

POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Matematica



Tesi di Laurea Magistrale

# Scambiabilità nei modelli di rischio di credito

**Relatore**

prof.ssa Patrizia Semeraro

**Candidato**

Erika Milletti

Anno Accademico 2020-2021

# Sommario

Questa tesi studia il concetto di scambiabilità nei modelli di rischio di credito. Nella prima parte vengono definite le due principali misure di rischio, Value-at-Risk ( $\text{VaR}_\alpha$ ) ed Expected-Shortfall ( $\text{ES}_\alpha$ ) ad un opportuno livello di confidenza  $\alpha$ , mostrandone vantaggi e limitazioni. Successivamente, sono presentati i mixture models come strumenti di calcolo per le misure di rischio di cui sopra, dando particolare importanza al modello misto di Bernoulli. Dal momento che il numero di insolvenze, condizionatamente alla variabile di mistura, segue una distribuzione binomiale, verrà considerata la sua approssimazione con una distribuzione di Poisson, arrivando così al modello misto di Poisson. Verrà dunque mostrato come i principali modelli standard industriali, quali CreditRisk<sup>+</sup> e CreditMetrics/KMV, possano essere scritti come modelli di mistura. Il concetto di scambiabilità viene introdotto per formalizzare matematicamente la nozione di gruppi omogenei, oggetto di studio nella tesi. Nella parte centrale dell'elaborato è presente un parallelismo tra i due modelli che parte con il definire un modello misto ad un fattore di Bernoulli e Poisson scambiabile, e arriva a dimostrare quale distribuzione segue il numero di insolvenze in entrambi i casi. Chiudono questa parte degli esempi numerici sul calcolo del VaR per il CreditRisk<sup>+</sup> usando entrambi i modelli e proponendo così un confronto. Viene infine introdotto il concetto di parziale scambiabilità con l'obiettivo di mostrare che, prendendo un modello misto di Bernoulli a due fattori, quindi con variabile latente bidimensionale, il vettore costituito dalle sequenze finite che rappresentano il modello di mistura soddisfa la definizione di parziale scambiabilità.

# Indice

<b>1</b>	<b>Variabili multidimensionali</b>	<b>7</b>
1.1	Definizioni preliminari . . . . .	7
1.2	Variabili casuali multidimensionali discrete . . . . .	8
1.3	Variabili casuali multidimensionali continue . . . . .	8
1.4	Distribuzione normale multivariata . . . . .	12
1.4.1	Normale bivariata . . . . .	12
1.4.2	Normale multivariata standardizzata . . . . .	13
1.4.3	Funzione generatrice dei momenti . . . . .	14
1.4.4	Proprietà della normale multidimensionale . . . . .	15
<b>2</b>	<b>Misure di rischio</b>	<b>19</b>
2.1	Definizioni preliminari . . . . .	19
2.1.1	Assiomi di coerenza . . . . .	19
2.2	Value-at-Risk . . . . .	20
2.2.1	Le limitazioni del VaR . . . . .	23
2.3	Expected-Shortfall . . . . .	27
<b>3</b>	<b>Mixture models e VaR nei modelli di rischio di credito</b>	<b>33</b>
3.1	Fondamenti per modellare il rischio di credito . . . . .	33
3.2	Mixture models . . . . .	34
3.2.1	CreditRisk <sup>+</sup> . . . . .	37
3.3	Relazione con i modelli a variabili latenti . . . . .	37
3.3.1	CreditMetrics/KMV come modelli misti di Bernoulli . . . . .	38
3.4	Modelli scambiabili ad un fattore in portafogli omogenei . . . . .	39
3.4.1	Modelli ad un fattore . . . . .	39
3.4.2	Bernoulli e Poisson mixture models scambiabili . . . . .	39
3.4.3	Correlazione . . . . .	40
3.4.4	CreditRisk <sup>+</sup> con distribuzione beta-binomiale . . . . .	41
3.4.5	CreditRisk <sup>+</sup> con distribuzione binomiale negativa . . . . .	43
3.5	Calcolo del VaR per la classe di modelli in $\mathcal{E}_d(p, \rho)$ . . . . .	44
3.5.1	Notazione . . . . .	44
3.5.2	Calibrazione del rischio con un modello misto di Bernoulli: esempi	44
3.5.3	Scenario 1 . . . . .	45
3.5.4	Scenario 2 . . . . .	48

3.5.5	Calibrazione del rischio con un modello misto di Poisson: esempi . .	51
3.5.6	Scenario 1 . . . . .	52
3.5.7	Scenario 2 . . . . .	55
3.6	Modello a due fattori e scambiabilità parziale . . . . .	58
<b>A</b>	<b>Variabili latenti e copule</b>	<b>63</b>
<b>B</b>	<b>Teorema di rappresentazione di De Finetti</b>	<b>67</b>

# Introduzione

Il problema principale alla base della gestione moderna del rischio è la misurazione del rischio. Il rischio di credito rappresenta uno dei fattori cruciali nella determinazione dei prezzi e dei rendimenti delle attività finanziarie. Esso può essere definito come: *"l'eventualità che una delle parti di un contratto non onori gli obblighi di natura finanziaria assunti, causando un perdita per la controparte creditrice"* (Ammann Manuel, 2001). La tesi non si focalizzerà sui motivi che rendono i debitori insolventi, ma piuttosto sull'entità della perdita, modellata come numero di defaults, attraverso un modello misto di Bernoulli. Vedremo che, in portafogli omogenei, il numero di insolvenze, condizionatamente alla variabile di mistura, segue una distribuzione binomiale. L'obiettivo della tesi è partire da questo esempio e cercare di fare un avanzamento in due direzioni. La prima consisterà nell'andare a vedere cosa succede, sotto le stesse condizioni, se utilizziamo un modello misto di Poisson ad un fattore scambiabile. Da qui seguirà un parallelismo e un confronto tra i due modelli che culminerà in degli esempi numerici. La seconda invece, che terminerà l'elaborato, proporrà una breve introduzione teorica del concetto di parziale scambiabilità abbandonando i modelli ad un fattore e considerando modelli di mistura in cui le variabili latenti sono dei vettori, senza riportare degli esempi specifici. In particolare:

nel capitolo 1, verrà proposto un richiamo teorico delle variabili casuali multidimensionali, oggetto di studio nella tesi, e in particolare sulla distribuzione normale multivariata che poi verrà utilizzata nel capitolo 3 per i modelli a variabili latenti (vedi sezione 3.3).

Nel capitolo 2, invece, si entrerà nel vivo del lavoro partendo dalla definizione di misura di rischio e gli assiomi secondo cui una misura di rischio è coerente. Verrà poi introdotta la principale misura di rischio, il *Value-at-Risk* ( $\text{VaR}_\alpha$ ) (definizione 15), dove  $\alpha$  è un parametro fondamentale che rappresenta un opportuno livello di confidenza e definisce la percentuale di protezione dal rischio di perdita, presentandone le due principali limitazioni: non è una misura di rischio coerente e non tiene conto dell'entità delle perdite al di fuori del livello di confidenza  $\alpha$ . Per questo verrà proposta una misura di rischio non sostitutiva, ma piuttosto complementare, *l'Expected-Shortfall* ( $\text{ES}_\alpha$ ) (definizione 17), che tiene conto delle perdite eccedenti il VaR e quindi permette di contenere maggiormente il rischio fornendo una chiara indicazione sul capitale da possedere per fronteggiare eventuali situazioni estreme. Infine verrà mostrata un'elegante formula analitica con cui è possibile rappresentare l' $\text{ES}_\alpha$ .

Nel capitolo 3, verranno introdotti i mixture models come strumenti di calcolo espliciti per le misure di rischio di cui sopra. Particolare rilevanza verrà data al *Bernoulli mixture model* (definizione 18) e, dal momento che il default è un evento raro, si considererà la

sua approssimazione, il *Poisson mixture model* (definizione 19). In entrambe le definizioni, gli eventi di default hanno una struttura di indipendenza condizionata, con variabile condizionante che rappresenta il vettore dei fattori comuni o delle variabili macroeconomiche, al centro di studio di questi modelli. Verrà mostrato proprio come i principali modelli industriali finanziari, quali il CreditRisk<sup>+</sup> e CreditMetrics, possano essere scritti come modelli di mistura. Nell'analisi si farà riferimento a portafogli omogenei (assumendo le probabilità di default uguali per tutte le compagnie). Il modo corretto di formalizzare matematicamente la nozione di omogeneità sarà quello di assumere il vettore degli indicatori di default *scambiabile* (definizione 21). Ci sarà, poi, una sezione dedicata ai modelli ad un fattore, arrivando ad ottenere dei modelli misti di Bernoulli e Poisson scambiabili (vedi sezione 3.4.2). Seguirà da qui un parallelismo tra i due modelli che, nel caso di un modello misto di Bernoulli scambiabile ad un fattore, porterà alla dimostrazione che il numero di defaults segue una distribuzione beta-binomiale, con la variabile mista che è distribuita come una beta. Nel secondo caso, invece, si arriverà alla conclusione che il numero di compagnie insolventi segue una distribuzione binomiale negativa, con forma distributiva sottostante di tipo gamma. Nella sezione 3.5 verrà quindi proposto il calcolo esplicito del VaR usando entrambi i modelli, conoscendo sia la probabilità di insolvenza che il coefficiente di correlazione, evidenziandone così le diversità. Chiuderà il capitolo e l'elaborato la sezione 3.6 nella quale verrà proposta una piccola introduzione del concetto di *parziale scambiabilità* (definizione 22), generalizzazione della scambiabilità. Tale generalizzazione è resa necessaria dal fatto che, nella loro forma più generale, le variabili latenti sono dei vettori e dunque la nozione di sequenza omogenea viene sostituita da portafogli omogenei a blocchi. In questa sede si farà riferimento a due sole sequenze. Si mostrerà che, in un modello misto di Bernoulli a due fattori, poichè il teorema di De Finetti con variabile latente bivariata non è più valido, supponendo verificata la rappresentazione (3.30), il vettore costituito dalle due sequenze finite è parzialmente scambiabile.

# Capitolo 1

## Variabili multidimensionali

In questo capitolo viene presentato un breve richiamo sulle variabili multidimensionali utilizzate nel resto della tesi. In particolare, nella sezione 1.1 verranno date le definizioni principali di funzione di ripartizione congiunta e marginale per un vettore aleatorio  $\mathbf{X} \in \mathbb{R}^d$ . Seguirà poi, nelle sezioni 1.2 e 1.3, un parallelismo di defizioni tra variabili casuali multidimensionali continue e discrete e verranno dimostrate le principali proprietà del valore atteso e della matrice di covarianza per un vettore aleatorio.

Nella sezione 1.4 verrà definita la distribuzione normale multivariata con la relativa funzione di densità. Seguirà poi lo studio nel caso specifico  $d = 2$  (normale bivariata) e successivamente verrà definita la funzione generatrice dei momenti. Chiuderà il capitolo un elenco delle principali proprietà della distribuzione normale multidimensionale.

### 1.1 Definizioni preliminari

Le variabili casuali multidimensionali sono  $d$ -ple ordinate di variabili aleatorie unidimensionali definite sullo stesso spazio di probabilità. Esse vengono indicate nel seguente modo:

$$\mathbf{X} = (X_1, \dots, X_d)'$$

dove  $d$  rappresenta la dimensione del vettore.

Posto che il vettore  $\mathbf{x} = (x_1, \dots, x_d)'$  corrisponda ad una possibile determinazione della variabile  $\mathbf{X}$ , è possibile dare la seguente definizione:

**Definizione 1.** La funzione di distribuzione ( o ripartizione ) congiunta  $F(\mathbf{x})$  è data dalla probabilità del verificarsi dell'evento  $(X_1 \leq x_1) \cap \dots \cap (X_d \leq x_d)$ :

$$F(\mathbf{x}) = F(x_1, \dots, x_d) = P(X_1 \leq x_1, \dots, X_d \leq x_d).$$

Analogamente:

**Definizione 2.** Se  $X_1, \dots, X_k$  sono elementi di un vettore aleatorio  $d$ -dimensionale ( $k < d$ ), la funzione di ripartizione marginale è data dalla probabilità di verificarsi l'evento  $(X_1 \leq x_1) \cap \dots \cap (X_k \leq x_k)$ :

$$F(x_1, \dots, x_k) = \lim_{x_{k+1}, \dots, x_d \rightarrow \infty} F(x_1, \dots, x_d) = F(x_1, \dots, x_k, \infty, \dots, \infty).$$

## 1.2 Variabili casuali multidimensionali discrete

Una variabile aleatoria  $d$ -dimensionale si dice discreta se può assumere un numero finito o un'infinità numerabile di  $d$ -ple di valori.

**Definizione 3.** Se  $\mathbf{X} = (X_1, \dots, X_d)'$  è un vettore multidimensionale discreto, la funzione di probabilità di  $\mathbf{X}$  è definita da

$$p(x_1, \dots, x_d) = P(X_1 = x_1, \dots, X_d = x_d).$$

Dunque  $p(x_1, \dots, x_d)$ , detta anche funzione di massa, restituisce la probabilità congiunta associata ad una determinata  $d$ -pla di valori.

**Definizione 4.** La funzione di probabilità marginale corrisponde alla probabilità associata a un sottoinsieme di  $k$  componenti del vettore aleatorio  $\mathbf{X}$  ( con  $k < d$  ) ed è data dall'espressione seguente:

$$p(x_1, \dots, x_k) = \sum_{x_{k+1}} \cdots \sum_{x_d} p(x_1, \dots, x_k, \dots, x_d).$$

Dunque la probabilità marginale di un certo sottoinsieme di  $k$  componenti di  $\mathbf{X}$  fornisce la probabilità della  $k$ -pla  $(x_1, \dots, x_k)$  qualunque siano le determinazioni assunte da  $(X_{k+1}, \dots, X_d)$ .

La funzione di probabilità condizionata restituisce la probabilità associata ad un sottoinsieme di  $k$  componenti di  $\mathbf{X}$  dato che i restanti  $d - k$  assumano un valore noto:

$$\begin{aligned} & p(x_1, \dots, x_k | x_{k+1}, \dots, x_d) = \\ & = P(X_1 = x_1, \dots, X_k = x_k | X_{k+1} = x_{k+1}, \dots, X_d = x_d) = \frac{p(x_1, \dots, x_d)}{p(x_{k+1}, \dots, x_d)}. \end{aligned}$$

## 1.3 Variabili casuali multidimensionali continue

Tutte le definizioni date nel paragrafo precedente hanno un corrispettivo con riferimento alle variabili aleatorie multidimensionali continue, ovvero dotate di funzione di densità congiunta.

**Definizione 5.** Dato un vettore aleatorio  $\mathbf{X}$  associato alla funzione di ripartizione  $F$ , se esiste una funzione  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  tale che per qualsiasi  $\mathbf{x} \in \mathbb{R}^d$  valga:

$$F(\mathbf{x}) = F(x_1, \dots, x_d) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} f(\mathbf{t}) dt$$

essa è detta funzione di densità congiunta di  $\mathbf{X}$ .

La funzione densità congiunta soddisfa le seguenti proprietà:



1.  $f(\mathbf{x}) \geq 0$ ;
2.  $\int_{\mathbb{R}^d} f(\mathbf{x}) dx = 1$ ;
3. per qualsiasi  $\mathbf{x} = (x_1, \dots, x_d)'$  punto di continuità per  $f$ , si ha:

$$f(x_1, \dots, x_d) = \frac{\partial^d F(x_1, \dots, x_d)}{\partial x_1, \dots, \partial x_d}.$$

**Definizione 6.** La generica  $k$ -pla  $(X_1, \dots, X_k)$ , con  $k < d$ , ha funzione di densità marginale data dall'espressione:

$$f(x_1, \dots, x_k) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_k, \dots, x_d) dx_{k+1}, \dots, dx_d.$$

La funzione di densità di  $X_1, \dots, X_k$  condizionata a  $X_{k+1} = x_{k+1}, \dots, X_d = x_d$  è data da

$$f(x_1, \dots, x_k | x_{k+1}, \dots, x_d) = \frac{f(x_1, \dots, x_d)}{f(x_{k+1}, \dots, x_d)}.$$

**Definizione 7.** Se  $F(x_1, \dots, x_d)$  è la funzione di distribuzione congiunta di  $\mathbf{X}$  e le  $F_i(x_i)$  sono le distribuzioni marginali delle componenti unidimensionali  $X_i$ , con  $i = 1, \dots, d$ , queste ultime sono dette stocasticamente indipendenti se

$$F(\mathbf{X}) = \prod_{i=1}^d F_i(x_i) \tag{1.1}$$

oppure, nel caso in cui  $\mathbf{X}$  possieda funzione densità, se:

$$f(\mathbf{x}) = \prod_{i=1}^d f_i(x_i).$$

L'indipendenza stocastica di  $(X_1, \dots, X_d)'$  implica quella di tutti i possibili accoppiamenti delle sue componenti unidimensionali. Non vale il viceversa.

Le seguenti definizioni valgono in analogia per  $\mathbf{X}$  variabile casuale multidimensionale discreta o continua.

**Definizione 8.** Si consideri un generico vettore aleatorio  $d$ -dimensionale  $\mathbf{X} = (X_1, \dots, X_d)'$ . Si definisce il vettore media di  $\mathbf{X}$ :

$$E(\mathbf{X}) := (E(X_1), \dots, E(X_d))' \tag{1.2}$$

ovvero il vettore costituito dai valori attesi delle singole componenti di  $\mathbf{X}$ .

Le principali proprietà sono le seguenti:

1. Il valore atteso del trasposto di un vettore aleatorio è uguale al trasposto del valore atteso del vettore stesso:

$$E(\mathbf{X}') = E(\mathbf{X})'. \quad (1.3)$$

2. **Linearità.** Per  $\mathbf{X}$  vettore casuale  $d$ -dimensionale, una matrice  $B \in \mathbb{R}^{k \times d}$  e un vettore  $\mathbf{b} \in \mathbb{R}^k$  vale

$$E(B\mathbf{X} + \mathbf{b}) = BE(\mathbf{X}) + \mathbf{b}. \quad (1.4)$$

Proviamo la (1.4):

$$\begin{aligned} E(B\mathbf{X} + \mathbf{b}) &= (E(B_{k1}X_1 + b_1), \dots, E(B_{kd}X_d + b_k))' = \\ &= (B_{k1}E(X_1) + b_1, \dots, B_{kd}E(X_d) + b_k)' = B(E(X_1), \dots, E(X_d))' + (b_1, \dots, b_k)' = \\ &= BE(\mathbf{X}) + \mathbf{b} \end{aligned}$$

avendo usato le proprietà di linearità e omogeneità del valore atteso di una variabile unidimensionale.

**Definizione 9.** La matrice di covarianza di un vettore aleatorio  $\mathbf{X}$   $d$ -dimensionale è data dalla seguente espressione:

$$\text{cov}(\mathbf{X}) := E((\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))') = E(\mathbf{X}\mathbf{X}') - E(\mathbf{X})E(\mathbf{X}'). \quad (1.5)$$

La definizione precedente ha senso purchè valga

$$E(X_i^2) < \infty, \quad \forall i = 1, \dots, d.$$

Il perchè è facilmente intuibile se si considera che gli elementi della matrice di covarianza  $\Sigma = \text{cov}(\mathbf{X})$  sono dati da

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1d} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{d1} & \sigma_{d2} & \dots & \sigma_d^2 \end{pmatrix}$$

Dunque  $\Sigma$  è una matrice quadrata simmetrica (la simmetria di  $\Sigma$  discende dalla simmetria dell'operazione di covarianza; infatti  $\sigma_{ij} = \text{cov}(X_i, X_j) = \text{cov}(X_j, X_i) = \sigma_{ji}$ ) in  $\mathbb{R}^{d \times d}$  che contiene le varianze  $\sigma_{ii} = \sigma_i^2 = E(X_i - E(X_i))^2$  di ciascuna componente  $X_i$  sulla diagonale principale, mentre al di fuori di questa si trovano tutte le possibili covarianze  $\sigma_{ij} = \sigma_{ji} = E((X_i - E(X_i))(X_j - E(X_j)))$ , se  $i \neq j$ , tra le coppie di componenti  $(X_i, X_j)$ .

**Proposizione 1.** La matrice di covarianza  $\Sigma$  di un vettore casuale  $\mathbf{X} = (X_1, \dots, X_d)'$  è semidefinita positiva.

*Dimostrazione.* Preso  $\mathbf{x} = (x_1, \dots, x_d)' \in \mathbb{R}^d$ , vale:

$$x' \Sigma x = \sum_{i,j=1}^N \sigma_{ij} x_i x_j = \sum_{i,j=1}^N \text{cov}(X_i, X_j) x_i x_j = \sum_{i,j=1}^N \text{cov}(x_i X_i, x_j X_j)$$

dove nell'ultima uguaglianza abbiamo usato la linearità della covarianza in entrambi gli argomenti. Quindi:

$$x' \Sigma x = \text{cov}\left(\sum_{i=1}^d x_i X_i, \sum_{j=1}^d x_j X_j\right) = \text{var}(W) \geq 0$$

dove  $W = \sum_{i=1}^d x_i X_i$ . □

**Proposizione 2.** *Sia  $\mathbf{X}$  un vettore aleatorio  $d$ -dimensionale. Per ogni matrice  $B \in \mathbb{R}^{k \times d}$  e ogni vettore  $\mathbf{b} \in \mathbb{R}^k$  vale:*

$$\text{cov}(B\mathbf{X} + \mathbf{b}) = B \text{cov}(\mathbf{X}) B'. \quad (1.6)$$

*Dimostrazione.*

$$\begin{aligned} \text{cov}(B\mathbf{X} + \mathbf{b}) &= E(((B\mathbf{X} + \mathbf{b}) - E(B\mathbf{X} + \mathbf{b}))((B\mathbf{X} + \mathbf{b}) - E(B\mathbf{X} + \mathbf{b}))') = \\ &= E((B\mathbf{X} + \mathbf{b} - BE(\mathbf{X}) - \mathbf{b})(B\mathbf{X} + \mathbf{b} - BE(\mathbf{X}) - \mathbf{b})') \end{aligned}$$

avendo applicato la (1.5) e le proprietà del valore atteso. Semplificando e mettendo in evidenza la  $B$  si ottiene che

$$\text{cov}(B\mathbf{X} + \mathbf{b}) = E(B(\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))' B').$$

In definitiva, usando l'omogeneità del valore atteso, si arriva a:

$$\text{cov}(B\mathbf{X} + \mathbf{b}) = BE((\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))') B' = B \text{cov}(\mathbf{X}) B'. \quad \square$$

La *matrice di correlazione*  $\rho(\mathbf{X})$  può essere definita introducendo il vettore standardizzato  $\mathbf{Y}$ , in modo che  $Y_i = \frac{X_i}{\sqrt{\text{var}(X_i)}}$  per ogni  $i$  e ponendo

$$\rho(\mathbf{X}) := \text{cov}(\mathbf{Y}). \quad (1.7)$$

Se utilizziamo l'operatore  $P$  per  $\rho(\mathbf{X})$ , allora l'elemento  $(i, j)$  della matrice è

$$\rho_{ij} = \rho(X_i, X_j) = \frac{\text{cov}(X_i, X_j)}{\sqrt{\text{var}(X_i)\text{var}(X_j)}}.$$

## 1.4 Distribuzione normale multivariata

**Definizione 10.**  $\mathbf{X} = (X_1, \dots, X_d)'$  ha una distribuzione normale multivariata o Gaussiana se

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + A\mathbf{Z} \quad (1.8)$$

dove  $\stackrel{d}{=}$  rappresenta l'uguaglianza nel senso delle distribuzioni.  $\mathbf{Z} = (Z_1, \dots, Z_k)'$  è un vettore di variabili casuali normali standard univariate, cioè  $Z_i \sim \mathcal{N}(0,1)$  con  $i = 1, \dots, k$ , mentre  $A \in \mathbb{R}^{d \times k}$  e  $\boldsymbol{\mu} \in \mathbb{R}^d$  sono una matrice e un vettore di termini costanti, rispettivamente.

Il vettore media di questa distribuzione è dato da

$$E(\mathbf{X}) = E(\boldsymbol{\mu} + A\mathbf{Z}) = \boldsymbol{\mu} + AE(\mathbf{Z}),$$

avendo usato la (1.4). Dal momento che  $E(\mathbf{Z}) = \mathbf{0}$ , si ha:

$$E(\mathbf{X}) = \boldsymbol{\mu}.$$

Per quanto riguarda la matrice di covarianza, usando la (1.6), abbiamo:

$$\text{cov}(\mathbf{X}) = \text{cov}(\boldsymbol{\mu} + A\mathbf{Z}) = A\text{cov}(\mathbf{Z})A'.$$

In questo caso,  $\text{cov}(\mathbf{Z}) = \text{diag}(\sigma_{Z_1}^2, \dots, \sigma_{Z_k}^2) = \text{diag}(1, \dots, 1) = \mathbf{I}_d$ , quindi:

$$\Sigma = \text{cov}(\mathbf{X}) = AA'.$$

**Definizione 11.** Un vettore aleatorio  $d$ -dimensionale  $\mathbf{X}$  ha legge di distribuzione normale multivariata se l'espressione della sua densità è data da

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\}, \quad \mathbf{x} \in \mathbb{R}^d, \quad (1.9)$$

dove  $|\Sigma|$  denota il determinante di  $\Sigma$ ,  $E(\mathbf{X}) = \boldsymbol{\mu}$  e  $\text{cov}(\mathbf{X}) = \Sigma$ . Formalmente  $\mathbf{X} \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$ .

### 1.4.1 Normale bivariata

Nel caso particolare in cui  $d = 2$ , possiamo scrivere la matrice di covarianza come

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix} \quad (1.10)$$

dove  $\rho = \frac{\text{cov}(X_1, X_2)}{\sigma_1\sigma_2} = \frac{\sigma_{12}}{\sigma_1\sigma_2}$  è il coefficiente di correlazione lineare tra le componenti di  $\mathbf{X} = (X_1, X_2)$ . L'inversione della precedente matrice porta a

$$\Sigma^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{pmatrix} \sigma_2^2 & -\sigma_1\sigma_2\rho \\ -\sigma_1\sigma_2\rho & \sigma_1^2 \end{pmatrix} = \frac{1}{1-\rho^2} \begin{pmatrix} \sigma_1^{-2} & -\rho(\sigma_1\sigma_2)^{-1} \\ -\rho(\sigma_1\sigma_2)^{-1} & \sigma_2^{-2} \end{pmatrix} \quad (1.11)$$

avendo nell'ultimo passaggio portato all'interno della matrice  $\sigma_1^2\sigma_2^2$ . Da quest'ultima uguaglianza arriviamo all'espressione della funzione di densità della normale bivariata:

$$\begin{aligned}
 f(x_1, x_2) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \\
 &\times \exp\left\{-\frac{1}{2(1-\rho^2)} \underbrace{(x_1 - \mu_1, x_2 - \mu_2) \begin{pmatrix} \sigma_1^{-2} & -\rho(\sigma_1\sigma_2)^{-1} \\ -\rho(\sigma_1\sigma_2)^{-1} & \sigma_2^{-2} \end{pmatrix} \begin{pmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{pmatrix}}_{(x_1 - \mu_1)[\sigma_1^{-2}(x_1 - \mu_1) - \rho(\sigma_1\sigma_2)^{-1}(x_2 - \mu_2)] + (x_2 - \mu_2)[- \rho(\sigma_1\sigma_2)^{-1}(x_1 - \mu_1) + \sigma_2^{-2}(x_2 - \mu_2)]}\right\} \\
 &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \\
 &\times \exp\left\{-\frac{1}{2(1-\rho^2)} \left[ \left(\frac{x_1 - \mu_1}{\sigma_1}\right)^2 - \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \left(\frac{x_2 - \mu_2}{\sigma_2}\right)^2 \right]\right\} \quad (1.12)
 \end{aligned}$$

### 1.4.2 Normale multivariata standardizzata

Supponiamo  $\mathbf{X} \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$ . Sia  $\Lambda$  la matrice degli autovalori di  $\Sigma$  e  $V$  la corrispondente matrice modale (ovvero costituita dagli autovettori ortonormali). La diagonalizzazione di  $\Sigma$  è data pertanto da:

$$\Sigma = V\Lambda V' = V\Lambda^{1/2}(V\Lambda^{1/2})'. \quad (1.13)$$

Prendendo l'inversa di entrambi i membri e ricordando che  $V' = V^{-1}$  si ottiene che:

$$\Sigma^{-1} = V\Lambda^{-1}V' = V\Lambda^{-1/2}(V\Lambda^{-1/2})'. \quad (1.14)$$

Consideriamo ora il vettore  $\mathbf{Z}$  ottenuto mediante la trasformazione lineare  $\mathbf{Z} = \Lambda^{-1/2}V'(\mathbf{x} - \boldsymbol{\mu})$  (si osservi come tale trasformazione ricordi in un certo senso l'operazione di standardizzazione delle normali univariate). Essendo  $\mathbf{x} - \boldsymbol{\mu} = V\Lambda^{1/2}\mathbf{Z}$ , la forma quadratica che compare nell'esponente della densità (1.9) si trasforma in:

$$\begin{aligned}
 (\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) &= (V\Lambda^{1/2}\mathbf{Z})' \Sigma^{-1} (V\Lambda^{1/2}\mathbf{Z}) = \\
 &= (V\Lambda^{1/2}\mathbf{Z})' V\Lambda^{-1/2} (V\Lambda^{-1/2})' (V\Lambda^{1/2}\mathbf{Z})
 \end{aligned}$$

dove nell'ultima uguaglianza abbiamo sostituito la (1.14). Effettuando le operazioni di trasposizione e ricordando le proprietà di  $V$  si arriva a:

$$\begin{aligned}
 (\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) &= \mathbf{Z}' \Lambda^{1/2} V' V \Lambda^{-1/2} \Lambda^{-1/2} V' V \Lambda^{1/2} \mathbf{Z} = \\
 &= \mathbf{Z}' \Lambda^{1/2} \Lambda^{-1/2} \Lambda^{-1/2} \Lambda^{1/2} \mathbf{Z} = \mathbf{Z}' \mathbf{Z}. \quad (1.15)
 \end{aligned}$$

Inoltre, il determinante della matrice jacobiana della trasformazione inversa è dato da:

$$|J| = \left| \frac{\partial}{\partial \mathbf{z}} (V\Lambda^{1/2}\mathbf{Z} + \boldsymbol{\mu}) \right| = |V\Lambda^{1/2}| = |\Sigma|^{1/2}. \quad (1.16)$$

Infine, tenendo conto della (1.9), (1.15) e (1.16), la funzione di densità di  $\mathbf{Z}$  diventa:

$$\begin{aligned} f_{\mathbf{Z}}(\mathbf{z}) &= {}^1|J|f_{\mathbf{X}}(\boldsymbol{\mu} + V\Lambda^{1/2}\mathbf{z}) = \\ &= |\Sigma|^{1/2} \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(V\Lambda^{1/2}\mathbf{z})'\Sigma^{-1}(V\Lambda^{1/2}\mathbf{z})\right\} = \frac{1}{(2\pi)^{d/2}} \exp\left\{-\frac{1}{2}\mathbf{z}'\mathbf{z}\right\} \end{aligned} \quad (1.17)$$

ovvero  $\mathbf{Z} \sim \mathcal{N}_d(\mathbf{0}, \mathbf{I}_d)$ .

### 1.4.3 Funzione generatrice dei momenti

La funzione generatrice dei momenti è individuata dal seguente valore atteso:

$$M_{\mathbf{X}}(\mathbf{t}) = E(e^{\mathbf{t}'\mathbf{X}}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp(\mathbf{t}'\mathbf{x})f(\mathbf{x}) d\mathbf{x}, \quad \forall \mathbf{t} \in \mathbb{R}^d. \quad (1.18)$$

Andando a sostituire la funzione densità si ottiene:

$$M_{\mathbf{X}}(\mathbf{t}) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left\{\mathbf{t}'\mathbf{x} - \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\} d\mathbf{x}.$$

Per risolvere l'integrale trasformiamo la variabile di integrazione tramite  $\mathbf{x} - \boldsymbol{\mu} = \Sigma\mathbf{t} + \mathbf{y}$ . Tenendo conto che la trasformazione inversa è  $\mathbf{y} = \mathbf{x} - \boldsymbol{\mu} - \Sigma\mathbf{t}$  e che lo Jacobiano  $J = \frac{\partial}{\partial \mathbf{y}}(\boldsymbol{\mu} + \Sigma\mathbf{t} + \mathbf{y}) = \mathbf{I}_d$ , si ha:

$$\begin{aligned} M_{\mathbf{X}}(\mathbf{t}) &= \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left\{\mathbf{t}'(\mathbf{y} + \boldsymbol{\mu} + \Sigma\mathbf{t}) - \frac{1}{2}(\mathbf{y} + \Sigma\mathbf{t})'\Sigma^{-1}(\mathbf{y} + \Sigma\mathbf{t})\right\} d\mathbf{y} = \\ &= \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left\{\mathbf{t}'\boldsymbol{\mu} + \mathbf{t}'\Sigma\mathbf{t} + \mathbf{t}'\mathbf{y} - \frac{1}{2}\mathbf{y}'\Sigma^{-1}\mathbf{y} - \frac{1}{2}\mathbf{y}'\mathbf{t} - \frac{1}{2}\mathbf{t}'\mathbf{y} - \frac{1}{2}\mathbf{t}'\Sigma'\mathbf{t}\right\}. \end{aligned}$$

Applicando le operazioni di trasposizione, la simmetria della matrice di covarianza e portando fuori dall'integrale i termini che non dipendono da  $\mathbf{y}$ , si ottiene:

$$\begin{aligned} M_{\mathbf{X}}(\mathbf{t}) &= \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left\{\mathbf{t}'\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}'\Sigma\mathbf{t}\right\} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\mathbf{y}'\Sigma^{-1}\mathbf{y}\right\} d\mathbf{y}}_{(2\pi)^{d/2}|\Sigma|^{1/2}} = \\ &= \exp\left\{\mathbf{t}'\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}'\Sigma\mathbf{t}\right\} \end{aligned} \quad (1.19)$$

dove nell'ultimo integrale è stata applicata la proprietà di normalizzazione.

---

<sup>1</sup>Sia  $\mathbf{X} \in \mathbb{R}^d$  con funzione densità  $f_{\mathbf{X}}(\mathbf{x})$  e sia  $\mathbf{Y}$  ottenuto da  $\mathbf{X}$  tramite la trasformazione biunivoca e regolare  $\mathbf{Y} = g(\mathbf{X})$ . Poichè  $g$  è una trasformazione biunivoca, vale:  $\int_B f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} = \int_A f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \int_{B=g(A)} f_{\mathbf{X}}(g^{-1}(\mathbf{y}))|J| d\mathbf{y}$ , dove A e B appartengono ai supporti di  $\mathbf{X}$  e  $\mathbf{Y}$  in  $\mathbb{R}^d$ , rispettivamente. Si ricava così la funzione densità congiunta della  $\mathbf{Y}$ :  $f_{\mathbf{Y}}(\mathbf{y}) = |J|f_{\mathbf{X}}(g^{-1}(\mathbf{y}))$ .

Possiamo osservare che:

- (i)  $M_{\mathbf{X}}(\mathbf{t})$  è ben definita in quanto esiste per ogni  $\mathbf{t} \in \mathbb{R}^d$ ;
- (ii)  $M_{\mathbf{X}}(\mathbf{t})$  dipende in modo univoco da  $\boldsymbol{\mu}$  e  $\Sigma$ . In altre parole, ad ogni valore della coppia  $(\boldsymbol{\mu}, \Sigma)$  resta associata una ed una sola f.g.m (funzione generatrice dei momenti) e quindi una ed una sola legge di distribuzione normale in  $\mathbb{R}^d$ .

Quando  $\Sigma$  non è invertibile non esiste ovviamente l'espressione della densità data nella definizione 11: in questo caso parleremo di distribuzione normale  $d$ -dimensionale *singolare* intesa come legge di distribuzione del vettore  $\mathbf{X}$  la cui f.g.m è esattamente la (1.19).

#### 1.4.4 Proprietà della normale multidimensionale

**Teorema 1 (Cramér-Wold).** Sia  $\mathbf{X} \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$  e siano  $B \in \mathbb{R}^{k \times d}$ ,  $\mathbf{b} \in \mathbb{R}^k$ . Allora:

$$B\mathbf{X} + \mathbf{b} \sim \mathcal{N}_k(B\boldsymbol{\mu} + \mathbf{b}, B\Sigma B') \quad (1.20)$$

ovvero, qualsiasi vettore ottenuto come combinazione lineare di un vettore aleatorio avente distribuzione normale multivariata ha a sua volta distribuzione normale multivariata.

*Dimostrazione.* Sia  $\mathbf{Y}$  il vettore aleatorio definito dalla trasformazione lineare  $\mathbf{Y} = B\mathbf{X} + \mathbf{b}$ . Poichè

$$\begin{aligned} M_{\mathbf{Y}}(\mathbf{t}) &= E[\exp(\mathbf{t}'\mathbf{Y})] = E[\exp(\mathbf{t}'B\mathbf{X} + \mathbf{t}'\mathbf{b})] = \underbrace{E[\exp(\mathbf{t}'B\mathbf{X})]}_{M_{\mathbf{X}}(B'\mathbf{t})} \exp(\mathbf{t}'\mathbf{b}) = \\ &= \exp\left[\mathbf{t}'B\boldsymbol{\mu} + \frac{1}{2}\mathbf{t}'B\Sigma B'\mathbf{t}\right] \exp(\mathbf{t}'\mathbf{b}) = \\ &= \exp\left[\mathbf{t}'(B\boldsymbol{\mu} + \mathbf{b}) + \frac{1}{2}\mathbf{t}'(B\Sigma B')\mathbf{t}\right]. \end{aligned} \quad (1.21)$$

Dalla (1.19) e (1.21) si può concludere che

$$\mathbf{Y} \sim \mathcal{N}_k(B\boldsymbol{\mu} + \mathbf{b}, B\Sigma B').$$

La  $k$ -dimensionalità discende dal fatto che  $B\boldsymbol{\mu} \in \mathbb{R}^k$  e  $B\Sigma B' \in \mathbb{R}^{k \times k}$ . □

**Corollario 1.** Un vettore casuale formato da un sottoinsieme delle componenti di un vettore casuale avente distribuzione normale multivariata ha a sua volta distribuzione normale multivariata.

*Dimostrazione.* Siano

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

con  $\mathbf{X}_1 \in \mathbb{R}^{d_1}$ ,  $\mathbf{X}_2 \in \mathbb{R}^{d_2}$  tali che  $d_1 + d_2 = d$ . Applicando la (1.20) con  $\mathbf{b} = \mathbf{0}$  e  $B = (\mathbf{I}_{d_1} \quad \mathbf{0})$  si ottiene che

$$\mathbf{X}_1 \sim \mathcal{N}_{d_1}(\boldsymbol{\mu}_1, \Sigma_{11}) \quad \mathbf{X}_2 \sim \mathcal{N}_{d_2}(\boldsymbol{\mu}_2, \Sigma_{22})$$

□

**Teorema 2.** *Le distribuzioni condizionate di sottoinsiemi degli elementi di un vettore casuale  $\mathbf{X} \sim \mathcal{N}_d(\boldsymbol{\mu}, \Sigma)$  sono a loro volta normali multivariate.*

*Dimostrazione.* Siano

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{pmatrix} \quad \boldsymbol{\mu}_{\mathbf{X}} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \quad \Sigma_{\mathbf{X}} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

con  $\mathbf{X}_1 \in \mathbb{R}^{d_1}$ ,  $\mathbf{X}_2 \in \mathbb{R}^{d_2}$  tali che  $d_1 + d_2 = d$ . Considero la trasformazione lineare  $\mathbf{Y} = A\mathbf{X}$  dove

$$A = \begin{pmatrix} \mathbf{I}_{d_1} & \mathbf{0} \\ -\Sigma_{21}\Sigma_{11}^{-1} & \mathbf{I}_{d_2} \end{pmatrix}$$

I parametri della distribuzione del vettore aleatorio  $\mathbf{Y}$  sono dunque

$$\boldsymbol{\mu}_{\mathbf{Y}} = A\boldsymbol{\mu}_{\mathbf{X}} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ -\Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2 \end{pmatrix} \quad (1.22)$$

e

$$\begin{aligned} \Sigma_{\mathbf{Y}} = \text{cov}(\mathbf{Y}) &= \text{cov}(A\mathbf{X}) = A\Sigma_{\mathbf{X}}A' = \begin{pmatrix} \mathbf{I}_{d_1} & \mathbf{0} \\ -\Sigma_{21}\Sigma_{11}^{-1} & \mathbf{I}_{d_2} \end{pmatrix} \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{pmatrix} \mathbf{I}_{d_1} & -(\Sigma_{11}^{-1})'\Sigma_{21}' \\ \mathbf{0} & \mathbf{I}_{d_2} \end{pmatrix} \\ &= \begin{pmatrix} \Sigma_{11} & \mathbf{0} \\ \mathbf{0} & \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{pmatrix} \end{aligned} \quad (1.23)$$

Dalla (1.22) e (1.23) si deduce che il vettore  $\mathbf{Y}$  è costituito a sua volta da due componenti indipendenti

$$\mathbf{Y}_1 \sim \mathcal{N}_{d_1}(\boldsymbol{\mu}_1, \Sigma_{11}), \quad \mathbf{Y}_2 \sim \mathcal{N}_{d_2}(-\Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2, \Sigma_{2|1})$$

dove  $\Sigma_{2|1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ . Possiamo quindi esprimere la densità congiunta di  $\mathbf{Y}$  attraverso le densità di  $\mathbf{Y}_1$  e  $\mathbf{Y}_2$ :

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= f_{\mathbf{Y}_1}(\mathbf{y}_1)f_{\mathbf{Y}_2}(\mathbf{y}_2) = \frac{|\Sigma_{11}|^{-1/2}}{(2\pi)^{d_1/2}} \exp\left\{-\frac{1}{2}(\mathbf{y}_1 - \boldsymbol{\mu}_1)'\Sigma_{11}^{-1}(\mathbf{y}_1 - \boldsymbol{\mu}_1)\right\} \times \\ &\times \frac{|\Sigma_{2|1}|^{-1/2}}{(2\pi)^{d_2/2}} \exp\left\{-\frac{1}{2}(\mathbf{y}_2 + \Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)'\Sigma_{2|1}^{-1}(\mathbf{y}_2 + \Sigma_{21}\Sigma_{11}^{-1}\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)\right\}. \end{aligned}$$

Grazie a questa fattorizzazione possiamo ricavare la densità  $f_{\mathbf{X}}(\mathbf{x})$  di  $\mathbf{X} = A^{-1}\mathbf{Y}$ . Infatti, osservando che  $|J| = \left|\frac{\partial}{\partial \mathbf{x}}(A\mathbf{x})\right| = |A| = \mathbf{I}_{d_1}\mathbf{I}_{d_2} = 1$ , si ha:

$$f_{\mathbf{X}}(\mathbf{x}) = |J|f_{\mathbf{Y}}(A\mathbf{x}) = \underbrace{\frac{|\Sigma_{11}|^{-1/2}}{(2\pi)^{d_1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}_1 - \boldsymbol{\mu}_1)'\Sigma_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1)\right\}}_{f_{\mathbf{X}_1}(\mathbf{x}_1)} \times$$



$$\times \frac{|\Sigma_{21}|^{-1/2}}{(2\pi)^{d_2/2}} \exp\left\{-\frac{1}{2}[\mathbf{x}_2 - \Sigma_{21}\Sigma_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1) - \boldsymbol{\mu}_2]' \Sigma_{21}^{-1} [\mathbf{x}_2 - \Sigma_{21}\Sigma_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1) - \boldsymbol{\mu}_2]\right\}$$

$f_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2)$

avendo sostituito a  $\mathbf{y}_1 = \mathbf{x}_1$  e  $\mathbf{y}_2 = -\Sigma_{21}\Sigma_{11}^{-1}\mathbf{x}_1 + \mathbf{x}_2$ . Per la definizione di densità condizionata possiamo quindi concludere che

$$\mathbf{X}_2|\mathbf{X}_1 \sim \mathcal{N}_{d_2}(\Sigma_{21}\Sigma_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1) + \boldsymbol{\mu}_2, \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}). \quad (1.24)$$

□



# Capitolo 2

## Misure di rischio

L'obiettivo di questo capitolo è quello di introdurre le due principali misure di rischio usate dalla maggior parte delle industrie moderne finanziarie, mostrandone vantaggi e limitazioni. Nella sezione 2.1, dopo aver definito una misura di rischio, verranno elencati gli assiomi di coerenza. Nella sezione 2.2 verrà definito il Value-at-Risk mostrando, in un esempio, che si tratta di una misura di rischio non coerente.

Nella sezione 2.3 verrà infine introdotto l'Expected-Shortfall come misura di rischio complementare al VaR dimostrando che, a differenza del Value-at-Risk, è una misura di rischio coerente.

### 2.1 Definizioni preliminari

Fissato uno spazio di probabilità  $(\Omega, \mathcal{F}, P)$ , sia  $L^0(\Omega, \mathcal{F}, P)$  l'insieme delle variabili aleatorie finite su questo spazio. I rischi finanziari sono rappresentati da un cono convesso  $\mathcal{M} \subseteq L^0(\Omega, \mathcal{F}, P)$ . Ricordiamo che  $\mathcal{M}$  è un cono convesso se dati  $L_1 \in \mathcal{M}$  e  $L_2 \in \mathcal{M}$  allora  $L_1 + L_2 \in \mathcal{M}$  e  $\lambda L_1 \in \mathcal{M}$ , per ogni  $\lambda > 0$ .

**Definizione 12 (Misura di rischio).** Dato un cono convesso  $\mathcal{M}$  di variabili casuali, si definisce misura di rischio con dominio  $\mathcal{M}$  la funzione  $\varphi : \mathcal{M} \rightarrow \mathbb{R}$ .

#### 2.1.1 Assiomi di coerenza

**Assioma 1 (Invarianza per traslazione).** Per ogni  $L \in \mathcal{M}$  e per ogni  $l \in \mathbb{R}$ , una misura di rischio invariante per traslazione soddisfa:  $\varphi(L + l) = \varphi(L) + l$ .

**Assioma 2 (Subadditività).** Per ogni  $L_1, L_2 \in \mathcal{M}$  una misura di rischio subadditiva soddisfa:  $\varphi(L_1 + L_2) \leq \varphi(L_1) + \varphi(L_2)$ .

**Assioma 3 (Omogeneità positiva).** Per ogni  $L \in \mathcal{M}$  e per ogni  $\lambda > 0$ , una misura di rischio omogenea positiva soddisfa:  $\varphi(\lambda L) = \lambda \varphi(L)$ .

**Assioma 4 (Monotonia).** Presi  $L_1$  e  $L_2 \in \mathcal{M}$  tali che  $L_1 \leq L_2$ , una misura di rischio monotona soddisfa:  $\varphi(L_1) \leq \varphi(L_2)$ .

Osserviamo che la subadditività e l'omogeneità positiva implicano che il funzionale  $\varphi$  sia convesso in  $\mathcal{M}$ , ovvero:

$$\varphi(\lambda L_1 + (1 - \lambda)L_2) \leq \lambda\varphi(L_1) + (1 - \lambda)\varphi(L_2).$$

Infatti, presi  $L_1$  ed  $L_2 \in \mathcal{M}$  ne considero la combinazione convessa del tipo  $\lambda L_1 + (1 - \lambda)L_2$ . La subadditività implica che:

$$\varphi(\lambda L_1 + (1 - \lambda)L_2) \leq \varphi(\lambda L_1) + \varphi((1 - \lambda)L_2) = \lambda\varphi(L_1) + (1 - \lambda)\varphi(L_2)$$

dal momento che  $\varphi(\lambda L) = \lambda\varphi(L)$ .

**Definizione 13 (Misura di rischio coerente).** Data una misura di rischio  $\varphi$  il cui dominio include il cono convesso  $\mathcal{M}$ , una misura di rischio si dice coerente se soddisfa gli assiomi 1-4.

**Definizione 14 (Distribuzione di perdita).** Consideriamo un dato portafoglio come collezione di azioni o obbligazioni. Indichiamo il suo valore al tempo  $t$  con  $V(t)$  e assumiamo che la variabile aleatoria  $V(t)$  sia osservabile al tempo corrente. Fissato un orizzonte temporale  $[t, t + \Delta t]$ , la perdita di credito nel portafoglio è data da

$$L_{[t, t + \Delta t]} := -(V(t + \Delta t) - V(t)). \quad (2.1)$$

La distribuzione di  $L_{[t, t + \Delta t]}$  è chiamata distribuzione di perdita.

## 2.2 Value-at-Risk

Il Value-at-Risk (VaR) è probabilmente la misura di rischio più utilizzata dalle istituzioni finanziarie. Consideriamo un portafoglio costituito da assets rischiosi, fissiamo un orizzonte temporale e indichiamo con  $F_L(l) = P(L \leq l)$  la funzione di ripartizione della distribuzione di perdita  $L$ . L'obiettivo è quello di definire una statistica basata su  $F_L$ , la quale misura la probabilità di contenere il rischio di credito del nostro portafoglio su un intervallo di tempo  $[t, t + \Delta t]$ . Un ovvio candidato è la perdita massima possibile data da

$$\inf\{l \in \mathbb{R} : F_L(l) = 1\}.$$

Tuttavia nella maggior parte dei modelli di interesse il supporto di  $F_L$  è illimitato e quindi l'estremo inferiore sarebbe direttamente infinito. Dunque il VaR sostituisce la perdita massima con la probabilità di perdita massima potenziale tenendo conto sia di uno specifico orizzonte temporale che di un intervallo di confidenza predefinito, indicato con  $\alpha$ .

**Definizione 15 (Value-at-Risk).** Dato un certo livello di confidenza  $\alpha \in (0,1)$ , si definisce  $\text{VaR}_\alpha$  del nostro portafoglio al livello  $\alpha$  il più piccolo numero  $l$  tale che la probabilità che la perdita  $L$  superi  $l$  sia non più grande di  $1 - \alpha$ . Formalmente:

$$\text{VaR}_\alpha := \inf\{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\}. \quad (2.2)$$

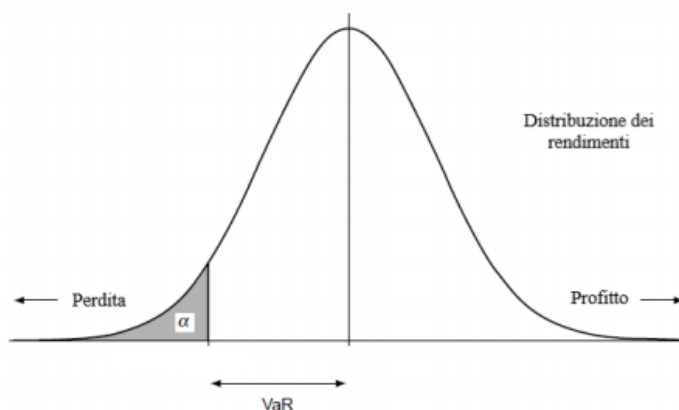


Figura 2.1: VaR

Tipici valori di  $\alpha$  sono  $\alpha = 0,95$  o  $\alpha = 0,99$ . Nel rischio di credito e operativo l'orizzonte temporale è solitamente un anno. Il VaR si caratterizza per essere una tecnica semplice e di immediata comprensione, che ha il pregio di poter essere applicata ad un numero elevato di rischi. Può essere considerato come uno strumento che, attraverso valori monetari, esprime la misura di rischio a cui un particolare individuo, detentore di un certo capitale o attività, è soggetto.

Affermare, per esempio, che il valore del VaR ad un livello di confidenza del 95 % e in un arco temporale di 10 giorni sia pari a 300 €, corrisponde a dire che nei prossimi 10 giorni, nel 95 % dei casi, la perdita massima che si può subire è di 300 €. Quindi è facilmente intuibile che per affrontare le possibili perdite che potrebbero presentarsi, l'individuo dovrà dotarsi di un capitale pari o superiore a 300 €.

**Osservazione 1.** Indichiamo con  $\mu$  la media della distribuzione di perdita. Si definisce il VaR medio

$$VaR_{\alpha}^{mean} := VaR_{\alpha} - \mu.$$

Se l'orizzonte temporale  $\Delta$  è un giorno, allora ci si riferisce al VaR medio come "*guadagni giornalieri a rischio*". La distinzione tra il VaR ordinario e il VaR medio è di piccola rilevanza nel rischio di mercato in cui l'orizzonte temporale è molto breve e la media  $\mu$  è vicina allo zero. Comincia ad essere rilevante nel caso del rischio di credito in cui l'arco temporale è più lungo.

**Definizione 16 (Inversa generalizzata e funzione quantile).**

(i) Data una funzione crescente  $T : \mathbb{R} \rightarrow \mathbb{R}$ , l'inversa generalizzata di  $T$  è definita da  $T^{\leftarrow}(y) := \inf\{x \in \mathbb{R} : T(x) \geq y\}$ .

(ii) Data una funzione di ripartizione  $F$ , l'inversa generalizzata  $F^{\leftarrow}$  è detta funzione quantile di  $F$ . Fissato  $\alpha \in (0,1)$ , l' $\alpha$ -quantile di  $F$  è dato da

$$q_{\alpha}(F) := F^{\leftarrow}(\alpha) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}. \quad (2.3)$$

Per una variabile casuale  $X$  con funzione di ripartizione  $F$  usiamo la notazione alternativa  $q_\alpha(X) := q_\alpha(F)$ . Se  $F$  è continua e strettamente crescente, abbiamo semplicemente che

$$q_\alpha(F) = F_X^{-1}(\alpha), \quad (2.4)$$

ovvero l'unico numero reale  $q$  tale che  $F(q) = P(L \leq q) = \alpha$ , mentre  $F^{-1}$  è l'inversa di  $F$ . In termini probabilistici, quindi, il VaR non è altro che il quantile della distribuzione della perdita  $L$ .

**Proposizione 3.** *Il quantile gode della proprietà di linearità:*

$$q_\alpha(aX + b) = aq_\alpha(X) + b. \quad (2.5)$$

*Dimostrazione.* Basta ricordare la (2.4), per cui

$$q_\alpha(aX + b) = F_{aX+b}^{-1}(\alpha).$$

Infatti:

$$\begin{aligned} q_\alpha(aX + b) = aq_\alpha(X) + b &\Leftrightarrow F_{aX+b}^{-1}(\alpha) = aF_X^{-1}(\alpha) + b \\ &\Leftrightarrow F_{aX+b}(F_{aX+b}^{-1}(\alpha)) = F_{aX+b}(aF_X^{-1}(\alpha) + b) \\ &\Leftrightarrow \alpha = P(aX + b \leq aF_X^{-1}(\alpha) + b) \\ &\Leftrightarrow \alpha = P(X \leq F_X^{-1}(\alpha)) \\ &\Leftrightarrow \alpha = F_X(F_X^{-1}(\alpha)). \end{aligned}$$

□

Dalla definizione del VaR come  $\alpha$ -quantile segue che tale misura di rischio è invariante per traslazione, omogenea positiva e monotona. L'invarianza per traslazione e l'omogeneità positiva discendono direttamente dalla precedente proposizione. La monotonia, invece, deriva dall'osservare che, se  $L_1 \leq L_2$  allora

$$P(L_1 \leq l) \geq P(L_2 \leq l) \quad \forall l.$$

Infatti

$$P(L_1 \leq l) = F_{L_1}(l) \geq F_{L_2}(l) = P(L_2 \leq l),$$

ma, dalla definizione del VaR come  $\alpha$ -quantile, si ha che  $F_{L_1}(l) = (q_\alpha(L_1))^{-1}$  e  $F_{L_2}(l) = (q_\alpha(L_2))^{-1}$  da cui

$$(q_\alpha(L_1))^{-1} \geq (q_\alpha(L_2))^{-1}.$$

Invertendo la relazione, si ottiene quindi

$$q_\alpha(L_1) \leq q_\alpha(L_2)$$

da cui la monotonia del VaR.

**Lemma 1.** *Un punto  $x_0 \in \mathbb{R}$  è l' $\alpha$ -quantile di una qualche funzione di ripartizione  $F$  se e solo se le seguenti due condizioni sono soddisfatte:*

1.  $F(x_0) \geq \alpha$ ;
2.  $F(x) < \alpha \quad \forall x < x_0$ .

**Esempio 1** (*Calcolo del VaR per una distribuzione normale*). Supponiamo che la variabile  $L$  con funzione di ripartizione  $F_L$  segua una distribuzione normale con media  $\mu$  e varianza  $\sigma^2$ , cioè  $L \sim \mathcal{N}(\mu, \sigma^2)$ . La funzione densità è data da

$$f(l) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(l-\mu)^2}{2\sigma^2}\right\}$$

da cui integrando si ottiene la funzione di ripartizione. Allora, dato un certo livello di confidenza  $\alpha \in (0,1)$  si ha

$$\text{VaR}_\alpha = \mu + \sigma\Phi^{-1}(\alpha) \quad (2.6)$$

dove  $\Phi$  denota la funzione di ripartizione di una normale standard e  $\Phi^{-1}(\alpha) = F_Z^{-1}(\alpha)$  è l' $\alpha$ -quantile di  $\Phi$  con  $Z \sim \mathcal{N}(0,1)$ . Dal momento che  $F_L$  è strettamente crescente, dal precedente lemma dobbiamo solo mostrare che  $F_L(\text{VaR}_\alpha) = \alpha$ . Quindi:

$$F_L(\text{VaR}_\alpha) = P(L \leq \text{VaR}_\alpha) = P\left(\frac{L-\mu}{\sigma} \leq \frac{\text{VaR}_\alpha - \mu}{\sigma}\right)$$

avendo standardizzato la variabile  $L$ , di modo che  $Z = \frac{L-\mu}{\sigma} \sim \mathcal{N}(0,1)$ . Essendo  $\frac{\text{VaR}_\alpha - \mu}{\sigma} = q_\alpha\left(\frac{L-\mu}{\sigma}\right) = q_\alpha(Z) = \Phi^{-1}(\alpha)$  si ottiene

$$F_L(\text{VaR}_\alpha) = P(Z \leq \Phi^{-1}(\alpha)) = \Phi(\Phi^{-1}(\alpha)) = \alpha.$$

da cui la (2.6).

### 2.2.1 Le limitazioni del VaR

Il VaR è stato ampiamente criticato come misura di rischio sulla base del fatto che ha scarse proprietà di aggregazione. Tale critica trova le sue origini nel lavoro di [4], i quali hanno mostrato che il VaR non è una misura di rischio coerente dal momento che viola la proprietà di subadditività che una misura di rischio dovrebbe avere. La non-subadditività significa che se noi abbiamo due distribuzioni di perdita  $L_1$  ed  $L_2$  con funzioni di ripartizione  $F_{L_1}$  e  $F_{L_2}$  e indichiamo la distribuzione di perdita del portafoglio misto con  $L = L_1 + L_2$  con funzione di ripartizione  $F_L$ , allora non necessariamente si ha che  $q_\alpha(F_L) \leq q_\alpha(F_{L_1}) + q_\alpha(F_{L_2})$ , cioè il VaR del portafoglio misto non è limitato superiormente dai VaR dei singoli portafogli. Questa proprietà è condizionata dal fatto che solitamente qualsiasi portafoglio risente, in maniera più o meno ampia, dell'effetto della diversificazione. Un portafoglio infatti, possiede al suo interno un numero elevato di titoli. Investendo in tale portafoglio, si potrebbe pensare che la misura del rischio dipenda solo da quella dei singoli titoli e da quanto si investe su ciascuno di essi. In realtà non è così: entra in gioco un altro fattore, la correlazione, cioè il modo in cui i rendimenti dei titoli si influenzano l'un l'altro. Tra questi ve ne saranno alcuni che possiedono una correlazione

molto bassa, se non negativa. Ciò implica che, se un titolo registra un rendimento negativo, il secondo, correlato negativamente con il primo, registrerà invece un rendimento positivo, così da eliminare la perdita prodotta dal primo titolo. Questa caratteristica permette di ridurre il rischio complessivo del portafoglio. Quindi, se si decide di aggiungere un titolo al portafoglio, secondo questa proprietà, il rischio del nuovo portafoglio non può essere più elevato della somma dei singoli titoli (vecchio e nuovo) che lo compongono. Tale assioma, nel momento in cui viene calcolato il VaR, viene rispettato solo se sussistono determinate condizioni e ipotesi.

Per esempio, il VaR è subadditivo se le variabili sono indipendenti e distribuite normalmente. Siano  $L_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$  e  $L_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ . Come è ben noto, la somma di due variabili normali indipendenti, è una variabile normale avente come valor medio la somma dei valori medi e come varianza la somma delle varianze, cioè  $L_1 + L_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ . Facendo riferimento all'esempio 1 si ha:

$$VaR_\alpha(L_1) = \mu_1 + \sigma_1 \Phi^{-1}(\alpha)$$

e

$$VaR_\alpha(L_2) = \mu_2 + \sigma_2 \Phi^{-1}(\alpha)$$

da cui

$$VaR_\alpha(L_1 + L_2) = (\mu_1 + \mu_2) + (\sigma_1^2 + \sigma_2^2)^{\frac{1}{2}} \Phi^{-1}(\alpha).$$

Quindi

$$VaR_\alpha(L_1) + VaR_\alpha(L_2) - VaR_\alpha(L_1 + L_2) = \Phi^{-1}(\alpha)[(\sigma_1 + \sigma_2) - (\sigma_1^2 + \sigma_2^2)^{\frac{1}{2}}] > 0$$

da cui la subadditività.

Mostriamo ora, attraverso un esempio preso da [2], che il VaR non è una misura di rischio subadditiva e quindi non coerente.

**Esempio 2 (Non coerenza del VaR).** Si consideri un portafoglio costituito da  $d = 50$  obbligazioni insolubili. Assumiamo che i defaults delle differenti obbligazioni siano indipendenti. La probabilità di default è la stessa per tutte le obbligazioni ed è uguale al 2%. Il valore di rimborso è pari a 100. Se il default non si verifica, tale ammontare viene rimborsato al tempo  $T = t + \Delta t$ , altrimenti il denaro non viene restituito. Al tempo corrente  $t$  il prezzo di ciascuna obbligazione è pari a 95. Quindi la perdita dell'obbligazione  $i$ , con  $i = 1, \dots, 50$ , è descritta dalla seguente variabile aleatoria:

$$L_i := -(100(1 - Y_i) - 95) = 100Y_i - 5$$

dove  $100(1 - Y_i)$  rappresenta il valore del portafoglio al tempo  $T$ , mentre 95 al tempo  $t$ .  $Y_i$  è un indicatore di default che vale 1 se il default si presenta e 0 altrimenti. Quindi  $(L_i)_{1 \leq i \leq 50}$  forma una sequenza di variabili casuali iid con

$$\begin{cases} P(L_i = -5) = P(Y_i = 0) = 0,98 \\ P(L_i = 95) = P(Y_i = 1) = 0,02. \end{cases}$$



Ora vogliamo confrontare due portafogli A e B. Il primo è completamente concentrato e consiste in 100 unità dell'obbligazione 1. Il secondo, invece, è diversificato ed è composto da 2 unità di ciascuna obbligazione. Andiamo a calcolare il VaR al livello di confidenza  $\alpha = 0,95$  per entrambi i portafogli. Per il portafoglio A la perdita è data da  $L = 100L_1$  da cui:

$$VaR_{0,95}(L) = 100VaR_{0,95}(L_1) = 100VaR_{0,95}(100Y_1 - 5) = -500$$

perchè la  $P(L_1 \leq -5) = P(Y_1 = 0) = 0,98 > 0,95$ , mentre la  $P(L_1 \leq l) = 0 < 0,95$  per ogni  $l < -5$ . Questo significa che anche dopo un prelievo di 500, il portafoglio è ancora accettabile lavorando con un VaR al livello di confidenza del 95%. Per il portafoglio B, invece, la perdita è descritta dalla variabile casuale

$$L = \sum_{i=1}^{50} 2L_i = \sum_{i=1}^{50} (200Y_i - 10) = 200 \sum_{i=1}^{50} Y_i - 500$$

da cui

$$VaR_{0,95}(L) = 200VaR_{0,95}\left(\sum_{i=1}^{50} Y_i\right) - 500.$$

La somma  $M = \sum_{i=1}^{50} Y_i$  ha una distribuzione binomiale con probabilità di successo  $p = 0,02$ . Si può calcolare numericamente che  $VaR_{0,95}(M) = q_\alpha(M) = 3$ , poichè  $P(M \leq 3) \approx 0,98 > 0,95$ , mentre  $P(M \leq 2) \approx 0,92 < 0,95$ . Quindi si ottiene:

$$VaR_{0,95}(L) = 600 - 500 = 100.$$

In questo caso la banca ha bisogno di un rischio capitale aggiuntivo di 100. Abbiamo mostrato che il VaR del portafoglio A è minore di quello del portafoglio B nonostante un'intuizione economica suggerisca il contrario dal momento che il portafoglio B è completamente diversificato. L'insensatezza del risultato è dovuta alla mancanza di subadditività del VaR. Infatti, per ogni misura di rischio coerente  $\varphi$ , che dipende esclusivamente dalla distribuzione di  $L$ , dovremmo avere:

$$\varphi\left(\sum_{i=1}^{50} 2L_i\right) \leq \sum_{i=1}^{50} 2\varphi(L_i) \leq \varphi(L_1) \sum_{i=1}^{50} 2 = 100\varphi(L_1) = \varphi(100L_1).$$

Quindi, il fatto che il  $VaR_{0,95}$  sia più piccolo per il portafoglio A piuttosto che per il portafoglio B mostra che, in generale, non è subadditivo.

La seconda limitazione riguarda il fatto che si tratta di una misura che non tiene conto della dimensione delle perdite. Infatti, nonostante permetta di calcolare la probabilità di perdita massima potenziale in un orizzonte temporale fissato in relazione ad un certo livello di confidenza, non fornisce alcuna informazione sulla perdita che potrebbe verificarsi nel restante intervallo di casi non presi in considerazione da  $\alpha$ .

È noto, per esempio, che la perdita massima potenziale che si può subire nei prossimi 10 giorni, al 95%, è di 300 €. Ma nel restante 5% dei casi, la perdita di che entità potrebbe essere? Il VaR non risponde a questa domanda. Nell'eventualità che si verifichi quel 5% di probabilità non tenuta in considerazione dal VaR, e la perdita subita è di 400 € anziché 300 €, tale circostanza può ancora essere accettata, ma se la perdita si attesta a 5000 € le cose cambiano.

Si parla di "tail effect" nel caso in cui le perdite oltre il VaR, cioè quelle che possono concentrarsi nelle code delle funzioni di ripartizione, sono significative. Graficamente questa mancanza di informazione del VaR si traduce nei tre grafici sottostanti, presi da [1]:

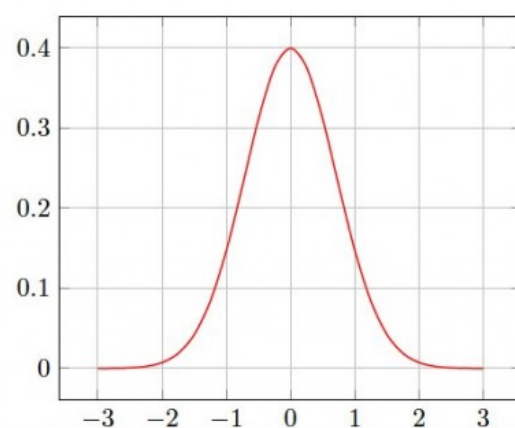


Figura 2.2: assenza di "tail effect"

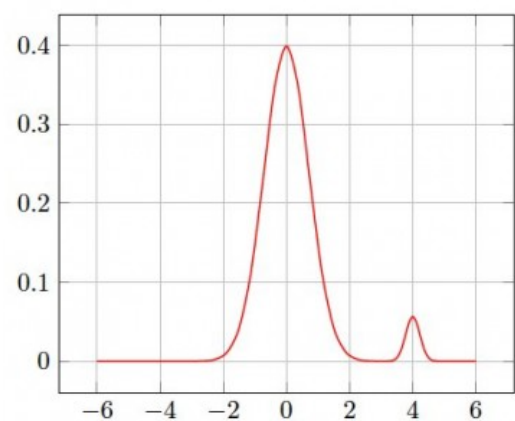


Figura 2.3: "tail effect" moderato

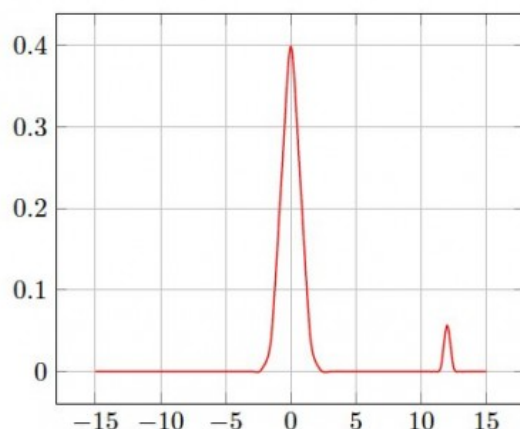


Figura 2.4: "tail effect" estremo

Quello che il VaR non dice è cosa aspettarsi "mediamente" da quelle perdite che superano tale valore. Nei tre scenari raffigurati magari il VaR con un certo livello di confidenza è lo stesso, ma tali valori medi sono molto diversi. Nel primo caso la perdita oltre il VaR è distribuita uniformemente su tutti i valori, nel secondo invece osserviamo una concentrazione di probabilità intorno al valore 4 e nel terzo e ultimo scenario intorno al valore 12. Quindi, utilizzando solo il Value-at-Risk come stima delle perdite inattese e capitale necessario, nel terzo caso ignorerei una pericolosa concentrazione delle perdite oltre il VaR intorno ad un valore elevato. Storicamente, una volta individuati i limiti del VaR, si sono sviluppate delle metodologie di calcolo del rischio che non sono da considerarsi alternative ad esso, ma piuttosto complementari. Noi ci limiteremo solo ad analizzarne una, ovvero l'Expected-Shortfall.

## 2.3 Expected-Shortfall

**Definizione 17 (Expected-Shortfall).** Detto  $VaR_\alpha$  il Value-at-Risk di un portafoglio su un orizzonte temporale fissato ad un livello di confidenza  $\alpha$ , si definisce Expected-Shortfall( $ES_\alpha$ ) o Conditional-VaR( $CVaR_\alpha$ ):

$$ES_\alpha := E(L|L \geq VaR_\alpha) \quad (2.7)$$

dove per l'appunto  $L$  è la variabile aleatoria che indica la perdita del portafoglio. Quindi può essere interpretata come il valore atteso delle perdite eccedenti il VaR. Pertanto si tratta di definire una media condizionata dal fatto che le perdite siano situate nella parte di coda non ricompresa dall'intervallo di confidenza definito dal VaR.

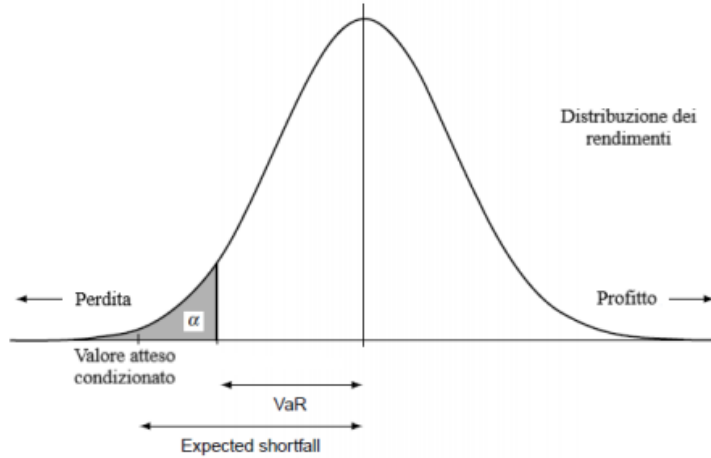


Figura 2.5: Expected-Shortfall

L'Expected-Shortfall è una misura che fornisce una chiara indicazione sull'ammontare di capitale da possedere per fronteggiare eventuali situazioni estreme. Esso si rappresenta anche mediante un'elegante formula analitica riconducibile intuitivamente al teorema della media.

**Proposizione 4.** Per una variabile casuale integrabile  $L$  con funzione di ripartizione  $F_L$  e per ogni  $\alpha \in (0,1)$  abbiamo:

$$ES_\alpha = \frac{1}{1-\alpha} \int_\alpha^1 VaR_\beta d\beta. \quad (2.8)$$

In altri termini, l' $ES_\alpha$  non è altro che la media di tutti i  $VaR_\beta$  con  $\beta \geq \alpha$ .

*Dimostrazione.* Consideriamo la variabile aleatoria  $L$  definita sullo spazio di probabilità  $(\Omega, \mathcal{F}, P)$  e sia  $A \in \mathcal{F}$  tale che  $P(A) \neq 0$ . Supponiamo che esista la funzione  $f_{L|A} : \mathbb{R} \rightarrow \mathbb{R}^+$  tale che

$$F_{L|A}(l|A) = \int_{-\infty}^l f_{L|A}(x|A) dx.$$

In altre parole stiamo assumendo che la distribuzione condizionata di  $L$  all'evento  $A$  ammette densità e  $f_{L|A}$  è proprio la funzione densità di probabilità condizionata. Nel nostro caso specifico assumiamo che  $A = \{x \in \Omega | a < L(x) < b\}$ . Quindi abbiamo che

$$F_{L|A}(l|A) = F_{L|A}(l|a < L \leq b) = P(L \leq l | a < L \leq b) = \frac{P(L \leq l \cap a < L \leq b)}{P(a < L \leq b)}$$

ovvero

$$F_{L|A}(l|a < L \leq b) = \begin{cases} 0 & x < a \\ \frac{F_L(l) - F_L(a)}{F_L(b) - F_L(a)} & x \in [a, b) \\ 1 & x \geq b \end{cases}$$

Poichè  $F_{L|A}$  ammette densità, allora

$$f_{L|a < L \leq b}(l|a < L \leq b) = \frac{dF_{L|A}}{dl}(l|A) = \begin{cases} 0 & x \leq a, x > b \\ \frac{f_L(l)}{F_L(b) - F_L(a)} & x \in (a, b] \end{cases}$$

Facendo tendere  $b \rightarrow \infty$ , la precedente espressione diventa:

$$f_{L|a < L}(l|a < L) = \frac{f_L(l)}{1 - F_L(a)}.$$

Detta  $f_L(l)$  la funzione densità di probabilità delle perdite, la (2.7) si può riscrivere come:

$$ES_\alpha = \frac{1}{P(L \geq VaR_\alpha)} \int_{VaR_\alpha}^{\infty} l f_L(l) dl.$$

Essendo la  $P(L \geq VaR_\alpha) = 1 - \alpha$  e facendo un cambio di varibiale chiamando  $l(\beta) = VaR_\beta$  si ha che  $dl(\beta) = \frac{dVaR_\beta}{d\beta} d\beta$ , da cui

$$ES_\alpha = \int_\alpha^1 \frac{VaR_\beta}{1 - \alpha} f_L(VaR_\beta) \frac{dVaR_\beta}{d\beta} d\beta. \quad (2.9)$$

Osservando che  $f_L(VaR_\beta) \frac{dVaR_\beta}{d\beta} = \frac{d}{d\beta} [F_L(VaR_\beta)]$ , possiamo riscrivere la precedente relazione come:

$$ES_\alpha = \int_\alpha^1 \frac{VaR_\beta}{1 - \alpha} \frac{d}{d\beta} [F_L(VaR_\beta)] d\beta \quad (2.10)$$

ma, dal momento che  $F_L(VaR_\beta) = P(L \leq VaR_\beta) = \beta$ , per definizione la (2.10) diventa

$$ES_\alpha = \frac{1}{1 - \alpha} \int_\alpha^1 VaR_\beta d\beta.$$

□

**Lemma 2.** Per una sequenza  $(L_i)_{i \in \mathbb{N}}$  di variabili casuali iid con funzione di ripartizione  $F$ , si ha:

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^{\lfloor n(1-\alpha) \rfloor} L_{i,n}}{n(1-\alpha)} = ES_\alpha$$

dove  $L_{1,n} \geq \dots \geq L_{n,n}$  sono gli ordinamenti statistici di  $L_1, \dots, L_n$  e  $\lfloor n(1-\alpha) \rfloor$  rappresenta il più grande intero che non supera  $n(1-\alpha)$ .

In altre parole, l'Expected-Shortfall al livello di confidenza  $\alpha$  può essere pensato come il limite della media di  $\lfloor n(1-\alpha) \rfloor$  ordinamenti statistici superiori di un campione di dimensione  $n$  distribuzioni di perdita.

**Proposizione 5.** *L'Expected-Shortfall è una misura di rischio coerente.*

*Dimostrazione.* L'invarianza per traslazione, l'omogeneità positiva e la proprietà di monotonia discendono direttamente dalla rappresentazione (2.8) e dalla proprietà della funzione quantile. Infatti, presi  $a, b \in \mathbb{R}$  si ha:

$$\begin{aligned} ES_\alpha(aL + b) &= \frac{1}{1-\alpha} \int_\alpha^1 q_u(aL + b) du = \frac{1}{1-\alpha} \int_\alpha^1 aq_u(L) + b du \\ &= \frac{1}{1-\alpha} a \int_\alpha^1 q_u(L) du + \frac{1}{1-\alpha} \int_\alpha^1 b du = aES_\alpha(L) + b, \end{aligned}$$

da cui l'invarianza per traslazione e l'omogeneità. La monotonia, invece, discende dalla monotonia dell'integrale. Resta da dimostrare la subadditività.

Consideriamo una generica sequenza di variabili casuali  $L_1, \dots, L_n$  con statistiche sugli ordini associate  $L_{1,n} \geq, \dots, \geq L_{n,n}$  e osserviamo che, per ogni  $1 \leq d \leq n$ , abbiamo:

$$\sum_{i=1}^d L_{i,d} = \sup\{L_{i_1} + \dots + L_{i_d} : 1 \leq i_1 < \dots < i_d \leq d\}.$$

Ora consideriamo due variabili casuali  $L$  e  $\tilde{L}$  con funzione di ripartizione congiunta  $F$  e una sequenza di vettori aleatori bivariati iid  $(L_1, \tilde{L}_1), \dots, (L_n, \tilde{L}_n)$  con stessa funzione di ripartizione  $F$ . Scrivendo  $(L + \tilde{L})_i := L_i + \tilde{L}_i$  e  $(L + \tilde{L})_{i,n}$  per l'ordinamento statistico di  $(L + \tilde{L})_1, \dots, (L + \tilde{L})_n$ , si ha:

$$\begin{aligned} \sum_{i=1}^d (L + \tilde{L})_{i,n} &= \sup\{(L + \tilde{L})_{i_1} + \dots + (L + \tilde{L})_{i_d} : 1 \leq i_1 < \dots < i_d \leq d\} \\ &\leq \sup\{L_{i_1} + \dots + L_{i_d} : 1 \leq i_1 < \dots < i_d \leq d\} \\ &\quad + \sup\{\tilde{L}_{i_1} + \dots + \tilde{L}_{i_d} : 1 \leq i_1 < \dots < i_d \leq d\} \\ &= \sum_{i=1}^d L_{i,n} + \sum_{i=1}^d \tilde{L}_{i,n}. \end{aligned}$$

Ponendo  $d = \lceil n(1-\alpha) \rceil$  e mandando  $n \rightarrow \infty$ , dal precedente lemma si ottiene che  $ES_\alpha(L + \tilde{L}) \leq ES_\alpha(L) + ES_\alpha(\tilde{L})$ . □

**Esempio 3** (*Calcolo dell' Expected-Shortfall per una distribuzione gaussiana*).

Supponiamo che la variabile  $L$  con funzione di ripartizione  $F_L$  segua una distribuzione normale con media  $\mu$  e varianza  $\sigma^2$ , cioè  $L \sim \mathcal{N}(\mu, \sigma^2)$ . Fissato  $\alpha \in (0,1)$  abbiamo

$$ES_\alpha = \mu + \sigma \frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha}, \tag{2.11}$$

dove  $\phi$  è la funzione densità della distribuzione normale standard. Anche in questo caso la dimostrazione è elementare. Riscriviamo la definizione di  $ES_\alpha$  andando a standardizzare la variabile  $L$ :

$$\frac{ES_\alpha - \mu}{\sigma} = E\left(\frac{L - \mu}{\sigma} \mid \frac{L - \mu}{\sigma} \geq \frac{VaR_\alpha - \mu}{\sigma}\right)$$

da cui

$$ES_\alpha = \mu + \sigma E\left(\frac{L - \mu}{\sigma} \mid \frac{L - \mu}{\sigma} \geq q_\alpha\left(\frac{L - \mu}{\sigma}\right)\right).$$

Ponendo  $Z = \frac{L - \mu}{\sigma} \sim \mathcal{N}(0,1)$  si ha

$$ES_\alpha(Z) = \mu + \sigma E(Z|Z \geq \Phi^{-1}(\alpha)).$$

Usando la definizione di valore atteso condizionato ad un evento si ottiene:

$$E(Z|Z \geq \Phi^{-1}(\alpha)) = \frac{1}{P(Z \geq \Phi^{-1}(\alpha))} \int_{\Phi^{-1}(\alpha)}^{\infty} z\phi(z) dz.$$

Essendo  $P(Z \geq \Phi^{-1}(\alpha)) = 1 - P(Z < \Phi^{-1}(\alpha)) = 1 - \alpha$  e integrando per parti, si arriva a:

$$E(Z|Z \geq \Phi^{-1}(\alpha)) = \frac{1}{1 - \alpha} [z\Phi(z)]_{\Phi^{-1}(\alpha)}^{\infty} - \frac{1}{1 - \alpha} \int_{\Phi^{-1}(\alpha)}^{\infty} \Phi(z) dz = \frac{1}{1 - \alpha} [-\phi(z)]_{\Phi^{-1}(\alpha)}^{\infty} = \frac{\phi(\Phi^{-1}(\alpha))}{1 - \alpha}$$

da cui la (2.11).





## Capitolo 3

# Mixture models e VaR nei modelli di rischio di credito

In questo capitolo l'obiettivo è quello di mostrare come le misure di rischio introdotte precedentemente possono essere calcolate nella pratica. Il metodo che noi usiamo è fondato sull'osservazione che tutti i modelli standard industriali, quali CreditRisk<sup>+</sup> e CreditMetrics/KMV, possono essere rappresentati come modelli misti di Bernoulli.

Dopo aver introdotto i fondamenti del calcolo del rischio, nella sezione 3.2 verranno definiti i modelli di mistura: i modelli misti di Bernoulli e di Poisson. Nella sezione 3.3 verrà proposto un collegamento con i modelli a variabili latenti e un risultato che mostrerà come ciascuno di essi possa essere scritto come modello misto di Bernoulli.

Nella sezione 3.4 verrà semplificata la notazione considerando modelli ad un fattore in portafogli omogenei e, in particolare, si studieranno i modelli di mistura di cui sopra applicati a questo caso.

Nella sezione 3.5 verranno proposti degli esempi numerici sul calcolo del VaR per un Bernoulli/Poisson mixture model scambiabile, mostrando come tali valori si inseriscono negli intervalli del Value-at-Risk forniti da [7].

Infine, nella sezione 3.6 verrà considerato un modello misto di Bernoulli con variabile latente bidimensionale, ottenendo così un portafoglio costituito da due blocchi omogenei e vedremo che, soddisfatta la rappresentazione come modello di mistura, il vettore è parzialmente scambiabile.

### 3.1 Fondamenti per modellare il rischio di credito

Prendiamo un portafoglio di  $d$  imprese e fissiamo un orizzonte temporale  $[t, t + \Delta t]$ , dove  $\Delta t$  è tipicamente un anno. Come discusso precedentemente, l'insolvenza della compagnia  $i$ -esima,  $i = 1, \dots, d$ , è descritta dall'indicatore di default  $Y_i$ , che è una variabile casuale di Bernoulli in  $\{0,1\}$ . Indichiamo con  $Y_i = 1$  lo stato di default e  $Y_i = 0$  uno stato di crescente qualità del credito. Assumiamo che al tempo  $t$  ciascuna delle compagnie si trovi in uno stato di non default.

Per modellare la perdita di credito in un portafoglio di  $d$  imprese, consideriamo la somma

delle perdite individuali:

$$L_d := \sum_{i=1}^d e_i \cdot l_i \cdot Y_i,$$

dove:

- $l_i \in (0,1]$  rappresenta il tasso di perdita della compagnia  $i$ -esima e  $\sum_{i=1}^d l_i = 1$ ;
- $e_i > 0$  indica l'ammontare dovuto dall'obbligazionista che, senza perdita di generalità, normalizziamo ad 1.

Nel nostro caso assumiamo  $l_i = \frac{1}{d}$ , ovvero lo stesso per tutte le ditte. Per pesi uguali,  $L_d = \frac{S_d}{d}$ , dove:

$$S_d = \sum_{i=1}^d Y_i$$

rappresenta il numero di insolvenze. Quindi, tramite le semplificazioni apportate, per analizzare la perdita complessiva  $L_d$  è sufficiente studiare il comportamento della somma  $S_d$ .

## 3.2 Mixture models

In un modello misto la probabilità di insolvenza di una compagnia si assume che dipenda da un insieme di fattori comuni, che sono interpretati come *variabili macroeconomiche*. Dati questi fattori, i defaults delle diverse imprese sono indipendenti. Il vettore  $\mathbf{Y} = (Y_1, \dots, Y_d)'$  è il vettore casuale degli indicatori di default per il portafoglio nell'arco temporale di interesse, mentre  $p(\mathbf{y}) = P(Y_1 = y_1, \dots, Y_d = y_d)$ , con  $\mathbf{y} \in \{0,1\}^d$ , è la funzione di massa. Le probabilità di default saranno indicate con  $p_i = P(Y_i = 1)$ ,  $i = 1, \dots, d$ . Gli eventi di default saranno definiti dalla correlazione tra loro esistente. Dal momento che la varianza per una variabile di Bernoulli è data da:

$$\text{var}(Y_i) = E(Y_i^2) - E(Y_i)^2 = E(Y_i) - p_i^2 = p_i - p_i^2$$

si ottiene che, prese le compagnie  $i$  e  $j$ , con  $i \neq j$ , il coefficiente di correlazione è dato da:

$$\begin{aligned} \rho(Y_i, Y_j) &= \frac{\text{cov}(Y_i, Y_j)}{\sqrt{\text{var}(Y_i)\text{var}(Y_j)}} = \frac{E(Y_i Y_j) - E(Y_i)E(Y_j)}{\sqrt{\text{var}(Y_i)\text{var}(Y_j)}} \\ &= \frac{E(Y_i Y_j) - p_i p_j}{\sqrt{(p_i - p_i^2)(p_j - p_j^2)}}. \end{aligned} \quad (3.1)$$

Come già spiegato nel capitolo 2, la correlazione tra i diversi investimenti delle  $d$  imprese che stiamo considerando rappresenta il legame tra loro esistente. Di fatto, può assumere un valore compreso tra -1 e 1. Nel primo caso, due attività si muoveranno in senso opposto, mentre nel secondo la relazione sarà di massima sintonia. Nel mezzo vi sono tutte le sfumature. L'obiettivo non è quello di costruire portafogli che presentino un buon bilanciamento tra i diversi livelli di correlazione. Infatti, si considera un portafoglio già dato e noti diversi livelli di correlazione e la probabilità di insolvenza, se ne calcolerà la misura di rischio.

**Definizione 18 (Bernoulli mixture model).** Dato  $h < d$  e un vettore casuale  $h$ -dimensionale  $\Psi = (\Psi_1, \dots, \Psi_h)$ , il vettore degli indicatori di default  $\mathbf{Y} = (Y_1, \dots, Y_d)'$  segue un modello misto di Bernoulli con vettore dei fattori comuni  $\Psi$ , se esistono delle funzioni  $Q_i : \mathbb{R}^h \rightarrow [0,1]$ ,  $1 \leq i \leq d$ , tali che, condizionatamente a  $\Psi$ ,  $\mathbf{Y}$  sia un vettore di variabili casuali indipendenti di Bernoulli con  $P(Y_i = 1 | \Psi) = Q_i(\Psi)$ .

Per  $\mathbf{y} = (y_1, \dots, y_d)'$  in  $\{0,1\}^d$  abbiamo che:

$$P(\mathbf{Y} = \mathbf{y} | \Psi) = \prod_{i=1}^d Q_i(\Psi)^{y_i} (1 - Q_i(\Psi))^{1-y_i} \quad (3.2)$$

e la distribuzione non condizionata di  $\mathbf{Y}$  si ottiene integrando la legge del vettore  $\Psi$ . In altre parole,  $Y_i | \Psi \sim \mathcal{B}(Q_i(\Psi))$ , dove, dalla definizione, le  $Q_i$  sono variabili aleatorie indipendenti che assumono valori in  $\{0,1\}$ . In particolare, la probabilità di default dell'impresa  $i$ -esima è data da:

$$\begin{aligned} p_i &= P(Y_i = 1) = E(Y_i) = E(E(Y_i | \Psi)) \\ &= E(1 \cdot P(Y_i = 1 | \Psi) + 0 \cdot P(Y_i = 0 | \Psi)) = E(P(Y_i = 1 | \Psi)) = E(Q_i(\Psi)). \end{aligned} \quad (3.3)$$

Dal momento che il default è un evento raro, analizziamo la possibilità di approssimare il numero di insolvenze  $S_d$ , che condizionatamente a  $\Psi$  è una binomiale, con variabili aleatorie di Poisson. Utilizziamo la notazione  $\tilde{Y}_i \in \{0,1,2, \dots\}$  per contare le variabili che danno il numero di insolvenze dell'impresa  $i$ -esima.

**Osservazione 2.** In generale, la distribuzione di Poisson può essere utilizzata per approssimare una distribuzione binomiale di parametri  $n$  e  $p$ , quando il numero  $n$  delle prove è grande e la probabilità di successo  $p$  è piccola. Nel caso in esame, le prove corrispondono al numero delle compagnie  $d$  e la probabilità di successo è rappresentata dalla probabilità di insolvenza della compagnia  $i$ -esima. Si può infatti dimostrare che la distribuzione di Poisson è il limite della distribuzione binomiale per  $d \rightarrow \infty$  e con  $\lambda = dp$ :

---

<sup>1</sup>Per tutte le variabili aleatorie  $X$  e  $Y$ , con  $X$  avente valore atteso finito, si ha:  $E(X) = E(E(X|Y))$ .

$$\lim_{d \rightarrow \infty} \binom{d}{k} p^k (1-p)^{d-k} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Tale risultato è di estrema importanza non solo per la definizione di seguito, ma soprattutto perchè permetterà di confrontare successivamente il calcolo del VaR in un modello misto di Bernoulli con un modello misto di Poisson.

**Definizione 19 (Poisson mixture model).** Siano  $h$  e  $\Psi$  come nella definizione 18. Il vettore casuale  $\tilde{\mathbf{Y}} = (\tilde{Y}_1, \dots, \tilde{Y}_d)'$  segue il modello misto di Poisson con vettore dei fattori comuni  $\Psi$ , se esistono delle funzioni  $\Lambda_i : \mathbb{R}^h \rightarrow (0, \infty)$  tali che, condizionatamente a  $\Psi$ , il vettore aleatorio  $\tilde{\mathbf{Y}}$  è un vettore di variabili casuali di Poisson indipendenti di parametro  $\Lambda_i(\Psi)$ .

Per  $\tilde{\mathbf{y}} = (\tilde{y}_1, \dots, \tilde{y}_d)'$  in  $\{0, 1, 2, \dots\}^d$  abbiamo che:

$$P(\tilde{\mathbf{Y}} = \tilde{\mathbf{y}} | \Psi) = \prod_{i=1}^d \exp(-\Lambda_i(\Psi)) \frac{\Lambda_i(\Psi)^{\tilde{y}_i}}{\tilde{y}_i!}$$

e la distribuzione non condizionata di  $\tilde{\mathbf{Y}}$  si ottiene integrando sulla legge del vettore  $\Psi$ . In altre parole  $\tilde{Y}_i | \Psi \sim \mathcal{P}(\Lambda_i(\Psi))$ . Se  $\tilde{\mathbf{Y}}$  segue un modello misto di Poisson e definiamo gli indicatori  $Y_i := I_{\{\tilde{Y}_i \geq 1\}}$ , cioè:

$$Y_i = \begin{cases} 1 & \tilde{Y}_i \geq 1 \\ 0 & \tilde{Y}_i = 0, \end{cases} \quad (3.4)$$

allora  $\mathbf{Y}$  segue un modello misto di Bernoulli e le variabili di mistura sono legate dalla seguente relazione:

$$Q_i(\Psi) = 1 - \exp(-\Lambda_i(\Psi)). \quad (3.5)$$

La (3.5) discende direttamente dalla relazione tra le definizioni 18 e 19. Infatti, dalla (3.4):

$$\begin{aligned} Q_i(\Psi) &= P(Y_i = 1 | \Psi) = P(\tilde{Y}_i \geq 1 | \Psi) = 1 - P(\tilde{Y}_i = 0 | \Psi) \\ &= 1 - \frac{\Lambda_i(\Psi)^0}{0!} \exp(-\Lambda_i(\Psi)) = 1 - \exp(-\Lambda_i(\Psi)), \end{aligned}$$

da cui la (3.5).

Definiamo ora la variabile  $\tilde{S}_d := \sum_{i=1}^d \tilde{Y}_i$  che conta il numero di compagnie insolventi. La sua distribuzione, condizionatamente a  $\Psi$ , soddisfa:

$$P(\tilde{S}_d = k | \Psi) = \exp\left(-\sum_{i=1}^d \Lambda_i(\Psi)\right) \frac{(\sum_{i=1}^d \Lambda_i(\Psi))^k}{k!}. \quad (3.6)$$

Quindi,  $\tilde{S}_d | \Psi \sim \mathcal{P}(\sum_{i=1}^d \Lambda_i(\Psi))$ . Vedremo nel caso di modelli ad un fattore come calcolare la distribuzione non condizionata.

### 3.2.1 CreditRisk<sup>+</sup>

Il CreditRisk<sup>+</sup> è un modello industriale per il rischio di credito. Oggi è usato in modo sempre più intensivo per contenere il pericolo di mancato pagamento totale o parziale dei crediti commerciali. Esso può essere rappresentato come un modello misto di Bernoulli dove la distribuzione degli indicatori di default è data da:

$$P(Y_i = 1|\Psi) = Q_i(\Psi) \quad \& \quad Q_i(\Psi) = 1 - \exp(-\mathbf{w}'_i\Psi).$$

$\Psi = (\Psi_1, \dots, \Psi_h)'$ , con  $h < d$ , è il vettore dei fattori macroeconomici le cui componenti sono distribuite con una distribuzione gamma e  $\mathbf{w}_i = (w_{i1}, \dots, w_{ih})'$  è un vettore di pesi costanti. Chiaramente questo modello ha la forma specificata nella definizione 18. Osserviamo, tuttavia, che il CreditRisk<sup>+</sup> è spesso presentato come modello misto di Poisson, nel senso della definizione 19. In questa più comune rappresentazione si assume che, condizionatamente a  $\Psi$ , il default della compagnia  $i$ , con  $i = 1, \dots, d$ , si presenti indipendentemente dalle altre ditte con un'intensità di Poisson data da:

$$\Lambda_i(\Psi) = \mathbf{w}'_i\Psi,$$

da cui nuovamente la (3.5). La formulazione del CreditRisk<sup>+</sup> come modello misto di Poisson ha la caratteristica che il numero di insolvenze nel portafoglio è uguale alla distribuzione della somma di variabili indipendenti binomiali negative, che vedremo nel seguito.

## 3.3 Relazione con i modelli a variabili latenti

A prima vista, i modelli a variabili latenti( vedi Appendice A) e i modelli misti di Bernoulli sembrano molto diversi. Tuttavia, come è stato già osservato da [6], nel caso speciale del CreditRisk<sup>+</sup> e CreditMetrics si tratta di differenze legate di più alla rappresentazione e interpretazione che alla sostanza matematica.

**Definizione 20 (Modello a variabili latenti).** Un vettore a variabili latenti  $\mathbf{X}$  ha una struttura di indipendenza condizionata  $h$ -dimensionale con variabile condizionante  $\Psi$  se esiste un qualche  $h < d$  e un vettore casuale  $h$ -dimensionale  $\Psi = (\Psi_1, \dots, \Psi_h)'$  tale che, condizionatamente a  $\Psi$ , le variabili aleatorie  $(X_i)_{1 \leq i \leq d}$  siano indipendenti.

La seguente condizione assicura che un modello a variabili latenti possa essere scritto come un modello misto di Bernoulli.

**Proposizione 6.** *Si consideri un vettore a variabili latenti  $\mathbf{X}$   $d$ -dimensionale e un vettore casuale  $\Psi$  con  $h < d$ . Allora le seguenti condizioni sono equivalenti:*

- (i)  $\mathbf{X}$  ha una struttura di indipendenza condizionata  $h$ -dimensionale con variabile condizionante  $\Psi$ ;
- (ii) per ogni scelta delle soglie  $w_{i1}$ , con  $1 \leq i \leq d$ , gli indicatori di default  $Y_i = 1_{\{X_i \leq w_{i1}\}}$  seguono un modello misto di Bernoulli con vettore dei fattori  $\Psi$ . Le probabilità di default condizionate sono date da

$$Q_i(\Psi) = P(X_i \leq w_{i1}|\Psi).$$

*Dimostrazione.* Supponiamo che la (i) valga. Per ogni  $\mathbf{y} \in \{0,1\}^d$  abbiamo:

$$\begin{aligned} P(Y_1 = y_1, \dots, Y_d = y_d | \Psi) &= \\ &= P((-1)^{1-y_1} X_1 < (-1)^{1-y_1} w_{11}, \dots, (-1)^{1-y_d} X_d < (-1)^{1-y_d} w_{d1} | \Psi) \\ &= \prod_{i=1}^d P((-1)^{1-y_i} X_i < (-1)^{1-y_i} w_{i1} | \Psi). \end{aligned}$$

Quindi, condizionatamente a  $\Psi$ , le  $Y_i$  sono variabili causali di Bernoulli indipendenti con  $Q_i(\Psi) := P(X_i < w_{i1} | \Psi)$ .  $\square$

### 3.3.1 CreditMetrics/KMV come modelli misti di Bernoulli

In entrambi i casi, consideriamo la variabile latente  $\mathbf{X} = (X_1, \dots, X_d)'$  con una distribuzione normale multivariata e assumiamo, senza perdita di generalità, che  $\mathbf{X}$  abbia media nulla, cioè  $\mathbf{X} \sim \mathcal{N}_d(\mathbf{0}, \Sigma)$ . Supponiamo inoltre che  $\mathbf{X}$  dipenda da fattori macroeconomici in accordo al *modello a fattori lineari* (vedi [2]). Si assuma che le componenti di  $\mathbf{X}$  possano essere scritte come:

$$X_i = \sum_{j=1}^h a_{ij} \Theta_j + \sigma_i \varepsilon_i \quad i = 1, \dots, d, \quad (3.7)$$

per qualche  $h < d$ , un vettore casuale  $h$ -dimensionale  $\Theta \sim N_h(\mathbf{0}, \Omega)$  e delle variabili aleatorie iid  $\varepsilon_1, \dots, \varepsilon_d$  che seguono una distribuzione normale standard e che sono a loro volta indipendenti da  $\Theta$ . Infine gli elementi  $a_{ij}$  sono i pesi di una matrice  $A \in \mathbb{R}^{d \times h}$  la cui calibrazione è ottenuta usando argomenti economici combinati con analisi statistiche sui rendimenti delle attività. La matrice di covarianza, sempre dal modello a fattori lineari, è della forma:

$$\Sigma = \text{cov}(\mathbf{X}) = A\Omega A' + \text{diag}(\sigma_1^2, \dots, \sigma_d^2).$$

Definiamo ora il vettore delle soglie  $\mathbf{w}_i := (w_{11}, w_{21}, \dots, w_{d1})$ . Il default della compagnia  $i$  viene modellato come occorrenze, cioè:

$$Y_i = 1 \iff X_i \leq w_{i1} \quad 1 \leq i \leq d. \quad (3.8)$$

Semplici calcoli mostrano che questa costruzione definisce un modello misto di Bernoulli. Ponendo  $\Theta = \Psi$  e usando (3.7) e (3.8) si ottiene che, condizionatamente a  $\Psi$ , le  $Y_i$  sono indipendenti dal momento che le  $\varepsilon_i$  sono iid. Quindi, in accordo alla proposizione 6:

$$Q_i(\Psi) = P(Y_i = 1 | \Psi) = P(X_i \leq w_{i1} | \Psi) = P(\mathbf{a}'_i \Theta + \sigma_i \varepsilon_i \leq w_{i1} | \Psi)$$

dove nell'ultima uguaglianza abbiamo sostituito la definizione di  $X_i$  riscritta in forma compatta, cioè  $X_i = \mathbf{a}'_i \Theta + \sigma_i \varepsilon_i$ , considerando così la prima riga della matrice  $A$  e facendo variare l'indice  $j = 1, \dots, h$  in accordo al modello soglia. Quindi:

$$P(Y_i = 1 | \Psi) = P\left(\varepsilon_i \leq \left(\frac{w_{i1} - \mathbf{a}'_i \Psi}{\sigma_i}\right) | \Psi\right) = \Phi\left(\left(\frac{w_{i1} - \mathbf{a}'_i \Psi}{\sigma_i}\right)\right) \quad (3.9)$$

dove  $\Phi$  è la funzione di ripartizione della normale standard.

## 3.4 Modelli scambiabili ad un fattore in portafogli omogenei

Nella loro forma più generale, i modelli sopra descritti sono completamente eterogenei con differenti probabilità di default e differenti correlazioni tra le controparti. Per evidenziare dal punto di vista matematico le differenze tra i modelli è utile ricondursi al caso di portafogli omogenei. Inoltre, i portafogli eterogenei sono estremamente difficili da calibrare ed è pratica comune segmentare grandi portafogli in un piccolo numero di gruppi omogenei. Per semplificare l'analisi, assumeremo che il vettore degli indicatori di default  $\mathbf{Y}$  sia *scambiabile*.

**Definizione 21 (Scambiabilità).** Consideriamo una sequenza di variabili casuali  $(Y_1, \dots, Y_d)$  definita sullo spazio di probabilità  $(\Omega, \mathcal{F}, P)$ . Tale sequenza si dice scambiabile se, per ogni permutazione  $(\sigma(1), \dots, \sigma(d))$  di  $\{1, \dots, d\}$ , le distribuzioni di  $(Y_1, \dots, Y_d)$  e  $(Y_{\sigma(1)}, \dots, Y_{\sigma(d)})$  coincidono. Formalmente:

$$P(Y_1 = y_1, \dots, Y_d = y_d) = P(Y_{\sigma(1)} = y_{\sigma(1)}, \dots, Y_{\sigma(d)} = y_{\sigma(d)}). \quad (3.10)$$

Una sequenza infinita  $\{Y_i\}_{i \geq 1}$  si dice scambiabile se le sequenze finite sono scambiabili per ogni  $d \in \mathbb{N}$ .

Quindi, qualsiasi valore di una sequenza finita o infinita è probabile quanto qualsiasi permutazione di quei valori, cioè la distribuzione di probabilità congiunta è invariante sotto il gruppo simmetrico.

### 3.4.1 Modelli ad un fattore

In alcune situazioni pratiche è utile considerare modelli ad un fattore, in cui gli indicatori  $Y_i$  dipendono da una sola variabile mista. Il motivo non è solo legato alla maggiore semplicità nei calcoli, ma è anche utile per calibrare i modelli che, in più fattori, presentano mancanza di informazioni. Sia  $\Psi$  una variabile aleatoria a valori in  $\mathbb{R}$ , e  $Q_i(\Psi) : \mathbb{R} \rightarrow [0,1]$ ,  $\Lambda_i(\Psi) : \mathbb{R} \rightarrow (0, \infty)$  sono funzioni tali che, condizionatamente a  $\Psi$ , il vettore degli indicatori di default  $\mathbf{Y}(\Psi)$  sia un vettore di variabili casuali indipendenti di Bernoulli(Poisson) con

$$P(Y_i = 1 | \Psi) = Q_i(\Psi) \quad P(\tilde{Y}_i = k | \Psi) = \exp(-\Lambda_i(\Psi)) \frac{\Lambda_i(\Psi)^k}{k!}. \quad (3.11)$$

### 3.4.2 Bernoulli e Poisson mixture models scambiabili

Un modello misto di Bernoulli scambiabile si ottiene nella (3.11) quando le funzioni  $Q_i$  sono identiche, dal momento che il vettore casuale  $\mathbf{Y}$  è scambiabile. È conveniente introdurre la variabile aleatoria  $Q := Q_i(\Psi)$ , per ogni  $i = 1, \dots, d$ , e indichiamo la funzione di ripartizione con  $G(q)$ . Quindi, nel caso di modelli ad un fattore in portafogli omogenei la probabilità di insolvenza marginale, condizionata a  $Q = q$ , è data da:

$$P(Y_i = 1 | Q = q) = q,$$

mentre la somma  $S_d$  segue una distribuzione binomiale di parametri  $d$  e  $q$ , cioè:

$$P(S_d = k|Q = q) = \binom{d}{k} q^k (1 - q)^{d-k} \quad k \geq 0. \quad (3.12)$$

Per ottenere la distribuzione non condizionata, integriamo sulla legge di  $Q$ :

$$P(S_d = k) = \binom{d}{k} \int_0^1 q^k (1 - q)^{d-k} dG(q). \quad (3.13)$$

Analogamente, un modello misto di Poisson scambiabile si ottiene quando le funzioni  $\Lambda_i$  sono identiche, dal momento che il vettore aleatorio  $\tilde{\mathbf{Y}}$  è scambiabile. Anche in questo caso è conveniente introdurre la variabile  $\Lambda := \Lambda_i(\Psi)$ , per ogni  $i$ , e indichiamo la legge associata con  $G(\lambda)$ . Quindi, la probabilità di insolvenza marginale, condizionata a  $\Lambda = \lambda$ , è data da:

$$P(\tilde{Y}_i = k|\Lambda = \lambda) = e^{-\lambda} \frac{\lambda^k}{k!},$$

cioè  $\tilde{Y}_i|\Lambda = \lambda \sim \mathcal{P}(\lambda)$ . La distribuzione condizionata della somma, invece, dal momento che i margini  $\Lambda_i$  sono tutti uguali e quindi nella (3.6) possiamo sommarli, sarà:

$$P(\tilde{S}_d = k|\Lambda = \lambda) = e^{-d\lambda} \frac{(d\lambda)^k}{k!} \quad k \geq 0, \quad (3.14)$$

cioè  $\tilde{S}_d|\Lambda = \lambda \sim \mathcal{P}(d\lambda)$ . Per ottenere la distribuzione non condizionata, integriamo sulla legge di  $\Lambda$ :

$$P(\tilde{S}_d = k) = \int_0^\infty P(\tilde{S}_d = k|\Lambda = \lambda) dG(\lambda) = \int_0^\infty \frac{(d\lambda)^k}{k!} e^{-d\lambda} dG(\lambda). \quad (3.15)$$

Quindi, nel modello misto di Poisson, a differenza di quello di Bernoulli in cui la variabile mista  $Q$  nel passaggio dalla marginale  $Y_i$  alla somma  $S_d$  rimane la stessa, qui bisogna fare attenzione ad un particolare: nel passaggio dalla marginale alla somma il parametro diventa  $d\lambda$ , che dunque compare nell'integrale, mentre si condiziona sempre rispetto alla marginale di parametro  $\lambda$ . Questa particolarità sarà fondamentale nell'esempio che vedremo di seguito.

### 3.4.3 Correlazione

Nel modello di Bernoulli scambiabile introduciamo la notazione  $\pi$  per indicare la probabilità che la compagnia  $i$ , con  $i = 1, \dots, d$ , fallisca. Quindi:

$$\pi = P(Y_i = 1) = E(Y_i) = E(Q)$$

avendo usato la (3.3) nella versione dei modelli ad un fattore.

Allo stesso modo:

$$\pi_k := P(Y_1 = 1, \dots, Y_k = 1) = E(E(Y_1 \cdot \dots \cdot Y_k|Q)).$$



Infatti, dal momento che le  $Y_i$ , condizionate a  $Q$ , sono indipendenti, si ha:

$$\begin{aligned}\pi_k &= E(E(Y_1|Q) \cdot E(Y_2|Q) \cdot \dots \cdot E(Y_k|Q)) = E(P(Y_1 = 1|Q) \cdot P(Y_2 = 1|Q) \cdot \dots \cdot P(Y_k = 1|Q)) \\ &= E(\underbrace{Q \cdot Q \cdot \dots \cdot Q}_{k \text{ volte}}) = E(Q^k).\end{aligned}\quad (3.16)$$

Quindi, il coefficiente di correlazione (3.1) diventa:

$$\rho_Y := \rho(Y_i, Y_j) = \frac{\pi_2 - \pi^2}{\pi - \pi^2}, \quad i \neq j \quad (3.17)$$

con  $\pi_2 = E(Y_i Y_j)$  momento di ordine 2. Possiamo osservare che si tratta di una semplice funzione del primo e secondo ordine delle probabilità di default. Come conseguenza, in un modello misto di Bernoulli ad un fattore scambiabile si ottiene:

$$\rho_Y = \frac{\pi_2 - \pi^2}{\pi - \pi^2} = \frac{E(Q^2) - E(Q)^2}{E(Q) - E(Q)^2} = \frac{\text{var}(Q)}{E(Q)(1 - E(Q))} \geq 0.$$

Allo stesso modo, possiamo eseguire calcoli uguali per un modello misto di Poisson scambiabile ottenendo:

$$\begin{aligned}\tilde{\pi} &= P(\tilde{Y}_i \geq 1) = P(Y_i = 1) \\ &= E(Q) = E(1 - \exp(-\Lambda)) \approx E(1 - (1 - \Lambda)) = E(\Lambda)\end{aligned}$$

avendo utilizzato lo sviluppo di Taylor dell' esponenziale e la relazione tra i due indicatori  $Y_i$  e  $\tilde{Y}_i$ . Analogamente, dalla (3.16):

$$\tilde{\pi}_k = P(\tilde{Y}_1 \geq 1, \dots, \tilde{Y}_k \geq 1) = P(Y_1 = 1, \dots, Y_k = 1) \approx E(\Lambda^k).$$

### 3.4.4 CreditRisk<sup>+</sup> con distribuzione beta-binomiale

La distribuzione di  $Q$  del CreditRisk<sup>+</sup> nella versione ad un fattore, per portafogli omogenei, è data dalla (3.5) con le variabili di mistura tutte uguali dove  $\Lambda \sim \text{Gamma}(a, b)$  ha funzione densità:

$$f(\lambda) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp(-b\lambda) \quad (3.18)$$

e  $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$  è la funzione Gamma di Eulero.

Dapprima dimostreremo che questa versione semplificata del CreditRisk<sup>+</sup> è equivalente a modellare la variabile mista  $Q$  con una distribuzione Beta di parametri  $a$  e  $b$ , e poi vedremo come svolgere gli stessi conti con la distribuzione Gamma di partenza. Infatti, in generale, se  $Q \sim \text{Beta}(a, b)$  allora ha densità:

$$f(q) = \frac{q^{a-1}(1-q)^{b-1}}{\beta(a,b)} \quad (3.19)$$

dove

$$\beta(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1} dx$$

è la funzione Beta. Il collegamento tra la densità appena scritta e la (3.18) si ottiene ponendo  $g(q) = \frac{f(-\ln(1-q))}{(1-q)}$ . Andando a sostituire si ha:

$$g(q) = \frac{\frac{b^a}{\Gamma(a)}(-\ln(1-q))^{a-1} \exp(-b(-\ln(1-q)))}{(1-q)}$$

da cui

$$g(q) = \frac{\frac{b^a}{\Gamma(a)}(-\ln(1-q))^{a-1}(1-q)^b}{(1-q)}.$$

Per  $q$  piccoli possiamo usare l'approssimazione  $-\ln(1-q) \approx q$  e si arriva così a:

$$g(q) = \frac{b^a}{\Gamma(a)} q^{a-1}(1-q)^{b-1} \quad a, b > 0.$$

Tale relazione è estremamente vicina alla distribuzione beta che ha densità (3.19). Questa osservazione è di estrema importanza, dal momento che permette di calcolare il  $\text{VaR}_\alpha$  per il  $\text{CreditRisk}^+$  utilizzando una distribuzione mista beta, una delle più usate nei modelli misti di Bernoulli scambiabili ad un fattore.

#### Esempio 4 (Distribuzione mista beta).

Assumiamo che  $Q \sim \text{Beta}(a, b)$ , per qualche  $a, b > 0$ . La funzione densità è data dalla (3.19). Dalla (3.16) otteniamo:

$$\pi_k = E(Q^k) = \frac{1}{\beta(a,b)} \int_0^1 q^k q^{a-1}(1-q)^{b-1} dq = \frac{\beta(a+k,b)}{\beta(a,b)} \quad k = 1, 2, \dots$$

Utilizzando la formula ricorsiva per la funzione beta  $\beta(a+1, b) = \frac{a}{a+b}\beta(a, b)$ , si ottiene che:

$$\pi_k = \prod_{j=0}^{k-1} \frac{a+j}{a+b+j}$$

da cui  $\pi = \frac{a}{a+b}$ ,  $\pi_2 = \frac{\pi(a+1)}{(a+b+1)}$  e  $\rho_Y = \rho = (a+b+1)^{-1}$ . La variabile aleatoria  $S_d$  ha una distribuzione beta-binomiale. Infatti, applicando la (3.13), si ha:

$$\begin{aligned} P(S_d = k) &= \binom{d}{k} \frac{1}{\beta(a,b)} \int_0^1 q^k (1-q)^{d-k} q^{a-1}(1-q)^{b-1} dq \\ &= \binom{d}{k} \frac{1}{\beta(a,b)} \int_0^1 q^{a+k-1}(1-q)^{b+d-k-1} dq = \binom{d}{k} \frac{\beta(a+k, b+d-k)}{\beta(a,b)}, \end{aligned} \quad (3.20)$$

cioè  $S_d \sim \text{BeBin}(d, a, b)$ .

### 3.4.5 CreditRisk<sup>+</sup> con distribuzione binomiale negativa

Come detto in precedenza, il CreditRisk<sup>+</sup> può essere visto anche come modello misto di Poisson, un'approssimazione del modello misto di Bernoulli(condizionatamente). Vedremo ora, sempre nel caso in cui la forma distributiva sottostante è una, un risultato molto importante a livello pratico.

#### Esempio 5 (Distribuzione mista gamma-poisson).

Assumiamo che  $\Lambda \sim \text{Gamma}(a, b)$ , per qualche  $a, b > 0$ . La funzione densità è data dalla (3.18). Dalla (3.16) otteniamo:

$$\tilde{\pi}_k = E(\Lambda^k) = \frac{b^a}{\Gamma(a)} \int_0^\infty \lambda^{a-1} e^{-b\lambda} \lambda^k d\lambda = \frac{b^a}{\Gamma(a)} \frac{1}{b^{a-1}} \frac{1}{b^{k+1}} \int_0^\infty x^{a-1} e^{-x} x^k dx = \frac{1}{b^k} \frac{\Gamma(a+k)}{\Gamma(a)}$$

avendo sostituito  $b\lambda = x$ . Utilizzando la formula ricorsiva della Gamma  $\Gamma(a+1) = a\Gamma(a)$  otteniamo quindi:

$$\tilde{\pi}_k = \frac{1}{b^k} \prod_{j=0}^{k-1} (a+j)$$

da cui  $\tilde{\pi} = \frac{a}{b}$  e  $\tilde{\pi}_2 = \tilde{\pi} \frac{(a+1)}{b}$ . I momenti di cui sopra fanno riferimento alla distribuzione marginale.

La proposizione 10.20 di [5] afferma che, se  $\tilde{Y}_i | \Lambda = \lambda \sim \mathcal{P}(\lambda)$  e  $\Lambda \sim \text{Gamma}(a, b)$ , allora le marginali  $\tilde{Y}_i \sim \text{BinNeg}(a, \frac{b}{b+1})$  con  $\tilde{p} := \frac{b}{b+1}$ . Proviamo ora a dimostrare che anche la somma non condizionata  $\tilde{S}_d$  segue una distribuzione binomiale negativa. Per prima cosa andiamo a sostituire esplicitamente nella (3.15) la funzione densità della gamma ottenendo:

$$P(\tilde{S}_d = k) = \int_0^\infty \frac{b^a}{\Gamma(a)} \frac{(d\lambda)^k}{k!} e^{-d\lambda} \lambda^{a-1} e^{-b\lambda} d\lambda = \frac{b^a}{\Gamma(a)k!} \int_0^\infty d^k \lambda^{a+k-1} e^{-(b+d)\lambda} d\lambda.$$

Sostituendo  $u = (b+d)\lambda$ , l'integrale diventa:

$$\begin{aligned} \int_0^\infty d^k \left( \frac{u}{b+d} \right)^{a+k-1} e^{-u} \frac{1}{(b+d)} du &= \int_0^\infty d^k u^{a+k-1} (b+d)^{-(a+k)} e^{-u} du \\ &= \frac{d^k}{(b+d)^{a+k}} \Gamma(a+k). \end{aligned}$$

Quindi si ottiene:

$$P(\tilde{S}_d = k) = \left( \frac{b}{b+d} \right)^a \left( \frac{d}{b+d} \right)^k \frac{\Gamma(a+k)}{k! \Gamma(a)}. \quad (3.21)$$

Usando la relazione  $\Gamma(a+k) = (a+k-1) \cdots a \Gamma(a)$ , si ottiene che la (3.21) è uguale alla funzione di massa di una variabile casuale binomiale negativa con  $\frac{\Gamma(a+k)}{k! \Gamma(a)} = \binom{k+a-1}{k}$  e  $\tilde{p}_s := \frac{b}{b+d}$ . Quindi  $\tilde{S}_d \sim \text{BinNeg}(a, \frac{b}{b+d})$ .

### 3.5 Calcolo del VaR per la classe di modelli in $\mathcal{E}_d(p, \rho)$

Dal momento che i modelli di mistura sono scambiabili( ciò deriva direttamente dalla forma della funzione di ripartizione congiunta per variabili iid), andiamo ora a calcolare il  $\text{VaR}_\alpha$  per il modello standard industriale CreditRisk<sup>+</sup> considerando modelli misti ad un fattore in portafogli omogenei, ovvero equicorrelati e con uguale probabilità di insolvenza, per vedere come quest' ultimo si colloca negli intervalli forniti da [7]. Il Value-at-Risk verrà calcolato nel caso in cui non solo è nota la probabilità di default  $p$  delle diverse controparti, ma quando è dato anche il coefficiente di correlazione  $\rho$ . Il fatto che la correlazione sia nota, permette di calcolare esplicitamente il VaR della distribuzione della perdita totale  $S_d$ .

#### 3.5.1 Notazione

Indichiamo con  $\mathcal{E}_d(p, \rho)$  la classe delle distribuzioni  $d$ -dimensionali di Bernoulli scambiabili con stessa distribuzione marginale e correlazione  $\rho$  nota.

Dal momento che stiamo assumendo scambiabilità, per il teorema di De Finetti( vedi Appendice B), il vettore  $\mathbf{Y} = (Y_1, \dots, Y_d)'$  segue una distribuzione di Bernoulli multivariata e le probabilità di insolvenza  $p_i$  sono tutte uguali a  $p$ . Osserviamo, inoltre, che la correlazione  $\rho$  tra due variabili di Bernoulli  $Y_i \sim \mathcal{B}(p)$  e  $Y_j \sim \mathcal{B}(p)$  è data dal momento del secondo ordine  $\pi_2 = E(Y_i Y_j)$ . La (3.17) diventa quindi:

$$\rho = \frac{\text{cov}(Y_i, Y_j)}{\sqrt{\text{var}(Y_i)\text{var}(Y_j)}} = \frac{\pi_2 - p^2}{\sqrt{p^2(1-p)^2}}$$

da cui

$$\pi_2 = \rho p(1-p) + p^2. \tag{3.22}$$

#### 3.5.2 Calibrazione del rischio con un modello misto di Bernoulli: esempi

Consideriamo un portafoglio  $P$  costituito da  $d = 100$  compagnie. Assumiamo che  $\mathbf{Y} \in \mathcal{E}$ , con  $\mathcal{E} := \mathcal{E}_{100}$ . Analizzeremo l'andamento della perdita, conoscendo uno specifico valore del coefficiente di correlazione  $\rho$ . Ci occupiamo della classe  $\mathcal{E}(p, \rho)$  nei due scenari  $p = 0,3\%$  e  $p = 1,7\%$  per i tre livelli di correlazione  $\rho = \frac{1}{6}; \frac{1}{2}; \frac{5}{6}$ . Come punto di riferimento scegliamo un modello misto di Bernoulli scambiabile, il modello misto beta, e calcoliamo il  $\text{VaR}_\alpha$  per  $\alpha = 0,90$ ,  $\alpha = 0,95$  e  $\alpha = 0,99$  per poi vedere come quest' ultimo si inserisce negli intervalli dei limiti dati. I risultati riportati nelle tabelle di seguito sui limiti del VaR quando il coefficiente di correlazione è noto sono presi da [7]. Indicando con  $S_\beta$  il numero di defaults della distribuzione mista beta, dalla (3.13) si ha:

$$P(S_\beta = k) = \binom{d}{k} \int_0^1 q^k (1-q)^{d-k} dG(q)$$

dove  $G(q)$  rappresenta la funzione di ripartizione della variabile mista  $Q \sim \text{Beta}(a, b)$ .  
Quindi, dall'esempio 4:

$$P(S_\beta = k) = \binom{d}{k} \frac{\beta(a+k, b+d-k)}{\beta(a, b)}. \quad (3.23)$$

Dal momento che stiamo considerando un modello misto di Bernoulli scambiabile:

$$\pi = p = E(Q) \quad \pi_2 = E(Q^2).$$

Andiamo a stimare i parametri  $a$  e  $b$  della distribuzione beta uguagliando i momenti teorici con quelli effettivi della distribuzione, cioè risolvendo il seguente sistema:

$$\begin{cases} p = \frac{a}{a+b} \\ \pi_2 = \frac{p(a+1)}{(a+b+1)} \end{cases} \quad (3.24)$$

### 3.5.3 Scenario 1

**Caso 1:**  $p=0,3\%$ ,  $\rho = \frac{1}{6}$ . Risolvendo il sistema (3.24), dopo aver calcolato  $\pi_2 = 5,075 \times 10^{-4}$ , si ottiene che  $a = 0,015$  e  $b = 4,98$ . Il calcolo del  $\text{VaR}_\alpha$  sarà dato da :

$$\text{VaR}_\alpha(S_\beta) = \inf\{k \in \mathbb{R} : P(S_\beta \leq k) \geq \alpha\} \quad (3.25)$$

dove, dalla (3.23):

$$P(S_\beta = k) = \binom{100}{k} \frac{\int_0^1 q^{a+k-1} (1-q)^{b+d-k-1} dq}{\int_0^1 q^{a-1} (1-q)^{b-1} dq}. \quad (3.26)$$

Calcolando numericamente questi integrali con i valori di  $a$  e  $b$  si ottiene che, per  $k = 0$ ,  $P(S_\beta \leq 0) = 0,954 > \alpha$  per  $\alpha = 0,90$  e  $\alpha = 0,95$ . Quindi in questi due casi,  $\text{VaR}_\alpha(S_\beta) = 0$ . Per  $\alpha = 0,99$ ,  $k = 0$  non è più il VaR. Bisogna trovare il più piccolo  $k$  per cui valga la (3.25). Svolgendo numericamente gli integrali, si trova che, per  $k = 9$  la disuguaglianza è soddisfatta. Infatti, dalla (3.26) si ha:

$$k = 1 \quad P(S_\beta = 1) = \binom{100}{1} \frac{\int_0^1 q^{0,015} (1-q)^{102,98} dq}{\int_0^1 q^{-0,985} (1-q)^{3,98} dq} = 0,01376$$

$$k = 2 \quad P(S_\beta = 2) = \binom{100}{2} \frac{\int_0^1 q^{1,015} (1-q)^{101,98} dq}{\int_0^1 q^{-0,985} (1-q)^{3,98} dq} = 6,71 \times 10^{-3}$$

$$k = 3 \quad P(S_\beta = 3) = \binom{100}{3} \frac{\int_0^1 q^{2,015} (1-q)^{100,98} dq}{\int_0^1 q^{-0,985} (1-q)^{3,98} dq} = 4,328 \times 10^{-3}$$

$$k = 4 \quad P(S_\beta = 4) = \binom{100}{4} \frac{\int_0^1 q^{3,015} (1-q)^{99,98} dq}{\int_0^1 q^{-0,985} (1-q)^{3,98} dq} = 3,14 \times 10^{-3}$$

$$\begin{aligned}
 k = 5 \quad P(S_\beta = 5) &= \binom{100}{5} \frac{\int_0^1 q^{4,015} (1-q)^{98,98} dq}{\int_0^1 q^{-0,985} (1-q)^{3,98} dq} = 2,418 \times 10^{-3} \\
 k = 6 \quad P(S_\beta = 6) &= \binom{100}{6} \frac{\int_0^1 q^{5,015} (1-q)^{97,98} dq}{\int_0^1 q^{-0,985} (1-q)^{3,98} dq} = 1,937 \times 10^{-3} \\
 k = 7 \quad P(S_\beta = 7) &= \binom{100}{7} \frac{\int_0^1 q^{6,015} (1-q)^{96,98} dq}{\int_0^1 q^{-0,985} (1-q)^{3,98} dq} = 1,6 \times 10^{-3} \\
 k = 8 \quad P(S_\beta = 8) &= \binom{100}{8} \frac{\int_0^1 q^{7,015} (1-q)^{95,98} dq}{\int_0^1 q^{-0,985} (1-q)^{3,98} dq} = 1,34 \times 10^{-3} \\
 k = 9 \quad P(S_\beta = 9) &= \binom{100}{9} \frac{\int_0^1 q^{8,015} (1-q)^{94,98} dq}{\int_0^1 q^{-0,985} (1-q)^{3,98} dq} = 1,15 \times 10^{-3}
 \end{aligned}$$

Sommando tutte queste probabilità si ottiene che la  $P(S_\beta \leq 9) \approx 0,9903 > 0,99$ , per cui il  $\text{VaR}_{0,99}(S_\beta) = 9$ . I risultati sono riassunti nella seguente tabella:

$\alpha$	$\text{Min VaR}_\alpha$	$\text{Max VaR}_\alpha$	$\beta - \text{VaR}_\alpha$
0,90	0	2	0
0,95	0	5	0
0,99	1	22	9

Tabella 3.1:  $\text{VaR}_\alpha$  del numero di defaults per  $\mathcal{E}(0,3\%, \frac{1}{6})$  per una Bernoulli multivariata

**Caso 2:  $p=0,3\%$ ,  $\rho = \frac{1}{2}$ .** Dopo aver calcolato  $\pi_2 = 1,5 \times 10^{-3}$ , risolvendo il sistema si trovano  $a = 0,003$  e  $b = 0,997$ . In questo caso, si può osservare che, per  $k = 0$ , la  $P(S_\beta \leq 0) = 0,984 > \alpha$  per  $\alpha = 0,90$  e  $\alpha = 0,95$ . Quindi, il  $\text{VaR}_\alpha(S_\beta) = 0$ . Per calcolare il VaR della distribuzione mista beta quando  $\alpha = 0,99$ , bisogna iterare  $k$ . Senza riscrivere esplicitamente gli integrali, troviamo che, quando:

$$k = 1 \quad P(S_\beta = 1) = 2,95 \times 10^{-3}$$

$$k = 2 \quad P(S_\beta = 2) = 1,47 \times 10^{-3}$$

$$k = 3 \quad P(S_\beta = 3) = 9,89 \times 10^{-4}$$

$$k = 4 \quad P(S_\beta = 4) = 7,42 \times 10^{-4}.$$

Sommando tutte queste probabilità si ottiene:  $P(S_\beta \leq 4) \approx 0,9901 > 0,99$ . Quindi il  $\text{VaR}_{0,99}(S_\beta) = 4$ , essendo il più piccolo  $k$  che verifica la disuguaglianza. I risultati sono riportati nella seguente tabella:

$\alpha$	$\text{Min VaR}_\alpha$	$\text{Max VaR}_\alpha$	$\beta - \text{VaR}_\alpha$
0,90	0	1	0
0,95	0	3	0
0,99	0	21	4

Tabella 3.2:  $\text{VaR}_\alpha$  del numero di defaults per  $\mathcal{E}(0,3\%, \frac{1}{2})$  per una Bernoulli multivariata

**Caso 3:**  $\mathbf{p=0,3\%}$ ,  $\rho = \frac{5}{6}$ . Dopo aver calcolato  $\pi_2 = 2,5015 \times 10^{-3}$ , otteniamo  $a = 0,0006$  e  $b = 0,199$ . In questo ultimo caso possiamo osservare che  $k = 0$  è il più piccolo intero che soddisfa la disuguaglianza  $P(S_\beta \leq k) \geq \alpha$  per ogni valore di  $\alpha$ . Infatti:

$$P(S_\beta = 0) = 0,997,$$

quindi il  $\text{VaR}_\alpha(S_\beta) = 0$ . I risultati sono riportati nella seguente tabella:

$\alpha$	$\text{Min VaR}_\alpha$	$\text{Max VaR}_\alpha$	$\beta - \text{VaR}_\alpha$
0,90	0	0	0
0,95	0	1	0
0,99	0	7	0

Tabella 3.3:  $\text{VaR}_\alpha$  del numero di defaults per  $\mathcal{E}(0,3\%, \frac{5}{6})$  per una Bernoulli multivariata

### 3.5.4 Scenario 2

**Caso 1:**  $\mathbf{p} = 1,7\%$ ,  $\boldsymbol{\rho} = \frac{1}{6}$ . Dati  $p$  e  $\rho$  si trova che  $\pi_2 = 3,07 \times 10^{-3}$  da cui, risolvendo il sistema, otteniamo  $a = 0,085$  e  $b = 4,915$ . Per calcolare il  $\text{VaR}_\alpha$  della distribuzione mista beta, osserviamo che, per  $\alpha = 0,90$ , il più piccolo  $k$  che soddisfa:  $P(S_\beta \leq k) \geq 0,90$  è  $k = 5$ . Infatti, risolvendo numericamente la (3.26), per:

$$k = 0 \quad P(S_\beta = 0) = 0,765$$

$$k = 1 \quad P(S_\beta = 1) = 0,0625$$

$$k = 2 \quad P(S_\beta = 2) = 0,032$$

$$k = 3 \quad P(S_\beta = 3) = 0,0218$$

$$k = 4 \quad P(S_\beta = 4) = 0,016$$

$$k = 5 \quad P(S_\beta = 5) = 0,013.$$

Quindi la  $P(S_\beta \leq 5) \approx 0,91 > \alpha = 0,90$ , da cui  $\text{VaR}_{0,90}(S_\beta) = 5$ . Per  $\alpha = 0,95$  dobbiamo continuare ad iterare  $k$ , ottenendo che per  $k = 11$  la disuguaglianza è soddisfatta. Infatti, per:

$$k = 6 \quad P(S_\beta = 6) = 0,010$$

$$k = 7 \quad P(S_\beta = 7) = 8,61 \times 10^{-3}$$

$$k = 8 \quad P(S_\beta = 8) = 7,31 \times 10^{-3}$$

$$k = 9 \quad P(S_\beta = 9) = 6,3 \times 10^{-3}$$

$$k = 10 \quad P(S_\beta = 10) = 5,398 \times 10^{-3}$$

$$k = 11 \quad P(S_\beta = 11) = 4,77 \times 10^{-3}.$$



Sommando tutti questi valori alla  $P(S_\beta \leq 5)$  si ottiene che:  $P(S_\beta \leq 11) \approx 0,952 > 0,95$ , da cui  $\text{VaR}_{0,95}(S_\beta) = 11$ . Per  $\alpha = 0,99$  il  $\beta - \text{VaR}_{0,99}$  è preso da [7]. I risultati sono riassunti nella seguente tabella:

$\alpha$	$\text{Min VaR}_\alpha$	$\text{Max VaR}_\alpha$	$\beta - \text{VaR}_\alpha$
0,90	0	16	5
0,95	1	25	11
0,99	2	55	29

Tabella 3.4:  $\text{VaR}_\alpha$  del numero di defaults per  $\mathcal{E}(1,7\%, \frac{1}{6})$  per una Bernoulli multivariata

**Caso 2:**  $\mathbf{p} = 1,7\%$ ,  $\boldsymbol{\rho} = \frac{1}{2}$ . In questo caso  $\pi_2 = 8,64 \times 10^{-3}$ , da cui  $a = 0,017$  e  $b = 0,983$ . Calcoliamo il  $\text{VaR}_\alpha$  della distribuzione mista beta per  $\alpha = 0,90$  e  $\alpha = 0,95$ . Osserviamo che, per  $k = 0$ , la  $P(S_\beta \leq 0) \approx 0,915 > 0,90$ , da cui  $\text{VaR}_{0,90}(S_\beta) = 0$ . Per  $\alpha = 0,95$ , si trova che il più piccolo  $k$  per cui la (3.25) è soddisfatta è  $k = 5$ . Svolgendo numericamente gli integrali troviamo che, per:

$$k = 1 \quad P(S_\beta = 1) = 0,0156$$

$$k = 2 \quad P(S_\beta = 2) = 7,9 \times 10^{-3}$$

$$k = 3 \quad P(S_\beta = 3) = 5,32 \times 10^{-3}$$

$$k = 4 \quad P(S_\beta = 4) = 4,02 \times 10^{-3}$$

$$k = 5 \quad P(S_\beta = 5) = 3,21 \times 10^{-3}.$$

Sommando le probabilità di cui sopra si ottiene che la  $P(S_\beta \leq 5) \approx 0,951 > 0,95$ , per cui  $\text{VaR}_{0,95}(S_\beta) = 5$ . Per  $\alpha = 0,99$  il  $\beta - \text{VaR}_{0,99}$  è preso da [7]. I risultati sono riportati nella seguente tabella:

$\alpha$	$Min VaR_\alpha$	$Max VaR_\alpha$	$\beta - VaR_\alpha$
0,90	0	9	0
0,95	0	25	5
0,99	1	93	57

Tabella 3.5:  $VaR_\alpha$  del numero di defaults per  $\mathcal{E}(1,7\%, \frac{1}{2})$  per una Bernoulli multivariata

**Caso 3:**  $\mathbf{p} = 1,7\%$ ,  $\boldsymbol{\rho} = \frac{5}{6}$ . Calcolando  $\pi_2 = 0,014$  troviamo  $a = 3,4 \times 10^{-3}$  e  $b = 0,197$ . Andando a calcolare la (3.26) per  $k = 0$  si ottiene che la  $P(S_\beta \leq 0) = 0,97 > \alpha$  per  $\alpha = 0,90$  e  $\alpha = 0,95$ , per cui  $VaR_\alpha(S_\beta) = 0$  in entrambi i casi. Per  $\alpha = 0,99$  il  $\beta - VaR_{0,99}$  è preso da [7]. I risultati sono riassunti di seguito:

$\alpha$	$Min VaR_\alpha$	$Max VaR_\alpha$	$\beta - VaR_\alpha$
0,90	0	3	0
0,95	0	8	0
0,99	61	100	94

Tabella 3.6:  $VaR_\alpha$  del numero di defaults per  $\mathcal{E}(1,7\%, \frac{5}{6})$  per una Bernoulli multivariata

**Osservazione 3.** Dai conti svolti, possiamo vedere come si comporta il  $VaR_\alpha$  della distribuzione mista beta quando  $p$  e  $\rho$  sono dati. In particolare, si può osservare come il  $\beta - VaR_\alpha$  sia più vicino, addirittura uguale, al limite inferiore quando la probabilità di default è piccola. Al contrario, al crescere di  $p$  e del coefficiente di correlazione  $\rho$  il VaR tende a stabilizzarsi nei pressi del limite superiore. Possiamo inoltre osservare che la differenza diventa sostanziale soprattutto all'aumentare del livello di confidenza  $\alpha$ : in questo caso, all'aumentare di  $p$ , cresce significamente il VaR giustificato anche dal fatto che si sta richiedendo una percentuale di protezione molto più alta. Ad ogni modo la larghezza dell'intervallo tra massimo e minimo, insieme alla posizione del  $\beta - VaR_\alpha$  al suo interno, fornisce una chiara indicazione di quanto possa andare male l'andamento del rischio, quando la correlazione è data, e come stringente sia la scelta di una opportuna distribuzione multivariata all'interno di tale calibrazione.

### 3.5.5 Calibrazione del rischio con un modello misto di Poisson: esempi

Eseguiamo ora gli stessi conti andando però a considerare un modello misto di Poisson scambiabile ad un fattore. Indicando con  $S_\Gamma$  il numero di defaults della distribuzione gamma-poisson, dall'esempio 5 si ha:

$$P(S_\Gamma = k) = \left(\frac{b}{b+d}\right)^a \left(\frac{d}{b+d}\right)^k \binom{k+a-1}{k}. \quad (3.27)$$

Ora dobbiamo stimare i parametri  $a$  e  $b$  della binomiale negativa. Per la marginale  $\tilde{Y}_i$  si ha:

$$E(\tilde{Y}_i) = \frac{a(1-\tilde{p})}{\tilde{p}} = \frac{a(b+1)}{b} \left(1 - \frac{b}{b+1}\right) = \frac{a}{b} = E(\Lambda)$$

e

$$\begin{aligned} E(\tilde{Y}_i^2) &= E(\tilde{Y}_i)^2 + \text{var}(\tilde{Y}_i) = \frac{a^2}{b^2} + \frac{a(1-\tilde{p})}{\tilde{p}^2} \\ &= \frac{a^2}{b^2} + a \left(1 - \frac{b}{b+1}\right) \frac{(b+1)^2}{b^2} = \frac{a^2}{b^2} + \frac{a}{b^2}(b+1) \\ &= \frac{a}{b^2}(a+b+1) = E(\Lambda) + E(\Lambda^2) \end{aligned}$$

dove per l'appunto  $\Lambda$  è la variabile di mistura per il modello misto di Poisson ed è distribuita come una Gamma di parametri  $a$  e  $b$ . Tuttavia non possiamo usare la marginale per trovare i parametri poiché, se fissiamo i primi due momenti, stiamo di fatto fissando anche la correlazione. Ad ogni modo, i conti fatti sulla marginale  $\tilde{Y}_i$  possono essere ripetuti sulla somma  $\tilde{S}_d$ , dal momento che anche quest'ultima segue una distribuzione binomiale negativa ed utilizzeremo proprio la somma per trovare  $a$  e  $b$ . Infatti:

$$E(\tilde{S}_d) = \frac{a(1-\tilde{p}_s)}{\tilde{p}_s} = \frac{a(b+d)}{b} \left(1 - \frac{b}{b+d}\right) = d \frac{a}{b}$$

e

$$\begin{aligned} E(\tilde{S}_d^2) &= E(\tilde{S}_d)^2 + \text{var}(\tilde{S}_d) = d^2 \frac{a^2}{b^2} + \frac{a(1-\tilde{p}_s)}{\tilde{p}_s^2} \\ &= d^2 \frac{a^2}{b^2} + a \left(1 - \frac{b}{b+d}\right) \frac{(b+d)^2}{b^2} = d^2 \frac{a^2}{b^2} + d \frac{a}{b^2}(b+d) \\ &= d \frac{a}{b^2}(ad+b+d). \end{aligned}$$

Ora, bisogna imporre che:

$$E(S_d) = pd = E(\tilde{S}_d) \quad E(S_d^2) = pd + d(d-1)\pi_2 = E(\tilde{S}_d^2),$$

ovvero che sia uguale al momento secondo di una somma di Bernoulli (vedi [7]), dove  $\pi_2$  è dato dalla (3.22). Quindi, in definitiva, per trovare  $a$  e  $b$  dobbiamo risolvere il seguente sistema:

$$\begin{cases} pd = d \frac{a}{b} \\ pd + d(d-1)\pi_2 = d \frac{a}{b^2}(ad+b+d). \end{cases} \quad (3.28)$$

Anche in questo caso ci occuperemo di calcolare esplicitamente il VaR nei due scenari  $p = 0,3\%$  e  $p = 1,7\%$  per i tre livelli di correlazione  $\rho = \frac{1}{6}; \frac{1}{2}; \frac{5}{6}$  andando a vedere come si colloca negli intervalli forniti già nelle tabelle precedenti. Per  $p = 1,7\%$  il calcolo esplicito è stato effettuato solo per  $\alpha = 0,90$  e  $\alpha = 0,95$ , perchè per  $\alpha = 0,99$  richiede un numero di iterazioni troppo elevato per cui servirebbe l'ausilio di un algoritmo numerico. La scelta di considerare probabilità di default piccole è necessaria, in quanto rende l'approssimazione di cui sopra il più precisa possibile.

### 3.5.6 Scenario 1

**Caso 1:**  $\mathbf{p} = 0,3\%$ ,  $\boldsymbol{\rho} = \frac{1}{6}$ . Andando a risolvere il sistema (3.28) troviamo  $a = 0,018$  e  $b = 6,08$ . L'obiettivo è quello di trovare il più piccolo  $k$  tale che la  $P(S_\Gamma \leq k) \geq \alpha$ . Per  $\alpha = 0,90$  il più piccolo  $k$  che soddisfa la disuguaglianza è  $k = 0$ . Infatti, per:

$$k = 0 \quad P(S_\Gamma = 0) = \left(\frac{6,08}{106,08}\right)^{0,018} = 0,9498$$

quindi il  $\text{VaR}_{0,90}(S_\Gamma) = 0$ . Iterando  $k$ :

$$k = 1 \quad P(S_\Gamma = 1) = \left(\frac{6,08}{106,08}\right)^{0,018} \left(\frac{100}{106,08}\right) 0,018 = 0,016$$

da cui la  $P(S_\Gamma \leq 1) \approx 0,966 > 0,95$ . Quindi  $\text{VaR}_{0,95}(S_\Gamma) = 1$ . Per  $\alpha = 0,99$  bisogna continuare ad iterare  $k$  ottenendo che, per:

$$k = 2 \quad P(S_\Gamma = 2) = \left(\frac{6,08}{106,08}\right)^{0,018} \left(\frac{100}{106,08}\right)^2 \frac{\Gamma(2,018)}{2!\Gamma(0,018)} = 7,733 \times 10^{-3}$$

$$k = 3 \quad P(S_\Gamma = 3) = \left(\frac{6,08}{106,08}\right)^{0,018} \left(\frac{100}{106,08}\right)^3 \frac{\Gamma(3,018)}{3!\Gamma(0,018)} = 4,904 \times 10^{-3}$$

$$k = 4 \quad P(S_\Gamma = 4) = \left(\frac{6,08}{106,08}\right)^{0,018} \left(\frac{100}{106,08}\right)^4 \frac{\Gamma(4,018)}{4!\Gamma(0,018)} = 3,488 \times 10^{-3}$$

$$k = 5 \quad P(S_\Gamma = 5) = \left(\frac{6,08}{106,08}\right)^{0,018} \left(\frac{100}{106,08}\right)^5 \frac{\Gamma(5,018)}{5!\Gamma(0,018)} = 2,642 \times 10^{-3}$$

$$k = 6 \quad P(S_\Gamma = 6) = \left(\frac{6,08}{106,08}\right)^{0,018} \left(\frac{100}{106,08}\right)^6 \frac{\Gamma(6,018)}{6!\Gamma(0,018)} = 2,083 \times 10^{-3}$$

$$k = 7 \quad P(S_\Gamma = 7) = \left(\frac{6,08}{106,08}\right)^{0,018} \left(\frac{100}{106,08}\right)^7 \frac{\Gamma(7,018)}{7!\Gamma(0,018)} = 1,688 \times 10^{-3}$$

$$k = 8 \quad P(S_\Gamma = 8) = \left(\frac{6,08}{106,08}\right)^{0,018} \left(\frac{100}{106,08}\right)^8 \frac{\Gamma(8,018)}{8!\Gamma(0,018)} = 1,396 \times 10^{-3}$$

$$k = 9 \quad P(S_\Gamma = 9) = \left(\frac{6,08}{106,08}\right)^{0,018} \left(\frac{100}{106,08}\right)^9 \frac{\Gamma(8,018)}{9!\Gamma(0,018)} = 1,172 \times 10^{-3}.$$

Quindi, sommando tutte queste probabilità si ottiene:  $P(S_\Gamma \leq 9) \approx 0,991 > 0,99$  per cui il  $\text{VaR}_{0,99}(S_\Gamma) = 9$ .

I risultati sono riassunti di seguito:

$\alpha$	$Min VaR_\alpha$	$Max VaR_\alpha$	$\Gamma - VaR_\alpha$
0,90	0	2	0
0,95	0	5	1
0,99	1	22	9

Tabella 3.7:  $\text{VaR}_\alpha$  del numero di defaults per  $\mathcal{E}(0,3\%, \frac{1}{6})$  per una Poisson multivariata

**Caso 2:**  $\mathbf{p} = 0,3\%$ ,  $\boldsymbol{\rho} = \frac{1}{2}$ . In questo caso troviamo che  $a = 6,097 \times 10^{-3}$  e  $b = 2,03$ . Per  $\alpha = 0,90$ ,  $\alpha = 0,95$  il più piccolo  $k$  che verifica la disuguaglianza è  $k = 0$ . Infatti, per:

$$k = 0 \quad P(S_\Gamma = 0) = 0,9764$$

da cui il  $\text{VaR}_\alpha(S_\Gamma) = 0$  per  $\alpha = 0,90$ ,  $\alpha = 0,95$ . Per  $\alpha = 0,99$ , quando  $k = 6$  la disuguaglianza è soddisfatta. Infatti:

$$k = 1 \quad P(S_\Gamma = 1) = 5,835 \times 10^{-3}$$

$$k = 2 \quad P(S_\Gamma = 2) = 2,877 \times 10^{-3}$$

$$k = 3 \quad P(S_\Gamma = 3) = 1,885 \times 10^{-3}$$

$$k = 4 \quad P(S_\Gamma = 4) = 1,389 \times 10^{-3}$$

$$k = 5 \quad P(S_\Gamma = 5) = 1,091 \times 10^{-3}$$

$$k = 6 \quad P(S_\Gamma = 6) = 8,918 \times 10^{-4}.$$

Quindi, sommando tutte queste probabilità si ottiene:  $P(S_\Gamma \leq 6) \approx 0,9904 > 0,99$ , da cui il  $\text{VaR}_{0,99}(S_\Gamma) = 6$ .

$\alpha$	$\text{Min VaR}_\alpha$	$\text{Max VaR}_\alpha$	$\Gamma - \text{VaR}_\alpha$
0,90	0	1	0
0,95	0	3	0
0,99	0	21	6

Tabella 3.8:  $\text{VaR}_\alpha$  del numero di defaults per  $\mathcal{E}(0,3\%, \frac{1}{2})$  per una Poisson multivariata

**Caso 3:**  $\mathbf{p} = 0,3\%$ ,  $\boldsymbol{\rho} = \frac{5}{6}$ . Risolvendo il sistema otteniamo:  $a = 3,647 \times 10^{-3}$  e  $b = 1,216$ . Anche in questa ultima situazione,  $k = 0$  è il valore del VaR per  $\alpha = 0,90$ ,  $\alpha = 0,95$ . Infatti, per:

$$k = 0 \quad P(S_\Gamma = 0) = 0,984$$

e quindi il  $\text{VaR}_\alpha(S_\Gamma) = 0$ . Per  $\alpha = 0,99$  la disuguaglianza è soddisfatta per  $k = 3$ . Infatti:

$$k = 1 \quad P(S_\Gamma = 1) = 3,545 \times 10^{-3}$$

$$k = 2 \quad P(S_\Gamma = 2) = 1,758 \times 10^{-3}$$

$$k = 3 \quad P(S_\Gamma = 3) = 1,16 \times 10^{-3}.$$

Sommando questa probabilità alle precedenti si ha:  $P(S_\Gamma \leq 3) \approx 0,9905 > 0,99$ , da cui  $\text{VaR}_{0,99}(S_\Gamma) = 3$ .

I risultati sono riportati nella tabella seguente:

$\alpha$	$Min VaR_\alpha$	$Max VaR_\alpha$	$\Gamma - VaR_\alpha$
0,90	0	0	0
0,95	0	1	0
0,99	0	7	3

Tabella 3.9:  $VaR_\alpha$  del numero di defaults per  $\mathcal{E}(0,3\%, \frac{5}{6})$  per una Poisson multivariata

### 3.5.7 Scenario 2

**Caso 1:**  $\mathbf{p} = 1,7\%$ ,  $\boldsymbol{\rho} = \frac{1}{6}$ . Risolvendo il sistema (3.28) troviamo  $a = 0,105$  e  $b = 6,181$ . Per  $\alpha = 0,90$  il più piccolo  $k$  che soddisfa la disuguaglianza è  $k = 5$ . Infatti, per:

$$k = 0 \quad P(S_\Gamma = 0) = 0,742$$

$$k = 1 \quad P(S_\Gamma = 1) = 0,073$$

$$k = 2 \quad P(S_\Gamma = 2) = 0,0382$$

$$k = 3 \quad P(S_\Gamma = 3) = 0,0252$$

$$k = 4 \quad P(S_\Gamma = 4) = 0,0184.$$

$$k = 5 \quad P(S_\Gamma = 5) = 0,0143.$$

Sommando tutti questi valori si ottiene:  $P(S_\Gamma \leq 5) \approx 0,911 > 0,90$ . Quindi, in questo caso,  $VaR_{0,90}(S_\Gamma) = 5$ . Per  $\alpha = 0,95$  dobbiamo continuare ad iterare  $k$ , ottenendo per  $k = 10$  verificata la (3.25). Infatti, per:

$$k = 6 \quad P(S_\Gamma = 6) = 0,0114$$

$$k = 7 \quad P(S_\Gamma = 7) = 9,386 \times 10^{-3}$$

$$k = 8 \quad P(S_\Gamma = 8) = 7,851 \times 10^{-3}$$

$$k = 9 \quad P(S_{\Gamma} = 9) = 6,659 \times 10^{-3}$$

$$k = 10 \quad P(S_{\Gamma} = 10) = 5,71 \times 10^{-3}$$

da cui la  $P(S_{\Gamma} \leq 10) \approx 0,952 > 0,95$ . Quindi, in questo caso,  $\text{VaR}_{0,95}(S_{\Gamma}) = 10$ . I risultati sono riassunti nella seguente tabella:

$\alpha$	$\text{Min VaR}_{\alpha}$	$\text{Max VaR}_{\alpha}$	$\Gamma - \text{VaR}_{\alpha}$
0,90	0	16	5
0,95	1	25	10

Tabella 3.10:  $\text{VaR}_{\alpha}$  del numero di defaults per  $\mathcal{E}(1,7\%, \frac{1}{6})$  per una Poisson multivariata

**Caso 2:**  $\mathbf{p} = 1,7\%$ ,  $\boldsymbol{\rho} = \frac{1}{2}$ . Troviamo  $a = 0,0349$  e  $b = 2,057$ . Per  $\alpha = 0,90$  si ha:

$$k = 0 \quad P(S_{\Gamma} = 0) = 0,873$$

$$k = 1 \quad P(S_{\Gamma} = 1) = 0,0298$$

da cui la  $P(S_{\Gamma} \leq 1) \approx 0,903 > 0,90$  e quindi il  $\text{VaR}_{0,90}(S_{\Gamma}) = 1$ . Per  $\alpha = 0,95$  troviamo che il più piccolo  $k$  che verifica la disuguaglianza è  $k = 7$ . Infatti:

$$k = 2 \quad P(S_{\Gamma} = 2) = 0,015$$

$$k = 3 \quad P(S_{\Gamma} = 3) = 0,01$$

$$k = 4 \quad P(S_{\Gamma} = 4) = 7,479 \times 10^{-3}$$

$$k = 5 \quad P(S_{\Gamma} = 5) = 5,914 \times 10^{-3}$$

$$k = 6 \quad P(S_{\Gamma} = 6) = 4,862 \times 10^{-3}$$



$$k = 7 \quad P(S_\Gamma = 7) = 4,108 \times 10^{-3}$$

Quindi, la  $P(S_\Gamma \leq 7) \approx 0,9504 > 0,95$  e dunque il  $\text{VaR}_{0,95}(S_\Gamma) = 7$ . Nella tabella 3.11 sono riportati i risultati.

$\alpha$	$Min VaR_\alpha$	$Max VaR_\alpha$	$\Gamma - VaR_\alpha$
0,90	0	9	1
0,95	0	25	7

Tabella 3.11:  $\text{VaR}_\alpha$  del numero di defaults per  $\mathcal{E}(1,7\%, \frac{1}{2})$  per una Poisson multivariata

**Caso 3:**  $\mathbf{p} = 1,7\%$ ,  $\boldsymbol{\rho} = \frac{5}{6}$ . In questo caso otteniamo  $a = 0,0213$  e  $b = 1,253$ . Partendo da  $\alpha = 0,90$ , la disuguaglianza è verificata quando  $k = 0$ . Infatti, dalla (3.27):

$$k = 0 \quad P(S_\Gamma = 0) = 0,911$$

da cui il  $\text{VaR}_{0,90}(S_\Gamma) = 0$ . Per  $\alpha = 0,95$  bisogna continuare ad iterare  $k$  ottenendo il risultato quando  $k = 4$ . Infatti:

$$k = 1 \quad P(S_\Gamma = 1) = 0,0191$$

$$k = 2 \quad P(S_\Gamma = 2) = 9,665 \times 10^{-3}$$

$$k = 3 \quad P(S_\Gamma = 3) = 6,431 \times 10^{-3}$$

$$k = 4 \quad P(S_\Gamma = 4) = 4,798 \times 10^{-3}$$

e quindi, sommando questi valori al precedente risultato otteniamo che la  $P(S_\Gamma \leq 4) \approx 0,951 > 0,95$ , da cui  $\text{VaR}_{0,95}(S_\Gamma) = 4$ . I risultati sono riportati nella seguente tabella:

$\alpha$	$Min VaR_\alpha$	$Max VaR_\alpha$	$\Gamma - VaR_\alpha$
0,90	0	3	0
0,95	0	8	4

Tabella 3.12:  $VaR_\alpha$  del numero di defaults per  $\mathcal{E}(1,7\%, \frac{5}{6})$  per una Poisson multivariata

**Osservazione 4.** I risultati sui limiti del VaR considerati valgono per la classe di variabili di Bernoulli multivariate scambiabili. Sicuramente, il calcolo per delle Poisson condurrà a dei VaR-bounds diversi. Possiamo osservare che, sia per  $p = 0,3\%$  che per  $p = 1,7\%$ , il calcolo del  $\Gamma - VaR_\alpha$  nei due modelli conduce pressochè allo stesso risultato per tutti e tre i livelli di correlazione. Si discosta leggermente considerando una precisione più elevata, cioè  $\alpha = 0,99$ . Tale differenza è resa più evidente all'aumentare del coefficiente di correlazione  $\rho$ . Ad ogni modo, per i casi esaminati, il VaR rientra nei limiti già noti. Non sappiamo se il calcolo del Value-at-Risk potrebbe condurre a considerazioni diverse per  $\alpha = 0,99$  che, per  $p = 1,7\%$ , non è stato calcolato poichè servirebbe l'ausilio di un algoritmo numerico. Quindi per i conti svolti, nella calibrazione delle misure di rischio, sembrerebbe che il modello misto di Poisson sia più conservativo rispetto a quello di Bernoulli.

### 3.6 Modello a due fattori e scambiabilità parziale

Tutto quello che abbiamo trattato finora presuppone l'uso di modelli di mistura in cui la forma distributiva sottostante è una sola. Attraverso questa semplificazione importante, essenziale in ambito pratico, il teorema di De Finetti assicura che sequenze infinite scambiabili possano essere rappresentate come modelli misti. Tuttavia, nella loro forma più generale, le variabili latenti sono dei vettori. In questa ultima sezione, supponiamo di avere un modello di mistura in cui le variabili latenti sono due. In questo caso, la definizione di scambiabilità totale cade e avremo a che fare con portafogli che sono costituiti da due gruppi omogenei. Questo concetto è generalizzabile, se assumiamo che le variabili siano parzialmente scambiabili, ad un portafoglio che ha  $K > 2$  blocchi omogenei. L'obiettivo, infatti, sarà quello di mostrare, supponendo che valga la rappresentazione come modello di mistura, che il vettore delle variabili macroeconomiche sia parzialmente scambiabile secondo la seguente definizione:

**Definizione 22 ( Parziale scambiabilità).** Si considerino due sequenze di variabili casuali finite  $X_1, \dots, X_{d_1}$  e  $Y_1, \dots, Y_{d_2}$  definite sullo spazio di probabilità  $(\Omega, \mathcal{F}, P)$ . Il vettore  $(\mathbf{X}, \mathbf{Y})$  è detto parzialmente scambiabile se, per ogni permutazione  $(\sigma_1(1), \dots, \sigma_1(d_1))$  di  $\{1, \dots, d_1\}$  e  $(\sigma_2(1), \dots, \sigma_2(d_2))$  di  $\{1, \dots, d_2\}$ , le distribuzioni di

$(X_1, \dots, X_{d_1})$  e  $(Y_1, \dots, Y_{d_2})$  coincidono rispettivamente con quelle di  $(X_{\sigma_1(1)}, \dots, X_{\sigma_1(d_1)})$  e  $(Y_{\sigma_2(1)}, \dots, Y_{\sigma_2(d_2)})$ . Formalmente:

$$P(X_1 = x_1, \dots, X_{d_1} = x_{d_1}, Y_1 = y_1, \dots, Y_{d_2} = y_{d_2}) \quad (3.29)$$

$$= P(X_{\sigma_1(1)} = x_{\sigma_1(1)}, \dots, X_{\sigma_1(d_1)} = x_{\sigma_1(d_1)}, Y_{\sigma_2(1)} = y_{\sigma_2(1)}, \dots, Y_{\sigma_2(d_2)} = y_{\sigma_2(d_2)}).$$

sotto l'ipotesi che  $d_1 + d_2 = d$ , ovvero la dimensione della sequenza di partenza.

Come per il teorema di De Finetti, definiamo due misure di probabilità associate alle sequenze. Siano  $\pi_{p_1}(1) = p_1$ ,  $\pi_{p_1}(0) = 1 - p_1$  per la sequenza  $X_1, \dots, X_{d_1}$  e  $\pi_{p_2}(1) = p_2$ ,  $\pi_{p_2}(0) = 1 - p_2$  per la sequenza  $Y_1, \dots, Y_{d_2}$  che in questo caso però, essendo le variabili indipendenti, coincidono con le distribuzioni marginali, e quindi:  $P(X_i = 1) = \pi_{p_1}(1) = p_1$ ,  $P(Y_i = 1) = \pi_{p_2}(1) = p_2$ .

Consideriamo un modello misto di Bernoulli della forma (3.2) a due fattori dove  $\Psi = (\Psi_1, \Psi_2)$  è il vettore delle variabili latenti con legge  $G$  e :

$$\begin{aligned} & P(X_1 = x_1, \dots, X_{d_1} = x_{d_1}, Y_1 = y_1, \dots, Y_{d_2} = y_{d_2} | (\Psi_1, \Psi_2)) \\ &= P(X_1 = x_1, \dots, X_{d_1} = x_{d_1} | (\Psi_1, \Psi_2)) P(Y_1 = y_1, \dots, Y_{d_2} = y_{d_2} | (\Psi_1, \Psi_2)) \end{aligned}$$

dal momento che il vettore  $(\mathbf{X}, \mathbf{Y})$  è costituito da variabili aleatorie indipendenti. Quindi,  $X_i | (\Psi_1, \Psi_2) \sim \mathcal{B}(p_1)$ , con  $i = 1, \dots, d_1$ , e  $Y_j | (\Psi_1, \Psi_2) \sim \mathcal{B}(p_2)$ , con  $j = 1, \dots, d_2$ . La distribuzione non condizionata si ottiene integrando sulla legge di  $\Psi$ , cioè:

$$\begin{aligned} & P(X_1 = x_1, \dots, X_{d_1} = x_{d_1}, Y_1 = y_1, \dots, Y_{d_2} = y_{d_2}) \\ &= \int \int \prod_{i=1}^{d_1} \pi_{p_1}(x_i) \prod_{j=1}^{d_2} \pi_{p_2}(y_j) dG(p_1, p_2). \end{aligned} \quad (3.30)$$

Il vettore  $(\mathbf{X}, \mathbf{Y})$  costituito dalle due sequenze  $(X_1, \dots, X_{d_1})$  e  $(Y_1, \dots, Y_{d_2})$  è parzialmente scambiabile nel senso della definizione 22? Dobbiamo cioè verificare che permutando gli indici delle due sequenze, la legge rimane sempre la stessa, ovvero la (3.30). Infatti:

$$\begin{aligned} (3.30) &= \int \int \prod_{i=1}^{d_1} \pi_{p_1}(x_{\sigma_1(i)}) \prod_{j=1}^{d_2} \pi_{p_2}(y_{\sigma_2(j)}) dG(p_1, p_2) \\ &= P(X_{\sigma_1(1)} = x_{\sigma_1(1)}, \dots, X_{\sigma_1(d_1)} = x_{\sigma_1(d_1)}, Y_{\sigma_2(1)} = y_{\sigma_2(1)}, \dots, Y_{\sigma_2(d_2)} = y_{\sigma_2(d_2)}). \end{aligned}$$

Come possiamo osservare, la prima variabile è indipendente dalla misura di probabilità  $\pi_{p_2}$  e la seconda da  $\pi_{p_1}$ . Quindi, qualsiasi permutazione degli indici  $\{1, \dots, d_1\}$  e  $\{1, \dots, d_2\}$  darà luogo alla stessa legge, poichè non farà altro che invertire l'ordine dei prodotti all'interno dell'integrale. Dunque è verificata la definizione di parziale scambiabilità.



# Conclusioni

Come spiegato nell'introduzione, la tesi opera in due direzioni principali. La prima consiste nel confrontare i due modelli di CreditRisk<sup>+</sup>: il modello misto di Bernoulli scambiabile ad un fattore in cui la somma dei defaults, condizionatamente al fattore macroeconomico, è una binomiale, e il modello misto di Poisson osservando che la distribuzione di Poisson è il limite della binomiale, che dunque l'approssima (l'approssimazione è precisa solo quando si considerano probabilità di insolvenza piccole). La seconda invece, di cui è stata proposta una piccola introduzione teorica senza esempi esplicativi, abbandona lo studio dei modelli ad un fattore e analizza modelli di mistura in cui, nella loro forma più generale, le variabili latenti sono multidimensionali e quindi dei vettori. Per quanto riguarda la prima direzione, invece, abbiamo dapprima calcolato il VaR con un modello misto di Bernoulli, il modello misto beta, vedendo come quest'ultimo si collocasse all'interno dei bounds forniti da [7]. È emerso che, tale valore, quando la probabilità di default è piccola, si stabilizza nei pressi del limite inferiore. Con l'aumentare di  $p$ , del coefficiente di correlazione e soprattutto del livello di confidenza  $\alpha$ , invece, si avvicina sempre di più al limite superiore. Tali ranges, sicuramente, non sono più verificati nel caso di un modello misto di Poisson. Essi non sono stati calcolati non essendo oggetto di studio della tesi ma, dopo aver effettuato gli stessi conti in questo ultimo scenario, considerando come modello misto di Poisson il modello misto gamma-Poisson, dalle analisi è emerso che il VaR comunque rientra nei bounds forniti in entrambi gli scenari considerati e per tutti e tre i livelli di correlazione scelti. Infatti, in entrambi i casi, il calcolo del Value-at-Risk conduce pressochè agli stessi risultati. Esso si discosta leggermente all'aumentare del coefficiente di correlazione  $\rho$ . Ad ogni modo, sebbene valgano le stesse considerazioni fatte per il modello misto di Bernoulli, il VaR, in questo caso, si avvicina al limite superiore più velocemente. Questo lascia pensare che il modello misto di Poisson sia più conservativo rispetto a quello di Bernoulli, giustificato anche dal fatto che si tratta di una sua approssimazione.



# Appendice A

## Variabili latenti e copule

**Definizione 23.** Una copula è una distribuzione multivariata con distribuzioni marginali normali standard.

Utilizziamo la notazione  $C(\mathbf{u}) = C(u_1, \dots, u_d)$  per le funzioni di ripartizione multivariate che sono copule. Quindi  $C$  è una funzione della forma  $C : [0,1]^d \rightarrow [0,1]$ . Le seguenti tre proprietà caratterizzano una copula  $C$ :

1.  $C(u_1, \dots, u_d)$  è crescente in ogni componente  $u_i$ ;
2.  $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$  per ogni  $i \in \{1, \dots, d\}$   $u_i \in [0,1]$ ;
3. per ogni  $(a_1, \dots, a_d), (b_1, \dots, b_d) \in [0,1]^d$  con  $a_i \leq b_i$  abbiamo:

$$\sum_{i_1=1}^2 \dots \sum_{i_d=1}^2 (-1)^{i_1+\dots+i_d} C(u_{1i_1}, \dots, u_{di_d}) \geq 0,$$

dove  $u_{j1} = a_j$  e  $u_{j2} = b_j$  per ogni  $j \in \{1, \dots, d\}$ .

Supponiamo che il vettore casuale  $\mathbf{X} = (X_1, \dots, X_d)'$  abbia funzione di ripartizione congiunta  $F$  con distribuzioni marginali  $F_1, \dots, F_d$ . Se applichiamo l'opportuna trasformazione di probabilità ad ogni componente otteniamo il vettore trasformato  $(F_1(X_1), \dots, F_d(X_d))$  la cui funzione di ripartizione è definita dalla copula che indichiamo con  $C$ . Segue che

$$\begin{aligned} F(x_1, \dots, x_n) &= P(F_1(X_1) \leq F_1(x_1), \dots, F_d(X_d) \leq F_d(x_d)) = & (A.1) \\ &= C(F_1(x_1), \dots, F_d(x_d)), \end{aligned}$$

o alternativamente  $C(u_1, \dots, u_n) = F(F_1^{\leftarrow}(u_1), \dots, F_d^{\leftarrow}(u_d))$ , dove  $F_i^{\leftarrow}$ , con  $i = 1, \dots, d$  indica l'inversa generalizzata della funzione di ripartizione  $F$ .

L'importanza delle copule nello studio delle distribuzioni multivariate è riassunto dal seguente elegante teorema che mostra, innanzitutto, come tutte le funzioni di ripartizione multivariate contengano copule, e poi che le copule possono essere usate nella congiunzione tra funzioni di ripartizione univariate e multivariate.

**Teorema 3 (Teorema di Sklar).** *Sia  $F$  la funzione di ripartizione congiunta con marginali  $F_1, \dots, F_d$ . Allora esiste una copula  $C : [0,1]^d \rightarrow [0,1]$  tale che per ogni  $x_1, \dots, x_d$  in  $[-\infty, \infty]$  vale la (A.1). Se le marginali sono continue, allora  $C$  è unica. Al contrario, se  $C$  è una copula e  $F_1, \dots, F_d$  sono delle funzioni di ripartizione, allora la funzione  $F$  data dalla (A.1) è una funzione di ripartizione congiunta con marginali  $F_1, \dots, F_d$ .*

Nei modelli a variabili latenti, sia  $S_i$  una variabile casuale che rappresenta un indicatore di stato per la compagnia  $i$  al tempo  $t + \Delta t$ . Scriveremo:

$$Y_i = 1 \iff S_i = 0 \quad Y_i = 0 \iff S_i > 0.$$

Prima di mostrare il collegamento tra modelli a variabili latenti e copule, diamo la seguente definizione:

**Definizione 24.** Sia  $\mathbf{X} = (X_1, \dots, X_d)'$  un vettore casuale  $d$ -dimensionale e sia  $W \in \mathbb{R}^{d \times h}$  una matrice deterministica di elementi  $w_{ij}$  tali che, per ogni  $i$ , gli elementi della riga  $i$ -esima formino un insieme di soglie crescenti che soddisfano  $w_{i1} < \dots < w_{ih}$ . Posto  $w_{i0} = -\infty$  e  $w_{i(h+1)} = \infty$  si ha

$$S_i = j \iff w_{ij} \leq X_i \leq w_{i(j+1)}, \quad j \in \{0, \dots, h\}, i \in \{1, \dots, d\}. \quad (\text{A.2})$$

Allora  $(\mathbf{X}, W)$  definisce un modello a variabili latenti per il vettore di stato  $\mathbf{S} = (S_1, \dots, S_d)'$ .

In altre parole, le variabili latenti sono variabili che non sono osservate direttamente ma che piuttosto sono dedotte da altre variabili osservate, per l'appunto le soglie deterministiche.

Ci riferiamo ad  $\mathbf{X}$  come il vettore delle variabili critiche e indichiamo la sua funzione di ripartizione marginale con  $F_i(x) = P(X_i \leq x)$ . L' $i$ -esima riga di  $W$  contiene le soglie critiche per la ditta  $i$ .

Dalla definizione, il default (corrispondente all'evento  $S_i = 0$ ) si presenta se  $X_i \leq w_{i1}$  cosicchè la probabilità di default della compagnia  $i$  è data da  $\bar{p}_i = F_i(w_{i1})$ . Nel contesto dei modelli descritti è importante distinguere tra il coefficiente di correlazione di default  $\rho(Y_i, Y_j)$  delle due compagnie con  $i \neq j$  dalla correlazione delle varibiali critiche  $X_i, X_j$ . Date le probabilità di default,  $\rho(Y_i, Y_j)$  è determinato da  $E(Y_i Y_j)$  come scritto in (3.1) e in un modello a variabili latenti

$$E(Y_i Y_j) = P(X_i \leq w_{i1}, X_j \leq w_{j1}).$$

**Definizione 25.** *Siano  $(\mathbf{X}, W)$  e  $(\tilde{\mathbf{X}}, \tilde{W})$  due modelli a variabili latenti per i vettori di stato  $\mathbf{S}$  e  $\tilde{\mathbf{S}}$ . I due modelli si dicono equivalenti se  $\mathbf{S} \stackrel{d}{=} \tilde{\mathbf{S}}$ .*

Diamo ora un semplice criterio per l'equivalenza tra due modelli a variabili latenti in termini delle distribuzioni marginali del vettore  $\mathbf{S}$  e della copula di  $\mathbf{X}$ .



**Lemma 3.** Siano  $(\mathbf{X}, W)$  e  $(\tilde{\mathbf{X}}, \tilde{W})$  due modelli soglia con vettori di stato  $\mathbf{S} = (S_1, \dots, S_d)'$  e  $\tilde{\mathbf{S}} = (\tilde{S}_1, \dots, \tilde{S}_d)'$ , rispettivamente. I due modelli sono equivalenti se valgono le seguenti condizioni:

(i) le distribuzioni marginali dei vettori casuali  $\mathbf{S}$  e  $\tilde{\mathbf{S}}$  coincidono, cioè:

$$P(S_i = j) = P(\tilde{S}_i = j), \quad j \in \{1, \dots, h\}, i \in \{1, \dots, d\}.$$

(ii)  $\mathbf{X}$  e  $\tilde{\mathbf{X}}$  ammettono la stessa copula  $C$ .

*Dimostrazione.* In accordo alla definizione 25,  $\mathbf{S} \stackrel{d}{=} \tilde{\mathbf{S}}$  se e solo se per ogni  $j_1, \dots, j_d \in \{1, \dots, h\}$  vale:

$$\begin{aligned} &P(w_{1j_1} < X_1 \leq w_{1(j_1+1)}, \dots, w_{dj_d} < X_d \leq w_{d(j_d+1)}) = \\ &= P(\tilde{w}_{1j_1} < \tilde{X}_1 \leq \tilde{w}_{1(j_1+1)}, \dots, \tilde{w}_{dj_d} < \tilde{X}_d \leq \tilde{w}_{d(j_d+1)}). \end{aligned}$$

Dalla teoria standard della misura, questa uguaglianza vale se e solo se, per ogni  $j_1, \dots, j_d \in \{1, \dots, h\}$ , si ha:

$$P(X_1 \leq w_{1j_1}, \dots, X_d \leq w_{dj_d}) = P(\tilde{X}_1 \leq \tilde{w}_{1j_1}, \dots, \tilde{X}_d \leq \tilde{w}_{dj_d}).$$

Dal teorema di Sklar (sulle copule) questo è equivalente a

$$C(F_1(w_{1j_1}), \dots, F_d(w_{dj_d})) = C(\tilde{F}_1(\tilde{w}_{1j_1}), \dots, \tilde{F}_d(\tilde{w}_{dj_d})),$$

dove  $C$  è la copula di  $\mathbf{X}$  e  $\tilde{\mathbf{X}}$ , avendo usato la condizione (ii). La condizione (i) implica che  $F_i(w_{ij}) = \tilde{F}_i(\tilde{w}_{ij})$  per ogni  $j \in \{1, \dots, h\}$  e  $i \in \{1, \dots, d\}$ .  $\square$



## Appendice B

# Teorema di rappresentazione di De Finetti

La proprietà della scambiabilità è strettamente legata all'uso di variabili casuali iid nei modelli statistici. Infatti, una sequenza di variabili aleatorie che sono iid, condizionate a qualche forma distribuzionale sottostante, è scambiabile. Questa implicazione discende direttamente dalla struttura della funzione di probabilità congiunta della forma iid. Il viceversa può essere stabilito attraverso il teorema di rappresentazione di Bruno De Finetti il quale afferma che, in qualsiasi sequenza infinita di variabili aleatorie scambiabili, esse sono condizionatamente indipendenti e identicamente distribuite. In realtà, il teorema originario affermava la tesi solo per variabili indicatrici casuali, anche se poi è stato dimostrato valere per ogni sequenza di variabili aleatorie. Ciò significa che sequenze infinite di variabili scambiabili possono essere considerate in modo equivalente come sequenze di variabili casuali condizionatamente iid, basate su una forma distributiva sottostante.

Per enunciare il teorema di rappresentazione di De Finetti andiamo ad introdurre  $\pi_p$  come misura di probabilità di Bernoulli su  $\{0,1\}$  data da  $\pi_p(1) = p$  e  $\pi_p(0) = 1 - p$ . Se le variabili  $Y_i$  sono iid in accordo a  $\pi_p$ , allora  $P(Y_i = 1) = \pi_p(1) = p$  e  $P(Y_i = 0) = \pi_p(0) = 1 - p$ .

**Teorema 4 (teorema di rappresentazione di De Finetti).** *Sia  $\{Y_i\}_{i \geq 1}$  una sequenza infinita di variabili casuali scambiabili che assumono valori in  $\{0,1\}$ . Allora esiste una legge  $G$  in  $[0,1]$  tale che per ogni  $d$  e per ogni sequenza  $y_1, \dots, y_d$  in  $\{0,1\}^d$  vale:*

$$P(Y_1 = y_1, \dots, Y_d = y_d) = \int \prod_{i=1}^d \pi_p(y_i) dG(p). \quad (\text{B.1})$$

In altre parole il teorema afferma che sequenze infinite scambiabili di Bernoulli possono essere modellate come un modello misto di Bernoulli scambiabile infinito.

Per la dimostrazione del teorema vedere [8].

**Osservazione 5.** Si noti che la tesi espressa dal teorema di De Finetti non vale per sequenze finite di variabili casuali scambiabili. Per esempio sia

$$P(Y_1 = 1, Y_2 = 0) = P(Y_1 = 0, Y_2 = 1) = \frac{1}{2}$$

e

$$P(Y_1 = 0, Y_2 = 0) = P(Y_1 = 1, Y_2 = 1) = 0.$$

La coppia  $(Y_1, Y_2)$  è scambiabile. Tuttavia, non esiste una misura di probabilità  $G$  che soddisfi il teorema. Infatti, se esistesse, si avrebbe

$$0 = P(Y_1 = 1, Y_2 = 1) = \int_0^1 p^2 dG(p)$$

ma allora  $G$  avrebbe densità di probabilità 1 nel punto 0, quindi non possiamo avere contemporaneamente

$$0 = P(Y_1 = 0, Y_2 = 0) = \int_0^1 (1 - p)^2 dG(p).$$

Quindi, nel caso specifico di modelli misti di Bernoulli ad un fattore, la (3.13) è una riformulazione del teorema di De Finetti, nel senso che la misura  $G$  è interpretata come legge del singolo fattore economico  $\Psi$ , che è una variabile unidimensionale. Stiamo, cioè, richiedendo che  $\mathbf{Y}$  soddisfi l'assunzione di De Finetti in modo che  $(Y_1, \dots, Y_d)$  siano  $d$  elementi di una sequenza scambiabile. In particolare, sono essi stessi scambiabili.

# Bibliografia

- [1] Conti, G.,(2017), *Le misure di rischio nelle banche: l'expected shortfall*.
- [2] Frey, R., McNeil, A.,(2002), *VaR and expected shortfall in portfolios of dependent credit risks: Conceptual and practical insights*.
- [3] Frey, R., McNeil, A.,(2003), *Dependent Defaults in Models of Portfolio Credit Risk*.
- [4] Artzner, P., Delbaen, F., Eber, J., Heath, D., (1999), *Coherent measures of risk*.
- [5] Embrechts, P., Frey, R., McNeil, A., (2005), *Quantitative Risk Management*.
- [6] Gordy, M.,(2000), *A comparative anatomy of credit risk models*. *Journal of Banking and Finance* 24:119-149.
- [7] Fontana, R., Luciano, E., Semeraro, P.,(2019), *Model risk in Credit Risk*.
- [8] Kirsh, W.,(2018), *An elementary proof of de Finetti's Theorem*.
- [9] Pollice, A., (a.a 2012-2013), *Statistica multivariata*  
[https://www.uniba.it/ricerca/dipartimenti/dse/dipartimento/personale/personale-docente/pollice/stat\\_mult/disp1.pdf](https://www.uniba.it/ricerca/dipartimenti/dse/dipartimento/personale/personale-docente/pollice/stat_mult/disp1.pdf).
- [10] Pollice, A., (a.a 2012-2013), *Statistica multivariata*  
[https://www.uniba.it/ricerca/dipartimenti/dse/dipartimento/personale/personale-docente/pollice/stat\\_mult/disp2.pdf](https://www.uniba.it/ricerca/dipartimenti/dse/dipartimento/personale/personale-docente/pollice/stat_mult/disp2.pdf).