UNIVERSITÀ DEGLI STUDI "ROMA TRE"
FACOLTÀ DI S.M.F.N.

TESI DI LAUREA IN MATEMATICA
presentata da

Valeria Romani

Stima del VaR di un portafoglio con tecniche di Importance Sampling

Relatore Prof.ssa Lucia Caramellino

Firma del Candidato Firma del Relatore

Anno Accademico 2001-2002

Classificazione AMS: 65C05, 91B28, 60F10.
Parole chiave: VaR (Valore al Rischio); approssimazione Delta-Gamma; simulazioni Monte Carlo; metodi di riduzione della varianza: *Importance Sampling*.

Sintesi

In questa tesi viene studiato un metodo Monte Carlo efficiente, di **Importance Sampling**, per il calcolo del **Valore a Rischio**, brevemente **VaR**, di un portafoglio.

Il VaR rappresenta un oggetto fondamentale per quantificare e trattare il rischio di un portafoglio. Intuitivamente, un portafoglio è l'insieme dei titoli (eventualmente derivati) nei quali si investe denaro. Supponiamo che in un mercato siano presenti m titoli, con prezzi $S_1, ..., S_m$. Un portafoglio su questi titoli è dato da una combinazione lineare di titoli derivati dai sottostanti processi di prezzo $S_1(t), ..., S_m(t)$, ad esempio opzioni call o put, opzioni digitali, o asiatiche, o quant'altro. Non entreremo qui nei dettagli; al momento, è sufficiente dire che il **valore** $V(t)$ **di un portafoglio** è tipicamente una funzione dell'istante t e dei valori osservati al tempo t dei processi di prezzo:

$$V(t) = V(t, S_1(t), \ldots, S_m(t)).$$

Sia quindi dato un portafoglio con valore $V(t)$. Fissato un istante t ed un intervallo temporale di osservazione, di ampiezza Δt, si definisce la **perdita** L come la variazione del portafoglio tra gli istanti t e $t + \Delta t$, ovvero

$$L = V(t) - V(t + \Delta t).$$

Ma cos'è il VaR? Data una probabilità p, il VaR, è l' $(1 - p)$-esimo quantile della distribuzione della perdita[1]:

$$\mathbb{P}(L > \mathrm{VaR}) = p.$$

Fissato p, il VaR rappresenta quindi la minima perdita che si può tollerare con probabilità p. Tipicamente, p è una quantità molto piccola (dell'ordine dell'1%), quindi presumibilmente la soglia definita dal VaR è piuttosto grande.

[1]In effetti, dalla definizione seguirebbe che il VaR è funzione del quantile p, quindi la notazione corretta dovrebbe essere VaR_p. Nella pratica però si omette la dipendenza dal quantile p.

In pratica il VaR dà agli utenti una misura sommaria del rischio di mercato. Ad esempio: una banca potrebbe dire che il VaR giornaliero di un portafoglio di mercato è di 35 milioni di euro a livello di confidenza del 99%. Ciò significa che c'è solo una possibilità su 100, in condizioni "normali" di mercato, che ci sia una perdita del valore del portafoglio maggiore di 35 milioni di euro.
Vediamo brevemente a cosa serve il VaR.

1. Il VaR viene utilizzato per avvertire i direttori di aziende del rischio che si corre nel commercio e nelle operazioni di investimento.

2. Il VaR può essere usato per porre posizioni limite per i commercianti e far decidere loro dove allocare le risorse limitate del capitale.

3. Il VaR viene utilizzato per un'interpretazione migliore del rischio.

Il VaR può essere adottato in massa dalle istituzioni finanziarie e dagli utenti finali preoccupati dei derivati. In generale il VaR può beneficiare ogni istituzione con esposizione al rischio finanziario: istituzioni finanziarie, regolatori, associazioni non finanziarie, asset managers.

In questo lavoro si vuole studiare un metodo numerico Monte Carlo per la valutazione del VaR.

In generale un portafoglio ha un'espressione piuttosto complicata, a volte decisamente intrattabile. Per tale ragione, il primo passo è quello di utilizzare la cosiddetta **approssimazione Delta-Gamma**, che in sostanza significa approssimare L con lo sviluppo di Taylor al secondo ordine del portafoglio:

$$L = V(t) - V(t+\Delta t) \approx -\frac{\partial V}{\partial t}\,\Delta t - \sum_{i=0}^{m} \frac{\partial V}{\partial S_i}\Big(S_i(t+\Delta t) - S_i(t)\Big)$$
$$-\frac{1}{2}\sum_{i,j=0}^{m} \frac{\partial^2 V}{\partial S_i \partial S_j}\Big(S_i(t+\Delta t) - S_i(t)\Big)\Big(S_j(t+\Delta t) - S_j(t)\Big).$$

Le quantità

$$\frac{\partial V}{\partial t}, \quad \mathrm{grad}_S V, \quad \mathrm{Hess}_S V$$

sono particolarmente importanti in Finanza, perché danno la sensibilità del portafoglio in relazione a variazioni del tempo e del prezzo, e rappresentano tre possibili **greche**: la prima prende il nome di **Theta**, la seconda è la **Delta** mentre l'ultima si chiama **Gamma**. Detto allora $\Delta S = [S(t+\Delta t) - S(t)]$ il vettore che identifica le variazioni dei prezzi (**fattori di rischio**), nell'intervallo di tempo $[t, t + \Delta t]$, l'approssimazione Delta-Gamma garantisce che

$$L \approx a_0 + a^T \Delta S + \Delta S^T A \Delta S,$$

dove

$$a_0 = -\text{Theta}\,\Delta t, \quad a = -\text{Delta}, \quad A = -\frac{1}{2}\text{Gamma}.$$

Va detto che anche il problema del calcolo delle greche è assolutamente non banale in Finanza: esse sono note solo per alcuni speciali modelli per il processo di prezzo S e per alcuni titoli derivati che compongono il portafoglio, in tutti gli altri casi sono calcolabili numericamente, con metodi più o meno efficienti. In ogni caso, qui ci poniamo nella situazione in cui i dati sono le greche ed un modello probabilistico (cioè una distribuzione) per la variazione dei prezzi ΔS. L'approssimazione Delta-Gamma consente quindi di facilitare molto le operazioni per il calcolo di L tramite simulazioni Monte Carlo. Nel corso del nostro lavoro abbiamo voluto stimare il VaR attraverso tecniche di riduzione della varianza. Facciamo un tipico esempio di Monte Carlo standard e vediamo per quale motivo è così importante poter introdurre tecniche di riduzione della varianza, per noi un metodo Monte Carlo con Importance Sampling.

Sia L la perdita. Prendendo valida per L la formula che si ottiene con l'approssimazione Delta-Gamma, e note sia le greche che una distribuzione per i fattori di rischio ΔS, la perdita L si può ritenere simulabile. Ora, fissiamo il quantile p, ad esempio $p = 0.01$: il problema è il calcolo del VaR, ossia della soglia x tale che $\mathbb{P}(L > x) = 0.01$. Tipicamente, quello che si fa è fissare qualche valore per x, stimare $\mathbb{P}(L > x)$ con metodi Monte Carlo e poi muoversi intorno ad x in modo che la stima sia vicina a 0.01. La stima Monte Carlo standard per $\mathbb{P}(L > x)$ è data dal tipico stimatore legato alla Legge dei Grandi Numeri:

$$\mathbb{P}(L > x) \approx \frac{1}{N} \sum_{k=1}^{n} I_{(L^{(k)} > x)}$$

dove N denota il numero di simulazioni, $L^{(k)}$ la k-esima simulazione della perdita L e I la funzione indicatrice. Questo algoritmo per il calcolo di $\mathbb{P}(L > x)$, anche se semplice da implementare, risulta altamente inefficiente in quanto gli eventi $\{L^{(k)} > x\}$ sono poco probabili, giacché si cerca x affinché $\mathbb{P}(L > x) = 0.01$. Inoltre, la varianza dello stimatore è dello stesso ordine della quantità da stimare. Usando il Teorema del Limite Centrale, segue allora che per avere una buona stima del VaR, saremmo costretti ad effettuare molte simulazioni e ciò comporterebbe una sostanziosa complessità computazionale, dunque una maggiore lentezza del programma. Per tale motivo introduciamo l'**Importance Sampling**. Questo metodo, attraverso la teoria delle trasformate di Laplace, opera un opportuno cambio di misura, detto **cambio di misura esponenziale**, in modo tale che, sotto questa nuova misura, gli eventi di interesse non sono più rari. Con l'Importance Sampling quindi l'evento $\{L > x\}$ non è più raro e di conseguenza non occorrono più troppe simulazioni per avere una buona stima di ciò che si cerca.

In questa tesi, seguendo i lavori [6] e [7] di Glasserman, Heidelberger e Shahabuddin, abbiamo studiato due possibili modelli per i fattori di rischio: il caso in cui ΔS assume una **distribuzione gaussiana** $\mathrm{N}(0, \Sigma)$, che è un modello a code sottili, e poi abbiamo analizzato un particolare modello con le code pesanti, ovvero con ΔS avente una distribuzione **t di Student**, decisamente più realistico in Finanza, come osservato da numerevoli autori.

Entriamo ora nei dettagli di quanto studiato in questa tesi, dividendo nei due casi di code sottili e code pesanti.

1. Modello Gaussiano

Supponiamo qui che $\Delta S \sim \mathrm{N}(0, \Sigma)$. Come già accennato, l'intento dell'Importance Sampling è quello di operare un cambio di misura esponenziale. Effettuata l'approssimazione Delta-Gamma per L, per semplificare ulteriormente i calcoli, abbiamo espresso tale approssimazione nel seguente modo:

$$L \approx a_0 + a^T \Delta S + \Delta S^T A \Delta S = a_0 + Q,$$

dove Q è una forma quadratica diagonalizzata:

$$Q = \sum_{i=1}^{m} (b_i \hat{Z}_i + \lambda_i \hat{Z}_i^2) \tag{1}$$

con $\hat{Z}_i \sim \mathrm{N}(0, 1)$ indipendenti e dove i vettori b e λ si scrivono in termini della Delta e della Gamma. L'idea dell'Importance Sampling è la seguente.

Siano date due variabili aleatorie X e Y a valori in $\Re^d$, rispettivamente con densità f_X e f_Y. Per semplicità, supponiamo che $f_Y > 0$. Fissata una funzione misurabile $g : \Re^d \to \Re^q$, allora per ogni boreliano A di $\Re^q$, possiamo scrivere

$$\mathbb{P}(g(X) \in A) = \mathbb{E}[I_{g(X) \in A}] = \mathbb{E}\left[I_{g(Y) \in A} \frac{f_X(Y)}{f_Y(Y)}\right] \tag{2}$$

Indicando con L_X e L_Y le leggi di X e Y, la (2) garantisce che la misura L_X è assolutamente continua rispetto a L_Y e la derivata di Radon-Nikodym vale

$$\frac{dL_X}{dL_Y}(y) = \frac{f_X(y)}{f_Y(y)}.$$

L'Importance Sampling consiste nell'operare un cambio di misura come quello appena visto: si parte da un evento del tipo $\{g(X) \in A\}$, che ha bassa probabilità di verificarsi, quindi si passa a considerare un'ulteriore v.a. Y tale che l'evento $\{g(Y) \in A\}$ è invece ragionevolmente probabile. Volendo stimare $\mathbb{P}(g(X) \in A)$ con metodi Monte Carlo, la (2) consente di usare un ulteriore stimatore costruito a partire da Y:

$$\mathbb{P}(g(X) \in A) = \mathbb{E}\left[I_{g(Y)\in A}\,\frac{dL_X}{dL_Y}(Y)\right] \approx \frac{1}{N}\sum_{k=1}^{N} I_{g(Y^{(k)})\in A}\,\frac{dL_X}{dL_Y}(Y^{(k)})$$

dove N denota il numero di simulazioni e $Y^{(k)}, \ldots, Y^{(k)}$ sono N copie indipendenti di Y. L'idea è quindi di scegliere la variabile aleatoria Y in modo tale che $\mathbb{P}(g(Y) \in A)$ non sia trascurabile, cosicché gli eventi $\{g(Y^{(k)}) \in A\}$ si verifichino con una certa frequenza. In tal modo, la varianza di questo stimatore risulta più piccola, a causa di una minore dispersione nei risultati.

Torniamo alla perdita $L = a_0 + Q$, dove Q è la forma quadratica presente in (1), dunque Q è funzione di un vettore aleatorio $\hat{Z}$ di gaussiane standard. Indicando con Z un ulteriore vettore gaussiano, seguendo l'idea appena vista e in particolare la (2), nel nostro caso possiamo scrivere

$$\mathbb{P}(L > x) = \mathbb{E}[I_{L>x}\ell(Z)]$$

dove $\ell(Z)$ denota la derivata di Radon-Nikodym della legge di $\hat{Z}$ rispetto alla legge di Z. Se Z ha media μ e matrice di varianza-covarianza B (e ricordando che $\hat{Z} \sim \mathrm{N}(0, I_{m\times m})$), si ha

$$\frac{dL_{\hat{Z}}}{dL_Z}(z) = \ell(z) = \frac{\exp(\frac{-1}{2}z^T z)}{(\det(B))^{-\frac{1}{2}}\exp(-\frac{1}{2}(z-\mu)^T B^{-1}(z-\mu))}.$$

Occorre ora scegliere i parametri μ e B. L'operazione che andiamo ora ad eseguire viene detta **twisting esponenziale**. Sia $\theta \geq 0$ un parametro, detto parametro *twist*, e siano:

$$\begin{cases} B(\theta) = (I - 2\theta\Lambda)^{-1} \\ \mu(\theta) = \theta B(\theta) b \end{cases},$$

dove $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$, con $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$, è la matrice degli autovalori di $\tilde{C}' A \tilde{C}$, con $\tilde{C}$ tale che $\tilde{C}\tilde{C}' = \Sigma$. Allora usando questa parametrizzazione alla media e alla matrice di varianza-covarianza, si dimostra che, per ogni θ tale che $1 - 2\theta\lambda_1 > 0$, la funzione di verosimiglianza diventa:

$$\ell(Z) = \exp\left(-\theta(b^T Z + Z^T \Lambda Z)\right) = \exp\left(-\theta Q + \psi(\theta)\right),$$

dove qui Q va intesa come funzione del vettore Z e la funzione $\psi(\theta)$ è data da:

$$\psi(\theta) = \sum_{i=1}^{m} \left(\frac{1}{2} \frac{(\theta b_i)^2}{1 - 2\theta\lambda_i} - \log(1 - 2\theta\lambda_i) \right).$$

Seguendo il procedimento brevemente riassunto in precedenza, determiniamo così un nuovo stimatore per $\mathbb{P}(L > x)$:

$$\mathbb{P}(L > x) \approx \frac{1}{N} \sum_{k=1}^{N} I_{a_0 + Q^{(k)} > x}\, \ell(Z^{(k)})$$

dove le $Z^{(k)}$ sono gaussiane indipendenti, di media $\mu(\theta)$ e matrice di varianza-covarianza $B(\theta)$, e $Q^{(k)}$ sono i valori che assume la forma quadratica Q in corrispondenza a $Z^{(k)}$. La quantità

$$T_\theta(Z) = I_{(L>x)} \cdot \ell(Z)$$

con $Z \sim \mathrm{N}(\mu(\theta), B(\theta))$, è detta **stimatore di Importance Sampling**. Osserviamo che lo stimatore di Importance Sampling è non distorto. Inoltre, si verifica che, detta $\mathbb{E}_\theta$ la media effettuata sotto la nuova misura $\mathbb{P}_\theta$, si ha

$$\mathbb{E}_{\theta_x}[L] = x.$$

Ma allora, sotto $\mathbb{P}_{\theta_x}$ l'evento $\{L > x\}$ non è più un evento raro: ci troviamo quindi nella situazione descritta in precedenza.

Da ora in poi, poiché lo stimatore di Importance Sampling dipende da θ, il nostro scopo è quello di studiare le proprietà di $T_\theta(Z)$ e in particolare di trovare il θ migliore, rendendo la varianza dello stimatore più piccola possibile. Osserviamo però che studiare direttamente la varianza è di fatto impraticabile, per via della difficile espressione che si ottiene. Per tale ragione, come normalmente si procede nell'Importance Sampling, si preferisce minimizzare non tanto la varianza ma una sua stima opportuna. Infatti, si verifica facilmente che

$$\mathrm{Var}_\theta(T_\theta(Z)) \le e^{2\psi(\theta) - 2\theta(x - a_0)} - \mathbb{P}(L > x)^2,$$

allora il θ ottimale si sceglie come quello che minimizza l'espressione a destra nella formula sopra scritta. Si dimostra che tale valore per θ risulta essere quello che risolve l'equazione

$$\psi'(\theta) = x - a_0.$$

Sia θ_x questo parametro *twist*: $\psi'(\theta_x) = x - a_0$.

Usando pesantemente la teoria delle trasformate di Laplace e sia la teoria che tecniche di dimostrazione provenienti dalle Grandi Deviazioni, è stato possibile dimostrare i seguenti risultati.

- **Comportamento asintotico in x di $\mathbb{P}(L > x)$.**

  Se $\lambda_1 > 0$ allora
  $$\lim_{x\to+\infty} \frac{\log \mathbb{P}(L > x)}{x} = -\frac{1}{2\lambda_1}$$

- **Comportamento asintotico in x del momento secondo dello stimatore di Importance Sampling.**

  Sia
  $$m_2(x, \theta) = \mathbb{E}_\theta[T_\theta^2(Z)] \leq e^{2\psi(\theta) - 2\theta(x - a_0)}$$
  il momento secondo di $T_\theta(Z)$. Allora,
  $$\lim_{x\to+\infty} \frac{\log m_2(x, \theta_x)}{x} = -\frac{1}{\lambda_1}.$$

In particolare, questi risultati consentono di affermare che lo stimatore di Importance Sampling è **asintoticamente ottimale**, nel senso che

$$\limsup_{x\to+\infty} \frac{1}{x} \log \mathbb{E}_{\theta_x}\left[T_{\theta_x}^2(Z)\right] = \inf_{\{\theta\in(0,1/(2\lambda_1))\}} \limsup_{x\to+\infty} \frac{1}{x} \log \mathbb{E}_\theta\left[T_\theta^2(Z)\right].$$

Inoltre, tutto questo ci ha permesso di evidenziare che

$$\mathrm{Var}_{\theta_x}(T_{\theta_x}(Z)) \approx e^{-2cx}, \qquad \text{con } c = \frac{1}{2\lambda_1}, \text{ se } x \to +\infty.$$

Ora, se consideriamo invece $\tilde{T}(Z)$ lo stimatore Monte Carlo standard, ovvero $\tilde{T}(Z) = I_{L>x}$, abbiamo

$$\mathrm{Var}(\tilde{T}(Z)) = \mathbb{P}(L > x) - \mathbb{P}^2(L > x) \approx \mathbb{P}(L > x),$$

dunque

$$\mathrm{Var}(\tilde{T}(Z)) \approx e^{-cx}, \qquad \text{con } c = \frac{1}{2\lambda_1}, \text{ se } x \to +\infty.$$

Ma allora la varianza dello stimatore di Importance Sampling, con la scelta $\theta = \theta_x$, converge a zero per $x \to +\infty$ più velocemente della varianza dello stimatore Monte Carlo standard.

Ricordiamo che nella pratica si è interessati a stimare $\mathbb{P}(L > x)$ per una certa gamma di valori di x, ad esempio si potrebbe essere interessati a trovare x in modo tale che $\mathbb{P}(L > x)$ cada in un certo intervallo [0.01,0.05]. Sappiamo che il VaR è definito come quantile della distribuzione della perdita L, quindi a priori non si conosce il valore di x. Procedendo con l'Importance Sampling è naturale chiedersi quale sia la sua efficienza quando usiamo il parametro *twist* θ_x per stimare $\mathbb{P}(L > y)$ per $y \neq x$. A tal proposito abbiamo mostrato che l'Importance Sampling con parametro θ_x è asintoticamente ottimale per la stima di $\mathbb{P}(L > x)$ in una vasta gamma di valori per y vicini a x. Tale proprietà è nota come **robustezza** del metodo; matematicamente si esprime nel seguente modo:

- **Robustezza.**

  Se $L = a_0 + Q, \quad \lambda_1 > 0$ *e* $y_x \to +\infty$*, in modo tale che*

  $$\limsup_{x\to+\infty}\left(\frac{x}{y_x}\right) < +\infty,$$

  allora il parametro twist θ_x *definito da* $\psi'(\theta_x) = x - a_0$ *è asintoticamente ottimale anche per la stima di* $\mathbb{P}(L > y_x)$*, cioè:*

  $$\lim_{x\to+\infty}\frac{\log m_2(y_x,\theta_x)}{y_x} = -\frac{1}{\lambda_1}.$$

In seguito abbiamo assunto un nuovo punto di vista: abbiamo fissato x e studiato il comportamento asintotico al tendere all'infinito del numero m di fattori di rischio. Si tratta questo di un problema interessante ad esempio per le banche, che possiedono portafogli costituiti da un alto numero di titoli. In tal caso, usando teoremi di convergenza specifici, è stato possibile dimostrare il prossimo risultato. Nel seguito, sottolineeremo la dipendenza da m piuttosto che da x.

- **Comportamento asintotico per $m \to \infty$.**

  Sia

  $$x_m = \sum_{i=1}^{m}\lambda_i + x_{std}\sqrt{\sum_{i=1}^{m}(b_i^2 + 2\lambda_i^2)} \tag{3}$$

  con x_{std} *numero fissato, e* θ_m *il parametro twist in corrispondenza al valore* $x = x_m$*:* $\psi'_m(\theta_m) = x_m$*. Se*

  $$\lim_{m\to\infty}\frac{\left[\sum_{i=1}^{m}\left(|b_i|^3 + |\lambda_i|^3\right)\right]^2}{\left[\sum_{i=1}^{m}\left(b_i^2 + \lambda_i^2\right)\right]^3} = 0,$$

  allora

  $$\lim_{m\to\infty}\mathbb{P}(Q_m > x_m) = 1 - \Phi(y) \qquad e \;\lim_{m\to\infty} m_2(x_m,\theta_m) = e^{y^2}(1 - \Phi(2y)),$$

  dove Φ *denota la funzione di distribuzione di una normale standard.*

Abbiamo poi considerato una combinazione di Importance Sampling con un campionamento stratificato, allo scopo di ottenere un'ulteriore riduzione della varianza. Risultati numerici (Glasserman, Heidelberger e Shahabuddin [6]) mostrano che l'efficienza dell'Importance Sampling combinato con un campionamento stratificato aumenta al diminuire di $\mathbb{P}(L > x)$. In breve, "stratificare" significa considerare una variabile Y a valori in uno spazio E ed $\{S_j\}_{j=1,\dots,k}$ una partizione (**strati**) di E. Ovviamente, se si sceglie $k = 1$ significa non procedere ad alcuna forma di stratificazione, quindi i risultati che si ottengono rientrano nella casistica studiata in precedenza. Indichiamo con p_j la probabilità che la variabile aleatoria ausiliaria Y appartenga allo strato S_j. Consideriamo $n_j = [q_j\, n]$ campionamenti di Y nello strato S_j (dunque q_j è la proporzione dei campionamenti nello strato S_j) e definiamo

$$F_{jn}(x) = 1 - \frac{1}{n_j} \sum_{i=1}^{n_j} \ell_{ij}\, I L_{ij} > x \quad \text{e} \quad F_n(x) = \sum_{j=1}^{k} F_{jn}(x),$$

dove L_{ij} denota la perdita del campione i nello strato j e ℓ_{ij} la relativa verosimiglianza osservata. È facile vedere che la distribuzione (aleatoria) F_n stima la distribuzione vera F della perdita L. Ora, fissata una probabilità p, se definiamo il quantile

$$\alpha = F^{-1}(1 - p),$$

ed il quantile (aleatorio)

$$\alpha_n = \inf \{x : F_n(x) \geq 1 - p\},$$

ci si chiede se e come α_n stima α. È evidente l'importanza di tali questioni per il nostro problema del calcolo del VaR: α è proprio il VaR! I risultati che si ottengono si possono riassumere come segue. Indicheremo con $\sigma^2(\alpha)$ la seguente quantità:

$$\sigma^2(\alpha) = \sum_{j=1}^{k} p_j^2\, \sigma_j^2(\alpha)/q_j,$$

dove $\sigma_j^2(\alpha)$ è la varianza di $\ell_{ij}\, I_{L_{ij}>x}$ e $\ell_{ij}\, I_{L_{ij}>x}$ sono variabili aleatorie i.i.d. sotto $\mathbb{P}_\theta$. Osserviamo che qui ℓ_{ij} viene calcolato sotto un θ generico positivo.

Esponiamo alcuni risultati asintotici, notando che sono stati ottenuti sotto la misura $\mathbb{P}_\theta$.

- **Comportamento asintotico dei quantili.**

  Supponiamo che la densità f di L sia continua e positiva in un intorno di α, che $\mathbb{E}_\theta[\ell_{ij}^3] < \infty$ e che $\sigma(\alpha) > 0$. Allora,

- **Teorema Limite Centrale:**

$$\sqrt{n}(\alpha_n - \alpha) \longrightarrow_{n\to+\infty} \mathrm{N}\left(0, \frac{\sigma^2(\alpha)}{f^2(\alpha)}\right), \ \textit{in legge.}$$

- **Convergenza forte:**

$$\alpha_n \longrightarrow_{n\to\infty} \alpha \qquad q.c.$$

Il risultato di convergenza forte è originale, essendo stato dimostrato per la prima volta in questa tesi.

2. Modello di Student

In questo caso abbiamo eseguito uno studio parallelo a quello fatto nel caso gaussiano, sviluppando però un altro modello per i fattori di rischio ΔS. Tale nuovo modello è basato su una distribuzione con le "code spesse": abbiamo preso in considerazione in particolare la distribuzione t di Student. Dopo aver dato alcuni cenni proprio su tale legge, assumendo ancora una volta sempre un portafoglio di tipo quadratico, abbiamo affrontato il problema della trasformata di Laplace per variabili aleatorie di questo tipo. In verità, le trasformate di Laplace in questo caso non sono propriamente un "problema": semplicemente, non esistono. Quindi lo studio parallelo cui accennavamo è stato effettuato tramite opportune manipolazioni su questo tipo di variabili aleatorie. Ma, una volta effettuate, abbiamo potuto davvero riproporre le tecniche viste in precedenza. Abbiamo poi sviluppato uno studio volto ad analizzare le proprietà del conseguente stimatore di Importance Sampling. Vediamo ora un po' più in dettaglio quanto è stato fatto.

Anche qui supponiamo $L \approx a_0 + Q$, con Q funzione quadratica con forma diagonalizzata

$$Q = \sum_{j=1}^{m}(\lambda_j X_j^2 + b_j X_j),$$

dove però questa volta assumiamo le X_j con distribuzione di Student. Abbiamo ora a che fare con un punto di vista finanziario più realistico, ovvero quello in cui i fattori di rischio $\Delta S(= X)$ sono delle variabili aleatorie con una distribuzione dalle code spesse. La non-esistenza della trasformata di Laplace di ciascuna X_j non consente di operare da subito un cambio di misura esponenziale, con l'introduzione del parametro *twist*, come visto nel caso gaussiano. Ciò nonostante anche qui, ci siamo proposti di fare un cambio di misura. Prima però abbiamo operato un cambio di variabile che ci ha consentito di lavorare su una quantità alternativa a Q. Vediamo come.

Una variabile aleatoria X su $\Re^m$ con distribuzione t di Student con ν gradi di libertà e matrice associata Σ, si può rappresentare nel seguente modo:

$$X \stackrel{d}{=} \frac{Z}{\sqrt{\frac{Y}{\nu}}}$$

dove Z e Y sono variabili aleatorie indipendenti, con Z gaussiana su $\Re^m$ a media nulla e con matrice di varianza-covarianza Σ, e Y di tipo chi-quadro con ν gradi di libertà. La densità di probabilità di X è data dalla formula

$$f_X(x) = \frac{\Gamma\left(\frac{\nu}{2}+\frac{n}{2}\right)}{(\nu\pi)^{\frac{n}{2}}\Gamma(\nu/2)\det(\Sigma)^{\frac{1}{2}}}\left(1+\frac{x^t\Sigma^{-1}x}{\nu}\right)^{-\frac{\nu}{2}-\frac{n}{2}}, \quad x \in \Re^m,$$

dunque le code della distribuzione vanno a zero a velocità polinomiale (da cui "code spesse"). A questo punto abbiamo considerato il suddetto cambio di variabile. Innanzitutto abbiamo denotato con $\tilde{Q}$ la quantità

$$\tilde{Q} = \sum_{i=1}^{m}\left(\lambda_i \frac{Z_i^2}{Y/\nu} + b_i \frac{Z_i}{\sqrt{Y/\nu}}\right),$$

osservando che $Q \stackrel{d}{=} \tilde{Q}$. Definendo

$$Q_x := \left(\frac{Y}{\nu}\right)(\tilde{Q} - x),$$

si ha ovviamente

$$\mathbb{P}(Q \leq x) = \mathbb{P}(Q_x \leq 0) \equiv F_x(0).$$

Tale semplice proprietà si rivela cruciale: si mostra facilmente che la distribuzione di Q_x non ha le code grasse, che per di più vanno a zero a velocità esponenziale, proprietà che finalmente consente di usare le trasformate di Laplace e le conseguenti proprietà.

Fatto questo, e considerata sempre l'approssimazione Delta-Gamma per il portafoglio, è stato possibile definire un cambio di misura esponenziale, per poi procedere a sviluppare un metodo di Importance Sampling. Riassumiamo qui il cambio di misura e le principali proprietà.

- **Cambio di misura nel modello di Student.**

  Siano ϕ_x e ϕ_Y le trasformate di Laplace rispettivamente di Q_x e di Y. Siano inoltre $\psi_x = \log\phi_x$ e $\psi_Y = \log\phi_Y$. Per $\theta < \frac{1}{2\lambda_1}$ sia

  $$\alpha(\theta) = -\frac{\theta x}{\nu} + \frac{2}{2\nu}\sum_{j=1}^{m}\frac{\theta^2 b_j^2}{1-2\lambda_i}.$$

- **Cambio di misura esponenziale.**
  Per ogni θ tale che $\alpha(\theta) < 1/2$, allora la misura $\mathbb{P}_\theta$ definita da

$$d\mathbb{P}_\theta = \exp(\theta Q_x - \psi_x(\theta))d\mathbb{P}$$

  è una misura di probabilità e

$$\mathbb{P}(L > y) = \mathbb{E}_\theta \left[e^{-\theta Q_x + \psi_x(\theta)} I_{L>y} \right] = \mathbb{E}_\theta \left[e^{-\theta(\frac{Y}{\nu})(Q-x)+\psi_x(\theta)} I_{L>y} \right].$$

  Inoltre, se $\mathbb{P}_\alpha$ è la misura definita da

$$\frac{d\mathbb{P}_\alpha}{d\mathbb{P}} = e^{\alpha Y - \psi_Y(\alpha)}$$

  con $\alpha < \frac{1}{2}$ e se $\mathbb{P}_\theta$ è la misura definita sopra, allora

$$\mathbb{P}_\theta = \mathbb{P}_{\alpha(\theta)}.$$

- **Distribuzione di X sotto la nuova misura $\mathbb{P}_\theta$.**
  Sotto $\mathbb{P}_\theta$, X ha la stessa distribuzione di $\frac{Z}{\sqrt{\frac{Y}{\nu}}}$, dove Y ha legge $\Gamma\left(\frac{\nu}{2}, \frac{1-2\alpha(\theta)}{2}\right)$ e condizionatamente a Y, le variabili aleatorie $Z_1, ..., Z_m$ sono indipendenti e gaussiane con medie e varianze date da

$$\mu_j(\theta) = \frac{\theta b_j \sqrt{\frac{Y}{\nu}}}{1 - 2\theta\lambda_j}, \quad e \quad \sigma^2(\theta) = \frac{1}{1 - 2\theta\lambda_j}, \quad j = 1, \dots, m.$$

Questi risultati si rivelano particolarmente utili, ad esempio per le simulazioni perché danno una semplice regola di simulazione del vettore X di Student sotto la nuova misura.

Ora, presa $\mathbb{P}_\theta$ come nuova misura di riferimento, possiamo definire anche nel modello t di Student uno stimatore di Importance Sampling e, come fatto anche nel caso gaussiano, ne abbiamo studiato alcune proprietà di rilievo. Tenendo conto del cambio di misura, lo stimatore in questione è basato sulla quantità

$$T_\theta(X) = I_{Q>x}\, e^{-\theta Q_x + \psi_x(\theta)}.$$

Anche qui si propone ovviamente il problema della scelta del θ ottimale. Procedendo come nel caso gaussiano, si considera dapprima la seguente stima del momento secondo $m_2(\theta, x)$:

$$m_2(\theta, x) = \mathbb{E}_\theta \left[e^{-2\theta Q_x + 2\psi_x(\theta)} I_{(Q>x)} \right] = \mathbb{E} \left[e^{-\theta Q_x + \psi_x(\theta)} I_{(Q>x)} \right] \le e^{\psi_x(\theta)},$$

e poi l'ottimizzazione su θ si ottiene minimizzando la quantità a destra dell'espressione precedente: il θ_x ottimale risolve l'equazione

$$\psi_x'(\theta_x) = 0.$$

Osserviamo che qui si tratta di un minimo vincolato, quindi non è proprio detto che θ_x sia un punto stazionario per ψ_x. Ed in effetti, nelle simulazioni numeriche che abbiamo effettuato in alcuni casi abbiamo avuto problemi nel calcolo del θ ottimale.

Le proprietà asintotiche, per $x \to \infty$, relative al modello di Student mostrano come, sotto ipotesi opportune, lo stimatore di Importance Sampling verifichi la cosiddetta proprietà di **limitatezza dell'errore relativo**, cioè:

$$\limsup_{x \to +\infty} \frac{m_2(\theta, x)}{\mathbb{P}(Q > x)^2} < \infty. \tag{4}$$

Riassumiamo qui di seguito i principali risultati.

- **Comportamento asintotico del momento secondo, I.**

  Se $\lambda_1 > 0$, *per ogni* θ *positivo fissato tale che* $\psi_x(\theta) < 0$, *esiste una costante* $c(\theta)$ *per la quale,*
  $$m_2(\theta, x) \leq c(\theta)\mathbb{P}(Q > x)x^{-\frac{\nu}{2}},$$
  per ogni x *grande abbastanza.*

- **Limitatezza dell'errore relativo, I.**

  Se $\lambda_j > 0$, *per ogni* $j = 1, ..., m$, *allora lo stimatore*
  $$T_\theta(X) = e^{-\theta Q_x + \psi_x(\theta)}\, I_{(Q>x)}$$
  di $\mathbb{P}(Q > x)$ *ha errore relativo limitato, cioè vale la (4).*

- **Comportamento asintotico del momento secondo, II.**

  Se θ_x *è tale che* $\psi_x'(\theta_x) = 0$ *e* λ_1 *è positivo, allora esiste una costante* K, *per la quale,*
  $$m_2(\theta_x, x) \leq K\mathbb{P}(Q > x)x^{-\frac{\nu}{2}},$$
  per ogni x *grande abbastanza.*

- **Limitatezza dell'errore relativo, II.**

  Se θ_x *è tale che* $\psi_x'(\theta_x) = 0$ *e se* $\lambda_j > 0$, *per ogni* $j = 1, ..., m$, *allora anche*
  $$T_{\theta_x}(X) = e^{-\theta Q_x + \psi_x(\theta_x)}\, I_{(Q>x)}$$
  ha errore relativo limitato, cioè vale la (4).

Una volta studiati teoricamente i due modelli di riferimento, abbiamo effettuato uno studio numerico per verificare l'efficienza anche pratica dell'Importance Sampling.

3. Risultati numerici

Nell'ultima parte del lavoro abbiamo riportato alcuni risultati numerici. In particolare, facendo riferimento ai portafogli indicati negli articoli di Glasserman Heidelberger e Shahabuddin [6] e [7], attraverso il pacchetto statistico dell'applicativo *Matlab* e attraverso il linguaggio *Pascal*, abbiamo effettuato simulazioni Monte Carlo standard e Monte Carlo con Importance Sampling.

Abbiamo considerato portafogli formati da opzioni di acquisto (**call**) e di vendita (**put**). Le formule delle greche sono state prese dalle formule provenienti dal **modello di Black e Scholes**, effettuando ovviamente delle opportune ulteriori approssimazioni. Abbiamo eseguito i programmi sia nel caso gaussiano, sia in quello t di Student, operando 80000 simulazioni, e abbiamo potuto appurare l'efficienza del metodo con Importance Sampling rispetto a Monte Carlo standard, soprattutto in termini di velocità di esecuzione del programma, che sottolineiamo è sostanzialmente dello stesso ordine per entrambe le procedure.

Come già detto, abbiamo individuato la funzione quadratica Q, nella sua forma diagonalizzata per poi calcolare numericamente $\mathbb{P}(Q+a_0 > x)$ e relativa varianza, applicando nei due casi, come detto, sia Monte Carlo standard che Monte Carlo con Importance Sampling. Riguardo proprio a questo metodo abbiamo dovuto cercare il valore ottimale θ_x, di cui abbiamo ampiamente parlato, risolvendo l'equazione $\psi'(\theta) = x - a_0$ oppure $\psi'_x(\theta) = 0$. Per tale motivo, nei nostri programmi abbiamo utilizzato il metodo di Newton per la ricerca degli zeri di una funzione. In un portafoglio in particolare (portafoglio a.7), a causa della sua complessità, per individuare θ_x, abbiamo dovuto ricorrere a una funzione dell'applicativo *Matlab*.

Riassumiamo nelle seguenti tabelle i risultati ottenuti.

3.1 Modello gaussiano

I portafogli considerati sono:

1. **Portafoglio a.1**, maturità $T = 0.5$ anni, costituito da 10 *call* e 5 *put at-the-money* in posizione corta su ogni titolo.

2. **Portafoglio a.2**, maturità $T = 0.5$ anni, costituito da 10 *call* e 5 *put at-the-money* in posizione corta sui primi 5 titoli, 10 *call at-the-money* in posizione lunga e 5 *put at-the-money* in posizione corta, sui rimanenti 5 titoli.

3. **Portafoglio a.3**, maturità $T = 0.10$ anni, per il resto identico al portafoglio a.1.

4. **Portafoglio a.4**, maturità $T = 0.10$ anni, per il resto identico al portafoglio a.2.

5. **Portafoglio a.5** (*Delta-Hedged*), maturità $T = 0.10$ anni, costituito da 10 *call at-the-money* in posizione corta sui primi 5 titoli, 5 *call at-the-money* in posizione lunga sui rimanenti 5 titoli. La posizione e il numero di *put* sui primi e sui secondi 5 titoli sono scelti in modo tale che Delta è uguale a zero.

6. **Portafoglio a.6**, equivalente al portafoglio a.5, ma nel quale abbiamo applicato l'IS solo sui titoli corrispondenti agli autovalori positivi.

7. **Portafoglio a.7**, maturità $T = 0.5$ anni, costituito da 50 *call at-the-money* e 50 *put at-the-money* in posizione corta su 10 titoli sottostanti con valore iniziale $S = [100, 50, 30, 100, 80, 20, 50, 200, 150, 10]$.

   La quantità x_{std} presente nelle tabelle è presa dalla formula (3).

   (a) Monte Carlo standard

Portafoglio	x_{std}	$\mathbb{P}(L > x)$	Varianza
a.1	2.59	1.00%	0.099
a.2	2.30	1.01%	0.010
a.3	2.75	1.15%	0.011
a.4	2.32	0.92%	0.009
a.5	2.89	1.09%	0.011
a.7	3.40	1.15%	0.012

   Tabella 1: Monte Carlo standard-caso gaussiano.

   Da questi primi risultati notiamo che la varianza è molto alta: dello stesso ordine della probabilità. Questo implica che gli intervalli di confidenza sono molto grandi.

   (b) Monte Carlo con Importance Sampling

Portafoglio	x_{std}	$\mathbb{P}(L > x)$	Varianza
a.1	2.59	1.00%	0.00033
a.2	2.30	1.00%	0.00029
a.3	2.75	1.11%	0.00050
a.4	2.32	1.00%	0.00031
a.5	2.89	1.11%	0.00055
a.6	2.89	1.12%	0.00073
a.7	3.20	1.11%	0.00331

Tabella 2: Monte Carlo con Importance Sampling-caso gaussiano.

Osserviamo che nella Tabella 2, le varianze hanno subito una diminuizione dell'ordine di 10^{-2} ad esclusione dell'ultimo portafoglio, dove comunque c'è stata una diminuizione pari a 10^{-1} (ricordiamo che per il portafoglio a.7, ci sono state delle difficoltà per il calcolo del θ ottimale. Per tale ragione abbiamo dovuto utilizzare una funzione di *Matlab*, in quanto si trattava di calcolare un minimo vincolato che numericamente è difficle).

3.1 Modello di Student

I portafogli considerati sono:

1. **Portafoglio b.1**, maturità $T = 0.5$ anni, costituito da 10 *call* e 5 *put at-the-money* in posizione corta su ogni titolo.

2. **Portafoglio b.2**, maturità $T = 0.5$ anni, costituito da 10 *call* e 5 *put at-the-money* in posizione lunga su ogni titolo.

3. **Portafoglio b.3**, maturità $T = 0.10$ anni, equivalente al portafoglio b.1.

4. **Portafoglio b.4**, maturità $T = 0.10$ anni, equivalente al portafoglio b.2.

5. **Portafoglio b.5** (*Delta-Hedged*), maturità $T = 0.10$ anni, equivalente al portafoglio b.3, ma con la posizione e il numero di *put* sui 10 titoli ottenuti ponendo Delta uguale a zero.

Nelle tabelle che seguono, indichiamo con Varianza(L), la varianza relativa alla quantità $\mathbb{P}(L > x)$, Varianza(Q), quella di $\mathbb{P}(Q + a_0 > x)$.

(a) Monte Carlo standard

Portafoglio	x	$\mathbb{P}(L > x)$	Varianza(L)	$\mathbb{P}(Q + a_0 > x)$	Varianza(Q)
b.1	311	1.05%	0.0103	1.20%	0.0118
b.2	145	0.96%	0.0094	1.25%	0.0124
b.3	469	0.95%	0.0094	1.57%	0.0154
b.4	148.8	1.00%	0.0101	0.90%	0.0088
b.5	617	1.02%	0.0101	1.67%	0.0164

Tabella 3: Monte Carlo standard-caso di Student.

(b) Monte Carlo con Importance Sampling

Portafoglio	x	$\mathbb{P}(L > x)$	Varianza(L)	$\mathbb{P}(Q + a_0 > x)$	Varianza(Q)
b.1	311	1.01%	0.0010	1.16%	0.0015
b.2	145	1.02%	0.0016	1.30%	0.0022
b.3	469	0.99%	0.0098	1.60%	0.0157
b.4	148.8	0.97%	0.0033	0.86%	0.0028
b.5	617	1.05%	0.0005	1.69%	0.0014

Tabella 4: Monte Carlo con Importance Sampling-caso di Student.

Dai risultati della Tabella 4 osserviamo che nel modello di Student, abbiamo ottenuto una diminuizione della varianza dell'ordine di 10^{-1}, tranne che per il portafoglio b.3, dove non riuscendo a trovare il θ ottimale, nemmeno con *Matlab*, il programma pone $\theta = 0$ e di conseguenza, poiché un metodo Monte Carlo con Importance Sampling con $\theta = 0$ non è altro che un metodo Monte Carlo standard, le varianze ottenute hanno lo stesso ordine.
In generale comunque in entrambi i modelli, confrontando il metodo Monte Carlo standard con il metodo Monte Carlo con Importance Sampling, si ha una diminuizione della varianza nel secondo caso.

Riassumiamo infine brevemente com'è organizzata questa tesi.

I Capitoli 1 e 2 sono introduttivi, il primo sulle trasformate di Laplace e sulle Grandi Deviazioni, il secondo sul calcolo stocastico e sulla finanza. La bibliografia di riferimento di questi due capitoli è: per il Capitolo 1, Baldi [1], Bucklew [2], Dembo e Zeutoni [5], Grimmett[11]; per il Capitolo 2, Baldi [1], Jorion [13], Lamberton e Lapeyre [14].

Nel Capitolo 3 abbiamo analizzato il metodo dell'Importance Sampling nel modello gaussiano. In particolare: il cambio di misura esponenziale, lo stimatore di Importance Sampling, la proprietà dell'ottimalità asintotica, la robustezza del metodo, il campionamento stratificato e il comportamento asintotico in relazione al numero dei fattori di rischio. La bibliografia di riferimento è: Bucklew [2], Chung [3], Glasserman, Heidelberger e Shahabuddin[6], [8], [9] e [10], Jorion [13].

Nel Capitolo 4 abbiamo trattato il modello di Student, in particolare: le generalità sulla distribuzione t di Student, la rappresentazione in termini di code sottili, il cambio di misura esponenziale, l'analisi dello stimatore e la robustezza del metodo. I relativi riferimenti bibliografici sono: Glasserman, Heidelberger e Shahabuddin [6], [7] e [9].

Infine nell'ultimo capitolo, ci siamo occupati della parte numerica, facendo simulazioni Monte Carlo standard e Monte Carlo con Importance Sampling. Ci siamo riferiti alla seguente bibliografia: Glasserman, Heidelberger e Shahabuddin [6] e [7], Imhof [12].

Bibliografia

[1] P. Baldi, *Equazioni differenziali stocastiche e applicazioni.* Pitagora Editrice, 2000.

[2] J. A. Bucklew, *Large deviation techniques in decision, simulation, and estimation.* John Wiley & Sons, Inc., New York, 1990.

[3] K. Chung, *A course in probability theory.* Academic Press, New York, second edition, 1974.

[4] G. Dall'Aglio, *Calcolo delle probabilità.* Zanichelli, 1991.

[5] A. Dembo & O. Zeitouni, *Large deviation techniques.* Jones and Barlett Publishers, 1993.

[6] P. Glasserman, P. Heidelberger, P. Shahabuddin, *Variance reduction techniques for estimating value-at-risk.* Computer Science/Mathematics, New York, 1999.

[7] P. Glasserman, P. Heidelberger, P. Shahabuddin, *Portfolio value-at-risk with heavy-tailed risk factors.* Columbia business school, 2000.

[8] P. Glasserman, P. Heidelberger, P. Shahabuddin, *Importance Sampling and stratification for Value-at-Risk.* Y.S. Abu-Mostafa, B. LeBaron, A.W. Lo, A.S. Weigend editors, Computational Finance 1999, MIT Press, Cambridge, MA, 1999.

[9] P. Glasserman, P. Heidelberger, P. Shahabuddin, *Asymptotically optimal Importance Sampling and stratification for pricing path-dependent options.* Mathematical Finance, 9:117-152, 1999.

[10] P. Glasserman, P. Heidelberger, P. Shahabuddin, *Stratification issues in estimating value-at-risk.* In Proceedings of the 1999 Winter Simulation Conference, 1999.

[11] G. R. Grimmett, D. R. Stirzaker, *Probability and random processes*. Oxford university press, second edition, 1995.

[12] J. Imhof, *Computing the distribution of quadratic forms in normal variables*. Biometrika, 48:419-426, 1961.

[13] P. Jorion, *Value at Risk*. Mc Graw Hill, New York, 1997.

[14] D. Lamberton & B. Lapeyre, *Introduction to stocastic calculus applied to finance*. Chapman & Hall, Berlin, Heidelberg, New York, 1996.