corso	5
1.2	Fattori di rischio e P&L	6
1.3	Il Value-at-Risk	10
1.3.1	Un esempio di calcolo del VaR	12
1.3.2	Alcune proprietà del VaR	14
1.4	L'Expected Shortfall	16
2	Analisi statistica dei fattori di rischio	18
2.1	Statistiche descrittive	18
2.2	Eventi estremi	20
2.3	Analisi dinamica di una serie storica	31
3	Modelli econometrici	34
3.1	Modelli ARMA	34
3.2	Modelli ARCH-GARCH	35
3.3	Variazioni sul tema GARCH	37
3.4	La volatilità implicita	38
4	Il modello RiskMetricsTM	40
4.1	Fattori di rischio	40
4.2	Stime della volatilità per titoli azionari e obbligazionari	41
4.2.1	Titoli azionari	42
4.2.2	Mapping di titoli obbligazionari	42
5	Il VaR di un portafoglio	44
5.1	Il caso univariato	44
5.2	Metodi esatti ed approssimati	45

5.3	Simulazioni Monte Carlo	47
5.4	Simulazione storica	48
5.5	Il caso multivariato	48
6	Backtesting	52
7	Stress testing	55
7.1	Lo Stress testing storico	55
7.2	Lo Stress testing parametrico	56
8	Teoria dei valori estremi	58
8.1	Modelli per i massimi	58
8.2	Modelli per le code	59
9	Altri modelli per la stima del rischio di mercato	61
9.1	Il modello di Frey, McNeil	61
9.2	Il modello di Barone-Adesi, Giannopoulos, Vosper	63
10	Complementi	66
10.1	Simulazione di una variabile aleatoria	66
10.2	Misure di rischio coerenti	67

Elenco delle figure

1.1	Densità della P&L e misure di rischio.	7
1.2	Il percentile al 5%.	11
2.1	Tasso di cambio Peso-Dollaro (1997-2002)	22
2.2	densità empirica dei rendimenti del MIB30 (1997-2002)	24
2.3	Simulazioni di percentili estremi	27
2.4	QQ-plot di dati simulati)	29
2.5	QQ-plot dei rendimenti dell'S&P500 (1997-2002)	30
2.6	Rendimenti del MIB30 (1997-2002)	31
2.7	Autocorrelazioni dei rendimenti del MIB30 (1997-2002)	33

Prefazione

Queste note sono state originate da alcuni corsi che abbiamo tenuto sulla misurazione del rischio di mercato e del Value-at-Risk presso il *Master in finanza per la banca e per l'assicurazione* dell'Università di Roma La Sapienza e presso la Scuola di S.Anna. L'obiettivo dei corsi era quello di fornire gli strumenti operativi per costruire un sistema di misurazione del rischio, offrendo una introduzione all'uso del Value at Risk e dello standard di riferimento del mercato RiskMetricsTM ed alcuni cenni su possibili alternative. Lo spirito dei corsi, che speriamo si ritrovi anche in queste note, è stato quello del fare e formare in finanza, puntando a condurre lo studente attraverso un percorso critico in cui costruire le principali tecniche operative. Un caloroso ringraziamento va ai professori Massimo De Felice e Franco Moriconi per averci invitato a tenere tali corsi.

Perugia, maggio 2003

Flavio Angelini, Stefano Herzel

Capitolo 1

Preliminari

1.1 Piano del corso

Lo scopo del corso è valutare il rischio che il valore di un dato portafoglio possa diminuire a causa di mutamenti imprevisti di alcuni fattori di mercato. Solitamente, il cambiamento di valore di un portafoglio viene definito tramite la Profit and Loss (P&L). A determinare i possibili valori della P&L intervengono i fattori di rischio. Per costruire un sistema di misurazione del rischio è necessario modellare correttamente i fattori di rischio e studiare la conseguente distribuzione di P&L. Una volta definito un modello probabilistico per la P&L si possono definire vari indicatori con lo scopo di valutare la rischiosità della posizione. In questo corso ci concentremo su quelle misure che attualmente sono più utilizzate nello standard operativo: il Value-at-Risk (VaR) e la perdita attesa oltre il Var (Expected Shortfall).

L'adozione di un particolare modello non può prescindere da una preliminare osservazione delle caratteristiche dei fattori di rischio. Mostriamo i principali strumenti atti ad individuare tali caratteristiche assieme ad alcuni esempi (cap. 2.1). Studieremo quindi alcuni modelli statistici, cercando di evidenziare se e come essi siano in grado di rappresentare correttamente gli aspetti più rilevanti dei fattori di rischio (cap. 3). Un capitolo a parte (cap. 4) verrà dedicato al modello RiskMetricsTM in quanto esso rappresenta attualmente lo standard di riferimento.

Una volta definito un modello per i fattori di rischio passeremo al problema della valutazione di una misura di rischio per un portafoglio finanziario (cap. 5). Per effettuare tale valutazione è necessario utilizzare le funzioni di

pricing che permettono di passare dai fattori ai prezzi e ricorrere a varie metodi, sia di simulazione che di approssimazione. La correttezza della misura prodotta alla fine di questo processo viene poi valutata attraverso tecniche di backtesting (cap. 6).

Un sistema per il controllo del rischio può essere utilizzato anche per analizzare il comportamento del nostro portafoglio rispetto a scenari particolarmente sfavorevoli, in quello che viene definito Stress testing (cap. 7).

Qualsiasi sistema per la valutazione del rischio (e per lo Stress testing) dovrebbe essere in grado di modellare correttamente quegli eventi che si verificano raramente ma che hanno un potenziale di notevole pericolosità per il valore del portafoglio. Per trattare dal punto di vista probabilistico tali eventi è stata elaborata la Teoria dei valori estremi (cap. 8). Una delle principali critiche che vengono poste al modello RiskMetricsTM è che, a causa dell'ipotesi semplificatrice di innovazioni normali, non modella correttamente gli eventi estremi: illustreremo due modelli che cercano di rispondere a tale critica (cap. 9).

1.2 Fattori di rischio e P&L

Sia V_t il valore in t di un portafoglio. Il suo valore V_T a una data futura T è aleatorio; la variabile aleatoria $\Delta V = V_T - V_t$ rappresenta il cambiamento di valore della posizione e si chiama *Profit and Loss* (P&L).

Il nostro obiettivo è quello di illustrare i sistemi più utilizzati per la stima della distribuzione di probabilità P&L.

La distribuzione P&L viene utilizzata per il calcolo di diverse misure di rischio, quali il VaR, l'expected shortfall, la deviazione standard ...

La Figura 1.2 mostra un esempio di densità per un P&L con alcune misure di rischiosità.

Alla base di ogni sistema di valutazione del rischio ci sono i *fattori di rischio*, che sono i determinanti fondamentali del valore di un portafoglio. Tali fattori possono essere equity, tassi di cambio, commodity e tassi d'interesse.

- Equity: portafogli azionari, contratti forward su azioni, opzioni e titoli derivati su azioni e su indici azionari.
- Tassi di cambio: posizioni in valuta estera, contratti forward e opzioni su tassi di cambio, cross currency swap.

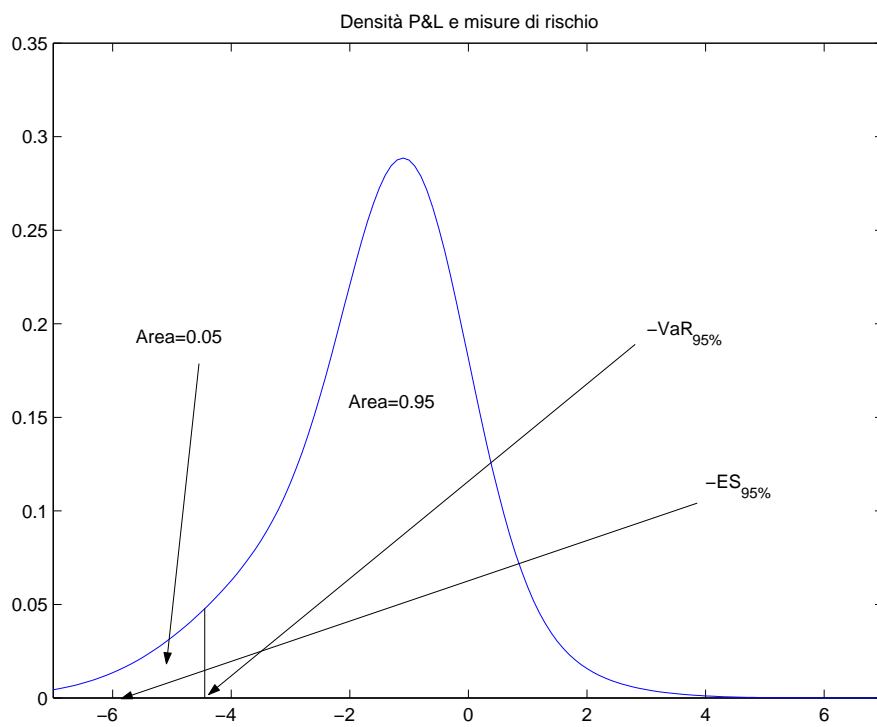


Figura 1.1: Densità della P&L e misure di rischio.

- Commodity: futures su commodities (grano, petrolio), opzioni su commodities, opzioni su futures su commodity.
- Tassi d'interesse: posizioni su titoli obbligazionari, swap (IRS), opzioni su bond, cap e floor, swaptions.
- Altri fattori: volatilità implicita, credit spread. In generale ogni parametro presente in una formula di pricing.

Sia P_t il valore (mark to market) di un fattore di rischio in t . Definiamo:

- tasso di rendimento

$$I_{t,T} := \frac{P_T - P_t}{P_t}$$

- rendimento logaritmico (log-return)

$$R_{t,T} := \log\left(\frac{P_T}{P_t}\right) = \log(1 + I_{t,T})$$

Osservazione: Sviluppando la funzione $\log(1+x)$ intorno al punto $x=0$ si ottiene:

$$\log(1+x) = x - x^2 + o(x^3) \quad x \rightarrow 0$$

quindi $R_{t,T}$ e $I_{t,T}$ sono molto vicini se $I_{t,T}$ è piccolo (in particolare, $R_{t,T} < I_{t,T}$).

Implementazione 1.2.1 *Nel workspace dati.mat è presente una matrice P che contiene otto serie storiche di prezzi di titoli finanziari, dove le righe sono osservazioni giornaliere, e le colonne rappresentano i diversi titoli. Per il calcolo delle serie storiche dei tassi di rendimento e dei rendimenti logaritmici si può procedere con i seguenti comandi*

```
>>I=diff(P,1,1)./P(1:end-1,:);
```

```
>>R=diff(log(P),1,1);
```

Si può verificare che le due grandezze sono generalmente molto vicine e che il rendimento logaritmico è sempre minore del corrispondente tasso di rendimento. Si provi, ad esempio, i comandi

```
>>i=1;
```

```
>>plot(I(:,i)-R(:,i))
```

Ora osserviamo come si compongono i rendimenti nel tempo: siano $t \leq s \leq T$. Si ha:

$$I_{t,T} = (1 + I_{t,s})(1 + I_{s,T}) - 1$$

$$R_{t,T} = R_{t,s} + R_{s,T}$$

Per un un portafoglio composto da m titoli con pesi w_i (con $\sum_{i=1}^m w_i = 1$) si ha

$$I_{t,T} = \sum_{i=1}^m w_i I_{t,T}^i$$

$$R_{t,T} = \log \sum_{i=1}^m w_i e^{R_{t,T}^i}$$

Quindi i rendimenti logaritmici sono più vantaggiosi per calcolare il rendimento in un intervallo di tempo più lungo, meno per calcolare il rendimento di un portafoglio. Tuttavia, dato che, come abbiamo osservato, rendimenti logaritmici e tassi di rendimento sono solitamente molto simili tra loro, si utilizza spesso la proprietà di linearità anche per la composizione di un portafoglio. Ad esempio, RiskMetrics, che utilizza i rendimenti logaritmici, li compone linearmente per ottenere il rendimento di un portafoglio.

Per determinare la P&L della posizione si procede in due passi:

1. generazione degli scenari futuri dei fattori di rischio
2. valutazione della P&L corrispondente applicando opportuni modelli di pricing

Per generare gli scenari dei fattori di rischio ci sono due modi:

1. parametrico: si specifica la probabilità che il fattore assuma un certo valore. Bisogna fare ipotesi sulla distribuzione.
2. non parametrico: si guarda al comportamento passato e si assume che si comporterà in maniera simile nel futuro.

1.3 Il Value-at-Risk

Una volta definita una distribuzione per la P&L si può procedere a valutare la rischiosità della posizione. Una delle misure di rischio più usate è il Value-at-Risk (VaR), che si definisce come la massima perdita che si può subire in un certo orizzonte temporale (di solito un giorno o due settimane lavorativi) con un livello di probabilità p assegnata. Può essere interpretato come il capitale che occorre accantonare per essere coperti da ogni possibile perdita con un dato livello di confidenza.

In altri termini, il VaR è un percentile della distribuzione P&L (con il segno cambiato). Se p è un livello di confidenza e $v_{(1-p)}$ l' $(1-p)$ -esimo percentile della distribuzione di probabilità di ΔV ,

$$Pr(\Delta V \leq v_{(1-p)}) = 1 - p.$$

Il VaR al livello p

$$\text{VaR}_p = -v_{(1-p)}.$$

Come livello di confidenza in genere sono raccomandati i valori 0.95 o 0.99.

Il VaR come definito sopra non fornisce un'indicazione immediata della rischiosità della posizione in quanto non è posto in relazione con il valore corrente di questa. A tal scopo si può utilizzare un VaR percentuale espresso come rapporto tra VaR e valore della posizione

$$\text{VaR}\% = \frac{\text{VaR}}{V_t}.$$

Sia X una variabile aleatoria con funzione di ripartizione $\Phi(x) := Pr(X \leq x)$.

Dato $0 \leq \alpha \leq 1$, il percentile x_α è definito da

$$Pr(X \leq x_\alpha) = \alpha.$$

Ovvero

$$x_\alpha = \Phi^{-1}(\alpha)$$

Se Z è una normale standard i suoi percentili sono mostrati nella Tabella 1.3.

Se $X \sim N(\mu, \sigma^2)$, $X \sim \mu + \sigma Z$, si ha:

$$x_\alpha = \mu + \sigma z_\alpha$$

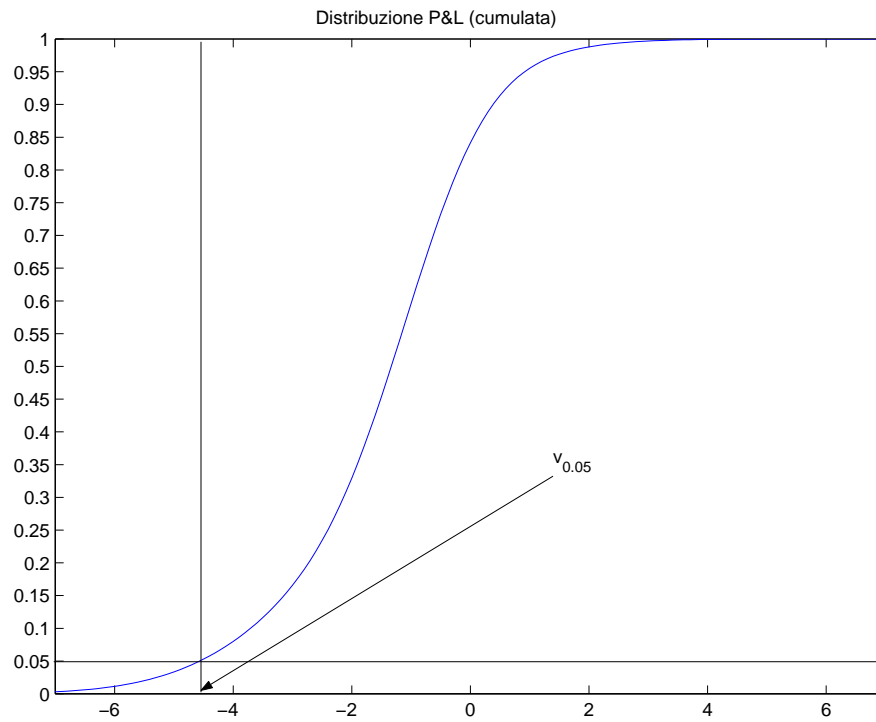


Figura 1.2: Il percentile al 5%.

α	0.05	0.025	0.01	0.95	0.975	0.99
z_α	-1.6449	-1.96	-2.3263	1.6449	1.96	2.3263

Tabella 1.1: Percentili della normale standard.

Se $\mathbf{X} \sim N(\mu, \Sigma)$ normale multivariata e δ è un vettore colonna numerico, consideriamo $X = \delta' \mathbf{X}$. Si ha $X \sim N(\delta' \mu, \delta' \Sigma \delta)$. Dunque

$$x_\alpha = \delta' \mu + \sqrt{\delta' \Sigma \delta} z_\alpha$$

Se $Y = F(X)$ e F è monotona crescente, si ha:

$$y_\alpha = F(x_\alpha).$$

Infatti

$$Pr(Y \leq F(x)) = Pr(X \leq x).$$

Se F è decrescente

$$y_\alpha = F(x_{1-\alpha}).$$

$$\begin{aligned} Pr(Y \leq F(x)) &= Pr(X \geq x) \\ &= 1 - Pr(X \leq x) \end{aligned}$$

Implementazione 1.3.1 *Il toolbox di statistica di MATLAB contiene funzioni che calcolano i percentili di distribuzioni note. Ad esempio, i percentili della normale si trovano con la funzione norminv.*

Si controlli la tabella 1.3 con i seguenti comandi

```
>>alpha=[0.05,0.025,0.01,0.95,0.975,0.99];
>>norminv(alpha,0,1)
```

1.3.1 Un esempio di calcolo del VaR

In questo primo semplice esempio mostreremo come calcolare il VaR di un portafoglio sotto due ipotesi differenti sulla distribuzione della P%L. Supponiamo di avere nel portafoglio $\Delta = 50$ mila azioni ABC che hanno valore oggi di 10 Euro ciascuna. Il valore della posizione è lineare nel fattore di rischio e il suo valore attuale è $V_t = 500000$ Euro. Vogliamo calcolare il VaR giornaliero al 95% e al 99%. Tratteremo due casi classici: tasso di rendimento o rendimento logaritmico normale.

Caso 1: Tasso di rendimento normale

Sia

$$I_{t,t+1} \sim \mu + \sigma Z,$$

con Z normale standard, $\sigma > 0$. Sia $\mu = 0, \sigma^2 = 0.0001$. In questo caso, μ e σ sono espressi su base giornaliera dato che rappresentano media e deviazione standard del rendimento giornaliero. Si ha

$$\Delta V = \Delta \cdot P_t \cdot I_{t,t+1} = V_t \cdot I_{t,T}$$

e dunque

$$\text{VaR}_p = -V_t \cdot (\mu + \sigma z_{(1-p)})$$

dove $z_{(1-p)}$ è il percentile della normale standard.

Quindi

$$\text{VaR}_{95\%} = -500000 \cdot (0.01 \cdot (-1.65)) = 8224$$

$$\text{VaR}_{99\%} = -500000 \cdot (0.01 \cdot (-2.32)) = 11631$$

Si noti che si può anche calcolare il VaR per un orizzonte temporale più lungo $T > t + 1$ supponendo che il rendimento $I_{t,T}$ sia una normale $(\mu(T), \sigma(T)^2)$, dove si deve pensare che $\mu(T)$ e $\sigma(T)$ siano su base $(T - t)$ giorni.

Caso 2: Log-return normale

Supponiamo che i log-return giornalieri della ABC siano normali iid:

$$R_{t,t+1} \sim N(\mu, \sigma^2),$$

con μ e σ espressi su base giornaliera. Sia $\mu = 0, \sigma^2 = 0.0001$. Abbiamo

$$R_{t,T} = \sum_{j=t+1}^T R_{j-1,j} \sim N((T-t)\mu, (T-t)\sigma^2).$$

In questo caso la distribuzione del rendimento su un orizzonte temporale più lungo si determina facilmente da quella del rendimento giornaliero ed è ancora una normale con parametri ottenuti in modo semplice dai parametri giornalieri. Si ha $\Delta V = V_t(e^{R_{t,T}} - 1)$ e dunque

$$\text{VaR}_p = -V_t \cdot (e^{(T-t)\mu + \sqrt{T-t}\sigma z_{(1-p)}} - 1)$$

I valori del VaR sono:

$$\text{VaR}(1g)_{95\%} = 8157$$

$$\text{VaR}(1g)_{99\%} = 11497$$

In questo caso possiamo calcolare anche il VaR a 10 giorni;

$$\text{VaR}(10g)_{95\%} = 25343$$

$$\text{VaR}(10g)_{99\%} = 35462$$

Implementazione 1.3.2 *Costruiamo delle funzioni che implementino i calcoli sopra descritti. La prima deve prendere come input il numero Δ di quote possedute del titolo, il prezzo corrente P di ciascuna, i parametri μ e σ^2 , un livello di probabilità p e restituire in output il VaR al livello p e, eventualmente, il valore attuale della posizione. La seconda, `varlin2`, prenderà gli stessi input e restituirà gli stessi output, ma si potrà introdurre l'orizzonte temporale T espresso in giorni. Si confronti la prima funzione con la funzione `vardeltanorm` che calcola il VaR per un portafoglio composto da più strumenti e dipendente da più fattori con il metodo dell'approssimazione lineare (si vedano i Paragrafi 5.2 e 5.5). Con i comandi*

```
>>v=varlin2(50000,10,0,0.0001,[0.95,0.99])
```

```
>>v=varlin2(50000,10,0,0.0001,[0.95,0.99],10)
```

si verifichino i calcoli del Caso 2 e si proceda analogamente per verificare i calcoli del Caso 1 con la funzione costruita.

1.3.2 Alcune proprietà del VaR

Il VaR_p è una funzione a valori reali definita sullo spazio delle variabili aleatorie.

E' facile vedere, utilizzando la definizione di percentile, che soddisfa le seguenti proprietà.

1. $X \geq 0 \Rightarrow \text{VaR}_p(X) \leq 0$;
 $X \geq Y \Rightarrow \text{VaR}_p(X) \leq \text{VaR}_p(Y)$;
2. $\lambda \geq 0, \text{VaR}_p(\lambda X) = \lambda \text{VaR}_p(X)$;
3. $k \in R, \text{VaR}_p(X + k) = \text{VaR}_p(X) - k$;

Tali proprietà sono facilmente interpretabili dal punto di vista economico ricordando che il VaR rappresenta il capitale da accantonare come margine di sicurezza.

In generale il VaR non verifica la relazione

$$\text{VaR}_p(X + Y) \leq \text{VaR}_p(X) + \text{VaR}_p(Y).$$

Tale proprietà sarebbe desiderabile per un sistema di risk management: per controllare il rischio totale assunto è sufficiente imporre dei limiti sui rischi presi da ciascun settore. Una misura di rischio che soddisfa le tre proprietà

precedenti insieme alla subadditività si dice coerente (Artzner et al. [1998]). Nel Capitolo 10 riportiamo alcuni approfondimenti sulle misure coerenti.

Ora mostreremo che il VaR non è una misura coerente, cioè che non sempre verifica

$$\text{VaR}_p(X + Y) \leq \text{VaR}_p(X) + \text{VaR}_p(Y).$$

Una banca fa due prestiti di 50 ciascuno a due clienti. I clienti possono restituire la cifra per intero o fare default e non restituire nulla. Supponiamo che gli eventi default siano indipendenti, ciascuno con probabilità pari a 0.008.

Siano X ed Y le variabili aleatorie che rappresentano la P&L della banca verso ciascun cliente, cioè $X = 0$ se il primo cliente paga il debito, $X = -50$ se fa default. Si ha:

- $X = 0$ e $Y = 0$ con probabilità $1 - 0.008$;
- $X = -50$ e $Y = -50$ con probabilità 0.008 .

Ne segue che, se per esempio $p = 99\%$,

$$\text{VaR}_p(X) = \text{VaR}_p(Y) = 0.$$

Consideriamo ora la variabile $X + Y$. La diversificazione del portafoglio dovrebbe portare ad una minore rischiosità. Si ha:

- $X + Y = 0$ con probabilità $(1 - 0.008)^2$;
- $X + Y = -100$ con probabilità $(0.008)^2$;
- $X + Y = -50$ con probabilità $2 \times 0.008 \times (1 - 0.008) \approx 1.6\%$.

Per cui,

$$\text{VaR}_p(X + Y) = 50.$$

Dunque

$$\text{VaR}_p(X + Y) > \text{VaR}_p(X) + \text{VaR}_p(Y) \quad (50 > 0 + 0).$$

Quindi in questo caso la proprietà di subadditività non è verificata; tuttavia la relazione di subadditività può essere verificata su particolari distribuzioni. Il caso classico è quando $\begin{pmatrix} X \\ Y \end{pmatrix}$ è una normale bivariata. Allora

$$\text{VaR}_p(X + Y) \leq \text{VaR}_p(X) + \text{VaR}_p(Y)$$

Infatti

1. $X + Y$ è normale (in questo caso)
2. $\text{VaR}_p(X) = -\sigma_X z_{1-p} = \sigma_X z_p$ (dove, per semplicità, si è supposto media nulla)
- 3.

$$\begin{aligned}\sigma_{X+Y} &= \sqrt{\sigma_X^2 + 2\rho\sigma_X\sigma_Y + \sigma_Y^2} \\ &\leq \sigma_X + \sigma_Y\end{aligned}$$

Concludiamo con una osservazione: abbiamo appena utilizzato il fatto che la somma di due marginali di una distribuzione normale multivariata è normale. Spesso si sente dire: la somma di due normali è normale! questo in generale è falso, come dimostra il seguente esempio: sia

$$Z \sim N(0, 1)$$

$$Y = Z\mathbf{1}_{|Z|<1} - Z\mathbf{1}_{|Z|>1}$$

È facile vedere che

$$P(Y < x) = P(Z < x)$$

quindi $Y \sim N(0, 1)$. Ma $Y + Z = 2Z\mathbf{1}_{|Z|<1}$ che ovviamente non è normale.

1.4 L'Expected Shortfall

L'Expected Shortfall o valore atteso oltre il VaR, è stato introdotto recentemente in finanza. È una misura di rischio che descrive la media delle perdite che eccedono il VaR. Fornisce informazioni sulle code della distribuzione P&L.

$$\text{ES}_p := -\mathbf{E}[\Delta V \mid \Delta V \leq -\text{VaR}_p]$$

L'Expected Shortfall è una misura di rischio coerente (si veda il §(10.2)).

Ne mostreremo ora un esempio di calcolo nel caso semplice della distribuzione normale. Se $Z \sim N(0, 1)$, $n(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$ è la funzione di densità. Si ha:

$$\begin{aligned}\text{ES}_p &:= -\mathbf{E}[Z \mid Z \leq -\text{VaR}_p(Z)] \\ &= -\mathbf{E}[Z \mid Z \leq z_{1-p}]\end{aligned}$$

$$\begin{aligned}
&= -\frac{\int_{-\infty}^{z_{1-p}} xn(x)dx}{Pr(Z \leq z_{1-p})} \\
&= -\frac{[n(x)]_{-\infty}^{z_{1-p}}}{1-p} \\
&= \frac{n(z_{1-p})}{1-p}
\end{aligned}$$

Se $X \sim N(\mu, \sigma^2)$, $X \sim \mu + \sigma Z$,

$$\begin{aligned}
ES_p &= -\mathbf{E}[X \mid X \leq -\text{VaR}_p(X)] \\
&= -\mathbf{E}[\mu + \sigma Z \mid \mu + \sigma Z \leq \mu + \sigma z_{1-p}] \\
&= -\mu \mathbf{E}[1 \mid Z \leq z_{1-p}] - \sigma \mathbf{E}[Z \mid Z \leq z_{1-p}] \\
&= -\mu + \frac{\sigma}{1-p} n(z_{1-p})
\end{aligned}$$

Per ulteriori approfondimenti sull' expected shortfall e sulle misure coerenti rimandiamo al Capitolo 10.

Capitolo 2

Analisi statistica dei fattori di rischio

Nel capitolo precedente sono stati identificati alcuni strumenti finanziari come titoli azionari, prezzi di zero coupon bond, tassi di cambio come fattori di rischio alla base della costruzione di un sistema di valutazione della distribuzione P&L di un portafoglio. Per costruire un modello di valutazione del rischio che sia basato su solide basi statistiche, appare quindi naturale e necessario uno studio delle caratteristiche tipiche delle serie finanziarie. Tale studio viene effettuato sia con un'analisi statica, utilizzando gli strumenti classici della statistica descrittiva, sia con un'analisi dinamica, tenendo conto del succedersi temporale delle serie storiche. Nell'ambito della misurazione del rischio di mercato, una particolare attenzione deve essere dedicata allo studio delle code della distribuzione, cioè alla possibilità o meno che si verifichino eventi estremi.

2.1 Statistiche descrittive

Consideriamo una serie storica di n rendimenti logaritmici giornalieri

$$R_i := R_{i-1,i}, \quad i = t - n + 1, \dots, t$$

dove t è la data di oggi. Una prima analisi delle caratteristiche principali delle serie può essere effettuata con le seguenti statistiche:

- La media

$$\hat{\mu} := \frac{\sum_{i=t-n+1}^t R_i}{n}.$$

La media è un indicatore della tendenza della serie.

- La deviazione standard

$$\hat{\sigma}_t = \sqrt{\frac{\sum_{i=t-n+1}^t (R_i - \hat{\mu})^2}{n-1}}.$$

La deviazione standard è un primo indicatore della variabilità, e quindi della rischiosità, della serie. Per volatilità della serie si intende la versione annualizzata della deviazione standard $\sqrt{N}\sigma_t$, dove N è il numero di giorni in un anno (regola della radice). La volatilità su base annua viene dunque interpretata come la deviazione standard del rendimento annuale. Tale interpretazione è giustificata dal punto di vista teorico solo se si fanno determinate ipotesi sul processo in esame. Ciò accade, ad esempio, nel modello proposto da RiskMetrics, come mostreremo nel seguito.

- L'asimmetria

$$\frac{\sum_{i=t-n+1}^t \left(\frac{R_i - \hat{\mu}}{\hat{\sigma}}\right)^3}{n}.$$

L'asimmetria indica lo scostamento della serie standardizzata dalla media. La distribuzione normale ha asimmetria nulla.

- La curtosi

$$\frac{\sum_{i=t-n+1}^t \left(\frac{R_i - \hat{\mu}}{\hat{\sigma}}\right)^4}{n}.$$

La curtosi è una prima misura della grossezza della coda; indica la presenza di movimenti estremi. La curtosi della distribuzione normale è tre. Si dice che una variabile aleatoria ha le code grosse, quando valori estremi appaiono con frequenza maggiore di quanto previsto dalla distribuzione normale. Dunque una curtosi significativamente più grande di tre indica code più grosse della distribuzione normale.

Attraverso l'analisi statica delle serie finanziarie si riscontrano con una certa frequenza una serie di fatti:

- Nella maggior parte delle osservazioni la media risulta trascurabile;
- L'asimmetria è spesso significativa (soprattutto nel mercato azionario, per il quale è generalmente negativa);

- La curtosi è spesso maggiore di tre, primo indizio del fatto che l'ipotesi di normalità è raramente giustificata .

Con tali statistiche descrittive, si ottiene un'indicazione sulle caratteristiche delle serie storiche. Due esempi interessanti sono i seguenti (Duffie, Pan [1997]):

1. I rendimenti giornalieri dell' S&P500 dal 1986 al 1996 hanno volatilità annua 15.9%, asimmetria -4.8 e curtosi 111.
2. I rendimenti giornalieri del cambio Dollaro-Peso dal 1986 al 1996 hanno volatilità annua 18.9%, asimmetria -8.1 e curtosi 217.5.

Implementazione 2.1.1 *Data la matrice di prezzi P del workspace `dati.mat`, come visto nell'Implementazione 1.2.1, si può calcolare la corrispondente matrice dei rendimenti logaritmici R . Utilizzando apposite funzioni MATLAB si possono calcolare le statistiche descrittive. Ad esempio, il comando*

```
>>devstd=std(R,0,1)
```

restituisce un vettore `devstd` composto dalle deviazioni standard di ciascuna delle otto serie. I comandi

```
>>r=R(:,8);
```

```
>>[mean(r),std(r),skewness(r),kurtosis(r)]
```

restituiscono le statistiche descrittive dei rendimenti logaritmici dell'indice S&P500 dal 1997 al 2002.

2.2 Eventi estremi

Abbiamo detto che la curtosi è indice di code grosse, cioè di possibilità non trascurabile di eventi estremi. Un evento estremo si può definire come un evento che dista dalla media un certo numero di deviazioni standard. Un sistema di misurazione del rischio di mercato si deve occupare in modo particolare di tali eventi. Nel caso di una normale (μ, σ) la probabilità di avere un valore minore di $\mu - 5\sigma$ è $2.86 \cdot 10^{-7}$ (cioè si dovrebbe verificare due o tre volte su dieci milioni). Dalla serie storica dei rendimenti logaritmici dell' S&P500 si riscontra che un tale evento si è verificato cinque volte in dieci anni tra il 1986 e il 1996 (Duffie, Pan [1997]), ovvero cinque volte su circa duemilaseicento. Questo è un forte indizio del fatto che l'ipotesi di normalità non viene soddisfatta da questa serie storica. La Figura 2.1 mostra una serie

piuttosto estrema, quella dei rendimenti logaritmici del tasso di cambio Peso Messicano contro Dollaro. Le linee orizzontali corrispondono ai valori $\mu - 5\sigma$ e $\mu + 5\sigma$, dove μ e σ sono rispettivamente la media e la deviazione standard della serie. Le volte che una di tale bande viene superata dovrebbe essere confrontata con il valore $2.86 \cdot 10^{-7}$.

Implementazione 2.2.1 *Per determinare la probabilità che una variabile normale (μ, σ) sia minore di $\mu - c\sigma$, si può utilizzare*

```
>>c=5;
```

```
>>normcdf(-c)
```

Costruiamo una funzione `plotss` che produce grafici di serie storiche, con la possibilità di inserire le bande di confidenza sopra descritte. La funzione prende come input una matrice di serie storiche, un argomento flag con il quale richiedere le bande (ad esempio $w = 1$) e un argomento c (opzionale con default pari a 5) nel quale si determina l'ampiezza delle bande. Come output restituisce i grafici delle singole serie storiche, inserendo come titolo della figura il numero della serie. Nel caso gli argomenti w e c non siano specificati, la funzione restituisce i grafici delle serie senza bande. Si provino i seguenti comandi

```
>>plotss(R)
```

```
>>plotss(R,1,3)
```

```
>>plotss(R(:,4),1)
```

L'ultimo comando restituisce la Figura 2.1.

Formalizziamo il concetto di code grosse. Si dice che una variabile aleatoria X , ha la coda sinistra più grossa di una variabile aleatoria Y se esiste u tale che

$$P(X \leq x) > P(Y \leq x) \quad \forall x < u.$$

Analogamente si dice che X ha la coda destra più grossa di Y se esiste u tale che

$$P(X \geq x) > P(Y \geq x) \quad \forall x > u.$$

Comunemente si dice che una variabile aleatoria ha le code grosse se ha le code, sia sinistra che destra, più grosse della normale. La grossezza di una coda sinistra (destra) di una distribuzione si misura con l'ordine di grandezza con cui la sua funzione di distribuzione tende a zero per $x \rightarrow -\infty$ (la sua funzione complementare tende a zero per $x \rightarrow \infty$).

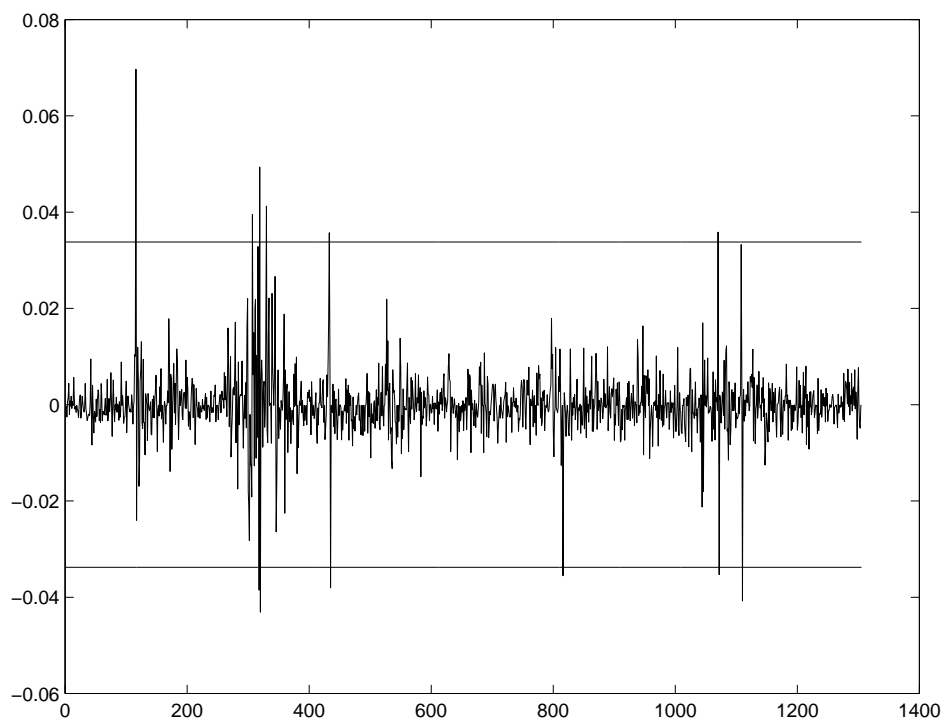


Figura 2.1: Rendimenti logartimici del tasso di cambio Peso Messicano contro dollaro (1997-2002). E' evidenziata la bande corrispondente a cinque deviazioni standard. Se la distribuzione fosse normale dovrebbe uscire con probabilità $2 \cdot 10^{-7}$.

Concentriamoci sulle code sinistre e supponiamo $X \leq 0$. Dato che

$$EX^n = \int_{-\infty}^0 x^n dF(x),$$

se $F(x) \sim x^{-\alpha}$ per $x \rightarrow -\infty$, i momenti di ordine maggiori o uguali a α sono infiniti. Quindi la grossezza delle code può misurarsi, in modo equivalente, con il numero dei momenti finiti. La normale ha momenti finiti per ogni n .

La tabella che segue riporta alcune distribuzioni di probabilità. Tra queste osserviamo che la $\chi^2(n, \mu)$ e la log-normale hanno momenti di ogni ordine m ; la T di Student(v) per ogni $m \leq v - 1$, la Pareto(k, α) per ogni $m < \alpha$.

variabile	densità ($x > 0$)
$\chi^2(n, \mu)$	$\frac{x^{n/2-1} e^{-x/2}}{2^{n/2} \Gamma(n/2)}$
log-normale(μ, σ)	$\frac{1}{\sqrt{2\pi}\sigma x} e^{-(\log(x)-\mu)^2/(2\sigma^2)}$
Student(v)	$\frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi}\Gamma(v/2)} (1+x^2)^{-\frac{v+1}{2}}$
Pareto(k, α)	$\frac{-\alpha k^\alpha}{(k+x)^{\alpha+1}}$

Torniamo ora all'analisi statica delle serie storiche. Data la serie R_i , $i = t - n + 1, \dots, t$, si può definire la funzione di distribuzione empirica, o cumulata,

$$F_n(x) = \frac{\sum_{i=1}^n \mathbf{1}_{R_i \leq x}}{n}$$

dove $\mathbf{1}_A$ è la funzione indicatrice dell'evento A .

Sotto opportune ipotesi, la funzione di distribuzione empirica tende quasi ovunque alla funzione di distribuzione teorica (Legge dei grandi numeri). Quindi l'utilizzo della distribuzione empirica può essere giustificato dal punto di vista teorico se la serie dei dati è abbastanza lunga e se si fa un'ipotesi di indipendenza. Si può costruire un'istogramma che fornisce una rappresentazione della funzione di densità. Si osserva che la distribuzione empirica dei rendimenti azionari è spesso leptocurtica (alta e magra in mezzo e grassa sulle code), come nella Figura 2.2.

Implementazione 2.2.2 *La funzione `distemp` restituisce il grafico della distribuzione empirica di una serie storica. Restituisce inoltre il grafico della densità empirica e lo confronta con quello di una normale con stessa media e varianza della serie. Il comando*

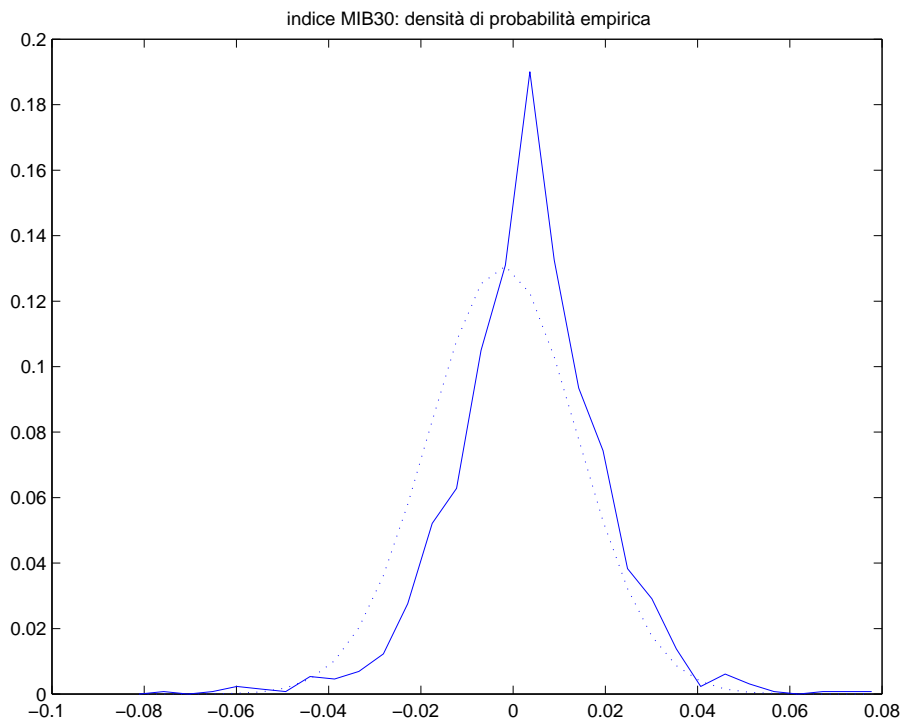


Figura 2.2: Densità empirica dei log-rendimenti giornalieri del MIB30 dal 16-5-97 al 16-5-02 e densità della normale con la stessa media e varianza. Osservare la caratteristica distribuzione leptocurtica.

```
>>distemp(R(:,7),30)
```

restituisce la Figura 2.2, insieme al grafico della funzione di distribuzione empirica.

Una volta definita la distribuzione empirica, possiamo calcolare il suo percentile, detto percentile empirico. Per determinare il percentile empirico ordiniamo la serie R_i , $i = t - n + 1, \dots, t$:

$$R_{1,n} \leq \dots \leq R_{n,n}.$$

$R_{i,n}$ si dice statistica d'ordine i -esima. Si ha

$$F_n(R_{i,n}) = \frac{i}{n}.$$

Dunque $R_{i,n}$ è una possibile scelta per il percentile empirico al livello di probabilità $\frac{i}{n}$. La scelta non è univoca dato che per ogni $x \in [R_{i,n}, R_{i+1,n})$ si ha che $F_n(x) = \frac{i}{n}$. $R_{i,n}$ rappresenta una scelta pessimistica, $R_{i+1,n}$ ottimistica. Si può optare per una scelta intermedia, come il valore medio tra i due. In generale, assegnato un livello di probabilità p , non esiste sempre i tale che $p = \frac{i}{n}$ e dunque il percentile empirico non è determinato. In tal caso si prende come i la parte intera di np . Le possibili scelte per il percentile ricadono nell'intervallo $[R_{i,n}, R_{i+1,n})$. Il percentile empirico può essere usato come stima del percentile teorico r_α , dato che

$$R_{i,n} \rightarrow r_\alpha \quad q.o.$$

quando $\frac{i}{n} \rightarrow \alpha$ (Legge dei grandi numeri).

Implementazione 2.2.3 *Costruiamo una funzione percentile che prende come input una serie di dati e un livello di probabilità e come output restituisce il percentile empirico. Si può utilizzare tale funzione per calcolare il percentile empirico al 5% dei rendimenti del MIB30 con il seguente comando*

```
>>p=percentile(R(:,7),0.05)
```

Tali risultati possono venire confrontati con quelli forniti dalla funzione `prctile` di MATLAB.

Le stime di percentili estremi (come quelli al livello 0.05 e 0.01) fatte attraverso i percentili empirici sono piuttosto inaffidabili, soprattutto se la distribuzione ha le code grosse. Questo fatto è stato osservato da Kupiec [1995] che riporta delle stime sulle varianze dello stimatore. Noi possiamo verificare tale affermazione attraverso le simulazioni. Osserviamo che per simulare un percentile empirico di una variabile aleatoria X con distribuzione F non è necessario simulare dati dalla X , ma è sufficiente estrarre il percentile empirico \hat{u} di una distribuzione uniforme su $(0, 1)$ e calcolare $F^{-1}(\hat{u})$ (perchè?).

Il test è il seguente: si estrae il percentile empirico da un campione simulato di n osservazioni da una distribuzione teorica. Si ripete tale procedura per n test volte. Abbiamo così ottenuto una simulazione di numerosità n della distribuzione del percentile empirico, di cui si può calcolare media e varianza. La media, all'aumentare del campione, si avvicina al percentile teorico. Si può inoltre confrontare la varianza ottenuta con la formula contenuta in Kupiec [1995]. La Figura 2.3 riporta i risultati ottenuti nel caso di percentili al 1% di tre distribuzioni T di Student con gradi di libertà due, sei e venti,

con $n_{sim}=3000$ e $n_{test}=1000$. E' evidente come i percentili estratti dalla distribuzione con le code più grosse (quella con due gradi di libertà) sono più sparsi degli altri. Questo dimostra che quando si vuole stimare un percentile empirico da una distribuzione con le code grosse bisogna considerare i dati con cautela perchè lo stimatore ha una notevole variabilità.

Implementazione 2.2.4 *Costruiamo una funzione `kupiectest` che produca il test sopra descritto. La funzione deve prendere come input la distribuzione di probabilità (per la scelta della quale si può vedere l'help della funzione MATLAB `icdf`), il livello di probabilità, la numerosità del campione da cui estrarre il percentile empirico, il numero dei percentili empirici estratti e i parametri della distribuzione scelta. Come output deve restituire il campione simulato dei percentili empirici e il percentile teorico della distribuzione. Lo script `kupiecfg` produce la Figura 2.3 e restituisce media e varianza delle tre simulazioni. Per esercizio si possono costruire script analoghi che confrontino il comportamento di diverse distribuzioni.*

Uno strumento grafico che permette di confrontare i percentili di due variabili aleatorie è il QQ-plot. Serve per determinare la pesantezza delle code di una distribuzione rispetto a un'altra. In genere si confrontano i percentili empirici con i percentili di una distribuzione teorica di riferimento Φ (ad esempio la normale). Il QQ-plot è il grafico dei punti $(R_{i,n}, \Phi^{-1}(\frac{i}{n}))$, ovvero $(F_n^{-1}(\frac{i}{n}), \Phi^{-1}(\frac{i}{n}))$. Se i dati fossero estrazioni dalla distribuzione Φ di riferimento, il grafico apparirebbe lineare, come mostrato nella Figura 2.4. Si notino gli scostamenti per i percentili più estremi, in accordo con quanto osservato da Kupiec. Per interpretare il QQ-plot si procede come segue: supponiamo che il punto corrispondente al percentile i/n sia sopra la diagonale, cioè $R_{i,n} < \Phi^{-1}(\frac{i}{n})$, e che tali valori siano negativi. Indichiamo con R la distribuzione empirica (i dati) e con Z la distribuzione con cui ci si vuole confrontare. Allora:

$$Pr(R \leq R_{i,n}) = Pr(Z \leq \Phi^{-1}(\frac{i}{n})) \geq Pr(Z \leq R_{i,n}).$$

Quindi, se da un certo punto in poi per valori negativi, il grafico si situa sopra la diagonale, la distribuzione empirica di R ha coda sinistra più grossa della distribuzione di riferimento Z . Analogamente si ragiona per quanto riguarda la coda destra: nel caso in questione, con $R_{i,n} < \Phi^{-1}(\frac{i}{n})$, si avrebbe che la distribuzione empirica ha la coda destra più fine della distribuzione di

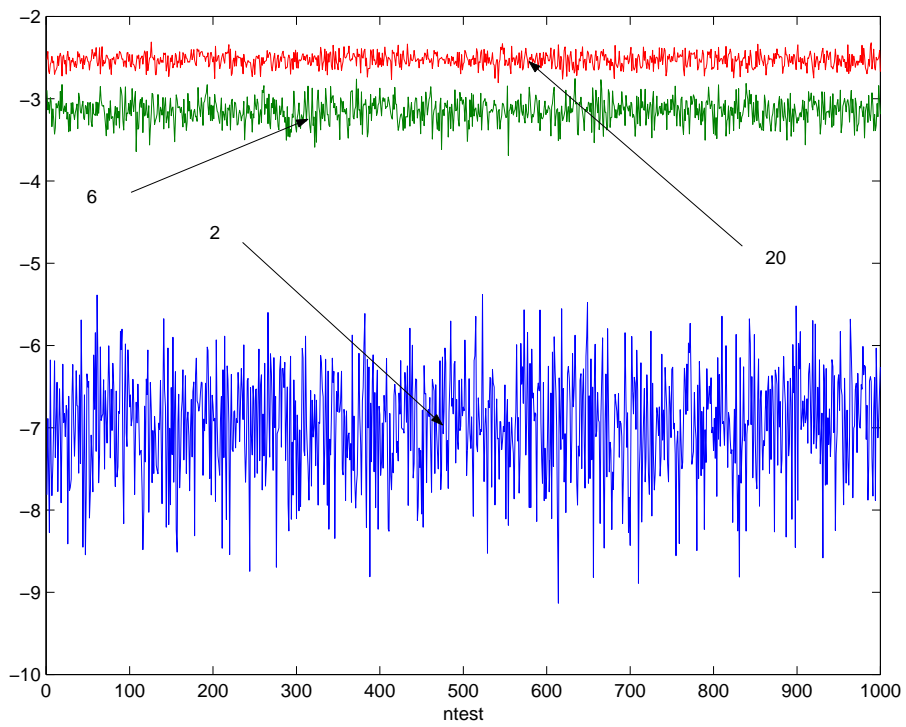


Figura 2.3: Mille simulazioni di percentili empirici al 1% generati da tre distribuzioni T di Student con due, sei e venti gradi di libertà. La varianza aumenta con il diminuire dei gradi di libertà, cioè con l'ingrossarsi della coda della distribuzione.

riferimento (verificare). Nei dati finanziari, soprattutto se confrontati con la distribuzione normale, si riscontra usualmente la tipica forma a S, come si osserva nella Figura 2.5. Ciò vuol dire che le serie finanziarie hanno code sia destre che sinistre più grosse della distribuzione normale.

Implementazione 2.2.5 *Costruiamo la funzione `qqplot` per produrre il QQ-plot dei dati contro una distribuzione da assegnare (si può utilizzare la funzione MATLAB `icdf`). MATLAB ha la funzione `qqplot`, ma non confronta dati con una distribuzione teorica, bensì con altri dati. Sia \mathbf{r} la serie storica dei rendimenti logaritmici di un titolo, ad esempio i rendimenti del S&P500 dal 1997 al 2002. La Figura 2.5 si produce con i seguenti comandi*

```
>>r=R(:,8);
```

```
>>qqplot(r,'norm',0,1)
```

Si provi inoltre a confrontare estrazioni da una distribuzione data con la distribuzione stessa (ad esempio la normale standard) con i seguenti comandi, che riproducono grafici analoghi alla Figura 2.4

```
>>r=normrnd(0,1,1000,1);
```

```
>>qqplot(r,'norm',0,1)
```

Per esercizio si provino comandi analoghi utilizzando la T di Student con v gradi di libertà e si noti che, all'aumentare della pesantezza della coda, il grafico appare più sparso nei percentili estremi, in accordo con il test della Figura 2.3.

Infine, si può calcolare l'Expected Shortfall empirico, ovvero l'Expected Shortfall della distribuzione empirica. Si ha:

$$ES_{\frac{i}{n}} = -\frac{\sum_{k=1}^n R_k \mathbf{1}_{R_k \leq R_{i,n}}}{i}.$$

Più in generale, dato un livello di probabilità α ,

$$ES_{\alpha} = -\frac{\sum_{k=1}^n R_k \mathbf{1}_{R_k \leq x_{\alpha}}}{\sum_{k=1}^n \mathbf{1}_{R_k \leq x_{\alpha}}},$$

dove x_{α} è il corrispondente percentile empirico della serie R_i .

Implementazione 2.2.6 *Costruiamo una funzione `es` che prenda come input una serie di dati e un livello di probabilità e come output restituisca l'Expected Shortfall empirico. Tale funzione utilizzerà la funzione `percentile` costruita nell'implementazione (2.2.3).*

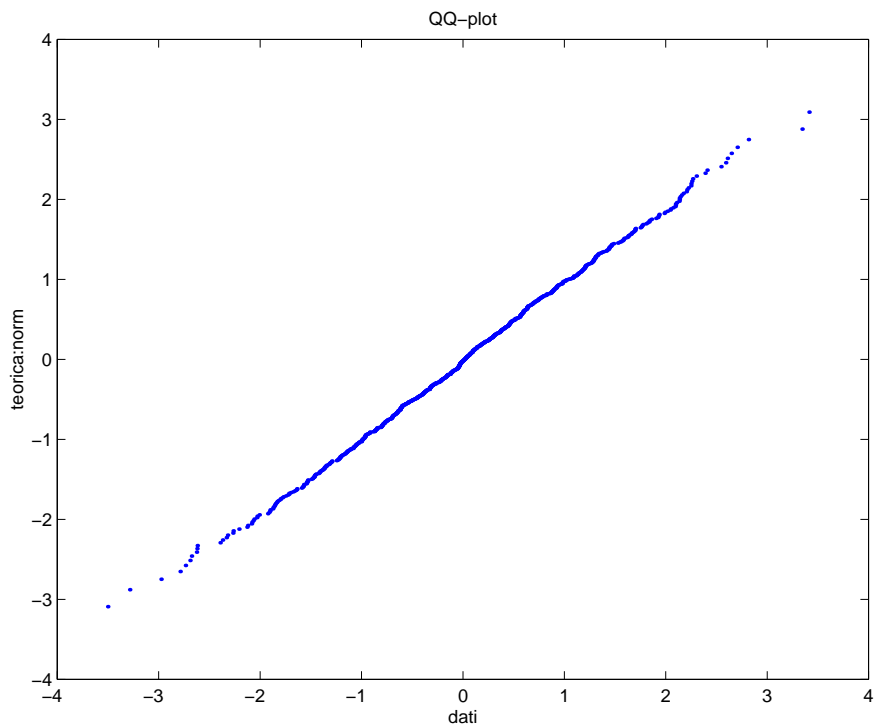


Figura 2.4: QQ-plot di estrazioni dalla normale standard contro la stessa distribuzione. Il grafico è lineare, con scostamenti nei percentili estremi.

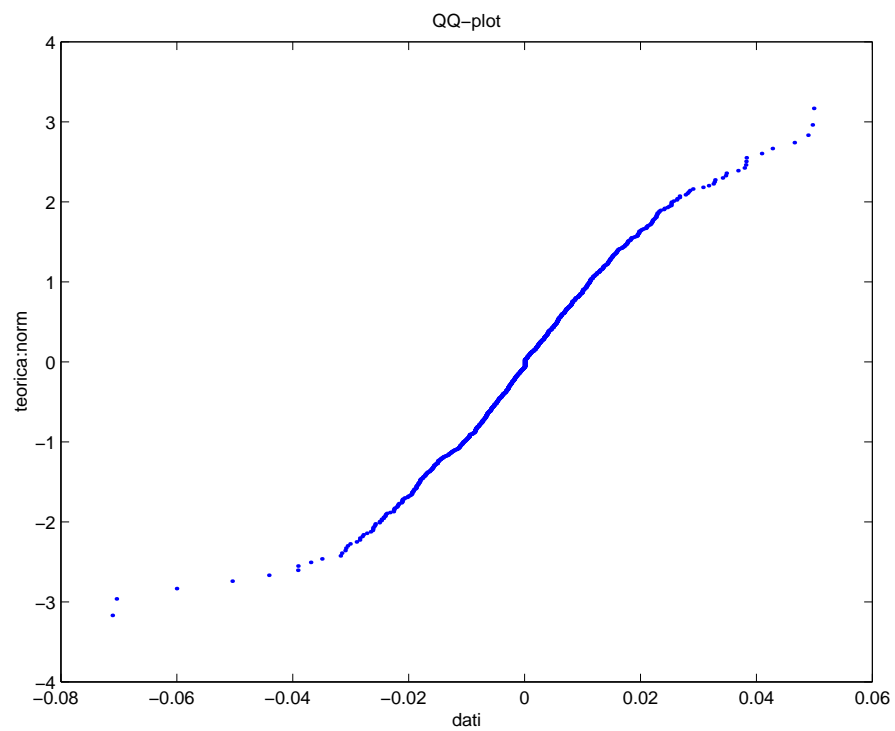


Figura 2.5: QQ-plot dei log-rendimenti dell'indice S&P500 (1997-2002) rispetto alla normale standard. Si noti la tipica forma ad S che mostra come le code della distribuzione empirica siano più pesanti rispetto a quelle della normale.

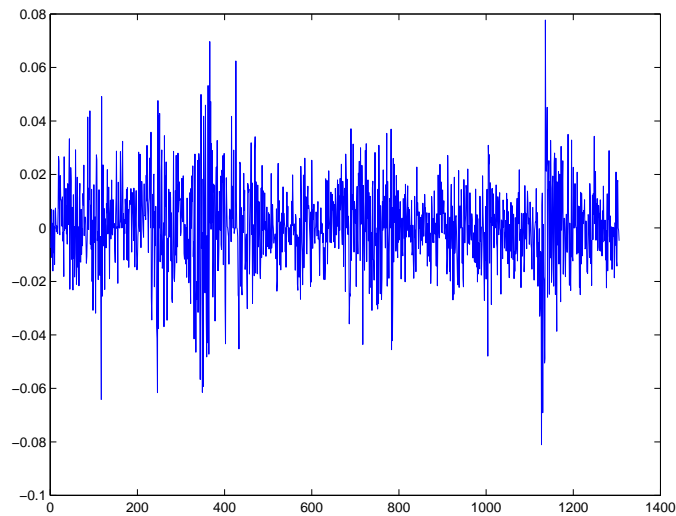


Figura 2.6: Rendimenti logaritmici del MIB30 dal 1997 al 2002.

2.3 Analisi dinamica di una serie storica

Dal punto di vista della dinamica della serie storica, per quanto riguarda le serie del mercato azionario, si riscontrano solitamente i seguenti fatti (che infatti vengono spesso chiamati fatti tipici). I primi si colgono osservando direttamente la serie storica.

- La volatilità cambia nel tempo.
- I valori estremi appaiono insieme (effetto clustering).

Come esempio si veda la Figura 2.6 che mostra la serie dei rendimenti logaritmici del MIB30 tra il 1997 ed il 2002.

Altri fatti tipici dal punto di vista dinamico si colgono attraverso la funzione di autocorrelazione. Essi sono:

- i rendimenti hanno bassa correlazione.
- la correlazione dei quadrati e dei valori assoluti è sempre significativa e positiva.

Tali fatti sono illustrati nella Figura 2.7 dove sono riportate le autocorrelazioni dei rendimenti del MIB30 e dei loro quadrati. Le bande di confidenza presenti nella figura indicano l'intervallo dentro al quale la correlazione ottenuta può essere considerata nulla. Si noti che l'analisi dinamica mostra che le serie finanziarie presentano fenomeni di dipendenza temporale.

Il fatto che le autocorrelazioni dei quadrati dei rendimenti siano significative anche per lag temporali molto lunghi è stato interpretato da molti studiosi come un'evidenza di un effetto di memoria lunga delle serie finanziarie, cioè eventi che si sono verificati anche molto tempo fa continuano ad influenzare l'andamento dei prezzi (più precisamente la loro volatilità). Tale interpretazione è però giustificata solo sotto ipotesi di stazionarietà della serie storica considerata. Infatti Mikosch e Starica [2000] hanno mostrato che serie non autocorrelate, ma non stazionarie, possono produrre una autocorrelazione dei quadrati significativa e positiva, come quella osservata sui mercati finanziari. Dato che l'ipotesi di stazionarietà ci sembra difficilmente difendibile quando si tratta di serie economiche, siamo propensi a scartare la spiegazione di tali grafici come un effetto dovuto alla memoria lunga.

Implementazione 2.3.1 *Si può effettuare l'analisi dinamica appena descritta. Per visualizzare il grafico di una serie si può utilizzare la funzione di MATLAB `plot`. Ad esempio la Figura 2.6 si ottiene con i seguenti comandi*

```
>>r=R(:,7);
```

```
>>plot(r)
```

Alternativamente si può utilizzare la funzione `plotss` costruita nell'Implementazione 2.2.1, la quale può essere conveniente nel caso si abbia una matrice di serie storiche. Inoltre si può utilizzare la funzione `acf` del toolbox VARIM. La Figura 2.7 è riprodotta dai primi tre dei seguenti comandi

```
>>r=R(:,7);
```

```
>>acf(r,50)
```

```
>>acf(r.^2,50)
```

```
>>acf(abs(r),50)
```

che restituiscono le autocorrelazioni rispettivamente della serie, dei suoi quadrati e dei suoi valori assoluti, fino a un lag temporale di cinquanta giorni.

Nel capitolo successivo verranno descritti alcuni modelli econometrici adatti ad interpretare i fatti tipici sopra descritti.

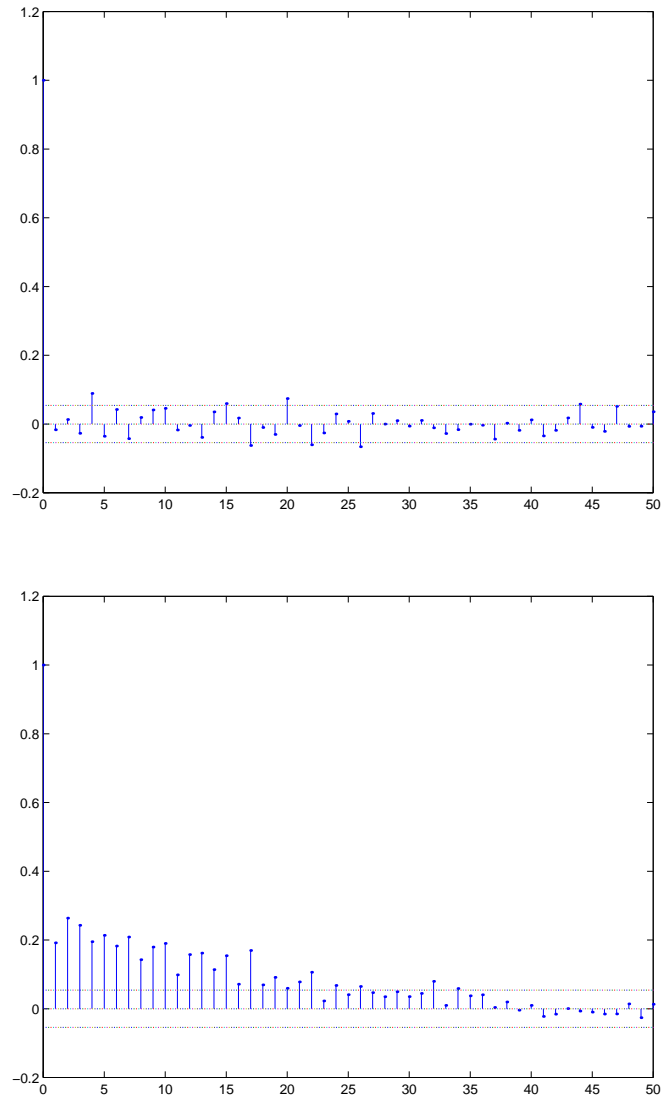


Figura 2.7: Autocorrelazioni della serie storica dei rendimenti del MIB30 (in alto) e dei suoi quadrati (in basso).

Capitolo 3

Modelli econometrici

L'analisi statistica del capitolo precedente ha evidenziato alcune caratteristiche empiriche delle serie finanziarie. Ora mostreremo alcuni approcci proposti per la modellizzazione statistica delle serie storiche. Dapprima definiremo i modelli ARMA, mostrando che non sono in grado di spiegare il tipico comportamento della funzione di autocorrelazione evidenziato precedentemente. I modelli ARCH e G-ARCH sono invece consistenti con tale caratteristica.

3.1 Modelli ARMA

In questa sede tratteremo solo un particolare tipo di processo Auto Regressive Moving Average (ARMA), rimandando il lettore interessato a Taylor [????]. Un processo ARMA(1,1) è definito da

$$R_{t+1} - \mu = a(R_t - \mu) + \varepsilon_{t+1} + b\varepsilon_t$$

Si dimostra che è stazionario se $|a| < 1$.

Se anche $|b| < 1$ può rappresentarsi come un AR(∞):

$$R_{t+1} - \mu = (a + b) \sum_{i=0}^{\infty} (-b)^{i-1} (R_{t-i} - \mu) + \varepsilon_{t+1}$$

quindi il rendimento in $t + 1$ dipende (in modo via via più smorzato) da tutti i rendimenti precedenti.

La varianza incondizionata è

$$E(R - \mu)^2 = \frac{1 + 2ab + b^2}{1 - a^2}$$

La varianza condizionata (o prevista al tempo t) è

$$E_t(R_{t+k} - \mu)^2 = 1 + (a + b)^2 \left(\frac{1 + a^{2k}}{1 - a^2} + (R_t - \mu)^2 \right)$$

La autocorrelazione di R è:

$$\rho_k = \frac{(1 + ab)(a + b)}{1 + 2ab + b^2} a^{k-1}, \quad k \geq 1$$

La autocorrelazione di R^2 è dell'ordine di ρ_k^2 (Taylor, pag.56). Quindi un ARMA ha bassa correlazione tra i quadrati dei rendimenti, contrariamente a ciò che si osserva sui dati reali. Pertanto, come tutti i modelli lineari, non è adatto per modellare serie con varianze correlate.

3.2 Modelli ARCH-GARCH

I modelli ARCH (Autoregressive Conditional Heteroscedastic) furono introdotti da Engle [1982]. L'idea è di modellare la varianza condizionale in modo auto-regressivo.

Il modello ARCH(1) è così definito:

$$R_{t+1} = \mu + \sigma_t \varepsilon_{t+1}$$

$$\sigma_t^2 = \delta + \alpha(R_t - \mu)^2$$

$\delta, \alpha > 0$.

Al tempo t sono noti R_t e σ_t . L'idea principale del modello è di rendere il livello di volatilità attuale dipendente dal rendimento di ieri.

Successivamente (Bollerslev, 1986) viene introdotto il modello (Generalized ARCH) o, più comunemente, GARCH. Il modello GARCH(1,1) è definito come

$$\sigma_t^2 = \delta + \alpha(R_t - \mu)^2 + \beta\sigma_{t-1}^2$$

$\delta, \alpha, \beta > 0$.

In tale modello la volatilità cambia e il suo livello attuale dipende dal rendimento di ieri, ma anche dalla volatilità di ieri. Questo permette di modellare il fatto che ci sono dei periodi di alta volatilità seguiti da periodi di bassa volatilità. (effetto di Volatility clustering).

La condizione di stazionarietà è $\alpha + \beta \leq 1$

Il coefficiente β indica la persistenza degli shock di volatilità, Il coefficiente α indica la reattività della volatilità rispetto ai movimenti di mercato. Il coefficiente δ rappresenta il livello atteso di σ_{t+k} per k grande.

La varianza incondizionata di un GARCH(1,1) è

$$E(R - \mu)^2 = \frac{\delta}{1 - \alpha - \beta}$$

quindi se $\alpha + \beta < 1$ la varianza è finita.

La varianza condizionata (o prevista al tempo t) è

$$\begin{aligned} E_t(R_{t+k} - \mu)^2 &= E_t\sigma_{t+k}^2 \\ &= \delta \left(\frac{1 - (\alpha + \beta)^k}{1 - (\alpha + \beta)} \right) + (\alpha + \beta)^k \sigma_t^2 \end{aligned}$$

Per $k \rightarrow \infty$ si ottiene la varianza incondizionata.

Supponiamo per semplificare le formule (senza perdere di generalità), $\mu = 0$. Osserviamo che

$$\begin{aligned} E_t R_{t+1} R_{t+1+k} &= E_t \sigma_t \epsilon_{t+1} \sigma_{t+k} \epsilon_{t+1+k} \\ &= E_t \sigma_t \epsilon_{t+1} \sigma_{t+k} E_{t+k} \epsilon_{t+1+k} \\ &= 0 \end{aligned}$$

Cioè i rendimenti di un processo GARCH sono incorrelati, come si osserva in gran parte delle serie finanziarie!

Se R_{t+1} è un GARCH(1,1),

$$R_{t+1}^2 = \delta + (\alpha + \beta)R_t^2 + w_{t+1} - \beta w_t$$

dove w_{t+1} è white noise. Quindi R_{t+1}^2 è un ARMA(1,1). Le autocorrelazioni della serie dei rendimenti al quadrato sono

$$\rho_k \approx (\alpha + \beta)^{k-1}, \quad k \geq 1$$

Quindi anche le autocorrelazioni dei quadrati dei rendimenti hanno un comportamento simile a quello osservato per le serie finanziarie. Pertanto possiamo concludere che, dal punto di vista dell'autocorrelazione, e se si mantiene l'ipotesi di fondo di stazionarietà (vedi capitolo precedente), il GARCH è un buon modello per le serie finanziarie.

3.3 Variazioni sul tema GARCH

Tra le numerose proposte citiamo:

1. Asymmetric GARCH (A-GARCH)

$$\sigma_t^2 = \delta + \alpha(R_t - \mu + \gamma)^2 + \beta\sigma_{t-1}^2$$

con $\delta > 0$ e $\alpha, \beta \geq 0$. È caratterizzato dal fatto che, se $\gamma < 0$, i rendimenti negativi hanno una maggiore influenza sulla volatilità di quelli positivi. Per questo motivo risulta appropriato per serie finanziarie, specialmente di tipo azionario. È stato infatti utilizzato da Barone-Adesi et al. [????] e da Frey e McNeil [???] per applicazioni di calcolo di VaR.

2. Integrated GARCH (I-GARCH):

È un GARCH(1,1) in cui $\alpha + \beta = 1$, cioè (ponendo $\beta = \lambda$)

$$\sigma_t^2 = \delta + (1 - \lambda)(R_t - \mu)^2 + \lambda\sigma_{t-1}^2$$

Se $\delta = 0$ e $\mu = 0$ si ottiene il modello RiskMetrics. In questo caso:

$$\sigma_t^2 = (1 - \lambda) \sum_{i=0}^{\infty} \lambda^i R_{t-i}^2$$

Risk Metrics pone $\lambda = 0.94$ per dati giornalieri, $\lambda = 0.97$ per dati mensili. Maggiore λ , maggiore l'importanza che viene data ai rendimenti più lontani nel tempo.

Come casi particolari delle relazioni viste prima otteniamo:

1. Ha le code molto grosse: infatti la varianza incondizionata non esiste (è infinita).
2. L'autocorrelazione dei quadrati è costante per ogni lag temporale: Effetto long memory.
3. Si ha:

$$E_t \sigma_{t+k}^2 = E_t \sigma_{t+k-1}^2$$

cioè la varianza prevista è costante. Da questo segue:

$$E_t\left(\sum_{i=1}^k R_{t+i} - k\mu\right)^2 = k\sigma_t^2$$

cioè la varianza condizionata del rendimento $R_{t,t+k}$ si ottiene moltiplicando per k la varianza giornaliera. Quindi in questo caso la regola della radice per la volatilità è giustificata.

Stimando i parametri di un GARCH su serie finanziarie lunghe si trova spesso che $\alpha + \beta \approx 1$. Questo ha portato diversi analisti a sostenere la presenza di un effetto di long memory. Tuttavia, anche in questo caso, occorre fare attenzione all'ipotesi di stazionarietà, difficilmente sostenibile per periodi lunghi (vedi Mikosch, Starica [2000]).

Ora intendiamo occuparci della distribuzione condizionata di un GARCH(1,1). Supponiamo di avere osservato una serie storica di n rendimenti R_{t-n+1}, \dots, R_t e supponiamo che le innovazioni siano iid e normali. Stimiamo i parametri del modello $\hat{\mu}, \hat{\delta}, \hat{\alpha}, \hat{\beta}$. Calcoliamo il valore di σ_t

$$\sigma_t^2 = \hat{\delta} + \hat{\alpha}(R_t - \hat{\mu})^2 + \hat{\beta}\sigma_{t-1}^2.$$

Quindi, condizionatamente all'informazione in t ,

$$R_{t+1} = \hat{\mu} + \sigma_t \varepsilon_{t+1},$$

è normale. Condizionatamente all'informazione in t (oggi), si ha che

$$R_{t+1} = \hat{\mu} + \sigma_t \varepsilon_{t+1},$$

è normale, ed il suo percentile $1 - p$ è

$$\hat{r}_{1-p} = \hat{\mu} + \sigma_t z_{1-p}$$

con z_{1-p} percentile della normale standard.

Se $k > 1$ la funzione di distribuzione di R_{t+k} è più complicata e non è immediato calcolarne i percentili. Si può utilizzare la simulazione.

3.4 La volatilità implicita

Tutti gli approcci visti in precedenza stimano i parametri del modello, ed in particolare la volatilità, dalla serie storica osservata. In alcuni casi però,

quando si dispone di un mercato sufficientemente liquido di titoli derivati, è possibile utilizzare i prezzi di tali titoli per stimare la volatilità. Si dice allora che si sta facendo una stima su cross section.

Supponiamo che C_t^M sia il prezzo di mercato di un titolo derivato e sia C^{BS} la formula di Black-Scholes

$$C_t^M = C^{BS}(S_t, \sigma, \tau, K, r)$$

invertendo rispetto a σ si trova la volatilità implicita σ_t^I .

La volatilità implicita tiene conto delle valutazioni attuali del mercato, quindi in molti casi ha proprietà previsive migliori rispetto alle stime storiche

Se si ipotizza che la volatilità sia un processo stocastico indipendente da S (sotto le probabilità risk-neutral Q), allora

$$C_t = \mathbf{E}_t^Q [C^{BS}(S_t, \bar{\sigma}_{t,T}, \tau, K, r)]$$

dove

$$\bar{\sigma}_{t,T} = \sqrt{\frac{\sigma_t^2 + \dots + \sigma_{t+T-1}^2}{T-t}}$$

Per le opzioni at the money si ha

$$C^{BS}(\sigma) \approx \alpha \sigma$$

quindi:

$$C_t^M = C^{BS}(\sigma_t^I) \approx \alpha \sigma_t^I$$

e

$$C_t^M = \mathbf{E}_t^Q C^{BS}(\bar{\sigma}_{t,T}) = \mathbf{E}_t^Q \alpha \bar{\sigma}_{t,T}$$

allora:

$$\sigma_t^I = \mathbf{E}_t^Q \bar{\sigma}_{t,T}$$

Questo per le opzioni at the money. Per quanto riguarda le altre opzioni, la convessità della volatilità implicita (al variare dello strike) può essere un'indicazione del livello di incertezza della volatilità.

Concludiamo osservando che, se è vero che la volatilità implicita è una approssimazione del valore atteso della volatilità futura (sotto determinate ipotesi), è però anche vero che tale valore atteso deve essere preso con riferimento alla probabilità risk-neutral. Come sappiamo tale probabilità è una misura che potrebbe non avere nulla a che fare con la probabilità oggettiva. Essa è esclusivamente legata al pricing e va utilizzata con cautela quando si intende misurare altre quantità come il VaR.

Capitolo 4

Il modello RiskMetricsTM

Per la sua importanza come standard di riferimento, trattiamo in particolare il modello Risk-metrics. Cominceremo con le ipotesi fatte sulla distribuzione dei fattori di rischio e vedremo poi alcuni approcci per la stima del VaR di un portafoglio sia nel caso univariato che nel caso multivariato.

4.1 Fattori di rischio

In questo modello si pone

$$R_t = \sigma_t \varepsilon_t,$$

quindi la media giornaliera μ_t viene posta uguale a zero. La volatilità si stima tramite un EWMA (Exponentially Weighted Moving Average)

$$\sigma_t^2 = \frac{(1 - \lambda)}{(1 - \lambda^n)} \sum_{i=0}^{n-1} \lambda^i R_{t-i}^2$$

RiskMetrics pone, per tutte le serie storiche, $\lambda = 0.94$ (dati giornalieri) o $\lambda = 0.97$ (dati mensili).

È sostanzialmente un I-GARCH con $\delta = \mu = 0$

$$\sigma_t^2 = (1 - \lambda)R_t^2 + \lambda\sigma_{t-1}^2.$$

Per un orizzonte più lungo RiskMetrics assume

$$R_{t,T} = \sigma_t \varepsilon \sqrt{T - t},$$

$\varepsilon \sim N(0, 1)$, cioè la regola della radice.

Per un I-GARCH si ha, per ogni k ,

$$E_t \sigma_{t+k}^2 = E_t \sigma_{t+k-1}^2$$

da cui, iterando

$$E_t R_{t+k}^2 = \sigma_t^2$$

(la varianza del rendimento giornaliero è costante). Da questo segue

$$E_t \left(\sum_{i=1}^k R_{t+i} \right)^2 = k \sigma_t^2$$

cioè la varianza (condizionata all'informazione in t) del rendimento $R_{t,t+k}$ si ottiene moltiplicando per k la varianza giornaliera. La regola della radice per la volatilità è giustificata.

Notiamo però che $R_{t,t+k}$ non è normale, anche se viene trattato come tale nelle applicazioni proposte.

Il modello univariato RiskMetrics si estende al caso di più fattori di rischio.

$$\mathbf{R}_{t,T} \sim N(0, \Sigma_t(T-t))$$

dove $\mathbf{R}_{t,T}$ è il vettore composto dai rendimenti dei fattori.

Le stime delle deviazioni standard dei singoli fattori e delle correlazioni sono ottenute in maniera analoga alle stime del caso univariato.

Se $\Sigma_t = (\Sigma_t(i, j))$,

$$\Sigma_t(i, j) = \frac{(1-\lambda)}{(1-\lambda^n)} \sum_{k=0}^{n-1} \lambda^k R_{t-k}^i R_{t-k}^j.$$

4.2 Stime della volatilità per titoli azionari e obbligazionari

Per limitare il numero dei fattori di rischio e quindi dei parametri da stimare, RiskMetrics utilizza alcune tecniche specifiche per titoli azionari e obbligazionari.

4.2.1 Titoli azionari

Risk Metrics stima la volatilità di titoli azionari a partire da quella dell'indice di mercato.

Nel CAPM, introducendo considerazione economiche, il rendimento di un singolo titolo dipende linearmente dal rendimento dell'indice di mercato I_M . Supponendo che i rendimenti attesi siano trascurabili

$$I = (\beta I_M + \varepsilon).$$

La volatilità di I è $\beta \cdot \sigma_M$, dove si è ignorato σ_ε^2 (può essere eliminato per diversificazione).

Esempio : Sia $\sigma_M = 0.25$ la volatilità del MIB30. Si calcoli la volatilità di un titolo del mercato italiano avente $\beta = 0.7$.

4.2.2 Mapping di titoli obbligazionari

Riskmetrics utilizza 14 maturity (dette vertici) per i titoli obbligazionari (da 1 mese a 30 anni).

Per calcolare il VaR di un titolo con maturity diversa dai vertici si utilizza il Cash flow mapping, scomponendolo in un portafoglio di zero coupon bond con maturity nei vertici e in una posizione Cash.

L'idea è di preservare il valore e la sensibilità del titolo rispetto a variazioni dei tassi di interesse.

Esempio. Consideriamo uno zcb unitario con scadenza in t . I vertici adiacenti a t sono t_L e t_R , con

$$t_L < t < t_R$$

Lo zcb unitario viene scomposto in un portafoglio composto da W_L quote di zcb con maturity t_L , W_R quote di zcb con maturity t_R e una rimanenza C di denaro contante. Si procede nel seguente modo:

- si calcola il tasso z_t per interpolazione lineare

$$z_t = \alpha z_L + (1 - \alpha) z_R$$

- si calcola il valore attuale dello z.c.b. e si imposta l'equazione del portafoglio

$$V_t = e^{-z_t t} = W_L e^{-z_L t_L} + W_R e^{-z_R t_R} + C$$

- si impone che la sensibilità dello zcb rispetto a variazioni dei tassi z_L e z_R sia la stessa di quella del portafoglio

$$\frac{\partial V_t}{\partial z_L} = -\alpha t e^{-z_L t} = -W_L t_L e^{-z_L t_L}$$

da cui si ottiene

$$W_L = \alpha \frac{t}{t_L} e^{-z_L t} e^{z_L t_L}$$

e, analogamente

$$W_R = (1 - \alpha) \frac{t}{t_R} e^{-z_R t} e^{z_R t_R}$$

- La posizione Cash viene determinata per differenza

$$C = -\frac{(t - t_L)(t_R - t)}{t_R t_L} e^{-z t}$$

Capitolo 5

Il VaR di un portafoglio

Supponiamo di avere un portafoglio composto da m posizioni di valore V^i che sono funzione dei fattori di rischio $\mathbf{P} = (P^1, \dots, P^N)$

$$V^i = V^i(\mathbf{P})$$

Supponiamo di conoscere la distribuzione di \mathbf{P} , ad come proposto nel capitolo precedente. Supponiamo inoltre di conoscere anche le funzioni di pricing V^i . Per calcolare misure di rischio come il VaR, l'expected shortfall, la deviazione standard, dobbiamo stimare la distribuzione di probabilità P&L:

$$\begin{aligned} P\&L &= \sum_{i=1}^m \Delta V^i \\ &= \sum_{i=1}^m (V^i(\mathbf{P}_T) - V^i(\mathbf{P}_t)) \end{aligned}$$

Tratteremo separatamente i casi univariato e multivariato e cominceremo naturalmente dal caso più semplice, quello in cui la P&L dipende da un solo fattore di rischio.

5.1 Il caso univariato

Questo caso si verifica ad esempio quando dobbiamo valutare la distribuzione P&L di:

1. un portafoglio di opzioni su un'azione o indice.

2. un portafoglio di opzioni scritte su uno zero coupon bond.
3. una opzione su tasso di cambio.
4. una posizione cash in valuta estera.

Ci sono vari metodi per il calcolo della distribuzione. La scelta normalmente dipende dalla funzione di pricing V .

Per fissare le idee ci riferiremo costantemente al seguente Esempio guida, applicandovi via via le varie metodologie per stimare la distribuzione e calcolare il VaR a un giorno e a due settimane.

Esempio guida: Consideriamo un portafoglio composto da q opzioni call europee at-the-money, scadenza un anno, scritte sul MIB30, il quale ha valore corrente (la data odierna è il 17-5-02) di 31667. Si supponga

$$R_{t,T} \sim N(0, \sigma^2(T - t))$$

La stima EWMA al 17-5-02 fornisce $\sigma = 0.0165$, cui corrisponde una volatilità del 26.13% su base annua (moltiplicando per $\sqrt{250}$). Come strumento di pricing utilizzeremo la formula di Black-Scholes con $r = 0.05$.

5.2 Metodi esatti ed approssimati

Il primo caso trattato è quello in cui la la funzione di pricing $V(P)$ è monotona. Vedremo come in questo caso sia possibile calcolare esattamente (cioè senza ricorrere ad approssimazioni), la funzione di ripartizione della P&L ed in particolare anche il VaR.

Se la funzione di pricing $V(P)$ è monotona (crescente o decrescente) e ne conosciamo l'inversa, possiamo calcolare la funzione di ripartizione di $\Delta V := V(P_{t+\Delta T}) - V(P_t)$.

In particolare, se V è crescente

$$\begin{aligned} Pr(\Delta V(P) \leq v) &= Pr(P \leq \Delta V^{-1}(v)) \\ &= F(\Delta V^{-1}(v)) \end{aligned}$$

dove F è la funzione di ripartizione di P_T .

Se V è decrescente

$$Pr(\Delta V(P) \leq v) = 1 - F(\Delta V^{-1}(v)).$$

Osserviamo che per calcolare il percentile v_α di ΔV non serve conoscere l'inversa di ΔV , basta conoscere il percentile p_α di P_T .

Se ΔV è crescente

$$Pr(\Delta V \leq v) = Pr(P \leq \Delta V^{-1}(v))$$

quindi $v_\alpha = \Delta V(p_\alpha)$.

Se ΔV è decrescente: $v_\alpha = \Delta V(p_{1-\alpha})$

Applicando il metodo della funzione monotona per calcolare il VaR dell'esempio guida si trova: $VaR(1g) = 360$; $VaR(10g) = 1088$.

Nel caso in cui la funzione di pricing non sia monotona si può ricorrere ad alcuni metodi di approssimazione. Il più semplice di tutti è ovviamente l'**approssimazione lineare**.

Si procede attraverso uno sviluppo in serie intorno al livello attuale del fattore di rischio P_t :

$$\Delta V \approx V'(P_t)(P_T - P_t) = \delta I_{t,T}$$

dove $\delta = V'(P_t)P_t$.

RiskMetrics utilizza l'approssimazione $R_{t,T} \approx I_{t,T}$, pertanto

$$\Delta V \approx \delta R_{t,T}$$

Per il calcolo del VaR, supponendo (come RiskMetrics) che

$$R \sim N((T-t)\mu, \sqrt{T-t}\sigma)$$

si ha

$$\Delta V \sim N(\delta\mu(T-t), |\delta|\sigma\sqrt{T-t})$$

quindi

$$VaR_p = -(\delta\mu(T-t) + |\delta|\sigma\sqrt{T-t}z_{(1-p)})$$

$z_{(1-p)}$ percentile della normale standard.

Da notare che l'approssimazione di Taylor è valida per piccole variazioni, mentre per il VaR ci interessano variazioni significative dell'indice.

Ci potrebbero inoltre essere problemi quando questa approssimazione viene applicata su posizioni fortemente non lineari, come ad esempio uno straddle (call e put scritte sullo stesso sottostante).

Esercizio: applicare il metodo della approssimazione lineare per calcolare il VaR dell'esempio guida. Confrontare i risultati ottenuti con quelli esatti dell'esercizio precedente.

Quando l'approssimazione lineare non è soddisfacente (posizioni con convessità non trascurabile, come uno Straddle), si può arrivare al secondo ordine dello sviluppo di Taylor, ricorrendo alla **approssimazione quadratica**.

$$\begin{aligned}\Delta V &\approx V'(P_t)(\Delta P) \\ &+ \frac{1}{2}V''(P_t)(\Delta P)^2 \\ &\approx \delta R_{t,T} + \frac{1}{2}\gamma R_{t,T}^2\end{aligned}$$

dove $\gamma = V''(P_t)P_t^2$.

In questo caso, anche se $R_{t,T}$ è normale, ΔV non lo è, quindi per calcolare il percentile si può utilizzare:

- il metodo dei momenti, cioè calcolare i primi 4 momenti di ΔV e poi determinare una distribuzione che abbia gli stessi momenti (e di cui si sa calcolare il percentile);
- le simulazioni

5.3 Simulazioni Monte Carlo

Quando non si dispone di una formula analitica per il percentile e non si vuole ricorrere a tecniche di approssimazione si può ricorrere alle simulazioni.

Avendo ipotizzato una distribuzione probabilistica per $R_{t,T}$ si procede come segue:

1. Si produce un valore \bar{R} estratto da tale distribuzione. (Si veda il Capitolo 10 per come procedere per simulare una variabile aleatoria da una distribuzione assegnata)
2. Si calcola $\bar{P}_T = P_t e^{\bar{R}}$.
3. Si calcola $\Delta V(\bar{P}_T) = V(\bar{P}_T) - V(P_t)$.
4. Si ripete un numero sufficientemente elevato di volte per trovare la distribuzione empirica di ΔV .

Per il calcolo del VaR si prende il percentile empirico.

Computazionalmente è il metodo più oneroso ma dà i risultati più precisi.

Esercizio: applicare il metodo della simulazione Monte Carlo per calcolare il VaR dell'esempio guida.

5.4 Simulazione storica

In questo caso si procede sempre tramite simulazioni, però non si ipotizza una distribuzione per $R_{t,T}$, utilizzando un metodo non parametrico. I valori \bar{R} si estraggono dalla serie storica del titolo. Si procede come segue:

1. Si produce un valore \bar{R} estratto dalla serie storica giornaliera.
2. Si calcola $\bar{P}_T = P_t e^{\bar{R}\sqrt{T-t}}$ (tempo espresso in giorni).
3. Si calcola $\Delta V(\bar{P}_T) = V(\bar{P}_T) - V(P_t)$.
4. Si ripete un numero sufficientemente elevato di volte per trovare la distribuzione empirica di ΔV .

Anche qui si considerano i percentili empirici.

I vantaggi di questo approccio sono che

- non si devono fare ipotesi sulla distribuzione. Si limita il rischio di modello;
- si catturano in una certa misura le caratteristiche della distribuzione, quali la grossezza delle code.

I punti deboli sono:

- in realtà non è vero che non si facciano ipotesi sulla distribuzione: si sta in realtà assumendo che i rendimenti siano tra loro indipendenti.
- si suppone che il passato rappresenti sufficientemente bene il futuro: anche eventi accaduti molto tempo fa hanno la stessa probabilità di accadere di eventi recenti.

Esercizio: applicare il metodo della simulazione storica per calcolare il VaR dell'esempio guida.

5.5 Il caso multivariato

Ora vedremo come trattare il caso, indubbiamente più realistico, di una posizione dipendente da più fattori di rischio: anche in questo caso faremo riferimento ad un esempio guida:

Esempio guida 2 Consideriamo un investitore europeo che possiede, il 17-5-02, un portafoglio composto da $q = 1000$ quote dell'indice S&P500, il cui valore corrente è $S_t = 1106.59$. Il portafoglio è soggetto, oltre che al rischio di mercato, anche al rischio di cambio. L'obiettivo è valutare il VaR al 95% a 1 giorno e a 2 settimane della posizione.

Sia X il tasso di cambio, ovvero il valore in Euro di 1 USD. Si ha $X_t = 1.09$.

Il valore della posizione è $V = qS \cdot X$. Il capitale investito oggi è $V_t = q \cdot 1098 \cdot 1.1$

Sia $P^1 = S$ e $P^2 = X$.

$$\mathbf{R}_{t,T} \sim N(0, \Sigma \sqrt{T-t})$$

La stima della matrice Σ fatta il 17-5-02 fornisce

$$\sigma_1 = \sqrt{\Sigma_{1,1}} = 0.22$$

$$\sigma_2 = \sqrt{\Sigma_{2,2}} = 0.076$$

$$\rho = \frac{\Sigma_{1,2}}{\sigma_1 \sigma_2} = 0.49$$

Un primo metodo di calcolo è fornito dalla **Approssimazione lineare**.

Per ciascuna posizione si ha

$$\Delta V \approx \sum_{i=1}^N \frac{\partial V}{\partial P^i} \Delta P^i \approx \sum_{i=1}^N \delta_i R_{t,T}^i$$

dove $\delta_i = P_t^i \frac{\partial V}{\partial P^i}$, e utilizzando l'approssimazione $R_{t,T}^i \approx I_{t,T}^i$. In notazione matriciale

$$\Delta V \approx \delta' \mathbf{R}_{t,T}$$

Dunque

$$\begin{aligned} P\&L &= \sum_{i=1}^m \Delta V^i \\ &\approx \sum_{i=1}^m \delta'_i \mathbf{R}_{t,T} \\ &= \delta'_{P_t} \mathbf{R}_{t,T} \end{aligned}$$

Se $\mathbf{R}_{t,T} \sim N((T-t)\mu, (T-t)\Sigma)$, si ha

$$P\&L \sim N(\delta' \mu (T-t), \delta' \Sigma \delta (T-t))$$

Quindi per il calcolo del VaR si ha

$$\text{VaR}_p = -(\delta'\mu(T-t) + \sqrt{\delta'\Sigma\delta(T-t)}z_{(1-p)})$$

Esercizio: calcolare il VaR dell'esempio guida utilizzando l'approssimazione lineare.

Un'altra possibilità è quella di utilizzare le **Simulazioni Monte Carlo**. Consideriamo il caso

$$\mathbf{R}_{t,T} \sim N(\mu(T-t), \Sigma(T-t)).$$

1. Si produce un valore $\bar{\mathbf{R}}$ estratto da una distribuzione $N(0, \Sigma)$.
2. Si calcola $\bar{\mathbf{P}}_T = \mathbf{P}_t e^{\bar{\mathbf{R}}\sqrt{T-t} + \mu(T-t)}$. Si è qui utilizzata una notazione adatta al MATLAB.
3. Si calcola $\Delta V^i(\bar{\mathbf{P}}_T) = V^i(\bar{\mathbf{P}}_T) - V^i(\mathbf{P}_t)$ per ogni strumento.
4. Si calcola $P\&L = \sum_{i=1}^m \Delta V^i$
5. Si ripete un numero sufficientemente elevato di volte per trovare la distribuzione empirica di ΔV

Per il calcolo del VaR si prende il percentile empirico.

Esercizio: applicare il metodo della simulazione Monte Carlo per calcolare il VaR dell'esempio guida.

Anche in quest caso, l'alternativa alla simulazione Monte Carlo è data dalla **simulazione storica**

Nel caso multifattoriale abbiamo N serie storiche di rendimenti di lunghezza n che possiamo inserire in una matrice \mathcal{R} di N righe e n colonne.

1. Si prende una colonna $\bar{\mathbf{R}}$ estratta a caso da \mathcal{R} .
2. Si calcola $\bar{\mathbf{P}}_T = \mathbf{P}_t e^{\bar{\mathbf{R}}\sqrt{T-t}}$ (tempo espresso in giorni) per ogni fattore di rischio.
3. Si calcola $\Delta V^i(\bar{\mathbf{P}}_T) = V^i(\bar{\mathbf{P}}_T) - V^i(\mathbf{P}_t)$.
4. Si calcola $P\&L = \sum_{i=1}^m \Delta V^i$
5. Si ripete un numero sufficientemente elevato di volte per trovare la distribuzione empirica di ΔV

Si considerano i percentili empirici.

Ulteriore vantaggio: non c'è bisogno di stimare la dipendenza tra i vari fattori.

Esercizio: applicare il metodo della simulazione storica per calcolare il VaR dell'esempio guida.

Capitolo 6

Backtesting

Il backtesting è un processo di verifica dei modelli adottati per stimare il VaR ed è raccomandato dal Comitato di Basilea. Nel metodo comunemente utilizzato si considerano n osservazioni passate del valore del portafoglio. Per ognuna di queste si producono, con il modello, le stime del VaR. Si conta il numero di volte in cui tale stima è stata superata.

Se, per esempio, calcoliamo il VaR al 95%, ci aspettiamo che, se il modello usato funziona bene e se abbiamo un numero sufficientemente elevato di osservazioni, le volte in cui le perdite hanno superato la soglia del VaR siano all'incirca il cinque per cento.

Infatti, sia p^* il livello di confidenza del VaR e siano X_1, \dots, X_n delle variabili aleatorie binomiali, tali che X_i vale 1 se il VaR il giorno i è stato superato e 0 altrimenti. Si ha $P(X_i = 1) = 1 - p^*$ e $P(X_i = 0) = p^*$. La media di X_i è $1 - p^*$. Per la legge dei grandi numeri

$$\frac{X_1 + \dots + X_n}{n} \rightarrow 1 - p^* \quad \text{q.o.}$$

Rimandiamo lo studente interessato a Duffie, Pan [1997] per alcuni risultati sulla velocità di convergenza.

Se la percentuale è maggiore potrebbe significare che stiamo sottostimando il VaR, altrimenti sovrastimando.

Come decidere se accettare il modello o meno?

Come osservato in precedenza, simulando n osservazioni da variabili aleatorie note si possono ottenere percentuali di superamento dei percentili con errori anche rilevanti, che dipendono da n e dalla coda della distribuzione (Kupiec [1995]). In pratica si considerano serie storiche non troppo lunghe.

Infatti, il Comitato di Basilea richiede il backtest del VaR al 99 % sulle ultime 250 osservazioni. Questo pone seri dubbi sulla efficacia statistica delle stime. Ad ogni modo, il risultato viene classificato in tre zone a secondo del numero di violazioni: verde (0-4), giallo (5-9), rosso (10-).

Per la verifica del modello si possono usare test statistici, detti test di ipotesi. Si vogliono determinare le regioni di accettazione e di rigetto delle ipotesi. Nel nostro caso l'ipotesi è che il VaR al livello p^* sia quello stimato.

In generale, se $X = X(\theta)$ è la variabile aleatoria usata per la verifica, si vuole avere una statistica test T tale che la regione di accettazione dell'ipotesi nulla sia:

$$\{x : T(x) \leq r_c\}$$

dove r_c è un percentile tale che

$$P(T(X) \leq r_c; \theta^*) = c.$$

x indica il vettore di dati usati per la verifica, $\theta = \theta^*$ è l'ipotesi nulla.

L'ipotesi è rigettata se i nostri dati sono fuori la regione di accettazione. Si commette un errore di tipo I (rigetto dell'ipotesi nulla quando vera) con probabilità $1 - c$.

Ora descriveremo un test importante: **Il test PF (proportion of failures)**

Consideriamo il test sulla percentuale dei superamenti del VaR. Sia x il numero di superamenti su n osservazioni. Consideriamo un livello di confidenza c (tipicamente $c = 0.95$). L'ipotesi nulla è che la probabilità di un superamento sia $p = p^*$. Per quali valori di x , su n osservazioni, commettiamo un errore di tipo I con probabilità $1 - c$? Si considera la variabile $X(p) = X_1 + \dots + X_n$. Si ha:

$$P(X(p) = x) = \text{binom}(n, x)(1 - p)^x p^{n-x}$$

Si introduce la statistica test log-likelihood ratio

$$\begin{aligned} LR(x) = & - 2 \log[((1 - p^*)^x (p^*)^{n-x}] \\ & + 2 \log[(1 - x/n)^{n-x} (x/n)^x] \end{aligned}$$

che converge, per $n \rightarrow \infty$ e sotto l'ipotesi nulla $p = p^*$, a una chi quadro con un parametro di libertà.

Dai percentili della chi quadro si ricavano le regioni di accettazione al livello di confidenza c :

$$\{x : LR(x) \leq r_c\}$$

$$P(LR(X(p)) \leq r_c; p^*) = c.$$

Kupiec riporta una tabella per vari livelli di confidenza del VaR e vari n . Ad esempio, con 255 osservazioni, per il livello di confidenza del VaR al 95%, si è portati ad accettare il modello, con probabilità di errore di tipo I del 5%, quando $6 < x < 21$, ovvero con percentuali da 2.3% a 8.2%.

Osservazioni:

1. le regioni di accettazione si assottigliano, dal punto di vista delle percentuali, al crescere di n ;
2. per pochi dati ($n = 255$) e livelli di confidenza alti ($p^* = 0.99$) la regione di accettazione è $x < 7$, per cui non siamo in grado di decidere se il modello sovrastima il rischio.

Per maggiori approfondimenti sui test statistici si rimanda a Crnkovic, Drachman [1996] oltre che a Kupiec [1995].

Capitolo 7

Stress testing

Lo stress testing è una metodologia che calcola il valore del P&L nel caso di eventi estremi particolarmente sfavorevoli. Per generare gli scenari di stress testing si procede in due modi

- storico
- parametrico

Il materiale di questo capitolo è tratto da (RiskMetrics [2001], capitolo 4)

7.1 Lo Stress testing storico

Risponde alla domanda: cosa succederebbe al mio portafoglio se si ripettesse uno scenario simile alla crisi x ?

Si procede selezionando un intervallo temporale da t a T , calcolando i log-return

$$r = \log(P_T/P_t)$$

e infine calcolando il P&L del nostro portafoglio se tali rendimenti si ripetessero

$$P\&L = V(Pe^r) - V(P)$$

Esempio: la crisi Russa del 1998.

Supponiamo di avere un portafoglio di 3000 USD investito in parti uguali sui mercati Brasiliano, Indonesiano e Polacco.

Vogliamo vedere quale sarebbe l'impatto sul nostro portafoglio di una crisi simile a quella Russa.

Prendiamo l'intervallo di tempo tra 1 Luglio 1998 e 30 Agosto 1998. I log-return in quel periodo sono stati:

	Equity		FX	
Brasile	Bovespa	-48,19%	BRL	-1,34%
Indonesia	JSE	-36,47%	IDR	22,60%
Polonia	WIG	-41,24%	PLN	-10,19%

La P&L per ciascuna posizione è

$$\Delta V = V[e^{r^1+r^2} - 1]$$

dove $V = 1000USD$, r^1 è il log-return dell'indice azionario (in valuta locale) e r^2 è il log-return del tasso di cambio rispetto al dollaro.

I risultati dello stress-test sono:

Nazione	ΔV (in USD)
Brasile	-350,59
Indonesia	-129,58
Polonia	-402,11
Totale	-922,29

7.2 Lo Stress testing parametrico

Si ipotizzano scenari estremi per alcuni fattori di rischio (fattori di base) e si determina, sulla base di un modello parametrico, i valori più realistici possibile per gli altri fattori (fattori periferici).

Occorre un modello per tutti i fattori, ad esempio, se \mathbf{r}_1 sono i fattori periferici e \mathbf{r}_2 i fattori di base:

$$\begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} \sim N \left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}; \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right)$$

da cui si trova

$$E[\mathbf{r}_1 | \mathbf{r}_2] = \mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (\mathbf{r}_2 - \mu_2)$$

Vediamo ora un esempio di stress testing parametrico sul portafoglio precedente.

Si prendono come fattori di base i tassi di cambio e si ipotizzano rendimenti logaritmici pari a -10% per ciascuno di essi. I fattori periferici sono

quindi i rendimenti degli indici di mercato. Applicando la matrice di varianza covarianza a pag. 36 di RiskMetrics (2001) si trova

Indice	log-ret
Bovespa	-8,59%
JSE	-1,83%
WIG	-0,57%

A questo punto si procede come sopra calcolando il P&L

Nazione	ΔV (in USD)
Brasile	-174,09
Indonesia	-116,31
Polonia	-105,12
Totale	-395,52

Capitolo 8

Teoria dei valori estremi

La Teoria dei valori estremi (EVT) fornisce metodi e strumenti per la modellizzazione di:

1. Valori massimi in un certo intervallo di tempo (es: annuali o mensili).
2. Code di distribuzioni.

Si può utilizzare per:

- stress testing;
- stime di percentili;
- stime di misure di rischio oltre il VaR, come
 - la distribuzione condizionata oltre il VaR

$$P(R > x \mid R > r_p)$$

- il valore atteso oltre il VaR (*expected shortfall*)

$$\mathbf{E}[R \mid R > r_p].$$

8.1 Modelli per i massimi

Consideriamo R_1, R_2, \dots delle variabili aleatorie i.i.d. con funzione di distribuzione F , $R_i \geq 0$. Siano $M_n = \max(R_1, \dots, R_n)$ i massimi di queste variabili.

Un Teorema fondamentale di EVT descrive le possibili funzioni di distribuzione che sono limiti di massimi (scalati e normalizzati). Analogamente al teorema del limite centrale per le somme.

$$P((M_n - b_n)/a_n \leq x) \approx H_\xi(x)$$

$$H_\xi(x) = \begin{cases} \exp(-(1 + \xi x)_+^{-1/\xi}) & \xi \neq 0, \\ \exp(-e^{-x}) & \xi = 0 \end{cases}$$

funzioni di distribuzione GEV (*Generalised Extreme Value*).

Il parametro ξ è detto parametro di forma.

$\xi > 0$ Fréchet. È il caso più importante per le applicazioni.

$\xi = 0$ Gumbel.

$\xi < 0$ Weibull.

Teorema: Le distribuzioni con le code grosse hanno massimi che si distribuiscono come una GEV con parametro $\xi > 0$ (Fréchet).

Le GEV modellano la distribuzione di massimi di variabili aleatorie.

Applicazioni:

- studio dei massimi campionari a blocchi (settimanali, bisettimanali, mensili, annuali);
- stress testing.

8.2 Modelli per le code

Cominciamo con il definire alcune distribuzioni che saranno utili nel seguito:

Le funzioni di distribuzione GPD (*Generalised Pareto Distribution*) sono le seguenti

$$G_{\xi,\beta}(x) = \begin{cases} 1 - (1 + \xi x/\beta)_+^{-1/\xi} & \xi \neq 0, \\ 1 - e^{-x/\beta} & \xi = 0 \end{cases}$$

$\beta > 0$.

$\xi > 0$ Pareto.

$\xi = 0$ Esponenziale.

$\xi < 0$ Pareto di tipo II.

La Pareto è una distribuzione con la coda grossa. Il parametro ξ misura la pesantezza della coda. Maggiore ξ più pesante è la coda. La curtosi di $G_{\xi,\beta}$ è infinita per $\xi \geq \frac{1}{4}$.

Se i massimi di R convergono a una distribuzione GEV con parametro ξ , la $G_{\xi,\beta}$ approssima la funzione di distribuzione delle eccedenze:

$$P(R - u \leq x \mid R > u) \approx G_{\xi,\beta}(x)$$

con $\beta = \beta(u)$ una funzione positiva di u e u sufficientemente grande.

Le GPD sono alla base del metodo POT (Peaks over Thresholds), il quale permette di stimare percentili: si basa sulla semplice relazione, valida per $x \geq u$:

$$P(R > x) = P(R > x \mid R > u) \cdot P(R > u)$$

Siano R_1, R_2, \dots, R_n sono i nostri dati.

Stimiamo $P(R > u)$ empiricamente come $N_u =$ numero dei dati superiori ad u diviso n .

allora si dimostra che

$$P(R > x \mid R > u) \approx 1 - G_{\xi,\beta}(x - u).$$

e quindi si ottiene:

$$P(R > x) \approx \frac{N_u}{n} (1 - G_{\xi,\beta}(x - u))$$

Ora vediamo come utilizzare il metodo POT per la stima dei percentili:

$$\hat{r}_p = u + \frac{\beta}{\xi} \left(\left(\frac{n}{N_u} (1 - p^*) \right)^{-\xi} - 1 \right)$$

Si ottiene anche una stima semplice del valore atteso oltre il VaR:

$$\mathbf{E}[R \mid R > \hat{r}_p] = \hat{r}_p \left(\frac{1}{1 - \xi} + \frac{\beta - \xi u}{\hat{r}_p (1 - \xi)} \right)$$

Per un esempio di applicazione della metodologia POT consideriamo i rendimenti del titolo FIAT dal 1986 al 1999.

Osseviamo che le code non troppo pesanti: infatti si stima un coefficiente ξ intorno al valore 0.15.

Da ciò seguono questi valori dei percentili (in parentesi quelli che si otterrebbero con ipotesi di normalità)

95-esimo percentile $\approx 3.9\%$ (N=3.2%)

99-esimo percentile $\approx 6.5\%$ (N=4.6%)

995-esimo percentile $\approx 7.8\%$ (N=5.1%)

999-esimo percentile $\approx 11.4\%$ (N=6.1%)

Capitolo 9

Altri modelli per la stima del rischio di mercato

In questo capitolo illustreremo due modelli che sono stati proposti recentemente e costituiscono una possibile alternativa alla metodologia RiskMetrics. Le ipotesi semplificatrici alla base del modello RiskMetrics costituiscono allo stesso tempo un punto di forza ed una debolezza. Nelle proposte che illustreremo nel seguito vedremo come rimuovere l'ipotesi di normalità delle innovazioni utilizzando la teoria dei valori estremi (modello di Frey, McNeal [1998]) e come procedere in modo semi-parametrico, non facendo ipotesi sulla distribuzione delle innovazioni ma usando un approccio di simulazione storica (modello BAGV [2000]). Da notare che entrambi gli approcci utilizzano un modello GARCH alla base della loro procedura di stima.

9.1 Il modello di Frey, McNeil

Il metodo POT si basa sull'assunzione che i nostri dati siano estrazioni indipendenti da una stessa distribuzione. Come abbiamo visto, questa ipotesi si adatta difficilmente ai dati finanziari.

L'idea è di filtrare la dipendenza attraverso un modello parametrico opportuno e di utilizzare poi la EVT.

Considerano un misto AR(1) + GARCH(1,1):

$$R_t = \phi R_{t-1} + \sigma_t \varepsilon_t$$
$$\sigma_t^2 = \delta + \alpha (R_{t-1} - \mu_{t-1})^2 + \beta \sigma_{t-1}^2$$

$\delta, \alpha, \beta > 0$.

L'obiettivo è produrre stime di VaR e di expected shortfall.

Fissiamo un valore n che rappresenta quanti dati passati considerare (ad esempio $n = 1000$). Alla fine del giorno t abbiamo i dati R_{t-n+1}, \dots, R_t . Eseguiamo i seguenti passi:

1. Calibratura del modello. Stimiamo $\hat{\phi}, \hat{\delta}, \hat{\alpha}, \hat{\beta}$.
2. Stimiamo le serie storiche dei valori attesi e volatilità condizionali $(\mu_{t-n+1}, \dots, \mu_t, \sigma_{t-n+1}, \dots, \sigma_t)$.
3. Calcoliamo μ_{t+1} e σ_{t+1}

$$\begin{aligned}\mu_{t+1} &= \hat{\phi} R_t, \\ \sigma_{t+1}^2 &= \hat{\delta} + \hat{\alpha}(R_t - \mu_t)^2 + \hat{\beta}\sigma_t^2.\end{aligned}$$

4. Estraiamo i residui risultanti

$$(e_{t-n+1}, \dots, e_t) = \left(\frac{R_{t-n+1} - \mu_{t-n+1}}{\sigma_{t-n+1}}, \dots, \frac{R_t - \mu_t}{\sigma_t} \right).$$

Si osserva che: residui possono essere considerati iid (infatti le acf dei rendimenti e dei loro quadrati sono trascurabili). Essi mostrano ancora code pesanti (come si osserva attraverso un QQ-plot contro la normale).

È quindi giustificato l'utilizzo del metodo POT applicato ai residui con il quale stimano $\hat{\varepsilon}_{1-p}$ e $E(\varepsilon \mid \varepsilon < \hat{\varepsilon}_{1-p})$.

Nel loro lavoro Frey e McNeil applicano il modello a serie storiche del DAX, SP500, BMW. Per un orizzonte più lungo adottano una procedura Montecarlo. Fanno simulazioni dal modello dinamico con innovazioni con code pesanti. Osservano che la regola della radice quadrata sembra non essere sufficiente.

Dal punto di vista empirico sembra che il modello spiega bene i dati. La procedura di backtesting eseguita sulle stime di VaR dà buoni risultati.

I limiti del modello sono principalmente due: così come viene proposto non è utilizzabile direttamente per portafogli e dal punto di vista teorico non tutta la loro procedura è giustificabile.

9.2 Il modello di Barone-Adesi, Giannopoulos, Vosper

L'idea alla base di questo modello è molto semplice: filtrare la dipendenza con un GARCH(1,1) su ciascuno delle componenti separatamente e utilizzare i residui per generare una serie di simulazioni storiche indipendenti.

Il modello è un ARMA-A-GARCH(1,1):

$$R_t = \phi R_{t-1} + \theta \sigma_{t-1} \varepsilon_{t-1} + \sigma_t \varepsilon_t$$

e

$$\sigma_t^2 = \delta + \alpha (R_{t-1} - \phi R_{t-2} - \theta \sigma_{t-2} \varepsilon_{t-2} + \gamma)^2 + \beta \sigma_{t-1}^2$$

oppure

$$\sigma_t^2 = \delta + \alpha (\sigma_{t-1} \varepsilon_{t-1} + \gamma)^2 + \beta \sigma_{t-1}^2$$

Cominciamo con il vedere come procedere nel caso di un solo titolo attraverso un metodo di simulazione storica per R_{t+1}

Ricordiamo che in t sono noti R_t e σ_{t+1} :

$$\sigma_{t+1}^2 = \hat{\delta} + \hat{\alpha} (\sigma_t e_t + \hat{\gamma})^2 + \hat{\beta} \sigma_t^2$$

Si procede seguendo i seguenti passi:

1. Stima dei coefficienti del modello $(\hat{\theta}, \hat{\delta}, \hat{\alpha}, \hat{\gamma}, \hat{\beta}, \hat{\phi})$
2. Calcolo della serie storica delle innovazioni standardizzate

$$e_1, e_2, \dots, e_t$$

(NB: se le ipotesi del modello sono verificate, le e_i sono realizzazioni di una $N(0, 1)$).

3. Estrazione casuale di una innovazione e dalla serie storica: sarà l'innovazione per il periodo successivo. $\hat{\varepsilon}_{t+1} := e$
4. Calcolo della realizzazione di R_{t+1}

$$\hat{R}_{t+1} := \hat{\phi} R_t + \hat{\theta} \sigma_t e_t + \sigma_{t+1} \hat{\varepsilon}_{t+1}$$

Ripetendo la procedura N volte (con reimbussolamento) si ottiene un insieme di estrazioni i.i.d.

Osserviamo che, rispetto al modello di Frey,McNeal la differenza principale consiste nel fatto che la distribuzione delle innovazioni non viene stimata. In questo caso si utilizza la distribuzione empirica. E' quello che si dice un metodo non parametrico.

Per simulare il valore di R_{t+k} , dato che non se ne conosce la distribuzione condizionata, occorre produrre dati simulati anche per le variabili σ_{t+j} , $j = 2, \dots, k$. L'algoritmo di simulazione è il seguente:

1. Estrazione casuale di k residui standardizzati e_1, \dots, e_k
2. A partire da e_1 , calcolare \dot{R}_{t+1} secondo lo schema precedente. Calcolare anche il corrispondente $\dot{\sigma}_{t+2}$.
3. A partire da e_2 , da \dot{R}_{t+1} e da $\dot{\sigma}_{t+2}$, calcolare \dot{R}_{t+2} secondo lo schema precedente.
4. Continuare fino a k ...

Fin qui, come detto, non sembrano esserci grandi differenze rispetto al modello di Frey,McNeal. Come osservato in precedenza, un punto debole del modello di Frey,McNeal è che non si applica direttamente al caso multivariato. Nel caso di BAGV ciò invece risulta molto semplice. Vediamo dunque il caso di un portafoglio composto da n titoli. In questo caso si stimano separatamente n modelli GARCH. Per una simulazione vengono estratti gli n residui realizzati nelle stesse k date. Si procede quindi come nel caso precedente.

Il vantaggio è che non occorre stimare la covarianza tra i fattori, infatti essa si conserva nella scelta contemporanea dei residui.

Estrazione delle k date: t_1, t_2, \dots, t_k e degli n residui:

$$\begin{aligned} \dot{e}^1 &:= \{e_{t_1}^1, e_{t_2}^1, \dots, e_{t_k}^1\} \\ \dot{e}^2 &:= \{e_{t_1}^2, e_{t_2}^2, \dots, e_{t_k}^2\} \\ &\vdots \\ \dot{e}^n &:= \{e_{t_1}^n, e_{t_2}^n, \dots, e_{t_k}^n\} \end{aligned}$$

Procedere come sopra per ottenere

$$\dot{R}_{t+k}^1, \dot{R}_{t+k}^2, \dots, \dot{R}_{t+k}^n$$

Il lavoro di BAGV presenta un'applicazione del modello a un portafoglio composto da tre contratti futures trattati al LIFFE e da una opzione call su uno di essi.

È stato presentato anche un lavoro di backtesting (BAGV [2000]).

E' possibile pensare al metodo BAGV combinato con EVT: una volta ottenuta con il modello BAGV la distribuzione P&L del portafoglio su un certo orizzonte, si applica il metodo POT.

Capitolo 10

Complementi

In questo capitolo di appendice abbiamo raccolto alcuni elementi complementari al corso.

10.1 Simulazione di una variabile aleatoria

Per simulare una variabile aleatoria univariate X con distribuzione $F(x)$ è sufficiente avere a disposizione un simulatore di una variabile aleatoria uniforme U tra 0 e 1. Infatti si ha:

$$X \sim F^{-1}(U)$$

dove U è una variabile aleatoria uniforme $(0, 1)$. Infatti:

$$\begin{aligned} P(F^{-1}(U) \leq x) &= P(U \leq F(x)) \\ &= F(x) = P(X \leq x) \end{aligned}$$

Nota: Si utilizza se si dispone di una formula per $F^{-1}(x)$, non va bene per simulare una normale.

Per simulare una distribuzione normale multivariata di dimensione d

$$\mathbf{R} \sim N(\mu, \Sigma)$$

è sufficiente simulare una normale standard Z di dimensione d . Infatti, dato che Σ è definita positiva, si ha

$$\Sigma = L'L,$$

allora

$$\mathbf{R} \sim L'Z + \mu$$

La matrice L può ottenersi tramite la decomposizione di Cholesky. Esempio (caso $d = 2$):

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$
$$L = \begin{pmatrix} \sigma_1\sqrt{1-\rho^2} & 0 \\ \rho\sigma_1 & \sigma_2 \end{pmatrix}$$

L'esempio che segue è una variabile aleatoria con le code grosse. Si ottiene con una mistura di normali. Può servire per effettuare simulazioni da una distribuzione con curtosi assegnata.

$$X := ZN_1 + (1 - Z)N_2$$

dove $Z = 1$ con prob. p e $Z = 0$ con prob. $1 - p$, $N_1 = N(\mu_1, \sigma_1)$, $N_2 = N(\mu_2, \sigma_2)$, Z, N_1, N_2 indipendenti.

$$EX = p\mu_1 + (1 - p)\mu_2$$

$$VarX = p\sigma_1^2 + (1 - p)\sigma_2^2$$

$$EX^4 = 3(p\sigma_1^4 + (1 - p)\sigma_2^4)$$

Esercizio: eseguire una simulazione di una mistura di normali con $\mu_1 = 0, \sigma_1 = 0.5, \mu_2 = 0, \sigma_2 = 4, p = 0.95$. La varianza è 1.0375, la curtosi 35.84.

10.2 Misure di rischio coerenti

Il materiale di questo paragrafo è tratto in gran parte dal lavoro di Artzner et al. [1998].

Un funzionale $\rho : L^\infty \rightarrow R$ è una misura di rischio coerente se soddisfa le seguenti condizioni:

1. $X \geq 0 \Rightarrow \rho(X) \leq 0$;
2. $\lambda \geq 0, \rho(\lambda X) = \lambda\rho(X)$;
3. $k \in R, \rho(X + k) = \rho(X) - k$;

$$4. \Rightarrow \rho(X + Y) \leq \rho(X) + \rho(Y).$$

Esempi di misure coerenti:

Sia \mathcal{P} un insieme di misure di probabilità Q tutte assolutamente continue rispetto a P .

Q si può identificare con la derivata di Radon-Nykodim (la densità) che è un elemento di $L^1(\Omega, P)$.

Definiamo:

$$\rho(X) = \sup\{E_Q[-X] \mid Q \in \mathcal{P}\}$$

ρ soddisfa le proprietà 1), 2), 3), 4).

Viceversa ogni misura di rischio che soddisfa 1), 2), 3), 4) è di questo tipo.

Casi particolari:

$$1. \mathcal{P} = \{Q \mid Q \text{ misura di probabilità}\}$$

$$\rho(X) = \sup(-X)$$

In questo caso si ha:

$$\rho(X) \leq 0 \iff X \geq 0$$

$$2. \mathcal{P} = \{Q\}$$

$$\rho(X) = E_Q[-X]$$

$$3. \mathcal{P}_\alpha = \{Q \mid \frac{dQ}{dP} \leq \frac{1}{\alpha}\}$$

$$\alpha \leq 1.$$

Se $\alpha = 1$ allora $\mathcal{P}_1 = \{P\}$ e siamo nel caso precedente.

Denotiamo con ρ_α la misura di rischio associata a \mathcal{P}_α

Il seguente risultato mostra che l'expected shortfall è l'unica misura di rischio coerente se X è una distribuzione continua.

Teorema: Se X ha funzione di distribuzione continua si ha:

$$\rho_\alpha(X) = E_P[-X \mid -X \geq \text{VaR}_{1-\alpha}(X)]$$

E' immediato osservare che $\rho_\alpha(X) \geq \text{VaR}_{1-\alpha}(X)$, ma che rapporto c'è tra VaR ed expected shortfall? Ovviamente dipende dalla distribuzione di X , come mostra il seguente esempio.

Consideriamo 150 clienti indipendenti con probabilità di default p .

$-X$ = numero di default.

$$P(Z = k) = \binom{150}{k} p^k (1-p)^{150-k}.$$

Se $p = 1.2\%$ e $\alpha = 1\%$:

$$\text{VaR}_{1-\alpha}(X) = 5$$

L'expected shortfall è $\rho_\alpha(X) = 6.287$.

Quindi, se la distribuzione non ha le code grosse, in particolare se X è normale,

$$\rho_\alpha \sim \text{VaR}_{1-\alpha}$$

- Se la distribuzione di X ha le code grosse, cioè $1 - F(x) = P(X > x) \in$

$\mathcal{R}_{-\beta}$

($\iff F \in MDA(\text{Fréchet})$)

allora $\rho_\alpha(X) \sim \frac{\beta}{\beta-1} \text{VaR}_{1-\alpha}$ Quindi, se β è vicino ad 1, $\rho_\alpha(X)$ si discosta da $\text{VaR}_\alpha(X)$ in maniera significativa.

Bibliografia

- [1] C. Alexander (1998), Volatility and correlation: measurement, models and applications, Risk management and analysis.
- [2] F. Angelini (1999), An analysis of Italian financial data using extreme value theory, Giornale dell'Istituto Italiano degli Attuari, volume LXV, n. 1-2 (2002).
- [3] P. Artzner, F. Delbaen, J.M. Eber, D. Heath, (1998) Coherent measures of risk, Math. Fin. 9, 203-228.
- [4] G. Barone-Adesi, K. Giannopoulos, L. Vosper (1998) VaR without correlation for portfolios of derivative securities, preprint.
- [5] G. Barone-Adesi, K. Giannopoulos, L. Vosper (2000) Filtering historical simulation. Backtest analysis, preprint.
- [6] T. Bollerselv (1986), Generalised autoregressive conditional heteroskedastic, Journal of Economics, 31, 307-327.
- [7] C. Crnkovic, J. Drachman (1996), Quality control, RISK, 9, (september 1996), 138-143.
- [8] D. Duffie, J. Pan (1997), An overview of Value at Risk, Journal of Derivatives, Spring 1997, 7-48.
- [9] P. Embrechts, C. Klüpperberg, T. Mikosch (1997), Modelling extremal events in insurance and finance, Springer-Verlag, Berlino.
- [10] R.F. Engle (1995) ARCH: Selected readings, Advanced Text in Econometrics, Oxford University Press.

- [11] R. Frey, A. McNeil (1998): Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach , Journal of Empirical Finance, 7, 271-300.
- [12] A. McNeil (1998) On extreme and crashes, RISK, January 1998.
- [13] P. Kupiec (1995) Techniques for verifying the accuracy of risk measurement models, Journal of Derivatives, Winter, 73-84.
- [14] P. Jorion (1997) Value at risk, McGraw-Hill.
- [15] RiskMetrics (1996), RiskMetrics technical document, <http://www.riskmetrics.com>.
- [16] RiskMetrics (2001), Return to Risk Metrics: The Evolution of a Standard, <http://www.riskmetrics.com>.