UNIVERSITÁ DEGLI STUDI DI PADOVA

FACOLTÁ DI SCIENZE STATISTICHE

**CORSO DI LAUREA SPECIALISTICA
IN SCIENZE STATISTICHE, ECONOMICHE,
FINANZIARIE E AZIENDALI**

TESI DI LAUREA SPECIALISTICA

La selezione del portafoglio tra robustezza e ambiguitá: il modello Media-Varianza Monotono.

Relatore: Ch.mo Prof. MARCO FERRANTE

Laureanda: GENNY FRANCESCHETTO

ANNO ACCADEMICO 2005-2006

"Non è vero che abbiamo poco tempo:
la verità è che ne perdiamo molto."

Seneca

Indice

Introduzione

Il nostro obiettivo in questa relazione è dare una soluzione al problema che affligge ogni investitore: dati alcuni titoli rischiosi e non, qual'è la combinazione ottimale che permette di ottenere maggiori guadagni?
Se si potesse conoscere con certezza il valore futuro di ogni titolo rischioso, basterebbe investire sul titolo che rende di più, come ci insegna la teoria dell'utilità attesa ([5],[9], [1]). Secondo questa teoria, infatti, ogni individuo conosce le alternative che ha a disposizione, le valuta, e sceglie quella che gli dà utilità maggiore.

Purtroppo, poichè non è possibile sapere con esattezza cosà accadrà nel futuro, dobbiamo limitarci a prevedere l'andamento dei titoli. Ma una previsione, per definizione, può non essere esatta. E' importante valutare quanto il valore da noi previsto si scosta dal vero valore assunto dal titolo, e, ancora più importante, è valutare la possibilità che il nostro titolo possa subire delle perdite.
Se un titolo ha un andamento molto variabile, può portare alti guadagni così come potrebbe portare grosse perdite. Investire o no dipende da quanto siamo disposti a rischiare, ovvero dall'avversione al rischio: più siamo propensi al rischio e più siamo disposti ad accettare la possibilità di subire perdite sapendo, però, che c'è anche la possibilità che ci siano grossi guadagni.
Ma come valutare il rischio? Un primo modo è usare la varianza come misura del rischio applicando il modello Media-Varianza ([3], [5], [12], [13]).

Con il modello Media-Varianza, Harry Markowitz ([16]) ottenne nel 1990 il Premio Nobel per aver trovato un sistema per risolvere il trade-off tra alti profitti e bassi rischi. Il modello fu studiato da Markowitz già negli anni '50: col suo primo lavoro aveva trovato il portafoglio efficiente in media-varianza focalizzandosi in particolare sul singolo periodo, apportando così un contributo significativo alla gestione del *risk management* grazie all'introduzione di una metodologia scientifica e quantitativa per lo studio e l'analisi del rischio finanziario.

Sull'argomento introdotto da Markowitz sono stati effettuati molti approfondimenti in studi successivi fino ad arrivare a considerare diverse combinazioni di titoli rischiosi e non rischiosi e diversi periodi. L'estensione al multiperiodo ha una formulazione leggermente diversa rispetto al singolo periodo. In particolare, nel multiperiodo, invece che trattare media ($\mathbb{E}[w]$) e varianza ($Var[w]$) del portafoglio come quantità separate per cercare poi una relazione tra esse, si considera direttamente l'utilità attesa della ricchezza finale ($\mathbb{E}[U(w)]$).
L'utilizzo di una funzione quadratica quale la varianza, ha però delle limitazioni. Per esempio ha un punto di massimo oltre il quale l'aumento della ricchezza provoca una diminuzione dell'utilità; un altro svantaggio stà nel fatto che con una funzione quadratica l'avversione al rischio cresce al crescere della ricchezza cosicchè i titoli rischiosi dovrebbero essere considerati beni inferiori.
Per questi motivi si sono cercate varie alternative al modello media-varianza, tra cui i modelli media-varianza pesata, media-semivarianza e media-downside risk ([8], [11]) e i modelli di avversione all'ambiguità ([10], [7], [14]) Questi ultimi prendono in considerazione la possibilità che il modello usato, che ipotizziamo essere P, non sia correttamente specificato, ovvero esista un altro modello Q, cioè un'altra legge che descrive i dati.

In questa relazione ci soffermeremo su un particolare modello di avversione all'ambiguità: il modello Media-Varianza Monotono ([15]). Verrà proposto un modello di selezione del portafoglio basato su una classe di preferenze che coincidono con le preferenze media-varianza nel loro dominio di monotonicità e differiscono dove le prefenze media-varianza non sono monotone. E' possibile, infatti, identificare una restrizione nel dominio al fine di renderle monotone. Per effettuare tale restrizione, consideriamo il funzionale concavo $J : B_o(\Sigma) \rightarrow \mathbb{R}$ dato da $J(\psi) = \int \psi dq - (\theta/2)Var(\psi)$ e prendiamo l'insieme $\mathcal{M}$ nel quale il differenziale di Gateaux di J è positivo. L'insieme in cui le preferenze media-varianza non violano l'assioma di monotonicità, è l'insieme chiuso e convesso $\mathcal{G}_\theta = \{f \in F : u(f) \in \mathcal{M}\}$ che chiameremo *dominio di monotonicità*.

Nel primo capitolo spiegheremo la teoria dell'utilità attesa cercando di capire perchè non risulta appropriata per risolvere problemi di selezione del portafoglio, passando attraverso il paradosso di Ellsberg e gli indici di avversione al rischio. Nel secondo capitolo verrà introdotto il modello Media-Varianza uniperiodale di Markowitz del quale esamineremo il caso con soli titoli rischiosi e il caso con titoli sia rischiosi sia privi di rischio. Andremo poi ad esplorare i modelli che apportano delle varianti al modello Media-Varianza classico parlando dei modelli media-varianza pesata, media semivarianza, media downside risk e i modelli di avversione all'ambiguità nel capitolo terzo. Nel quarto capitolo, invece, introdurremo il modello Media-Varianza Monotono valutando le principali differenze col modello non monotono e sviluppando alcuni esempi. Nell'ultimo capitolo andremo ad applicare il modello monotono a dati reali e ne analizzeremo i risultati traendone opportune conclusioni.

Parte I

Dall'Utilità Attesa al Modello Media-Varianza

Capitolo 1

Utilità Attesa

1.1 Teoria dell'Utilità Attesa

Per definire un sistema di selezione del portafoglio è importante stabilire una relazione d'ordine che ci consenta di ordinare i titoli e le loro possibili combinazioni. Questa relazione d'ordine definisce la struttura delle preferenze del consumatore. L'ipotesi di base è che il consumatore abbia un comportamento razionale, ovvero:

- il consumatore conosce tutte le alterne che ha a disposizione e alloca il proprio reddito su un paniere di n beni senza alcuna incertezza;

- il consumatore è in grado di ordinare le alternative;

- una volta individuata l'alternativa migliore, la sceglie;

- il consumatore spende tutto il reddito disponibile.

Il principio di razionalità, però, non è sufficiente per ordinare le scelte del consumatore tra alternative di investimento che implicano il rischio di perdite.
Per poter ordinare la scelta tra alternative rischiose si fa riferimento alla *teoria dell'utilità attesa*, secondo la quale la scelta tra alternative rischiose può essere rappresentata confrontando i valori attesi di una funzione, detta *funzione di utilità*.
Se a e b sono due alternative rischiose, la teoria dell'utilità attesa afferma che:

$$a \succ b \Leftrightarrow \mathbb{E}[U(a)] > \mathbb{E}[U(b)]$$

dove $\succ$ indica la preferenza e $U(\cdot)$ è la funzione di utilità.
La principale caratteristica della funzione di utilità è che essa deve essere crescente, a meno di una trasformazione affine: se $U(\cdot)$ è un sistema di preferenze, anche

$V(\cdot) = a + b\,U(\cdot)$ è una funzione di utilità che rappresenta un sistema di preferenze. Questi modelli di rappresentazione delle preferenze sono costruiti sulla base di alcuni assiomi che devono essere soddisfatti:

1. *Assioma di completezza.* Quando vengono proposte due combinazioni, il consumatore è sempre in grado di dire quale delle due preferisce;

2. *Assioma di transitività.* Se l'alternativa a è preferibile all'alternativa b, e l'alternativa b è preferibile alla c, allora a è preferibile a c, ovvero
$$a \succ b, b \succ c \Rightarrow a \succ c;$$

3. *Assioma di non sazietà.* Il consumatore non è mai sazio, nel senso che una quantità maggiore sarà sempre preferibile rispetto a una quantità minore;

4. *Assioma di indipendenza.* Se l'alternativa a è preferita all'alternativa b, allora qualsiasi mistura con un terza alternativa non modifica l'ordine delle preferenze
$$a \succ b \Leftrightarrow \alpha\,a + (1-\alpha)\,c \succ \alpha\,b + (1-\alpha)\,c$$
dove se a e b sono sue alternative rischiose e $\alpha\,a + (1-\alpha)b$ è una mistura tra alternative rischiose.

La teoria economica sulle scelte del consumatore si propone di spiegare come un individuo prende decisioni razionali in una situazione in cui le risorse che ha disposizione sono limitate. Tutte le risorse sono scarse, nel senso che non sono disponibili in quantità sufficiente per soddisfare tutti i bisogni di tutti gli individui. Per risorse non intendiamo solo le risorse naturali come il petrolio e l'aqua, ma anche le risorse umane, come il lavoro, e le risorse di capitale. Un'importante conseguenza di questa condizione di scarsità è che gli individui devono scegliere in un insieme limitato di possibilità. La scelta di avere una quantità maggiore di un bene comporta necessariamente la rinuncia a una certa quantità di altri beni.
Per capire il comportamento del consumatore, quindi, dobbiamo innanzitutto capire *cosa vuole* e rappresentare i suoi gusti. Poi dobbiamo conoscere il budget che ha a disposizione per sapere *cosa può fare* e rappresentare i vincoli a cui deve sottostare a causa della sua ricchezza limitata. Infine andremo a mettere insieme le preferenze del consumatore e i limiti a cui deve sottostare per determinare quale tra le possibili alternative gli assicurerà la massima soddisfazione. Per questo ultimo passaggio ci si avvale proprio della funzione di utilità.
La teoria dell'utilità attesa è un buon sistema per tentare di stabilire un ordine di preferenze, ma, come vedremo di seguito, non sempre risulta appropriata per risolvere problemi di gestione del portafoglio.

1.2 Paradosso di Ellsberg

Preferiamo investire su titoli per i quali possediamo grandi quantità di informazioni che ci permettono di valutare meglio la probabilità di successo, oppure la quantità di informazioni che abbiamo a disposizione non influenza le nostre scelte? Uno dei maggiori limiti della teoria dell'utilità attesa è proprio legato alla quantità di informazioni di cui disponiamo sulle alternative rischiose, perchè se la quantità di informazioni influisce in qualche modo sulle preferenze, allora la teoria dell'utilità attesa ci può condurre a risultati paradossali. Tale effetto è noto come *Paradosso di Ellsberg*, dal nome dell'autore che per primo evidenziò una certa incongruenza tra i risultati di diversi esperimenti da lui effetturati nel 1961.

Proviamo a spiegare questo paradosso con un esempio. Supponiamo di dover scegliere tra investire 1.000 euro in azioni COMIT o investire 1.000 euro in azioni ENI. Il rischio è rappresentato da tre scenari (H,M,L) a ognuno dei quali è associato un possibile rendimento dei titoli. Supponiamo poi di non avere sufficienti informazioni circa le probabilità che si verifichino i tre stati di natura, ovvero sappiamo con certezza solo che lo stato H si verifica con probabilità 1/3 ma non sappiamo come distribuire la rimanente probabilità (2/3) tra lo stato M e lo stato L. Ecco i possibili scenari:

	Probabilità	ENI	COMIT
Stato H	1/3	0.6	0
Stato M	?	0	0
Stato L	?	0	0.6

Se preferiamo investire su scelte rischiose per le quali però conosciamo la probabilità di successo, allora il titolo ENI è preferibile al titolo COMIT:

$$\text{ENI} \succ \text{COMIT}$$

Vediamo ora cosa succede se disponiamo di informazioni diverse. Ipotizziamo di investire il 50% della ricchezza in azioni ENI o in azioni COMIT, e di investire il rimanente 50% in un altro titolo, il MIB30. Gli scenari che si presentano sono i seguenti:

	Probabilità	ENI+ MIB30	COMIT+MIB30
Stato H	1/3	0.3	0
Stato M	?	0.3	0.3
Stato L	?	0	0.3

In questo caso, se scegliamo l'investimento per il quale conosciamo la probabilità di successo, la nostra preferenza cadrà su COMIT+MIB30:

$$\text{ENI+MIB30} \prec \text{COMIT+MIB30}$$

Il Paradosso di Ellsberg ci permette di trarre interessanti conclusioni circa l'utilizzo della teoria dell'utilità attesa per la scelta tra alternative rischiose. Per esempio, se il soggetto fosse indifferente alla qualità delle informazioni di cui dispone, il paradosso non sussisterebbe.
La preferenza per alternative per le quali si dispone di più (o meno) informazione non può essere rappresentata utilizzando la teoria dell'utilità attesa; per questo motivo sono stati sviluppati nuovi modelli (che vedremo nei prossimi capitoli) che propongono valide alternative alla teoria dell'utilità attesa.

1.3 Rischio e Incertezza

Riportiamo di seguito un'antica distinzione, sviluppata da Knight nel 1920, tra rischio e incertezza. Con il temine *rischio* si vuole descrivere una situazione nella quale si è in grado di assegnare una probabilità a ciascuno stato di natura, mentre con *incertezza* viene indicata una situazione nella quale non può essere fatta una precisa assegnazione di probabilità.
E' importante considerare che la teoria dell'utilità attesa è efficace per gestire situazioni di rischio, ma non per gestire situazioni di incertezza.

1.4 Indici di avversione al rischio

Abbiamo visto che la teoria dell'utilità attesa non consente di ordinare la scelta tra alternative di investimento per le quali disponiamo di informazioni di qualità diversa e che non ci permette di gestire situazioni nelle quali non riusciamo ad assegnare un probabilità precisa ad ogni stato di natura. Ciò nonostante, la teoria dell'utilità attesa è ampiamente utilizzata in campo finanziario.
Nella teoria dell'utilità attesa, l'ordinamento delle scelte rischiose avviene sulla base di due funzioni:

- la *distribuzione di probabilità*, rispetto alla quale si calcola il *valore atteso*;
- la *funzione di utilità*, che deve essere monotona crescente.

In base alla forma della funzione di utilità, è possibile valutare l'attitudine del consumatore rispetto a situazioni di rischio. Più precisamente, diremo che il soggetto è *avverso al rischio* se di fronte alla scelta tra una somma certa $U(w)$, dove w rappresenta la ricchezza, e una lotteria che gli promette un valore medio dello stesso importo ($\mathbb{E}[U(w+z)]$), sceglie la somma certa:

$$U(w) < \mathbb{E}[U(w+z)]$$

con $\mathbb{E}(z) = 0$.
Possiamo dire, quindi, che il soggetto è avverso al rischio se la sua funzione di utilità è tale per cui il valore atteso della funzione stessa sia minore della funzione del valore atteso:

$$\mathbb{E}[U(w)] < U(\mathbb{E}[w])$$

che equivale ad assumere che la funzione di utilità sia concava (vedere *Teorema di Jensen*, appendice B.2), ovvero che la sua derivata seconda sia negativa ($U'' < 0$).
Si può misurare il grado di avversione al rischio tramite l'utilizzo di *indici di avversione al rischio* che misurano la concavità della funzione di utilità. L'indice più noto è l'*indice di avversione assoluta al rischio* di *Arrow-Pratt* (ARA):

$$ARA = -\frac{U''(w)}{U'(w)}$$

L'*indice di tolleranza al rischio* è invece definito come $\frac{1}{ARA}$.
Il caso più semplice di funzione concava è una *funzione quadratica* del tipo:

$$U(w) = w - \frac{\theta}{2}w^2$$

con $b > 0$.
Una funzione di utilità quadratica è di facile utilizzo, ma ha dei grossi svantaggi. Il primo è che la funzione quadratica ha un punto di massimo oltre il quale l'aumento della ricchezza provoca un diminuzione dell'utilità; in questo modo si viola un requisito importante della funzione di utilità secondo il quale il consumatore preferisce *più* a *meno*. Un altro svantaggio è il fatto che l'avversione al rischio cresce al crescere della ricchezza, cosicchè i titoli rischiosi hanno la caratteristica di *beni inferiori*.
Queste considerazioni hanno portato alla ricerca di funzioni di utilità più realistiche, come per esempio la *funzione di utilità esponenziale* (CARA, *Constant Absolute Risk Aversion*):

$$U(w) = a - \exp(-\theta w)$$

che garantisce che un'avversione al rischio indipendente dalla ricchezza. Infatti, se utilizziamo una funzione CARA, abbiamo che:

$U'(w) = -[-\theta \exp(-\theta w)] = \theta \exp(-\theta w)$
$U''(w) = \theta[-\theta \exp(-\theta w)] = -\theta^2 \exp(-\theta w)$

da cui

$$ARA = -\frac{-\theta^2 \exp(-\theta w)}{\theta \exp(-\theta w)} = \theta$$

Un altro tipo di funzione utilizzata è la *funzione di utilità a potenza* (CRRA, *Constant Relative Risk Aversion*):

$$U(w) = \frac{w^\gamma - 1}{\gamma}$$

Per questo tipo di funzione di utilità si possono considerare due casi interessanti:

i) se $\gamma = 1$, la funzione di utilità è lineare e il consumatore è *neutrale al rischio*;

ii) se γ tende a zero otteniamo, invece, una funzione di utilità logaritmica che è largamente utilizzata in modelli di tipo dinamico.

Si possono rappresentare anche funzioni di utilità che ipotizzano un andamento iperbolico dell'avversione al rischio e che utilizzano l'indice di tolleranza al rischio. Sono funzioni di tipo HARA, (*Hyperbolic Absolute Risk Aversion*):

$$U(w) = \frac{1-\gamma}{\gamma}\left[\frac{\beta w}{1-\gamma} + \eta\right]^\gamma$$

per cui l'indice di tolleranza al rischio diventa

$$\frac{1}{ARA} = \frac{w}{1-\gamma} + \frac{\eta}{\beta}$$

Le funzioni di utilità HARA sono in grado di rappresentare diversi atteggiamenti del consumatore verso il rischio e vengono usate, come le funzioni CRRA, per svolgere analisi di tipo dimanico.

1.5 Introduzione al Modello Media-Varianza

Nella gestione del portafoglio assume una particolare importanza la valutazione del trade-off tra rendimento e rischio che si può misurare attraverso l'utilizzo di un modello, detto *modello media-varianza*, che utilizza esclusivamente i primi due momenti della distribuzione di probabilità.

Per l'applicazione del modello media-varianza, assumiamo di utilizzare una funzione di utilità quadratica:

$$U(w) = w - \frac{\theta}{2}w^2$$

Utilizzando la formula di scomposizione della varianza tale che

$$Var[w] = \mathbb{E}[w^2] - \mathbb{E}[w]^2$$

possiamo ottenere il valore dell'utilità attesa:

$$\begin{aligned}
\mathbb{E}[U(w)] \quad &= \mathbb{E}[w] - \tfrac{\theta}{2}\mathbb{E}[w^2] = \\
&= \mathbb{E}[w] - \tfrac{\theta}{2}\mathbb{E}[w]^2 - \tfrac{\theta}{2}Var[w] = \\
&= U(\mathbb{E}[w]) - \tfrac{\theta}{2}Var[w]
\end{aligned}$$

Si può notore come il livello di utilità attesa è direttamente proporzionale al valore atteso della ricchezza e inversamente proporzionale alla varianza della ricchezza.
Se si utilizza una funzione di utilità quadratica, quindi, la rappresentazione delle scelte in termini di media e varianza è esattamente riconducibile alla teoria dell'utilità attesa. Purtroppo, come visto sopra, una funzione quadratica presenta dei limiti che ne impediscono un buon utilizzo. Ricordiamo anche che il modello media-varianza implica l'utilizzo solo del momento primo e secondo della distribuzione di probabilità. Questo va bene se i momenti superiori al secondo sono funzione dei momenti primo e secondo, come accade nel caso della distribuzione normale. Altrimenti il modello media-varianza va bene solo come approssimazione.

Capitolo 2

Modello Media-Varianza

2.1 Media-Varianza e Utilità Attesa

Un operatore economico è interessato a conoscere il profilo che può attendersi dalla sua scelta di investimento, ma è anche interessato a valutarne la rischiosità. La teoria dell'utilità attesa insegna a scegliere grandezze aleatorie tenendo conto di tutte le loro caratteristiche, ma in modo soggettivo, cioè sulla base della propria personale funzione di utilità. Se vogliamo avere una valutazione oggettiva del rischio possiamo ricorrere all'operatore *varianza*. La varianza è una misura naturale della rischiosità, intesa come probabilità di un esito lontano dal valore atteso.
Se decidiamo di usare i primi due momenti della distribuzione, cioè la media e la varianza, allora andiamo ad applicare il criterio *media-varianza* che consiste nello scegliere, a parità di varianza, la grandezza con valore atteso maggiore oppure, a parità di media, la grandezza con varianza minore.
L'analisi media-varianza e l'utilità attesa sono due approcci molto diversi per trattare il problema della selezione del portafoglio, di conseguenza il portafoglio ottimale determinato con l'utilità attesa generalmente non è efficiente in media-varianza. Infatti, utilizzando la formula di Taylor di ordine 2 con punto di origine in $\mathbb{E}[w]$ (vedere appendice A.6), possiamo sviluppare la funzione di utilità $U(w)$ come segue:

$$U(w) = U(\mathbb{E}[w]) + U'(\mathbb{E}[w])(w - \mathbb{E}[w]) + \frac{1}{2}U''(\mathbb{E}[w])(w - \mathbb{E}[w])^2 + R_2(w)$$

con $R_2(w)$ resto della formula di Taylor dato da:

$$R_2(w) = \sum_{k=3}^{+\infty} \frac{1}{k!} \cdot U^{(k)}(\mathbb{E}[w]) \cdot (w - \mathbb{E}[w])^k$$

Quindi la funzione di utilità prende in considerazione tutti i momenti della distribuzione di w. Con il criterio media-varianza andiamo, invece, a considerare solo i primi due momenti.
E' importante notare come il portafoglio determinato col metodo media-varianza e il portafoglio che ha maggiore utilità attesa coincidono solo nel caso in cui usiamo una funzione di utilità quadratica del tipo:

$$U(w) = w - \frac{\theta}{2}\, w^2 \tag{2.1}$$

Se la funzione di utilità è quadratica, lo sviluppo in serie di Taylor di ordine 2 non contiene il resto:

$$U(w) = U(\mathbb{E}[w]) + U'(\mathbb{E}[w])(w - \mathbb{E}[w]) + \frac{1}{2}\, U''(\mathbb{E}[w])(w - \mathbb{E}[w])^2$$

Poichè $\sum w - \mathbb{E}[w] = 0$, ciò che resta dello sviluppo in serie di Taylor di ordine 2 della funzione di utilità quadratica può essere facilmente riportato, salvo approssimazioni, alla 2.1. Questo mostra come un portafoglio efficiente secondo il criterio dell'utilità attesa è efficiente anche in media-varianza se la funzione di utilità è quadratica.
Il criterio media-varianza, per come è formulato, non permette di scegliere tra due grandezze che abbiano contemporaneamente sia valore atteso sia varianza diverse, ovvero, una grandezza aleatoria non è più descritta mediante un solo parametro numerico (ad esempio il valore atteso), ma da due parametri: media e varianza. Ogni grandezza può essere quindi rappresentata come un punto su un piano cartesiano nel quale sull'asse delle ascisse è riportata la varianza e sull'asse delle ordinate la media. L'insieme dei punti che stanno sulla linea che collega il punto più in alto (che ha la media maggiore tra tutti i punti del piano) con il punto più a sinistra (che ha la varianza minore tra tutti i punti) rappresenta la *frontiera efficiente*.
La scelta finale tra alternative efficienti richiede che in presenza di due variabili casuali X e Y tali che $\mathbb{E}[X] > \mathbb{E}[Y]$ e $Var[X] > Var[Y]$, l'operatore stabilisca se la differenza di valore medio tra X e Y sia per lui sufficiente a compensare la differenza tra le varianze. Questa scelta è puramente soggettiva in quanto interviene una maggiore o minore avversione al rischio da parte del decisore. E' opportuno, quindi considerare le *curve di indifferenza* individuali, ossia le curve sulle quali stanno i punti che rappresentano grandezze che l'operatore ritiene ugualmente desiderabili. Sovrapponendo il fascio delle curve d'indifferenza alla frontiera efficiente, si determina l'alternativa in assoluto migliore: quella corrispondente al punto della frontiera che si trova sulla curva di indifferenza di livello più alevato.

2.2 Analisi Media-Varianza nel singolo periodo

Iniziamo analizzando un generico caso di modello media-varianza considerando n titoli e un solo periodo. Utilizziamo la seguente notazione:

- x_v è il capitale investito nel titolo v;
- $x \in \mathbb{R}^n$ è il vettore dei portafogli di ampiezza $n \times 1$;
- $r \in \mathbb{R}^n$ è il vettore casuale dei guadagni, ovvero dei tassi di interesse;
- $m(x) = r'x$ è il rendimento, ;
- con e indichiamo un vettore composto da tutti 1, per cui $e'x = \sum_{i=1}^{n} x_i$.

Supponiamo inoltre che r sia generato da una distribuzione di probabilità congiunta tale che i primi due momenti siano i seguenti:

- media: $\mathbb{E}[r] = \overline{r}$
- matrice di varianza-covarianza: $\Sigma = \mathbb{E}[(r - \overline{r})(r - \overline{r})'] = \mathbb{E}[r'r] - \overline{r} \cdot \overline{r}'$

Il rendimento di un portafoglio è legato alla media dei guadagni, mentre il rischio è legato alla varianza. Più in dettaglio sarà:

$$\begin{aligned}
&\textit{Rendimento atteso} \Rightarrow && \mathbb{E}[m(x)] && = \mathbb{E}[r'x] = \overline{r}'x \\
&\textit{Rischio} \Rightarrow && \sigma^2[m(x)] && = \sigma^2[r'x] = \\
& && && = \mathbb{E}[r'x - \mathbb{E}(r'x)^2] = \\
& && && = \mathbb{E}[x'(r - \overline{r})(r - \overline{r})'x] = \\
& && && = x'\Sigma x
\end{aligned}$$

Dal differenziale totale dell'espansione di Taylor del secondo ordine, ci ricaviamo le curve di indifferenza tra rendimento e rischio:

$$\frac{\mathrm{d}\, m}{\mathrm{d}\, \sigma} = -\frac{U''}{U'} \cdot \sigma = ARA \cdot \sigma$$

Per un soggetto avverso al rischio, le curve di indifferenza sono crescenti e la loro inclinazione aumenta all'aumentare del rischio. La forma della curva di indifferenza, invece, dipende dall'indice di avversione al rischio.

Il problema che ci poniamo ora è di riuscire a determinare una combinazione di titoli che ci permetta di ottenere il più elevato livello di rendimento per ogni livello di rischio, ovvero il minor livello di rischio per ogni livello di rendimento.
Possiamo operare seguendo due strade: massimizzando il valore atteso di una funzione di utilità quadratica, oppure minimizzando il rischio. Le due formulazioni sono equivalenti.
Se decidiamo di seguire il primo modo proposto, ci troveremo con la seguente formulazione:

$$\begin{aligned} \max_x \quad & \mu\, m(x) - \frac{1}{2}\, \sigma^2[m(x)] \\ \text{s.v.} \quad & e'x = 1 \end{aligned} \tag{2.2}$$

dove *s.v.* sta per *sotto il vincolo*.
Un'alternativa è cercare di minimizzare il rischio fissando il livello di rendimento:

$$\begin{aligned} \min_x \quad & \frac{1}{2}\, \sigma^2[m(x)] \\ \text{s.v.} \quad & e'x = 1 \\ & m(x) = m \end{aligned} \tag{2.3}$$

Nei prossimi paragrafi esamineremo alcuni casi del modello media-varianza, in particolare vedremo cosa succede se siamo in presenza di soli titoli rischiosi e cosa accade invece se è presente almeno un titolo privo di rischio.

2.3 Portafoglio con soli titoli rischiosi

2.3.1 Caso con due titoli rischiosi

Partiamo con due titoli rischiosi x_i $(i = 1, 2)$, con:

$$\begin{array}{ll} \text{Rendimento:} & m(x_i) = m_i; \\ \text{Volatilità:} & \sigma[m(x_i)] = \sigma_i. \end{array}$$

La parte di ricchezza investita nel primo titolo è pari ad α, mentre quella investita nel secondo titolo è $1 - \alpha$. La combinazione dei due titoli ci fornisce il portafoglio $t = \alpha x_1 + (1 - \alpha)x_2$, con rendimento atteso $\mathbb{E}[m_t]$ e volatilità σ_t come segue:

$$\mathbb{E}[m_t] = \alpha\mathbb{E}[m_1] + (1 - \alpha)\mathbb{E}[m_2]$$

$$\sigma_t = \sqrt{\alpha^2\sigma_1^2 + (1 - \alpha)^2\sigma_2^2 + 2\alpha(1 - \alpha)\rho\,\sigma_1\sigma_2}$$

dove $\rho = cov(m_1, m_1)/\sigma_1\sigma_2$ è il coefficiente di correlazione.
Poichè $\rho \in [-1, 1]$, analizziamo i due casi estremi: $\rho = 1$ e $\rho = -1$.

Perfetta correlazione positiva: $\rho = 1$

La volatilità è data da:

$$\sigma_t = \alpha\sigma_1 + (1-\alpha)\sigma_2$$

Ricavando $(1-\alpha)$, otteniamo un relazione lineare tra rendimento e rischio:

$$\mathbb{E}[m_t] = \mathbb{E}[m_1] + \frac{\mathbb{E}[m_2 - m_1]}{\sigma_2 - \sigma_1}(\sigma_t - \sigma_1)$$

Se poniamo il rendimento sull'asse delle ascisse e la volatilità sull'asse delle ordinate, il portafoglio sta sulla retta che congiunge il punto $A = (\sigma_1, \mathbb{E}[m_1])$ e $B = (\sigma_2, \mathbb{E}[m_2])$.

Perfetta correlazione negativa: $\rho = -1$

In questo caso la volatilità può assumere sia valori positivi, sia negativi:

$$\sigma_t = \pm[\alpha\sigma_1 - (1-\alpha)\sigma_2]$$

Notiamo che esiste un portafoglio a volatilità nulla, ovvero privo di rischio:

$$\alpha_0 = \frac{\sigma_2}{\sigma_1 + \sigma_2}$$

$$\begin{aligned} r_0 \quad &= \mathbb{E}[m_1] + (1-\alpha_0)\mathbb{E}[m_2 - m_1] = \\ &= \mathbb{E}[m_1] + \sigma_1 \tfrac{\mathbb{E}[m_2 - m_1]}{\sigma_1 + \sigma_2} \end{aligned}$$

Per quanto riguarda la volatilità, possiamo tenere la radice positiva o la radice negativa e valutare i due sottocasi: $\alpha < \alpha_0$ e $\alpha > \alpha_0$.

radice negativa: $\alpha < \alpha_0$

$$\sigma_t = -[\alpha\sigma_1 - (1-\alpha)\sigma_2] = \sigma_2 - \alpha(\sigma_1 + \sigma_2) = \sigma_2\Big[1 - \frac{\alpha}{\alpha_0}\Big] > 0$$

$$\mathbb{E}[\rho_t] = m_0 + \frac{\mathbb{E}[m_2 - m_1]}{\sigma_1 + \sigma_2}\sigma_t$$

Il portafoglio sta sulla retta con inclinazione positiva che unisce l'intercetta $I = (0, m_0)$ con il punto $A = (\sigma_1, \mathbb{E}[m_1])$.

radice positiva: $\alpha > \alpha_0$

$$\sigma_t = +[\alpha\sigma_1 - (1-\alpha)\sigma_2] = \alpha(\sigma_1 + \sigma_2) - \sigma_2 = \sigma_2\Big[\frac{\alpha}{\alpha_0} - 1\Big] > 0$$

$$\mathbb{E}[m_t] = m_0 - \frac{\mathbb{E}[m_2 - m_1]}{\sigma_1 + \sigma_2}\sigma_t$$

Il portafoglio sta sulla retta con inclinazione negativa che unisce l'intercetta $I = (0, m_0)$ con il punto $B = (\sigma_2, \mathbb{E}[m_2])$.

Figura 2.1: Grafico rischio-rendimento nel modello con due titoli rischiosi: casi estremi. (c.n.= correlazione negativa)

Nella figura 2.1 si può osservare la retta segnata da un linea continua che rappresenta il caso in cui ci sia correlazione positiva ($\rho = 1$), e le due rette tratteggiate che rappresentano il caso di correlazione negativa ($\rho = -1$) con distinti i due sottocasi: se $\alpha < \alpha_0$ la retta ha un'inclinazione positiva, mentre se $\alpha > \alpha_0$ avremo un retta con inclinazione negativa.

2.3.2 Caso con più titoli rischiosi

Supponiamo di avere un vettore $x_r \in \mathbb{R}^n$ con n titoli rischiosi e poniamo le seguenti *ipotesi*:

$A1$) la matrice di varianza e covarianza Σ è definita positiva;

$A2$) il rendimento $\overline{r}$ non è multiplo di e.

Con $A1$ assumiamo che tutti i titoli siano rischiosi (in quanto $\Sigma \neq 0$), mentre con $A2$ si vuole evitare di incorrere in situazioni degeneri.
Dalla $A1$ ricaviamo le seguenti costanti:

$$\omega = e'\Sigma^{-1}e$$
$$\beta = e'\Sigma^{-1}\overline{r}$$
$$\gamma = \overline{r}'\Sigma^{-1}\overline{r}$$
$$\delta = \omega\gamma - \beta^2$$

Per cercare il portafoglio ottimale, iniziamo utilizzando una formulazione simile alla 2.2:

$$\max_x \mu(\overline{r}'x_r) - \frac{1}{2}x_r'\Sigma x_r$$
$$\text{s.v. } e'x_r = 1$$

Risolviamo con l'ausilio del Lagrangiano:

$$\begin{aligned} L(x_r, \lambda; \mu) \quad &= \tfrac{1}{2}x_r'\Sigma x_r - \mu(\overline{r}'x_r) - \lambda(e'x_r - 1) = \\ &= \tfrac{1}{2}\sigma^2(m) - \mu(\overline{r}'x_r) - \lambda(e'x_r - 1) \end{aligned}$$

dove μ è il *moltiplicatore di rendimento* e λ è il *moltiplicatore di budget*.
Troviamo una soluzione unica:

$$x_r = \Sigma^{-1}(\lambda e + \mu\overline{r})$$

$$\lambda = \frac{1 - \mu\beta}{\omega}$$

da cui ricaviamo

$$m = \lambda\beta + \mu\gamma = \frac{\beta + \mu\delta}{\omega}$$

In alternativa alla formulazione 2.2, possiamo minimizzare il rischio seguendo la 2.3:

$$\min_x \frac{1}{2}\, x_r'\Sigma x_r$$

$$\begin{aligned} \text{s.v.} \quad & e'x_r = 1 \\ & \overline{r}'x_r = m \end{aligned}$$

Anche qui applichiamo il Lagrangiano:

$$\begin{aligned} L(x_r, \lambda, \mu; \delta) \quad &= \tfrac{1}{2}x_r'\Sigma x_r - \mu(\overline{r}'x_r - m) - \lambda(e'x_r - 1) = \\ &= \tfrac{1}{2}\sigma^2(m) - \mu(\overline{r}'x_r - m) - \lambda(e'x_r - 1) \end{aligned}$$

e otteniamo una soluzione unica:

$$x_r = \Sigma^{-1}(\lambda e + \mu\overline{r})$$

$$\lambda = \frac{\gamma - \beta m}{\delta}$$

$$\mu = \frac{\omega m - \beta}{\delta}$$

La formulazione 2.2 da cui deriva il parametro μ, e la formulazione 2.3 da cui deriva il parametro m, sono equivalenti se e solo se $\mu = m$ *ottimale*, ovvero se $m = \mu$ *ottimale*.
Infatti:

$$\lambda = \frac{1 - \mu\beta}{\omega} = \frac{\delta - \omega\beta m + \beta^2}{\omega m} = \frac{\omega\gamma - \omega\beta m}{\omega m} = \frac{\gamma - \beta m}{m}$$

Ne deriva che il rischio ottimale è una funzione quadratica di m:

$$\sigma^2(m) = \frac{\omega\rho^2 - 2\beta m + \gamma}{\delta} = \frac{\mu^2\delta + 1}{\omega}$$

Il punto di minimo della parabola definisce il rischio minimo ottenibile. In questo caso i valori saranno:

$$\sigma^2(\hat{m}) = \frac{1}{\omega} \qquad \text{e} \qquad \hat{m} = \frac{\beta}{\omega}$$

da cui ricaviamo che

$$\hat{x} = \frac{\Sigma^{-1}}{\omega} \qquad \hat{\lambda} = \frac{1}{\omega} \qquad \hat{\mu} = 0$$

Il grafico di $\sigma^2(m)$ rappresenta la *frontiera efficiente*.

2.4 Portafoglio con titoli rischiosi e non rischiosi

2.4.1 Caso con un titolo rischioso e uno privo di rischio

Inseriamo ora nel nostro problema un titolo non rischioso e partiamo dall'ipotesi di avere complessivamente solo due titoli. Ci troveremo quindi con un titolo privo di rischio x_p che avrà un rendimento $m(x_p) = m_p$, e un titolo rischioso x_r con rendimento $m(x_r) = m_r$ e varianza $\sigma^2(x_r) = \sigma_r^2$. Il titolo privo di rischio avrà varianza nulla proprio perchè è privo di rischio.
Indichiamo poi con α la quota di ricchezza investita nel titolo privo di rischio, quindi $1-\alpha$ sarà la quota destinata al titolo rischioso. Il portafoglio t è dato dalla combinazione dei due titoli, ovvero $t = \alpha x_p + (1-\alpha)x_r$, e avrà rendimento atteso $\mathbb{E}[m_t]$ e volatilità σ_t definiti come segue:

$$\mathbb{E}[m_t] = \alpha m_p + (1-\alpha)\mathbb{E}[m_r]$$

$$\sigma_t = (1-\alpha)\sigma_r$$

La relazione tra rendimento atteso e volatilità è di tipo lineare:

$$\mathbb{E}[m_t] = m_p + \frac{\mathbb{E}[m_p - m_r]}{\sigma_r}\sigma_t$$

L'intercetta è data dal rendimento del titolo privo di rischio, mentre l'inclinazione è rappresentata dal rapporto tra eccesso di rendimento del titolo rischioso rispetto al titolo privo di rischio e la volatilità del titolo rischioso. Il portafoglio con la combinazione ottimale di titoli sarà nel punto dove la retta delle combinazioni è tangente alla curva di indifferenza, ovvero:

$$ARA\sigma_t = \frac{\mathbb{E}[m_r - m_p]}{\sigma_r}$$

e la quota di ricchezza investita nel titolo rischioso sarà:

$$\begin{aligned} 1-\alpha^* &= \tfrac{\mathbb{E}[m_r - m_p]}{\sigma_r}\tfrac{1}{ARA\sigma} = \\ &= \tfrac{\mathbb{E}[m_r - m_p]}{\sigma_r{}^2 ARA} = \\ &= \tfrac{\mathbb{E}[m_r - m_p]}{\sigma_r{}^2}\left[\tfrac{w}{1-\gamma}\tfrac{\eta}{\beta}\right] \end{aligned}$$

dove $\frac{w}{1-\gamma}\frac{\eta}{\beta}$ è l'indice di tolleranza al rischio, $\frac{1}{ARA}$.
Partendo da una funzione di tipo ARA abbiamo ottenuto una funzione HARA che

garantisce una relazione lineare tra ricchezza (w) e domanda di titoli richiosi. La relazione tra avversione al rischio e ricchezza, infatti, è direttamente legata alla domanda di titoli rischiosi.

2.4.2 Caso con n-1 titoli rischiosi e un titolo privo di rischio

Supponiamo di avere un titolo non rischioso (x_p) e $n-1$ titoli rischiosi ($x_r \in \mathbb{R}^{n-1}$), con relativi rendimenti m_p e m_r. La matrice di varianza e covarianza relativa ai titoli rischiosi sarà Σ; il titolo privo di rischio avrà volatilità nulla ($\sigma_p = 0$).
Assumiamo le seguenti *ipotesi*:

$A1$) la matrice di varianza-covarianza Σ è definita positiva;

$A3$) $\overline{r}_r \neq e' r_p$

Con la prima assunzione imponiamo che tutti gli $n-1$ titoli rischiosi siano veramente rischiosi; la seconda assunzione ci serve (come la $A2$) per escludere situazioni degeneri.
Per trovare il portafoglio ottimale procediamo in modo analogo al caso in cui avevamo n titoli rischiosi. Possiamo scegliere di utilizzare il valore atteso della funzione di utilità quadratica, come nella formulazione 2.2:

$$\max_x \mu(\overline{r}'_r x_r + r_p x_p) - \frac{1}{2} x' \Sigma x$$

$$\text{s.v. } e' x_r + x_p = 1$$

Oppure possiamo minimizzare il rischio seguendo la formulazione 2.3:

$$\min_x \frac{1}{2} x' \Sigma x$$

$$\begin{aligned} \text{s.v.} \quad & e' x_r + x_p = 1 \\ & \overline{r}'_r x_r + r_p x_p = m \end{aligned}$$

Troviamo un'unica soluzione:

$$x_r = \Sigma^{-1}(\lambda e + \mu \overline{r}_r) = \mu \Sigma^{-1}(\overline{r}_r - r_p\, e)$$

$$x_p = 1 - \mu(\beta - r_p\, \omega)$$

$$\lambda = -r_p\, \mu$$

$$\mu = \frac{m - r_p}{\delta_p}$$

con $\delta_p = (\overline{r}_r - r_p\, e)'\Sigma^{-1}(\overline{r}_r - r_p\, e)$.
Il rischio ottimale dunque risulta essere il seguente:

$$\sigma^2(\rho) = \frac{(m - r_p)^2}{\delta_p} = \mu^2\delta^2$$

2.5 Teorema di separazione dei fondi

Secondo il *teorema di separazione dei fondi*, il vettore ottimale dei pesi (x) si può ottenere dalla combinazione lineare di due portafogli:

$$x^* = \lambda_1\beta\frac{\Sigma^{-1}r}{\beta} + \lambda_2\omega\frac{\Sigma^{-1}e}{\omega}$$

da cui possiamo definire i due portafogli

$$x_1 = \frac{\Sigma^{-1}r}{\beta} = \frac{\Sigma^{-1}r}{e'\Sigma^{-1}r} \qquad x_2 = \frac{\Sigma^{-1}e}{\omega} = \frac{\Sigma^{-1}e}{e'\Sigma^{-1}e}$$

I due portafogli rispettano il vincolo $e'x = 1$ e i loro rispettivi rendimenti attesi e varianze sono:

$$\mathbb{E}[m_1] = r'x_1 = \frac{r'\Sigma^{-1}r}{e'\Sigma^{-1}r} = \frac{a}{\beta}$$

$$\sigma_1^2 = x_1'\Sigma^{-1}x_1 = \frac{\gamma}{\beta^2}$$

$$\mathbb{E}[m_2] = \Sigma^{-1}r'x_2 = \frac{r'\Sigma^{-1}e}{e'\Sigma^{-1}e} = \frac{\beta}{\omega}$$

$$\sigma_2^2 = x_2'\Sigma^{-1}x_2 = \frac{1}{\omega}$$

Si può notare come il portafoglio 2 è quello di minima varianza, mentre il portafoglio 1 è definito da una retta che passa per l'origine e per il portafoglio di minima varianza. Infatti, x_2 gode della seguente proprietà:

$$cov(m_v, m_2) = x_v'\Sigma^{-1}x_2 = x_v'\Sigma\frac{\Sigma^{-1}e}{\omega} = \frac{1}{\omega} = \sigma_2^2$$

ovvero, la covarianza di un qualsiasi portafoglio col portafoglio di minima varianza è pari alla varianza di quest'ultimo.
Applicando le condizioni del primo ordine otteniamo:

$$\mathbb{E}[m_t] = \beta\lambda_1\mathbb{E}[m_1] + \omega\lambda_2\mathbb{E}[m_2] = \gamma\lambda_1 + \beta\lambda_2$$

con $\beta\lambda_1 + \omega\lambda_2 = 1$.

Poichè

$$\alpha^* = \beta\lambda_1 = \frac{\beta(\omega\mathbb{E}[\rho_t] - \beta)}{\gamma\omega - \beta^2} = \frac{\beta\omega\left(\mathbb{E}[m_t] - \frac{\beta}{\omega}\right)}{\beta\omega\left(\frac{\gamma}{\beta} - \frac{\beta}{\omega}\right)} = \frac{\mathbb{E}[m_t] - \mathbb{E}[m_2]}{\mathbb{E}[m_1] - \mathbb{E}[m_2]}$$

si può facilmente ricavare la combinazione lineare di portafogli che ci permette di raggiungere il vettore ottimale dei pesi:

$$x^* = \alpha^* x_1 + (1 - \alpha^*) x_2$$

che è ciò che volevamo ottenere.

Parte II

Varianti del modello Media-Varianza

Capitolo 3

Estensioni del modello Media-Varianza

3.1 Alternative per la misura del rischio

Nel 1950 Harry Markowitz (si veda [16]) introdusse un modello media-varianza per la selezione del portafoglio nel singolo periodo, usando la varianza come misura del rischio. Purtroppo l'utilizzo della varianza, e quindi di una funzione di utilità quadratica, comporta grosse limitazioni come abbiamo già visto nella sezione 1.4. Sorge dunque la necessità di cercare valide alternative alla varianza per misurare il rischio.

Jin, Ya e Zhou [8] propongono alcuni modelli nei quali vengono prese in considerazione diverse formulazioni per la misura del rischio. Questi modelli sono delle estensioni al seguente problema generale:

$$\min_X \mathbb{E}[f(X - \mathbb{E}[X])] \tag{3.1}$$

$$\begin{aligned} \text{s.v.} \quad & \mathbb{E}[X] = z \\ & \mathbb{E}[\rho X] = x_0 \end{aligned}$$

dove x_0 è la ricchezza iniziale.

Il problema media-varianza di Markowitz è un caso speciale del problema 3.1 con $f(x) = x^2$. Nel modello classico, infatti, si vuole trovare il portafoglio efficiente che consente di minimizzare la varianza, che può essere scritta come

$$Var[x] = \mathbb{E}[x^2] - (\mathbb{E}[x])^2$$

da cui ricaviamo che il problema di ottimizzazione di Markowitz è:

$$\min_x \mathbb{E}[f(x - \mathbb{E}[x])] = \min_x \mathbb{E}[(x - \mathbb{E}[x])^2] = \min_x Var[x]$$

che è facilmente riconducibile al problema 2.3.
Nei paragrafi che seguono andremo a considerare vari modelli. Inizieremo con un modello media-varianza pesato, nel quale si danno pesi diversi a rendimenti diversi; questo modello ammette una soluzione esplicita. Analizzeremo poi un modello media-semivarianza, nel quale, pur non essendoci una soluzione ottimale, è comunque possibile ottenere una soluzione asintotica in quanto la varianza pesata converge alla semivarianza nel caso in cui il peso assegnato al livello minimo di varianza tende a zero. Un ulteriore modello che prenderemo in considerazione è il modello media-downside risk nel quale solo il rendimento minimo è soggetto a penalizzazione.

3.1.1 Modello media-varianza pesato

Nel modello media-varianza classico si usa la varianza come misura del rischio assegnando lo stesso peso sia ai rendimenti più alti sia a quelli minori. Un modello media-varianza pesato, invece, assegna pesi diversi a diversi livelli di rendimenti. Ipotizziamo di avere una funzione $f(x) = \alpha x_+^2 + \beta x_-^2$ con:

- $x_+ = \max(x, 0)$, parte positiva di x;
- $x_- = -\min(x, 0)$, parte negativa di x.

Ci si può ricondurre al modello media-varianza classico ponenedo $\alpha = \beta$.
Dato $Y = X - \mathbb{E}[X]$, il problema di ottimizzazione è:

$$\min_Y \mathbb{E}[\alpha Y_+^2 + \beta Y_-^2] \tag{3.2}$$

$$\begin{aligned} \text{s.v.} \quad & \mathbb{E}[Y] = 0 \\ & \mathbb{E}[\rho Y] = y_0 \end{aligned}$$

Il problema 3.2 è equivalente al seguente:

$$\min_Y \mathbb{E}[\alpha Y_+^2 + \beta Y_-^2 - 2(\lambda - \mu\rho)Y] \tag{3.3}$$

Troveremo un'unica soluzione ottimale al problema 3.3 data da:

$$Y^* = \frac{(\lambda - \mu\rho)_+}{\alpha} - \frac{(\lambda - \mu\rho)_-}{\beta}$$

dove la coppia (λ, μ) è soluzione unica del seguente sistema:

$$\begin{cases} \frac{\mathbb{E}[(\lambda-\mu\rho])_+}{\alpha} - \frac{\mathbb{E}[(\lambda-\mu\rho])_-}{\beta} = 0 \\ \frac{\mathbb{E}[\rho(\lambda-\mu\rho])_+}{\alpha} - \frac{\mathbb{E}[\rho(\lambda-\mu\rho])_-}{\beta} = y_0 \end{cases} \tag{3.4}$$

Poichè i problemi 3.2 e 3.3 sono equivalenti, avranno la stessa soluzione Y^*.

3.1.2 Modello media-semivarianza

Nel modello media-semivarianza viene penalizzato solo il rendimento più basso. Useremo quindi una funzione del tipo $f(x) = x_-^2$.
Il problema di ottimo diventa:

$$\min \mathbb{E}[Y_-^2] \tag{3.5}$$

$$\text{s.v.} \quad \mathbb{E}[Y] = 0$$
$$\mathbb{E}[\rho Y] = y_0$$

Il problema 3.5 non ha una soluzione ottimale, ma ponendo $\beta = 1 - \alpha$ possiamo considerarlo come una variante del problema 3.2, quindi la soluzione del problema 3.5 può essere vista come variante della soluzione del problema 3.2, ovvero:

$$Y^*(\alpha) = \frac{\big(\lambda(\alpha) - \mu(\alpha)\rho\big)_+}{\alpha} - \frac{\big(\lambda(\alpha) - \mu(\alpha)\rho\big)_-}{\beta}$$

Y^* soddisfa i vincoli $\mathbb{E}[Y(\alpha)] = 0$ e $\mathbb{E}[\rho Y(\alpha)] = y_0$ ed è perciò soluzione ammissibile per il problema 3.5.

3.1.3 Modello media-downside risk

Il modello downside risk considera, come misura del rischio, la deviazione minima dei rendimenti rispetto alla media. Supponendo che $f(x) \geq 0$ e che $f(x) = 0$ $\forall\, x \in \mathbb{R}^+$ con $f(x)$ strettamente decrescente $\forall\, x \in \mathbb{R}^-$, il problema di ottimizzazione per il modello media-downside risk è:

$$\min \mathbb{E}[f(Y)] \tag{3.6}$$

$$\text{s.v.} \quad \mathbb{E}[Y] = 0$$
$$\mathbb{E}[\xi Y] = y_0$$

dove ξ è una costante positiva. Il problema 3.6 ammette soluzione solo se $y_0 = 0$. La semivarianza, affrontata nel problema 3.5, può essere vista come un caso particolare del downside risk.

3.2 Modelli di avversione all'ambiguità

Negli ultimi anni c'è stata una discussione in macroeconomia sulle ipotesi di aspettative razionali, secondo le quali tutti gli agenti hanno la stessa distribuzione di probabilità. Per descrivere questo comportamento razionale degli agenti, possiamo utilizzare una classe di modelli avversi all'ambiguità come i modelli *Multiplier Priors Preferences* (conosciuti anche come *Maximin Expected Utility*) che sono identificati dalla seguente funzione criterio:

$$U(f) = \min_{Q \in C} \int u(f)\, \mathrm{d}Q$$

dove C è un sottoinsieme convesso dell'insieme Δ di tutte le probabilità degli stati. Questo tipo di modelli, così come definiti da Kelsey [10], prevedono che l'azione a è preferibile all'azione b se e solo se il minimo possibile valore dell'utilità attesa di a è maggiore del valore atteso minimo di b.
Possiamo riscrivere le *Multiplier Priors Preferences*, usando una formulazione più generica:

$$U(f) = \min_{Q \in \Delta} \left\{ \int u(f)\, \mathrm{d}Q + \delta_C(Q) \right\}$$

con

$$\delta_C(Q) = \begin{cases} 0 & \text{se } Q \in C \\ \infty & \text{altrimenti} \end{cases}$$

In realtà gli agenti possono avere modelli diversi e sono consapevoli della possibilità che il loro modello sia non correttamente specificato. Questa incertezza nel modello è dovuta all'ambiguità causata da una qualità scadente delle informazioni iniziali. Entra quindi in gioco un approccio alle preferenze di tipo robusto introdotto da Hansen e Sargent [7] nel quale si considera la possibilità che il modello dell'agente non sia quello corretto ma sia solo un'approssimazione del vero modello.
Si arriva così a specificare una classe di modelli, le *Multiplier Preferences*, con una funzione criterio così definita:

$$U(f) = \min_{Q \in \Delta} \left\{ \int u(f)\, \mathrm{d}Q + \theta R(Q||P) \right\}$$

dove:

- Δ è l'insieme di tutte le probabilità;
- θ, parametro positivo, riflette il peso che l'agente dà alla possibilità che P possa non essere la vera legge che governa il fenomeno;

- $R(Q||P) : \Delta \to [0, \infty)$ è l'entropia relativa.

Poichè P potrebbe non essere il vero modello, possiamo prendere in considerazione un altro possibile modello, Q, la cui verosimiglianza è misurata con l'entropia relativa, $R(Q||P)$. In particolare $\theta R(Q||P)$ è una funzione indicatrice di analisi convessa e può essere vista, insieme a $\delta_C : \Delta \to [0, \infty)$, come una estensione degli indici di ambiguità.
Un modello *Multiplier Preferences* può essere quindi così generalizzato:

$$U(f) = \min_{Q \in \Delta} \left\{ \int u(f)\, \mathrm{d}Q + c(Q) \right\}$$

con $c : \Delta \to [0, \infty)$ funzione convessa.
Le *Multiplier Preferences* sono un caso particolare di una classe più ampia di preferenze, chiamata *Divergence Preferences* (vedere appendice, sezione C.4), nella quale rientrano anche le *preferenze media-varianza* introdotte da Markowitz:

$$U(f) = \int f\, \mathrm{d}P - \frac{1}{2\theta} Var(f)$$

Si può dimostrare, e lo vedremo nel dettaglio con il modello media-varianza monotono, come nel dominio di monotonicità di $U(f)$, vale la seguente relazione:

$$\int f\, \mathrm{d}P - \frac{1}{2\theta} Var(f) = \min_{Q \in \Delta} \left\{ \int f\, \mathrm{d}Q + \theta C(Q||P) \right\}$$

dove $C(Q||P) : \Delta \to [0, \infty]$ è l'indice di concentrazione di Gini.

3.2.1 Indice di concentrazione di Gini

Le preferenze $\succeq^g \in F$ sono *preferenze di Gini* se:

$$f \succeq^g g \Leftrightarrow \min_{Q \in \Delta^\sigma(P)} \left\{ \int u(f) \mathrm{d}Q + \theta C^w(Q||P) \right\} \geq \min_{Q \in \Delta^\sigma(P)} \left\{ \int u(g) \mathrm{d}Q + \theta C^w(Q||P) \right\}$$

dove $C^w(Q||P) : \Delta \to [0, \infty]$ è l'indice di concentrazione di Gini pesato dato da:

$$C^w(Q||P) = \begin{cases} \int_S w(s) \frac{1}{2} \left(\frac{\mathrm{d}Q}{\mathrm{d}P}(s) - 1 \right)^2 \mathrm{d}P(s) & \text{se } Q \in \Delta^\sigma(P) \\ \infty & \text{altrimenti} \end{cases}$$

L'indice di concentrazione di Gini è un caso speciale della divergenza $D^w_\Phi(Q||P)$ con

$\Phi(t) = 2^{-1}(t-1)^2$, (vedere appendice, sezione C.4).
Le preferenze di Gini sono preferenze variazionali continue, con indice di avversità all'ambiguità definito come segue:

$$c^*(Q) = \theta C^w(Q||P) \qquad \forall\, Q \in \Delta$$

Il parametro θ può essere interpretato come coefficiente di avversione all'ambiguità.
L'indice di Gini è utile per studiare le preferenze media-varianza che sono date da:

$$f \succeq^{mv} g \Leftrightarrow \int u(f)\mathrm{d}P - \frac{1}{2\theta}Var[u(f)] \geq \int u(g)\mathrm{d}P - \frac{1}{2\theta}Var[u(g)]$$

con $\theta > 0$, $P \in \Delta^\sigma$; $Var(.)$ è la varianza calcolata con riferimento a P.
Se la funzione $u(\cdot)$ fosse una funzione neutrale al rischio, ovvero $u(x) = x$, allora

$$f \succeq^{mv} g \Leftrightarrow \int f\,\mathrm{d}P - \frac{1}{2\theta}Var(f) \geq \int g\,\mathrm{d}P - \frac{1}{2\theta}Var(g)$$

che è esattamente il modello introdotto da Markowitz.
Le preferenze di Gini che hanno una w uniforme, ovvero quando $w(s) = 1\ \forall\, s \in S$, prendono il nome di *preferenze media-varianza monotone*:

$$f \succeq^{mmv} g \Leftrightarrow \min_{Q \in \Delta^\sigma(P)} \left\{ \int u(f)\,\mathrm{d}Q + \theta C(Q||P) \right\} \geq \min_{Q \in \Delta^\sigma(P)} \left\{ \int u(g)\,\mathrm{d}Q + \theta C(Q||P) \right\}.$$

Capitolo 4

Modello Media Varianza Monotono

4.1 Monotone Mean-Variance Preferences

La funzione di utilità del modello media-varianza è definita da:

$$U_\theta(f) = \mathbb{E}^P[f] - \frac{\theta}{2} Var^P[f]$$

dove P è una misura di probabilità e θ è l'indice di avversione al rischio. Consideriamo un esempio di selezione del portafoglio secondo una funzione di utilità media-varianza con $\theta = 2$ supponendo di dover scegliere tra due alternative f e g in base al seguente prospetto:

Stati di Natura	Probabilità	f	g
s_1	0.25	1	1
s_2	0.25	2	2
s_3	0.25	3	3
s_4	0.25	4	5

La funzione di utilità per l'alternativa f è data da:

$$U_2(f) = \mathbb{E}^P[f] - \frac{2}{2} Var^P[f]$$

con

$$\mathbb{E}^P[f] = \sum_{i=1}^{4} P(s_i) f_i \qquad \text{e} \qquad Var^P[f] = \sum_{i=1}^{4} P(s_i)(f_i - \mathbb{E}^P[f])^2$$

Quindi

$$\mathbb{E}^P[f] = 0.25(1) + 0.25(2) + 0.25(3) + 0.25(4) = 2.5$$

$$Var^P[f] = 0.25(1-2.5)^2 + ... + 0.25(4-2.5)^2 = 1.25$$

da cui

$$U_2(f) = 2.5 - 1.25 = 1.25$$

Allo stesso modo ci calcoliamo la funzione di utilità per l'alternativa g:
$\mathbb{E}^P[g] = 0.25(1) + 0.25(2) + 0.25(3) + 0.25(5) = 2.75$
$Var^P[g] = 0.25(1-2.75)^2 + ... + 0.25(5-2.75)^2 = 2.1875$
da cui

$$U_2(g) = 2.75 - 2.1875 = 0.5625$$

Abbiamo che $f \succ g$:

$$U_2(f) = 1.25 > U_2(g) = 0.5625$$

L'utilità dell'alternativa f è maggiore rispetto all'utilità dell'alternativa g, ma osservando bene i payoff risulta evidente che l'alternativa g è preferibile all'alternativa f. Questo esempio mostra come il modello media-varianza può portare a risultati paradossali. Introduciamo allora un nuovo medello di selezione del portafoglio: il modello media-varianza monotono.
Il modello monotono media-varianza (mmv) è un modello di selezione del portafoglio basato su una classe di preferenze che coincidono con il modello media-varianza (mv) per la parte dove il modello media-varianza ha significato economico e lo estende in maniera monotona negli altri casi.
Per questo modello definiamo la seguente funzione di utilità:

$$V_\theta(f) = \min_Q \left\{ \mathbb{E}^Q(f) + \frac{1}{2\theta} C(Q||P) \right\} \qquad \forall f \in \mathcal{L}^2(P)$$

con

$\theta =$ indice di avversione al rischio ($\theta > 0$);
$P, Q =$ misure di probabilità;
$C(Q||P) =$ indice di concentrazione di Gini (o χ^2 *distance*).

Sia $\mathcal{G}_\theta$ l'insieme delle funzioni di $\mathcal{L}^2(P)$ per le quali il differenziale di Gateaux di U_θ è positivo. Questo insieme è chiuso e convesso e lo definiremo il dominio di monotonicità di U_θ. Vale il seguente risultato:

$$\mathcal{G}_\theta = \left\{ f \in \mathcal{L}^2(P) : f - \mathbb{E}^P(f) \leq \frac{1}{\theta} \right\}$$

Proposizione 1 *Il dominio $\mathcal{G}_\theta$ coincide con* $\left\{f \in \mathcal{L}^2(P) : f - \mathbb{E}^P(f) \leq \frac{1}{\theta}\right\}$

Dimostrazione

Proviamo che se il differenziale di Gateaux di U_θ in f è positivo, allora

$$\mathcal{G}_\theta = \left\{f \in \mathcal{L}^2(P) : f - \mathbb{E}^P(f) \leq \frac{1}{\theta}\right\}$$

Partiamo considerando che

$$U_\theta(f) = \mathbb{E}[f] - \frac{\theta}{2} Var[f]$$

quindi

$$\begin{aligned} U_\theta(f+tv) &= \mathbb{E}[f+tv] - \tfrac{\theta}{2} Var[f+tv] = \\ &= \mathbb{E}[f] + t\mathbb{E}[v] - \tfrac{\theta}{2}\left\{Var[f] + t^2 Var[v] + 2Cov(f, tv)\right\} \end{aligned}$$

allora possiamo ricavare il differenziale di Gateaux per $U_\theta(f)$ lungo il versore v (vedere appendice, sezione B.1):

$$\begin{aligned} d^+U_\theta(f)(v) &= \lim_{t\to 0^+} \tfrac{U_\theta(f+tv) - U_\theta(f)}{t} = \\ &= \lim_{t\to 0^+} \tfrac{\mathbb{E}[f] + tE[v] - \frac{\theta}{2}Var[f] - \frac{\theta}{2}t^2 Var[v] - \frac{\theta}{2}\cdot 2Cov(f,tv) - \mathbb{E}[f] + \frac{\theta}{2}Var[f]}{t} = \\ &= \lim_{t\to 0^+}\left[\mathbb{E}[v] - \tfrac{\theta}{2} t Var[f] - \theta \tfrac{tCov(f,v)}{t}\right] = \\ &= \mathbb{E}[v] - \theta\ Cov(f,v). \end{aligned}$$

Essendo

$$Cov(f, v) = \mathbb{E}[f \cdot v] - \mathbb{E}[f]\mathbb{E}[v]$$

otteniamo che

$$d^+U_\theta(f)(v) = \mathbb{E}[v] - \theta\ (\mathbb{E}[f \cdot v] - \mathbb{E}[f]\mathbb{E}[v]).$$

Sapendo che:

- *se*$\mathcal{G}_\theta$ è un sottoinsieme di $\mathcal{L}^2$ dove il differenziale di Gateaux di U_θ é positivo, allora $\forall f \in \mathcal{L}^2$ ho che $d^+U_\theta(f)(v) \geq 0$;
- se $v \geq 0$, allora $d^+U_\theta(f)(v) \geq 0$;
- $\mathbb{E}[v] > 0$ se $v \geq 0$.

Possiamo procedere come segue:

$$\begin{aligned} d^+U_\theta(f)(v) \geq 0 \quad &\Rightarrow \quad \mathbb{E}[v] - \theta(\mathbb{E}[f \cdot v] - \mathbb{E}[f] \cdot \mathbb{E}[v]) \geq 0 \quad \Rightarrow \\ &\Rightarrow \quad \theta(\mathbb{E}[f \cdot v] - \mathbb{E}[f]\mathbb{E}[v]) \leq \mathbb{E}[v] \end{aligned}$$

Se dividiamo ogni termine per $\mathbb{E}[v]$, otteniamo

$$\mathbb{E}\left[f \cdot \frac{v}{\mathbb{E}[v]}\right] - \mathbb{E}[f] \leq \frac{1}{\theta}$$

$\forall v \geq 0, \mathbb{E}[v] > 0$.

Inoltre, ponendo $\mathbb{E}\left[f \cdot \frac{v}{\mathbb{E}[v]}\right] = f$, troviamo che

$$f - \mathbb{E}[f] \leq \frac{1}{\theta}$$

Quindi il differenziale di Gateaux è positivo quando $f - \mathbb{E}[f] \leq \frac{1}{\theta}$.

Essendo $\mathcal{G}_\theta$ il sottoinsieme di $\mathcal{L}^2(P)$ dove $d^+U_\theta(f)(v) \geq 0$, ricaviamo che:

$$\mathcal{G}_\theta = \left\{ f \in \mathcal{L}^2(P) : f - \mathbb{E}(f) \leq \frac{1}{\theta} \right\}$$

che è proprio quello che volevamo dimostrare. $\square$

Abbiamo visto come la funzione di utilità del modello media-varianza è data da:

$$U_\theta(f) = \mathbb{E}^P(f) - \frac{\theta}{2} Var^P(f)$$

e sappiamo che $U_\theta(f)$ è monotono $\forall f \in \mathcal{G}_\theta$, per come è stato definito il dominio di monotonicità.
Studiamo ora il funzionale *mmv* così definito:

$$V_\theta(f) = \min_Q \left\{ \mathbb{E}^Q(f) + \frac{1}{2\theta} C(Q||P) \right\} \qquad \forall f \in \mathcal{L}^2(P)$$

Teorema 2 *Il funzionale $V_\theta : \mathcal{L}^2(P) \to \mathrm{R}$ è il minimo funzionale monotono in $\mathcal{L}^2(P)$ che coincide con U_θ su $\mathcal{G}_\theta$, ovvero*

$$V_\theta(f) = \sup \{ U_\theta(g) : g \in \mathcal{G}_\theta, g \leq f \} \qquad \forall f \in \mathcal{L}^2(P)$$

inoltre

$$V_\theta(f) \geq U_\theta(f) \qquad \forall f \in \mathcal{L}^2(P).$$

Si noti che $V_\theta(f)$ possiede le seguenti proprietà:

i) il funzionale $V_\theta(f)$ è concavo e continuo;

ii) coincide con $U_\theta(f)$ in $\mathcal{G}_\theta$, dominio di monotonicità del funzionale *mv*;

iii) è il funzionale monotono minimale che estende il funzionale *mv* $U_\theta(f)$ fuori dal suo dominio di monotonicità $\mathcal{G}_\theta$;

iv) $V_\theta(f)$ è la migliore approssimazione monotona possibile per $U_\theta(f)$:

$$|V_\theta(f) - U_\theta(f)| \leq |V'_\theta(f) - U_\theta(f)| \qquad \forall f \in \mathcal{L}^2(P)$$

Per dimostrare il Teorema 2 ci serviremo dei seguenti tre lemma, del primo dei quali riportiamo solo l'enunciato (per la dimostrazione si veda [15]).

Lemma 3 *La funzione* $g_f : \mathbb{R} \to [0, +\infty)$ *è continua con*

$$F_f(z) = \lim_{\epsilon \to 0^+} \left[\frac{g_f(z+\epsilon) - g_f(z)}{\epsilon} \right] \qquad \forall z \in \mathbb{R}$$

dove F_f *è la derivata destra di* g_f, *e* $F_f(z)$ *è la derivata di* g_f *in ogni punto* z *dove* F_f *è continua. Inoltre se poniamo*

$$\zeta = essinf\,(f) = \sup\{\alpha \in \mathbb{R} : P(f < \alpha) = 0\}$$

allora

- g_f *è strettamente crescente nell'intervallo* $(\zeta, +\infty)$;
- $g_f \equiv 0$ *nell'intervallo* $(-\infty, \zeta]$;
- $\lim_{z \to \zeta^+} g_f(z) = 0^+$;
- $\lim_{z \to +\infty} g_f(z) = +\infty$.

Lemma 4 *Sia* $f \in \mathcal{L}^2(P) \setminus \mathcal{G}_\theta$ *e* $t \in \mathbb{R}$. *Allora:*

$$f \wedge t \in \mathcal{G}_\theta \Leftrightarrow g_f(t) \leq \frac{1}{\theta} \tag{4.1}$$

Dimostrazione

Possiamo definire $(f-k)^-$ come segue:

$$(f-k)^- = (k-f) \cdot \mathbb{I}_{(f \leq k)} = \begin{cases} -(f-k) & \text{se } f-k \leq 0 \\ 0 & \text{se } f-k > 0 \end{cases}$$

da cui otteniamo:

$$\begin{aligned} g_f(k) &= \textstyle\int (f-k)^- \mathrm{d}P \\ &= kP(f \leq k) - \textstyle\int f \cdot \mathbb{I}_{(f \leq k)} \mathrm{d}P = \\ &= \textstyle\int [k - (f \wedge k)] \mathrm{d}P = \\ &= \textstyle\int_{-\infty}^{k} P(f < t) \mathrm{d}t. \end{aligned}$$

Inoltre, se $f \wedge t \in \mathcal{G}_\theta$ con $f \notin \mathcal{G}_\theta$, allora $f \wedge t = t$, ovvero:

$$essup\,(f \wedge t) = t \qquad \forall f \notin \mathcal{G}_\theta$$

Partendo da

$$\begin{aligned}(f \wedge t) - \mathbb{E}[f \wedge t] &= f \cdot \mathbb{I}_{(f \leq t)} + t \cdot \mathbb{I}_{(f>t)} - t \cdot P(f > t) - \mathbb{E}[f \cdot \mathbb{I}_{(f\leq t)}] = \\ &= f \cdot \mathbb{I}_{(f \leq t)} + t \cdot \mathbb{I}_{(f>t)} - t(1 - P(f \leq t)) - \mathbb{E}[f \cdot \mathbb{I}_{(f\leq t)}] = \\ &= (f - t) \cdot \mathbb{I}_{(f\leq t)} + g_f(t)\end{aligned}$$

e considerando che $(f - t) \cdot \mathbb{I}_{(f\leq t)} < 0$, abbiamo che

$$g_f(t) \leq \frac{1}{\theta} \Rightarrow (f \wedge t) - \mathbb{E}[f \wedge t] \leq \frac{1}{\theta}$$

Ne deriva che, se $f \wedge t \in \mathcal{G}_\theta$, allora:

$$\begin{aligned}(f \wedge t) - \mathbb{E}[f \wedge t] \leq \tfrac{1}{\theta} \qquad &\Rightarrow essup\,(f \wedge t) - tP(f > t) - \mathbb{E}[f \cdot \mathbb{I}_{(f\leq t)}] \leq \tfrac{1}{\theta} \\ &\Rightarrow t - tP(f > t) - \mathbb{E}[f \cdot \mathbb{I}_{(f\leq t)}] \leq \tfrac{1}{\theta} \Rightarrow \\ &\Rightarrow tP(f \leq t) - \mathbb{E}[f \cdot \mathbb{I}_{(f\leq t)}] \leq \tfrac{1}{\theta} \Rightarrow \\ &\Rightarrow \textstyle\int (f - t)^- dP = g_f(t) \leq \tfrac{1}{\theta}\end{aligned}$$

Quindi $f \wedge t \in \mathcal{G}_\theta \Rightarrow g_f(t) \leq \frac{1}{\theta}$

Possiamo concludere che $g_f(t) \leq \frac{1}{\theta} \Leftrightarrow (f \wedge t) \in \mathcal{G}_\theta$ $\square$

Dai lemmi 3 e 4 deriva il seguente corollario.

Corollario 5 *Sia $f \notin \mathcal{G}_\theta$, allora:*

$$g_f^{-1}\left(\frac{1}{\theta}\right) = max\{t \in \mathbb{R} : f \wedge t \in \mathcal{G}_\theta\} = k \qquad \Rightarrow \qquad g_f(k) = \frac{1}{\theta}$$

Dimostrazione

Dal lemma 4 sappiamo che $f \wedge t \in \mathcal{G}_\theta \Rightarrow g_f(t) \leq \frac{1}{\theta}$ e con il lemma 3 abbiamo visto che $g_f(t)$ è una funzione strettamente crescente nell'intervallo $(\zeta, +\infty)$ con $\zeta = essinf\ (f)$. Quindi deve esistere un $\bar{t} \in \mathbb{R} : g_f(\bar{t}) = \frac{1}{\theta}$.
In particolare abbiamo che:

$$\bar{t} = \max\{t \in \mathbb{R} : f \wedge t \in \mathcal{G}_\theta\} = g_f^{-1}\left(\frac{1}{\theta}\right)$$

$\square$

Lemma 6 *Per ogni $\theta > 0$ abbiamo che:*

i. *$\mathcal{G}_\theta$ è convesso;*

ii. *U_θ è lineare in ogni sottospazio di $\mathcal{G}_\theta$;*

iii.
$$U^*(Y) = \begin{cases} -\frac{1}{2\theta}(\mathbb{E}[Y^2] - 1) & se\ \mathbb{E}[Y] = 1 \\ -\infty & altrimenti \end{cases} \tag{4.2}$$

con $Y \in \mathcal{L}^2(P)$;

iv. *U^* è strettamente concavo nell'insieme $\{Y \in \mathcal{L}^2(P) : \mathbb{E}[Y] = 1\}$.*

Dimostrazione

Se $\mathbb{E}[Y] = 1$, allora possiamo definire il funzionale $W : \mathcal{L}^2(P) \to \mathbb{R}$ dato da

$$\begin{aligned} W(f) &= \langle f, Y\rangle - U_\theta(f) = \\ &= \langle f, Y\rangle - \langle f, 1\rangle + \tfrac{1}{\theta}\langle f - \mathbb{E}[f], f - \mathbb{E}[f]\rangle \end{aligned}$$

Il funzionale W è convesso e il suo differenziale di Gateaux è :

$$\nabla W(f) = Y - 1 + \theta(f - \mathbb{E}[f])$$

Quindi

$$\begin{aligned} U_\theta^*(Y) &= \min_{f \in \mathcal{L}^2(P)} W(f) = \\ &= \langle -\tfrac{1}{\theta}Y, Y\rangle - \langle -\tfrac{\theta}{2}Y, 1\rangle + \tfrac{1}{\theta}Var\left[-\tfrac{1}{\theta}Y\right] = \\ &= -\tfrac{1}{\theta}\mathbb{E}[Y^2] + \tfrac{1}{\theta}\mathbb{E}[Y] + \tfrac{\theta}{2}\,\tfrac{1}{\theta^2}Var[Y] = \\ &= -\tfrac{1}{2\theta}\left(\mathbb{E}[Y^2] - 1\right) \end{aligned}$$

$\square$

Grazie ai tre lemma sopra presentati, possiamo passare alla dimostrazione del Teorema 2.

Dimostrazione Teorema 2

Per ogni $f \in \mathcal{G}_\theta$ si ha che

$$U_\theta(f) = \min_{Y:\mathbb{E}[Y]=1} \left\{ \mathbb{E}[fY] + \frac{1}{2\theta}(\mathbb{E}[Y^2]-1) \right\} = \min_{Q \in \mathcal{L}^2(P)} \left\{ \mathbb{E}^Q[f] + \frac{1}{2\theta} C(Q||P) \right\} = V_\theta(f)$$

Infatti, se $\mathbb{E}[Y]=1$, si ha che $\mathbb{E}[fY] = \mathbb{E}[Y]$ e che $C(Q||P) = \mathbb{E}[Y^2]-1$.

Proviamo ora a verificare se

$$V_\theta(f) = \sup \{U_\theta(g) : g \in \mathcal{G}_\theta, g \leq f\} \, \forall f \in \mathcal{L}^2(P)$$

Sia $U : E \to \mathbb{R}$ un funzionale concavo e continuo, con E spazio vettoriale ordinato. Sia E' il duale di E e sia E'_+ il cono di tutti i funzionali positivi, lineari e continui in E.
Sia il *superdifferenziale* di U rispetto a f definito come segue:

$$\partial U(f) = \left\{ f' \in E' : U(g) - U(f) \leq \langle g-f, f' \rangle \right\} \qquad \forall g \in E$$

Sia il funzionale $U^* : E' \to [-\infty, +\infty)$ la *coniugata di Fenchel* di U data da:

$$U^*(f') = \inf_{f \in E} \left\{ \langle f, f' \rangle - U(f) \right\} \qquad \forall f' \in E'$$

Allora $V(f)$ é il funzionale minimo monotono che domina U se

$$V(f) = \min_{f' \in E'_+} \left\{ \langle f, f' \rangle - U^*(f') \right\} = \sup\{U(g) : g \in E, g \leq f\} \qquad \forall f \in E$$

Sviluppando il superdifferenziale otteniamo:

$$\begin{aligned}
\partial U(f) &= \{f' \in E' : U(g) - U(f) \leq \langle g - f, f' \rangle\} = \\
&= \{f' \in E' : U(g) - U(f) \leq \langle g, f' \rangle - \langle f, f' \rangle\} \\
&= \{f' \in E' : [U(g) - \langle g, f' \rangle] - [U(f) - \langle f, f' \rangle] \leq 0\} = \\
&= \{f' \in E' : [U(f) - \langle f f' \rangle] - [U(g) - \langle g, f' \rangle] \geq 0\} = \\
&= \{f' \in E' : [\langle f, f' \rangle - U(f)] - [\langle g, f' \rangle - U(g)] \leq 0\} = \\
&= \{f' \in E' : \langle f, f' \rangle - U(f) \leq \langle g, f' \rangle - U(g)\}
\end{aligned}$$

Quindi, se partiamo da:

$$\begin{aligned}
V(f) &= \min_{f' \in E'_+} \{\langle f, f' \rangle - U^*(f')\} = \\
&= \min_{f' \in E'_+} \{\langle f, f' \rangle - inf_{f \in E}(\langle f, f' \rangle - U(f))\} = \\
&= \min_{f' \in E'_+} sup_{f \in E} \{\langle f, f' \rangle - \langle f, f' \rangle + U(f)\} = \\
&= sup_{f \in E}\{U(f)\}
\end{aligned}$$

e consideriamo che dal superdifferenziale abbiamo ottenuto che

$$\langle f, f' \rangle - U(f) \leq \langle g, f' \rangle - U(g)$$

arriviamo alla seguente conclusione:

$$V(f) = \min_{f' \in E'_+} \sup_{f \in E} \{\langle f, f' \rangle - \langle f, f' \rangle + U(f)\} \geq \min_{f' \in E'_+} \sup_{g \in E} \{\langle g, f' \rangle - \langle g, f' \rangle + U(g)\}$$

da cui

$$V(f) = \sup_{f \in E}\{U(f)\} \geq \sup_{g \in E}\{U(g)\}.$$

Quindi $V(f) = \sup\{U(g) : g \leq f\} \qquad \forall f \in E, \forall g \in E.$ $\square$

Il seguente teorema illustra la natura di V_θ. Un agente che presenta preferenze di tipo monotono (*mmv*), può essere considerato come se usasse preferenze *mv* in un prospetto di valutazione fuori dal dominio di monotonicità $\mathcal{G}_\theta$. Infatti se siamo all'interno del dominio di monotonicità, si verifica che le preferenze monotone coincidono con le preferenze non monotone, mentre fuori da $\mathcal{G}_\theta$ le preferenze relative al funzionale V_θ sono definte come da Teorema 7.

Teorema 7

$$V_\theta(f) = \begin{cases} U_\theta(f) & \textit{se } f \in \mathcal{G}_\theta \\ U_\theta(f \wedge k) & \textit{altrimenti} \end{cases}$$

con $k = \max\{t \in \mathbb{R} : f \wedge t \in \mathcal{G}_\theta\}$.

Dimostrazione

Vogliamo dimostrare che $V_\theta(f) = U_\theta(f \wedge k)$ se $f \notin \mathcal{G}_\theta$. Per farlo, definiamo prima le seguenti quantità:

- $Y^* = \theta(k-f)^- \cdot \mathbb{I}_{(f \le k)}$
- $(Y^*)^2 = \theta^2(k-f)^2 \cdot \mathbb{I}_{(f \le k)} = \theta^2[k^2 \cdot \mathbb{I}_{(f \le k)} + f^2 \cdot \mathbb{I}_{(f \le k)} - 2kf \cdot \mathbb{I}_{(f \le k)}]$
- $\mathbb{E}[(Y^*)^2] = \theta^2[k^2 P(f \le k) + \int f^2 \cdot \mathbb{I}_{(f \le k)} \mathrm{d}P - 2k \int f \cdot \mathbb{I}_{(f \le k)} \mathrm{d}P]$
- $\mathbb{E}[fY^*] = \mathbb{E}[\theta kf \cdot \mathbb{I}_{(f \le k)} - \theta f^2 \cdot \mathbb{I}_{(f \le k)}] = \theta k \int f \cdot \mathbb{I}_{(f \le k)} \mathrm{d}P - \theta \int f^2 \cdot \mathbb{I}_{(f \le k)} \mathrm{d}P$

Poichè $\mathbb{E}^Q(f) = \mathbb{E}[fY^*]$, possiamo scrivere $V_\theta(f)$ nel seguente modo:

$$V_\theta(f) = \mathbb{E}[fY^*] + \frac{1}{2\theta}\mathbb{E}[(Y^*)^2] - \frac{1}{2\theta}$$

Quindi:

$V_\theta(f) = \mathbb{E}[fY^*] + \frac{1}{2\theta}\mathbb{E}[(Y^*)^2] - \frac{1}{2\theta} =$

$$= \Big[\theta k \int f \cdot \mathbb{I}_{(f \le k)} \mathrm{d}P - \theta \int f^2 \cdot \mathbb{I}_{(f \le k)} \mathrm{d}P\Big] + \tfrac{1}{2\theta}\Big\{\theta^2\Big[k^2 P(f \le k) +$$

$$+\int f^2\cdot \mathbb{I}_{(f\le k)}\mathrm{d}P-2k\int f\cdot \mathbb{I}_{(f\le k)}\mathrm{d}P\Big]\Big\}-\tfrac{1}{2\theta}=$$

$$=\tfrac{\theta}{2}\Big\{2\Big[k\int f\cdot \mathbb{I}_{(f\le k)}\mathrm{d}P-\int f^2\cdot \mathbb{I}_{(f\le k)}\mathrm{d}P\Big]+k^2P(f\le k)+$$
$$+\int f^2\cdot \mathbb{I}_{(f\le k)}\mathrm{d}P-2k\int f\cdot \mathbb{I}_{(f\le k)}\mathrm{d}P\Big\}-\tfrac{1}{2\theta}=$$

$$=\tfrac{\theta}{2}\Big[k^2P(f\le k)-\int f^2\cdot \mathbb{I}_{(f\le k)}\mathrm{d}P\Big]-\tfrac{1}{2\theta}$$

da cui

$$V_\theta(f)=\frac{\theta}{2}\left[k^2P(f\le k)-\int f^2\cdot \mathbb{I}_{(f\le k)}\mathrm{d}P\right]-\frac{1}{2\theta} \tag{4.3}$$

Inoltre

$$U(f\wedge k)=\mathbb{E}[f\wedge k]-\frac{\theta}{2}Var[f\wedge k]$$

Proviamo ad esaminare ogni singola quantità:

$$\begin{aligned}\mathbb{E}[f\wedge k] &= \mathbb{E}[f\cdot \mathbb{I}_{(f\le k)}]+kP(f>k)=\\ &=\mathbb{E}[f\cdot \mathbb{I}_{(f\le k)}]+k[1-P(f\le k)]=\\ &=\mathbb{E}[f\cdot \mathbb{I}_{(f\le k)}]-kP(f\le k)\ +k=\\ &= \qquad -g_f(k) \qquad +k=\\ &= \qquad -\tfrac{1}{\theta} \qquad +k\end{aligned}$$

$$\begin{aligned}Var[f\wedge k] &= \mathbb{E}[(f\wedge k)^2]-[\mathbb{E}(f\wedge k)]^2=\\ &=\mathbb{E}\Big[\Big(f\cdot \mathbb{I}_{(f\le k)}+kP(f>k)\Big)^2\Big]-\Big[k-\tfrac{1}{\theta}\Big]^2=\\ &=\mathbb{E}\Big[f^2\cdot \mathbb{I}_{(f\le k)}+k[1-P(f\le k)]-2kf\cdot \mathbb{I}_{(f\le k)}\Big]-\Big(k^2+\tfrac{1}{\theta^2}-\tfrac{2k}{\theta}\Big)=\end{aligned}$$

$$= \int f^2 \cdot \mathbb{I}_{(f\leq k)} \mathrm{d}P + k^2 - k^2 P(f \leq k) - k^2 - \tfrac{1}{\theta^2} + \tfrac{2k}{\theta}$$

Quindi:

$$U(f \wedge k) = \mathbb{E}[f \wedge k] - \tfrac{\theta}{2} Var[f \wedge k] =$$

$$= \left(k - \tfrac{1}{\theta}\right) - \tfrac{\theta}{2}\left[\int f^2 \cdot \mathbb{I}_{(f\leq k)} dP - k^2 P(f \leq k) - \tfrac{1}{\theta^2} + \tfrac{2k}{\theta}\right] =$$

$$= \tfrac{1}{2\theta}\left[2\theta^2 k - 2\theta - \theta^3 \int f^2 - \tfrac{1}{\theta^2} - \tfrac{2k}{\theta} dP + \theta^3 k^2 P(f \leq k) + \theta - 2\theta^2 k\right] =$$

$$= \tfrac{1}{2\theta^2}\left[-\theta - \theta^3 \int f^2 \cdot \mathbb{I}_{(f\leq k)} \mathrm{d}P + \theta^3 k^2 P(f \leq k)\right] =$$

$$= -\tfrac{1}{2\theta} - \tfrac{\theta}{2} \int f^2 \cdot \mathbb{I}_{(f\leq k)} \mathrm{d}P + \tfrac{\theta}{2} k^2 P(f \leq k)$$

da cui

$$U(f \wedge k) = -\frac{1}{2\theta} - \frac{\theta}{2} \int f^2 \cdot \mathbb{I}_{(f\leq k)} \mathrm{d}P + \frac{\theta}{2} k^2 P(f \leq k) \tag{4.4}$$

Si può notare come la 4.4 corrisponde esattamente alla 4.3, quindi

$$U(f \wedge k) = V(f)$$

che è ciò che volevamo dimostrare. □

4.2 Selezione del portafoglio

Consideriamo ora l'allocazione di un portafoglio con $n+1$ titoli in un singolo periodo. Supponiamo che i primi n titoli siano rischiosi e che solo il $n+1^{mo}$ sia privo di rischio. Utilizziamo la seguente notazione:

- X_i con $i = 1, ..., n$ sono i primi n portafogli rischiosi;
- R indica il $n+1^{mo}$ portafoglio (privo di rischio);
- α è la precentuale di ricchezza investita in titoli rischiosi.

La ricchezza di fine periodo sarà data da:

$$W = R + \alpha(X - e'R)$$

Se utilizziamo preferenze $\succeq^{mmv}$ di tipo media-varianza monotono, il portafoglio sarà definito tramite il seguente funzionale:

$$V_\theta(W) = \min_{Q \in \Delta^2(P)} \left\{ \mathbb{E}^Q[W] + \frac{1}{2\theta} C(P||Q) \right\}$$

e il problema di portafoglio potrà essere scritto come segue:

$$\max_{\alpha \in R^n} \min_{Q \in \Delta^2(P)} \left\{ \mathbb{E}^Q[W] + \frac{1}{2\theta} C(P||Q) \right\}$$

4.3 Il portafoglio ottimale

Il portafoglio ottimale viene definito dal teorema 8.

Teorema 8 *Il vettore α^* è soluzione del problema di selezione del portafoglio se e solo se esiste $k^* \in \mathbb{R}$ tale che la coppia (α^*, k^*) soddisfa il seguente sistema di equazioni:*

$$\begin{cases} \theta P(W \leq k) Var^P[X|W \leq k]\alpha = \mathbb{E}^P[X - e'R|W \leq k] \\ \mathbb{E}^P[(W-k)^-] = \frac{1}{\theta} \end{cases} \tag{4.5}$$

Il portafoglio ottimale α^* viene determinato insieme a k^* risolvendo il sistema (4.5) in $n+1$ equazioni e $n+1$ incognite.
Con l'applicazione di un modello media-varianza avremo:

$$\alpha^*_{mv} = \frac{1}{\theta}(Var^P[X])^{-1}\mathbb{E}^P[X - e'R] \tag{4.6}$$

mentre con l'applicazione di un modello media-varianza monotono otteniamo:

$$\alpha^*_{mmv} = \frac{1}{\theta P(W \leq k)} Var^P[X|W \leq k^*]^{-1}\mathbb{E}^P[X - e'R|W \leq k^*] \tag{4.7}$$

Abbiamo ignorato la parte di distribuzione dove $W > k$ perchè se introduciamo dei rendimenti molto alti incrementiamo molto la varianza. Questo incremento viene controllato con l'utilizzo di un modello monotono nel quale si pone $P(W > k) = 0$. Vediamo ora che relazione c'è tra il portafoglio ottimale ottenuto con un modello media-varianza e il portafoglio ottimale ottenuto con un modello media-varianza monotono.

Proposizione 9 *Sia α^*_{mv} soluzione del problema di ottimo risolto utilizzando un modello media-varianza e sia α^*_{mmv} soluzione del problema di ottimo con l'applicazione di un modello media-varianza monotono. Allora si ha che:*

$$\alpha^*_{mmv} \geq \alpha^*_{mv} \geq 0$$

$$\alpha^*_{mmv} \leq \alpha^*_{mv} \leq 0$$

Inoltre, se fosse $P(W > k) > 0$ allora:

$$\alpha^*_{mmv} > \alpha^*_{mv} \qquad se \qquad \alpha^*_{mmv} > 0$$

$$\alpha^*_{mmv} < \alpha^*_{mv} \qquad se \qquad \alpha^*_{mmv} < 0$$

Un agente che investe con un modello media-varianza non monotono ha un portafoglio con maggiori perdite rispetto a un agente che investe secondo un modello media-varianza monotono perchè compra di più e vende di più, quindi un investitore *mv* rischia di più rispetto a un investitore *mmv*.
Questo perchè un portafoglio allocato secondo un modello media-varianza di tipo non monotono dà maggiori possibilità di guadagno rispetto allo stesso portafoglio allocato secondo l'utilizzo di un modello media-varianza monotono, a scapito, però, del rischio di subire, insieme a maggiori rendimenti, anche maggiori perdite. Proviamo a spiegare meglio quest'ultimo concetto riportando alcuni esempi.

4.4 Alcuni esempi di selezione del portafoglio

Consideriamo 5 stati di natura $s_1, s_2, \ldots, s_5$ ognuno dei quali è realizzabile con probabilità $P(s_i)$ con $i = 1, 2, \ldots, 5$ e supponiamo di avere un titolo rischioso X e un titolo non rischioso R che poniamo essere pari a 1 per ogni stato di natura. Per

semplicità, fissiamo il parametro θ pari a 10.
Il nostro obiettivo è trovare W_{mv} e W_{mmv} e, poichè $W = R + \alpha(X - e'R)$, abbiamo bisogno di calcolare α^*_{mv} e α^*_{mmv}. Per farlo dobbiamo definire prima le seguenti quantità:

$$\mathbb{E}^P[X] = \sum_{i=1}^{5} X_i \, P(s_i)$$

$$\mathbb{E}^P[X|A] = \sum_{i=1}^{5} X_i \, P(s_i|A)$$

$$\mathbb{E}^P[X - e'R] = \sum_{i=1}^{5} (X_i - 1) \, P(s_i)$$

$$\mathbb{E}^P[X - e'R|A] = \sum_{i=1}^{5} (X_i - 1) \, P(s_i|A)$$

$$Var^P[X] = \sum_{i=1}^{5} (X_i - \mathbb{E}^P[X])^2 \, P(s_i)$$

$$Var^P[X|A] = \sum_{i=1}^{5} (X_i - \mathbb{E}^P[X|A])^2 \, P(s_i|A)$$

dove con A indichiamo l'evento $W \leq k$. In particolare, sappiamo che

$$P(W \leq k) = P(A) = 1$$

perchè nel modello monotono la probabilità dell'evento $W > k$ è nulla.
Il valore di k lo ricaviamo dall'equazione 4.5, ovvero:

$$\mathbb{E}^P[(W - k)^-] = \frac{1}{\theta}$$

$$\begin{aligned}
\mathbb{E}^P[(W - k)^-] &= (k - W_1)P(s_1) + (k - W_2)P(s_2) + ... + (k - W_5)P(s_5) = \\
&= k[P(s_1) + P(s_2) + ... + P(s_5)] + [W_1P(s_1)W_2P(s_2) + ... + W_5P(s_5)] = \\
&= k - \mathbb{E}^P[W]
\end{aligned}$$

da cui ottengo

$$k = \frac{1}{\theta} + \mathbb{E}^P[W]$$

4.4.1 Esempio 1

Iniziamo considerando il caso in cui $P(s_i) = P(s_i|A)$. Abbiamo i seguenti dati di partenza:

Stati di natura	$P(s_i)$	$P(s_i\|A)$	R	X
s_1	0.1	0.1	1	0.97
s_2	0.2	0.2	1	0.99
s_3	0.4	0.4	1	1.01
s_4	0.2	0.2	1	1.03
s_5	0.1	0.1	1	1.05

Calcoliamo le quantità che ci servono per arrivare ad α^*_{mv} e α^*_{mmv}:

$\mathbb{E}^P[X] = (0.97)0.1 + (0.99)0.2 + (1.01)0.4 + (1.03)0.2 + (1.05)0.1 = 1.01$

$\mathbb{E}^P[X - e'R] = (0.97 - 1)0.1 + (0.99 - 1)0.2 + \ldots + (1.05 - 1)0.1 = 0.01$

$$\begin{aligned} Var^P[X] = {} & (0.97 - 1.01)^2 0.1 + (0.99 - 1.01)^2 0.2 + (1.01 - 1.01)^2 0.4 + \\ & + (1.03 - 1.01)^2 0.2 + (1.05 - 1.01)^2 0.1 = 0.00048 \end{aligned}$$

Poichè $P(s_i) = P(s_i|A)$, avremo che

$$Var^P[X] = Var^P[X|A]$$

$$\mathbb{E}^P[X] = \mathbb{E}^P[X|A]$$

$$\mathbb{E}^P[X - e'R] = \mathbb{E}^P[X - e'R|A]$$

Ora possiamo ricavarci α^*_{mv} e α^*_{mmv} che ci serviranno per trovare W_{mv} e W_{mmv}:

$\alpha^*_{mv} = \frac{1}{\theta}(Var^P[X])^{-1}\mathbb{E}^P[X - e'R] = \frac{1}{10}(0.00048)^{-1}0.01 = 2.0833$

$\alpha^*_{mmv} = \frac{1}{\theta P(A)}(Var^P[X|A])^{-1}\mathbb{E}^P[X - e'R|A] = \frac{1}{10}(0.00048)^{-1}0.01 = 2.0833$

quindi il vettore W sarà dato da:

$$W_{mv} = W_{mmv} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + 2.083 \begin{pmatrix} 0.97\text{-}1 \\ 0.99\text{-}1 \\ 1.01\text{-}1 \\ 1.03\text{-}1 \\ 1.05\text{-}1 \end{pmatrix} = \begin{pmatrix} 0.9375 \\ 0.9792 \\ 1.0208 \\ 1.0625 \\ 1.1041 \end{pmatrix}$$

Calcoliamo anche k per verificare che effettivamente ogni W sia minore di k.

$\mathbb{E}^P[W] = (0.9375)0.1 + (0.9792)0.2 + (1.0208)0.4 + (1.0625)0.2 + (1.1041)0.1 = 1.0208$

$k = (1/10) + \mathbb{E}^P[w] = 0.1 + 1.0208 = 1.1208$

In questo caso ogni valore di W è minore di 1.1208, quindi

$$P(W_i \leq k) = 1 \qquad \forall i = 1, 2, ..., 5$$

Possiamo riassumere i risultati in questa tabella:

Stati di natura	$P(s_i)$	$P(s_i \vert A)$	R	X	W_{mv}	W_{mmv}
s_1	0.1	0.1	1	0.97	0.9375	0.9375
s_2	0.2	0.2	1	0.99	0.9792	0.9792
s_3	0.4	0.4	1	1.01	1.0208	1.0208
s_4	0.2	0.2	1	1.03	1.0625	1.0625
s_5	0.1	0.1	1	1.05	1.1041	1.1041
$\alpha_{mv}* = 2.0833$						
$\alpha_{mmv}* = 2.0833$						
$k^* = 1.1208$						

4.4.2 Esempio 2

Consideriamo ora il caso in cui $P(s_i) \neq P(s_i|A)$ supponendo di partire dai seguenti dati:

Stati di natura	$P(s_i)$	R	X
s_1	0.1	1	0.97
s_2	0.2	1	0.99
s_3	0.4	1	1.01
s_4	0.2	1	1.03
s_5	0.1	1	1.10

Come nell'esempio 1, ci calcoliamo tutte le quantità che ci servono:

$$\mathbb{E}^P[X] = (0.97)0.1 + (0.99)0.2 + (1.01)0.4 + (1.03)0.2 + (1.10)0.1 = 1.015$$

$$\mathbb{E}^P[X - e'R] = (0.97 - 1)0.1 + (0.99 - 1)0.2 + \ldots + (1.10 - 1)0.1 = 0.015$$

$$\begin{aligned} Var^P[X] &= (0.97 - 1.015)^2 0.1 + (0.99 - 1.015)^2 0.2 + (1.01 - 1.015)^2 0.4 + \\ &+ (1.03 - 1.015)^2 0.2 + (1.10 - 1.015)^2 0.1 = 0.001105 \end{aligned}$$

Ora possiamo ricavare α^*_{mv}:

$$\alpha^*_{mv} = \frac{1}{\theta}(Var^P[X])^{-1}E^P[X - e'R] = \frac{1}{10}(0.001105)^{-1}0.015 = 1.3575$$

quindi il vettore W_{mv} sarà dato da:

$$W_{mv} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + 1.3575 \begin{pmatrix} 0.97\text{-}1 \\ 0.99\text{-}1 \\ 1.01\text{-}1 \\ 1.03\text{-}1 \\ 1.10\text{-}1 \end{pmatrix} = \begin{pmatrix} 0.9593 \\ 0.9864 \\ 1.0136 \\ 1.0407 \\ 1.1357 \end{pmatrix}$$

Calcoliamo il valore di k:

$$\mathbb{E}^P[W] = (0.9583)0.1 + (0.9864)0.2 + (1.0136)0.4 + (1.0407)0.2 + (1.1357)0.1 = 1.1204$$

$k = (1/10) + \mathbb{E}^P[w] = 0.1 + 1.0204 = 1.1204$

In questo caso $W_5 = 1.1657 > k$, quindi assegneremo probabilità zero a $P(s_5|W_5 \leq k)$ e di conseguenza anche le probabilità condizionate assegnate agli altri stati subiranno delle variazioni.
Se $P(s_5) = 0$, allora $\sum_{i=1} = 5 \cdot P(s_i) = 0.1 + 0.2 + 0.4 + 0.2 = 0.9$. Le probabilità condizionate assumarenno i seguenti valori:

$$
\begin{gathered}
P(s_1|A) = P(s_1)/0.9 = 0.1/0.9 = 0.1111 \\
P(s_2|A) = P(s_2)/0.9 = 0.2/0.9 = 0.2222 \\
P(s_3|A) = P(s_3)/0.9 = 0.4/0.9 = 0.4444 \\
P(s_4|A) = P(s_4)/0.9 = 0.2/0.9 = 0.2222 \\
P(s_5|A) = P(s_5)/0.9 = 0/0.9 = 0
\end{gathered}
$$

Usando le probabilità condizionate, ricaviamo la media condizionata e la varianza condizionata del titolo rischioso:

$\mathbb{E}^P[X|A] = (0.97)0.1111 + (0.99)0.2222 + (1.01)0.4444 + (1.03)0.2222 + (1.10)0 = 1.005$

$\mathbb{E}^P[X - e'R|A] = (0.97 - 1)0.1111 + (0.99 - 1)0.2222 + \ldots + (1.10 - 1)0 = 0.00556$

$$
\begin{aligned}
Var^P[X|A] &= (0.97 - 1.005)^2 0.1111 + (0.99 - 1.005)^2 0.2222 + \\
&+ (1.01 - 1.005)^2 0.4444 + (1.03 - 1.005)^2 0.2222 + (1.10 - 1.005)^2 0 = \\
&= 0.000335
\end{aligned}
$$

Calcoliamo ora il valore di α^*_{mmv} e quindi del vettore W_{mmv}:

$\alpha^*_{mmv} = \frac{1}{\theta P(A)}(Var^P[X|A])^{-1} E^P[X - e'R|A] = \frac{1}{10}(0.000335)^{-1} 0.00556 = 1.6597$

$$
W_{mmv} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + 1.6597 \begin{pmatrix} 0.97\text{-}1 \\ 0.99\text{-}1 \\ 1.01\text{-}1 \\ 1.03\text{-}1 \\ 1.10\text{-}1 \end{pmatrix} = \begin{pmatrix} 0.9503 \\ 0.9834 \\ 1.0166 \\ 1.0497 \\ 1.1657 \end{pmatrix}
$$

Possiamo riassumere i risultati in questa tabella:

Stati di natura	$P(s_i)$	$P(s_i\|A)$	R	X	W_{mv}	W_{mmv}
s_1	0.1	0.1111	1	0.97	0.9583	0.9503
s_2	0.2	0.2222	1	0.99	0.9864	0.9834
s_3	0.4	0.4444	1	1.01	1.0136	1.0166
s_4	0.2	0.2222	1	1.03	1.0407	1.0497
s_5	0.1	0	1	1.10	1.1357	1.1657
$\alpha^*_{mv} = 1.3575$						
$\alpha^*_{mmv} = 1.6597$						
$k = 1.1204$						

4.4.3 Esempio 3

Proviamo a osservare come cambiano α^*_{mv} e α^*_{mmv} al variare di X_5, cioè al variare del valore del rendimento del titolo rischioso al verificarsi dello stato s_5.
Partiamo dai seguenti dati:

Stati di natura	$P(s_i)$	R	X
s_1	0.1	1	0.97
s_2	0.2	1	0.99
s_3	0.4	1	1.01
s_4	0.2	1	1.03
s_5	0.1	1	X_5

Se assegnamo a X_5 valori sempre maggiori, partendo da $X_5 = 0$ fino a $X_5 = 3$, seguendo lo stesso procedimento adottato negli esempi precedenti otteniamo i valori di W_5^{mv} con relativo valore di k, i valori di W_5^{mmv}, di α^*_{mv} e di α^*_{mmv} come riportati di seguito:

X_5	W_5^{mv}	W_5^{mmv}	k	α^*_{mv}	α^*_{mmv}
$0,00$	$1,1040$	$1,1040$	$1,1099$	-0.1040	-0.1040
$0,30$	$1,1009$	$1,1009$	$1,1094$	-0.1441	-0.1441
$0,70$	$1,0862$	$1,0862$	$1,1072$	-0.2878	-0.2878
$0,90$	$1,0383$	$1,0383$	$1,1012$	-0.3831	-0.3831
$1,00$	$1,0000$	$1,0000$	$1,1082$	1.6393	1.6393
$1,05$	$1,1042$	$1,1042$	$1,1208$	2.0833	2.0833
$1,10$	$1,1357$	$1,1654$	$1,1204$	1.3575	1.6544
$1,20$	$1,1350$	$1,3309$	$1,1169$	0.6748	1.6544
$1,50$	$1,1233$	$1,8272$	$1,1135$	0.2466	1.6544
$2,00$	$1,1176$	$2,6544$	$1,1123$	0.1176	1.6544
$3,00$	$1,1144$	$4,3088$	$1,1117$	0.0572	1.6544

Fino a valori di X_5 pari a 1.05, tutti i valori che assume W_5^{mv} sono minori di k. Quindi per $X_5 \leq 1.05$ accade che la probabilità condizionata che si verifichi l'i^{mo} stato di natura corrisponde alla probabalità semplice, cioè per valori di X_5 fino a 1.05, si ha che $P(s_5) = P(s_5|A)$. Per $X_5 > 1.05$, invece, i valori assunti dalla variabile W_5^{mv} sono maggiori dei valori di k, quindi per $X_5 > 1.05$ si ha che $P(s_5) \neq P(s_5|A)$. Questo comporta che per ogni $X_5 \leq 1.05$, α^*_{mv} coincide con α^*_{mmv}, mentre per ogni $X_5 > 1.05$ abbiamo che $\alpha^*_{mv} < \alpha^*_{mv}$, come si vede dalla figura 4.1.

Figura 4.1: alpha

Per quanto riguarda la media e la varianza, condizionate e non, i risultati che otteniamo sono riportati in questa tabella:

X_5	$E[X]$	$E[X\|A]$	$Var[X]$	$Var[X\|A]$
0.00	0.905	0.905	0.091305	0.091305
0.30	0.935	0.935	0.045105	0.045105
0.70	0.975	0.975	0.008705	0.008705
0.90	0.995	0.995	0.001305	0.001305
1.00	1.005	1.005	0.000305	0.000305
1.05	1.010	1.010	0.000480	0.000480
1.10	1.015	1.005	0.001105	0.000335
1.20	1.025	1.005	0.003705	0.000335
1.50	1.055	1.005	0.022305	0.000335
2.00	1.105	1.005	0.089305	0.000335
3.00	1.205	1.005	0.358305	0.000335

La media e la varianza del vettore X variano per valori di X_5 maggiori di 1.05, come si può notare anche osserando le figure 4.2 e 4.3 rispettivamente. In particolare, la media e la media condizionata coincidono fino a $X_5 \leq 1.05$, mentre per valori di $X_5 > 1.05$ accade che la media semplice assume valori sempre maggiori rispetto alla media condizionata.

Figura 4.2: media

Figura 4.3: varianza

E' interessante notare come i valori di α^*_{mmv}, $\mathbb{E}[X|A]$ e $Var[X|A]$ restano costanti se assegnamo a X_5 valori superiori a 1.05. Questo succede perchè $P(s_5|W \leq k) = 0$ per ogni $X_5 > 1.05$, quindi qualsiasi valore maggiore di 1.05 che assegnamo a X_5 non andrà a influenzare in nessun modo la media e la varianza del modello media-varianza monotono. Riassumiamo quanto detto fin'ora nella seguente tabella:

$X_5 \leq 1.05$	$X_5 > 1.05$
$\alpha^*_{mv} = \alpha^*_{mmv}$	$\alpha^*_{mv} < \alpha^*_{mmv}$
$\mathbb{E}[X] = \mathbb{E}[X\|A]$	$\mathbb{E}[X] > \mathbb{E}[X\|A]$
$Var[X] = Var[X\|A]$	$Var[X] > Var[X\|A]$

Possiamo concludere che un portafoglio allocato secondo un modello non monotono porta guadagni maggiori rispetto a un portafoglio allocato secondo un modello di tipo monotono, ma comporta anche una maggiore varianza nei rendimenti e quindi il rischio di subire perdite più consistenti.

Capitolo 5

Esempio: Google Inc.SpA. e Fastweb SpA

5.1 Obiettivo

Nel seguente esempio applichiamo il modello media-varianza monotono a un portafoglio composto da tre titoli di cui uno non rischioso e due rischiosi. Per osservare la variazione percentuale del valore dei titoli, prendiamo due titoli reali: Google Inc.SpA e Fastweb SpA.
Per prima cosa raccogliamo i dati e calcoliamo la variazione percentuale del valore dei titoli. Poi li raggruppiamo in gruppi che siano omogenei al loro interno ed eterogenei tra di loro. Andremo in fine ad applicare il modello ai gruppi così ottenuti.

5.2 Raccolta dei dati

Prendiamo la variazione percentuale del valore giornaliero delle azioni di Google Inc.SpA e di Fastweb SpA in un periodo di tempo compreso tra il 19 agosto 2004 e il 14 marzo 2006 (394 giorni).
Creiamo una matrice che chiamiamo *dati*, con due colonne (Google e Fastweb) e 394 righe riferite ai valori che assumono i due titoli nel periodo di riferimento.

Abbiamo un totale di 394 dati, che possono essere così riassunti:

	Google	*Fastweb*
min	$0,8663$	$0,9102$
max	$1,0926$	$1,0745$
media	0.9971	1.0030
varianza	0.000651	0.000356

Partendo dal valore massimo e minimo che i due titoli assumono nel periodo di riferimento, decidiamo di suddividere le azioni google in quattro gruppi (0,86-0,92,0,92-0,98, 0,98-1,04,1,04-1,10) e le azioni Fastweb in tre gruppi (0,90-0,96,0,96-1,02,1,02-1,08) ottenendo così una ripartizione del vettore finale in 12 sezioni.
Riportiamo i comandi utilizzati col programma R per ottenere la suddivisione dl vettore dei titoli nei 12 gruppi:

```
A <- dati[dati$ GOOGLE <0.92,]
A1 <- A[A$ FASTWEB <0.96,]
A2 <- A[A$FASTWEB>=0.96 & A$FASTWEB<1.02,]
A3 <- A[A$FASTWEB>=1.02,]

B <- dati[dati$GOOGLE>=0.92 & dati$GOOGLE<0.98,]
B1 <- B[B$ FASTWEB <0.96,]
B2 <- B[B$FASTWEB>=0.96 & B$FASTWEB<1.02,]
B3 <- B[B$FASTWEB>=1.02,]

C <- dati[dati$GOOGLE>=0.98 & dati$GOOGLE<1.04,]
C1 <- C[C$ FASTWEB <0.96,]
C2 <- C[C$FASTWEB>=0.96 & C$FASTWEB<1.02,]
C3 <- C[C$FASTWEB>=1.02,]

D <- dati[dati$GOOGLE>=1.04,]
D1 <- D[D$ FASTWEB <0.96,]
D2 <- D[D$FASTWEB>=0.96 & D$FASTWEB<1.02,]
D3 <- D[D$FASTWEB>=1.02,]
```

Come valore di riferimento per ciascun gruppo prendiamo la media aritmetica del gruppo stesso:

gruppo	Google	Fastweb
$A1$	$-$	$-$
$A2$	$0,8918$	$0,9931$
$A3$	$0,8975$	$1,0273$
$B1$	$-$	$-$
$B2$	$0,9666$	$0,9976$
$B3$	$0,9659$	$1,0512$
$C1$	$0,9949$	$0,9458$
$C2$	$1,0024$	$0,9996$
$C3$	$1,0039$	$1,0302$
$D1$	$1,0715$	$0,9599$
$D2$	$1,0549$	$1,0008$
$D3$	$1,0649$	$1,0275$

Ci interessa conoscere la probabilità con cui ciascuna coppia di valori (Google,Fastweb) può appartenere ad un determinato gruppo. Calcoliamo quindi il numero di coppie contenute in ogni gruppo per poter poi ricavare la probabilità di appartenenza:

gruppo	n.coppie		Probabilità
$A1$	0	$\Rightarrow$	$0/394 = 0$
$A2$	3	$\Rightarrow$	$3/394 = 0,0076$
$A3$	1	$\Rightarrow$	$1/394 = 0,0025$
$B1$	0	$\Rightarrow$	$0/394 = 0$
$B2$	62	$\Rightarrow$	$62/394 = 0,1574$
$B3$	10	$\Rightarrow$	$10/394 = 0,0254$
$C1$	6	$\Rightarrow$	$6/394 = 0,0152$
$C2$	250	$\Rightarrow$	$250/394 = 0,6345$
$C3$	46	$\Rightarrow$	$46/394 = 0,1168$
$D1$	1	$\Rightarrow$	$1/394 = 0,0025$
$D2$	13	$\Rightarrow$	$13/394 = 0,033$
$D3$	2	$\Rightarrow$	$2/394 = 0,0051$

Per comodità, nella nostra analisi non consideriamo i gruppi che non contengono nessuna coppia, ovvero i gruppi $A1$ e $B1$.

Riportiamo di seguito i valori che ci interessano:

s_i	$P(s_i)$	Google	Fastweb
1	$0,0076$	$0,8918$	$0,9931$
2	$0,0025$	$0,8975$	$1,0273$
3	$0,1574$	$0,9666$	$0,9976$
4	$0,0254$	$0,9659$	$1,0512$
5	$0,0152$	$0,9949$	$0,9458$
6	$0,6345$	$1,0024$	$0,9996$
7	$0,1168$	$1,0039$	$1,0302$
8	$0,0025$	$1,0715$	$0,9599$
9	$0,0330$	$1,0549$	$1,0008$
10	$0,0051$	$1,0649$	$1,0275$

Con s_i indichiamo l'i^{mo} stato di natura, ovvero l'i^{mo} gruppo, mentre con $P(s_i)$ indichiamo la probabilità che una coppia di valori possa appartenere all'i^{mo} gruppo.

5.3 Elaborazione dei dati

Ipotizziamo di avere un portafoglio con un titolo non rischioso R e due titoli rischiosi X_1 e X_2. Scegliamo due titoli reali: il primo titolo (X_1) è Google Inc.SpA, il secondo (X_2) è Fastweb S.p.A. Le probabilità assegnate a ciascun stato di natura per i tre titoli sono riportate nella tabella che segue:

s_i	$P(s_1)$	R	X_1	X_2
s_1	0.0076	1	0.8918	0.9931
s_2	0.0025	1	0.8975	1.0273
s_3	0.1574	1	0.9666	0.9976
s_4	0.0254	1	0.9659	1.0512
s_5	0.0152	1	0.9949	0.9458
s_6	0.6345	1	1.0024	0.9996
s_7	0.1168	1	1.0039	1.0302
s_8	0.0025	1	1.0715	0.9599
s_9	0.0330	1	1.0549	1.0008
s_{10}	0.0051	1	1.0649	1.0275

Calcoliamo tutte le quantità che ci servono con riferimento a X_1 e a X_2:

$$\mathbb{E}^P[X_1] = (0.8918)0.0076 + (0.8975)0.0025 + (0.9666)0.1574 + (0.9659)0.0254 + \ldots + \\ +(1.0549)0.0330 + (1.0649)0.0051 = 0.9971$$

$$\mathbb{E}^P[X_1 - e'R] = (0.8918 - 1)0.0076 + (0.8975 - 1)0.0025 + (0.9666 - 1)0.1574 + \ldots + \\ +(1.0549 - 1)0.0330 + (1.0648 - 1)0.0051 = -0.00298$$

$$Var^P[X_1] = (0.8918 - 0.9971)^2 0.0076 + (0.8975 - 0.9971)^2 0.0025 + \ldots + \\ +(1.0549 - 0.9971)^2 0.0330 + (1.0649 - 0.9971)^2 0.0051 = 0.000450$$

$$\mathbb{E}^P[X_2] = (0.9931)0.0076 + (1.0273)0.0025 + (0.9676)0.1574 + (1.0333)0.0254 + \ldots + \\ +(1.0008)0.0330 + (1.0275)0.0051 = 1.0034$$

$$\mathbb{E}^P[X_2 - e'R] = (0.9931 - 1)0.0076 + (1.0273 - 1)0.0025 + (0.9676 - 1)0.1574 + \ldots + \\ +(1.0008 - 1)0.0330 + (1.0275 - 1)0.0051 = 0.00352$$

$$Var^P[X_2] = (0.9931 - 1.0029)^2 0.0076 + (1.0273 - 1.0029)^2 0.0025 + \ldots + \\ +(1.0008 - 1.0029)^2 0.0330 + (1.0275 - 1.0029)^2 0.0051 = 0.000210$$

Ai fini di trovare α_{mv} ci calcoliamo anche la covarianza:

$$Cov(X_1, X_2) = \mathbb{E}\big[(X_1 - \mathbb{E}[X_1])P(X_1) \cdot (x_2 - \mathbb{E}[X_2]P(X_2))\big] = -0.000002$$

Ora possiamo ricavare α^*_{mv}:

$$\alpha_{mv} = \begin{pmatrix} \alpha^1_{mv} \\ \alpha^2_{mv} \end{pmatrix} = \frac{1}{\theta} \begin{pmatrix} Var^P[X_1] & Cov(X_1, X_2) \\ Cov(X_1, X_2) & Var^P[X_2] \end{pmatrix} \begin{pmatrix} \mathbb{E}^P[X_1 - R] \\ \mathbb{E}^P[X_2 - R] \end{pmatrix} =$$

$$= \frac{1}{10} \begin{pmatrix} 0.000450 & -0.000002 \\ -0.000002 & 0.000210 \end{pmatrix}^{-1} \begin{pmatrix} 0.00298 \\ 0.00352 \end{pmatrix} =$$

$$= \frac{1}{0.00000094} \begin{pmatrix} 0.000210 & 0.000002 \\ 0.000002 & 0.000450 \end{pmatrix} \begin{pmatrix} 0.00298 \\ 0.00352 \end{pmatrix} =$$

$$= \begin{pmatrix} -0.6548 \\ 1.6701 \end{pmatrix}$$

Dato α_{mv}, calcoliamo W_{mv}:

$$W_{mv} = R + \alpha_{mv}^1(X_1 - R) + \alpha_{mv}^2(X_2 - R) =$$
$$= \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - 0.6548 \begin{pmatrix} 0.8918 \text{ -1} \\ 0.8975 \text{ -1} \\ 0.9666 \text{ -1} \\ 0.9659 \text{ -1} \\ 0.9949 \text{ -1} \\ 1.0024 \text{ -1} \\ 1.0039 \text{ -1} \\ 1.0715 \text{ -1} \\ 1.0549 \text{ -1} \\ 1.0649 \text{ -1} \end{pmatrix} + 1.6701 \begin{pmatrix} 0.9931 \text{ -1} \\ 1.0273 \text{ -1} \\ 0.9976 \text{ -1} \\ 1.0512 \text{ -1} \\ 0.9458 \text{ -1} \\ 0.9996 \text{ -1} \\ 1.0302 \text{ -1} \\ 0.9599 \text{ -1} \\ 1.0008 \text{ -1} \\ 1.0275 \text{ -1} \end{pmatrix} = \begin{pmatrix} 1.0593 \\ 1.1127 \\ 1.0178 \\ 1.1078 \\ 0.9128 \\ 0.9977 \\ 1.0478 \\ 0.8862 \\ 0.9653 \\ 1.0034 \end{pmatrix}$$

Calcoliamo il valore di k:

$$\begin{aligned} \mathbb{E}^P[W] &= (1.0593)0.0076 + (1.1127)0.0025 + (1.0178)0.1574 + (1.1078)0.0254 + \ldots + \\ &+ (0.9653)0.0330 + (1.0034)0.0051 = 1.0076 \end{aligned}$$

$$k = (1/10) + \mathbb{E}^P[W] = 0.1 + 1.0070 = 1.1076$$

In questo caso $W_2 = 1.1127 > k$ e $W_4 = 1.1078 > k$, quindi assegneremo probabilità zero a $P(s_2|W_2 \leq k)$ e $P(s_4|W_4 \leq k)$ e di conseguenza anche le probabilità condizionate assegnate agli altri stati subiranno delle variazioni.
Se $P(s_2) = 0$ e $P(s_4) = 0$, allora

$$\sum_{i=1}^{10} P(s_i) = 0.0076 + 0.1574 + 0.0152 + 0.6345 + 0.1168 + 0.0025 + 0.0330 = 0.9721$$

Le probabilità condizionate assumaranno i seguenti valori:

$$\begin{aligned} P(s_1|A) &= 0.0076/0.9721 = 0.0078 \\ P(s_2|A) &= 0.0000/0.9721 = 0.0000 \\ P(s_3|A) &= 0.1574/0.9721 = 0.1619 \\ P(s_4|A) &= 0.0000/0.9721 = 0.0000 \\ P(s_5|A) &= 0.0152/0.9721 = 0.0156 \end{aligned}$$

$$
\begin{aligned}
P(s_6|A) &= 0.6345/0.9721 = 0.6527 \\
P(s_7|A) &= 0.1168/0.9721 = 0.1202 \\
P(s_8|A) &= 0.0025/0.9721 = 0.0026 \\
P(s_9|A) &= 0.0330/0.9721 = 0.0339 \\
P(s_{10}|A) &= 0.0051/0.9721 = 0.0053
\end{aligned}
$$

Usando le probabilità condizionate, ricaviamo la media e la varianza di X_1 e X_2:

$$
\begin{aligned}
\mathbb{E}^P[X_1|A] = {} & (0.8918)0.0078 + (0.8975)0.0000 + (0.9666)0.1619 + (0.9659)0.0000 + \ldots + \\
& +(1.0549)0.0339 + (1.0649)0.0053 = 0.9981
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}^P[X_1 - e'R|A] = {} & (0.8918 - 1)0.0078 + (0.8975 - 1)0.0000 + (0.9666 - 1)0.1619 + \ldots + \\
& +(1.0549 - 1)0.0339 + (1.0648 - 1)0.0053 = -0.00191
\end{aligned}
$$

$$
\begin{aligned}
Var^P[X_1|A] = {} & (0.8918 - 0.9971)^2 0.0078 + (0.8975 - 0.9971)^2 0.0000 + \ldots + \\
& +(1.0549 - 0.9971)^2 0.0339 + (1.0649 - 0.9971)^2 0.0053 = 0.000412
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}^P[X_2|A] = {} & (0.9931)0.0078 + (1.0273)0.0000 + (0.9676)0.1619 + (1.0333)0.0000 + \ldots + \\
& +(1.0008)0.0339 + (1.0275)0.0053 = 1.0022
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}^P[X_2 - e'R|A] = {} & (0.9931 - 1)0.0078 + (1.0273 - 1)0.0000 + (0.9676 - 1)0.1619 + \ldots + \\
& +(1.0008 - 1)0.0339 + (1.0275 - 1)0.0053 = 0.00215
\end{aligned}
$$

$$
\begin{aligned}
Var^P[X_2|A] = {} & (0.9931 - 1.0029)^2 0.0078 + (1.0273 - 1.0029)^2 0.0000 + \ldots + \\
& +(1.0008 - 1.0029)^2 0.0339 + (1.0275 - 1.0029)^2 0.0053 = 0.000160
\end{aligned}
$$

$$
Cov(X_1, X_2|A) = \mathbb{E}\big[(X_1 - \mathbb{E}[X_1])P(X_1) \cdot (x_2 - \mathbb{E}[X_2]P(X_2))\big] = 0.000045
$$

Possiamo ora trovare il valore di α^*_{mmv}:

$$
\begin{aligned}
\alpha_{mmv} = \begin{pmatrix} \alpha^1_{mmv} \\ \alpha^2_{mmv} \end{pmatrix} &= \frac{1}{\theta} \begin{pmatrix} Var^P[X_1|A] & Cov(X_1, X_2|A) \\ Cov(X_1, X_2|A) & Var^P[X_2|A] \end{pmatrix} \begin{pmatrix} \mathbb{E}^P[X_1 - R|A] \\ \mathbb{E}^P[X_2 - R|A] \end{pmatrix} = \\
&= \frac{1}{10} \begin{pmatrix} 0.000412 & 0.000045 \\ 0.000045 & 0.000160 \end{pmatrix}^{-1} \begin{pmatrix} -0.00191 \\ 0.00215 \end{pmatrix} = \\
&= \frac{1}{0.00000071} \begin{pmatrix} 0.000160 & -0.000045 \\ -0.000045 & 0.000412 \end{pmatrix} \begin{pmatrix} -0.00191 \\ 0.00215 \end{pmatrix} = \\
&= \begin{pmatrix} -0.5635 \\ 1.3609 \end{pmatrix}
\end{aligned}
$$

Dato α_{mmv}, calcoliamo W_{mmv}:

$$W_{mmv} = R + \alpha^1_{mmv}(X_1 - R) + \alpha^2_{mmv}(X_2 - R) =$$

$$= \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - 0.5635 \begin{pmatrix} 0.8918 \text{ -1} \\ 0.8975 \text{ -1} \\ 0.9666 \text{ -1} \\ 0.9659 \text{ -1} \\ 0.9949 \text{ -1} \\ 1.0024 \text{ -1} \\ 1.0039 \text{ -1} \\ 1.0715 \text{ -1} \\ 1.0549 \text{ -1} \\ 1.0649 \text{ -1} \end{pmatrix} + 1.3609 \begin{pmatrix} 0.9931 \text{ -1} \\ 1.0273 \text{ -1} \\ 0.9976 \text{ -1} \\ 1.0512 \text{ -1} \\ 0.9458 \text{ -1} \\ 0.9996 \text{ -1} \\ 1.0302 \text{ -1} \\ 0.9599 \text{ -1} \\ 1.0008 \text{ -1} \\ 1.0275 \text{ -1} \end{pmatrix} = \begin{pmatrix} 1.0515 \\ 1.0949 \\ 1.0155 \\ 1.0888 \\ 0.9291 \\ 0.9981 \\ 1.0389 \\ 0.9073 \\ 0.9701 \\ 1.0008 \end{pmatrix}$$

Possiamo riassumere i risultati in questa tabella:

s_i	$P(s_i)$	$P(s_i\|A)$	R	X_1	X_2	W_{mv}	W_{mmv}
s_1	0.0076	0.0078	1	0.8918	0.9931	1.0593	1.0515
s_2	0.0025	0.0000	1	0.8975	1.0273	1.1127	1.0949
s_3	0.1574	0.1619	1	0.9666	0.9976	1.0178	1.0155
s_4	0.0254	0.0000	1	0.9659	1.0512	1.1078	1.0888
s_5	0.0152	0.0156	1	0.9949	0.9458	0.9128	0.9291
s_6	0.6345	0.6527	1	1.0024	0.9996	0.9977	0.9981
s_7	0.1168	0.1202	1	1.0039	1.0302	1.0478	1.0389
s_8	0.0025	0.0026	1	1.0715	0.9599	0.8862	0.9073
s_9	0.0330	0.0339	1	1.0549	1.0008	0.9653	0.9701
s_{10}	0.0051	0.0053	1	1.0649	1.0275	1.0034	1.0008
$\alpha^*_{mv} = (-0.6548, 1.6701)$							
$\alpha^*_{mmv} = (-0.5635, 1.3609)$							
$k = 1.1706$							

Si può notare come i valori di W_{mmv} rientrano in un range più ristretto rispetto ai valori di W_{mv}, ovvero il portafoglio efficiente calcolato secondo il modello monotono porta ad avere rendimenti leggermente più bassi ma, in compenso, permette di subire perdite minori rispetto al portafoglio del modello media-varianza classico, ovvero:

$$\begin{array}{lcl} \text{se } W < 1 & \Rightarrow & W_{mv} < W_{mmv} \\ \text{se } W > 1 & \Rightarrow & W_{mv} > W_{mmv} \end{array}$$

Questo fenomeno lo si può osservare dalla figura 5.1, nella quale si evidenzia come varia il valore della ricchezza finale W nei 10 diversi stati di natura:

Figura 5.1: Ricchezza finale

La figura 5.1 mostra come per valori che rappresentano un incremento della ricchezza ($W > 1$), il modello *mv* porta ad incrementi maggiori rispetto al modello *mmv*, mentre per valori di W che rappresentano una perdita ($W < 1$) si ha che $W_{mv} < W_{mmv}$.

Conclusioni

La selezione del portafoglio secondo il modello Media-Varianza presentato da Markowitz permette di affrontare il trade off tra rischio e rendimento, ma potrebbe essere economicamente non significativa in quanto può verificarsi che le preferenze siano definite su un dominio non monotono. Per ovviare a questo inconveniente, Maccheroni, Marinacci, Rustichini e Togoba [15] hanno introdotto una versione monotona del modello Media-Varianza che coincide col modello classico quando le preferenze sono monotone e differisce nel caso in cui le preferenze media-varianza falliscano nella monotonicità. Il funzionale *mmv*, infatti, coincide col funzionale *mv* nel caso in cui le preferenze siano definite nel dominio di monotonicità $\mathcal{G}_\theta$, mentre fuori da $\mathcal{G}_\theta$ i due funzionali assumono forme diverse.
Abbiamo verificato come differiscono i due modelli applicando sui titoli Google Inc.SpA e Fastweb SpA sia il modello monotono sia quello non monotono. Ne è risultato che la ricchezza finale ottenuta applicando il modello monotono assume valori che spaziano in un range leggermente più ristretto rispetto ai valori ottenuti col modello non monotono. In particolare, il portafoglio allocato secondo un modello non monotono porta maggiori guadagni rispetto al portafoglio ottenuto con un modello monotono, ma comporta anche maggiori perdite in quanto prevede una maggiore varianza nei rendimenti.
In linea con quanto trovato da [15], possiamo concludere che un modello Media-Varianza dà rendimenti medi maggiori (vedi fig.4.2) rispetto a un modello Media-Varianza Montono, ma con una varianza maggiore (vedi fig.4.3). Quindi un modello di tipo monotono permette di allocare il portafoglio con il rischio di subire minori perdite rispetto a un modello non monotono, a scapito però di minori ritorni economici.

Appendice A

Analisi Matematica

A.1 Spazi metrici

Definizione 10 (Distanza) *Sia* $d : X \times X \to [0, \infty]$ *una funzione che ad ogni coppia* (x, y) *di punti di* X *associa un numero reale* $d(x, y) \geq 0$*, allora* d *è una* distanza su X *se si verificano le seguenti condizioni:*

1. $d(x, y) \geq 0$
2. $d(x, y) = 0$ *se e solo se* $x = y \qquad \forall x, y \in X$*;*
3. $d(x, y) = d(y, x) \qquad \forall x, y \in X$*;*
4. $d(x, y) \leq d(x, z) + d(y, z) \qquad \forall x, y \in X$

Definizione 11 (Disuguaglianza triangolare) *La condizione* $d(x, y) \leq d(x, z) + d(y, z)$ *prende il nome di* disuguaglianza triangolare.

Proposizione 12 *In uno spazio metrico ogni intorno circolare è un aperto, ogni unione di aperti è un aperto e l'intersezione di due aperti è un aperto. Inoltre, in uno spazio metrico l'unione di due chiusi è un chiuso e l'intersezione di due chiusi è un chiuso.*

Definizione 13 (Distanza euclidea) *Dato* $n \in \mathbb{N}$*, e dati* $x, y \in \mathbb{R}^n$ *di coordinate* $x = (x_1, x_2, ..., x_n)$ *e* $y = (y_1, y_2, ..., y_n)$*, si chiama* distanza euclidea *la seguente quantità:*

$$d_n(x, y) = \Big[\sum_{i=1}^{n} (x_i - y_i)^2 \Big]^{\frac{1}{2}}$$

A.2 Spazi vettoriali

Sia V uno *spazio vettoriale*, allora $\forall\ x, y \in V$ e $\forall\ \lambda, \mu \in \mathbb{R}$ valgono i seguenti assiomi:

1. $x + y = y + x$;
2. $(x + y) + z = x + (y + z)$;
3. $x + 0 = x$;
4. $x + (-x) = 0$;
5. $(\lambda + \mu)x = \lambda x + \mu x$
6. $\lambda(x + y) = \lambda x + \lambda y$
7. $0 \cdot x = 0$ e $1 \cdot x = x$

Definizione 14 *I vettori* $\{x_1, x_2, ..., x_n\}$ *costituiscono una base dello spazio vettoriale* V *se e solo se ogni vettore* x *si può esprimere come combinaizone lineare di* $x_1, x_2, ..., x_n$*:*

$$x = \lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_n x_n$$

Definizione 15 (Prodotto scalare) *Il* prodotto scalare *tra due elementi* $x = (x_1, x_2, ..., x_n)$ *e* $y = (y_1, y_2, ..., y_n)$ *di* $\mathbb{R}^n$ *è definito da:*

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

Proprietà del prodotto scalare:

i) $\langle x, y \rangle = \langle y, x \rangle$;

ii) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, x \rangle$;

iii) $\lambda \langle x, y \rangle = \langle \lambda x, y \rangle$;

iv) $\langle x, x \rangle \geq 0$;

v) $\langle x, x \rangle = 0$ se e solo se $x = 0$.

Definizione 16 (Modulo) *Si definisce* modulo *(o* norma euclidea*) di un vettore* $x \in \mathbb{R}^n$ *la quantità*

$$|x| = \langle x, x \rangle = \Big(\sum_{i=1}^{n} x_i^2 \Big)^{\frac{1}{2}}$$

A.3 Spazi normati

Definizione 17 *Per ogni elemento* $p \in [1, \infty)$ *la funzione* $||\cdot||_p : \mathbb{R}^n \to \mathbb{R}$ *è una* norma *su* $\mathbb{R}^n$.

Proprietà di una norma:

i) $||x|| = 0 \Leftrightarrow x = \underline{0}$;

ii) $||x + y|| \leq ||x|| + ||y||$;

iii) $||\alpha x|| = |\alpha| \cdot ||x||$.

Teorema 18 (Disuguaglianza di Minkowski) *Per ogni* $p \in [1, \infty)$ *e per ogni* $x, y \in \mathbb{R}^n$ *vale la seguente disuguaglianza:*

$$||x + y||_p \leq ||x||_p + ||y||_p \tag{A.1}$$

La A.1 è detta disuguaglianza di Minkowski

A.4 Spazi di Banach

Sia (X, d) uno spazio metrico. Si dice che x_k è una *successione di Cauchy* in X se $\forall \epsilon > 0$ esiste un $v \in \mathbb{N}$ tale che:

$$d(x_k, x_h) < \epsilon \qquad \forall\, h, k > v$$

Lo spazio metrico (X, d) si dice *completo* se ogni successione di Cauchy è convergente.

Definizione 19 (Convergenza) *Sia* I *un insieme di numeri reali e sia* $f_k : I \to \mathbb{R}$ *una successione di funzioni reali. Si dice che* f_k converge *in* I *verso la funzione* $f : I \to \mathbb{R}$ *se*

$$\lim_{k \to +\infty} f_k(x) = f(x) \qquad \forall\, x \in I$$

ovvero se $\forall\, \epsilon > 0$ *e* $\forall\, x \in I$ *esiste un* $v \in \mathbb{N}$ *tale che*

$$|f_k(x) - f(x)| < \epsilon \qquad \forall\, k > v$$

Definizione 20 (Spazio di Banach) *Sia* V *è uno spazio metrico normato con norma* $||\cdot||_V$. *Allora* V *è uno* spazio di Banach *se è completo come spazio metrico rispetto alla distanza*

$$d(x, y) = ||x - y||_V$$

dove $d(x, y)$ *è generata dalla norma* $||\cdot||_V$.

Definizione 21 (Operatori lineari) *A è un* operatore lineare *se*

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A x_1 + \alpha_2 A x_2$$

$\forall\, x_1, x_2 \in X$ *e* $\forall\, \alpha_1, \alpha_2 \in \mathbb{R}$. *Se X è uno spazio vettoriale normato, allora l'operatore lineare A è detto* limitato *se esiste una costante M tale che*

$$||Ax|| \leq M||x|| \qquad \forall\, x$$

Definizione 22 (Duale) *Si chiama* duale *(o* coniugato*) di X lo spazio X^* formato dai funzionali lineari limitati in uno spazio normato X.*

Proposizione 23 *Il duale X^* di un qualsiasi spazio normato X è uno spazio di Banach.*

Definizione 24 (Ordine parziale e totale) *Un insieme S è* parzialmente ordinato *se valgono le seguenti proprietà:*

1. *Reflessività: $a \leq a\ \forall\ a \in S$;*
2. *Assimetria debole: se $a \leq b$ e $b \leq a$, allora $a = b$;*
3. *Transitività: se $a \leq b$ e $b \leq c$, allora $a \leq c$.*

Un insieme S è totalmente ordinato *se insieme alle proprietà $1, 2, 3$ vale anche la seguente proprietà:*

4. *Comparabilità: Per ogni $a, b \in S$, si ha che $a \leq b$ o $b \leq a$.*

A.5 Spazi $\mathcal{L}^P$

Definizione 25 (Algebra) *Sia Ω un insieme non vuoto e sia $\mathcal{A}$ un sottoinsieme di Ω. L'insieme $\mathcal{A}$ è detto* algebra *se soddisfa le seguenti proprietà:*

(1) $0 \in \mathcal{A}$ *e* $\mathcal{A} \in \Omega$;

(2) se $B \in \mathcal{A}$ allora anche $B^c \in \mathcal{A}$ (con B^c complementare di B);

(3) se $B_1, B_2, ...B_n \in \mathcal{A}$ allora anche l'unione di tutti gli insiemi $B_i \in \mathcal{A}$ e l'intersezione di tutti gli insieme $B_i \in \mathcal{A}$, con $i = 1, 2, ..., n$

Inoltre $\mathcal{A}$ è detto $\sigma -$ algebra se soddisfa anche la seguente proprietà:

(4) se $B_1, B_2, \ldots \in \mathcal{A}$ allora anche l'unione di tutti gli insiemi $B_i \in \mathcal{A}$ e l'intersezione di tutti gli insieme $B_i \in \mathcal{A}$, con $i = 1, 2, \ldots, \infty$

Definizione 26 (spazio $\mathcal{L}^P$) *Sia $E \in \mathbb{R}^n$ uno spazio misurabile e sia $p \geq 1$ un numero reale, allora indichiamo con $\mathcal{L}^P$ la famiglia di tutte le funzioni f misurabili in E tali che $|f|^P$ sia sommabile:*

$$\mathcal{L}^P = \left\{ f : E \to \mathbb{R} \,\middle|\, f \text{ misurabile}, \int_E |f|^P \mathrm{d}x < +\infty \right\}$$

A.6 Sviluppo in serie di Taylor

Sia $f(x)$ una funzione reale definita in un intervallo (a, b) di $\mathbb{R}$ e sia $x_0 \in (a, b)$. Se f è dotata di derivate di ogni ordine nell'intervallo (a, b), allora f può essere sviluppata in *serie di Taylor* in (a, b), ovvero può essere sviluppata in serie di potenze di un punto iniziale x_0 nell'intervallo (a, b):

$$\begin{aligned} f(x) &= \textstyle\sum_{j=o}^{\infty} \frac{f^{(j)}(x_0)}{j!}(x - x_0)^j = \\ &= f(x_0) + f'(x_0)(x - x_0) + f''(x_0)\tfrac{(x-x_0)^2}{2!} + \ldots + f^{(j)}(x_0)\tfrac{(x-x_0)^j}{j!} + \ldots \end{aligned} \tag{A.2}$$

L'equazione A.2 è detta *serie di Taylor* di f, ovvero f è *sviluppabile in serie di Taylor in (a,b).*

Teorema 27 (Sviluppo in serie di Taylor di ordine K) *Sia $f : (a, b) \to \mathbb{R}$ una funzione derivabile K volte nell'intervallo (a, b), allora f è* sviluppabile in serie di Taylor di ordine K*:*

$$\begin{aligned} f(x) &= \textstyle\sum_{k=o}^{K} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k = \\ &= f(x_0) + f'(x_0)(x - x_0) + f''(x_0)\tfrac{(x-x_0)^2}{2!} + \ldots + f^{(K)}(x_0)\tfrac{(x-x_0)^K}{K!} \end{aligned} \tag{A.3}$$

Appendice B

Cenni di Analisi Convessa

B.1 Monotonia duale di Fenchel

Definizione 28 (Differenziale di Gateaux) *Sia $\psi : E \to \mathbb{R}$ un funzionale concavo e continuo, con E spazio vettoriale ordinato. Sia E' il duale di E e sia E'_+ il cono di tutti i funzionali positivi, lineari e continui in E.*
La derivata direzionale di ψ in x è:

$$d^+\psi(x)(v) = \lim_{t \to 0^+} \frac{\psi(x+tv) - \psi(x)}{t} \qquad \forall\, v \in E$$

Se il imite esiste per ogni $v \in E$, allora ψ è Gateaux differenziabile *in x e il funzionale $\Delta\psi(x) : v \to d^+(x)(v)$ è il* differenziale di Gateaux *di ψ.*

Definizione 29 (Superdifferenziale) *Il* superdifferenziale *di ψ rispetto a f è definito come segue:*

$$\partial\psi(x) = \left\{x' \in E' : \psi(y) - \psi(x) \leq \langle y - x, x' \rangle\right\} \qquad \forall\, y \in E$$

Definizione 30 (Coniugata di Fenchel) *La* coniugata di Fenchel *di ψ è data dal funzionale $\psi^* : E' \to (-\infty, +\infty)$ definito come segue:*

$$\psi^*(x') = \inf_{x \in E} \left\{\langle x, x' \rangle - \psi(x)\right\} \qquad \forall\, x' \in E'$$

B.2 Funzioni Concave e Disuguaglianza di Jensen

Definizione 31 (Insieme convesso) *Una funzione $f : C \to \mathbb{R}$, dove C è un insieme convesso in uno spazio vettoriale, è detta* concava *se:*

$$f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha) f(y)$$

$\forall\, x, y \in C$ *e* $\forall\, \alpha \in [0,1]$.
Una funzione f *è detta* convessa *se* $-f$ *è concava.*

Definizione 32 (Retta di sostegno) *Una retta passante per un punto* $(z, f(z))$ *appartenente al grafico di* f *è detta* retta di sostegno *se il suo grafico sta interamente al di sopra del grafico di* f*, cioè se esiste almeno un numero* $\lambda = \lambda(z)$ *tale che:*

$$f(t) \leq f(z) + \lambda \cdot (t - z) \qquad \forall\, t \in C$$

Teorema 33 (Funzione concava) *Una funzione è concava se e solo se esiste almeno una retta di sostegno per ogni punto del suo dominio.*

Dimostrazione

(i) Sia f concava. Allora:

$$\begin{aligned} f(x + (1-\alpha)(y-x)) \geq f(x) + (1-\alpha)[f(y) - f(x)] \Rightarrow \\ \Rightarrow \frac{f(x + (1-\alpha)(y-x)) - f(x)}{1-\alpha} \geq f(y) - f(x) \Rightarrow \\ \Rightarrow \frac{f(x + (1-\alpha)(y-x)) - f(x)}{(1-\alpha)(y-x)}(y-x) \geq f(y) - f(x). \end{aligned}$$

Se facciamo tendere $1-\alpha$ a 0 e se f è differenziabile, allora si ottiene:

$$f'(x) \cdot (y-x) \geq f(y) - f(x).$$

Se f non è differenziabile, basta prendere un qualsiasi λ:

$$\lambda \cdot (y-x) \geq f(y) - f(x).$$

(ii) Ponendo $\lambda = \alpha x - (1-\alpha)y$, abbiamo:

$$f(x) \leq f(\alpha x + (1-\alpha)y) + \lambda(1-\alpha)(x-y)$$

$$f(y) \leq f(\alpha x + (1-\alpha)y) + \lambda\alpha(y-x).$$

da cui si ottiene :

$$\alpha f(x) + (1-\alpha)f(y) \leq f(\alpha x + (1-\alpha)y)$$

Quindi possiamo concludere che f è *concava*, come volevamo dimostrare.

Teorema 34 (Disuguaglianza di Jensen) *Se f è concava, allora:*

$$\mathbb{E}[f(x)] \leq f(\mathbb{E}[x]) \qquad \forall\, x \in \mathbb{R}$$

Questo risultato è noto come disuguaglianza di Jensen.

Osservazione 35 *Dal teorema 33 deriviamo che un funzione concava (o convessa) è univocamente determinata dalla famiglia delle sue rette di sostegno. La funzione, infatti, coincide con l'estremo inferiore (rispettivamente, superiore) delle rette di sostegno, $\forall\, x \in C$. Quindi f è concava (o convessa) se e solo se esiste un insieme S tale che:*

$$f(x) = \inf_{l \in S} l(x) \qquad \Big(f(x) = \sup_{l \in S} l(x) \Big).$$

B.3 Funzionali lineari e sublineari

Una funzione F definita su uno spazio vettoriale V e a valori reali, $F : V \to \mathbb{R}$, è un *funzionale*. Un funzionale è detto:

(i) *lineare* se $F(\alpha x + \beta y) = \alpha F(x) + \beta F(y) \quad \forall\, x, y \in V$ e $\forall\, \alpha, \beta \in \mathbb{R}$;

(ii) *positivamente omogeneo* se $F(\alpha x) = \alpha F(x) \quad \forall\, x \in V$ e $\forall\, \alpha \geq 0$;

(iii) *subadditivo* se $F(x + y) \leq F(x) + F(y) \quad \forall\, x, y \in V$;

(iv) *sublineare* se è sia positivo omogeneo sia subadditivo, cioè se $F(\alpha x + \beta y) \leq \alpha F(x) + \beta F(y) \quad \forall\, x, y \in V$ e $\forall\, \alpha, \beta \geq 0$;

(v) *positivo* se $x \geqq 0$ implica $F(x) \geq 0$ ed è *strettamente positivo* se $x \geq 0$ implica $F(x) > 0$;

(vi) *monotono* se $y \geqq x$ implica $F(y) \geq F(x)$ ed è *strettamente monotono* se $y \geq x$ implica $F(y) > F(x)$.

Osservazione 36 *Per ogni funzionale sublineare $F : V \to \mathbb{R}$ si ha che:*

$$F(x) \geq -F(-x) \qquad \forall\, x \in V$$

Infatti:

$$0 = F(0) = F(x - x) \leq F(x) + F(-x) \Rightarrow F(x) \geq -F(-x)$$

Inoltre:

$$F(y) = F(y - x + x) \leq F(y - x) + F(x) \Rightarrow F(y - x) \geq F(y) - F(x)$$

B.4 Teoremi di separazione

Definizione 37 (Funzionale di Minkowski) *Sia $K \in \mathbb{R}^m$ un insieme convesso che contiene lo* 0. *Il funzionale definito da:*

$$p_K(x) = \inf \left\{ \alpha > 0 : \frac{x}{\alpha} \in K \right\}$$

è detto funzionale di Minkowski relativo a K.

Teorema 38 *Un funzionale di Minkowski è sempre finito e sublineare.*

Osservazione 39 *Un funzionale di Minkowski, essendo sublineare, è continuo. Valgono le seguenti proprietà:*

a) L'interno di K *è* $int(K) = \{x \in \mathbb{R}^m : p_K(x) < 1\}$*;*

b) La chiusura di K *è* $\overline{K} = \{x \in \mathbb{R}^m : p_K(x) \leq 1\}$*;*

c) La frontiera di K *è* $\partial K = \{x \in \mathbb{R}^m : p : K(x) = 1\}$.

Osservazione 40 *L'ipotesi che* 0 *sia punto interno dell'insieme convesso K serve a garantire che, $\forall x$, l'insieme $\{\alpha > 0 : \frac{x}{\alpha} \in K\}$ non sia vuoto e quindi che sia $p_K(x) < +\infty \ \forall x \in \mathbb{R}^m$. Se eliminiamo l'ipotesi che $0 \in K$, ci troveremo a definire un funzionale $p^*_K = \inf\{\alpha > 0 : \frac{x}{\alpha} \in K\}$ detto* funzionale generalizzato di Minkowski *relativo a K. Se* 0 *non è interno a K può succedere che $p^*_K = +\infty$ per qualche $x \in \mathbb{R}^m$.*

Teorema 41 (Teorema di Separazione) *Siano A e B due insiemi convessi e disgiunti in $\mathbb{R}^m$, almeno uno dei quali, supponiamo A, abbia interno non vuoto. Allora esiste un funzionale lineare non nullo $F : R^m \to \mathbb{R}$ che separa A e B, cioè esiste un funzionale tale che:*

$$F(x) \leq c \leq F(y) \qquad \forall\, x \in A, \forall\, y \in B, c \in \mathbb{R}$$

Osservazione 42 *La separazione è detta* impropria *quando $F(x) = c = F(y)\ \forall\, x \in A$ e $\forall\, y \in B$.*

B.5 Proprietà dei funzionali sublineari

Teorema 43 (Funzionali lineari) *Tutti i funzionali sublineari $F : \mathbb{R}^m \to \mathbb{R}$ sono del tipo:*

$$F(x) = \max_{\phi \in \mathcal{L}} \phi(x)$$

dove $\mathcal{L}$ è un insieme convesso e compatto (cioè chiuso e limitato) di funzionali lineari $\phi : \mathbb{R}^m \to \mathbb{R}$.

Osservazione 44 *Se $\mathcal{L}$ fosse soltanto chiuso e convesso, potrebbe succedere che $F(x) = -\infty \ \forall x \in \mathbb{R}^m$.*

Osservazione 45 *Se e solo se $\mathcal{L}_1 \leq \mathcal{L}_2$, allora:*

$$F_1(x) = \max_{\phi \in \mathcal{L}_1} \phi x \leq F_2(x) = \max_{\phi \in \mathcal{L}_2} \phi x \qquad \forall x \in \mathbb{R}^m$$

Teorema 46 *Un funzionale sublineare $F : V \to \mathbb{R}$ è* monotono *se e solo se*

$$x \geqq 0 \qquad \Rightarrow \qquad -F(-x) \geq 0$$

ed è strettamente monotono *se e solo se*

$$x \geq 0 \qquad \Rightarrow \qquad -F(-x) > 0$$

Dimostrazione.

Se F è monotono, allora:

$$x \geqq 0 \Rightarrow -x \leqq 0 \Rightarrow F(-x) \leq F(0) = 0 \Rightarrow -F(-x) \geq 0.$$

Viceversa,

$$\begin{aligned} y \geqq x \Rightarrow (y - x) \geqq 0 \quad & \Rightarrow -F(-(y - x)) \geq 0 \Rightarrow \\ & \Rightarrow F(x - y) \leq 0 \Rightarrow \\ & \Rightarrow F(x) - F(y) \leq F(x - y) \leq 0 \Rightarrow \\ & \Rightarrow F(y) \geq F(x). \end{aligned}$$

Teorema 47 (Funzionale sublineare positivo) *Un funzionale sublineare $F(x) = \max_{\phi \in \mathcal{L}} \phi(x)$ è* positivo *(o* strettamente positivo*) se e solo se $\mathcal{L}$ contiene almeno un funzionale lineare positivo (o strettamente positivo.)*

Teorema 48 (Funzionale sublineare monotono) *Un funzionale sublineare $F(x) = \max_{\phi \in \mathcal{L}} \phi(x)$ è* monotono *(o* strettamente monotono*) se e solo se $\mathcal{L}$ contiene solo funzionali lineari monotoni (o strettamente monotoni).*

B.6 Ottimizzazione Lineare

Siano A una matrice $m \times n$, $c \neq 0$ un vettore riga $\in \mathbb{R}^n$, $b \neq 0$ un vettore colonna $\in \mathbb{R}^m$ e $\xi \in \mathbb{R}^n$. Allora il problema :

$$\min_{\xi} c \cdot \xi \qquad \text{con } A\xi \geq b, \xi \geq 0 \tag{B.1}$$

è detto *problema di ottimizzazione lineare*.
L'insieme $\{\xi \in \mathbb{R}^m : A\xi \geq b, \xi \geq 0\}$, contenente i vincoli, è detto *regione ammissibile* e ogni punto in esso contenuto è detto *ammissibile*. Se la regione ammissibile è vuota, allora il problema di ottimizzazione lineare non ammette soluzione.

Definizione 49 (Duale) *Si chiama* duale *del problema (B.1) il seguente:*

$$\max_{\eta} c \cdot \eta b \qquad \textit{con } \eta A \leq c, \eta \geq 0 \tag{B.2}$$

Anche il problema (B.2) è un *problema di ottimizzazione lineare.*

Teorema 50 *1) Condizione necessaria e sufficiente affinchè un problema di ottimizzazione lineare ammetta soluzione è che il suo duale ammetta soluzione;*

2) Condizione necessaria e sufficiente affinchè $\hat{\xi}$ e $\hat{\eta}$ siano soluzioni ottime per i problemi (B.1) e (B.2) è che:

$$c\hat{\xi} = \hat{\eta} b$$

ovvero che

$$\min_{\xi} c\xi = \max_{\eta} \eta b.$$

B.7 Ottimizzazione Non Lineare

Siano $f : X \to \mathbb{R}$ e $g : X \to \mathbb{R}^m$, con $X \in \mathbb{R}^n$, rispettivamente una funzione scalare e una funzione vettoriale e sia b un vettore colonna $\in \mathbb{R}^m$.
Il problema:

$$\max f(x) \qquad \text{con } g(x) \leq b \tag{B.3}$$

è detto *problema di ottimizzazione non lineare.*. L'insieme $A = \{x : g(x) \leq b\}$ contenente il vincolo, è detto *regione ammissibile.*
Si distinguono due casi:

1. problemi di *ottimizzazione concava*, quando f è una funzione concava e g è una funzione convessa;

2. problemi di *ottimizzazione differenziabile*, quando f e g sono entrambe differenziabili (almeno nei punti di massimo).

Lemma 51 (Ottimizzazione concava) *Sia $\phi : X \to \mathbb{R}^k$ con $X \in \mathbb{R}^n$ convesso, una funzione vettoriale concava.*
Se $\phi(x) > 0$ è impossibile per $x \in X$, allora esiste un vettore riga $\hat{\lambda} \geq 0$ in $\mathbb{R}^k$ tale che $\hat{\lambda} \cdot \phi(x) \leq 0$ per ogni $x \in X$.

Appendice C

Preferenze e indici di concentrazione

C.1 Robustezza e Ambiguità

Assiomi.

Proprietà delle preferenze $\succeq$:

- A.1 *Ordine semplice.* Se $f, g, h \in F$, allora

$$f \succeq g, g \succeq h \Rightarrow f \succeq h$$

- A.2 *Indipendenza debole.* Se $f, g, \in F$, $x, y \in X$ e $\alpha \in (0,1)$, allora

$$\alpha f + (1-\alpha)x \succeq \alpha g + (1-\alpha)x \Rightarrow \alpha f + (1-\alpha)y \succeq \alpha g + (1-\alpha)y$$

- A.2' *Indipendenza.* (E' la versione più forte dell'assioma $A.2$). Se $f, g \in F$, $x \in X$ e $\alpha \in (0,1)$, allora

$$f \succeq g \Leftrightarrow \alpha f + (1-\alpha)x \succeq \alpha g + (1-\alpha)x$$

- A.3 *Continuità.* Se $f, g, h \in F$, allora gli insiemi

$$\{\alpha \in [0,1] : \alpha f + (1-\alpha)g \succeq h\}$$

  e

$$\{\alpha \in [0,1] : h \succeq \alpha f + (1-\alpha)g\}$$

  sono insiemi chiusi;

- A.4 *Monotonicità.* Se $f, g \in F$ e $f(s) \succeq g(s)$ $\forall s \in S$, allora $f \succeq g$;

- A.5 *Avversione all'incertezza.* Se $f, g \in F$ e $\alpha \in (0, 1)$, allora

$$f \sim g \Rightarrow \alpha f + (1 - \alpha) g \succeq f$$

  L'assioma $A.5$ può essere interpretato come assioma di avversione all'ambiguità;

- A.6 *Non degenericità.* Se $f, g \in F$, allora $f \succeq g$.

Teorema 52 *Sia $\succeq$ una relazione binaria su F, allora le seguenti condizioni sono equivalenti:*

(i) $\succeq$ *soddisfa gli assiomi $A.1 - A.6$;*

(ii) esiste una funzione affine non costante $u : X \to \mathbb{R}$ ed esiste una funzione limitata, convessa, semicontinua $c : \Delta \to [0, \infty)$ tale che $\forall\, f, g \in F$ si ha che

$$f \succeq g \Leftrightarrow \min_{Q \in \Delta} \left\{ \int u(f) \mathrm{d}Q + c(Q) \right\} \geq \min_{Q \in \Delta} \left\{ \int u(g) \mathrm{d}Q + c(Q) \right\} \qquad \text{(C.1)}$$

Per ogni u c'è un unico minimale $c^* : \Delta \to [0, \infty]$ che soddisfa la C.1 tale che:

$$c^*(Q) = \sup_{f \in F} \left\{ u(x_f) - \int u(f) \mathrm{d}Q \right\} \qquad \text{(C.2)}$$

Definizione 53 (Preferenza variazionale) *Una preferenza $\succeq$ è detta* variazionale *se soddisfa gli assiomi $A.1 - A.6$*

C.2 Attitudine all'ambiguità

Definizione 54 *Date due preferenze $\succeq_1$ e $\succeq_2$, si dice che $\succeq_1$ è più avversa all'ambiguità rispetto a $\succeq_2$ se, $\forall\, f \in F$ e $\forall\, x \in X$, si ha che:*

$$f \succeq_1 x \Rightarrow f \succeq_2 x$$

Proposizione 55 *Ogni preferenza variazionale è avversa all'ambiguità.*

Proposizione 56 *Dati due preferenze variazionali $\succeq_1$ e $\succeq_2$, le seguenti condizioni sono equivalenti:*

(i) $\succeq_1$ *è pià avverso all'ambiguità rispetto a* $\succeq_2$*;*

(ii) $u_1 \approx u_2$ *e* $c_1^* \leq c_2^*$,
con c^* *indice di avversione all'ambiguità.*

Una condizione necessaria affinchè siano avverse all'ambiguità è che esibiscano un comportamento tipo Ellsberg, ovvero che non siano probabilisticamente sofisticate.

C.3 Sofisticazione Probabilistica

Definizione 57 (Invarianza rispetto all'ordinamento) *Una relazione di preferenza del tipo* $\succeq$ *è detta* invariante rispetto all'ordinamento *(con riferimento a* P*) quando*

$$P(s \in S : x = f(s)) = P(s \in S : x = g(s)) \qquad \forall x \in X \Rightarrow f \sim g$$

Definizione 58 (Sofisticazione probabilistica) *Una relazione di preferenza del tipo* $\succeq$ *è detta* probabilisticamente sofisticata *quando*

$$P(s \in S : x \succeq f(s)) \leq P(s \in S : x \succeq g(s)) \qquad \forall x \in X \Rightarrow f \succeq g$$

C.4 Divergence preferences

Sia $P \in \Delta^\sigma$ una misura di probabilità. Data una funzione sigma-misurabile $w : S \to \mathbb{R}$ tale che $\inf_{s \in S} w(s) > 0$ e $\int w \, \mathrm{d}P = 1$ e data una funzione convessa e continua $\Phi : \mathbb{R}_+ \to \mathbb{R}_+$ tale che $\Phi(1) = 0$ e $\lim_{t \to \infty} \Phi(t)/f = \infty$, allora la Φ - *divergenza* pesata di $Q \in \Delta$ con riferimento a P è data da:

$$D_\Phi^w(Q||P) = \begin{cases} \int_S w(s)\Phi\left(\frac{\mathrm{d}Q}{\mathrm{d}P}(s)\right)\mathrm{d}P(s) & \text{se } Q \in \Delta^\sigma(P) \\ \infty & \text{altrimenti} \end{cases} \tag{C.3}$$

dove w sono i pesi (normalizzati). Se $w(s) = 1 \; \forall s \in S$, ovvero se w è uniforme, allora possiamo scrivere $D_\Phi^w(p||q) = D_\Phi(Q||P)$ che misura la divergenza tra distribuzioni. Le due forme più note di divergenza sono:

- *entropia relativa*:	$\Phi(t) = t\ln t - t + 1$
- *indice di concentrazione di Gini*:	$\Phi(t) = 2^{-1}(t-1)^2$

Le preferenze definite dal seguente funzionale $V : F \to \mathbb{R}$ sono dette preferenze di divergenza:

$$V(f) = \min_{Q \in \Delta^\sigma} \left\{ \int u(f)\mathrm{d}Q + \theta D_\Phi^w(Q||P) \right\}$$

Definizione 59 (Divergence preferences) *Una preferenza* $\succeq$ *è detta* preferenza di divergenza *se*

$$f \succeq g \Leftrightarrow \min_{Q \in \Delta^\sigma(P)} \left\{ \int u(f) \mathrm{d}Q + \theta D^w_\Phi(Q||P) \right\} \geq \min_{Q \in \Delta^\sigma(P)} \left\{ \int u(g) \mathrm{d}Q + \theta D^w_\Phi(Q||P) \right\}$$

dove $\theta > 0$, $u : X \to R$ *è una funzione affine,* $D^w_\Phi(.||P) : \Delta \to [0, \infty)$ *è la* $\Phi -$ *divergenza pesata data dall'equazione C.3.*

Teorema 60 *Le preferenze di divergenza sono preferenze variazionali continue, con indice di avversione all'ambiguità dato da:*

$$c^*(p) = \theta D^w_\Phi(Q||P) \qquad\qquad \forall Q \in \Delta$$

Queste preferenze sono probabilisticamente sofisticate se w *è uniforme.*

Tutte le *divergence preferences* sono preferenze variazionali probabilisticamente sofisticate e continue, come da teorema 60, e sono avverse all'ambiguità, come da proposizioni 55 e 56.

C.4.1 Misure di entropia

Definizione 61 (Entropia) *Una preferenza* $\succeq$ *è detta* preferenza entropica *se*

$$f \succeq^r g \Leftrightarrow \min_{Q \in \Delta^\sigma(P)} \left\{ \int u(f) \mathrm{d}Q + \theta R^w(Q||P) \right\} \geq \min_{Q \in \Delta^\sigma(P)} \left\{ \int u(g) \mathrm{d}Q + \theta R^w(Q||P) \right\}$$

dove $\theta > 0$, $P \in \Delta^\sigma$, $u : x \to R$ *è una funzione affine e* $R^w(.||P) : \Delta \to [0, \infty]$ *è l'entropia relativa pesata definita come segue:*

$$R^w(Q||P) = \begin{cases} \int_S w(s) \left(\frac{\mathrm{d}Q}{\mathrm{d}P}(s) ln \frac{\mathrm{d}Q}{\mathrm{d}P}(s) - \frac{\mathrm{d}Q}{\mathrm{d}P}(s) + 1 \right) \mathrm{d}P(s) & se\ Q \in \Delta^\sigma(P) \\ \infty & altrimenti \end{cases}$$

L'entropia $R^w(Q||P)$ è un caso particolare della divergenza $D^w_\Phi(Q||P)$ definita dall'equazione C.3, dove $\Phi(t) = t\ln t - t + 1$, quindi, dal teorema 60, le preferenze entropiche sono preferenze variazionali continue, con indice di ambiguità dato da:

$$c^*(Q) = \theta R^w(Q||P), \qquad \forall Q \in \Delta$$

Se w è uniforme, le preferenze $\succeq^r$ sono probabilisticamente sofisticate.
Nel caso ci fosse mancanza di uniformità nelle preferenze entropiche, allora esse produrrebbero un comportamento tipo Ellsberg e sarebbero avverse all'ambiguità, quindi se w non è uniforme le preferenze entropiche non sono probabilisticamente sofisticate.

Bibliografia

[1] J. Payne Bigelow. *Consistency of Mean-Variance Analysis and Expectation Utility Analysis.* Economics Letters, vol.43, pp.187-192, (1993).

[2] M.Bonato. *Convex Measure of Risk: Comparision of Different Approaches and Applications to Utility Maximization and Pricing Rules.* Tesi di Laurea Specialistica, Facoltà di Scienze Statistiche, Università Degli Studi di Padova, (2004).

[3] F.Cacciafesta. Lezioni di Matematica Finanziaria classica e moderna. G.Giappichelli Editore, (2001).

[4] E.Castagnoli, G.Favero, F.Maccheroni. Marchetto, il prezzatore perfetto. Edizioni Cakuntala, (2005).

[5] Cherubini, Dalla Lunga. Il Rischio Finanziario. McGrow-Hill, (2001).

[6] N. Fusco, P. Marcellini, C. Sbordone *Analisi Matematica Due* Liguori Editore, (2004)

[7] L.P.Hansen, T.J.Sargent. *Robust Control and Model Uncertainty.* American Economic Review Papers and Proceedings, 91, 2, pp 60-66, (2001).

[8] H.Jin, Jia An Yan, Xun Yu Zhou. *Continuous-Time Mean-Risk Portfolio Selection.* Elsevier SAS, (2005).

[9] M.L.Katz, H.S.Rosen. Microeconomia. McGrow-Hill.

[10] David Kelsey. *Maximin Expected Utility and Weight of Evidence.* Oxford Economic Paper 46, pp.425-444, (1994).

[11] A.l.Lavassani, Xun Li, G.D.Vidovic, A.Ware. *Dynamic Portfolio Selection under Capital-at-Risk.* The Matematical and Computional Finance Laboratory, University of Calgary, (Luglio 2003).

[12] Andrew E.B.Lim, Xun Li, Xun Yu Zhou. *Dynamic Mean-Variance Portfolio Selection with No-Shorting Constraints.* Society for Industrial and Applied Mathematics, vol.40, n.5, pp.1540-1555, (2002).

[13] Andrew E.B.Lim, Xun Yu Zhou. *Mean-Variance Portfolio Selection with Random Parameters in a Complete Market.* Mathematics of Operations Research, vol.27, n.1, pp.101-120, (Febbraio 2002).

[14] F.Maccheroni, M.Marinacci, A.Rustichini. *Ambiguity Aversion, Robustness, and the Variational Representation of Preferences.* (Maggio 2005).

[15] F.Maccheroni, M.Marinacci, A.Rustichini, M.Toboga. *Portfolio Selection with Monotone Mean-Variance Preferences.* ICER Working Paper - Applied Mathematich Series, vol.27, (2005).

[16] H.Markowitz. *Portfolio Selection.* J.Finance, 7, pp.77-91, (1952).

[17] Rockafellar, R.T.. Convex Analysis. Princeton University Press, (1970).

[18] Marc C.Steinbach. *Markowitz Revisited: Mean-Variance Models in Financial Portfolio Analysis.* Society for Industrial and Applied Mathematics, vol.43, n.1, pp.31-85, (2001).