

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

SCUOLA DI SCIENZE
Corso di Laurea Triennale in Matematica

Misure di rischio nelle scienze sociali

Tesi di Laurea in Probabilità e Statistica Matematica

Relatore:
Chiar.mo Prof.
Andrea Pascucci

Presentata da:
Federico Falcone

Anno Accademico 2024/2025

*A tutte le persone che mi vogliono bene
ed a tutte quelle che me ne vorranno...*

Introduzione

In questa tesi esamineremo alcune misure di rischio ed il loro legame con le scienze sociali concentrandoci nello specifico sulle misure di rischio al ribasso. Analizzeremo il loro uso in economia e finanza ma vedremo che esse vengono utilizzate anche in sociologia. Daremo un quadro generale dell'atteggiamento che un individuo può avere nei confronti del rischio introducendo la teoria decisionale della dominanza stocastica con un accenno a quella dell'utilità: le due teorie sono infatti strettamente collegate tra loro. La teoria della dominanza stocastica in particolare è una teoria utilizzata per prendere decisioni in condizioni di rischio dove la singola preferenza dell'individuo può essere rappresentata attraverso una sua intrinseca funzione di utilità. Ciascuna di queste esprime una caratteristica dell'individuo in termini di relazione di preferenza e di rapporto con il rischio. Le relazioni di preferenza possono essere anche descritte più semplicemente da tre ordini di dominanza stocastica, ognuno relativo ad un determinato individuo con una certa funzione di utilità. Successivamente tratteremo, attraverso teoremi e proposizioni, le principali proprietà che definiscono una "buona" misura di rischio ossia convessità e coerenza: esse sono verificate attraverso specifici assiomi. Per quanto riguarda la coerenza faremo anche vedere la sua equivalenza con un insieme di rischi accettabili. Passeremo quindi ad esaminare le principali misure di rischio al ribasso, come il valore a rischio, il valore a rischio condizionale ed i momenti parziali inferiori. Essendo il valore a rischio condizionale una variante del valore a rischio, entrambe hanno gli stessi metodi di calcolo: un metodo parametrico (metodo di varianza-covarianza) e due non parametrici (metodo di simulazione storica e di Monte Carlo). La principale differenza tra le due misure di rischio risiede nel fatto che esse si concentrano su parti diverse della distribuzione e ciò fa sì che abbiano differenti criticità e punti di forza. Un'altra differenza è rappresentata dalla coerenza: il valore a rischio non è coerente, mentre, sotto particolari condizioni, il valore a rischio condizionale lo è e questo porta a preferire uno all'altro in determinate situazioni. Per quanto riguarda i momenti parziali inferiori, vedremo come essi rientrano in parte anche nella categoria delle misure di povertà poichè possono fornire un'indicazione di vulnerabilità alla povertà in un gruppo di persone. L'ultima sezione della tesi è dedicata all'analisi di alcuni indici importanti in economia e finanza quali gli indici di Sharpe (ex ante ed ex post) e gli indici alpha e beta. Mentre i primi fungono da indicatori di convenienza per un investimento (la versione ex ante viene usata prima di

effettuare un investimento mentre quella ex post viene usata a posteriori) ed hanno un legame più immediato con il rischio, i secondi sono usati per valutare il rendimento e la volatilità di un portafoglio di investimenti rispetto ad un indice di riferimento fissato ed in generale hanno un legame più indiretto con esso.

Indice

| | |
|---|-----------|
| Introduzione | i |
| 1 Dominanza stocastica | 1 |
| 2 Introduzione alle misure di rischio | 9 |
| 2.1 Misure di rischio coerenti | 9 |
| 2.1.1 Equivalenza con insieme di rischi accettabili | 11 |
| 2.2 Misure di rischio convesse | 12 |
| 3 Misure di rischio al ribasso | 15 |
| 3.1 Valore a rischio | 16 |
| 3.1.1 Metodo di varianza-covarianza | 17 |
| 3.1.2 Metodo di simulazione storica | 20 |
| 3.1.3 Metodo di simulazione Monte Carlo | 23 |
| 3.2 Valore a rischio condizionale | 26 |
| 3.3 Momenti parziali inferiori | 35 |
| 4 Indici di rischio | 43 |
| 4.1 Indice di Sharpe | 43 |
| 4.2 Indici alpha e beta | 48 |
| Appendice | 51 |
| Elenco delle figure | 59 |
| Bibliografia | 61 |

Capitolo 1

Dominanza stocastica

La teoria della dominanza stocastica è una teoria decisionale di ordinamento stocastico che permette di prendere decisioni in condizioni di rischio ed incertezza ed affermare, date due lotterie (o prospetti) μ e ν , ossia due distribuzioni associate a degli esiti, quale delle due è preferita all'altra. Queste decisioni si basano su un sistema di preferenze di esiti e di atteggiamenti nei confronti del rischio da parte dei vari investitori.

Nonostante ci siano più ordini di dominanza stocastica, ci concentreremo solo sui primi due e daremo solo la definizione di quello di terzo ordine (Unser, 1998; Ogryczak, Opolska-Rutkowska, 2004; Ma, Wong, 2010 [27][16][11]): i primi due ordini definiscono un ordine di preferenza che si basa sulla gestione del rischio e sui possibili guadagni.

Definizione 1.1 (Dominanza stocastica di primo ordine). Siano X, Y due v.a. e $F_X(x)$, $F_Y(x)$ rispettivamente le loro CDF. Si dice che X domina Y sotto la *dominanza stocastica di primo ordine* (si indica $X \succeq_{FSD} Y$) se

$$F_X(x) \leq F_Y(x),$$

per ogni $x \in \mathbb{R}$.

Dalla definizione di funzione di ripartizione, la dominanza stocastica di primo ordine ci dà un'indicazione sulla probabilità di ottenere un rendimento inferiore ad una certa soglia di riferimento x per ogni suo valore. Se quindi vale $X \succeq_{FSD} Y$ si preferirà X ad Y in ogni modello, a prescindere dal livello di rischio considerato, in quanto la distribuzione associata ad X avrà una probabilità minore rispetto a quella associata ad Y di scendere sotto un certo valore fissato.

La dominanza stocastica di secondo ordine riguarda invece le decisioni prese in condizioni di rischio.

Definizione 1.2 (Dominanza stocastica di secondo ordine). Siano X, Y due v.a. e $F_X(x)$, $F_Y(x)$ rispettivamente le loro CDF. Si dice che X domina Y sotto la *dominanza stocastica di secondo ordine* (si indica $X \succeq_{SSD} Y$) se

$$\int_{-\infty}^x F_X(t)dt \leq \int_{-\infty}^x F_Y(t)dt,$$

per ogni $x \in \mathbb{R}$.

Se $X \succeq_{SSD} Y$ allora X verrà preferita ad Y da tutti i modelli avversi al rischio che si concentrano su grandi esiti.

Esempio 1.3. Supponiamo X, Y v.a. entrambe con distribuzione normale. Se Y ha una maggiore varianza o un minore valore atteso rispetto ad X , allora $X \succeq_{SSD} Y$.

Osservazione 1.4. La definizione precedente si può anche dare in termini di valore atteso condizionato sfruttando la seguente uguaglianza (Ogryczak, Ruszczyński, 1999 [17]):

$$\int_{-\infty}^x F_X(t)dt = P(X \leq x)\mathbb{E}[x - X|X \leq x] = \mathbb{E}[\max\{x - X, 0\}].$$

Paragonando le definizioni di dominanza stocastica di primo e secondo ordine avremo la relazione

$$X \succeq_{FSD} Y \Rightarrow X \succeq_{SSD} Y.$$

La dominanza stocastica di primo e secondo ordine è inoltre preservata quando la v.a. è moltiplicata per scalari o sommata ad altre v.a (Wolfstetter, 1996 [30]).

Proposizione 1.5. Date tre v.a. non negative X, Y e W e due scalari $a > 0$ e $b \geq 0$ allora

$$\begin{aligned} X \succeq_{FSD} Y &\Rightarrow (aX + bW) \succeq_{FSD} (aY + bW), \\ X \succeq_{SSD} Y &\Rightarrow (aX + bW) \succeq_{SSD} (aY + bW). \end{aligned}$$

Alternativamente la dominanza stocastica può essere espressa anche tramite i quantili. Ricordando che $q_\alpha(X) = \inf\{x \in \mathbb{R} : F_X(x) \geq \alpha\}$, allora X domina Y sotto la FSD se e solo se

$$q_\alpha(X) \geq q_\alpha(Y),$$

per ogni $\alpha \in (0, 1)$. Mentre X domina Y sotto la SSD se e solo se

$$\int_0^\beta q_\alpha(X)d\alpha \geq \int_0^\beta q_\alpha(Y)d\alpha,$$

per ogni $\beta \in (0, 1]$.

Diamo ora la definizione di dominanza stocastica di terzo ordine che ci sarà utile in seguito.

Definizione 1.6 (Dominanza stocastica di terzo ordine). Siano X, Y due v.a. e $F_X(x), F_Y(x)$ rispettivamente le loro CDF. Si dice che X domina Y sotto la *dominanza stocastica di terzo ordine* (si indica $X \succeq_{TSD} Y$) se

$$\int_{-\infty}^x \int_{-\infty}^v F_X(t) dt dv \leq \int_{-\infty}^x \int_{-\infty}^v F_Y(t) dt dv,$$

con $t, v \in \mathbb{R}$ e vale inoltre

$$\mathbb{E}[F_X] \geq \mathbb{E}[F_Y].$$

La rappresentazione di una relazione di preferenza su un insieme S è data da quella che si chiama *funzione di utilità*. Essa può essere vista come una funzione che associa ad ogni elemento x in S un determinato valore di utilità (ossia un indice di soddisfazione individuale) in \mathbb{R} (Mas-Colell, Whinston, Green, 1995; Fishburn, 1970 [12][7]).

Definizione 1.7 (Funzione di utilità). Dato un insieme S , la rappresentazione di un ordine di preferenza \succ è data da una funzione

$$U : S \longrightarrow \mathbb{R},$$

detta *funzione di utilità* ed è tale per cui per ogni $x, y \in S$ valga

$$x \succ y \iff U(x) > U(y),$$

o equivalentemente

$$x \succeq y \iff U(x) \geq U(y).$$

Notiamo che le funzioni di utilità sono strettamente crescenti e non uniche.

Proposizione 1.8. Data U una funzione di utilità, per ogni $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ funzione strettamente crescente, allora $\psi = \varphi \circ U$ è ancora una funzione di utilità per la stessa relazione di preferenza.

Dimostrazione. Dobbiamo solo provare l'unicità. Date due funzioni di utilità strettamente crescenti $U(x)$ e $V(x)$ che rappresentano lo stesso ordine di preferenza, cerchiamo una funzione strettamente crescente $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ tale che $V(x) = \varphi \circ U(x) = \varphi(U(x))$, per ogni $x \in S$. Identifichiamo φ come $\varphi(t) = V(U^{-1}(t))$. Più rigorosamente, per ogni $t > 0$ nell'immagine $U(S)$ di U , c'è un $x_t \in S$ tale che $t = U(x_t)$. Definiamo $\varphi(t) = V(x_t)$. Mostriamo che φ è ben definita. Se c'è un'altra $y \in S$ tale che $U(y) = U(x_t)$, dato che U e V rappresentano la stessa preferenza, allora $V(y) = V(x_t)$, il che significa che $\varphi(t) = V(y)$. Nonostante x_t potrebbe non essere unico per ogni t , il valore $\varphi(t)$ è unico. Allora, per ogni x ed y , abbiamo

$$\varphi(U(x)) = V(U^{-1}(U(x))) = V(x), \varphi(U(y)) = V(U^{-1}(U(y))) = V(y).$$

Se per assurdo φ non fosse strettamente crescente allora potremmo trovare due valori t_1 e t_2 , con $t_1 < t_2$ e $\varphi(t_1) \geq \varphi(t_2)$. Potremmo quindi trovare due elementi arbitrari x_1 e x_2 tale che $U(x_1) = t_1$ e $U(x_2) = t_2$. Avremmo quindi che $x_2 \succ x_1$ per U , ma per una proprietà della funzione di utilità si dovrà avere che

$$V(x_1) = \varphi(U(x_1)) = \varphi(t_1) \geq \varphi(t_2) = \varphi(U(x_2)) = V(x_2),$$

ossia $x_1 \succeq x_2$ per V , il che contraddice il fatto che U e V rappresentino la stessa preferenza. Da cui φ dev'essere strettamente crescente. \square

Con la definizione di funzione di utilità possiamo dare una definizione equivalente di dominanza stocastica di primo grado.

Osservazione 1.9. Date due lotterie μ e ν , diremo che μ domina ν sotto la dominanza stocastica di primo ordine se e solo se ogni investitore con una funzione di utilità u crescente, preferisce μ a ν , ossia vale:

$$\int u(x)d\mu(x) \geq \int u(x)d\nu(x).$$

Nonostante la semplicità della definizione, ci possono essere dei casi in cui la dominanza non è così intuitiva.

Esempio 1.10. Supponiamo che le distribuzioni precedenti descrivano due diversi guadagni monetari. Diremo che $\mu \succeq_{FSD} \nu$ se e solo se $\mu(x) \leq \nu(x)$ per ogni x . Per dimostrare ciò consideriamo $\lambda(x) = \mu(x) - \nu(x)$ e supponiamo $\lambda(\bar{x}) > 0$ per un certo \bar{x} . Definendo

$$u(x) = \begin{cases} 1, & \text{se } x > \bar{x} \\ 0, & \text{altrimenti} \end{cases}$$

possiamo vedere che $\int u(x)d\lambda(x) = -\lambda(\bar{x}) < 0$ e questo prova la condizione. Considerando ora l'altro caso ed integrando per parti si avrà che

$$\int u(x)d\lambda(x) = [u(x)\lambda(x)]_0^\infty - \int u'(x)\lambda(x)dx.$$

Dato che $\lambda(0) = 0$ e $\lambda(x) = 0$ per $x \leq \bar{x}$, il primo termine è zero. Allora

$$\int u(x)d\lambda(x) \geq 0 \iff \int u'(x)\lambda(x)dx \geq 0.$$

Quindi se $\lambda(x) \leq 0$ per ogni x e $u(\cdot)$ è monotona, allora il secondo termine è minore o uguale a zero e ciò dimostra la disuguaglianza.

Una definizione equivalente possiamo darla anche di dominanza stocastica di secondo grado.

Osservazione 1.11. Date due lotterie μ e ν , diremo che μ domina ν sotto la dominanza stocastica di secondo ordine se e solo se ogni investitore con una funzione di utilità u crescente e concava, preferisce μ a ν .

Successivamente vedremo che la concavità di u coincide con un'avversione al rischio, per cui la dominanza di secondo ordine si concentra sugli investitori contrari al rischio. Un caso particolare di funzione di utilità è quella di Von-Neumann Morgenstern (Von Neumann, Morgenstern, 1953 [28]).

Definizione 1.12 (Funzione di utilità di Von-Neumann Morgenstern). Sia μ una distribuzione di una v.a. X , sia S un insieme di esiti, $u : S \rightarrow \mathbb{R}$ una funzione strettamente crescente, allora viene definita *funzione di utilità di Von-Neumann Morgenstern* (abbreviata in vNM) la seguente

$$U(\mu) = \int u(x)d\mu(x).$$

Osservazione 1.13. La funzione di utilità di vNM non è altro che il valore atteso della funzione $u(x)$, ossia $\mathbb{E}[u(x)]$.

La funzione appena definita ha un ruolo chiave nella teoria dell'utilità attesa, teoria che si concentra sulle decisioni in condizioni di rischio e nella quale gli investitori indecisi sui prospetti tenderanno a scegliere quello con l'utilità attesa più alta, dove questa può essere vista come un valore atteso di una v.a. che descrive l'utilità del prospetto.

Definizione 1.14 (Prezzo equo/Premio equo). Data una distribuzione μ ed S un insieme di esiti, definiamo

$$m(\mu) := \int xd\mu(x),$$

prezzo equo se μ è la distribuzione dei rendimenti di un certo asset, mentre lo definiamo *premio equo* se μ è la distribuzione dei pagamenti ricevuti da un'assicurazione in un contratto.

Con questa definizione possiamo definire l'avversione al rischio in termini della funzione di utilità di vNM.

Definizione 1.15 (Avversione al rischio). Un individuo è *avverso al rischio* se un certo esito $\delta_{m(\mu)}$ è preferito ad un prospetto μ per il quale $\mathbb{E}[\mu] = m(\mu)$. Ossia $\delta_{m(\mu)} \succ \mu$ a meno che $\delta_{m(\mu)} = \mu$. Questa condizione si traduce attraverso la funzione di utilità di vNM nella disuguaglianza di Jensen

$$U(\delta_m(\mu)) = u\left(\int x d\mu(x)\right) \geq \int u(X) d\mu(x) = U(\mu),$$

dove $u : S \rightarrow \mathbb{R}$ è una funzione strettamente crescente e continua in S .

In generale un investitore può avere tre tipi di approccio al rischio: propensione, neutralità ed avversione. Questi approcci corrispondono ad un comportamento specifico della funzione di utilità u .

Osservazione 1.16. Nella teoria dell'utilità attesa l'avversione al rischio è matematicamente equivalente alla concavità della funzione u , la convessità della funzione u è legata alla propensione al rischio, mentre la neutralità al rischio corrisponderà ad una funzione u lineare (ossia contemporaneamente concava e convessa).

Un'indicazione più precisa riguardo il grado di concavità di u (e quindi il grado di avversione al rischio) ci viene fornita da Kenneth J. Arrow e John W. Pratt. Essi ci forniscono due misure di avversione al rischio: una assoluta ed una relativa.

Definizione 1.17 (Assoluta avversione al rischio). Sia x un valore definito di ricchezza, $u'(x)$ la sua utilità marginale (cioè l'incremento del livello di utilità) e $u''(x)$ il tasso di variazione dell'utilità marginale in relazione alla ricchezza. Viene definita *misura di assoluta avversione al rischio* (ARA)

$$A(x) := -\frac{u''(x)}{u'(x)}.$$

Osservazione 1.18. Sotto le ipotesi di avversione al rischio, la funzione $A(x)$ è positiva (la funzione $u(x)$ è concava e quindi si avrà $u''(x) < 0$).

L'ARA misura il tasso di decadimento dell'utilità marginale quando il valore di ricchezza aumenta di un euro; tuttavia, usando diverse valute, anche l'ARA cambierà. Per ovviare a questo problema si introduce una misura di relativa avversione al rischio, la quale misura il tasso di decrescita dell'utilità quando il valore di ricchezza aumenta di un punto percentuale.

Definizione 1.19 (Relativa avversione al rischio). Sia x un valore definito di ricchezza, $u'(x)$ la sua utilità marginale (cioè l'incremento del livello di utilità) e $u''(x)$ il tasso di variazione dell'utilità marginale in relazione alla ricchezza. Viene definita *misura di relativa avversione al rischio* (RRA)

$$R(x) := -\frac{xu''(x)}{u'(x)} = xA(x).$$

Forniamo ora due esempi classici di classi di funzioni di utilità.

Esempio 1.20. Un primo esempio è la classe di *funzioni di utilità con avversione al rischio assoluto costante* (CARA). Queste sono funzioni esponenziali della forma

$$u(x) = -\frac{e^{-ax}}{a},$$

dove $a \in \mathbb{R}^+$. Da una formulazione alternativa dell'ARA ($A(x) = (\log(u'))'(x)$), avremo che $A(x) = a$ per ogni x . Il fatto che l'ARA sia costante è utile quando bisogna prendere decisioni riguardo rischi la cui dimensione è invariante a variazioni di ricchezza.

Esempio 1.21. Il secondo esempio è la classe di *funzioni di utilità con avversione al rischio relativo costante* (CRRA) chiamate anche *funzioni di utilità con avversione al rischio assoluto iperbolica* (HARA). Tale insieme di funzioni è definito da

$$u(x) = \begin{cases} \frac{x^{1-\gamma}}{1-\gamma}, & \text{per } \gamma \geq 0, \gamma \neq 1 \\ \log(x), & \text{per } \gamma = 1 \end{cases}$$

Questo tipo di funzioni sono utili quando si prendono decisioni riguardo rischi la cui dimensione è proporzionale al livello di ricchezza. L'ipotesi che la relativa avversione al rischio sia costante semplifica numerosi problemi di macroeconomia e finanza.

Capitolo 2

Introduzione alle misure di rischio

Nell'ambito pratico si può pensare di associare ad ogni prospetto una v.a. X che indichi perdite o guadagni di questo. Concretamente si introduce quindi una misura di rischio, ossia una funzione $\xi : \mathcal{G} \rightarrow \mathbb{R}$ (dove \mathcal{G} è l'insieme delle variabili aleatorie che determinano il valore di rischio di una posizione in un istante di tempo), che associ ad ogni valore una sua rischiosità, in modo tale da presentare l'andamento del prospetto in termini di variazione del rischio. Le principali misure di rischio dipendono tuttavia soltanto dalla distribuzione a loro associata e tramite questa vengono analizzate. Si avrà quindi che date due v.a. X, Y vale che $\xi(X) = \xi(Y)$ se e solo se hanno la stessa distribuzione (cioè $F_X(x) = F_Y(x)$ per ogni $x \in \mathbb{R}$) (Frittelli, Rosazza Gianin, 2005 [8]).

2.1 Misure di rischio coerenti

Tra le misure di rischio più importanti dal punto di vista operativo ci sono le misure coerenti, che soddisfano una serie di assiomi che definiscono una “buona” misura (Artzner, Delbaen, Eber, Heath, 1998 [4]).

Il primo assioma che caratterizza una misura di rischio coerente è l'*invarianza per traslazioni*

$$(T) \quad \forall X \in \mathcal{G}, \forall \alpha \in \mathbb{R}, \xi(X + \alpha r) = \xi(X) - \alpha.$$

Questo assioma ci garantisce che il rischio di una certa v.a. X a cui viene aggiunto un valore α moltiplicato per il tasso di rendimento certo r è uguale al rischio di X meno α .

Osservazione 2.1. Se $\alpha = \xi(X)$ allora (T) diventa $\xi(X + \xi(X)r) = 0$. Vedremo che questa uguaglianza avrà un'interpretazione in termini di insieme di rischi accettabili associato a ξ .

Il secondo assioma che caratterizza una misura di rischio coerente è la *subadditività*

$$(S) \quad \forall X, Y \in \mathcal{G}, \xi(X + Y) \leq \xi(X) + \xi(Y).$$

Questo assioma afferma che se diversifichiamo il nostro portafoglio di investimenti, il rischio totale sarà sempre minore o uguale al rischio dei singoli investimenti.

Successivamente tratteremo il VaR, una misura di rischio che non soddisfa questa proprietà e noteremo come il non soddisfarla porterà a numerose criticità.

Il terzo assioma che caratterizza una misura di rischio coerente è la *positiva omogeneità*

$$(PH) \quad \forall X \in \mathcal{G}, \forall \lambda \geq 0, \xi(\lambda X) = \lambda \xi(X).$$

Questo assioma è legato alla liquidità di una posizione (ossia di una strategia di investimento): dato che la posizione λX è generalmente meno liquida di X , è ragionevole supporre che la rischiosità di assumere la posizione λX sia uguale ad assumere λ posizioni X regolarmente.

Il quarto ed ultimo assioma che caratterizza una misura di rischio coerente è la *monotonia*

$$(M) \quad \forall X, Y \in \mathcal{G}, \text{ se } X \leq Y \text{ allora } \xi(Y) \leq \xi(X).$$

La necessità di questo assioma appare evidente se consideriamo che i valori positivi di X rappresentanti i rischi indicano dei guadagni.

Definizione 2.2 (Misura di rischio coerente). Una misura di rischio ξ si definisce *coerente* se e solo se soddisfa gli assiomi (T), (S), (PH) ed (M).

Le misure di rischio coerenti sono buone misure con cui lavorare poichè sono associate ad un atteggiamento al rischio razionale. Nella pratica una misura coerente si definisce tale solo se “associa un valore maggiore ad un rischio maggiore”. In generale, sfruttando la proprietà che ogni combinazione lineare di misure di rischio coerenti è una misura di rischio coerente, si ha che queste fanno parte di una classe più ampia di misure di rischio chiamate misure spettrali di rischio (Acerbi, 2002 [1]).

Definizione 2.3 (Misura spettrale di rischio). Data una funzione $\phi : [0, 1] \rightarrow \mathbb{R}^+$ positiva, strettamente decrescente e tale che

- $\lim_{m \rightarrow 0} \phi(m) = \infty;$
- $\lim_{m \rightarrow 1} \phi(m) = 0;$
- $\int_0^1 |\phi(m)| dm = 1;$

ed una funzione di distribuzione $F_X(x)$ associata ad una v.a. X , si definisce *misura spettrale di rischio*

$$M_\phi(X) = - \int_0^1 F_X^{-1}(m)\phi(m)dm.$$

La funzione ϕ è chiamata *spettro di rischio* ed identifica l'atteggiamento al rischio da parte di un investitore: se ϕ è costante si avrà una neutralità alle perdite, mentre in generale si avrà un'avversione ad esse. Di queste misure di rischio spettrali fa parte il CVaR, misura che tratteremo in seguito.

2.1.1 Equivalenza con insieme di rischi accettabili

Un modo equivalente di interpretare le misure di rischio coerenti è tramite l'insieme dei rischi accettabili.

Definizione 2.4 (Insieme di rischi accettabili). Data una misura di rischio ξ ad essa viene associato un *insieme di rischi accettabili*

$$\mathcal{A}_\xi = \{X \in \mathcal{G} : \xi(X) \leq 0\}.$$

Viceversa, dato un insieme $\mathcal{A} \subset \mathcal{G}$, ad esso può venir associata una misura di rischio $\xi_{\mathcal{A}}(X) = \inf\{t : tr + X \in \mathcal{A}\}$ con r il tasso di rendimento certo.

Osservazione 2.5. Il cono di tutti gli elementi non negativi di \mathcal{G} si denota L_+ , mentre quello con tutti gli elementi negativi L_- .

Il legame tra le due nozioni è espresso dalle seguenti proposizioni.

Proposizione 2.6. Sia $\mathcal{A} \subset \mathcal{G}$ un insieme di rischi che soddisfa i seguenti assiomi:

- (1) $L_+ \subset \mathcal{A}$;
- (2) $\mathcal{A} \cap L_{--} = \emptyset$ dove

$$L_{--} = \{X : \forall \omega \in \Omega, X(\omega) < 0\},$$

dove Ω è l'insieme degli esiti di una v.a.;

- (3) \mathcal{A} è convesso;
- (4) \mathcal{A} è un cono convesso positivamente omogeneo;

allora $\xi_{\mathcal{A}}$ è una misura di rischio coerente e si ha $\mathcal{A}_\xi = \overline{\mathcal{A}}$.

Dimostrazione. Gli assiomi (2) e (3) assicurano che per ogni X , $\xi_{\mathcal{A}}(X)$ è un numero finito.

L'uguaglianza $\inf\{p : X + (\alpha + p) \cdot r \in \mathcal{A}\} = \inf\{q : X + qr \in \mathcal{A}\} - \alpha$ prova che $\xi_{\mathcal{A}}(X + r\alpha) = \xi(X) - \alpha$ e l'assioma (T) è soddisfatto.

La subadditività di $\xi_{\mathcal{A}}$ segue dal fatto che se $X + mr$ e $Y + nr$ appartengono entrambi a \mathcal{A} , deve valere anche per $X + Y + (m + n) \cdot r$ come mostrano (3) e (4).

Se $m > \xi_{\mathcal{A}}(X)$ allora per ogni $\lambda > 0$ abbiamo che $\lambda X + \lambda m \cdot r \in \mathcal{A}$ e questo prova che $\xi_{\mathcal{A}}(\lambda X) \leq \lambda m$. Se $m < \xi_{\mathcal{A}}(X)$ allora per ogni $\lambda > 0$ abbiamo che $\lambda X + \lambda m \cdot r \notin \mathcal{A}$ e questo prova che $\xi_{\mathcal{A}}(\lambda X) \geq \lambda \cdot m$. Si ha quindi che $\xi_{\mathcal{A}}(\lambda X) = \lambda \cdot \xi_{\mathcal{A}}(X)$.

La monotonia di $\xi_{\mathcal{A}}$ segue dal fatto che se $X \leq Y$ e $X + mr \in \mathcal{A}$ allora $Y + mr \in \mathcal{A}$ usando (3) e (1).

Per ogni $X \in \mathcal{A}$, $\xi_{\mathcal{A}}(X) \leq 0$ da cui $X \in \mathcal{A}_{\xi}$. La definizione precedente combinata con parte della dimostrazione assicurano che \mathcal{A}_{ξ} è chiuso, il che prova che $\mathcal{A}_{\xi} = \overline{\mathcal{A}}$. \square

Proposizione 2.7. Viceversa se ξ è una misura di rischio coerente su \mathcal{G} , \mathcal{A}_{ξ} è chiuso e soddisfa i precedenti assiomi ed inoltre si ha $\xi_{\mathcal{A}_{\xi}} = \xi$.

Dimostrazione. (S) e (PH) assicurano che ξ è una funzione convessa di \mathcal{G} , quindi continua, e che l'insieme $\mathcal{A}_{\xi} = \{X : \xi(X) \leq 0\}$ è chiuso, convesso ed un cono omogeneo.

(PH) implica che $\xi(0) = 0$. Insieme ad (M) questo assicura che l'insieme \mathcal{A}_{ξ} contiene L_+ .

Sia $X \in L_{--}$ con $\xi(X) < 0$. (M) assicura che $\xi(0) < 0$, che è una contraddizione. Se $\xi(X) = 0$, allora possiamo trovare un $\alpha > 0$ tale che $X + \alpha r \in L_{--}$, che implica, usando (T), la relazione $-\alpha \geq 0$, che è una contraddizione. Da cui si dovrà avere che $\xi(X) > 0$, quindi $X \notin \mathcal{A}_{\xi}$, che prova l'assioma (2).

Per ogni X sia δ ogni numero tale che $\xi_{\mathcal{A}_{\xi}}(X) < \delta$. Allora $X + \delta \cdot r \in \mathcal{A}_{\xi}$, da cui $\xi(X + \delta \cdot r) \leq 0$, quindi $\xi(X) \leq \delta$, che prova che $\xi(X) \leq \xi_{\mathcal{A}_{\xi}}(X)$, che è $\xi \leq \xi_{\mathcal{A}_{\xi}}$.

Per ogni X , sia δ ogni numero tale che $\delta > \xi(X)$, allora $\xi(X + \delta \cdot r) < 0$ e $X + \delta \cdot r \in \mathcal{A}_{\xi}$, da cui $\xi_{\mathcal{A}_{\xi}}(X + \delta \cdot r) \leq 0$. Questo prova che $\xi_{\mathcal{A}_{\xi}}(X) \leq \delta$ e che $\xi_{\mathcal{A}_{\xi}}(X) \leq \xi(X)$, da cui si ha che $\xi_{\mathcal{A}_{\xi}} \leq \xi$. \square

2.2 Misure di rischio convesse

Un'altra classe di misure di rischio degne di nota, sono quelle convesse. Esse hanno legami sia con le misure di rischio coerenti che con la teoria dell'utilità (Pelessoni, 2002; Artzner, Delbaen, Eber, Heath, 1998 [19][4]). Il rapporto con le misure di rischio coerenti viene fornito dall'assioma di *convessità*

$$(C) \quad \forall X, Y \in \mathcal{G}, \forall \lambda \in [0, 1], \xi(\lambda X + (1 - \lambda)Y) \leq \lambda \xi(X) + (1 - \lambda)\xi(Y).$$

Questo assioma ci dice che la diversificazione di un portafoglio non incrementa il rischio complessivo. Di questo ci si può convincere se pensiamo che la convessità è equivalente alla *quasi convessità*

$$\forall X, Y \in \mathcal{G}, \forall \lambda \in [0, 1], \xi(\lambda X + (1 - \lambda)Y) \leq \max\{\xi(X), \xi(Y)\}.$$

Definizione 2.8 (Misura di rischio convessa). Una misura di rischio ξ si definisce *convessa* se e solo se soddisfa l'assioma (C).

Osservazione 2.9. Se una misura di rischio convessa soddisfa gli assiomi (PH) e (S) allora è una misura di rischio coerente.

Una proposizione simile a quelle precedenti può essere enunciata anche per le misure di rischio convesse.

Proposizione 2.10. Se \mathcal{A}_ξ è convesso allora $\xi_{\mathcal{A}}$ è una misura di rischio convessa. Se \mathcal{A}_ξ è un cono (per ogni $X \in \mathcal{A}$ e per ogni $\lambda \geq 0$ si ha $\lambda X \in \mathcal{A}$) allora $\xi_{\mathcal{A}}$ soddisfa (PH) e, in particolare, se \mathcal{A}_ξ è un cono convesso, $\xi_{\mathcal{A}}$ è una misura di rischio coerente.

Dimostrazione. Supponiamo $X_1, X_2 \in \mathcal{G}$ e che $m_1, m_2 \in \mathbb{R}$ sono tali che $m_i + X_i \in \mathcal{A}_\xi \forall i \in \{1, 2\}$. Se $\lambda \in [0, 1]$, allora la convessità di \mathcal{A} implica che $\lambda(m_1 + X_1) + (1 - \lambda)(m_2 + X_2) \in \mathcal{A}_\xi$. Allora per un proprietà di $\xi_{\mathcal{A}}$ si ha,

$$\begin{aligned} 0 &\geq \xi_{\mathcal{A}}(\lambda(m_1 + X_1) + (1 - \lambda)(m_2 + X_2)) \\ &= \xi_{\mathcal{A}}(\lambda X_1 + (1 - \lambda)X_2) - (\lambda m_1 + (1 - \lambda)m_2), \end{aligned}$$

da cui segue la convessità di $\xi_{\mathcal{A}}$. Per la seconda parte della proposizione, otteniamo che $\xi_{\mathcal{A}}(\lambda X) \leq \lambda \xi_{\mathcal{A}}(X)$ per $\lambda \geq 0$ se \mathcal{A}_ξ è un cono. Per provare la disuguaglianza inversa, sia $m < \xi_{\mathcal{A}}(X)$, allora $m + X \notin \mathcal{A}_\xi$ e quindi $\lambda m + \lambda X \notin \mathcal{A}_\xi$ per $\lambda \geq 0$. Da cui si ha che $\lambda m < \xi_{\mathcal{A}}(\lambda X)$. \square

Il rapporto tra le misure di rischio convesse, la teoria dell'utilità e le misure di rischio coerenti è invece espresso da

Osservazione 2.11. Se ξ è una misura di rischio convessa, allora $\varphi(X) := -\xi(X)$ è chiamato funzionale di utilità concavo. Se ξ è coerente allora φ è chiamato funzionale di utilità coerente.

Invertendo il segno della misura di rischio ξ enfatizziamo l'utilità di una posizione piuttosto che il suo rischio.

Un esempio di come non sempre misure di rischio convesse e misure di rischio coerenti siano collegate è dato da

Esempio 2.12. Sia $u : \mathbb{R} \rightarrow \mathbb{R}$ una funzione di utilità strettamente crescente e concava. Fissiamo $x_0 \in \mathbb{R}$ e sia

$$\mathcal{A} := \{X \in \mathcal{G} : \mathbb{E}[u(X)] \geq u(x_0)\}.$$

X è accettata se e solo se $u^{-1}(\mathbb{E}[u(X)]) \geq x_0$. Inoltre la concavità di u implica che per $X, Y \in \mathcal{A}$ e $\lambda \in [0, 1]$ vale

$$\begin{aligned} \mathbb{E}[u(\lambda X + (1 - \lambda)Y)] &\geq \mathbb{E}[\lambda u(X) + (1 - \lambda)u(Y)] \\ &= \lambda \mathbb{E}[u(X)] + (1 - \lambda)\mathbb{E}[u(Y)] \\ &\geq u(x_0), \end{aligned}$$

da cui \mathcal{A} è convesso. La corrispondente misura di rischio convessa è data da

$$\xi_{\mathcal{A}}(X) = \inf\{m \in \mathbb{R} : \mathbb{E}[u(X + m)] \geq u(x_0)\}.$$

Consideriamo ora un caso speciale di funzione di utilità esponenziale $u(x) = 1 - e^{-\beta x}$ con $\beta > 0$, un tipo di funzione CARA, ed $x_0 = 0$. Abbiamo che

$$\begin{aligned} u'(x) &= \beta e^{-\beta x} > 0, \\ u''(x) &= -\beta^2 e^{-\beta x} < 0. \end{aligned}$$

La condizione $\mathbb{E}[u(X + m)] \geq u(x_0)$ ora diventa $1 - e^{-\beta m} \mathbb{E}[e^{-\beta X}] \geq 0$ o, equivalentemente, $m \geq \frac{1}{\beta} \log(\mathbb{E}[e^{-\beta X}])$. Da cui

$$\xi_{\mathcal{A}}(X) = \frac{1}{\beta} \log(\mathbb{E}[e^{-\beta X}]),$$

chiamata *misura di rischio entropica*. Questa misura di rischio è una possibile alternativa al VaR ed al CVaR (misure di rischio al ribasso che approfondiremo nel capitolo successivo) ed è un tipico esempio di misura di rischio convessa ma non coerente.

Capitolo 3

Misure di rischio al ribasso

La teoria delle misure di rischio al ribasso insieme alla teoria del portafoglio si è sviluppata a partire dalla pubblicazione di due articoli nel 1952. Il primo, scritto da Markowitz, dava un metodo per misurare il rischio ed il rendimento di un portafoglio. Egli usava inizialmente media, varianza, deviazione standard e covarianza per ottenere una frontiera efficiente sulla quale ogni portafoglio massimizzava il rendimento atteso per una determinata varianza o minimizzava la varianza per un dato rendimento atteso (criterio della media-varianza). Questo criterio si basava principalmente sull'attribuire una funzione di utilità ad ogni individuo. Il secondo articolo fu quello di Roy che non credeva fosse possibile associare ad ogni investitore una sua funzione di utilità. Roy chiamò rendimento di riferimento il minimo rendimento accettabile da parte dell'investitore, il quale, secondo lui, preferiva l'investimento con la minor probabilità di andare sotto questo livello. Massimizzando quindi il rapporto tra guadagni e variabilità, l'investitore sceglierà il portafoglio con la minor probabilità di scendere sotto una fissata soglia di riferimento (rischio al ribasso). Markowitz riconobbe l'importanza di quest'idea. Realizzò che l'investitore era interessato a minimizzare il rischio al ribasso per due motivi: il primo era l'importanza per l'investitore di monitorare questo tipo di rischio, il secondo era che le distribuzioni utilizzate potevano non essere sempre normali, mentre la misura di rischio al ribasso aiutava gli investitori a lavorare proprio con le distribuzioni non normali. Markowitz mostrò quindi come entrambi i metodi funzionassero quando le distribuzioni erano normali, mentre quando non lo erano solo usando la misura di rischio al ribasso si aveva la soluzione corretta (Nawrocki, 1999 [14]).

Con il tempo si notò che gli investitori si preoccupavano più di un andamento sotto le aspettative piuttosto che di uno sopra, ciò portò Markowitz a sviluppare delle prime misure di rischio al ribasso che sfruttavano la semivarianza e la semideviazione standard

$$\begin{aligned}\bar{\sigma}^2(X) &= \mathbb{E}[(\max\{\mathbb{E}[X] - X, 0\})^2], \\ \bar{\sigma}(X) &= \sqrt{\mathbb{E}[(\max\{\mathbb{E}[X] - X, 0\})^2]}.\end{aligned}$$

Data una v.a. X che indicasse l'andamento di un portafoglio, la semivarianza poteva essere calcolata sia dal valor medio di X sia da un valore fissato t , dando quindi agli investitori un riferimento più utile sull'andamento del loro portafoglio. Con la progressione della teoria finanziaria però questo tipo di misure di rischio al ribasso caddero in disuso a favore di misure più precise e facilmente implementabili. Tra queste ci sono misure collegate alla teoria dell'utilità, come i momenti parziali inferiori (LPMs) e misure più convenzionali come il valore a rischio (VaR) ed il valore a rischio condizionale (CVaR).

3.1 Valore a rischio

Il valore a rischio, abbreviato VaR, è una misura di rischio che stima le potenziali perdite di un portafoglio in un breve periodo di tempo. Essa calcola la probabilità che l'andamento di un portafoglio scenda sotto un dato livello di confidenza entro un certo periodo di tempo. Viene usata ad esempio dalle banche per gestire i portafogli di asset dei clienti: confrontando il VaR ottenuto con il capitale disponibile e la loro liquidità, possono assicurarsi che le potenziali perdite vengano economicamente coperte senza mettere a repentaglio i portafogli dei loro investitori.

Definizione 3.1 (Valore a rischio). Data una v.a. X che definisce il rendimento di un portafoglio, un livello di confidenza fissato $\alpha \in (0, 1)$ ed un periodo di tempo t , il *valore a rischio* di X è definito come

$$\text{VaR}_{\alpha,t}(X) = \min\{x \in \mathbb{R} : P(X \leq x) \geq \alpha\}.$$

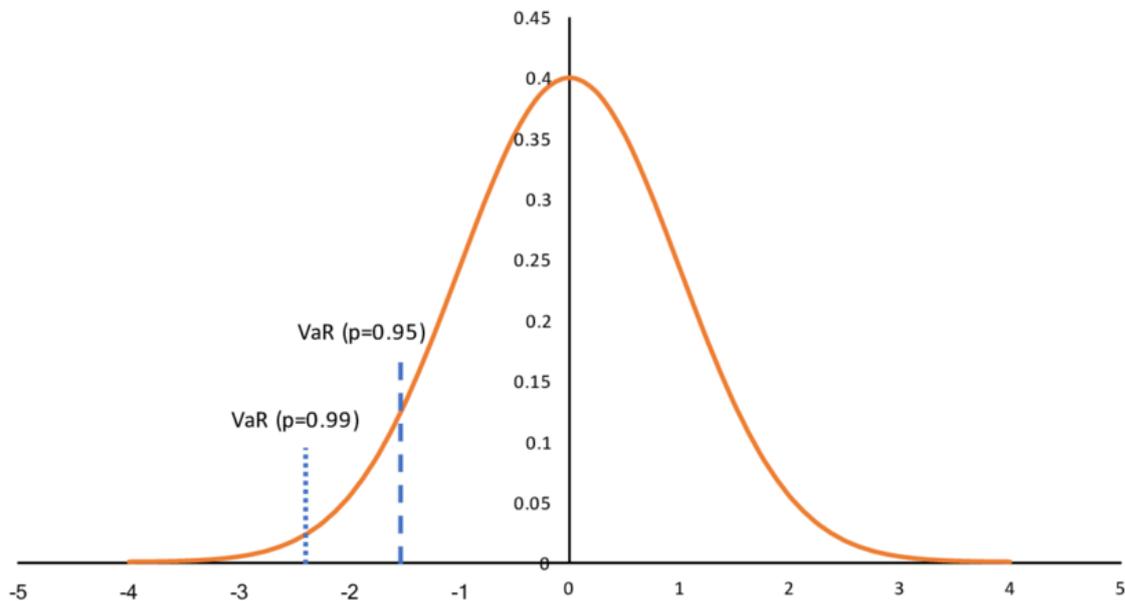
Osservazione 3.2. Il VaR può essere definito usando i quantili come il quantile inferiore di X di ordine α (o il 100α -esimo percentile) (Sarykalin, Serraino, Uryasev, 2008 [23]) o, usando la funzione di ripartizione di X (Pflug, 2000 [20]), come

$$\text{VaR}_{\alpha,t}(X) = F_X^{-1}(\alpha),$$

dove $F_X^{-1}(\alpha) := \min\{x : F_X(x) \geq \alpha\}$.

Il valore di riferimento α (indicato anche con p) è dato in percentuale ed è solitamente fissato sul 95-99% (per dare una maggiore garanzia agli investitori) mentre il lasso temporale t è di solito di ore, giorni o massimo qualche settimana.

Esempio 3.3. Se il VaR di un portafoglio calcolato in una settimana e con un livello di confidenza del 95% è di un milione di euro, allora significa che si avrà solo il 5% di probabilità che il valore del portafoglio scenda sotto quel valore nelle settimane successive.



Nella pratica il VaR può essere calcolato usando principalmente tre metodi: un primo metodo parametrico di varianza-covarianza che si basa su assunzioni fatte sulle distribuzioni dei rendimenti dei fattori di rischio, un secondo metodo di simulazione storica non parametrico che si basa sull'uso dei dati di ipotetici portafogli usando serie storiche ed un terzo metodo anch'esso non parametrico che sfrutta le simulazioni di Monte Carlo. Tutti e tre presentano vantaggi e svantaggi.

3.1.1 Metodo di varianza-covarianza

Il metodo parametrico di varianza-covarianza, anche chiamato delta normal, è il metodo più semplice per il calcolo del VaR: esso si basa sull'analisi dei singoli asset che compongono un portafoglio, studiandone separatamente i parametri di varianza e correlazione. Tramite questa suddivisione e alcune assunzioni fatte sulle distribuzioni dei rendimenti dei fattori di rischio, possiamo calcolare il VaR dell'intero portafoglio. Il metodo si fonda sulle ipotesi iniziali che i rendimenti dei fattori di rischio di tutti gli asset siano normalmente distribuiti e che deviazione standard e correlazione di questi siano costanti nel tempo. Come conseguenza si avrà che la distribuzione generale dei rendimenti del portafoglio sarà anch'essa normale (poichè combinazione lineare delle distribuzioni dei singoli asset) ed il calcolo del VaR sarà quindi più semplice. Una volta stimate varianza e covarianza degli asset tramite i dati sulle serie storiche di questi (ad esempio consultando il servizio *RiskMetrics* offerto da J.P.Morgan nel 1995¹) ed analizzando le matrici di correlazione e covarianza ottenute, si calcolerà il VaR.

¹RiskMetrics-Technical Document, J.P. Morgan, December 17, 1996; Zangari, P., 1996, An Improved Methodology for Computing VaR, J.P. Morgan RiskMetrics Monitor, Second Quarter 1996.

Il metodo di varianza-covarianza ha evidenti vantaggi e limitazioni. Tra i vantaggi ci sono la semplicità e velocità dal punto di vista implementativo e l'esclusione di possibili problematiche legate alla correlazione di più asset tra loro che semplificano notevolmente il calcolo del VaR. Dall'altra parte le ipotesi su distribuzioni normali ed invarianza nel tempo di correlazione e volatilità (deviazione standard) risultano alquanto irrealistiche, portando talvolta ad incongruenze tra il vero VaR ed il VaR effettivamente calcolato. Anche le stime di varianza e correlazione prese dalle serie storiche potrebbero non essere sempre attendibili. Per chiarezza riportiamo un esempio di applicazione.

Esempio 3.4. Sono dati due portafogli aventi entrambi tre azioni di tre compagnie differenti (M&M Ltd, WIPRO Ltd e ICICI Bank), i pesi di queste tuttavia sono diversi tra i portafogli. Per entrambi i portafogli vengono scelti due livelli di confidenza ($\alpha = 95\%$ ed $\alpha = 99\%$) due lassi di tempo (un giorno e dieci giorni) ed un periodo di test che va dal primo gennaio 2018 al 31 dicembre 2018². Per il portafoglio 1 i prezzi registrati a dicembre (closing prices) e i pesi delle tre azioni sono:

| Azioni | Prezzo (in rupie) | Peso |
|------------|-------------------|-------|
| M&M Ltd | 803,700 | 33,3% |
| WIPRO Ltd | 330,25 | 33,3% |
| ICICI Bank | 360 | 33,3% |

Per il portafoglio 2 i prezzi registrati a dicembre e i pesi delle tre azioni sono:

| Azioni | Prezzo (in rupie) | Peso |
|------------|-------------------|------|
| M&M Ltd | 803,700 | 50% |
| WIPRO Ltd | 330,25 | 30% |
| ICICI Bank | 360 | 20% |

Per calcolare la deviazione standard del portafoglio P usiamo la seguente formula:

$$\sigma_P = \sqrt{\sum_{i=1}^3 (W_i \sigma_i)^2 + 2 \left(\sum_{i=1}^3 \sum_{j=1}^3 (W_i \sigma_i \cdot W_j \sigma_j \cdot \rho_{ij}) \right)}$$

dove W_i e W_j sono i pesi delle azioni, σ_i e σ_j sono le deviazioni standard delle azioni e ρ_{ij} è il coefficiente di correlazione tra le azioni i e j .

L'andamento dei rendimenti, la varianza e la deviazione standard (volatilità) delle azioni sono:

²Variance-Covariance (Delta normal) approach of VaR Models: An Example From Bombay Stock Exchange, Finance&Marketing, Oriental Institute of Management, Vashi Navi Mumbai.

| Azioni | M&M Ltd | WIPRO Ltd | ICICI Bank |
|---------------------|----------|-----------|------------|
| Minimo | -0,06892 | -0,04319 | -0,06110 |
| Medio | 0,00034 | 0,00017 | 0,00067 |
| Massimo | 0,05151 | 0,04767 | 0,10277 |
| Varianza | 0,00028 | 0,00020 | 0,00039 |
| Deviazione standard | 0,01670 | 0,01409 | 0,01974 |

Da cui si avrà la matrice di correlazione

$$\rho = \begin{pmatrix} 1 & 0 & 0 \\ 0,44133 & 1 & 0 \\ 0,39205 & 0,40438 & 1 \end{pmatrix}$$

e la matrice di covarianza

$$\text{Cov} = \begin{pmatrix} 0,00028 & 1,89357 \cdot 10^{-5} & 0,00010 \\ 1,89357 \cdot 10^{-5} & 0,00020 & 3,9046 \cdot 10^{-6} \\ 0,00010 & 3,9046 \cdot 10^{-6} & 0,00039 \end{pmatrix}$$

Per trovare la volatilità del portafoglio i valori di covarianza tra le azioni sono moltiplicati per i singoli pesi e poi sommati tra loro. Da ciò otteniamo

| | Portafoglio 1 | Portafoglio 2 |
|---------------------|---------------|---------------|
| Varianza | 0,00012 | 0,00013 |
| Deviazione standard | 0,01101 | 0,01137 |

Per calcolare il VaR usiamo la seguente formula:

$$\text{VaR} = \sqrt{t} P \alpha \sigma$$

dove P è il valore del portafoglio (o posizione), α è il livello di confidenza (per il 95% α è 1,65, mentre per il 99% α è 2,33), σ è la deviazione standard (volatilità) e t è il lasso temporale considerato. Se supponiamo che P abbia un valore di un milione di rupie allora il VaR sarà

| | Portafoglio 1 | | Portafoglio 2 | |
|--------------------|---------------|-----------|---------------|-----------|
| | 95% | 99% | 95% | 99% |
| VaR (un giorno) | 18.170,51 | 25.658,96 | 18.764,9 | 26.498,31 |
| VaR (dieci giorni) | 57.460,2 | 81.140,76 | 59.339,82 | 83.795,02 |

Dai dati nella tabella notiamo che, considerando il VaR calcolato per un giorno, nel giorno seguente il portafoglio 1 ha solo il 5% di probabilità di scendere sotto le 18.171 rupie, mentre il portafoglio 2, con la stessa probabilità del 5%, può scendere sotto le 18.765 rupie. Discorso analogo vale per i dati relativi ai dieci giorni con un livello di confidenza del 99%. A causa delle poche azioni presenti nei portafogli e della loro correlazione, non così bassa da far diminuire il rischio complessivo, concludiamo quindi che non c'è una netta differenza tra i risultati dei due portafogli.

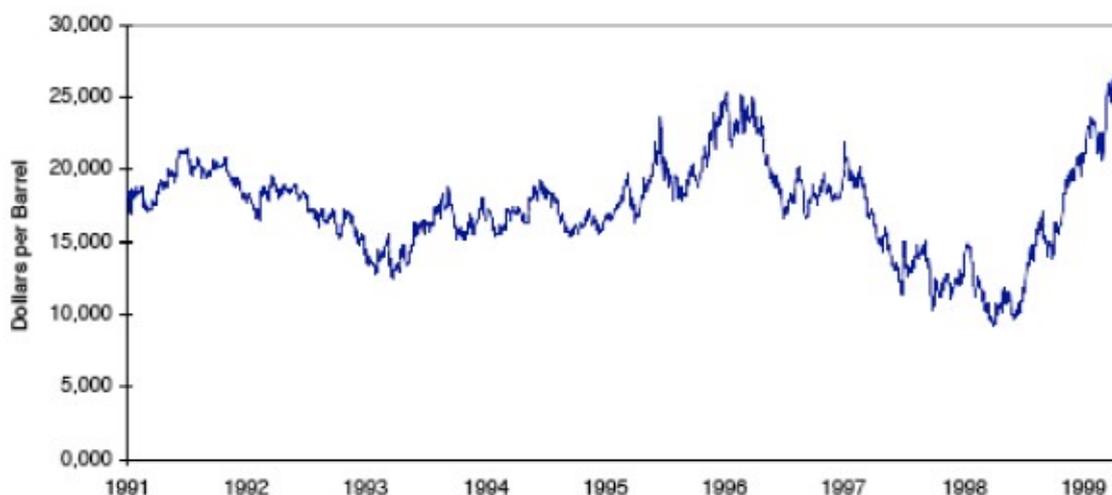
3.1.2 Metodo di simulazione storica

Il secondo metodo per il calcolo del VaR è il primo non parametrico che andremo ad analizzare. Il metodo di simulazione storica permette di valutare il VaR di vari portafogli stimandolo tramite la creazione di un'ipotetica serie storica di rendimenti. Questa è ottenuta simulando l'andamento del portafoglio tramite vecchi dati e tenendo in considerazione i cambiamenti di ogni periodo.

Anche questo metodo non è ovviamente esente da svantaggi. Il primo è che ad ogni prezzo giornaliero della serie storica viene dato ugual peso per il calcolo del VaR e questo rappresenta un problema nella sua misurazione in caso di trend nella variabilità. Un secondo è l'assunzione del periodo di analisi dei dati: preso uno specifico periodo di osservazione, non è detto che esso rispecchi l'andamento generale. Un terzo e più nascosto limite è il non supporre inizialmente una distribuzione dei rendimenti. Se questo da un lato può sembrare un punto a favore, dall'altro rappresenta un'ipotesi azzardata: si sta infatti assumendo che la distribuzione associata ai vecchi dati sia affidabile nella previsione di andamenti futuri dei rendimenti. Questa scelta può facilmente rivelarsi errata soprattutto se la contestualizziamo in un mercato soggetto a continue variazioni e volatilità in costante aggiornamento. Anche in questo caso diamo un esempio di applicazione.

Esempio 3.5. Cabedo e Moya fornirono un semplice esempio di applicazione del metodo di simulazione storica calcolando il VaR sui prezzi del petrolio³. Dai dati dal 1991 al 1999 ottennero i prezzi giornalieri del petrolio al barile della Brent Crude Oil. Separarono le variazioni di prezzo in valori positivi e negativi. Fissato un livello di confidenza del 99%, il VaR positivo fu definito come il cambio di prezzo nel 99-esimo percentile dei valori positivi, mentre il VaR negativo come il cambio di prezzo nel 99esimo percentile dei valori negativi. Per il periodo di osservazione, il VaR giornaliero al 99esimo percentile era dell'1% in entrambe le direzioni.

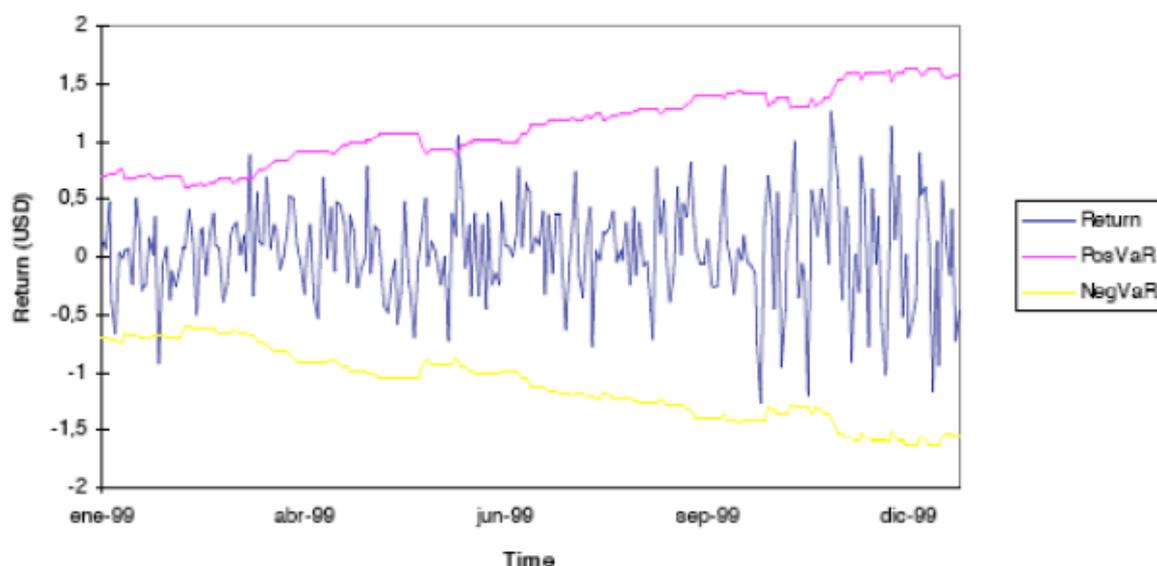
³J.D. Cabedo and I. Moya, 2003, Estimating oil price Value at Risk using the historical simulation approach, *Energy Economics*, v25, 239-253.



Grazie a questo esempio possiamo mettere in evidenza gli svantaggi esposti in precedenza. Attribuire lo stesso peso a tutti i dati a prescindere dall'anno considerato, significa non tenere conto della volatilità di questi e quindi delle fluttuazioni di mercato, che potrebbero fornirci un VaR falsato. Per quanto riguarda la seconda criticità, se il calcolo del VaR avesse preso in considerazione i dati dal 1999 al 2004 piuttosto che quelli dal 1991 al 1999, il gestore del portafoglio avrebbe dovuto confrontarsi con perdite molto più grandi del previsto. Questo ci suggerisce una limitazione più generale del VaR: la sottovalutazione (o sopravvalutazione) del rischio ed in generale la poca affidabilità nel monitoraggio di casi estremi.

Per risolvere parte delle problematiche di questo approccio presentiamo alcune modifiche che sono state apportate negli anni. Come prima cosa si può pensare di attribuire più peso ai dati più recenti piuttosto che lo stesso per tutti: questo perchè si suppone che questi forniscano un'indicazione più veritiera riguardo gli andamenti futuri. Nello specifico Boudoukh, Richardson e Whitelaw presentarono una variante della simulazione storica dove ai dati più recenti veniva dato un peso maggiore tramite un "fattore di decadenza". Se il fattore di decadenza era 0,90 e il dato più recente aveva un peso p allora il dato precedente ad esso avrebbe avuto un peso di $0,9 \cdot p$, quello ancora prima di $0,81 \cdot p$ e così via (l'approccio classico della simulazione storica può essere pensato come un caso particolare di questa variante con un fattore di decadenza uguale ad 1). Per testare la bontà di questa variante venne analizzato il VaR di un portafoglio di azioni in un periodo di 250 giorni prima e dopo il crollo di mercato del 19 ottobre 1987. Usando l'approccio classico, il VaR del portafoglio rimaneva invariato anche nei giorni successivi al crollo (perchè ad ogni giorno veniva attribuito lo stesso peso), mentre con il nuovo approccio si notò che con il passare dei giorni dopo il crollo, il VaR si adattava ai recenti accadimenti. Un'altra particolare modifica al metodo della simulazione storica la

troviamo in una variante dell'esperimento di Cabado e Moya illustrato precedentemente. Per stimare meglio il VaR, essi presero dei parametri da un modello di una serie storica e li utilizzarono per prevederne il comportamento. Nello specifico essi adottarono un modello autoregressivo a media mobile (ARMA) dei prezzi del petrolio dal 1991 al 1999 e lo usarono per predire il VaR con un livello di confidenza del 99%. Si accorsero che, con il nuovo metodo, i prezzi del petrolio nel 1999 rimanevano tra il VaR positivo ed il VaR negativo il 98,8% delle volte, contro il 97,7% del vecchio metodo. Uno dei motivi che li spinse ad adottare questa variante era la differenza di sensibilità del VaR effettivamente calcolato tra serie storiche e simulazioni: quello calcolato tramite le serie storiche era molto più sensibile alle variazioni di prezzo come dimostra il seguente grafico (l'allargamento del range nell'ultimo periodo corrisponde ad un aumento della volatilità dei prezzi).



Un'ultima miglioria per il calcolo del VaR venne suggerita da Hull e White i quali pensarono di aggiornare costantemente la volatilità affinché anche i dati più vecchi riflettessero i cambiamenti dati dall'aggiornamento. Se quindi per esempio la volatilità attuale era dello 0,8% e quella stimata tramite i dati dei 20 giorni precedenti era dello 0,6%, piuttosto che usare quest'ultima, essa si sarebbe ridimensionata in modo tale da restituire un effettivo cambiamento: un 1% sarebbe stato convertito in un 1,33%. Questo approccio richiede una costante stima della varianza (e quindi della volatilità) che può essere fornita da vari modelli statistici.

3.1.3 Metodo di simulazione Monte Carlo

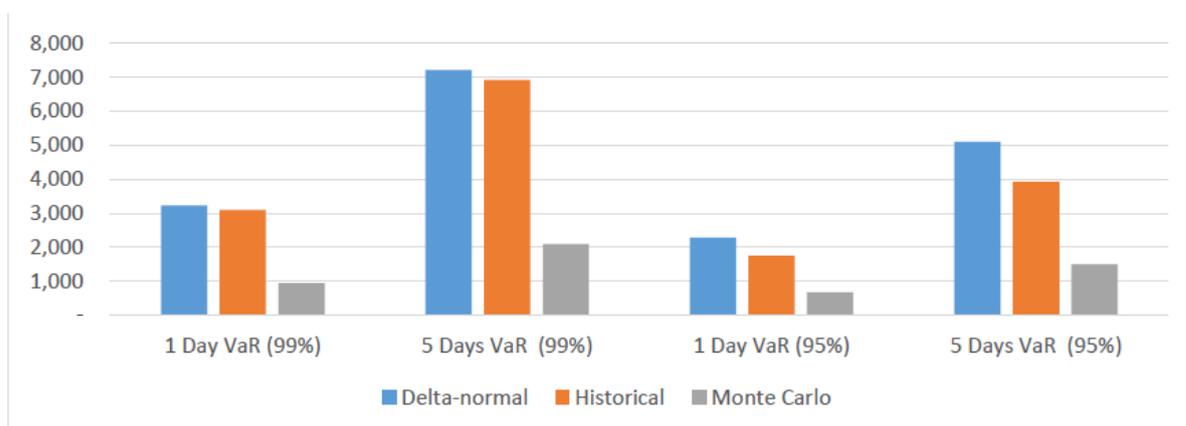
Il terzo ed ultimo metodo che andremo ad analizzare per il calcolo del VaR è il metodo delle simulazioni di Monte Carlo. Nonostante i primi passi ricordino il metodo di varianza-covarianza (si suddivide il portafoglio in vari asset per i quali vengono calcolate varianza e correlazione), per il calcolo di queste ci si affida ad una simulazione in cui vengono specificate le distribuzioni di ogni asset e i vari gradi di correlazione tra questi. La potenza di questo metodo risiede nel poter attribuire ad ogni asset una distribuzione diversa che all'occorrenza può essere modificata basandosi su supposizioni ed opinioni soggettive. Nelle simulazioni l'andamento del portafoglio riflette gli esiti dei vari asset e, dopo migliaia di simulazioni, si avrà una distribuzione dei valori del portafoglio che si utilizzerà per valutare il VaR. Nonostante la potenza delle simulazioni di Monte Carlo, molti preferiscono usare ancora i dati storici per fare assunzioni sulle distribuzioni, questo perchè le simulazioni non vengono considerate attendibili nella gestione di tanti asset e con correlazioni complesse. Il numero di asset su cui fare ipotesi ed il numero di simulazioni da compiere per ottenere l'andamento generale del portafoglio sono due delle maggiori criticità che si possono muovere al metodo. Nonostante ciò, se comparato in generale ai precedenti due metodi, questo risulta essere la via di mezzo ottimale tra questi, fornendo il giusto equilibrio tra assunzioni e libertà di scelta.

Come nei due approcci precedenti anche in questo le principali modifiche che si possono apportare puntano a colmare i difetti. Uno di questi è sicuramente il carico computazionale: anche pochi asset ma fortemente correlati tra loro possono far aumentare esponenzialmente il numero di simulazioni necessarie. Jamshidan e Zhu consigliano di far girare la simulazione su un numero ristretto di scenari e magari facendo sì che asset rischiosi influenzino solo una parte dei valori dello scenario e non tutti. Frye suggerisce invece di sottoporre alla simulazione scenari già preliminarmente perturbati per migliorare la stima finale del VaR. Si possono volendo anche unire i punti di forza del metodo di varianza-covarianza con quelli della simulazione di Monte Carlo andando ad annullarne le rispettive debolezze: abbinando cioè la velocità del primo con la flessibilità nella gestione del rischio del secondo si possono ottenere stime ottimali del VaR. Glasserman, Heidelberger e Shahabuddin impiegarono esattamente questa tecnica usando le approssimazioni fornite dal metodo di varianza-covarianza come guida per il processo di campionamento del metodo Monte Carlo, riportando così un sostanziale risparmio di tempo e risorse mantenendo tuttavia una buona precisione nella stima. Per concludere forniamo un confronto tra i tre approcci.

Esempio 3.6. Confrontiamo il VaR calcolato tramite i tre approcci su un portafoglio di 20 azioni nelle quali vengono investite 5.000 sterline ciascuna, con dei livelli di confidenza del 95% e 99% ed utilizzando orizzonti temporali di uno e cinque giorni⁴. Ovviamente il

⁴A comparison of Value at Risk methods in portfolios with linear and non-linear financial instruments, University of East London for the degree of Finance and Risk MSc, 2016.

VaR calcolato con un orizzonte temporale di cinque giorni è maggiore rispetto a quello di un giorno perchè, con il passare del tempo, le azioni sono soggette a più rischio. Con l'aumentare del livello di confidenza aumenta inoltre anche il VaR poichè un livello di confidenza maggiore richiede anche maggiore precisione. Notiamo come c'è minor discrepanza tra i valori calcolati con il 95% di confidenza in accordo a quanto sperimentato da Hendricks sulla calibrazione del VaR: egli, insieme ad altri analisti, si accorse di come la misurazione del rischio e le problematiche di coda grassa potessero essere aggiustate semplicemente abbassando il livello di confidenza. I due VaR più grandi sono quelli calcolati dal metodo di varianza-covarianza e dalla simulazione storica: il primo è principalmente basato sulla deviazione standard dei rendimenti, il secondo è basato invece sulla volatilità dei dati passati (se la volatilità dei dati passati è alta, il VaR sarà grande) e ciò suggerisce come tra i tre metodi quello da preferirsi sia il metodo di Monte Carlo che rappresenta la scelta più adatta in questo caso. In generale a prescindere dal metodo di calcolo utilizzato, se non vengono soddisfatte le ipotesi relative all'approccio, il VaR calcolato può essere errato e fuorviante, portando così a prendere decisioni sbagliate.



Notiamo come il VaR sia una misura di rischio che si concentra sulle perdite e non sui guadagni prendendo in considerazione solo il rischio di mercato senza considerare altri tipi di rischio: se infatti oltre a questo si considerassero anche il rischio politico o il rischio di liquidità, il VaR calcolato potrebbe essere molto diverso dal VaR reale. Un potenziale svantaggio del VaR è rappresentato anche dall'orizzonte temporale nel quale viene misurato: se tendenzialmente si limita a qualche giorno o poche settimane, esso può essere anche esteso a mesi o anni a patto però che si tenga conto di possibili sovrastrime o sottostime, il che rende il VaR una misura poco affidabile nelle stime in un lungo periodo. L'utilizzo su brevi periodi è dato dal fatto che le società di servizi finanziari sono interessate maggiormente alla copertura del rischio in un breve periodo di tempo; un altro motivo è la precisione sulle stime date dai metodi di varianza-covarianza e di simulazione storica. Un altro difetto del VaR è il non fornire una stima sulle effettive perdite:

quando cioè si è certi delle perdite, il VaR non dà un'indicazione riguardo la portata di queste, focalizzandosi solo sul valore di soglia e non tenendo conto delle perdite sotto tale valore (problema del rischio di coda). Variando anche di poco il livello di confidenza si può avere un VaR completamente diverso. Ciò può essere usato per manipolare la percezione di perdite e guadagni portando ad una sovraesposizione al rischio. Basak e Shapiro notarono infatti come gli investitori che usavano il VaR come misura di rischio investissero in portafogli più rischiosi del normale: essi si concentravano sull'evitare il rischio di scendere sotto la soglia di confidenza esponendosi così a grandi perdite nel caso i loro investimenti fossero andati male. Un dirigente consapevole delle limitazioni e delle potenzialità del VaR potrebbe sfruttare a suo vantaggio questa dinamica compiendo degli investimenti più rischiosi del previsto mantenendosi tuttavia sempre sotto il livello di soglia. Questo approccio si può tuttavia utilizzare solo quando si conoscono gli andamenti passati di un titolo e la sua volatilità (Damodaran, 2007 [5]). Oltre ai vari difetti, il VaR rappresenta in generale un esempio di misura non coerente poichè non soddisfa l'assioma di subadditività. Il non soddisfare l'assioma si traduce concretamente nella penalizzazione della diversificazione di un portafoglio ed in generale la mancanza di coerenza coincide con un atteggiamento nei confronti del rischio non razionale: non associa cioè un valore maggiore ad un rischio maggiore. La coerenza del VaR si può comunque avere in caso di v.a. con distribuzione normale.

Diamo infine un legame tra VaR e dominanza stocastica (Ma, Wong, 2010 [11]). Dal punto di vista della dominanza stocastica, X domina Y se X è meno rischiosa di Y in termini di guadagno: questo ci dice che X ha una probabilità più piccola di scendere sotto un certo valore critico x rispetto ad Y . In termini di valore a rischio questo si traduce nell'aver un VaR minore rispetto a quello di Y a patto che entrambe le v.a. siano relative allo stesso investimento iniziale W_0 . Questo concetto viene formalizzato dal seguente teorema.

Teorema 3.7. Siano X, Y due v.a., allora le seguenti affermazioni sono equivalenti:

- (1) $X \succeq_{FSD} Y$;
- (2) $\forall \alpha \in (0, 1), \text{VaR}_{\alpha,t}(X) \leq \text{VaR}_{\alpha,t}(Y)$;
- (3) $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ per ogni funzione di utilità u limitata e crescente;

Dimostrazione. Sia $X \succeq_{FSD} Y$. Per ogni $\alpha \in (0, 1)$ abbiamo:

$$\alpha \leq F_X(\zeta_X^-(\alpha)) \leq F_Y(\zeta_X^-(\alpha)),$$

dove $\zeta_X^-(\alpha) := \inf\{x : F_X(x) \geq \alpha\}$ e dove la prima disuguaglianza segue da una proprietà di $\zeta_X^-(\cdot)$ e la seconda dall'ipotesi (1). Dalla monotonia di $\zeta_X^-(\cdot)$ otteniamo

$$\zeta_Y^-(\alpha) \leq \zeta_Y^-(F_Y(\zeta_X^-(\alpha))) \leq \zeta_X^-(\alpha),$$

dove la seconda disuguaglianza è data da una proprietà di $\zeta_X^-(\cdot)$.

Al contrario, sia $\text{VaR}_{\alpha,t}(X) \leq \text{VaR}_{\alpha,t}(Y)$ per ogni $\alpha \in (0, 1)$. Notiamo che per ogni $x \in \mathbb{R}$ tale che $F_X(x) > 0$ il valore a rischio per X in $\alpha = F_X(x)$ è dato da

$$\text{VaR}_{F_X(x)}(X) = W_0 - \zeta_X^-(F_X(x)).$$

Abbiamo quindi che

$$x \geq \zeta_X^-(F_X(x)) \geq \zeta_Y^-(F_X(x)),$$

dove la prima disuguaglianza è data da una proprietà di $\zeta_X^-(\cdot)$. Dalla monotonia di $F_Y(\cdot)$ e da una proprietà di $\zeta_X^-(\cdot)$ otteniamo

$$F_Y(x) \geq F_Y(\zeta_Y^-(F_Y(x))) \geq F_X(x).$$

Data l'arbitrarietà di $x \in \mathbb{R}$, concludiamo che $X \succeq_{FSD} Y$. Dell'equivalenza tra (2) e (3) ci sono diverse dimostrazioni, vedere ad esempio Wong e Ma (2008). \square

L'equivalenza tra (2) e (3) informalmente ci dice che gli investitori preferiranno la X alla Y ogni qual volta la Y sarà più rischiosa della X in termini di VaR. Per definire un rapporto di rischio tra X ed Y esso dev'essere valido per ogni livello di confidenza α fissato, per cui il VaR non è sempre lo strumento migliore per definire i rapporti di rischio tra due portafogli.

3.2 Valore a rischio condizionale

Il valore a rischio condizionale (abbreviato CVaR) è una variante del VaR ed è definito come la media della distribuzione di coda relativa ad $\alpha \in (0, 1)$ (Sarykalin, Serraino, Uryasev, 2008 [23]).

Definizione 3.8 (Valore a rischio condizionale). Data una v.a. X che definisce il rendimento di un portafoglio, un livello di confidenza $\alpha \in (0, 1)$ ed un orizzonte temporale t , il *valore a rischio condizionale* di X è definito come:

$$\text{CVaR}_{\alpha,t}(X) := \int_{-\infty}^{\infty} x F_X^\alpha(x) dx,$$

dove

$$F_X^\alpha(x) = \begin{cases} 0, & \text{se } x < \text{VaR}_{\alpha,t}(X) \\ \frac{F_X(x) - \alpha}{1 - \alpha}, & \text{se } x \geq \text{VaR}_{\alpha,t}(X) \end{cases}$$

Una definizione alternativa di CVaR può essere data usando il VaR.

Osservazione 3.9. Sia $\text{CVaR}_{\alpha,t}^+(X)$ il *CVaR superiore* (o *expected shortfall*) definito come

$$\text{ES}_{\alpha,t}(X) = \text{CVaR}_{\alpha,t}^+(X) := \mathbb{E}[X|X > \text{VaR}_{\alpha,t}(X)],$$

allora il CVaR può essere definito come la media pesata del VaR e del CVaR superiore. Supponiamo che $F_X(\text{VaR}_{\alpha,t}(X)) < 1$, in modo tale che ci sia una possibilità di perdita maggiore del VaR, allora il CVaR viene definito come

$$\text{CVaR}_{\alpha,t}(X) := \lambda_\alpha(X)\text{VaR}_{\alpha,t}(X) + (1 - \lambda_\alpha(X))\text{CVaR}_{\alpha,t}^+(X),$$

dove

$$\lambda_\alpha(X) = \frac{F_X(\text{VaR}_{\alpha,t}(X)) - \alpha}{1 - \alpha}.$$

e questa è in $[0, 1]$.

Questa definizione alternativa deriva dal fatto che sia il VaR che il CVaR superiore sono due misure di rischio che non si comportano bene in caso di distribuzioni generali (entrambe infatti sono discontinue rispetto α). Il CVaR invece rappresenta una misura continua e convessa. Se $F_X(\text{VaR}_{\alpha,t}(X)) = 1$ e quindi $\text{VaR}_{\alpha,t}(X)$ può essere la massima perdita possibile, si ha

$$\text{CVaR}_{\alpha,t}(X) = \text{VaR}_{\alpha,t}(X).$$

In generale il CVaR è sempre maggiore del VaR.

Osservazione 3.10. Notiamo che per v.a. X che hanno funzione di distribuzione continua, il *CVaR inferiore*, definito come

$$\text{CVaR}_{\alpha,t}^- := \mathbb{E}[X|X \geq \text{VaR}_{\alpha,t}(X)],$$

che rappresenta il valore atteso delle perdite eccedenti il VaR, è uguale al CVaR. In caso di distribuzioni generali invece il CVaR inferiore è sia discontinuo rispetto ad α che non convesso. Dato che nella pratica si tende a lavorare con distribuzioni continue, la definizione di CVaR coinciderà spesso con quella di CVaR inferiore. Nelle definizioni di CVaR inferiore e superiore la v.a. X indica le perdite di un portafoglio, quindi se questa è positiva si ha una perdita, mentre se è negativa si ha un guadagno.

Dalla precedente formula possiamo ricavare un'altra rappresentazione alternativa del CVaR (Pflug, 2000 [20]), infatti, se $F_X(F_X^{-1}(\alpha)) = \alpha$ vale

$$\mathbb{E}[X|X \geq F_X^{-1}(\alpha)] = \frac{1}{1 - \alpha} \int_\alpha^1 F_X^{-1}(x)dx = \frac{1}{1 - \alpha} \int_\alpha^1 \text{VaR}_{x,t}(X)dx.$$

dove ricordiamo che $F_X^{-1}(\alpha) = \text{VaR}_{\alpha,t}(X)$.

Generalmente vale $\text{CVaR}_{\alpha,t}^-(X) \leq \text{CVaR}_{\alpha,t}(X) \leq \text{CVaR}_{\alpha,t}^+(X)$ con l'uguaglianza quando la distribuzione non ha salti in corrispondenza della soglia del VaR. Tuttavia nelle situazioni pratiche questi si presentano sempre, facendo sì che la catena di disuguaglianze sia stretta.

La vera particolarità di definire il CVaR come combinazione di CVaR superiore e VaR è che in questo modo si può “dividere il contributo” del CVaR. Data infatti una v.a. X con distribuzione associata $F_X(x)$, se quest'ultima ha una discontinuità a salto allora esisteranno due punti della distribuzione, in un intervallo di confidenza α , aventi VaR uguale. I due estremi dell'intervallo saranno

$$\alpha^- = F_X(\text{VaR}_{\alpha,t}^-(X)) \quad \text{e} \quad \alpha^+ = F_X(\text{VaR}_{\alpha,t}(X))$$

dove $F_X(\text{VaR}_{\alpha,t}^-(X)) = P(X < \text{VaR}_{\alpha,t}(X))$.

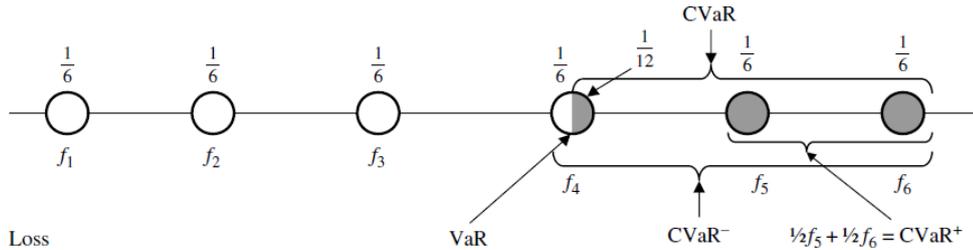
Quando $F_X(\text{VaR}_{\alpha,t}^-(X)) < \alpha < F_X(\text{VaR}_{\alpha,t}(X)) < 1$ allora $\text{VaR}_{\alpha,t}(X)$, che avrà una probabilità totale di $\alpha^+ - \alpha^-$, sarà diviso dal livello di confidenza α in due pezzi con probabilità $\alpha^+ - \alpha$ e $\alpha - \alpha^-$. Diamo due esempi che chiariscono questa spiegazione.

Esempio 3.11. Supponiamo di avere sei scenari di perdita $f_1 \dots f_6$ equiprobabili. Sia $\alpha = 7/12$, allora α “divide” il contributo del CVaR ossia

$$\lambda_\alpha(X) = \frac{F_X(\text{VaR}_{\alpha,t}(X)) - \alpha}{1 - \alpha} > 0,$$

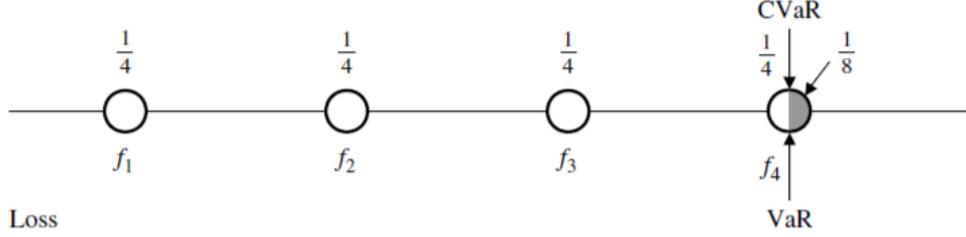
ed il CVaR è dato da

$$\text{CVaR}_{\alpha,t}(X) = \frac{1}{5} \text{VaR}_{\alpha,t}(X) + \frac{4}{5} \text{CVaR}_{\alpha,t}^+(X) = \frac{1}{5} f_4 + \frac{2}{5} f_5 + \frac{2}{5} f_6.$$



Nel secondo caso consideriamo quattro scenari $f_1 \dots f_4$ equiprobabili ed $\alpha = 7/8$ in modo tale che solo l'ultimo scenario venga diviso. In questo caso $\text{VaR}_{\alpha,t}(X) = \text{CVaR}_{\alpha,t}^-(X) = \text{CVaR}_{\alpha,t}(X) = f_4$, mentre il CVaR superiore non è definito. In questa situazione si può definire il CVaR attraverso il CVaR inferiore ed il CVaR superiore. Supponendo cioè la divisione precedente del VaR, si avrà quindi che

$$\text{CVaR}_{\alpha,t}(X) = \frac{\alpha^+ - \alpha}{\alpha^+ - \alpha^-} \frac{1 - \alpha^-}{1 - \alpha} \text{CVaR}_{\alpha,t}^-(X) + \frac{\alpha - \alpha^-}{\alpha^+ - \alpha^-} \frac{1 - \alpha^+}{1 - \alpha} \text{CVaR}_{\alpha,t}^+.$$



Teorema 3.12. In caso di distribuzione continua, il CVaR, al contrario del VaR, è una misura di rischio coerente. In particolare, sfruttando le proposizioni che caratterizzano una misura coerente, il CVaR soddisfa l'invarianza per traslazioni (T), la positiva omogeneità (PH), la convessità (C) e la monotonia (M).

Dimostrazione. (T) e (PH) sono ovvie dalla definizione di $\text{CVaR}_{\alpha,t}(X)$. Dimostriamo (C): sia a_i tale che

$$\text{CVaR}_{\alpha,t}(X) = a_i + \frac{1}{1-\alpha} \mathbb{E}[X_i - a_i]^+,$$

dove $[X_i - a_i]^+ = \max\{0, X_i - a_i\}$. Dalla convessità della funzione $x \rightarrow [x - a]^+$ abbiamo che

$$\begin{aligned} \text{CVaR}_{\alpha,t}(\lambda X_1 + (1-\lambda)X_2) &\leq \lambda a_1 + (1-\lambda)a_2 + \frac{1}{1-\alpha} \mathbb{E}[\lambda X_1 + (1-\lambda)X_2 - \lambda a_1 + (1-\lambda)a_2]^+ \\ &\leq \lambda a_1 + (1-\lambda)a_2 + \frac{\lambda}{1-\alpha} \mathbb{E}[X_1 - a_1]^+ + \frac{1-\lambda}{1-\alpha} \mathbb{E}[X_2 - a_2]^+ \\ &\leq \lambda \text{CVaR}_{\alpha,t}(X_1) + (1-\lambda) \text{CVaR}_{\alpha,t}(X_2). \end{aligned}$$

Infine la (M) segue dalla monotonia e dalla convessità della funzione $x \rightarrow [x - a]^+$. In particolare se la funzione di distribuzione delle perdite è continua, il CVaR è anche subadditivo. \square

Un'altra proprietà del CVaR degna di nota è data dalla seguente

Proposizione 3.13. Data X una v.a. assolutamente continua (avente quindi una densità), allora vale (Pflug, 2000 [20])

$$\mathbb{E}[X] = (1-\alpha) \text{CVaR}_{\alpha,t}(X) - \alpha \text{CVaR}_{(1-\alpha),t}(-X).$$

Dimostrazione. Dato che

$$\begin{aligned}
\text{CVaR}_{(1-\alpha),t}(-X) &= \mathbb{E}[-X | -X > \text{VaR}_{(1-\alpha),t}(-X)] \\
&= \mathbb{E}[-X | -X > -\text{VaR}_{\alpha,t}(X)] \\
&= -\mathbb{E}[X | X < \text{VaR}_{\alpha,t}(X)],
\end{aligned}$$

si nota che

$$\begin{aligned}
\mathbb{E}[X] &= \alpha \mathbb{E}[X | X < \text{VaR}_{\alpha,t}(X)] + (1 - \alpha) \mathbb{E}[X | X > \text{VaR}_{\alpha,t}(X)] \\
&= -\alpha \text{CVaR}_{(1-\alpha),t}(-X) + (1 - \alpha) \text{CVaR}_{\alpha,t}(X),
\end{aligned}$$

da cui la tesi □

Una relazione tra VaR e CVaR è data da

Osservazione 3.14. Sia X una v.a. non negativa, allora

$$\left(\frac{\mathbb{E}[X^n] - (1 - \alpha) \text{CVaR}_{\alpha,t}(X^n)}{\alpha} \right)^{\frac{1}{n}} \rightarrow \text{VaR}_{\alpha,t}(X).$$

Dimostrazione. La dimostrazione è basata sul fatto che per ogni v.a. non negativa vale

$$\lim_{n \rightarrow \infty} (\mathbb{E}[X^n])^{\frac{1}{n}} \rightarrow \inf\{u : P(X > u) = 0\},$$

dalla positiva omogeneità del CVaR si ha

$$-\text{CVaR}_{(1-\alpha),t}(-X) = \frac{1}{\alpha} (\mathbb{E}[X] - (1 - \alpha) \text{CVaR}_{\alpha,t}(X)).$$

Dall'altra parte si avrà

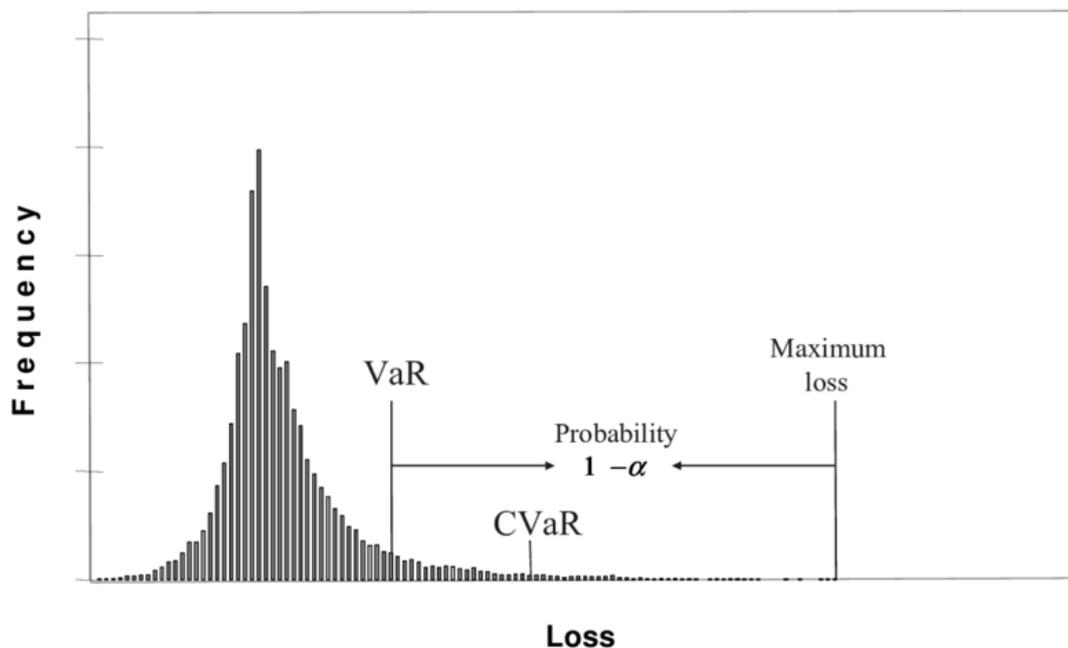
$$-\text{CVaR}_{(1-\alpha),t}(-X) = \mathbb{E}[X | X \leq \text{VaR}_{\alpha,t}(X)] = \sup\{a - \frac{1}{\alpha} \mathbb{E}[X - a]^- : a \in \mathbb{R}\},$$

dove $[X - a]^- = -\min\{0, X - a\}$. Dato che vale

$$(\mathbb{E}[X^n | X^n \leq \text{VaR}_{\alpha,t}(X^n)])^{\frac{1}{n}} = (\mathbb{E}[X^n | X \leq \text{VaR}_{\alpha,t}])^{\frac{1}{n}} \rightarrow \text{VaR}_{\alpha,t}(X),$$

da ciò segue la tesi. □

In caso di una distribuzione normale non ci sono differenze tra VaR e CVaR. Sostanziali diversità emergono invece quando si lavora con distribuzioni non normali. In generale, in questo caso, il CVaR colma la maggior parte delle lacune del VaR: oltre la coerenza, focalizzandosi sulle perdite oltre il VaR, si concentra sulla coda della distribuzione (in alcuni casi il CVaR è chiamato anche VaR di coda) cercando quindi di monitorare le grandi perdite (Yamai, Yoshiba, 2002 [31]). In particolare il CVaR trascura tutti gli esiti sopra l' α -quantile e dà lo stesso peso a tutti quelli sotto l' α -quantile.



Nonostante questo il CVaR necessita di più dati rispetto al VaR per ottenere una stima corretta delle perdite ed in caso di distribuzioni a coda grassa la stima del CVaR potrebbe non essere così precisa. La mancanza di precisione è da imputare ad andamenti imprevedibili e a grandi perdite, le quali possono portare ad un errore nel calcolo del CVaR maggiore rispetto a quello del VaR: nel caso di problemi di coda, infatti, variando anche di poco il livello di confidenza α , si potrà avere un valore del CVaR molto diverso.

Portiamo alcuni esempi di stima dove si mettono a confronto VaR e CVaR⁵. In tutti i casi la distribuzione considerata è sempre continua (quindi varrà $\text{CVaR}_{\alpha,t}^-(X) = \text{CVaR}_{\alpha,t}(X)$). Nel primo esempio la distribuzione dipende da un parametro α chiamato indice di stabilità. Questo descrive la coda della distribuzione: più è piccolo, più essa sarà grassa. Se $\alpha = 2$ la distribuzione quindi coinciderà con la normale, se $\alpha = 1$ allora coinciderà con la distribuzione di Cauchy.

Esempio 3.15. Consideriamo 10.000 simulazioni di Monte Carlo con un campione di 1.000 scenari e con $\alpha = 2.0, 1.9, \dots, 1.1$. Otteniamo quindi 10.000 stime di VaR e CVaR di cui calcoliamo media, deviazione standard ed intervallo di confidenza del 95%. Gli errori di stima di VaR e CVaR sono comparati attraverso la deviazione standard relativa (la deviazione standard divisa per la media). Notiamo che l'errore di stima del CVaR è maggiore rispetto a quello del VaR quando la distribuzione è a coda grassa: man mano

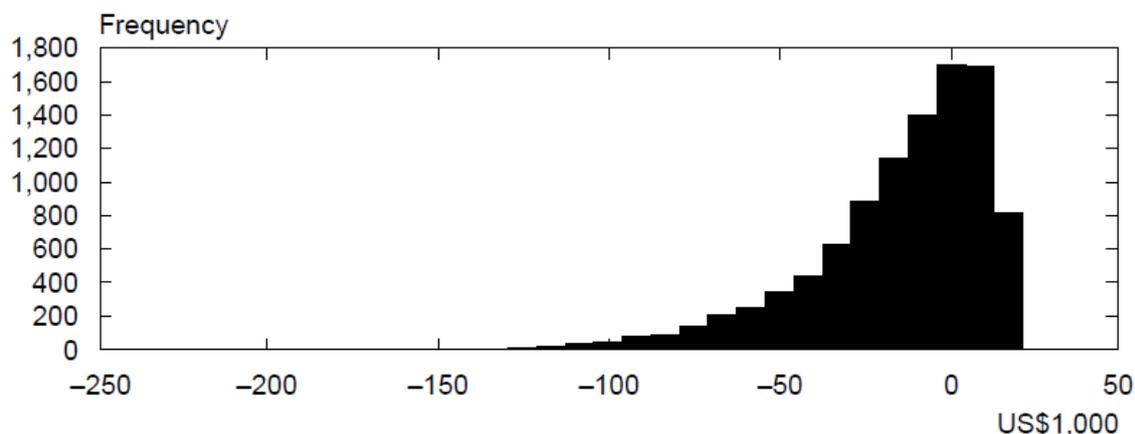
⁵Comparative Analyses of Expected Shortfall and Value-at-Risk: Their Estimation Error, Decomposition, and Optimization, Yasuhiro Yamai, Toshinao Yoshida, Monetary and Economic studies, January 2002.

che α si avvicina ad 1, la deviazione standard relativa del CVaR diventa molto più grande rispetto a quella del VaR. Dall'altra parte quando α si avvicina a 2 le due deviazioni standard relative sono quasi uguali. Nel caso di un livello di confidenza del 99% la differenza tra deviazioni standard relative di VaR e CVaR diventa ancora più marcata. Quando le distribuzioni sono a coda grassa, la probabilità di grandi ed infrequenti perdite è molto alta e, visto che il CVaR si focalizza sulla coda della distribuzione, esso sarà più sensibile ai fenomeni che la riguardano.

| α | Misura di rischio | Media | Deviazione standard | Deviazione standard relativa | Intervallo di confidenza |
|----------|-------------------|-------|---------------------|------------------------------|--------------------------|
| 2,0 | VaR | 1,64 | 0,07 | 0,04 | [1,51 1,77] |
| | CVaR | 2,05 | 0,08 | 0,04 | [1,90 2,21] |
| 1,9 | VaR | 1,70 | 0,08 | 0,04 | [1,55 1,85] |
| | CVaR | 2,42 | 0,80 | 0,33 | [2,06 3,14] |
| 1,8 | VaR | 1,77 | 0,09 | 0,05 | [1,60 1,95] |
| | CVaR | 2,90 | 1,81 | 0,63 | [2,28 4,20] |
| 1,7 | VaR | 1,86 | 0,11 | 0,06 | [1,67 2,08] |
| | CVaR | 3,53 | 3,84 | 1,09 | [2,58 5,60] |
| 1,6 | VaR | 1,98 | 0,13 | 0,07 | [1,75 2,26] |
| | CVaR | 4,39 | 8,34 | 1,90 | [2,96 7,62] |
| 1,5 | VaR | 2,15 | 0,16 | 0,08 | [1,86 2,50] |
| | CVaR | 5,67 | 19,31 | 3,41 | [3,48 10,71] |
| 1,4 | VaR | 2,38 | 0,21 | 0,09 | [2,02 2,82] |
| | CVaR | 7,71 | 48,95 | 6,35 | [4,16 15,76] |
| 1,3 | VaR | 2,68 | 0,26 | 0,10 | [2,22 3,25] |
| | CVaR | 11,46 | 139,60 | 12,19 | [5,10 25,13] |
| 1,2 | VaR | 3,08 | 0,34 | 0,11 | [2,49 3,85] |
| | CVaR | 19,79 | 463,10 | 23,40 | [6,48 42,45] |
| 1,1 | VaR | 3,65 | 0,46 | 0,13 | [2,86 4,67] |
| | CVaR | 44,41 | 1.866,40 | 42,03 | [8,59 81,44] |

Esempio 3.16. Nel secondo esempio prendiamo in considerazione tre azioni i cui rendimenti supponiamo abbiano una distribuzione normale multivariata con le loro relative opzioni di acquisto ed adottiamo una posizione corta. Fissiamo come periodo di riferimento un giorno e supponiamo che la volatilità sia costante per tutto il periodo. Per stimare VaR e CVaR usiamo il metodo parametrico e stimiamo la matrice di covarian-

za ottenuta tramite dati storici. L'andamento delle perdite e dei guadagni è dato dal seguente grafico.

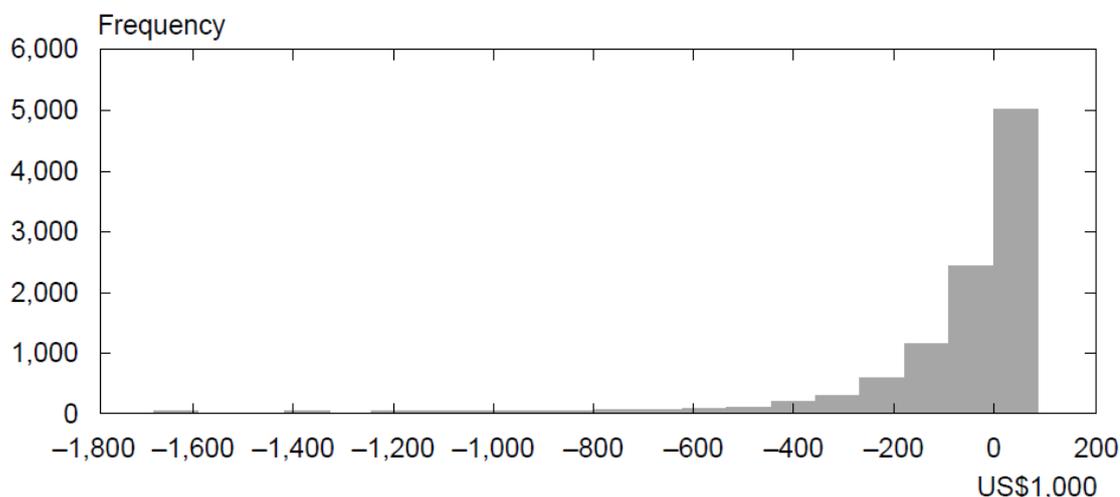


Il grafico presenta un'asimmetria negativa data dall'aver adottato una posizione corta per le opzioni di acquisto. Valutiamo ora gli errori nella stima di VaR e CVaR usando il metodo di Monte Carlo. In particolare prendiamo 1.000 simulazioni con dei campioni di 1.000, 10.000 e 100.000 scenari. Calcoliamo quindi media, deviazione standard e stime del VaR e del CVaR con un livello di confidenza del 95%. I risultati sono mostrati nella seguente tabella.

| Campione | Misura di rischio | Media | Deviazione standard | Deviazione standard relativa | Intervallo di confidenza |
|----------|-------------------|-------|---------------------|------------------------------|--------------------------|
| 1.000 | VaR | 68,33 | 3,34 | 0,0489 | [61,72 75,03] |
| | CVaR | 91,20 | 4,63 | 0,0508 | [82,44 100,30] |
| 10.000 | VaR | 68,18 | 1,05 | 0,0154 | [66,10 70,22] |
| | CVaR | 91,56 | 1,38 | 0,0151 | [88,97 94,26] |
| 100.000 | VaR | 68,15 | 0,33 | 0,0049 | [67,51 68,81] |
| | CVaR | 91,57 | 0,46 | 0,0050 | [90,73 92,50] |

Gli errori di stima di VaR e CVaR sono quasi uguali, questo perchè è stata scelta una determinata opzione adottando una strategia corta. Errori di stima analoghi si ottengono per un livello di confidenza del 99%.

Esempio 3.17. Consideriamo ora lo stesso portafoglio precedente ma con un'opzione *out of the money*⁶. Il grafico delle perdite e dei guadagni sarà il seguente.



In questo caso l'asimmetria negativa è ancora più accentuata poiché si è scelta l'opzione *out of the money*. Gli errori nella stima di VaR e CVaR sono calcolati usando gli stessi parametri del caso precedente.

| Campione | Misura di rischio | Media | Deviazione standard | Deviazione standard relativa | Intervallo di confidenza |
|----------|-------------------|--------|---------------------|------------------------------|--------------------------|
| 1.000 | VaR | 334,62 | 22,01 | 0,0658 | [293,70 380,86] |
| | CVaR | 500,46 | 35,00 | 0,0699 | [433,38 571,32] |
| 10.000 | VaR | 332,56 | 7,06 | 0,0212 | [318,53 346,88] |
| | CVaR | 502,01 | 11,17 | 0,0222 | [481,57 524,05] |
| 100.000 | VaR | 332,26 | 2,15 | 0,0065 | [328,14 336,70] |
| | CVaR | 502,19 | 3,45 | 0,0069 | [495,53 509,05] |

In questo caso l'errore di stima del CVaR è maggiore rispetto a quello del VaR poiché la distribuzione sarà a coda grassa. Anche se aumentiamo il livello di confidenza al 99% le differenze di errore nella stima saranno destinate ad aumentare.

In sintesi VaR e CVaR si concentrano su parti diverse della distribuzione e possono essere scelti in base ad interessi specifici. Se ad esempio si è propensi al rischio e non si

⁶Assumendo che la distribuzione dei rendimenti sia una normale multivariata e che la volatilità sia costante, la probabilità che i prezzi delle azioni crescano più dei prezzi d'esercizio è meno dello 0,01%.

hanno buoni modelli di stima della coda della distribuzione, allora si tenderà a scegliere il primo. Quando ad esempio si usa un metodo di simulazione storica esso probabilmente stimerà meglio il VaR poichè non sarà adatto ad individuare problematiche di coda grassa o pesante, di conseguenza il CVaR ottenuto non si potrà utilizzare per monitorare correttamente il rischio. Se invece ci si vorrà tutelare da grandi perdite, interessandosi quindi ai comportamenti sulla coda e potendola stimare correttamente, allora si preferirà il CVaR: quest'ultimo ha infatti proprietà più utili del VaR e può essere usato anche per problemi di ottimizzazione del portafoglio. Bisogna infine calcolare VaR e CVaR con opportuni livelli di confidenza e non con lo stesso per entrambi poichè essi si focalizzano su due parti diverse della distribuzione. Per avere una visione più ampia e chiara nella gestione del rischio di un asset, occorre quindi combinare le misure di VaR e CVaR.

Come già anticipato nella sezione sulle misure coerenti, il CVaR fa parte della famiglia delle misure spettrali di rischio. Vale infatti (Acerbi, 2002 [1])

$$\text{CVaR}_{\alpha,t}(X) = M_{\phi}(X) = - \int_0^1 F_X(m)\phi(m)dm, \quad \text{dove} \quad \phi(m) = \frac{1}{\alpha} \mathbb{1}_{(0 \leq m \leq \alpha)}.$$

La condizione di stretta decrescenza di ϕ permette al CVaR di attribuire un peso maggiore ad una perdita maggiore. In questo modo si può analizzare con più precisione il comportamento della distribuzione nella coda e quindi riconoscere subito determinate problematiche legate al rischio di coda. La funzione ϕ si occupa nel concreto di “pesare” il rischio al variare di m e la condizione $\|\phi\| = 1$ ci assicura che la somma dei pesi sia uguale ad 1. Affinchè un investitore possa comparare portafogli con diverso rischio e con un certo livello di confidenza dovrà quindi combinare il valore di $\phi(m)$ con quello di $\alpha \in (0, 1)$.

3.3 Momenti parziali inferiori

Un altro insieme di misure di rischio al ribasso che si focalizza interamente sui rendimenti di coda di una distribuzione è quello dei momenti parziali inferiori (in inglese *Lower Partial Moments* o abbreviato LPMs). Essi, introdotti inizialmente da Bawa nel 1975, similmente al CVaR, si concentrano sulla coda della distribuzione attribuendo però un maggior peso ai rendimenti negativi più grandi, dando quindi un significato intuitivo alle perdite in base all’impatto che esse hanno sul portafoglio. I LPMs non considerano solo l’entità dei rendimenti negativi ma anche la loro frequenza: un portafoglio che ha piccole perdite ma frequenti può infatti avere un LPM più grande rispetto ad uno con perdite poco frequenti ma più gravi. Viene fissata innanzitutto una soglia di riferimento z che divide le perdite dai guadagni e rappresenta una specie di “tolleranza” al rischio. Considerato $q = \max\{i : x_i < z\}$ allora questo indicherà il numero di eventi di deficit

al fronte degli eventi totali x_1, \dots, x_n . La soglia di riferimento può essere sia positiva che negativa: nel primo caso segnerà un rendimento accettabile sotto il quale è meglio non andare, nel secondo segnerà un limite oltre il quale le perdite rappresentano un problema (Hahn, Pfingsten, Wagner, 2002 [9]).

Definizione 3.18 (Momento parziale inferiore di ordine m). Dati degli eventi di deficit x_1, \dots, x_q e un valore fissato z , il *momento parziale inferiore di ordine m* viene definito come

$$\text{LPM}_m = \frac{1}{n} \sum_{i=1}^q (z - x_i)^m.$$

Alternativamente i LPMs di ordine m possono essere definiti anche introducendo una CDF continua F_X associata ad una v.a. X (Sunoj, Maya, 2008 [26]), come

$$\text{LPM}_m = \int_{-\infty}^z (z - x)^m dF(x), \quad (3.1)$$

dove z è la soglia fissata ed x è un esito della distribuzione in $(-\infty, z)$.

Esempio 3.19. Il momento parziale inferiore di ordine zero rappresenta la probabilità che si presenti un evento di deficit dati n eventi totali (probabilità di deficit) ossia la probabilità di avere un rendimento minore del target ($\text{LPM}_0 = q/n$).

Esempio 3.20. Il momento parziale inferiore di ordine uno non è altro che il deficit atteso, cioè la media dei rendimenti sotto il target (ai rendimenti sopra il target viene dato un deficit nullo). Se la soglia di riferimento coincide con il VaR allora questo sarà uguale all'*expected shortfall*.

Esempio 3.21. Il momento parziale inferiore di ordine due è il deficit medio al quadrato. Se $z = \mathbb{E}[X]$ allora questo coinciderà con la semivarianza.

Fu Bawa a scoprire il legame tra i LPMs e la dominanza stocastica. Egli concluse che la dominanza stocastica permetteva di attribuire ad un certo LPM una determinata relazione di dominanza stocastica. In particolare soddisfare le relazioni di prima, seconda e terza dominanza stocastica (rispettivamente FSD, SSD, TSD) coincideva con il massimizzare l'utilità per determinate classi di funzioni di utilità (Unser, 1998 [27]).
Date

$$\begin{aligned} U_1 &:= \{u(x) : u'(x) > 0 \forall x \in \mathbb{R}\}; \\ U_2 &:= \{u(x) : u'(x) > 0; u''(x) < 0 \forall x \in \mathbb{R}\}; \\ U_3 &:= \{u(x) : u'(x) > 0; u''(x) < 0; u'''(x) > 0 \forall x \in \mathbb{R}\}; \end{aligned}$$

allora massimizzare U_1 era equivalente alla FSD, massimizzare U_2 era equivalente alla SSD e massimizzare U_3 era equivalente alla TSD.

Tutte e tre identificano una specifica classe di investitori che devono prendere decisioni. U_1 identifica la classe di chi, a prescindere dall'atteggiamento nei confronti del rischio, preferisce ottenere più soldi, U_2 identifica solo chi è avverso al rischio, mentre U_3 oltre alla condizione di U_2 identifica chi preferisce una distribuzione con asimmetria positiva ed ha una funzione di assoluta avversione al rischio crescente o decrescente.

Queste tre classi di funzioni di utilità sono legate a determinate relazioni tra LPMs.

Proposizione 3.22. Date X, Y due v.a. e $F_X(x)$ e $F_Y(x)$ le rispettive CDF associate, allora F_X è preferita ad F_Y se e solo se valgono:

- $\forall u(x) \in U_1, \text{LPM}_{0, F_X}(t) \leq \text{LPM}_{0, F_Y}(t) \forall t \in \mathbb{R},$
 - $\forall u(x) \in U_2, \text{LPM}_{1, F_X}(t) \leq \text{LPM}_{1, F_Y}(t) \forall t \in \mathbb{R},$
 - $\forall u(x) \in U_3, \text{LPM}_{2, F_X}(t) \leq \text{LPM}_{2, F_Y}(t) \forall t \in \mathbb{R},$
- e per quest'ultima condizione vale inoltre $\mathbb{E}[X] \geq \mathbb{E}[Y].$

A seconda dei casi vengono scelti il valore di soglia t e l'ordine m . Nello specifico più si è avversi al rischio, più grande sarà il valore t e più grande sarà scelto il valore di m , dando quindi più peso agli esiti che si allontanano dal valore di soglia. A valori di m in $(0, 1)$ corrisponderà un atteggiamento di propensione al rischio, mentre se $m = 1$ si avrà una neutralità al rischio.

Mostriamo ora alcune relazioni tra i LPMs (Sunoj, Maya, 2008 [26]).

Da (3.1) segue che:

$$\text{LPM}_m = m \int_{-\infty}^t (t-x)^{m-1} F_X(x) dx,$$

e derivando (3.1) otteniamo

$$\frac{d\text{LPM}_m(t)}{dt} = m \int_{-\infty}^t (t-x)^{m-1} f(x) dx,$$

dove $f(x)$ è la funzione di densità associata alla v.a. X . Iterando il passaggio precedente m volte si ottiene

$$\frac{d^m \text{LPM}_m(t)}{dt^m} = m! F_X(t), \quad \text{dove} \quad F(t) = \int_{-\infty}^t f(x) dx,$$

o equivalentemente

$$F_X(t) = \frac{\text{LPM}_m^{(m)}(t)}{m!}, \tag{3.2}$$

dove $\text{LPM}_m^{(m)}(t)$ è la derivata m -esima di $\text{LPM}_m(t)$ rispetto t . Dall'ultima uguaglianza si nota che le derivate del LPM di ordine m determinano univocamente la CDF di una v.a.

Il rapporto tra due LPMs consecutivi è espresso dal seguente

Teorema 3.23. Data X v.a. non negativa, e la sua funzione di densità $f(t)$ tale che $\lim_{t \rightarrow 0} tf(t) = 0$ allora vale

$$\frac{\text{LPM}_m(t)}{\text{LPM}_{m-1}(t)} = \gamma t, \quad (3.3)$$

dove γ è una costante in $(0, 1)$ che caratterizza una distribuzione di potenza con CDF

$$F(t) = \left(\frac{t}{b}\right)^c, \quad \text{con } t \in (0, b) \text{ e } b, c > 0. \quad (3.4)$$

Dimostrazione. Supponiamo che valga la (3.3), usando la (3.1) abbiamo che

$$\int_0^t (t-x)^m f(x) dx = \gamma t \int_0^t (t-x)^{m-1} f(x) dx.$$

Usando $t = t - x + x$ nell'uguaglianza precedente e semplificando otteniamo che

$$(1 - \gamma) \int_0^t (t-x)^m f(x) dx = \gamma \int_0^t x(t-x)^{m-1} f(x) dx.$$

Derivando questa m volte ed usando la (3.2) otteniamo che

$$(1 - \gamma)F_X(t) = \gamma t f(t),$$

oppure

$$\lambda(t) = \frac{1 - \gamma}{\gamma t}.$$

Ora dalla proprietà di unicità di $\lambda(t)$ e sfruttando che

$$F(t) = \exp\left(-\int_t^\infty \lambda(x) dx\right),$$

allora $\lambda(t)$ corrisponde dalla distribuzione di potenza con CDF (3.4).

Al contrario se supponiamo la (3.4), sostituendo $f(t)$ nella (3.1) e semplificando otteniamo la (3.3). \square

Si possono anche dare delle caratterizzazioni delle distribuzioni esponenziale ed uniforme tramite i seguenti teoremi di cui non forniremo una dimostrazione.

Teorema 3.24. Sia X una v.a. non negativa, allora i LPMs di ordine m soddisfano

$$\text{LPM}_m(t) + \gamma \text{LPM}_{m-1}(t) = t^m,$$

con $\gamma > 0$ per ogni $t > 0$ se e solo se X ha legge Exp_λ con CDF

$$F(t) = 1 - e^{-\lambda t} \quad \text{con } t, \lambda > 0.$$

Teorema 3.25. Sia X una v.a. non negativa, allora i LPMs di ordine m soddisfano

$$\text{LPM}_m(t) = \gamma(t - a)^{m+1},$$

con $\gamma > 0$ per ogni $t > 0$ se e solo se X ha legge Unif_n con funzione di densità

$$f(t) = \frac{1}{b - a} \quad \text{con } t \in (a, b).$$

I LPMs giocano un ruolo importante anche nel rapporto tra rischio e disuguaglianze economiche. Facendo un ragionamento intuitivo le disuguaglianze economiche sorgono quando non tutti gli individui di una società guadagnano gli stessi soldi in un dato periodo, mentre una distribuzione dei rendimenti viene definita rischiosa quando ci sono eventi che hanno un valore diverso in uno stesso portafoglio. Sono stati condotti numerosi studi sulla relazione tra i due argomenti ed è stato scoperto che i LPMs forniscono informazioni importanti sulla vulnerabilità alla povertà. Questa fa parte di un contesto più ampio ossia quello relativo alle misure di povertà. Esse, insieme alle misure di rischio al ribasso, fissata una soglia t che divide gli eventi, focalizzano le loro analisi sugli esiti sotto questa; in particolare ignorano la maggior parte degli esiti di una distribuzione, concentrandosi prevalentemente sugli esiti di un singolo individuo o delle persone sotto una soglia di povertà fissata. A questi viene dato un notevole peso mentre non viene dato alcun peso a tutti quelli leggermente sopra la soglia. Per le misure di rischio al ribasso la t dividerà gli eventi critici con valori di portafoglio minori o uguali di questa dagli eventi non critici con valori di portafoglio maggiori, mentre per le misure di povertà, la t dividerà i poveri dai non poveri (Sunoj, Maya, 2008 [26]).

Data una v.a. X che rappresenta il reddito di una comunità, viene definito una reddito minimo rappresentato dalla soglia di povertà t : gli individui che hanno una reddito inferiore alla soglia ($x_i < t$) sono quindi definiti poveri. La v.a. ha valori $t - x$ e zero per gli individui il cui reddito è rispettivamente sotto o sopra la soglia t . In questo caso il LPM di ordine zero ci dà una proporzione dei poveri rispetto all'intera comunità. La loro distribuzione dei redditi è data da

$$F_X(x) = \begin{cases} \frac{\text{LPM}_0(x)}{\text{LPM}_0(t)}, & \text{se } X \leq t \\ 1, & \text{se } X > t \end{cases}$$

Nell'analisi del reddito, un importante indice che quantifica la povertà è l'*income gap ratio* (IGR), indicato con $\alpha(t)$ dove il divario di reddito di un individuo è $t - x$. L'IGR è un caso particolare di una famiglia di indici di povertà chiamati *indici di Foster-Greer-Thorbecke* (FGT) i quali si occupano di stimare il numero di poveri nel futuro. La famiglia di indici dipende da un parametro α positivo ed è definita come

$$\text{FGT}_\alpha = \frac{1}{N} \sum_{i=1}^H \left(\frac{t - x_i}{t} \right)^\alpha,$$

dove N è il numero di individui nella comunità ed H è il numero di poveri. Il parametro α indica il peso da dare agli individui: se questo è basso allora tutti gli individui verranno pesati più o meno identicamente mentre all'aumentare di α un peso sempre maggiore verrà attribuito agli individui più poveri. In generale più l'FGT è grande, più ci sarà povertà nella comunità. Un'altra misura importante in questo campo è il *reddito medio sotto il livello di povertà*, indicato con $\mu(t)$. Tutti gli indici sono definiti a partire dai LPMs:

$$\alpha(t) = \frac{\text{LPM}_1(t)}{t\text{LPM}_0(t)},$$

$$\mu(t) = t - \frac{\text{LPM}_1(t)}{\text{LPM}_0(t)}.$$

Dalle precedenti definizioni segue che:

$$\mu(t) = t(1 - \alpha(t)).$$

Diamo ora un esempio concreto di uso dei LPMs come indicatori di vulnerabilità alla povertà. La vulnerabilità può essere definita come la povertà prevista (Witt, Waibel, 2009 [29]), ossia la probabilità di diventare poveri nel futuro il che è equivalente alla definizione di LPM di ordine zero.

Esempio 3.26. Lo studio viene fatto analizzando quattro attività di sostentamento in una provincia del Camerun settentrionale nel maggio del 2008: queste attività sono la raccolta di sorgo, riso, miglio e la pesca. La comunità presa in esame conta 238 famiglie di lavoratori⁷. Fissata la soglia di povertà t come il 50% del reddito medio di 354USD, in questo studio vengono messe a confronto le misure di povertà (indicate dagli indici FGT di ordine zero, uno e due) e quelle di vulnerabilità (indicate dai LPMs di ordine zero, uno e due).

⁷Lower partial moments as a measure of vulnerability to poverty in Cameroon, Rudolf Witt, Hermann Waibel, Institute of Development and Agricultural Economics, Faculty of Economics and Management, Leibniz Universität Hannover, Discussion Paper No. 434, November 2009.

| Poveri | | | | | | |
|--|--------------|---------------------------------|----------------------------------|--------------------------------|------------------|---------------|
| | | Coltivatori di sorgo | Coltivatori di miglio | Coltivatori di riso | Pescatori | Totale |
| Numero | | 9 | 10 | 45 | 3 | 67 |
| Reddito medio | | 129,99 | 126,31 | 101,60 | 111,37 | 109,54 |
| Deviazione standard reddito | | 41,05 | 32,34 | 30,54 | 40,12 | 32,65 |
| FGT | $\alpha = 0$ | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| | $\alpha = 1$ | 30,03 | 50,69 | 60,44 | 35,63 | 67,48 |
| | $\alpha = 2$ | 3.675 | 4.016 | 7.311 | 6.288 | 6.285 |
| LPM | $m = 0$ | 0,80 | 0,85 | 0,91 | 0,84 | 0,88 |
| | $m = 1$ | 53,42 | 54,29 | 78,21 | 71,01 | 70,99 |
| | $m = 2$ | 5.306 | 5.084 | 8.426 | 7.620 | 7.472 |

| Non poveri | | | | | | |
|--|---------|---------------------------------|----------------------------------|--------------------------------|------------------|---------------|
| | | Coltivatori di sorgo | Coltivatori di miglio | Coltivatori di riso | Pescatori | Totale |
| Numero | | 82 | 17 | 45 | 27 | 171 |
| Reddito medio | | 438,59 | 364,09 | 393,77 | 631,38 | 449,83 |
| Deviazione standard reddito | | 163,52 | 91,61 | 77,97 | 191,20 | 138,23 |
| LPM | $m = 0$ | 0,11 | 0,11 | 0,07 | 0,06 | 0,09 |
| | $m = 1$ | 10,28 | 4,22 | 3,24 | 4,68 | 6,94 |
| | $m = 2$ | 2.526 | 422,47 | 301,52 | 669,11 | 1.439 |

| Poveri e non poveri | | | | | | |
|--|--------------|---------------------------------|----------------------------------|--------------------------------|------------------|---------------|
| | | Coltivatori di sorgo | Coltivatori di miglio | Coltivatori di riso | Pescatori | Totale |
| Numero | | 91 | 27 | 90 | 30 | 238 |
| Reddito medio | | 408,07 | 276,02 | 247,69 | 579,38 | 354,04 |
| Deviazione standard reddito | | 151,41 | 69,66 | 54,26 | 176,10 | 108,51 |
| FGT | $\alpha = 0$ | 0,10 | 0,37 | 0,50 | 0,10 | 0,28 |
| | $\alpha = 1$ | 4,65 | 18,78 | 37,71 | 6,56 | 19,00 |
| | $\alpha = 2$ | 363,50 | 1.487 | 3.656 | 628,82 | 1.769 |
| LPM | $m = 0$ | 0,18 | 0,38 | 0,49 | 0,14 | 0,31 |
| | $m = 1$ | 14,54 | 22,76 | 40,73 | 11,31 | 24,97 |
| | $m = 2$ | 2.801 | 2.149 | 4.364 | 1.364 | 3.137 |

Dai dati nell'ultima tabella notiamo che il 28% delle persone della comunità è povera con un divario medio di povertà di 64,5USD. La povertà non è inoltre equamente distribuita tra i vari lavoratori: mentre solo il 10% dei coltivatori di sorgo e dei pescatori ha un reddito sotto la soglia di povertà, i coltivatori di miglio e riso poveri sono rispettivamente il 37% ed il 50% dei totali. Le stesse differenze sono osservate nel divario medio di povertà dove i coltivatori di riso hanno il divario più grande tra i poveri (60,44USD a persona) ed in generale hanno un divario di 37,71USD rispetto alla comunità. In termini di vulnerabilità la probabilità di diventare poveri è del 31% con un deficit medio di circa 71USD. Anche in questo caso i coltivatori di riso sono i lavoratori più colpiti tra le categorie: quelli poveri hanno un deficit atteso di 78USD seguiti dai pescatori con un deficit di 71USD, mentre i coltivatori di sorgo e miglio hanno un deficit atteso molto più basso (53USD e 54USD rispettivamente). Se consideriamo i lavoratori non poveri la situazione cambia molto: in questo caso i soggetti più vulnerabili alla povertà sono i coltivatori di sorgo con l'11% di deficit medio ed il 10,3USD di deficit atteso mentre al contrario i coltivatori di riso sono quelli meno vulnerabili in termini di deficit atteso (3,24USD). Osserviamo che i pescatori non poveri hanno il reddito medio più alto (631,4USD) ma la grande variabilità di questo fa sì che questa categoria sia la più vulnerabile al rischio. Al contrario i coltivatori di riso non poveri hanno il reddito medio basso (relativamente a quello degli altri) ma, avendo una deviazione standard bassa del reddito, sono meno vulnerabili al rischio. Nonostante questo a causa del grande numero di coltivatori di riso poveri, la povertà media e l'incidenza della vulnerabilità sono le più grandi per questa categoria. Per quanto riguarda gli indici FGT, se li consideriamo come misura di povertà cronica, ciò ci dice che i coltivatori di riso e miglio soffrono di povertà cronica mentre quelli di sorgo ed i pescatori soffrono di una povertà temporanea. In conclusione i risultati sulla vulnerabilità mostrano che il rischio di ribasso è un rischio che prescinde dalla posizione delle famiglie della comunità rispetto alla soglia di povertà.

Capitolo 4

Indici di rischio

Oltre all'accenno agli indici di povertà, analizziamo ora alcuni indici importanti in economia e finanza per valutare la performance di un investimento.

4.1 Indice di Sharpe

Uno degli indici più importanti per studiare il valore e la razionalità di un investimento in un portafoglio è l'indice di Sharpe. Questo indice deriva dal *Capital Asset Pricing Model* (abbreviato CAPM) ossia un modello di equilibrio dei mercati finanziari che mette in relazione il rendimento di un titolo con la sua rischiosità e si basa, oltre che sul metodo di media-varianza di Markowitz, anche sullo studio delle funzioni di utilità del singolo investitore (Sharpe, 1964; Yang, 2021 [24][32]).

Definizione 4.1 (Indice di Sharpe ex ante). Dato R_p il rendimento del portafoglio, R_f il tasso d'interesse privo di rischio, l'*indice di Sharpe* è definito come

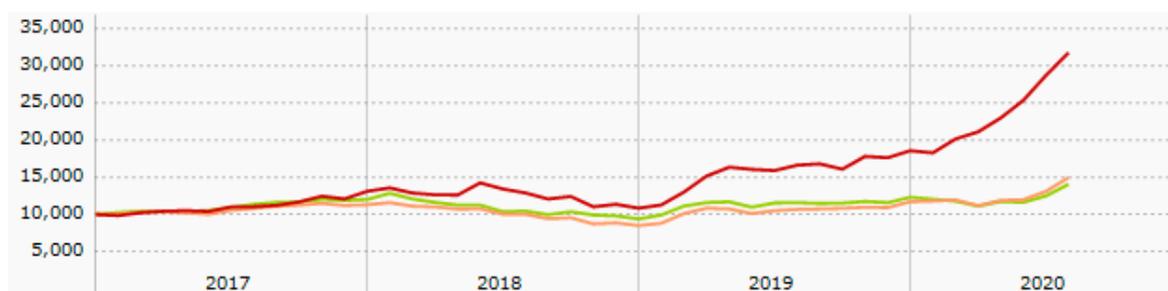
$$SR = \frac{\mathbb{E}[R_p - R_f]}{\sigma_d},$$

dove σ_d è la deviazione standard di $R_p - R_f$ (chiamato rendimento differenziale)(Sharpe, 1994 [25]).

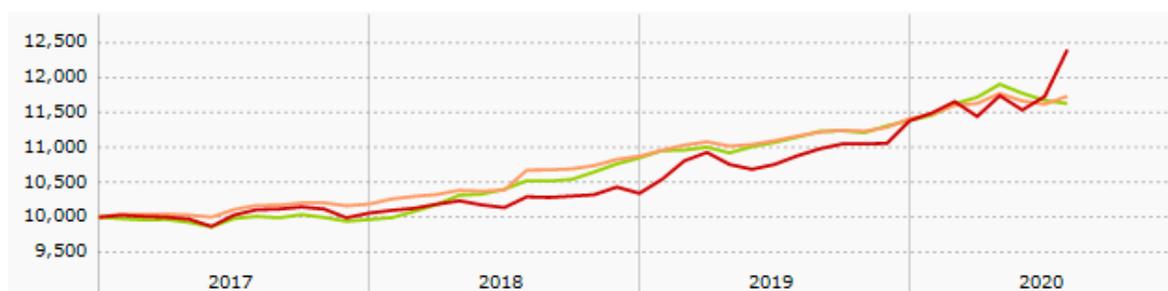
Questo indice di Sharpe viene anche chiamato indice di Sharpe *ex ante* dato che si utilizza prima di prendere una decisione. Successivamente esamineremo la sua variante *ex post* mettendone in evidenza il diverso utilizzo.

Come esempio di utilizzo dell'indice di Sharpe *ex ante* compariamo i dati di tre fondi corrispondenti a tre livelli di rischio: un fondo azionario, un fondo tradizionale ed un fondo monetario. Il primo ha il livello di rischio maggiore mentre l'ultimo il minore¹.

Esempio 4.2. Il primo fondo è il pacchetto azionario 002621 Lombarda China ed è un fondo ad alto rischio (varianza del 23,59%). La linea rossa indica il rendimento del fondo (ha un rendimento totale dell'85,98%) mentre quella verde indica il suo rendimento medio (4,43%). Nel complesso è un fondo redditizio ma molto rischioso, usato come investimento a breve termine. Il suo indice di Sharpe è 1,58.

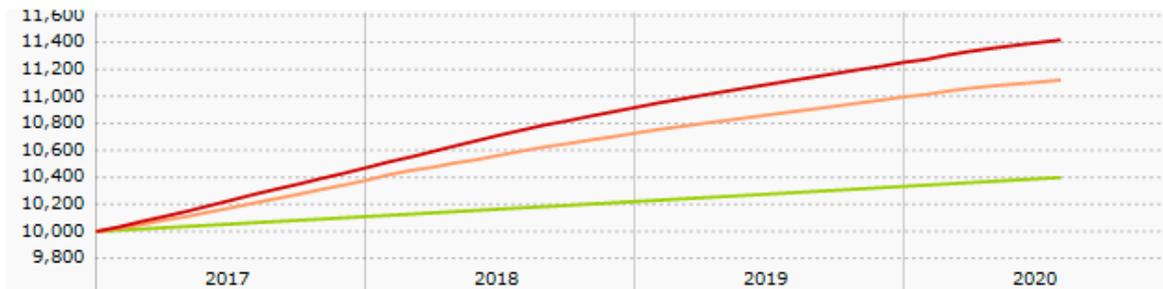


Il secondo fondo scelto è il China AMC Bond A/B 001001 ed è a rischio intermedio (varianza del 5%). Abbassandosi il rischio si abbassa anche il rendimento totale (solo del 9,84% nel 2020). Essendo molto stabile, gli investitori lo sceglieranno per un normale investimento piuttosto che uno a breve o lungo termine. Il suo indice di Sharpe è 1,17.



Il terzo ed ultimo fondo è un 000621 E Fund Cash Income ed è un fondo monetario a basso rischio (varianza dello 0,25%). Essendo quello più sicuro dove investire, è anche quello meno redditizio (solo l'1,65% nel 2020). Questo fondo va difficilmente in perdita quindi gli investitori possono investire grandi somme di denaro senza correre grandi rischi. Il suo indice di Sharpe è 9,71.

¹Analysis on CAPM and Sharpe Ratio in Market Investment, Zhou Yang, Atlantis Press, Advances in Economics, Business and Management Research, volume 166, Proceedings of the 6th International Conference on Financial Innovation and Economic Development (ICFIED), 2021.



Un alto indice di Sharpe di un investimento corrisponderà ad un investimento conveniente poichè, a parità di rischio, l'investimento con l'indice di Sharpe maggiore sarà quello più vantaggioso (Sharpe, 1994 [25]). Possiamo quindi concludere che il miglior fondo su cui investire sarà l'ultimo dei tre.

Nonostante la sua intuitività, l'indice di Sharpe nasconde alcune limitazioni. Innanzitutto la varianza (e quindi la deviazione standard) come misura di rischio non è accurata: se questa infatti è molto grande, l'indice può dare un'indicazione fuorviante su quale sia il miglior investimento. Come seconda limitazione esso risulta poco affidabile quando si vuole calcolare la convenienza di più investimenti o portafogli correlati tra loro. Risulta inaffidabile anche nel caso in cui il suo valore sia negativo: in questa situazione l'indicazione dell'indice risulta ancora una volta ingannevole, preferendo magari un fondo più rischioso ad uno meno rischioso a parità di rendimento dei due².

Per ovviare al problema della correlazione si può ragionare con un approccio più generale (Dowd, 2000 [6]). Supponiamo di avere un portafoglio di investimenti e di volerne compiere uno nuovo. Possiamo allora introdurre un nuovo indice di Sharpe relativo al portafoglio che include il nuovo investimento e confrontare i due indici: se rinominiamo quello già calcolato come $SR^{(v)}$ e quello nuovo come $SR^{(n)}$ allora il nuovo investimento sarà conveniente se e solo se varrà:

$$SR^{(n)} \geq SR^{(v)}.$$

Introducendo un nuovo indice di Sharpe scorreliamo il nuovo portafoglio dal vecchio non rischiando di avere problemi di affidabilità. In particolare, supponendo che il tasso d'interesse privo di rischio sia nullo, se il rendimento del nuovo investimento A (R_A) e quello del portafoglio senza questo ($R_p^{(v)}$) sono in proporzione di a su $1 - a$ allora il rendimento del nuovo portafoglio sarà:

$$R_p^{(n)} = aR_A + (1 - a)R_p^{(v)}. \quad (4.1)$$

²Vedere esempio *Interpreting the Sharpe ratio when excess returns are negative*, W. McLeod, G. Van Vuuren, Investment Analysts Journal, No. 59 (2004) a pag. 17-18-19.

Dalla supposizione precedente segue che le deviazioni standard del vecchio e del nuovo portafoglio sono uguali. Grazie a questo, possiamo riscrivere la disuguaglianza tra gli indici di Sharpe come

$$\frac{D^{(n)}}{\sigma_{R_p^{(n)}}} \geq \frac{D^{(v)}}{\sigma_{R_p^{(v)}}},$$

con $D^{(n)}$ e $D^{(v)}$ rispettivamente il differenziale nuovo e vecchio e $\sigma_{R_p^{(n)}}$ e $\sigma_{R_p^{(v)}}$ le relative deviazioni standard. Usando la (4.1) in combinazione con la disuguaglianza precedente possiamo riscrivere la relazione tra i rendimenti come

$$R_A \geq R_p^{(v)} + \left[\frac{\sigma_{R_p^{(n)}}}{\sigma_{R_p^{(v)}}} - 1 \right] \frac{R_p^{(v)}}{a}. \quad (4.2)$$

In questo modo l'investimento A risulta conveniente se e solo se il suo rendimento soddisfa la precedente disuguaglianza. Esso dovrà quindi soddisfare una diversa disuguaglianza a seconda della relazione che intercorrerà tra $\sigma_{R_p^{(n)}}$ e $\sigma_{R_p^{(v)}}$.

Un altro modo per scrivere la (4.2) sfrutta la seguente uguaglianza

$$\frac{\sigma_{R_p^{(n)}}}{\sigma_{R_p^{(v)}}} = \frac{\text{VaR}^{(n)}}{\text{VaR}^{(v)}},$$

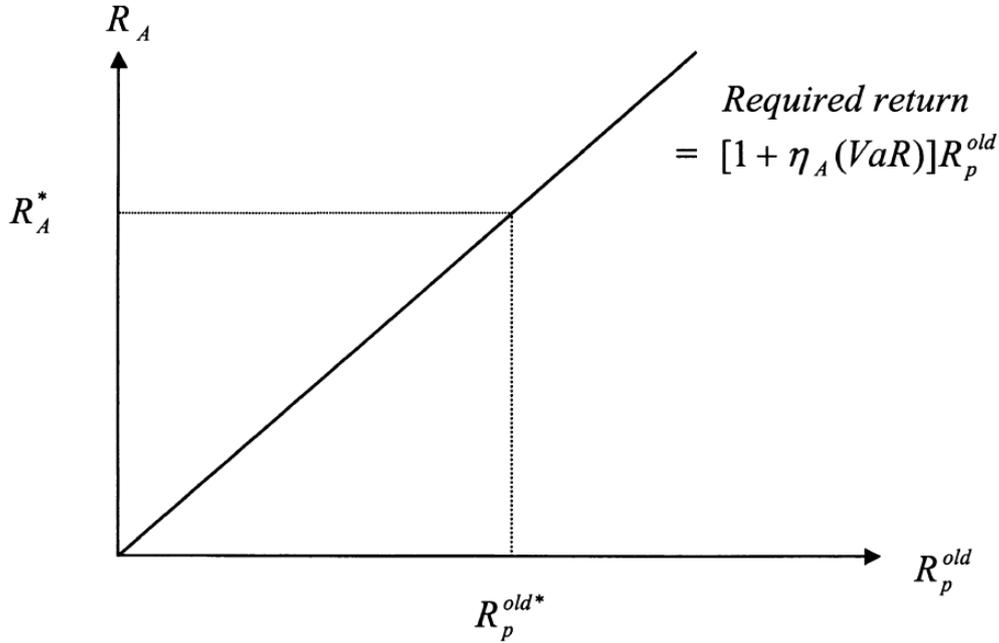
da cui si ottiene

$$R_A \geq R_p^{(v)} + \left[\frac{\text{VaR}^{(n)}}{\text{VaR}^{(v)}} - 1 \right] \frac{R_p^{(v)}}{a}.$$

Grazie alla precedente disuguaglianza il VaR ci fornisce un'altra indicazione sulla convenienza dell'investimento, infatti

$$R_A \geq R_p^{(v)} + \left[\frac{\Delta \text{VaR}}{\text{VaR}^{(v)}} \right] \frac{R_p^{(v)}}{a} = [1 + \eta_A(\text{VaR})] R_p^{(v)}, \quad \text{dove} \quad \Delta \text{VaR} = \text{VaR}^{(n)} - \text{VaR}^{(v)}$$

$\eta_A(\text{VaR})$ è chiamata *elasticità del VaR* ed è l'incremento percentuale del VaR dato dall'acquisizione della nuova posizione grazie al nuovo investimento A diviso la nuova posizione. La relazione tra R_A , $R_p^{(v)}$ e $\eta_A(\text{VaR})$ è data dal seguente grafico.



Se il rendimento atteso del vecchio portafoglio è $R_p^{(v)*}$ allora il rendimento atteso richiesto dall'investimento A affinché il suo acquisto non crei perdite sarà R_A^* . Ogni investimento con rendimento superiore o uguale a questo sarà conveniente, ogni investimento con un rendimento inferiore a questo non lo sarà.

L'elasticità del VaR ci permette di capire quanto convenga il nuovo investimento: se essa sarà positiva allora il nuovo investimento aggiungerà ulteriore rischio a quello esistente ed il rendimento richiesto dall'investimento A sarà maggiore di quello del vecchio portafoglio; se essa sarà negativa allora il nuovo investimento ridurrà il rischio complessivo ed il rendimento richiesto di A sarà minore del rendimento del vecchio portafoglio. Nel caso in cui l'elasticità del VaR sia minore di -1 allora il rendimento richiesto sarà negativo.

Mentre le discussioni teoriche si concentrano più sulla versione *ex ante*, l'indice *ex post* viene utilizzato principalmente nelle implementazioni pratiche (Sharpe, 1994; McLeod, Von Vuuren, 2004 [25][13]).

Definizione 4.3 (Indice di Sharpe ex post). Sia $R_p^{(t)}$ il rendimento di un portafoglio in un periodo di riferimento t e $R_f^{(t)}$ il tasso di interesse privo di rischio, allora definiamo la differenza come

$$D^{(t)} = R_p^{(t)} - R_f^{(t)},$$

e la sua deviazione standard come

$$\sigma_{D^{(t)}} = \sqrt{\frac{\sum_{t=1}^T (D^{(t)} - \mathbb{E}[D^{(t)}])^2}{T-1}},$$

con T un periodo di tempo fissato. Da cui l'*indice di Sharpe ex post* è definito come

$$SR_h = \frac{\mathbb{E}[D^{(t)}]}{\sigma_{D^{(t)}}}.$$

In generale sia l'indice *ex ante* che quello *ex post* sono dipendenti dal tempo. E' pratica comune "riscalare" ad un periodo di un anno due o più indici calcolati su un periodo maggiore in modo tale da avere un lasso di tempo uguale dove confrontare gli indici. Dato un indice di Sharpe $SR^{(1)}$ calcolato per un periodo di un giorno allora lo stesso indice calcolato per un periodo fissato T si ottiene tramite la seguente formula:

$$SR^{(T)} = \sqrt{T}SR^{(1)}.$$

Mentre l'indice *ex ante* è usato per valutare investimenti prima di compierli, quello *ex post* si usa per valutare la performance dopo l'evento.

Altri due indici che forniscono più indirettamente un'indicazione sul rischio sono gli indici alpha e beta.

4.2 Indici alpha e beta

Un altro indice utile per valutare la performance di un investimento è l'indice alpha. Esso confronta l'andamento dell'investimento con un indice di riferimento adatto (chiamato benchmark). Tramite il valore di alpha si può confrontare il rendimento dell'investimento rispetto l'andamento generale del mercato rappresentato dal benchmark ed affermare se questo rende più o meno del previsto (Alsaedi, 2022 [3]). Il valore dell'indice è espresso da un numero ma rappresenta una percentuale.

Esempio 4.4. Un valore alpha di un investimento di 1 mostra che il suo rendimento è maggiore rispetto alla media generale di mercato dell'1%.

Al contrario se il valore alpha è -1 allora vorrà dire che l'investimento avrà un rendimento minore rispetto la media di mercato dell'1%.

Se alpha è zero allora l'investimento sarà in linea con l'andamento generale medio del mercato.

L'indice alpha è spesso usato in combinazione con l'indice beta ed entrambi sono usati, insieme ad altri indici, nel *Capital Asset Pricing Model* come indicatori della performance di un portafoglio di investimenti.

L'indice, o coefficiente beta, misura invece la volatilità o il rischio relativo di un investimento rispetto alla performance di un benchmark. Questo indice permette agli investitori di valutare il rendimento di un investimento in base alla loro propensione al rischio e ad un indice di riferimento fissato. L'indice dà inoltre un'indicazione riguardo la rischiosità dell'investimento rispetto all'andamento generale del mercato. Il suo valore base è 1 ed indica che il rendimento dell'investimento è in linea con il mercato (e quindi con l'indice di riferimento). Se beta è positivo allora rendimento dell'investimento ed andamento del mercato saranno entrambi in crescita, se è negativo allora uno crescerà e l'altro diminuirà. Nel caso particolare in cui sia positivo ma minore di 1 allora l'investimento sarà in crescita ma non quanto l'andamento di mercato. Se è maggiore di 1 allora l'investimento avrà una volatilità maggiore rispetto a quella del mercato.

Esempio 4.5. Se l'investimento ha un indice beta di 1,1 allora il rendimento dell'investimento sarà soggetto ad una fluttuazione del 10% maggiore rispetto a quella del benchmark.

Come già anticipato alpha e beta giocano un ruolo fondamentale nel CAPM tanto da comparire nella formula principale che regola il rendimento di un portafoglio.

$$R_p = R_f + \beta(R_m - R_f) + \alpha, \quad (4.3)$$

dove R_p è il rendimento del portafoglio, R_f è il tasso d'interesse privo di rischio ed R_m è il rendimento di mercato di un dato benchmark. Alternativamente possiamo calcolare β tramite varianza e covarianza e, considerando R_p ed R_m come due v.a., abbiamo:

$$\beta = \frac{\text{cov}(R_p, R_m)}{\text{var}(R_m)},$$

dove $\text{cov}(R_p, R_m)$ è la covarianza tra R_p ed R_m e $\text{var}(R_m)$ è la varianza di R_m . Equivalentemente β si può riscrivere come

$$\beta = \rho_{p,m} \frac{\sigma_p}{\sigma_m},$$

dove $\rho_{p,m}$ è il coefficiente di correlazione di R_p ed R_m e σ_p e σ_m rispettivamente le deviazioni standard di R_p ed R_m .

Possiamo riscrivere l'equazione (4.3) isolando α :

$$\alpha = R_p - R_f - \beta(R_m - R_f).$$

Esempio 4.6. Supponiamo che il rendimento di un fondo sia del 30%, il tasso d'interesse privo di rischio sia l'8%, il rendimento del benchmark sia del 20% e l'indice beta sia 1,1, allora α è uguale a

$$\alpha = 0,30 - 0,08 - 1,1(0,20 - 0,08) = 0,088 = 8,8\%.$$

Ciò significa che il rendimento del fondo in questo caso è maggiore del rendimento del benchmark dell'8,8%.

Nella maggior parte dei casi alpha e beta si usano insieme, alcuni investitori tuttavia preferiscono usarli separatamente in modo tale da controllare le loro posizioni relative agli investimenti e massimizzare il rendimento di questi.

Appendice

Diamo definizioni, teoremi e proprietà relative a funzioni e grandezze che ci saranno utili (Pascucci, 2020 [18]).

Valore atteso

Il valore atteso di una variabile aleatoria è il suo integrale rispetto alla misura di probabilità.

Definizione 1 (Valore atteso). In uno spazio di probabilità (Ω, \mathcal{F}, P) , il *valore atteso* di una variabile aleatoria integrabile X è definito da:

$$\mathbb{E}[X] := \int_{\Omega} X dP = \int_{\Omega} X(\omega)P(d\omega),$$

dove ω è l'esito in Ω

In generale, il calcolo di un valore atteso definito come integrale astratto (sullo spazio Ω) non è particolarmente agevole: il seguente risultato mostra che è possibile esprimere il valore atteso di una v.a. X come integrale (sullo spazio euclideo \mathbb{R}^d) rispetto alla distribuzione μ_X della v.a. stessa.

Teorema 2 (Teorema del calcolo della media). Siano

$$X : \Omega \longrightarrow \mathbb{R}^d \quad \text{e} \quad f : \mathbb{R}^d \longrightarrow \mathbb{R}^N$$

rispettivamente una v.a. su (Ω, \mathcal{F}, P) con legge μ_X e una funzione \mathcal{B}_d -misurabile ($f \in m\mathcal{B}_d$). Allora $f \circ X \in L^1(\Omega, P)$ se e solo se $f \in L^1(\mathbb{R}^d, \mu_X)$ e in tal caso vale

$$\mathbb{E}[f(X)] = \int_{\mathbb{R}^d} f d\mu_X.$$

In particolare, se $\mu_X = \sum_{k=1}^{\infty} p_k \delta_{x_k}$ è una distribuzione discreta allora

$$\mathbb{E}[f(X)] = \sum_{k=1}^{\infty} f(x_k)p_k,$$

mentre se μ_X è assolutamente continua con densità γ_X allora si ha

$$\mathbb{E}[f(X)] = \int_{\mathbb{R}^d} f(x)\gamma_X(x)dx.$$

Varianza

Definizione 3 (Varianza). Sia $X \in L^2(\Omega, P)$ una v.a. reale. Si definisce *varianza* di X il numero reale non negativo

$$\text{var}(X) := \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

La radice della varianza $\sqrt{\text{var}(X)}$ è chiamata *deviazione standard* e si indica con σ . La deviazione standard è una media della distanza di X dal proprio valore atteso.

Osservazione 4. Per linearità, per ogni $a, b \in \mathbb{R}$ si ha

$$\text{var}(aX + b) = a^2\text{var}(X).$$

Inoltre, si ha

$$\text{var}(X) = 0$$

se e solo se

$$X \stackrel{\text{q.c.}}{=} \mathbb{E}[X]$$

Calcoliamo ora media e varianza di alcune v.a. discrete e continue

Esempio 5.

(i) se $X \sim \delta_{x_0}$ con $x_0 \in \mathbb{R}^d$ allora si ha

$$\begin{aligned} \mathbb{E}[X] &= \int_{\mathbb{R}^d} y\delta_{x_0}(dy) = x_0, \\ \text{var}(X) &= \int_{\mathbb{R}^d} (y - x_0)^2\delta_{x_0}(dy) = 0. \end{aligned}$$

(ii) se $X \sim \text{Unif}_n$ allora ha funzione di distribuzione $\gamma(k) = \frac{1}{n}$ per $k \in I_n$ e vale

$$\begin{aligned}\mathbb{E}[X] &= \sum_{k=1}^n k\gamma(k) = \frac{1}{n} \sum_{k=1}^n k = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}, \\ \text{var}(X) &= \mathbb{E}[X^2] - E[X]^2 \\ &= \sum_{k=1}^n k^2\gamma(k) - \left(\frac{n+1}{2}\right)^2 = \frac{1}{n} \sum_{k=1}^n k^2 - \left(\frac{n+1}{2}\right)^2 \\ &= \frac{1}{n} \cdot \frac{n(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}.\end{aligned}$$

(iii) se $X \sim \text{Unif}_{[a,b]}$ si ha

$$\begin{aligned}\mathbb{E}[X] &= \int_{\mathbb{R}} y \text{Unif}_{[a,b]}(dy) = \frac{1}{b-a} \int_a^b y dy = \frac{a+b}{2}, \\ \text{var}(X) &= \int_{\mathbb{R}} \left(y - \frac{a+b}{2}\right)^2 \text{Unif}_{[a,b]}(dy) \\ &= \frac{1}{b-a} \int_a^b \left(y - \frac{a+b}{2}\right)^2 dy = \frac{(b-a)^2}{12}.\end{aligned}$$

(iv) se $X \sim \text{Be}_p$ allora ha funzione di distribuzione γ definita da $\gamma(1) = p$, $\gamma(0) = 1-p$ e vale

$$\begin{aligned}\mathbb{E}[X] &= \sum_{k \in \{0,1\}} k\gamma(k) = 0 \cdot (1-p) + p = p, \\ \text{var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \sum_{k \in \{0,1\}} k^2\gamma(k) - p^2 = p(1-p).\end{aligned}$$

(v) se $X \sim \mathcal{N}_{\mu,\sigma^2}$ con $\sigma > 0$ allora

$$\begin{aligned}\mathbb{E}[X] &= \int_{\mathbb{R}} y \mathcal{N}_{\mu,\sigma^2}(dy) = \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} y e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy = \quad (\text{col cambio di variabile } z = \frac{y-\mu}{\sigma\sqrt{2}}) \\ &= \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} (\mu + z\sigma\sqrt{2}) e^{-z^2} dz = \frac{\mu}{\sqrt{\pi}} \int_{\mathbb{R}} e^{-z^2} dz = \mu.\end{aligned}$$

In modo analogo si vede che

$$\text{var}(X) = \int_{\mathbb{R}} (y-\mu)^2 \mathcal{N}_{\mu,\sigma^2}(dy) = \sigma^2.$$

Covarianza

Definizione 6 (Covarianza). La *covarianza* di due v.a. reali $X, Y \in L^2(\Omega, P)$ è il numero reale

$$\text{cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].$$

Esempio 7. Sia (X, Y) con densità

$$\gamma_{(X,Y)}(x, y) = ye^{-xy} \mathbb{1}_{\mathbb{R}_{\geq 0} \times [1,2]}(x, y).$$

Allora si ha

$$\mathbb{E}[X] = \iint_{\mathbb{R}^2} x \gamma_{(X,Y)}(x, y) dx dy = \log(2), \quad \mathbb{E}[Y] = \iint_{\mathbb{R}^2} y \gamma_{(X,Y)}(x, y) dx dy = \frac{3}{2},$$

e

$$\text{cov}(X, Y) = \iint_{\mathbb{R}^2} (x - \log(2)) \left(y - \frac{3}{2} \right) \gamma_{(X,Y)}(x, y) dx dy = 1 - \frac{3}{2} \log(2).$$

Usando le seguenti notazioni:

- $e_X := \mathbb{E}[X]$ per l'attesa di X ;
- $\sigma_{XY} := \text{cov}(X, Y) := e_{(X-e_X)(Y-e_Y)} = e_{XY} - e_X e_Y$ per la covarianza di X, Y ;
- $\sigma_X = \sqrt{\text{var}(X)}$ per la deviazione standard di X , dove

$$\text{var}(X) = \text{cov}(X, X) = e_{(X-e_X)^2} = e_{X^2} - (e_X)^2.$$

Osserviamo che:

(i) per ogni $c \in \mathbb{R}$ si ha

$$\text{var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] \leq \mathbb{E}[(X - c)^2]$$

e vale l'uguaglianza se e solo se $c = \mathbb{E}[X]$. Infatti

$$\begin{aligned} \mathbb{E}[(X - c)^2] &= \mathbb{E}[(X - e_X + e_X - c)^2] \\ &= \sigma_X^2 + 2 \underbrace{\mathbb{E}[X - e_X]}_{=0} (e_X - c) + (e_X - c)^2 \\ &= \sigma_X^2 + (e_X - c)^2 \geq \sigma_X^2. \end{aligned}$$

(ii) se $\sigma_X > 0$ si può sempre “normalizzare” la v.a. X ponendo

$$Z = \frac{X - e_X}{\sigma_X},$$

in modo che $\mathbb{E}[Z] = 0$ e $\text{var}(Z) = 1$.

(iii) vale

$$\text{var}(X + Y) = \text{var}(X) + \text{var}(Y) + 2\text{cov}(X, Y).$$

Se $\text{cov}(X, Y) = 0$ si dice che le v.a. X, Y sono *scorrelate*.

(iv) la covarianza $\text{cov}(\cdot, \cdot)$ è un operatore bilineare e simmetrico su $L^2(\Omega, P) \times L^2(\Omega, P)$, ossia per ogni $X, Y, Z \in L^2(\Omega, P)$ e $\alpha, \beta \in \mathbb{R}$ vale

$$\text{cov}(X, Y) = \text{cov}(Y, X) \quad \text{e} \quad \text{cov}(\alpha X + \beta Y, Z) = \alpha \text{cov}(X, Z) + \beta \text{cov}(Y, Z).$$

(v) per la disuguaglianza di Cauchy-Schwarz si ha

$$|\text{cov}(X, Y)| \leq \sqrt{\text{var}(X)\text{var}(Y)}$$

ossia

$$|\sigma_{XY}| \leq \sigma_X \sigma_Y \tag{4}$$

e si ha l'uguaglianza nella (4) se e solo se Y è funzione lineare di X nel senso che $Y \stackrel{\text{q.c.}}{=} \bar{a}X + \bar{b}$: nel caso in cui $\sigma_X > 0$, le costanti \bar{a} e \bar{b} sono date da

$$\bar{a} = \frac{\sigma_{XY}}{\sigma_X^2}, \quad \bar{b} = e_Y - e_X \frac{\sigma_{XY}}{\sigma_X^2}.$$

Correlazione

Definizione 8 (Correlazione). Siano $X, Y \in L^2(\Omega, P)$ tali che $\sigma_X, \sigma_Y > 0$. Il *coefficiente di correlazione* di X ed Y è definito da

$$\rho_{XY} := \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

Dalla (4) segue che $\rho_{XY} \in [-1, 1]$ e $|\rho_{XY}| = 1$ se e solo se $Y \stackrel{q.c.}{=} \bar{a}X + \bar{b}$: si noti che $\bar{a} > 0$ se $\rho_{XY} = 1$ e $\bar{a} < 0$ se $\rho_{XY} = -1$. Dunque il coefficiente di correlazione misura il grado di dipendenza lineare fra X ed Y . Sia ora $X = (X_1, \dots, X_d) \in L^2(\Omega, P)$ una v.a. a valori in \mathbb{R}^d . La *matrice di covarianza di X* è la matrice $d \times d$ simmetrica

$$\text{cov}(X) = (\sigma_{X_i X_j})_{i,j=1,\dots,d} = \mathbb{E} \left[\underbrace{(X - \mathbb{E}[X])}_{d \times 1} \underbrace{(X - \mathbb{E}[X])^T}_{1 \times d} \right],$$

dove M^T indica la trasposta della matrice M . Poichè

$$\langle \text{cov}(X)y, y \rangle = \mathbb{E} [|(X - \mathbb{E}[X])^T y|^2] \geq 0, \quad \text{con } y \in \mathbb{R},$$

la matrice di covarianza è semi-definita positiva. Si noti che gli elementi della diagonale sono le varianze $\sigma_{X_i}^2$ per $i = 1, \dots, d$. Se $\sigma_{X_i} > 0$ per ogni $i = 1, \dots, d$, definiamo la *matrice di correlazione* in modo analogo:

$$\rho(X) = (\rho_{X_i X_j})_{i,j=1,\dots,d}.$$

La matrice $\rho(X)$ è simmetrica, semi-definita positiva e gli elementi della diagonale sono uguali ad uno: per esempio nel caso $d = 2$, posto $\rho = \rho_{X_1 X_2}$ si ha

$$\rho(X) = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \quad \text{cov}(X) = \begin{pmatrix} \sigma_{X_1}^2 & \rho \sigma_{X_1} \sigma_{X_2} \\ \rho \sigma_{X_1} \sigma_{X_2} & \sigma_{X_2}^2 \end{pmatrix}.$$

Infine se A è una matrice costante $N \times d$ e $b \in \mathbb{R}^N$, allora la v.a. $Z := AX + b$ a valori in \mathbb{R}^N ha media

$$\mathbb{E}[Z] = A\mathbb{E}[X] + b,$$

e matrice di covarianza

$$\text{cov}(Z) = \mathbb{E} [(AX + b - \mathbb{E}[AX + b])(AX + b - \mathbb{E}[AX + b])^T] = A\text{cov}(X)A^T.$$

Asimmetria e curtosi

Mentre una distribuzione normale $\mathcal{N}_{\mu, \sigma^2}$ associata ad una v.a. X è caratterizzata da un grafico a forma campanulare, l'asimmetria è tipica di distribuzioni non normali. La curtosi invece riguarda sia distribuzioni simmetriche che distribuzioni asimmetriche e indica l'allontanamento di una distribuzione μ_X associata ad una v.a. X dalla distribuzione normale standardizzata. In particolare μ_X può avere un grafico più o meno "appuntito" rispetto a quello della normale di Gauss.

Definizione 9 (Momento di grado n). Data una v.a. X con la sua funzione di ripartizione (CDF, o funzione di distribuzione) associata $F_X(x)$ ed un valore di riferimento m , si definisce *momento di grado n* il seguente valore:

$$\mu_n = M_{n,m}(F_X(x)) = \mathbb{E}[(m - X)^n] = \int_{-\infty}^{\infty} (m - x)^n dF_X(x).$$

In particolare quando $m = \mathbb{E}[X]$ si parlerà di *ennesimo momento centrale* di X .

Esempio 10. Valore atteso e varianza sono rispettivamente momenti di grado uno e due.

Definizione 11 (Asimmetria). Data una v.a. X con $F_X(x)$ sua CDF associata e μ il suo valore atteso, *l'asimmetria* è il terzo momento centrale standardizzato definito come

$$\mu_3 = \mathbb{E} \left[\left(\frac{\mu - X}{\sigma} \right)^3 \right] = \int_{-\infty}^{\infty} (\mu - x)^3 dF_X(x).$$

Si parla di asimmetria positiva quando il grafico della distribuzione ha una coda allungata verso destra, mentre si parlerà di asimmetria negativa quando il grafico della distribuzione ha una coda allungata verso sinistra.

Definizione 12 (Curtosi). Data una v.a. X con $F_X(x)$ sua CDF associata e μ il suo valore atteso, la *curtosi* è il quarto momento centrale standardizzato, definito come

$$\mu_4 = \mathbb{E} \left[\left(\frac{\mu - X}{\sigma} \right)^4 \right] = \int_{-\infty}^{\infty} (\mu - x)^4 dF_X(x).$$

Il valore di curtosi di una distribuzione normale è 3 (distribuzione *normocurtica*). Se per una distribuzione generica μ_X esso è minore di 3 si parlerà di distribuzione *platicurtica*, se invece è maggiore di 3 la distribuzione sarà chiamata *leptocurtica*.

Quantili

Di particolare importanza per la teoria della dominanza stocastica e l'introduzione alle principali misure di rischio al ribasso sono i quantili. Essi vengono definiti a partire da una funzione di ripartizione di una v.a. X (Acerbi, Tasche, 2002 [2]).

Definizione 13 (Quantile). Sia X una v.a. ed $F_X(x)$ la sua CDF, dato $\alpha \in (0, 1)$ un valore di riferimento fissato, vengono definiti

$$\begin{aligned} q_\alpha(X) &= \inf\{x \in \mathbb{R} : F_X(x) \geq \alpha\} && \text{come } l'\alpha\text{-quantile inferiore di } X, \\ q^\alpha(X) &= \inf\{x \in \mathbb{R} : F_X(x) > \alpha\} && \text{come } l'\alpha\text{-quantile superiore di } X, \end{aligned}$$

dove α è detto ordine del quantile.

Notiamo che $q^\alpha(X) = \sup\{x \in \mathbb{R} : P(X \leq x) \leq \alpha\}$. Dall'inclusione

$$\{x \in \mathbb{R} : P(X \leq x) > \alpha\} \subset \{x \in \mathbb{R} : P(X \leq x) \geq \alpha\}$$

si evince che $q_\alpha(X) \leq q^\alpha(X)$. Inoltre è facile vedere che vale l'uguaglianza se e solo se $P(X \leq x) = \alpha$ per al più un x .

Osservazione 14. Dato un valore x allora condizioni equivalenti all'essere un α -quantile sono le seguenti:

- (i) $P(X \leq x) \geq \alpha$ e $P(X \geq x) \geq 1 - \alpha$;
- (ii) $F_X(x) \geq \alpha$ e $F_X(x-) \leq \alpha$ dove $F_X(x-) = \lim_{x \rightarrow q} F_X(x)$;

La prima condizione dell'osservazione si traduce informalmente nel "dividere" la distribuzione in un numero di parti uguali che dipende da α . Forniamo alcuni esempi per chiarire le idee.

Esempio 15. Il quantile di ordine $\alpha = 0,25$ (o venticinquesimo percentile o primo quartile) è un valore tale per cui il 25% dei valori della distribuzione siano più piccoli di questo (e conseguentemente il 75% siano più grandi).

Esempio 16. La mediana altro non è che il quantile di ordine $\alpha = 0,50$ (o cinquantesimo percentile o secondo quartile).

Elenco delle figure

Il seguente elenco riporta le fonti da cui sono state prese le figure. L'ordine è quello di apparizione.

- [1] A. Jobayed: *Evaluating the Predictive Performance of Value-at-Risk (VaR) Models on Nordic Market Indices*, Department of Finance and Statistics, Hanken School of Economics, Vaasa, pag. 13 (2017)
- [2] A. Damodaran: *Strategic Risk Taking: A Framework for Risk Management*, Pearson Prentice Hall, pag. 12 (2007)
- [3] A. Damodaran: *Strategic Risk Taking: A Framework for Risk Management*, Pearson Prentice Hall, pag. 15 (2007)
- [4] D.M. Neto: *A comparison of Value at Risk methods in portfolios with linear and non-linear financial instruments*, University of East London, pag. 32 (2016)
- [5] S. Sarykalin, G. Serraino, S. Uryasev: *Value-at-Risk vs. Conditional Value-at-Risk in Risk Management and Optimization*, Tutorials in Operations Research, INFORMS, pag. 274 (2008)
- [6] S. Sarykalin, G. Serraino, S. Uryasev: *Value-at-Risk vs. Conditional Value-at-Risk in Risk Management and Optimization*, Tutorials in Operations Research, INFORMS, pag. 275 (2008)
- [7] O. Shirokikh, A. Sorokin, V. Boginski: *A Note on Transmission Switching in Electric Grids with Uncertain Line Failures*, Energy Systems, pag. 4 (2013)
- [8] Y. Yamai, T. Yoshiba: *Comparative Analyses of Expected Shortfall and Value-at-Risk: Their Estimation Error, Decomposition, and Optimization*, Tokyo Metropolitan University, Research Papers in Economics, pag. 97 (2002)
- [9] Y. Yamai, T. Yoshiba: *Comparative Analyses of Expected Shortfall and Value-at-Risk: Their Estimation Error, Decomposition, and Optimization*, Tokyo Metropolitan University, Research Papers in Economics, pag. 99 (2002)

- [10] Z. Yang: *Analysis on CAPM and Sharpe Ratio in Market Investment*, Advances in Economics, Business and Management Research, volume 166, Proceedings of the 6th International Conference on Financial Innovation and Economic Development, pag. 5 (2021)
- [11] Z. Yang: *Analysis on CAPM and Sharpe Ratio in Market Investment*, Advances in Economics, Business and Management Research, volume 166, Proceedings of the 6th International Conference on Financial Innovation and Economic Development, pag. 6 (2021)
- [12] Z. Yang: *Analysis on CAPM and Sharpe Ratio in Market Investment*, Advances in Economics, Business and Management Research, volume 166, Proceedings of the 6th International Conference on Financial Innovation and Economic Development, pag. 6 (2021)
- [13] K. Dowd: *Adjusting for risk: An improved Sharpe ratio*, International Review of Economics and Finance 9, pag. 216 (2000)

Bibliografia

- [1] C. Acerbi: *Spectral Measures of Risk: A Coherent Representation of Subjective Risk Aversion*, Journal of Banking & Finance 26(7), 1505–1518 (2002)
- [2] C. Acerbi, D. Tasche: *On the Coherence of Expected Shortfall*, Journal of Banking & Finance 26(7), 1487–1503 (2002)
- [3] R. Alsaedi: *Alpha and Beta*, Global Scientific Journal, Volume 10, Issue 1, 581–602 (2022)
- [4] P. Artzner, F. Delbaen, J. Eber, D. Heath: *Coherent Measures Of Risk*, Mathematical Finance 9(3), 203–228 (1998)
- [5] A. Damodaran: *Strategic Risk Taking: A Framework for Risk Management*, Wharton School Publishing, First Edition (2007)
- [6] K. Dowd: *Adjusting for risk: An improved Sharpe ratio*, International Review of Economics and Finance 9, 209–222 (2000)
- [7] P.C. Fishburn: *Utility Theory For Decision Making*, Research Analysis Corporation, First Edition (1970)
- [8] M. Frittelli, E. Rosazza Gianin: *Law Invariant Convex Risk Measures*, Advances in Mathematical Economics, Vol. 7, 42–53 (2005)
- [9] C. Hahn, A. Pfingsten, P. Wagner: *An Empirical Investigation of the Rank Correlation Between Different Risk Measures*, SSRN Electronic Journal 6(4) (2002)
- [10] M. Leippold: *Value-at-Risk and Other Risk Measures*, SSRN Electronic Journal, Department of Banking and Finance, University of Zurich and the Swiss Finance Institute (2015)
- [11] C. Ma, W. Wong: *Stochastic Dominance and Risk Measure: A Decision-Theoretic Foundation for VaR and C-VaR*, European Journal of Operational Research, Volume 207, Issue 2, 927–935 (2010)

- [12] A. Mas-Colell, M.D. Whinston, J.R. Green: *Microeconomic theory*, Oxford University Press, First Edition (1995)
- [13] W. McLeod, G. Van Vuuren: *Interpreting the Sharpe ratio when excess returns are negative*, Investment Analysts Journal 59, 15–20 (2004)
- [14] D.N. Nawrocki: *A Brief History of Downside Risk Measures*, The Journal of Investing 8(3) (1999).
- [15] D.M. Neto: *A comparison of Value at Risk methods in portfolios with linear and non-linear financial instruments*, University of East London (2016)
- [16] W. Ogryczak, M. Opolska–Rutkowska: *On Mean–Risk Models Consistent with Stochastic Dominance*, Report of the Institute of Control & Computation Engineering, Warsaw University of Technology, Report nr: 04–08 (2004)
- [17] W. Ogryczak, A. Ruszczyński: *From Stochastic Dominance to Mean–Risk Models: Semideviations as Risk Measures*, European Journal of Operational Research 116(1), 33–50 (1999)
- [18] A. Pascucci *Teoria della probabilità: Variabili aleatorie e distribuzioni*, Springer (2020)
- [19] R. Pelessoni: *Misure di Rischio e Coerenza*, Atti della Scuola Estiva su Finanza Quantitativa, Auronzo di Cadore, Dipartimento di Matematica Applicata Università Ca’ Foscari, Venezia, 423–433 (2002)
- [20] G.C. Pflug: *Some Remarks On The Value–At–Risk And The Conditional Value–At–Risk*, Department of Statistics and Decision Support Systems, University of Vienna (2000)
- [21] S. Rockafellar, S. Uryasev: *Conditional Value–At–Risk for General Loss Distributions*, Journal of Banking & Finance 26(7), 1443–1471 (2001)
- [22] R. S Bendigeri: *Variance–Covariance (Delta normal) Approach of VaR Models: An Example From Bombay Stock Exchange*, Aarhat Multidisciplinary International Education Research Journal (AMIERJ), Volume VIII, Issue II (2019)
- [23] S. Sarykalin, G. Serraino, S. Uryasev: *Value–at–Risk vs. Conditional Value–at–Risk in Risk Management and Optimization*, Tutorials in Operations Research, INFORMS, 270–294 (2008)
- [24] W.F. Sharpe: *Capital Asset Prices: A Theory Of Market Equilibrium Under Conditions of Risk*, The Journal of Finance, Volume XIX, No. 3, 425–442 (1964)

- [25] W.F. Sharpe: *The Sharpe Ratio*, The Journal of Portfolio Management, 49–58 (1994)
- [26] S.M. Sunoy, S.S. Maya: *The role of lower partial moments in stochastic modeling*, METRON–International Journal of Statistics, vol. LXVI, n. 2, 223–242 (2008)
- [27] M. Unser: *Lower Partial Moments as Measures of Perceived Risk–An Experimental Study*, Bank–und Finanzwirtschaftliche Forschung (BAFIFO) Lehrstuhl für Betriebswirtschaftslehre insbesondere Finanzwirtschaft, Universität Bamberg, Diskussionsbeiträge Nr. 09 (1998)
- [28] J. Von Neumann, O. Morgenstern: *Theory of Games and Economic Behavior*, Princeton University Press, Third Edition (1953)
- [29] R. Witt, H. Waibel: *Lower partial moments as a measure of vulnerability to poverty in Cameroon*, Institute of Development and Agricultural Economics, Faculty of Economics and Management, Leibniz Universität Hannover, Discussion Paper No. 434 (2009)
- [30] E.G. Wolfstetter: *Stochastic Dominance: Theory and Applications*, Humboldt–Universität zu Berlin, Institut f. Wirtschaftstheorie I, Wirtschaftswissenschaftliche Fakultät, Berlin (1970)
- [31] Y. Yamai, T. Yoshihara: *Comparative Analyses of Expected Shortfall and Value-at-Risk: Their Estimation Error, Decomposition, and Optimization*, Monetary and Economic studies, 87–122 (2002)
- [32] Z. Yang: *Analysis on CAPM and Sharpe Ratio in Market Investment*, Atlantis Press, Advances in Economics, Business and Management Research, volume 166, Proceedings of the 6th International Conference on Financial Innovation and Economic Development (ICFIED), 5–8 (2021)
- [33] L. Zhiqiang, W. Jianming, Z. Junyu, L. Haitao: *The Statistical Meaning of Kurtosis and Its New Application to Identification of Persons Based on Seismic Signals*, Sensors 8(8), 5106–5119 (2008)

Ringraziamenti

Ringrazio il professor Pascucci per avermi aiutato nella stesura della tesi ed avermi trasmesso la passione per la probabilità e la statistica.

Ringrazio i miei genitori ed i miei parenti per il continuo sostegno anche nei momenti di difficoltà e smarrimento.

Ringrazio in particolare zia Anna, zio Ernesto ed Alessandro per avermi fatto vivere la magia di Bologna ancora per qualche anno.

Ringrazio Alessio, Samuele ed generale tutta la compagnia di amici di Modugno per il loro supporto e per i consigli.

Ringrazio Alberto, Carlo ed Enrico per i momenti di leggerezza e le risate.

Ringrazio Stefano, Giovanni, Fawzia, Arianna, Antonio, Vittorio, Marco e tutte le altre persone con cui ho stretto amicizia durante l'uni: nonostante non abbia potuto approfondire i rapporti con loro, hanno tutti lasciato inevitabilmente un ricordo positivo nel mio cuore.

Ringrazio tutti i coinquilini che ho conosciuto durante la convivenza: alcuni di loro mi hanno temprato psicologicamente ma con altri ho vissuto momenti indimenticabili. Un sentito grazie va ad Arianna per la sua positività, la sua energia e l'aver contribuito a creare un clima familiare in casa, ad Alessia per la sua pura ed infinita dolcezza ma soprattutto a Francesca per avermi aiutato a diventare una versione migliore di me stesso.

Ringrazio tutte le persone che hanno portato e che continuano a portare parte del loro affetto nella mia vita. Senza il loro sostegno non sarei qui.

Ringrazio infine me stesso per non essermi mai arreso nonostante i momenti di sconforto ed aver avuto una forte determinazione.