

Dipartimento di Impresa e Management

Cattedra: Metodi quantitativi per l'impresa

Il rischio di credito, modelli e applicazioni.
Analisi sulla probabilità di *default* delle imprese

Prof. Alessandro Calvia

RELATORE

Prof. Stefano Marzioni

CORRELATORE

Antonino Ruvolo (matricola 721771)

CANDIDATO

Anno Accademico 2020/2021

INDICE

INDICE	I
INDICE DELLE FIGURE.....	III
INDICE DELLE TABELLE	IV
INTRODUZIONE.....	1
CAPITOLO 1. IL RISCHIO DI CREDITO	3
1.1 Il concetto di rischio in generale	3
1.2 Il rischio di credito	6
1.3 Relazione tra Rischio di Credito e Rating	9
1.3.1 Definizione e obiettivi del rating.....	9
1.3.2 Tipologie di rating e modalità di emissione	10
1.4 Le componenti del rischio di credito.....	12
1.4.1 La Perdita attesa (EL)	14
1.4.1.1 La probabilità di default (PD)	15
1.4.1.2 Il tasso di perdita in caso di insolvenza (LGD)	15
1.4.1.3 L'esposizione al momento dell'insolvenza (EAD).....	17
1.4.2 La perdita inattesa (UL).....	18
1.5 Perché approfondire il concetto di rischio?	19
CAPITOLO 2. MODELLI PER LA STIMA DELLA <i>PROBABILITY OF DEFAULT</i>	21
2.1 Costruzione dei modelli.....	21
2.1.1 Lo sviluppo e la convalida dei modelli	22
2.1.2 Modelli per contratti attivamente negoziati e per contratti fuori standard.....	25
2.2 I modelli <i>Rating System</i>	28
2.2.1 <i>Credit Rating Migration</i>	28
2.3 I modelli <i>credit scoring</i>	31
2.3.1 Il modello di Altman.....	31
2.3.2 I modelli logit, probit e le reti neurali	35
2.4 Modelli alternativi.....	40
2.4.1 Il metodo RAPD	40
2.4.2 La logica <i>fuzzy</i>	40
2.5 Il modello di Merton e il modello KMV	42
2.5.1 L'approccio del modello di Merton.....	42
2.5.2 Il calcolo della <i>probability of default</i>	43
2.5.3 Il modello KMV	46
2.5.3.1 Stima dell' <i>asset value A</i> e dell' <i>asset volatility σ_A</i>	47
2.5.3.2 Stima della <i>distance to default</i>	47
2.5.3.3 Stima della <i>probability of default</i>	49

2.5.3.4	Calcolo della PD tramite la formula di Black e Scholes	51
2.6	I modelli in forma ridotta	54
2.6.1	L'intensità di default e il valore di recupero	54
2.6.2	Estensione al caso multivariato	58
2.7	Accordi internazionali e modelli sviluppati da banche.....	59
2.7.1	Il modello di Vasicek	59
2.7.2	Credit Risk Plus.....	61
2.7.3	Creditmetrics	65
CAPITOLO 3. ANALISI DELLA STIMA DELLA <i>PROBABILITY OF DEFAULT</i>		68
3.1	Caratteristiche del modello e presentazione dei campioni utilizzati	68
3.1.1	Significatività del modello.....	71
3.1.2	Presentazione dei campioni costruiti.....	73
3.2	Presentazione dei regressori e analisi del campione statunitense	76
3.2.1	Analisi e risultati	82
3.3	Considerazioni ulteriori tramite la costruzione di un campione europeo.....	89
3.4	Appendice: valori utilizzati per il calcolo dei regressori	94
3.4.1	Campione statunitense.....	94
3.4.2	Campione di verifica statunitense	97
3.4.3	Campione europeo	98
CONCLUSIONI		101
BIBLIOGRAFIA		103

Indice delle figure

<i>Figura 1 Distribuzione delle perdite e relativo valore atteso. Fonte: Gordon, D., "Aspects of Modeling Fraud Prevention of Online Financial Services", (2015).</i>	13
<i>Figura 2 Grafico della funzione logistica e della normale. Fonte: Brunello, P., "La probabilità di sofferenza nei finanziamenti alle famiglie", (2008).</i>	37
<i>Figura 3 Esempificazione della struttura multistrato di una rete neurale. Fonte: D'annunzio, N., Falavigna, G., "Modelli di analisi e previsione del rischio di insolvenza. Una prospettiva delle metodologie applicate", Ceris-Cnr, W.P. n° 17/2004.</i>	39
<i>Figura 4 Funzionamento dell'inferenza con logica fuzzy. Fonte: Cosma, S., Facchinetti, G., Mastroleo, G., Ferretti, R., "I sistemi di credit scoring: un'applicazione della logica fuzzy nella valutazione delle piccole e medie imprese" (2001).</i>	41
<i>Figura 5 Distribuzione del valore di mercato dell'attivo e probabilità di insolvenza. Fonte: "Introduzione al modello di Merton/KMV per la stima delle probabilità di insolvenza di società quotate", a cura di C. Epis</i>	44
<i>Figura 6 Payoff per gli azionari e per i creditori. Fonte: "Approcci matematico-statistici alla valutazione del rischio di credito". A cura di A. Matta</i>	45
<i>Figura 7 Valore dell'attivo e valore dell'equity. Fonte: "Sironi e Marsella (1999)"</i>	45
<i>Figura 8 Indicazione delle variabili rilevanti all'interno del modello KMV. Fonte "Crosbie (1997)"</i>	48
<i>Figura 9 Rappresentazione grafica di una tabella di frequenza. Fonte "Introduzione al modello di Merton/KMV per la stima delle probabilità di insolvenza di società quotate", a cura di C. Epis</i>	50
<i>Figura 10 Illustrazione di un processo di Poisson con indicazione dell'istante in cui si verifica l'insolvenza. Fonte: "Approcci matematico-statistici alla valutazione del rischio di credito", a cura di A. Matta.</i>	55
<i>Figura 11 Distribuzione delle perdite per insolvenza. Asimmetria positiva (default correlation $\neq 0$). Fonte: Hull, J. C., "Risk Management e Financial Institution". John Wiley & Sons (2018).</i>	65
<i>Figura 12 Rappresentazione grafica della correlazione lineare tra regressori. Fonte: propria elaborazione</i>	80

Indice delle tabelle

<i>Tabella 1 Sintesi delle diverse fattispecie di rischio indicate negli accordi di Basilea II. Fonte: "Basel Committee on Banking Supervision", 2002.....</i>	<i>4</i>
<i>Tabella 2 Sintesi delle principali tipologie di rischio. Fonte: rielaborazione delle categorie di rischio tratte da McNeil, A., Frey, R., & Embrechts, P. Quantitative Risk Management. Princeton: Princeton University Press (2015).</i>	<i>6</i>
<i>Tabella 3 Classificazione delle principali agenzie di rating. Fonte: Borsa Italiana.</i>	<i>12</i>
<i>Tabella 4 Probabilità di migrazione da un rating a un altro entro un anno. Fonte: Moody's.</i>	<i>29</i>
<i>Tabella 5 Media cumulata dei tassi di default (%). Fonte: Moody's.....</i>	<i>29</i>
<i>Tabella 6 Tabella di frequenza. Fonte "Estrella (2000)"</i>	<i>50</i>
<i>Tabella 7 Corrispondenza tra PD e classi di rating. Fonte: "Bellucci et al. (1997)"</i>	<i>51</i>
<i>Tabella 8 Tassi di insolvenza annui. Fonte: Moody's.....</i>	<i>63</i>
<i>Tabella 9 Distribuzione probabilistica del numero di insolvenze per $\mu=4$. Fonte: J. C. Hull, "Risk Management e Istituzioni Finanziarie", 2018.....</i>	<i>64</i>
<i>Tabella 10 CreditMetrics: esempio di tabella di transizione. Fonte: J. C. Hull, "Risk Management e Istituzioni Finanziarie", 2018 .</i>	<i>67</i>
<i>Tabella 11 Tabella di confronto tra risultati effettivi e previsti. Fonte: propria elaborazione.....</i>	<i>73</i>
<i>Tabella 12 Presentazione del campione statunitense. Fonte: Database Refinitiv.....</i>	<i>74</i>
<i>Tabella 13 Presentazione del campione di verifica statunitense. Fonte: Database Refinitiv.</i>	<i>75</i>
<i>Tabella 14 Presentazione del campione europeo. Fonte: Database Refinitiv.</i>	<i>76</i>
<i>Tabella 16 Dati per calcolo correlazione lineare tra regressori. Fonte: propria elaborazione.....</i>	<i>79</i>
<i>Tabella 15 Composizione del campione statunitense e valori regressori. Fonte: propria elaborazione.....</i>	<i>82</i>
<i>Tabella 17 Valori medi dei regressori nel campione statunitense. Fonte: propria elaborazione.</i>	<i>83</i>
<i>Tabella 18 Calcolo dei coefficienti e del p-value del modello 1 sul campione statunitense. Fonte: propria elaborazione.</i>	<i>83</i>
<i>Tabella 19 Calcolo dei coefficienti e del p-value del modello 2 sul campione statunitense. Fonte: propria elaborazione.</i>	<i>84</i>
<i>Tabella 20 Calcolo dei coefficienti e del p-value del modello 3 sul campione statunitense. Fonte: propria elaborazione.</i>	<i>84</i>
<i>Tabella 21 Calcolo dei coefficienti e del p-value del modello 4 sul campione statunitense. Fonte: propria elaborazione.</i>	<i>84</i>
<i>Tabella 22 Classificazione delle aziende secondo il modello costruito. Fonte: propria elaborazione.....</i>	<i>86</i>
<i>Tabella 23 Analisi degli errori per tipologia. Fonte: propria elaborazione.</i>	<i>86</i>
<i>Tabella 24 Classificazione delle aziende del campione di controllo secondo il modello costruito. Fonte: propria elaborazione.</i>	<i>88</i>
<i>Tabella 25 Analisi degli errori per tipologia sul campione di verifica. Fonte: propria elaborazione.....</i>	<i>88</i>
<i>Tabella 26 Valori medi dei regressori nel campione europeo. Fonte: propria elaborazione.</i>	<i>89</i>
<i>Tabella 27 Composizione del campione europeo e valori regressori. Fonte: propria elaborazione.</i>	<i>91</i>
<i>Tabella 28 Calcolo dei coefficienti e del p-value del modello 1 sul campione europeo. Fonte: propria elaborazione.</i>	<i>91</i>
<i>Tabella 29 Calcolo dei coefficienti e del p-value del modello 2 sul campione europeo. Fonte: propria elaborazione.</i>	<i>91</i>
<i>Tabella 30 Calcolo dei coefficienti e del p-value del modello 3 sul campione europeo. Fonte: propria elaborazione.</i>	<i>92</i>
<i>Tabella 31 Calcolo dei coefficienti e del p-value del modello 4 sul campione europeo. Fonte: propria elaborazione.</i>	<i>92</i>
<i>Tabella 32 Classificazione delle aziende secondo il modello costruito. Fonte: propria elaborazione.....</i>	<i>94</i>
<i>Tabella 33 Analisi degli errori per tipologia. Fonte: propria elaborazione.</i>	<i>94</i>
<i>Tabella 34 Valori di bilancio utilizzati per il calcolo dei regressori. Campione statunitense. Fonte: Database Refinitiv.</i>	<i>96</i>
<i>Tabella 35 Valori di bilancio utilizzati per il calcolo dei regressori. Campione di verifica statunitense. Fonte: Database Refinitiv. .</i>	<i>98</i>
<i>Tabella 36 Valori di bilancio utilizzati per il calcolo dei regressori. Campione di europeo. Fonte: Database Refinitiv.....</i>	<i>100</i>

Introduzione

Il rischio di credito è un tema sempre attuale. Un'esposizione eccessiva nei confronti di debitori che versino in situazioni di emergenza finanziaria può portare gli istituti finanziari verso il *default* e generare ricadute sull'intera economia. È facile allora immaginare quanto sia delicato il tema della stima del rischio di credito delle controparti.

Negli ultimi venti anni abbiamo avuto modo di assistere a due crisi rilevanti. La prima, iniziata nel 2007, ha dimostrato quali possano essere i danni causati dall'eccessiva erogazione di prestiti senza prestare adeguata attenzione all'analisi della probabilità di *default* della controparte. La seconda, tuttora in corso, è stata causata dalla pandemia e dal conseguente *lockdown* imposto da molti governi per arginare e combattere la diffusione del virus. Ciò ha tuttavia portato a una riduzione dei ricavi delle imprese e a una maggiore necessità di approfondite analisi prima della concessione di eventuali crediti da parte degli intermediari finanziari.

Per rispondere a queste esigenze, è necessario dotarsi di strumenti sempre più raffinati, modelli in grado di stimare efficacemente il livello di rischio di una impresa, lasciando poi ovviamente al *management* il compito di stabilire il livello di rischio ritenuto accettabile.

Lo scopo del presente lavoro è quello di evidenziare l'importanza del rischio di credito per l'intera economia e costruire un modello che possa provare a stimare la probabilità di *default* di un campione di imprese.

Il lavoro si articola in tre capitoli strutturati come segue. Il capitolo 1 si apre con la descrizione del concetto di rischio, fondamentale e ineludibile componente di qualsiasi attività d'impresa. Si prosegue con la definizione del rischio di credito e di *rating* e, in particolare, evidenziando come quest'ultimo possa fornire un aiuto nella comprensione del livello di rischio che potenzialmente si decide di sostenere. Si procede poi con la scomposizione del rischio di credito nelle sue componenti, esplicitando il concetto di perdita attesa, funzione dell'*exposure at default*, della *probability of default* e della *loss given default*, e di perdita inattesa, stimata dagli istituti finanziari per salvaguardarsi dalle perdite che in passato hanno ecceduto le perdite medie. Nell'ultimo paragrafo del capitolo 1, infine, si evidenziano le conseguenze economiche che una mancata attenzione nei confronti del rischio di credito può generare sia dal punto di vista degli *stakeholders* delle aziende sia da quello della società in senso lato.

Il capitolo 2 si apre con una panoramica delle *best practices* in tema di sviluppo e convalida dei modelli per lo studio di fenomeni economici. In seguito, ci si concentra sui modelli utilizzabili per la stima della *probability of default*. In particolare, si evidenziano differenti approcci all'analisi come il *credit rating migration*, basato sul concetto di *rating*, e il *credit scoring*, basato sull'idea dell'attribuzione di uno *score* alle aziende, permettendone l'associazione a una determinata affidabilità creditizia. Viene qui sottolineata l'importanza del modello logit, poi utilizzato per l'analisi dei dati presentata successivamente.

Negli ultimi tre paragrafi del capitolo vengono invece presentati i modelli in forma ridotta, i modelli strutturali (in particolare facendo riferimento sia al modello di Merton sia al modello KMV) e alcuni noti modelli sviluppati da istituti finanziari, come Credit Risk Plus, proposto negli anni '90 da Credit Suisse, e Creditmetrics, proposto nel '97 da J. P. Morgan. Per tutti i modelli vengono indicati la logica di base, la costruzione, le potenzialità e i limiti.

Infine, nel capitolo 3 si presenta il lavoro di analisi dati svolto. L'obiettivo è quello di sviluppare un modello logit tramite cui stimare la *probability of default* delle imprese. Ulteriore aspetto che si vuole mettere in evidenza è l'adattamento della capacità previsionale del modello al nuovo contesto socioeconomico creatosi in seguito al diffondersi del Covid-19. Infatti, viene costruito un primo campione di aziende che prende in considerazione dati di bilancio precedenti al 2020. Il campione di controllo viene poi costruito tramite dati relativi a bilanci relativi all'anno fiscale 2020.

L'analisi inizia con la scelta del tipo di società da inserire nel campione e del mercato di quotazione di riferimento. Si precede poi con la raccolta dei dati utili tramite il *database Refinitiv*. Viene così costruito un *training set* costituito da un campione di 58 aziende e il campione di verifica formato da 52 società, tutte quotate sul mercato statunitense. Inoltre, è stato costruito un ulteriore campione formato da 61 aziende quotate sui mercati europei.

Si procede poi alla presentazione dei cinque regressori selezionati per l'analisi e alla regressione logistica che porta al calcolo dei loro coefficienti.

Il modello statunitense e quello europeo vengono costruiti seguendo una logica di tipo *backward*, ossia partendo dai modelli con tutti i regressori individuati e rimuovendo di volta in volta quello con *p-value* maggiore. Otteniamo infine così dei modelli in cui tutti i regressori possono offrire significativa capacità di previsione.

CAPITOLO 1. Il rischio di credito

1.1 Il concetto di rischio in generale

Al concetto di rischio siamo tutti indubbiamente familiari, tuttavia, esso è noto per lo più come la possibilità che un evento si sviluppi in maniera negativa arrecandoci un danno. Difficilmente il rischio viene associato in primo luogo a qualcosa di positivo. A conferma di ciò, consultando il dizionario Treccani si ribadisce che il rischio sia l'“eventualità di subire un danno connessa a circostanze più o meno prevedibili”¹ mentre l'*Oxford English Dictionary* lo definisce come “a chance of bad consequences, loss or exposure to mischance”. Ancora una volta quindi sottolineiamo che queste definizioni sono deficitarie di un aspetto essenziale: il rischio porta con sé anche opportunità ossia i soggetti che decidono di sopportarlo lo fanno perché intravedono una possibilità di guadagno. Non a caso vedremo che sarà essenziale studiare dei modelli che permettano di quantificare il rischio per poterne apprezzare il valore e applicarlo alle transazioni economiche².

Esistono indubbiamente diversi tipi di rischio e in questa sezione vogliamo passare in rassegna i più importanti in ambito finanziario per poi concentrarci sul rischio di credito. Si vuole quindi evidenziare l'ampiezza del concetto di rischio e la sua vitalità non solo per i mercati finanziari e per i soggetti che quotidianamente vi operano ma, in generale, per l'intera società la cui salute finanziaria si basa sulla corretta comprensione e sul corretto controllo di tali rischi. Secondo quanto osservato da *McNeil, Frey ed Embrechts*³ possiamo distinguere principalmente i seguenti rischi:

1. Rischio di mercato: consiste nella variazione di valore di una posizione finanziaria o di un portafoglio dovuti a un cambiamento di variabili importanti del mercato da cui i sottostanti del portafoglio dipendono. Alcuni esempi possono essere i prezzi di azioni, obbligazioni e materie prime oppure i tassi di interesse e di cambio⁴. È prassi diffusa coprirsi da tali rischi tramite l'utilizzo di strumenti finanziari derivati che permettono di rendere neutrali le fluttuazioni del mercato;
2. Rischio operativo: viene definito come il rischio di subire delle perdite dovute a inadeguate procedure interne, errori commessi dal personale o perdite legate a eventi esterni all'organizzazione diversi dalle perdite attese. Questa è una definizione “in positivo” utilizzata anche dal *Solvency II Framework* che differisce dalla definizione “negativa” che era stata adottata nella regolamentazione degli accordi di Basilea I. Il rischio operativo veniva infatti inteso come qualsiasi rischio non classificabile come rischio di mercato o di credito. All'interno della famiglia del rischio operativo, secondo gli accordi di Basilea II, rientrano diverse fattispecie di rischio, di seguito brevemente riassunte.

¹ Confronta <https://www.treccani.it/vocabolario/rischio/>.

² Confronta F. Knight, “Risk Uncertainty and Profit”, 1921

³ Confronta McNeil, A., Frey, R., & Embrechts, P. *Quantitative Risk Management*. Princeton: Princeton University Press (2015).

⁴ Confronta <https://www.borsaitaliana.it/borsa/glossario/rischio-di-mercato.html>.

Frode interna	Perdita dovuta al comportamento fraudolento che miri a aggirare la regolamentazione interna della società o la legge, coinvolgendo almeno una parte "interna" della società.
Frode esterna	Perdita dovuta al comportamento fraudolento assunto con la finalità di aggirare la regolamentazione interna della società o la legge, coinvolgendo solo terze parti.
Sicurezza sul lavoro	Perdita relativa ad azioni non coerenti con leggi o accordi sull'occupazione, la salute o la sicurezza. Può derivare anche da risarcimenti per lesioni personali o per discriminazione.
Attività relative ai clienti, ai prodotti e al business	Perdita dovuta a un'insoddisfazione non intenzionale di un'obbligazione professionale verso uno specifico cliente oppure dovuta dalla natura o l'implementazione di un prodotto.
Danneggiamento fisico di un asset	Perdita conseguente al danneggiamento fisico di un asset in seguito a eventi naturali o di altra tipologia.
Interruzione dell'attività	Perdita conseguente all'interruzione dell'attività o del mancato funzionamento dei sistemi aziendali (come ad esempio <i>hardware</i> , <i>software</i> , telecomunicazioni o <i>utilities</i>).
Esecuzione, consegna e <i>process management</i>	Perdite relative a errori nelle attività di processamento delle transazioni con clienti e fornitori.

Tabella 1 Sintesi delle diverse fattispecie di rischio indicate negli accordi di Basilea II. Fonte: "Basel Committee on Banking Supervision", 2002

3. *Model risk*: è definito come la perdita potenziale cui una società va incontro come conseguenza di decisioni prese prevalentemente da *output* prodotti tramite modelli interni, dovuta al fatto che quei modelli possono avere errori di sviluppo, implementazione o anche di utilizzo⁵. I rischi legati ai modelli possono essere particolarmente rilevanti soprattutto se combinati ad altri eventi potenzialmente dannosi.

Un esempio che possa farne comprendere l'importanza è lo scandalo che coinvolse J. P. Morgan noto come "*The London Whale*" le cui conseguenze furono significative. Infatti, la banca andò incontro a perdita stimate in sei miliardi di dollari e fu multata per ulteriori 920 milioni. L'accaduto può essere sintetizzato come segue. Il *Chief Investment Officer* (CIO) era il responsabile degli investimenti della liquidità in eccesso derivante dai depositi. Per coprirsi da eventuali fasi di recessione del mercato furono acquistati degli strumenti finanziari derivati di copertura⁶. In una prima fase questi strumenti derivati erano detenuti a scopo di copertura ma ben presto il portafoglio divenne speculativo ed eccessivamente rischioso, incrementando notevolmente il suo valore. Il rischio legato ai modelli entra in gioco proprio nel 2012 quando la banca, al posto di cercare di ridimensionare il rischio, cambia la propria metodologia di calcolo del VaR. Tuttavia, si verifica un errore nei calcoli del modello che portano a una sottostima del debito di circa il 50%. Questo errore ha permesso al portafoglio di

⁵ Per maggiori approfondimenti si veda la Direttiva 2013/36/UE sui requisiti patrimoniali (CRD IV, articolo 3.1.11).

⁶ I contratti derivati acquistati da JP Morgan sono noti come *synthetic CDS derivatives*.

continuare a crescere fino a quando la banca fu duramente colpita dalla crisi dei debiti sovrani europei;

4. **Rischio di liquidità:** si definisce tale il rischio che un'entità non riesca a far fronte ai propri impegni di pagamento quando essi giungono a scadenza. Normalmente, una società è in grado di fronteggiare le proprie uscite di cassa in diversi modi. Ad esempio, mediante i flussi in entrata, le attività prontamente liquidabili oppure la propria capacità di ottenere credito. Tuttavia, le attività prontamente liquidabili possono presentare delle problematiche non sempre legate solo alla società. Può infatti accadere che sui mercati si manifestino tensioni che ne rendano difficoltosa, se non impossibile, la vendita o l'utilizzo come garanzia in cambio di fondi. In queste condizioni, il rischio di liquidità è strettamente legato alle condizioni di liquidità del mercato. Questo, tra l'altro, solleva un punto importante. Può essere un grave errore considerare i rischi come aspetti a sé stanti. Occorre invece tenere bene a mente le loro interazioni e le conseguenti problematiche che ne possono derivare. Per combattere tale rischio ogni società dovrebbe dotarsi di linee guida di governo del rischio di liquidità, ossia l'organizzazione e la cultura aziendale che delineano l'insieme dei principi, delle metodologie, delle norme e dei processi necessari a prevenire l'insorgere di situazioni di crisi di liquidità e prevedono che vengano sviluppati approcci prudenziali nella sua gestione con l'obiettivo di mantenere il profilo di rischio su livelli ritenuti adeguati⁷;
5. *Underwriting risk:* particolarmente rilevante in ambito assicurativo, si tratta del rischio cui va incontro il sottoscrittore relativo alla perdita o alla variazione sfavorevole del valore delle passività, derivante dalla possibilità che i premi non siano sufficienti a coprire i sinistri più le spese. Emerge in maniera prepotente l'importanza, in questo caso, di idonei modelli che possano rendere il management consapevole delle opportunità e dei rischi cui sta andando incontro con la necessità di ridurre al entro range accettabili il rischio stesso. In ambito assicurativo un errore di valutazione di questo rischio può causare enormi perdite finanche il fallimento della società;
6. **Rischio di credito:** è il rischio che il creditore non sia in grado di onorare i propri obblighi relativamente al pagamento degli interessi e del capitale. Per una trattazione più completa e adeguata di questa tipologia di rischio si rinvia al paragrafo successivo.

Si propone qui un breve riassunto delle varie tipologie di rischio brevemente introdotte nella sezione precedente.

⁷ Confronta <https://group.intesasanpaolo.com/en/>.

Tipologia di Rischio	Breve descrizione
1. Mercato	Cambiamento del valore di una posizione finanziaria o di un portafoglio legato al cambio di valore dei sottostanti.
2. Operativo	Perdite derivanti da inadeguatezza delle procedure interne, da errori commessi dal personale della società o, infine, da eventi esterni.
3. Modello	Rischio legati all'utilizzo di un modello errato e di una sua (anche parziale) errata applicazione.
4. Liquidità	Rischio che la controparte non sia in grado di onorare i propri impegni alla scadenza.
5. Sottoscrizione	Rischio legato alla errata stima dei premi in relazione alla rischiosità e gravità degli eventi assicurati.
6. Credito	Rischio che una variazione inattesa del merito creditizio di una controparte generi una corrispondente variazione inattesa del valore di mercato della posizione creditoria.

Tabella 2 Sintesi delle principali tipologie di rischio. Fonte: rielaborazione delle categorie di rischio tratte da McNeil, A., Frey, R., & Embrechts, P. *Quantitative Risk Management*. Princeton: Princeton University Press (2015).

1.2 Il rischio di credito

Il rischio di credito viene definito da Borsa Italiana come il rischio che il debitore non sia in grado di adempiere ai suoi obblighi di pagamento di interessi e di rimborso del capitale⁸. Per essere più specifici si potrebbe dire che il rischio di credito è rappresentato dal rischio che una variazione non attesa del merito creditizio⁹ di una controparte generi una variazione del valore della posizione creditoria. È evidente che se fosse possibile stimare a priori in maniera esatta e sicura la perdita creditizia avrebbe poco senso parlare di rischio di credito: l'alea risulta quindi essere un elemento essenziale e parte integrante di questo rischio¹⁰.

Il perimetro del rischio di credito è più ampio di quanto possa apparire a una prima analisi. Non ci si limita infatti alla sola possibilità di insolvenza della controparte ma anche il solo deterioramento del merito creditizio della stessa deve essere considerato a tutti gli effetti una manifestazione di tale categoria di rischio. Quanto fin qui analizzato fa riferimento al cosiddetto *full credit risk* ma si fa spesso riferimento ad altri due tipi di rischio di credito: il *delivery risk* e il *substitution risk*. Incontriamo il primo quando le controparti sono titolari di obbligazioni reciproche ed esse vanno eseguite contemporaneamente. Il rischio consiste allora nella possibilità che una delle controparti non adempia ai propri obblighi anche dopo che l'altra abbia effettuato la propria prestazione. Parliamo invece di *substitution risk* quando, in presenza di contratti a

⁸ Confronta <https://www.borsaitaliana.it/borsa/glossario/rischio-di-credito.html#:~:text=Rischio%20che%20il%20debitore%20non,e%20di%20rimborso%20del%20capitale.>

⁹ Il merito creditizio può essere pensato come l'affidabilità di un soggetto in termini economico-finanziari. Essa influisce sul rischio finanziario connesso alla transazione con il soggetto stesso.

¹⁰ Si veda anche Bielecki e Rutkowski, "Credit Risk", 2010.

prestazioni corrispettive¹¹, un contraente non effettua la propria prestazione e la parte *in bonis* sarà costretta a sopportare un minor guadagno o un maggior costo in diretta conseguenza dell'inadempimento contrattuale della controparte. In questo caso risulta plausibile la stipula di un nuovo contratto spesso a condizioni meno favorevoli per la parte solvente.

Quando parliamo di rischio di credito può essere utile approfondirne alcuni aspetti in quanto, come precedentemente anticipato, non si parla di rischio di credito solo in presenza di *default* della controparte.

Possiamo quindi anche introdurre i seguenti approfondimenti:

1. Il rischio di migrazione: bisogna osservare che il peggioramento del merito creditizio di una controparte non deve dar luogo necessariamente a una perdita economica immediata a meno che tale esposizione creditizia non abbia come origine un'attività negoziata in un mercato secondario liquido come potrebbe essere, ad esempio, quello dei *corporate bond*;
2. Il rischio di recupero: è legato alla possibilità che il tasso di recupero consuntivo connesso alle esposizioni nei confronti delle controparti risulti significativamente inferiore a quanto inizialmente stimato;
3. Il rischio di esposizione: esso rappresenta il rischio che la dimensione dell'esposizione nei confronti di una controparte aumenti inaspettatamente proprio in corrispondenza del periodo appena precedente il verificarsi dell'insolvenza;
4. Il rischio di spread: si fa in questo caso riferimento alla possibilità che, a parità di merito creditizio, aumenti il premio al rischio (*spread*) richiesto dal mercato di capitali. Situazione tipica in cui potremmo assistere al manifestarsi di tale rischio è una crisi di liquidità dei mercati o in un incremento dell'avversione al rischio degli investitori.

Inoltre, il rischio di insolvenza viene diviso in tre componenti principali da Sironi e Marsella¹²:

1. il tasso di perdita attesa: esso è il valore medio della distribuzione dei tassi di perdita. Tramite modelli matematico-statistici è possibile determinare questo valore con relativa semplicità ed è quindi relativamente facile da prevedere. Per tale motivo spesso non viene considerato come vera e propria identificazione del rischio;
2. La variabilità della perdita attorno al valor medio: il valore è relativo alla volatilità della perdita, quindi, non è possibile prevederlo *ex ante* con facilità. Pertanto, essa rappresenta la vera componente del rischio poiché solamente *ex post* sarà possibile quantificare la perdita superiore a quella stimata *ex ante*;

¹¹ Si tratta della categoria di contratti in cui le prestazioni dovute dalle parti sono tra loro connesse, al punto che l'una costituisce il corrispettivo dell'altra. Si parla in tal caso di nesso "sinallagmatico" tra le obbligazioni nascenti dal contratto.

¹² Si veda: Sironi A., Marsella M. (1998), "La Misurazione e la gestione del rischio di credito. Modelli, strumenti e politiche", Bancaria editrice.

3. l'effetto diversificazione: la diminuzione che il tasso di perdita inattesa subisce quando, nello stesso portafoglio, vengono inseriti "impieghi" con tassi di perdita inattesa correlati in modo imperfetto. Questa componente è l'unica potenzialmente controllabile che permette una previsione del rischio di credito. Infatti, le variazioni delle probabilità di insolvenza fra settori diversi ed aree geografiche diverse sono molto forti e diventano molto più interessanti nelle fasi di crisi economica. Dunque, un'adeguata politica di diversificazione permette alle imprese di ridurre significativamente, a parità di rendimento, il rischio globale del portafoglio.

In ogni caso, è sempre bene tenere a mente che il rischio di credito è una componente fondamentale di ciascun investimento e non può essere del tutto eliminato. Anzi, come anticipato all'inizio del capitolo, non si deve intendere il rischio come un'eventualità solo negativa, ossia una possibilità che ci venga arrecato un danno finanziario ma, al contrario, come un'opportunità nelle scelte di investimento. È chiaro che una volta stabilito il livello di rischio che si è in grado di sopportare sarà possibile cercare quelle opzioni di investimento che possano premiarci con il maggior rendimento possibile. Il rischio di credito è tuttavia influenzato da diversi fattori che possono essere o meno legati alla situazione del debitore. Facciamo in particolare riferimento all'andamento del ciclo economico per quegli aspetti esterni alla posizione del debitore, e al rischio emittente o rischio specifico per gli aspetti inerenti al debitore stesso. Risulta evidente che in periodi di espansione economica il rischio di credito, seppur da tenere in debitamente in considerazione, assume una rilevanza minore rispetto ai periodi di recessione.

In ambito finanziario, la tipologia di distinzione tra generico e specifico è molto diffusa e valida per tutti i tipi di investimento. Ciò vuol dire anche che esistono anche differenze tipiche che variano a seconda delle *asset class*, si veda ad esempio la differenza di rischio tra un investimento azionario e uno obbligazionario. Tuttavia, in questo paragrafo vogliamo approfondire la differenza tra rischio generico e specifico. L'andamento dell'economia globale con le variazioni continue e più o meno rilevanti dei prezzi di borsa ma anche altri aspetti come il grado di efficienza del sistema finanziario visto nel suo complesso contribuiscono alla definizione di rischio generico, detto anche rischio sistemico perché non eliminabile. In altre parole, facciamo riferimento a quella variabilità strettamente correlata e dipendente dalle fluttuazioni dei mercati, fluttuazioni che hanno quindi impatto sull'intera varietà dei titoli. Nelle fasi di recessione economica sarà quindi probabile che anche i nostri investimenti risentano della generale condizione economica. Il caso opposto è invece rappresentato dal rischio specifico in cui l'entità del rischio dipenderà in primo luogo e maggiormente delle caratteristiche peculiari da valutare caso per caso. Se il soggetto economico cui abbiamo prestato denaro si trovasse in difficoltà potremmo andare incontro a una riduzione del valore dell'investimento o al mancato percepimento di interessi. Se da un lato il rischio generico non è evitabile, il rischio specifico può essere più facilmente contrastato ricorrendo alla diversificazione. È quindi possibile ridurre il proprio legame, e di conseguenza la propria esposizione, all'andamento del singolo investimento

controbilanciandola con l'andamento degli altri investimenti presenti nel proprio portafoglio. Tuttavia, questo è un obiettivo non sempre da facile realizzazione in quanto i tradizionali concetti di "frammentazione" e "concentrazione" del rischio non sempre sono adeguati a riflettere la realtà. Quando parliamo di concentrazione facciamo riferimento al numero di crediti presenti in portafoglio e al loro peso in relazione al valore dell'intero portafoglio. Sembra ragionevole ritenere che all'aumentare del numero di crediti il portafoglio risulti maggiormente diversificato e quindi meno rischioso rispetto a un altro in cui sono detenuti pochi crediti di valore molto significativo. La concentrazione e la frammentazione, ossia semplicemente il suo contrario, sono concetti molto utilizzati e studiati anche con modelli abbastanza semplici basati ad esempio sull'indice di Hirschmann-Herfindahl¹³ e sull'indice di concentrazione delle quattro imprese¹⁴. Tuttavia, due portafogli con la stessa concentrazione sono diversamente diversificati se sono composti da controparti che si muovono in modo scarsamente correlato; da qui ritroviamo l'importanza della correlazione con il mercato.

1.3 Relazione tra Rischio di Credito e Rating

1.3.1 Definizione e obiettivi del rating

Il rating può essere definito come una sintetica valutazione della solvibilità dei titoli e delle imprese. È, in altre parole, un giudizio espresso da un soggetto che si possa ritenere esterno e soprattutto indipendente, ossia le agenzie di rating. Esse si esprimono quindi sulla capacità di un soggetto di poter onorare o meno i propri debiti tramite le risorse generate della propria attività. Essendo il giudizio legato a diversi fattori che mutano nel tempo deve essere soggetto a periodica revisione. Questa informazione può essere particolarmente preziosa perché facilmente reperibile, non costosa e di facile e immediata interpretazione. Il rating, infatti, fornito da un soggetto serio, indipendente e credibile può essere usato per monitorare il valore e la qualità degli investimenti nel seguente modo: un eventuale *downgrade* da parte di un'agenzia fornisce un chiaro segnale che i mercati e gli investitori possono facilmente cogliere e che, senza questo ausilio, magari non sarebbero riusciti a prevedere in tempo utile.

La solidità attribuita dalle agenzie di rating ha un impatto notevole sull'economia in quanto contribuisce a determinare il rendimento richiesto da un investimento che indubbiamente è funzione della sua rischiosità. Nel caso in cui la controparte non sia effettivamente in grado di ripagare i propri debiti o di corrispondere gli interessi maturati le agenzie di rating eseguono un *downgrading* ossia la riduzione del rating attribuito all'emittente.

¹³ L'indice Herfindahl-Hirschman è un indice che misura la concentrazione di mercato di un settore. L'indice viene ottenuto sommando le quote di mercato delle imprese del settore al quadrato: $HHI = \sum_{i=0}^n Q_i^2$ dove Q indica la quota di mercato dell'impresa *i*. Più è basso il valore dell'indice più il mercato è competitivo. Il valore massimo indica la condizione di monopolio.

¹⁴ L'indice di concentrazione delle quattro imprese è uguale alla somma delle quote di mercato delle quattro aziende più grandi operanti nel mercato ossia la somma del fatturato delle quattro società diviso il fatturato complessivo del mercato.

Il rating può anche aiutare gli investitori a controllare periodicamente il valore di un investimento nel tempo. In particolare, le banche tendono a tutelarsi dal rischio di credito con garanzie e fondi svalutazione ma devono comunque tenere in conto il valore del rating. Spesso, il creditore realizza azioni per fronteggiare il rischio di credito a prescindere dal livello di rating. Ciò può avvenire in diversi modi, ad esempio tramite il *risk-based pricing* ossia l'aggiustamento del prezzo delle attività finanziarie in funzione del rischio, in maniera tale da aggiungere il valore economico della probabilità di default del debitore sul costo della transazione. Un altro metodo comune consiste nei *collaterals* ossia semplicemente nel chiedere delle garanzie, reali o personali, a tutela della transazione, o ancora rifacendosi a delle forme assicurative contro il fallimento delle proprie controparti¹⁵.

Un importante problema può però sorgere nei casi in cui il creditore non ritenga verosimile che la copertura del rischio sia adeguata da un punto di vista qualitativo o che semplicemente non sia sufficiente da un punto di vista quantitativo. In questi casi è possibile andare in contro a fenomeni di contrazione del credito concesso, effetti che se estesi a sufficienza possono persino mettere in crisi l'intero sistema economico¹⁶.

1.3.2 Tipologie di rating e modalità di emissione

Esistono diverse tipologie di rating: in particolare ci focalizzeremo brevemente sui rating relativi al credito, al debito e al *country ceiling*. Il primo tipo di rating, ossia quello di credito, è già stato precedentemente introdotto. Parliamo quindi giudizio di un'entità indipendente sull'affidabilità del debito emesso da una società. Tuttavia, possiamo anche fare riferimento a un rating di credito internazionale qualora vogliamo valutare i rischi di un investitore che possieda dei titoli espressi nella valuta di un altro Stato e voglia, ad esempio, trasferirli nella valuta della propria nazione. Molto importanti risultano essere invece i rating sul debito delle nazioni. Infatti, spesso sono proprio gli Stati a sopportare debiti in misura estremamente rilevante e in valore assoluto di importi superiori a quelli di molte aziende. La logica rimane assolutamente la stessa in quanto le agenzie di rating provvedono a classificare i titoli di debito in funzione della capacità stimata degli Stati di poter disporre di risorse sufficienti per ripagare i debiti sui titoli emessi. Infine, troviamo il *country ceiling rating* che si pone come obiettivo la valutazione della rischiosità di un investimento in uno Stato che possa verosimilmente attuare delle misure volte a bloccare uscite di capitale dai propri confini. Per emettere il giudizio sulla qualità del debito di una società l'agenzia di rating avvia una procedura che prevede l'analisi delle caratteristiche economico-finanziarie della società in questione. Viene quindi analizzato il bilancio in tutte le sue componenti e vengono analizzati parametri come la redditività dell'azienda, la sua capacità di produrre risorse e reddito, la remunerazione del capitale, i flussi di cassa, i rapporti fra i mezzi propri e il debito e così via.

¹⁵ Si vedano a proposito i Credit Default Swap (CDS)

¹⁶ Si pensi in questo caso al fenomeno del *Credit Crunch* che ha avuto effetti rilevanti sull'intera economia mondiale durante la crisi del 2008.

Relativamente al processo di emissione del rating vogliamo adesso brevemente delineare le fasi salienti. Inizialmente è necessaria una prima parte in cui gli analisti dell'agenzia di rating valutano e confrontano i parametri della società sotto esame con quelli delle altre società del settore in cui essa opera. Risulta quindi essere necessario anche l'analisi delle caratteristiche del settore di riferimento e l'andamento del mercato. Chiaramente, non esistono solamente fattori quantitativi ma anche quelli qualitativi. Infatti, gli analisti vogliono valutare fattori chiave quali l'affidabilità della società, le capacità del consiglio di amministrazione e dei livelli inferiori di management o ancora la credibilità dei progetti che si vogliono realizzare in futuro e degli obiettivi che la società intende perseguire.

È evidente che se i dati a disposizione non sono accurati o se vi sono degli errori nello sviluppo di questa prima e delicata fase verrà inficiata la validità delle dichiarazioni dell'agenzia di rating. Per questo motivo, spesso sono le stesse società sotto esame che si impegnano a fornire le informazioni e la documentazione utili sulla base delle richieste della stessa agenzia di rating.

In sintesi, parliamo dunque di un'analisi economico-finanziaria, che permette di esaminare il bilancio, la redditività, la remunerazione del capitale, i flussi di cassa e la capacità di produrre risorse e reddito. Esaminati questi dati si passerà all'analisi di settore.

Prima di procedere alla seconda fase è opportuno sottolineare che le agenzie di rating, per avviare una corretta analisi delle attività e dell'affidabilità di una società consultano la "centrale dei rischi" della Banca d'Italia. Questa è un database attraverso cui la Banca d'Italia fornisce delle informazioni particolarmente rilevanti circa lo stato di indebitamento di alcuni soggetti.

La seconda fase prevede invece che gli analisti si rechino direttamente in azienda per incontrare il management e valutare le azioni intraprese: dopo aver incontrato ed essersi confrontati con i vertici aziendali, gli analisti sottopongono una raccomandazione ad un comitato di credito, composto da specialisti di settore, dal direttore generale e dagli esperti del credito dell'agenzia.

Generalmente, il processo di attribuzione del rating dura circa 90 giorni ma le tempistiche possono chiaramente subire variazioni in funzione di particolari necessità, scadenze o eventi.

Stabilita la valutazione ufficiale da assegnare, il rating è pubblicato tramite un comunicato stampa sul sito dell'agenzia e i principali mezzi di informazione finanziaria a livello mondiale.

Inoltre, è opportuno ricordare che le agenzie di rating possono anche emettere più rating relativi a una stessa società in relazione alla durata temporale: è possibile quindi avere dei rating di breve, medio o lungo periodo. Il rating è talmente significativo che per diversi operatori finanziari come la Banca Centrale Europea o i fondi pensione sono previsti divieti di acquisto di titoli di credito con rating inferiori a una certa soglia.

Le scale di valori con cui le principali agenzie di rating comunicano i propri pareri sono generalmente costituite da scale alfabetiche. Ad esempio, nel caso di Standard & Poor's i giudizi variano da "AAA", ossia il migliore dei giudizi indicante quindi il massimo grado di solvibilità di una società, fino alla D caratterizzante

una società insolvente. Moody's invece predilige una scala di valori di poco differente e indica con una C il rating peggiore. Considerando che gli operatori tengono fortemente in considerazione i rating, è opportuno precisare che una valutazione che arrivi fino al limite minimo della tripla B viene percepito come un investimento abbastanza sicuro su cui grandi operatori, come ad esempio le grandi banche d'affari, possono indirizzare le proprie risorse. Tuttavia, scendere al di sotto di tale soglia significherebbe per un titolo accumulare una rischiosità particolarmente significativa e, per tale motivo, spesso le obbligazioni con un rating inferiore alla tripla B vengono indicate come speculative. Di seguito una tabella che riepiloga le classificazioni delle principali agenzie di rating.

Standard and Poor's	Moody's	Fitch Ratings	Livello di rischio	
AAA	Aaa	AAA	Estremamente basso	Investment Grade
AA+	Aa1	AA+	Rischio quasi nullo	
AA	Aa2	AA		
AA-	Aa3	AA-		
A+	A1	A+	Rischio molto basso, possibili rischi futuri	
A	A2	A		
A-	A3	A-		
BBB+	Baa1	BBB+	Medio-basso	
BBB	Baa2	BBB		
BBB-	Baa3	BBB-		
BB+	Ba1	BB+	Medio-alto	Speculative Grade
BB	Ba2	BB		
BB-	Ba3	BB-		
B+	B1	B+	Alto	
B	B2	B		
B-	B3	B-		
CCC+	Caa	CCC	Elevato	
CCC	Ca			
CCC-	C			
D	-	DDD DD D	Default	

Tabella 3 Classificazione delle principali agenzie di rating. Fonte: Borsa Italiana.

1.4 Le componenti del rischio di credito

Come precedentemente anticipato, la presenza del rischio di credito espone gli investitori al pericolo che essi non possano trarre i benefici economici desiderati. Il rischio di credito non fa soltanto riferimento alla possibilità di default ma anche alla possibilità di deterioramento del livello di solvibilità della controparte. Per valutare attentamente il rischio e poter prendere le dovute cautele, oltre all'adempimento delle normative correlate, i prestatori, come ad esempio le banche, devono tenere in considerazione le componenti che costituiscono il rischio di credito. Tra esse troviamo le seguenti:

1. La *Probability of default* (PD);
2. La *Loss given default* (LGD);
3. L'*Exposure at default* (EAD);
4. La *Maturity* (M).

Queste grandezze, opportunamente studiate, permettono di preventivare quali potranno essere le perdite collegate all'investimento che posso essere di due tipi: perdite attese o *Expected Loss* (EL) e perdite inattese o *Unexpected Loss* (UL).

Possiamo considerare le realizzazioni delle perdite come delle variabili aleatorie la cui distribuzione è riportata in figura 1. Il valore atteso delle perdite è rappresentato dalla media della distribuzione e costituisce il limite superiore delle perdite attese potendo definire così le perdite di entità superiore come perdite inattese.

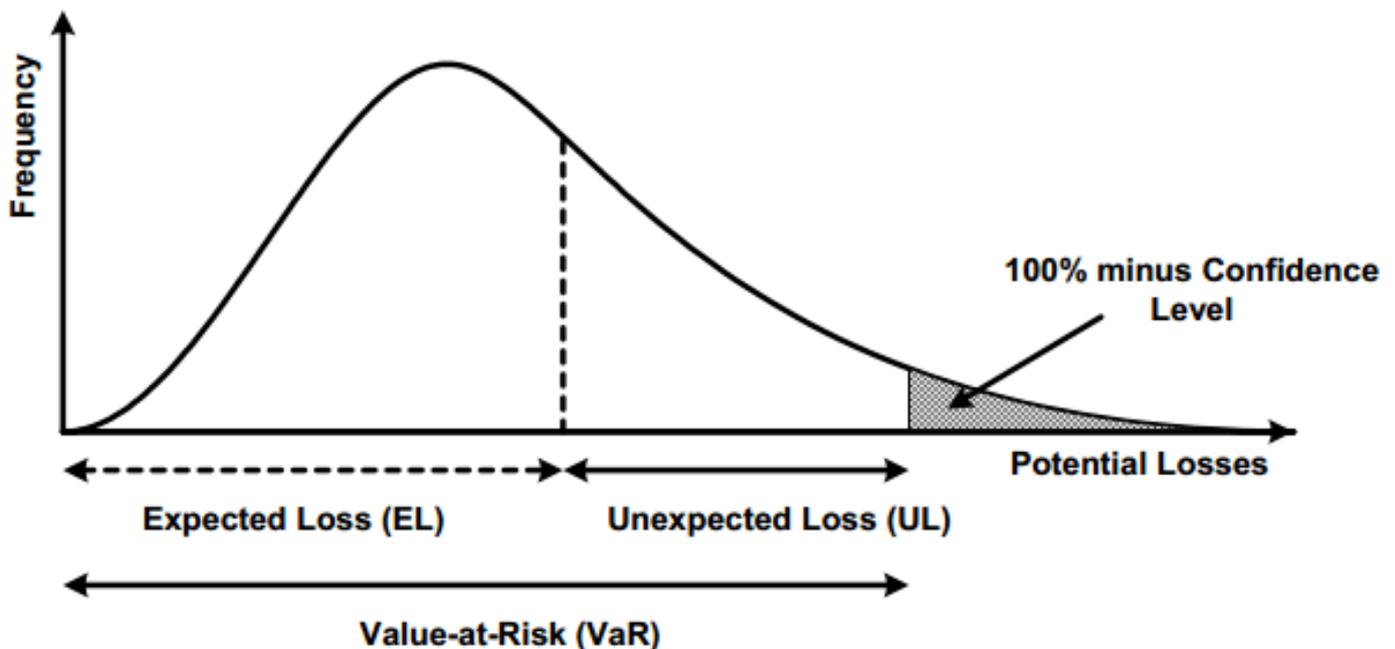


Figura 1 Distribuzione delle perdite e relativo valore atteso. Fonte: Gordon, D., "Aspects of Modeling Fraud Prevention of Online Financial Services", (2015).

Iniziamo adesso ad analizzare più nel dettaglio le varie componenti del rischio di credito. Il tasso di perdita atteso, ossia l'*Expected Loss Rate* (ELR), è il tasso che indica quale frazione del credito verrà persa tenendo in considerazione da un lato la perdita in caso di insolvenza e dall'altro la probabilità che questa si verifichi. Abbiamo allora che:

$$ELR^{17} = PD \cdot LGD$$

Su un orizzonte temporale limitato dobbiamo considerare anche la variabilità delle perdite. Essa quantifica il modo in cui le perdite tendono a discostarsi dalla media ossia dalla perdita attesa.

1.4.1 La Perdita attesa (EL)

Come anticipato all'inizio del paragrafo, la perdita attesa può essere definita come il valore medio della distribuzione delle perdite, considerata come una variabile aleatoria, che un soggetto suppone di subire su un singolo investimento o un portafoglio di investimenti¹⁸. Tuttavia, tale perdita, proprio in quanto "attesa", non può essere utilizzata per individuare l'effettivo grado di rischio. Infatti, se il calcolo avviene correttamente e risulta valido a consuntivo la perdita attesa non dovrebbe avere ripercussioni economiche poiché un soggetto economico come, ad esempio, una banca sarebbe in grado di caricare sul *pricing* la perdita attesa stimata *ex ante*.

La perdita attesa non risulta ridotta da una politica di diversificazione e trova copertura, nel caso in cui venga correttamente stimata *a priori*, nel conto economico. La perdita attesa può essere calcolata nel seguente modo:

$$EL = EAD \cdot ELR = EAD \cdot PD \cdot LGD$$

Possiamo quindi affermare che la perdita attesa sia funzione di tre elementi: *exposure at default*, *probability of default* e *loss given default*.

La perdita attesa è un concetto applicabile sia alla singola esposizione sia alla misurazione dell'intero portafoglio dei crediti. Per estendere l'applicazione a quest'ultimo caso sommare le singole esposizioni presenti nel portafoglio. In particolare:

$$EL = \sum_{i=1}^n EL_i = \sum_{i=1}^n PD_i \cdot LGD_i \cdot EAD_i$$

Possiamo riscrivere la stessa equazione in termini percentuali. Parliamo in questo caso di tasso di perdita atteso ossia *Expected loss rate* (ELR). Facciamo riferimento alla frazione del valore del portafoglio soggetta a rischio ottenuta dividendo la perdita attesa per il valore totale delle esposizioni in caso di default. In particolare, abbiamo che:

$$ELR = \frac{EL}{\sum_{i=1}^n EAD_i} = \sum_{i=1}^n w_i \cdot LGD_i \cdot PD_i$$

¹⁷ Questa formula fa riferimento al tasso di perdita attesa su una singola esposizione e non su un intero portafoglio. Inoltre, L'ELR può essere considerata come l'espressione percentuale dell'EAD.

¹⁸ Per la stesura dei paragrafi relativi alle componenti del rischio di credito si veda Nardon M., *Un'introduzione al rischio di credito*, Università Ca' Foscari, Venezia, n. 123/2004.

Nei paragrafi successivi saranno analizzate maggiormente nel dettaglio le componenti del rischio di credito.

1.4.1.1 La probabilità di default (PD)

La probabilità di default può essere descritta come la probabilità che i debitori appartenenti a una classe omogenea di soggetti con caratteristiche analoghe possano dichiarare default entro un dato orizzonte temporale, ad esempio di un anno. Per una maggiore affidabilità tali stime vengono successivamente confrontate con le frequenze relative di default che si sono effettivamente registrate. Tale confronto si pone come obiettivo quello di verificare *ex post* l'affidabilità dei modelli utilizzati per il calcolo di tale probabilità. Le banche utilizzano sistemi di rating, detti *borrower rating*, per classificare i debitori in determinate classi di merito che si differenziano per la rischiosità percepita. Come intuitivamente si può immaginare, i debitori collocati nelle classi di rating migliori si suppone possano dichiarare default con una probabilità più bassa rispetto ai soggetti collocati nelle classi di rating inferiori.

La stima della probabilità del default può avvenire con categorie di modelli diversi:

1. Modelli di scoring e rating: analisi discriminante, regressioni lineari, modelli logit e probit, reti neurali;
2. Modelli strutturali: il più famoso è il modello di Merton, proposto sulla base del modello di Black e Scholes;
3. Modelli in forma ridotta: questi tipi di modelli si differenziano dagli altri poiché non si pone come obiettivo quello di individuare e calcolare gli elementi che costituiscono l'evento default ma cerca di stimare la probabilità di default partendo dai dati di mercato. Pertanto, il verificarsi del default è considerato un evento esogeno al modello.

1.4.1.2 Il tasso di perdita in caso di insolvenza (LGD)

Possiamo definire il tasso di perdita in caso di insolvenza come la percentuale di credito che non sarà oggetto di recupero qualora si verifichi l'eventualità del default. Specularmente, il tasso di recupero (ossia il *recovery rate*, RR) indica invece la frazione recuperata ed è pertanto il complemento a uno del tasso di perdita.

Possiamo quindi scrivere che:

$$LGD = \frac{\textit{perdite}}{\textit{esposizione}} = 1 - \frac{\textit{credito recuperato}}{\textit{esposizione}} = 1 - RR$$

Un modello molto noto per il calcolo della LGD è il *LossCalc Model* sviluppato da Moody's nel 2005. Tramite questo modello si applica una regressione lineare multivariata che prende in considerazione alcuni fattori di rischio come, ad esempio, alcuni fattori macroeconomici ed elementi di rischio legati al settore di riferimento.

In questo modello possiamo osservare che la LGD viene calcolata come la somma del valore attuale dei *cash flow* nell'intervallo temporale di riferimento che dipende dalle pregresse esperienze di default riscontrate.

Come si evince dalla formula sotto riportata, indicando con t la data di default, con k la data di stipula del contratto e con s la data di scadenza, affermiamo che:

$$LGD_{k,t} = \frac{\sum_{s=k}^t \frac{IS_s + LP_s}{(1 + C_s)^{s-k+1}}}{B_k}$$

dove:

1. $LGD_{k,t}$ è il tasso di perdita in caso di default stimato al tempo t rispetto alla data di scadenza del contratto;
2. IS_s e LP_s rappresentano rispettivamente gli interessi e le perdite subite al tempo s , scontate al tasso di interesse composto C_s ;
3. B_k rappresenta gli importi ancora dovuti alla data k .

Notiamo quindi che questo modello tiene in considerazione tre date di riferimento: la data di stipula del contratto, la data di default e la data di fine del contratto.

Nel caso specifico degli istituti bancari è possibile che essi non utilizzino internamente un sistema di rating avanzato ma un sistema che permetta di calcolare la sola probabilità di default, allora il tasso di perdita in caso di insolvenza che la banca deve inserire come input è la *LGD regolamentare* che può avere due diverse misure:

1. $LGD = 45\%$ per i crediti di rango primario¹⁹;
2. $LGD = 75\%$ per i crediti subordinati.

Al contrario invece, quando viene utilizzato un approccio avanzato dei rating interni per la stima della LGD bisogna rispettare alcune imposizioni provenienti dalla regolamentazione in materia. Infatti, la stima della LGD deve essere:

1. Basata sulla definizione di default e perdita economica;
2. adatta a una fase recessiva del ciclo economico (*downturn LGD*) se questa risulti essere più prudente di quella basata sulla media di lungo periodo;
3. non negativa;
4. basata sui dati raccolti nel lungo periodo e su evidenze empiriche e quindi non solo su valutazioni discrezionali;
5. costantemente aggiornata non appena vi siano nuove e rilevanti informazioni disponibili e comunque con cadenza non inferiore all'anno;

¹⁹ Un credito è definito "di rango primario" quando la sua soddisfazione precede il soddisfacimento degli altri creditori dell'impresa. Si parla invece di "credito subordinato" quando la soddisfazione del credito dipende dal soddisfacimento degli altri creditori dell'impresa.

6. basata su una metodologia di stima che considera la storia passata dei recuperi. La LGD, detta in questo caso *workout LGD*, tiene conto dell'effetto finanziario del tempo e il tasso della LGD viene allora calcolato partendo da quattro variabili:

- la percentuale recuperata del credito nominale (RRN);
- i costi amministrativi sostenuti durante il processo di recupero (CA);
- il tempo di recupero;
- il costo finanziario connesso al tempo di recupero (i);

Possiamo a questo punto descrivere il tasso della *loss given default* come segue, alla luce delle variabili appena definite:

$$LGDR = 1 - \frac{RRN - CA}{(1 + i)^T}$$

Per concludere, la LGD dipende dai seguenti fattori:

1. caratteristiche del finanziamento: bisogna tenere in considerazione la presenza o meno di garanzie reali o finanziarie. Esse, infatti, impattano sull'importo potenziale del recupero. Quando dobbiamo stimare la LGD le garanzie sono il fattore che forse impatta maggiormente sul calcolo. Ovviamente, le garanzie legate ai prestiti devono essere rivalutate con una cadenza adeguata. Esse sono correlate negativamente alla LGD in quanto un miglioramento quantitativo o qualitativo delle garanzie permette di far decrescere il valore della LGD mentre costi e tempi di recupero contribuiscono all'innalzamento del suo valore;
2. caratteristiche del debitore: parliamo in particolar modo del paese e del settore di appartenenza;
3. caratteristiche del finanziatore: facciamo qui riferimento alle procedure seguite dal soggetto finanziatore. Esse impattano sulla durata del contenzioso e sull'importo del potenziale recupero;
4. fattori di natura esogena: alcuni fattori come il ciclo economico, i costi amministrativi interni ed esterni e i tassi di interesse possono avere un impatto significativo sull'importo effettivamente recuperato.

1.4.1.3 L'esposizione al momento dell'insolvenza (EAD)

Possiamo definire l'esposizione al momento dell'insolvenza come il valore totale cui un soggetto è esposto al momento dell'inadempienza della controparte. Soggetti particolarmente esposti come le banche calcolano il rischio tramite modelli basati su un approccio di rating interno (IRB). Spesso viene calcolato un valore di EAD per ogni prestito per poi utilizzare questi valori nel loro insieme per determinare il rischio di default complessivo.

Chiaramente l'EAD è un valore dinamico, che cambia ogni volta che una controparte onora il proprio debito.

1.4.2 La perdita inattesa (UL)

Dopo aver introdotto la perdita attesa (EL), ci vogliamo qui concentrare sulla perdita *inattesa* (UL). Generalmente, la banca, o qualsiasi altro datore di fondi, oltre al capitale accantonato per coprire le perdite attese, si salvaguarda anche per le perdite che in passato hanno ecceduto le perdite medie. Una misura efficace e spesso utilizzata per la misura dell'ordine di grandezza delle perdite attese è la deviazione standard della variabile di perdita \tilde{L} . Questa grandezza viene chiamata *unexpected loss* (UL) ed è definita come segue:

$$UL = \sqrt{V(\tilde{L})} = \sqrt{V[EAD \times SEV \times L]}$$

dove con $V(\tilde{L})$ si intende la varianza della perdita e SEV è la severity della posizione.

Dalla formula precedente può essere derivato un ulteriore risultato, se assumiamo che sia SEV che D siano due variabili aleatorie non correlate e che EAD sia costante nel tempo, allora l'equazione può essere riscritta nel seguente modo:

$$UL = EAD \sqrt{V[SEV] \times PD + LGD^2 \times PD(1 - PD)}$$

Le ipotesi su cui si basa questa assunzione sono poco verosimili, in quanto è assolutamente accettato che in media il tasso di recupero diminuisca al peggiorare delle condizioni economiche.

Quanto detto finora vale solo nel caso si consideri una sola posizione. Passiamo ora ad esaminare il caso di un portafoglio di crediti partendo dalla funzione di perdita di una singola posizione i -esima:

$$\tilde{L}_i = EAD_i \times SEV_i \times L_i$$

dove $L_i = \mathbf{1}_{D_i}$ e $p_{D_i} = PD_i$ con $i = 1, \dots, m$.

La perdita, considerando l'intero portafoglio, sarà data da:

$$\tilde{L}_P = \sum_{i=1}^m \tilde{L}_i = \sum_{i=1}^m EAD_i \times SEV_i \times L_i$$

La perdita attesa dell'intero portafoglio, dato il valore di \tilde{L}_P risulta essere:

$$EL_P = \sum_{i=1}^m EL_i = \sum_{i=1}^m EAD_i \times LGD_i \times PD_i$$

Infine, la perdita inattesa dell'intero portafoglio, date le ipotesi precedenti, sarà definita come:

$$UL_P = \sqrt{V[\tilde{L}_P]} = \sqrt{\sum_{i=1}^m \sum_{j=1}^m EAD_i \times EAD_j \times Cov[SEV_i \times L_i, SEV_j \times L_j]}$$

Se introduciamo un'ulteriore ipotesi, ovvero nel caso in cui tutte le SEV_i fossero costanti, ovvero valesse la seguente relazione $SEV_i = SEV$, per $i = 1, \dots, m$, allora la relazione diventa:

$$ULP^2 = \sum_{i,j=1}^m EAD_i \times EAD_j \times LGD_i \times LGD_j \times \sqrt{PD_i(1 - PD_i)PD_j(1 - PD_j)\rho_{ij}}$$

dove $\rho_{ij} = Corr[L_i, L_j] = Corr[\mathbf{1}_{D_i}, \mathbf{1}_{D_j}]$ indica la correlazione tra l'insolvenza della posizione i e quella della posizione j .

1.5 Perché approfondire il concetto di rischio?

In questo paragrafo verrà approfondita la tematica dell'importanza economica e, più in generale, sociale dell'attenzione nei confronti del rischio, di come sfruttarlo per cogliere le occasioni che può generare ma riducendolo entro livelli ritenuti gestibili. In passato, errate valutazioni legate al rischio hanno condotto a pesanti recessioni con ripercussioni a ogni livello sociale. La convenienza nell'investimento in *risk management* può essere affrontata da diversi punti di vista, partendo dalle istituzioni finanziarie, i suoi *stakeholders*, il suo management, la politica e il pubblico in generale. È evidente che ognuno di tali soggetti possa anche avere un punto di vista differente relativamente alla stessa questione. Vogliamo in questo paragrafo concentrarci su due punti di vista: quello della società in senso lato e quello degli *shareholders* delle aziende.

Partendo dalla prima prospettiva, è innegabile che la società moderna faccia massicciamente riferimento al sistema bancario e assicurativo, risultando quindi la sua stabilità di interesse assolutamente collettivo. I legislatori dei vari Paesi hanno cercato per questo motivo di fornire delle regolamentazioni adeguate che rispondessero a tale esigenza. Si pensi infatti agli accordi di Basilea o al *Solvency II Framework* tramite cui si è cercato di prevenire l'insolvenza di singole e importanti istituzioni, così da garantire in generale protezione anche ai loro clienti. Questo tipo di approccio viene talvolta definito come "approccio microprudenziale" in quanto ci si focalizza prevalentemente su istituzioni considerate come singoli. Tuttavia, ci si è sempre più concentrati su un altro obiettivo, ossia la riduzione del rischio sistemico, in particolare dopo la crisi del 2007-2009. Soprattutto da quel momento in poi la regolamentazione è passata maggiormente da un approccio microprudenziale a uno macroprudenziale. È quindi evidente che si voglia tutelare la stabilità dell'intero sistema economico ed è pertanto essenziale che si eviti che degli operatori importanti falliscano tuttavia, nel lungo periodo potrebbe anche essere meglio permetterlo. Infatti, ciò potrebbe fornire una lezione sull'importanza del *risk management* per le altre società e anche che non sempre la collettività può coprire gli errori di singole imprese anche quando questo comporti il loro fallimento. Resta tuttavia, il problema delle società passate alla storia come *too big to fail*, società talmente rilevanti per fatturato, numero di dipendenti e collegamenti con altre imprese che potrebbero trascinare con sé tutto il settore intaccando l'economia del Paese. Il rischio, in questo caso, è chiaramente legato al rischio sistemico. Il problema più rilevante potrebbe però risiedere nell'azzardo morale del *management* di queste società in quanto esso potrebbe essere tentato

di accettare un rischio eccessivamente elevato nella consapevolezza che, in caso di rischio di default, sarà comunque il governo a cercare una mediazione per evitare il coinvolgimento dell'intero sistema economico. Una seconda prospettiva da considerare è invece quella degli *shareholders*. È opinione diffusa che un adeguato *risk management* possa realisticamente contribuire all'aumento del valore di un'azienda e quindi dei suoi azionisti. Tra i benefici più rilevanti troviamo i seguenti (McNeil, Frey, & Embrechts, 2015). Il *risk management* può:

1. Ridurre l'impatto delle tasse. Il regime della tassazione generalmente può essere pensato come una funzione convessa dei profitti di una società. Quindi, riducendo la variabilità dei *cash flow* il *risk management* può contribuire al conseguimento di un utile dopo le imposte più alto;
2. avere effetti positivi sulle società che lo implementano adeguatamente in quanto esse ottengono un accesso al mercato dei capitali migliore rispetto agli investitori individuali;
3. può aumentare il valore della società in presenza di costi di bancarotta poiché la rende meno probabile. Tali costi consistono in costi diretti, come ad esempio i costi legali, e soprattutto indiretti. Questi possono comprendere i costi di liquidazione che possono essere particolarmente rilevanti in caso di presenza significativa di beni intangibili all'interno del bilancio della società, come costi di ricerca e sviluppo, avviamento elevato oppure *know-how*. Questo è probabilmente anche il motivo per cui si riscontra una correlazione tra l'investimento in *risk management* e quello in ricerca e sviluppo e altri intangibili. Inoltre, un aumento del rischio di bancarotta porta ulteriori problematiche per la società in quanto ha un effetto negativo sui dipendenti, sul *board* e sui partner commerciali. Questi soggetti saranno meno disposti a intraprendere relazioni di lungo periodo sapendo della presenza di un rischio di bancarotta rilevante;
4. può ridurre l'impatto dei costi esterni di finanziamento aumentando così il valore della società, dando il proprio contributo anche al raggiungimento dell'allocazione migliore degli investimenti. Infatti, è noto che in finanza aziendale i fondi esterni rappresentano un costo più elevato rispetto a quelli interni, costi attribuiti di solito al problema dell'asimmetria informativa tra il management e gli investitori azionari.

CAPITOLO 2. Modelli per la stima della *probability of default*

In questo capitolo affronteremo un tema particolarmente delicato. In primo luogo, verrà approfondito il tema della costruzione dei modelli in generale, le criticità che posso sopraggiungere e gli errori che spesso vengono commessi. Vedremo inoltre, quali sono le *best practice* e quali sono le caratteristiche essenziali che possono determinare il successo di un modello²⁰. Nei paragrafi successivi vedremo, invece, quali modelli sono stati effettivamente proposti per il calcolo della *probability of default* e quali sono i punti di forza e debolezza di ciascuno di essi.

2.1 Costruzione dei modelli

Essendo la realtà eccessivamente complessa non è possibile, o talvolta non è conveniente, cercare di studiare e analizzare i fenomeni economici e i dati a nostra disposizione così come essi si presentano. Dobbiamo quindi trovare un modo per semplificare la realtà senza però perdere di vista gli aspetti chiave dei fenomeni analizzati. Per questo motivo matematici, fisici ed economisti si impegnano nella costruzione di modelli, ossia semplificazioni della realtà, che riescano a cogliere gli aspetti salienti dei fenomeni studiati ma che allo stesso tempo non siano eccessivamente complicati in maniera tale che non risultino di difficile comprensione e applicazione. Una parte essenziale dei modelli è costituita dalle ipotesi di base circa il fenomeno da modellare. Questa è una diretta conseguenza del fatto che la realtà è troppo complessa per essere analizzata così per come la osserviamo ed è necessario introdurre delle semplificazioni. Risulta tuttavia essenziale che tali ipotesi siano profondamente comprese e soprattutto che vengano modificate quando non possano più essere considerate valide.

L'esigenza della costruzione dei modelli non viene solo dalle aziende, che sicuramente hanno necessità di analizzare una grandissima quantità di dati e devono prendere decisioni sulla base di tali informazioni con impatti significativi anche sulla loro stessa continuità aziendale, ma anche dalle istituzioni che spesso impongono determinati comportamenti di natura preventiva a seconda dei valori che emergono da specifici modelli. Per essere più chiari, si pensi alle autorità di vigilanza che impongono la costituzione e il mantenimento di un capitale regolamentare e, in particolare dopo la crisi del 2007, esse si sono mostrate sempre meno favorevoli all'utilizzo di modelli interni per la determinazione di questo tipo di capitale. Inoltre, si sono sempre più preoccupate di come le istituzioni finanziarie tendano a usare i propri modelli interni per finalità anche diverse dalla costituzione del capitale regolamentare. Questo comportamento da parte delle autorità di vigilanza fa emergere in maniera inequivocabile il problema essenziale che si trova alla base dell'utilizzo dei modelli ossia il *model risk* cui si è anche fatto cenno brevemente nel capitolo 1.

²⁰ Per la stesura di questo paragrafo si è fatto riferimento prevalentemente al libro "Risk Management e Istituzione Finanziarie" di John C. Hull, edizione italiana a cura di Emilio Barone (Hull, 2018).

Secondo l'SR 11-7 pubblicata nell'aprile 2011 dal *Board of Governors* della *Federal Reserve*, lettera di supervisione e regolamentazione contenenti le linee guida per un'efficace gestione del rischio di modello da parte delle banche, il rischio di modello è definibile come segue: *“la possibilità di conseguenze avverse derivanti da decisioni basate sulle stime e sui documenti prodotti da modelli non corretti o male utilizzati”*.

Ne deriva che le cause principali del rischio di modello sono due:

1. possono essere presenti degli errori fondamentali alla base del modello come, ad esempio, i dati utilizzati, le modalità di calcolo, le procedure numeriche o le ipotesi sottostanti;
2. il modello potrebbe essere utilizzato in maniera scorretta o inappropriata.

Un ulteriore problema è rappresentato dal fatto che quando più modelli sono collegati tra di loro o si basano sulle medesime ipotesi è possibile che il rischio di modello si amplifichi. Per tale motivo, è fondamentale monitorare le *performance* dei modelli per verificare quando uno di essi non produca più *output* attendibili. Infine, un'ultima precisazione. I modelli trattati in questo capitolo sono modelli relativi a fenomeni socioeconomici. Mentre i modelli usati, ad esempio, in fisica descrivono processi fisici e sono pertanto particolarmente accurati, i modelli usati in finanza devono descrivere delle variabili di mercato, di clienti, di fornitori, di debitori o altro. Facendo riferimento, quindi, a comportamenti umani la descrizione offerta dai modelli è solo una rappresentazione molto più approssimata della realtà rispetto ai modelli fisici. Inoltre, mentre i parametri usati nei modelli fisici spesso sono valori che non mutano nel tempo, quelli usati nei modelli economici vengono modificati con frequenza molto elevata.

2.1.1 Lo sviluppo e la convalida dei modelli

Parte fondamentale del modello è ovviamente il suo sviluppo. Quando una società costruisce un modello deve opportunamente documentarlo. Sebbene questa attività non sia spesso apprezzata dagli sviluppatori, risulta essere di particolare importanza per i seguenti motivi:

1. se il *model developer* termina la sua attività lavorativa nella società senza lasciare un'adeguata documentazione è difficile che altri soggetti vogliano assumersi la responsabilità di continuare con il lavoro e si corre il rischio di perdere il lavoro svolto;
2. molto spesso sono le stesse autorità di vigilanza a valutare i modelli e le procedure utilizzate dalle istituzioni finanziarie e la fase di controllo passa proprio dalla documentazione;
3. quando il modello risulta essere ben documentato allora il lavoro di convalida e revisione periodica risulta avvantaggiato.

Nonostante l'attività di documentazione spesso non sia considerata la più appagante da un punto di vista professionale risulta comunque di notevole importanza sia per spiegare nel dettaglio le ipotesi e le procedure numeriche utilizzate sia perché costringe i *model developer* a ripensare ai loro modelli, riesaminando il lavoro

svolto e inducendoli, talvolta, a effettuare delle modifiche. La verifica del modello assume ovviamente un rilievo notevole e deve coprirne ogni aspetto, dal codice al comportamento su vari input. In particolare, essi dovrebbero essere via via sempre più estremi in maniera tale da verificare il comportamento dei modelli nei casi limite. Può essere opportuno anche confrontare i risultati ottenuti con un modello con quelli che si sarebbero ottenuti con un modello evidentemente più complesso ma troppo lento dal punto di vista computazionale per poi essere effettivamente utilizzato.

Inoltre, il modo in cui le *performance* di un modello debbano essere verificate dipende molto dalla sua tipologia. Ad esempio, un modello basato sulla base di serie storiche potrebbe essere controllato con dati fuori dal campione. Se prendiamo, quindi, in considerazione un modello che ci aiuti a prendere delle decisioni circa delle concessioni di prestiti e questo modello è sviluppato sulla base di un dato campione risulta chiaro che non possiamo testare questo specifico modello facendo riferimento a quello stesso campione di dati.

Un aspetto che non deve essere sottovalutato nella gestione e nella convalida dei modelli è quello del *back-testing*. Ci chiediamo quindi come un modello avrebbe potenzialmente performato in una situazione passata. Il vantaggio di questo tipo di analisi è che i dati, essendo tutti consuntivi, sono reperibili spesso senza particolari difficoltà. Se utilizziamo un modello che, basato su specifiche ipotesi, riesca a prezzare delle opzioni ci dobbiamo allora chiedere se in passato questo modello avrebbe dato dei prezzi sufficientemente vicini a quelli che poi sono risultati a consuntivo e se ci avrebbe anche adeguatamente remunerato per il rischio sostenuto. Un errore che potrebbe essere sottovalutato è però quello di non considerare la modifica nei comportamenti degli operatori nel caso in cui essi siano a conoscenza del modello e ne intuiscono il funzionamento. Se un modello viene utilizzato per smascherare delle operazioni illecite con fini di riciclaggio di denaro è possibile che esso funzioni fino a quando i soggetti interessati non comprendano le ipotesi alla base del modello in questione, cambiando il loro comportamento per poterlo ingannare e restare così impuniti. Se, spesso tramite metodi "*trial and error*", i *model user* comprendono le basi del funzionamento di un modello, quest'ultimo rischia di diventare inutile o, ancor peggio, dannoso come strumento di supporto nel prendere delle decisioni.

Poiché l'importanza dei modelli nel tempo è diventata sempre più significativa sono nati nelle aziende dei comitati detti *model validation groups* il cui obiettivo è appunto la verifica della corrispondenza del funzionamento dei modelli con le aspettative della società a riguardo. Per garantire il corretto funzionamento di tali gruppi è necessario che i soggetti coinvolti siano del tutto indipendenti dal *model developer* e dal *model user*. Se l'indipendenza non viene garantita l'intera funzione di *model risk management* rischia di perdere la sua efficacia.

Talvolta, può essere una buona idea provare i modelli su scala ridotta così da avere una prima e informale idea sul loro funzionamento prima di procedere all'effettiva convalida iniziale. È chiaro, in ogni caso, che la

profondità e complessità del processo di convalida dipendono fortemente dall'entità dei rischi che dal modello possono derivare per la società che lo adotta.

La convalida non è una procedura da applicare una volta, subito dopo la stesura del modello e prima dell'applicazione, e poi da considerarsi conclusa ma, esattamente al contrario, è una prassi da applicare costantemente. È necessario, infatti, che ad ogni modifica delle condizioni di mercato e delle prassi operative la convalida venga applicata nuovamente per verificare l'eventuale impatto sulle *performance* del modello. Secondo la normativa SR 11-7 i modelli dovrebbero essere riesaminati dal *model validation group* non meno di una volta l'anno con procedura di convalida comprendente tre elementi fondamentali:

1. fase di valutazione della solidità del modello dal punto di vista concettuale, partendo dalle fasi iniziali dello sviluppo;
2. monitoraggio ininterrotto che includa test dei processi e confronto con *benchmark*;
3. analisi dei risultati comprendente anche il *back-testing*.

In particolare, relativamente all'ultimo punto, è lo stesso *model developer* che dovrebbe effettuare tale tipo di analisi. Vanno, inoltre, attentamente considerate le differenze tra i risultati attesi e quelli ottenuti a consuntivo.

Il motivo per cui tutte le modifiche apportate al modello debbano essere opportunamente documentate e verificate diviene più chiaro nel caso in cui consideriamo che il *model user* e il *model developer* siano la stessa persona. Tale soggetto potrebbe essere tentato di modificare il modello per migliorarne le *performance*. Tuttavia, per quanto questa possa sembrare una buona idea nell'immediato, si rischierebbe di arrivare a una situazione in cui il modello altro non sia che una seria indefinita di modifiche, senza più alcuna base teorica valida e rendendo il processo di ricostruzione dei passaggi, dal modello originale a quello attuale, un compito arduo se non supportato da opportuna documentazione. Inoltre, si corre anche il rischio che il *model developer* non sia imparziale in quanto preferisca un particolare tipo di modello oppure potrebbe essere tentato di svilupparne uno particolarmente complesso per fare emergere le sue doti professionali. Il compito di verificare che non sia così è affidato proprio alla procedura di convalida.

Un altro fattore di rilievo è la stabilità. Infatti, se una variazione di modesta entità di un dato *input* risulta in grado di determinare una grande variazione degli *output* il modello è da considerarsi instabile e aumenta il rischio del suo utilizzo. Nella procedura di convalida sarà allora sempre opportuno includere l'attività di *stress testing* per verificare il comportamento del modello quando vengono considerati *input* estremi.

Occorre ricordare che non tutti i modelli provengono dall'interno dell'organizzazione. Alcuni vengono direttamente acquistati dai *model vendors* ma il motivo di tale strategia può essere differente. Infatti, è possibile che il modello esterno sia acquistato per essere preso in considerazione come *benchmark* per un altro modello interno. In questi casi allora il modello esterno può essere parte integrante della procedura di convalida. Se invece il modello è acquistato per il semplice fatto che questa sia l'alternativa economicamente

più conveniente rispetto al suo sviluppo allora il modello esterno deve essere sottoposto alla stessa procedura di convalida ma essa può risultare particolarmente complessa poiché le competenze che hanno portato al suo sviluppo sono esterne alla società. Per tale motivo il *model vendor* deve necessariamente fornire tutta la documentazione necessaria affinché possano essere compresi profondamente tre aspetti essenziali:

1. il modello;
2. le verifiche empiriche cui è stato sottoposto;
3. i suoi limiti.

Inoltre, la società con molta probabilità dovrà delegare al fornitore anche le attività di monitoraggio, di analisi dei risultati e della loro trasmissione ai clienti.

Poiché il *model validation group* deve assumersi la responsabilità dei modelli anche quando questi sono sviluppati esternamente e poiché spesso nei modelli esterni non si dispone dell'accesso al codice alla base del modello (per quanto sia possibile comunque svilupparne un'analisi modificando i dati di *input* e verificando i risultati) è opportuno predisporre un "piano di emergenza", detto *contingency plan*, nel caso in cui il *model vendor* vada in fallimento o decida di non fornire più assistenza su uno specifico modello. Una *chance* è quella di chiedere che il codice del programma e tutto il resto della documentazione utile vengano posti come garanzia. Mettendo tali strumenti a disposizione del cliente in caso di fallimento o mancata continuazione dell'assistenza, la società cliente può acquisire degli strumenti che consentono di limitare il rischio legato al modello.

Tutte le procedure di convalida non devono essere sottovalutate anche nei casi in cui la strategia di investimento sembri semplice o quando c'è poco disaccordo sul mercato circa il prezzo di alcuni prodotti²¹.

2.1.2 Modelli per contratti attivamente negoziati e per contratti fuori standard

Nei casi in cui i contratti siano attivamente negoziati sul mercato non c'è effettivamente bisogno di un modello per la stima del loro prezzo. È direttamente il mercato stesso a fornire tale informazione. Tuttavia, spesso ci si imbatte nella situazione in cui dobbiamo valutare un contratto simile, ma non identico, a uno

²¹ Il caso Kidder Peabody, riportato da J. C. Hull in "Risk Management e Istituzioni Finanziarie" (2018), è significativo in questo senso. Un loro *trader* aveva pensato a una semplice strategia di investimento. Le banche d'investimento suddividono i titoli di Stato con cedole in *zero-coupon bonds*, chiamati *strips*, ciascuno dei quali dà diritto a uno dei pagamenti previsti dal titolo con cedole. La strategia consisteva nel comprare *strips* nel mercato *spot* e venderli nel mercato *forward*. Rinnovando le operazioni in scadenza riusciva a evitare che gli interessi venissero addebitati alle sue posizioni. Poiché i prezzi di un titolo che non offre redditi nel mercato *forward* sono sempre più alti rispetto allo *spot* il sistema informatico registrava costantemente profitti sulle operazioni. In realtà, la differenza di prezzo consiste nell'interesse pagato per finanziare l'acquisto *spot* dello *strip*. Inizialmente pensando di aver conseguito un profitto di \$100 milioni (e dopo un consistente *bonus* al *trader*), la società si rese conto di aver accumulato perdite per circa \$350 milioni.

standard. In questi casi risulta utile l'utilizzo di un modello per poter essere ragionevolmente convinti di ottenere una valutazione coerente con i prezzi di mercato forniti per contratti simili²².

Quando parliamo dei contratti attivamente negoziati è opportuno tenere a mente che il rischio più rilevante potrebbe derivare dalle coperture. Possiamo fare una prima distinzione tra le coperture "all'interno del modello", *within-model hedging*, e coperture "all'esterno del modello", *outside-model hedging*. Le prime hanno la caratteristica di tenere in conto il rischio determinato dalle variabili che sono considerate aleatorie dal modello. Le seconde, invece, tengono in conto il rischio determinato dalle variabili che il modello considera costanti o deterministiche. Nella prassi, i *traders* utilizzano frequentemente entrambi i tipi di coperture. In ogni caso, spesso le istituzioni finanziarie decidono di valutare l'efficacia delle operazioni di copertura scomponendo le variazioni giornaliere del valore del portafoglio. In particolare, le componenti sono date dalle variazioni derivanti da:

1. rischi che non sono stati coperti;
2. imperfezione del modello usato;
3. nuovi contratti negoziati durante il giorno.

Questa scomposizione, detta *profit and loss decomposition*, permette di analizzare più nel dettaglio le variazioni nel caso in cui queste siano troppo rilevanti e, soprattutto, permette di individuare le aree che richiedono maggiore attenzione. Occorre tenere in considerazione che spesso gli errori tendono ad avere segno opposto e quindi a compensarsi tra di loro.

Per quanto concerne invece i contratti fuori standard ossia contratti simili, ma non identici, a quelli attivamente negoziati sui mercati, il processo di valutazione è il seguente:

1. osservazione dei prezzi di mercato di contratti simili attivamente negoziati sui mercati;
2. stima dei parametri del modello e interpolazioni;
3. valutazione dei contratti fuori standard.

Mentre il rischio di modello è relativamente basso quando dobbiamo stimare i prezzi teorici di contratti attivamente negoziati, così non è quando parliamo di contratti fuori standard. Il modello potrebbe quindi avere un impatto rilevante sui prezzi stimati e sulle modalità di copertura. Un aspetto interessante da sottolineare è il fatto che siamo interessati alla "valutazione relativa", vale a dire che il nostro scopo è quello di valutare i contratti fuori standard coerentemente con i prezzi dei contratti simili attivamente negoziati.

²² Se dobbiamo valutare un'opzione che abbia il prezzo di esercizio o la vita residua diversi dalle opzioni negoziate sul mercato possiamo usare un modello in cui i parametri sono stimati sulla base dei prezzi delle opzioni attivamente negoziate sul mercato.

Quando si fanno tali tipi di stime, i soggetti coinvolti dovrebbero evitare l'utilizzo di un singolo modello poiché in tale modo si potranno stimare più prezzi secondo modalità diverse ed essere al contempo più coscienti sul rischio legato al modello utilizzato²³.

Al posto di utilizzare un ridotto numero di modelli, un'alternativa potrebbe essere quella di effettuare un'analisi che permetta di calcolare il peggiore e il migliore prezzo possibile. In pratica, si vuole rispondere alla seguente domanda: "quali sono le variazioni possibili dei prezzi teorici dei contratti fuori standard forniti da modelli calibrati sui prezzi di mercato dei contratti standard?". Uno dei metodi utilizzati per risolvere tale quesito è quello delle simulazioni Monte Carlo ponderate. Si procede con delle simulazioni Monte Carlo assegnando dei pesi alle traiettorie simulate. Tali pesi, chiaramente, sono scelti in maniera tale da restituire i valori corretti per i prezzi dei contratti standard.

In conclusione, vogliamo indicare gli elementi che potrebbero spiegare, in generale, il successo di un modello e gli errori tipici nello sviluppo degli stessi. In particolare, prendendo in considerazione il caso del modello Black-Scholes-Merton, il successo potrebbe essere spiegato come segue:

1. fornisce valutazioni coerenti con i prezzi di mercato;
2. è un efficace strumento di comunicazione;
3. è relativamente semplice da comprendere e applicare.

I modelli che, in generale, godono di queste proprietà possono avere maggiori *chance* di diffusione rispetto ad altri eccessivamente complessi o meno efficaci.

Relativamente agli errori di sviluppo, dobbiamo tenere a mente la finalità del modello ossia quello di riconoscere quali sono gli aspetti chiave di un fenomeno potendolo così semplificare senza perdere informazioni essenziali. Gli errori più frequenti potrebbero essere sostanzialmente due:

1. rendere il modello troppo complesso da comprendere o da applicare; potrebbe essere, ad esempio, il caso dell'utilizzo di un numero eccessivo di parametri (*overparameterization*);
2. ricercare un'aderenza eccessiva tra i prezzi teorici e quelli effettivi. In quest'ultimo caso si va incontro al fenomeno noto come *overfitting*. Anche quest'ultimo caso potrebbe essere una conseguenza del numero eccessivo di parametri. L'*overfitting* assume particolare rilievo in tema di "apprendimento automatico" quando un algoritmo viene "allenato" usando una banca dati di esempi, detti *training set*, magari utilizzando casi in cui i valori di *output* sono già noti. Si vuole che il modello possa imparare a prevedere anche *output* di casi non presenti nel *database* ma se quest'ultimo non è ampio e il tempo di apprendimento è lungo il modello potrebbe adattarsi a caratteristiche che sono specifiche

²³ Se una società ha utilizzato più modelli che hanno fornito prezzi diversi dovrebbe porsi il problema di quale prezzo applicare e di come riconoscere i profitti in bilancio. Anche se il modello che ha fornito il costo più basso fosse riconosciuto come il più affidabile bisognerebbe essere prudenti nella rilevazione dei profitti tenendo in debita considerazione anche i prezzi forniti dagli altri modelli.

del singolo *training set* fornendo risultati eccellenti nella previsione dei relativi output ma non essendo in grado di generalizzare altrettanto bene i risultati al di fuori di esso.

2.2 I modelli *Rating System*

Ci concentreremo da questo paragrafo in poi sui principali modelli utilizzati per la stima della *probability of default*. I modelli di *rating system* sono dei modelli utilizzati dalle agenzie di rating con lo scopo di analizzare la qualità del credito di specifiche aziende basandosi su dati sia quantitativi sia qualitativi. Il giudizio espresso dalle agenzie di *rating*, come già sottolineato nel capitolo 1, non è un giudizio che viene modificato a ogni minimo cambio di condizione dell'azienda esaminata. L'obiettivo è più quello di riconoscere il *trend* generale. Tuttavia, è possibile che delle banche commissionino un giudizio esterno e indipendente sul credito relativo a una specifica azienda ottenendo un parere relativo alla stima della LGD e della PD²⁴.

Oltre al ben noto meccanismo dei *rating* che classificano le posizioni debitorie tramite lettere che ne indicano in maniera intuitiva e immediata la qualità, un'altra opzione è quella indicata da Blumh, Overbeck e Wagner ossia la *quick and dirty solution*. Essi partono dalle frequenze storiche delle insolvenze, ne calcolano media e varianza tramite regressione lineare e ottengono una relazione tra le frequenze di default e la variazione del merito creditizio. Tale relazione si presenta di tipo esponenziale. Tale funzione viene usata per attribuire un valore di PD a ogni classe. I metodi che utilizzano la regressione hanno il pregio di normalizzare i dati non tenendo eccessivamente in conto i dati con poca rilevanza sul giudizio finale.

2.2.1 *Credit Rating Migration*

Nell'approccio detto *credit migration* ogni azienda è assegnata a una categoria di *rating* in un determinato tempo t . Le probabilità di spostarsi da una categoria di *rating* a un'altra in un dato orizzonte temporale, generalmente di un anno, possono essere calcolate e sono spesso rappresentate tramite una matrice. Un esempio può essere fornito dalla seguente matrice di probabilità calcolata da Moody's.

In questo caso, considerando le probabilità fornite dalla tabella, la probabilità di default a un anno di una società con rating A è da considerarsi lo 0,06% mentre se avessimo considerato il rating Caa allora tale probabilità sarebbe salita al 13,34%. Nella tabella viene considerata una correzione: infatti è possibile che il valore del *rating* venga ritirato (*rating withdrawal* indicato con WS). Un metodo semplice per aggiustare tutte le percentuali legate ai *rating* per tenere conto di questo aspetto è dividere le prime nove probabilità di ciascuna riga per uno meno l'ultima probabilità delle righe stesse (ossia uno meno il valore indicato nell'ultima colonna, quella denominata WR). Ciò implicitamente assume che il ritiro del *rating* non ci fornisce informazioni sulla probabilità di *upgrade*, *downgrade* o *default* di un determinato soggetto.

²⁴ Per un approfondimento sul tema del *rating* e il suo funzionamento si rimanda al capitolo 1.

Rating alla fine dell'anno (%)										
Rating iniziale	Aaa	Aa	A	Baa	Ba	B	Caa	Ca-C	Default	WR
Aaa	87,20	8,20	0,63	0,00	0,03	0,00	0,00	0,00	0,00	3,93
Aa	0,91	84,57	8,43	0,49	0,60	0,02	0,01	0,00	0,02	5,48
A	0,06	2,48	86,07	5,47	0,57	0,11	0,03	0,00	0,06	5,13
Baa	0,039	0,17	4,11	84,84	4,05	7,55	1,63	0,02	0,17	5,65
Ba	0,01	0,05	0,35	5,52	75,75	7,22	0,58	0,07	1,06	9,39
B	0,01	0,03	0,11	0,32	4,58	73,53	5,81	0,59	3,85	11,16
Caa	0,01	0,02	0,02	0,12	0,38	8,70	61,71	3,72	13,34	12,00
Ca-C	0,00	0,00	0,00	0,00	0,40	2,03	9,38	35,46	37,93	14,80

Tabella 4 Probabilità di migrazione da un rating a un altro entro un anno. Fonte: Moody's.

Le società di rating generalmente forniscono anche matrici con le probabilità di default cumulate relativi a periodi di tempo di lunghi di un anno. La matrice seguente può fornirne un esempio.

Scadenza in anni							
Rating iniziale	1	2	3	4	5	10	15
Aaa	0,00	0,01	0,01	0,04	0,11	0,50	0,93
Aa	0,02	0,07	0,14	0,26	0,38	0,92	1,75
A	0,06	0,20	0,41	0,63	0,87	2,48	4,26
Baa	0,180	0,50	0,89	1,37	1,88	4,70	8,62
Ba	1,11	3,08	5,42	7,93	10,18	19,70	29,17
B	4,05	9,60	15,22	20,13	24,61	41,94	52,22
Caa-C	16,45	27,87	36,91	44,13	50,37	69,48	79,18

Tabella 5 Media cumulata dei tassi di default (%). Fonte: Moody's.

Un aspetto interessante da notare è che esistono prove empiriche che le probabilità di default tendono a variare in sincronia con lo stato del ciclo economico. In altri termini, come anche l'intuito suggerisce, è più probabile che vi siano dei *downgrade* durante dei periodi di recessione e, viceversa, degli *upgrade* durante le fasi di espansione economica. I tassi di transizione che sono stimati da società come S&P e Moody's, come visto nella tabella 4 ad esempio, sono medie storiche calcolate su lunghi intervalli temporali che includono diversi cicli economici. I dati della tabella 4 sono calcolati prendendo in considerazione il periodo 1970-2012. In certi casi potremmo, quindi, essere interessati a una stima della probabilità di default legata a un tempo t preciso e non a una media storica. Possiamo ottenere tale valore procedendo con degli aggiustamenti ai tassi calcolati su medie storiche. Un esempio potrebbe essere dato dal modello *public-firm EDF (expected default frequency)* che utilizza i prezzi dell'*equity* come ulteriore fonte di informazione.

Quando parliamo delle transizioni dei *rating* possiamo anche fare riferimento alle catene di Markov. Consideriamo (R_t) un processo stocastico a tempo discreto definito in $t = 0, 1, \dots$ che può assumere valori nell'insieme $S = \{0, 1, \dots, n\}$ dove S , detto spazio degli stati, rappresenta la qualità del *rating* ossia più è alto il suo valore e migliore sarà il *rating*, indicando con 0 il valore legato al default. (R_t) rappresenta allora

l'evoluzione del rating di un soggetto nel tempo. Assumiamo adesso che (R_t) sia una catena di Markov. Questo vuol dire che la probabilità condizionata di transizione soddisfa la seguente condizione:

$$P(R_t = k \mid R_0 = r_0, R_1 = r_1, \dots, R_{t-1} = j) = P(R_t = k \mid R_{t-1} = j)$$

per ogni $t \geq 1$ e $j, r_0, r_1, \dots, r_{t-2}, k \in S$. In altri termini, la probabilità condizionata relativa alla transizione del *rating* di un determinato soggetto dipende solo dall'ultimo valore di *rating* $R_{t-1} = j$ e non dalle condizioni precedenti o di come tale soggetto sia arrivato allo stato j al tempo $t-1$.

Le assunzioni di base di Markov per la migrazione dei *rating* sono state tuttavia criticate. È ragionevole, infatti, pensare che vi siano problematiche legate a due aspetti: il *momentum* e la *stickiness*. Il primo fa riferimento alla probabilità che un soggetto che abbia recentemente subito un *downgrade* sia più esposto a un ulteriore abbassamento del *rating* rispetto a chi si trovi nella sua attuale situazione già da tempo. La seconda richiama invece il tipico comportamento delle agenzie di *rating* che tendono a evitare un *downgrade* fino a quando le circostanze non siano soprafacenti. Nonostante queste critiche, le assunzioni di base delle catene di Markov sono ugualmente molto utilizzate poiché permettono l'utilizzo di modelli relativamente semplici con una solida teoria alle spalle.

Infine, possiamo sottolineare che ci sono degli svantaggi nel modellare le transizioni dei *rating* secondo catene di Markov che considerano intervalli di tempo discreti. Nella realtà, i cambi di *rating* possono verificarsi su intervalli temporali irregolari. Generalmente, approssimiamo questi dati in intervalli di tempo regolari come, ad esempio, anni, trimestri o mesi ma facendo ciò accettiamo di perdere delle informazioni. I modelli discreti ignorano gli eventuali cambi di *rating* intermedi tra il tempo t e il tempo $t+1$. Se un soggetto passa dalla classe A alla BBB e poi da quest'ultima alla BB nell'arco di tempo $[t, t+1]$, verrà registrata solo la variazione da A verso BB, con la conseguente perdita delle informazioni intermedie. Per evitare che ciò si verifichi è comunque possibile modellare le transizioni dei *rating* trattando la variabile t continua. In questo caso le probabilità di transizione non sono direttamente modellate ma sono ottenute tramite i tassi di transizione. Immaginiamo che in un piccolo intervallo temporale δt la probabilità di transizione dal *rating* j al k sia data approssimativamente da $\lambda_{jk} \delta t$ per una costante $\lambda_{jk} > 0$ che rappresenta il tasso di transizione del *rating* tra j e k . Invece, la probabilità di rimanere del *rating* j è data da $1 - \sum_{k \neq j} \lambda_{jk} \delta t$. Se definiamo una matrice Λ che abbia come elementi al di fuori della diagonale principale λ_{jk} e sulla diagonale principale $-\sum_{k \neq j} \lambda_{jk}$, possiamo sintetizzare le probabilità di transizione per piccoli intervalli temporali δt nella matrice $(I_{n+1} + \Lambda \delta t)$, dove I_{n+1} è la matrice identità $(n+1) \times (n+1)$. Adesso consideriamo le transizioni nel periodo di tempo $[0, t]$ e indichiamo la corrispondente matrice delle probabilità di transazione con $P(t)$. Se dividiamo il periodo di tempo in N intervalli di dimensione $\delta t = t/N$, per N alti possiamo approssimare la matrice delle probabilità di transizione nel seguente modo:

$$P(t) \approx \left(I_{n+1} + \frac{\Lambda t}{N} \right)^N$$

che, per $N \rightarrow \infty$, converge nella cosiddetta matrice esponenziale di Λt :

$$P(t) = e^{\Lambda t}$$

Questa formula ci fornisce un modo per calcolare le probabilità di transizione per ogni orizzonte temporale t .

2.3 I modelli *credit scoring*

Gli analisti hanno da sempre cercato di costruire dei modelli per cercare di prevedere la crisi di impresa. Si cerca quindi di capire se vi siano degli elementi che, opportunamente studiati, possano fornire degli spunti utili per delle previsioni future. Una delle prime analisi svolte fu quella di Beaver nel 1966. Egli provò a considerare sei indicatori di bilancio per verificarne l'efficacia. Gli indicatori scelti furono i seguenti:

1. *Cash Flow Ratios*;
2. *Net Income Ratios*;
3. *Debt to Total Asset Ratios*;
4. *Liquid Asset to Total Asset Ratios*;
5. *Liquid Asset to Current Debt Ratios*;
6. *Turnover Ratios*.

Il modello di Beaver era basato su un'analisi univariata degli indicatori. Egli giunse alla conclusione che tra i vari indicatori di bilancio ci fosse una forte difformità nella capacità di previsione così come nella bontà di adattamento. Ad ogni modo, la precisione delle previsioni circa l'insolvenza delle società analizzate non fu particolarmente elevata, soprattutto utilizzando i *liquid asset to total asset ratios* e i *liquid asset to current debt ratios*. La precisione non era dunque sufficientemente accettabile e, nel 1968, Altman propose il suo modello *Z-score*.

2.3.1 Il modello di Altman

Il modello di Altman può essere utilizzato per la classificazione di un'osservazione in uno o più gruppi in relazione alle caratteristiche della prima osservazione stessa. Questo metodo è spesso utilizzato per fare delle previsioni in cui la variabile dipendente ha natura qualitativa: in questo caso, infatti, parliamo di società fallite o non fallite²⁵.

Il primo passaggio è quello di ottenere una combinazione lineare delle caratteristiche che permettono di discriminare i gruppi. Se vi sono delle caratteristiche, come ad esempio dei rapporti finanziari, che possono

²⁵ Per maggiori approfondimenti si veda Bottani, P., Cipriani, L., Serao F., "Il modello di Z-Score applicato alle PMI", amministrazione e finanza n. 1/2004.

essere calcolati per tutte le società facenti parte il campione dell'analisi si possono stimare dei coefficienti discriminanti. La funzione discriminante può essere espressa come segue:

$$Z = v_1x_1 + \dots + v_nx_n$$

dove v_1, \dots, v_n sono i coefficienti discriminanti e x_1, \dots, x_n sono le variabili indipendenti. Possiamo, tramite la funzione appena espressa, collegare i valori delle singole variabili in un risultato singolo che funge da discriminante e ci permette di classificare l'oggetto analizzato. I coefficienti discriminanti v_j con $j = 1, \dots, n$ sono preventivamente calcolati mentre le variabili indipendenti x_j con $j = 1, \dots, n$ sono dei valori attuali.

Nel modello di Altman, a differenza di quello di Beaver, l'analisi è di tipo multivariato. Basandosi su indicatori di bilancio il primo problema consisteva proprio nella scelta di tali indicatori. Una prima idea era stata quella di considerare gli indici che coprissero l'area finanziaria e strategica delle imprese. In un secondo momento, tuttavia, si è deciso di considerare solo le variabili ritenute rilevanti ai fini del fallimento, verificato ovviamente *ex post*, disponibili nei bilanci aziendali. Le variabili individuate erano ventidue e furono raggruppate in diverse categorie: liquidità, profittabilità, solidità e altre attività aziendali. Nella prima versione del modello, Altman prese in considerazione 66 imprese quotate di cui la metà sane e la parte rimanente fallite. Per l'individuazione effettiva delle variabili poi concretamente utilizzate nel modello, Altman seguì il seguente approccio:

1. osservazione della significatività statistica delle variabili indipendenti;
2. valutazione della correlazione tra le variabili ritenute maggiormente rilevanti;
3. verifica dell'accuratezza previsiva delle variabili;
4. giudizio finale dell'analista.

L'equazione cui alla fine giunse Altman è la seguente:

$$Z = 1,2x_1 + 1,4x_2 + 3,3x_3 + 0,6x_4 + 0,999x_5$$

Dove le variabili indipendenti sono:

1. x_1 : capitale circolante netto²⁶/totale attivo. Permette di relativizzare il capitale circolante netto all'intero attivo della società. Indica quanto la stessa disponga in termini di attività liquide nette rispetto al totale delle attività. Questo indice tiene quindi conto delle dimensioni dell'impresa e della sua liquidità, diretta e indiretta;
2. x_2 : utili non distribuiti/totale attivo. Questo indice aiuta indirettamente a tenere in considerazione anche l'età dell'impresa. È probabile, infatti, che le imprese giovani non abbiano avuto il tempo di

²⁶ Il Capitale circolante netto è l'ammontare di risorse a breve di una società, risorse che, quindi, finanziano l'attività operativa della stessa. La sola definizione può però trarre in inganno: nel capitale circolante vanno inclusi soltanto gli elementi operativi quindi attività e passività che abbiano natura non finanziaria, natura ricorrente nell'attività dell'azienda e natura monetaria (e non "contabile").

accumulare molti utili non distribuiti, per cui tale rapporto tende a essere più basso nelle imprese giovani, crescendo nel tempo. La scelta di questo rapporto non vuole essere una “discriminazione” dell’autore verso le imprese di più recente fondazione ma, al contrario, semplicemente riflette ed evidenzia il fatto che, empiricamente, le imprese giovani mostrano una possibilità di fallimento più alta rispetto a quelle più vecchie;

3. x_3 : EBIT²⁷/totale attivo. Indica una misura relativa del risultato aziendale prima dell’impatto degli oneri finanziari e delle tasse. Grazie all’effetto del denominatore, così come in tutte le altre variabili tranne la quarta, si tiene in considerazione la dimensione dell’impresa;
4. x_4 : Valore di mercato dell’*Equity/Book value* del debito totale. Il valore del numeratore è dato dal prodotto del numero di azioni emesse dalla società per il valore di mercato di un’azione. Per il denominatore si considera invece il *book value* ossia il valore contabile. Il debito totale include sia quello a breve sia quello a lungo termine. Con questo rapporto si vuole indicare quanto le attività possano diminuire prima che le passività le superino e le aziende diventino così insolventi. Questo concetto sarà poi ripreso nel modello di Merton;
5. x_5 : vendite/totale attivo. questo indicatore è particolarmente importante poiché indice il grado di efficienza nell’utilizzo dell’attivo. È il tasso di *turnover* delle vendite, esprime il numero di volte in cui il capitale investito si trasforma, nell’esercizio, in ricavi di vendita. È indicativo della capacità del *management* di rendere l’azienda competitiva.

I risultati della prima versione del modello di Altman lo portarono ad affermare che un’impresa potrà essere dichiarata sana se il suo *Z-Score* sia al di sopra del valore 2,99. Al contrario, per valori inferiori a 1,81 l’azienda viene dichiarata fallita mentre per valori intermedi vi è una zona di incertezza. Questo modello sembra funzionare abbastanza bene quando si fa riferimento ai dati di bilancio dell’ultimo anno. Mentre maggiore imprecisione si riscontra nell’applicazione su intervalli temporali più lunghi.

Come affermato sopra, nella prima versione del modello Altman prese in considerazione delle società quotate. Successivamente il modello venne esteso anche alle società non quotate portando alla nascita del modello *Z’-Score*. In esso, la differenza consiste nella quarta variabile indipendente x_4 . Infatti, non abbiamo più a disposizione il valore di mercato dell’*equity* e prenderemo allora in considerazione il suo *book value*. Insieme a tale modifica cambierà leggermente l’equazione e i percentili di riferimento, avendo adesso che l’azienda è considerata in crisi o in fallimento per valori minori di 1,23 mentre per valori superiori a 2,90 sarà considerata sana. Chiaramente, anche in questo caso i valori intermedi denotano l’area di incertezza.

²⁷ *Earnings before interest and taxes*. Indica il valore dell’utile operativo prima degli oneri finanziari e delle imposte.

Un'ulteriore versione del modello fu poi pensata per le aziende manifatturiere²⁸ o per quelle in cui una specifica attività abbia un peso molto rilevante sul bilancio. Il modello Z'' -Score può essere rappresentato come segue:

$$Z'' = 6,56x_1 + 3,26x_2 + 6,72x_3 + 1,05x_4$$

Questa volta le variabili indipendenti sono quattro e, rispetto alla prima versione, notiamo le seguenti differenze:

1. x_3 rappresenta il rapporto tra il reddito operativo e il totale dell'attivo;
2. x_4 indica il rapporto tra il *book value* dell'*equity* e le passività totali;
3. i valori dei parametri sono differenti;
4. l'insolvenza viene considerata per valori di Z'' minori di 1,1.

È possibile utilizzare l'analisi discriminante per arrivare a una stima della *probability of default* associata a una singola società. Infatti, se le variabili indipendenti sono distribuite secondo una normale multivariata, si è dimostrato che la PD può essere calcolata come segue:

$$PD = p(B|x_i) = \frac{1}{1 + \frac{1 - \pi_B}{\pi_B} e^{z_i - \alpha}}$$

dove:

1. z_i rappresenta lo *score* dell'impresa *i-esima*;
2. α rappresenta il valore limite che separa un'impresa sana da una insolvente nel modello, detto valore di *cut-off*;
3. π_B indica la probabilità di default *a priori*, ossia quella probabilità di default che non dipende dalla singola impresa ma da valori del settore o del mercato. Intuitivamente, più è basso tale valore e più sarà bassa la PD della singola impresa.

Dallo Z -Score di Altman deriva poi il modello EM -Score che si prefigge l'obiettivo di ottenere una valutazione del *rating* del debito delle società. Tutte le variabili utilizzate dal modello sono le stesse dello Z -Score ad eccezione delle vendite nette. Il valore massimo raggiungibile è di 8,15 e corrisponde perciò a un *rating* di AAA. Più tale valore decresce e più ci si avvicina al default dell'impresa cui corrisponde il valore zero nel modello.

I vantaggi dell'utilizzo di metodi come quelli appena analizzati sono molteplici:

²⁸ Nel 2004, uno studio di Bottani, Cipriani e Serao sperimentò l'applicazione del modello di Altman su un campione di PMI italiane. In particolare, furono prese in considerazione sessantasei aziende di cui trentatré già fallite nel 2002. I dati si basavano sui bilanci del 1999 e del 2000 (poiché empiricamente si nota che su periodi di tempo più lunghi il modello si mostra meno affidabile). La funzione discriminante permette di classificare le variabili in cinque indici di bilancio (relativi a redditività, liquidità, leva finanziaria, solvibilità e attività). Dai risultati di questo studio emerge che le aziende sono considerate in crisi con uno Z -Score minore di 4,864 mentre sono considerate sane con un valore maggiore di 8,105.

1. vengono prese in considerazione diverse caratteristiche comuni a più imprese e vengono analizzate le loro interazioni e influenze reciproche;
2. permette di considerare un numero abbastanza piccolo di variabili che tuttavia riesce a fornire informazioni significative;
3. consente di affrontare il problema della classificazione dell'intero insieme delle variabili piuttosto che esaminare le singole caratteristiche individualmente.

I limiti principali di questi modelli sono, in primo luogo, la mancanza di una teoria economica di base che possa dare autorevolezza e validità a questi studi e, in secondo luogo, l'impostazione di tipo lineare, la limitatezza dei dati e la logica *backward looking* dell'intero modello. I dati utilizzati, infatti, sono dati storici e non vi è modo di tenere in considerazione le aspettative future dell'azienda. Questo limita chiaramente il potere di previsione del modello.

Quando vengono utilizzati tali modelli possono essere commessi diversi tipi di errori. In particolare, gli errori cosiddetti del primo tipo riguardano la classificazione come sana di un'impresa insolvente. Al contrario, gli errori del secondo tipo riguardano la classificazione di un'impresa insolvente come sana.

I modelli che rispondono alla logica *forward looking* saranno esposti in seguito mentre nel prossimo paragrafo ci concentreremo sui modelli logit, probit e sulle reti neurali.

2.3.2 I modelli logit, probit e le reti neurali

Una possibile costruzione dei modelli può essere affidata a soluzioni legate alle regressioni lineari. Potremmo così determinare le probabilità di insolvenza e il loro peso. Tuttavia, la probabilità individuata dal modello potrebbe eccedere gli estremi $[0,1]$. Inoltre, i residui della regressione potrebbero avere varianze diverse poiché dipendono dalla variabile dipendente, andando incontro a problemi di eteroschedasticità. Una soluzione può essere fornita allora dai modelli logit e probit²⁹. Inoltre, questo risolve un'altra debolezza dei modelli, ossia la loro linearità.

I modelli probit e logit si basano sulla considerazione che vi sia una variabile non osservabile y_t^* determinabile tramite dei regressori nel seguente modo:

$$y_t^* = X_t\beta + \varepsilon_t, t = 1, 2, \dots, T \text{ con } E(\varepsilon_t) = 0, \forall t$$

Tuttavia, come anticipato, questa grandezza non risulta osservabile. Possiamo invece osservare la variabile dipendente y_t , con le seguenti caratteristiche:

$$y = \begin{cases} 1, & y_t^* > 0 \\ 0, & y_t^* \leq 0 \end{cases}$$

²⁹ Per maggiori dettagli sul modello logit si veda Grilli L., Rampichini C. (2014) Ordered Logit Model. In: Michalos A.C. (eds) Encyclopedia of Quality of Life and Well-Being Research. Springer, Dordrecht.

È interessante notare che la variabile y_t^* può essere moltiplicata per una qualsiasi costante non nulla e il modello non cambia:

$$\lambda y_t^* = \lambda X_t \beta + \lambda \varepsilon_t$$

Per tale motivo si può normalizzare il modello in maniera tale da ricavare un ε_t , termine di disturbo, con varianza unitaria:

$$E(\varepsilon^2) = \sigma^2 = 1$$

Il modello probit assume che i termini di disturbo ε_t si distribuiscano secondo una normale:

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

e

$$p(y_t = 1) = p(y_t^* > 0) = \Phi(X_t \beta)$$

in cui

$$\Phi(x) = \frac{1}{\sqrt{2\pi} \int_{-\infty}^x \exp\left\{-\frac{X^2}{2}\right\} dx}$$

che indica la funzione di ripartizione di una variabile normale standardizzata.

Secondo il modello logit, invece, i termini di disturbo ε_t seguono una distribuzione logistica.

$$F(\varepsilon_t) = \frac{\exp(\varepsilon_t)}{1 + \exp(\varepsilon_t)} = \frac{1}{1 + \exp(-\varepsilon_t)}$$

con la seguente funzione di ripartizione:

$$F(x) = \frac{e^x}{(1 + e^x)}$$

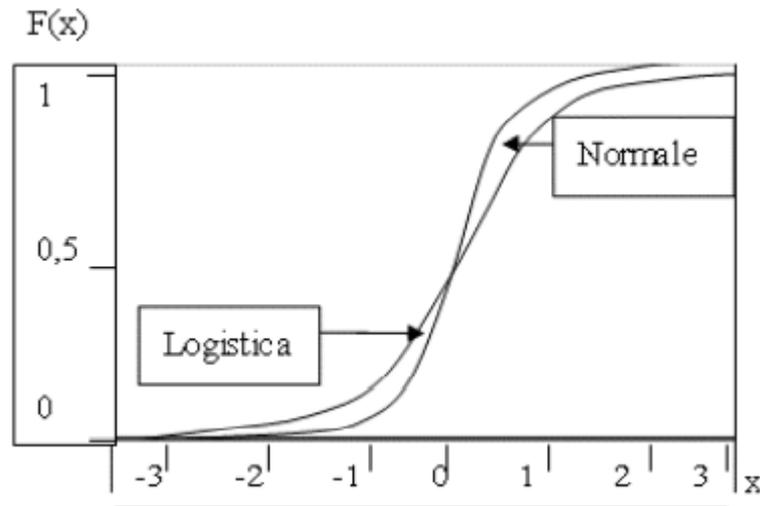


Figura 2 Grafico della funzione logistica e della normale. Fonte: Brunello, P., "La probabilità di sofferenza nei finanziamenti alle famiglie", (2008).

Possiamo vedere che la funzione logistica tende a zero per $x \rightarrow -\infty$ e tende a uno per $x \rightarrow \infty$. Vale 0,5 per $x=0$ e per valori molto vicini a zero ha un andamento simile a quello lineare.

Passiamo adesso a parlare dei modelli basati sulle reti neurali. Questi modelli si rifanno alle ricerche in ambito biologico, soprattutto su quelle relative alla struttura del cervello. Studiando il modo in cui l'uomo prende le decisioni o compie dei gesti si è cercato di comprendere la forma, le caratteristiche nonché l'organizzazione dei neuroni. L'obiettivo è quello di costruire dei nodi artificiali che siano in grado di simulare l'attività dei neuroni tramite una programmazione informatica. La connessione di più neuroni artificiali porta alla composizione della rete neuronale artificiale (*Artificial Neural Network, ANN*). Le reti neurali sono spesso usate in modo combinato con un altro "sistema intelligente", quello dell'algoritmo genetico (AG). Nel proseguo, spiegheremo il funzionamento di quest'ultimo proprio per l'importanza e la frequenza del loro utilizzo simultaneo. Quando analizziamo gli AG nell'analisi del rischio di credito abbiamo la possibilità di:

1. generare funzioni lineari;
2. generare *score* basati su specifiche regole.

Nel primo caso, possiamo pensare alla funzione genetica lineare nel modo seguente:

$$GLS = a_0 + a_1R_{n1} + a_2R_{j2} + \dots + a_nR_{rn}$$

Sarà poi l'algoritmo genetico a scegliere la costante a_0 , i coefficienti a_i e gli n indicatori tratti dalle n famiglie. Deve però essere l'analista a decidere i segni dei coefficienti a_i , il numero n delle famiglie e gli indicatori che appartengono a ciascuna famiglia. Se invece, come nel secondo punto, facciamo riferimento agli *score*, gli algoritmi genetici sono utilizzati per la produzione delle regole che si basano su dei *test* che fanno riferimento ai segni e al valore di chiusura degli indicatori.

Come anticipato, gli AG e le reti neurali vengono spesso usate insieme. Queste ultime possono essere considerate come una classe di modelli formati da diversi strati di unità elementari di elaborazione, detti *Processing Elements (PE)*, che hanno la funzione di elaborare informazioni tramite funzioni non lineari. Talvolta, ci si riferisce ad esse con il nome di neuroni o di nodi per analogia con il corrispettivo biologico.

Alcune interessanti caratteristiche delle reti neurali sono di seguito elencate:

1. sono potenzialmente adattive in quanto hanno capacità di apprendimento;
2. sono in grado di generalizzare;
3. sono resistenti al rumore³⁰;
4. sono resistenti alle lesioni³¹;
5. possono essere rappresentate simbolicamente con relativa semplicità.

Gli elementi essenziali della rete possono essere così indicati:

1. input/strati di attivazione (x_i);
2. connessioni/pesi (w_i);
3. stato di attivazione netto (Net);
4. funzione di attivazione ($f(Net)$);
5. valore di soglia/bias (θ);
6. output/stato di attivazione (y).

Per capire le effettive capacità di questo strumento dobbiamo comprendere la sua fase più importante ossia quella dell'apprendimento, o *training*, che ha luogo tramite un algoritmo che definisce la struttura della rete. Inoltre, è importante stabilire le leggi di apprendimento che stabiliscono come avvengono le variazioni alle connessioni della rete. Queste regole possono essere dei seguenti tipi:

1. apprendimento non supervisionato: la rete utilizza degli stimoli esterni fornendo delle risposte sulla base degli stessi. La rete si adatta da sola per avere una adeguata visione d'insieme;
2. apprendimento supervisionato: la rete, questa volta, interagisce indirettamente con il mondo reale attraverso sia gli stimoli sia delle risposte che vengono fornite da un intermediario esterno. In questo caso sarà necessario stabilire quale valore deve assumere ogni unità di *output* quando la rete riesce a classificare correttamente una osservazione relativa a un *input*. Bisogna, inoltre, presentare un *training set* ossia degli esempi di apprendimento di cui si conosce quale la corretta classificazione. Infine, bisogna valutare la risposta della rete: se questa è corretta si procede mentre se è errata

³⁰ Gli input che si ricevono nel mondo reale non sono "puliti" tuttavia la rete deve essere in grado di comprenderli ugualmente.

³¹ Se viene danneggiata una parte della rete essa è potenzialmente ancora in grado di operare.

bisogna calcolare l'errore tra la risposta della rete e quella corretta, modificando anche i pesi per poter ottenere in futuro la risposta desiderata.

In entrambi i casi sopra descritti, l'apprendimento avviene tramite più cicli (detti anche epoche) poiché gli *input* appartenenti al *training set* vengono ripresentati più volte. Avremo quindi un vettore y relativo ai valori di *output* e un vettore t relativo ai valori attesi. L'obiettivo sarà quindi quello di minimizzare la funzione di errore che ne esprime la differenza.

Applicando questi sistemi all'insolvenza delle imprese dovremmo individuare un insieme di imprese in default e un altro di aziende sane. Successivamente, dovremmo calcolarne i pesi discriminanti per poi applicare il modello a un altro insieme di imprese sane di cui si vuole verificare la possibilità di default. Quest'ultimo insieme è detto *validation set* poiché ci permette di controllare se la rete ha ben appreso, in quanto noi sappiamo che le aziende di quell'insieme sono ritenute sane ma la rete ignora questo dato.

Passiamo, infine, ad analizzare quali possono essere i punti deboli dei modelli basati sulle reti.

Innanzitutto, se il *training set* non viene scelto adeguatamente, la rete potrebbe non "imparare" dei casi della realtà che la renderanno potenzialmente incapace di riconoscere tali casi (ad esempio, una particolare situazione aziendale). Potremmo in questi casi incorrere in degli errori di valutazione. Inoltre, occorre considerare che la struttura della rete è "multistrato", motivo per cui esisteranno un certo numero di strati "nascosti" tra gli *input* e gli *output*. Infatti, se un problema ha delle soluzioni che si intersecano sarà necessario attivare un ulteriore strato nascosto per poter ricevere dalla rete la soluzione corretta.

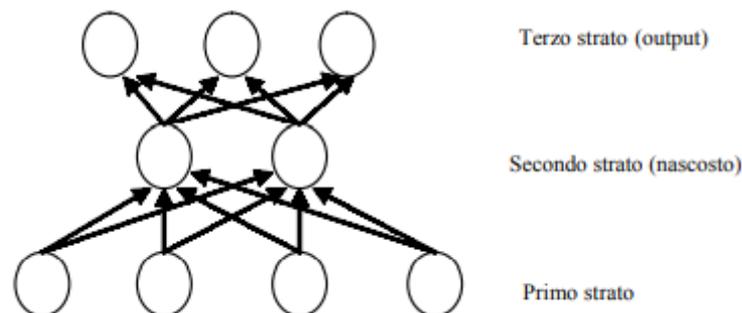


Figura 3 Esempificazione della struttura multistrato di una rete neurale. Fonte: D'annunzio, N., Falavigna, G., "Modelli di analisi e previsione del rischio di insolvenza. Una prospettiva delle metodologie applicate", Ceris-Cnr, W.P. n° 17/2004.

Infine, potremmo essere in presenza di vettori di *input* uguali che dovrebbero condurre a risultati diversi. La rete potrebbe non essere in grado di giungere ai corretti *output* poiché le mancano le informazioni discriminanti necessarie. Per evitare tali errori, sarà dunque necessario accertarsi che vi sia un numero adeguato di variabili che permettano alla rete di essere nelle condizioni di poter discriminare il fenomeno analizzato senza incorrere in errori di classificazione.

2.4 Modelli alternativi

In questo paragrafo analizzeremo dei modelli alternativi che negli anni sono stati proposti per superare le criticità esposte in precedenza.

2.4.1 Il metodo RAPD

Il metodo RAPD, ossia *Risk Analysis Probability of Default*, è una metodologia proposta da Montesi nel 2003 relativa al calcolo della probabilità di insolvenza. Tale metodo consta di quattro fasi:

1. costruzione del modello previsionale;
2. definizione e modulazione dello stato di default;
3. modellizzazione dell'incertezza;
4. simulazione Montecarlo per la stima della probabilità di default.

Con questo metodo non vogliamo ottenere il valore preciso di una specifica variabile nel futuro ma vogliamo una stima del *range* di valori possibili all'interno del quale si trovi il vero valore, in funzione di una distribuzione di probabilità.

Ogni prova genera allora uno scenario dell'azienda analizzata con relativo bilancio previsionale. La solvibilità aziendale viene stimata a seconda della situazione di incertezza ipotizzata nella simulazione. Quindi, la probabilità di default a un anno può essere calcolata come il numero di scenari in cui si realizza il default sul numero totale di simulazioni. In conclusione, sottolineiamo che la forza principale del modello è quella di essere *forward looking*.

2.4.2 La logica fuzzy

Da un punto di vista matematico, sia le reti neurali sia i sistemi *fuzzy* sono degli approssimatori di funzioni ossia sono in grado di approssimare i legami tra le variabili di *input* e quelle di *output* in modo anche non lineare³².

Se consideriamo un insieme A contenuto in un universo X, possiamo definire un insieme *fuzzy* A come un insieme di coppie ordinate

$$A = \{(x, \mu_A(x)) : x \in A, \mu_A(x) \in [0,1]\}$$

dove $\mu_A(x)$ è la funzione di appartenenza dell'insieme A. Definiamo "altezza" il massimo valore che può assumere la funzione di appartenenza dell'insieme A. Un insieme *fuzzy* è detto "normale" se il valore della

³² Per la redazione di questo paragrafo si è fatto riferimento in particolare a Cosma, S., Mastroleo, G., Facchinetti, G., Ferretti R., "I sistemi di credit scoring: un'applicazione della logica fuzzy nella valutazione delle piccole e medie imprese", ottobre 2001 e a D'Annunzio, N., Falavigna, G., "Modelli di analisi e previsione del rischio di insolvenza. Una prospettiva delle metodologie applicate", working paper n. 17/2004, Ceris-Cnr (2004).

sua altezza è pari a uno. In ogni caso, se l'insieme non è normale è possibile normalizzarlo sostituendo μ_A con:

$$\tilde{\mu}_A(x) = \frac{\mu_A(x)}{\max_{x \in A} \mu_A(x)}$$

Il dominio di un insieme *fuzzy* A è il dominio di $\mu_A(x)$, mentre il suo supporto è un sottoinsieme di X in cui $\mu_A(x)$ risulta positiva.

Un insieme *fuzzy* è detto convesso se $\forall \lambda \in [0,1]$ e $\forall x_1, x_2 \in A$:

$$\mu_A[\lambda x_1 + (1 - \lambda)x_2] \geq \min \{\mu_A(x_1), \mu_A(x_2)\}$$

Un numero *fuzzy* è un insieme *fuzzy* definito in R, convesso e normalizzato.

Un SEF, sistema esperto basato sulla logica *fuzzy*, può essere considerato un'evoluzione dei modelli Z-Score. I vantaggi ottenuti possono essere notevoli. Ad esempio, il funzionamento potrebbe richiedere minori tempi e minori dispendi di energie. Inoltre, il numero di regole per il loro funzionamento è inferiore di quello richiesto per i altri modelli. Anche i costi sono generalmente inferiori poiché richiedono tempi di sviluppo inferiori e possono essere implementati su postazioni a basso costo.

Un sistema esperto *fuzzy* è formato da un certo numero di insiemi *fuzzy* nei domini delle variabili di *input* e di *output*, da operatori logici e da regole qualitative. Per effettuare un'inferenza di tipo *fuzzy* bisogna analizzare le categorie delle variabili di *input*, collegarle alle categorie di *output* tramite blocchi di regole. Infine, è necessario unire le risposte ottenute per ottenere l'*output* finale. Di seguito un'immagine riepilogativa dell'inferenza operata in logica *fuzzy*.

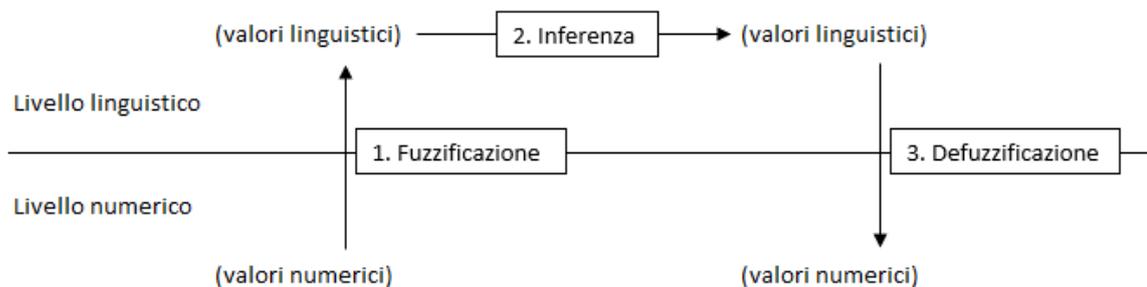


Figura 4 Funzionamento dell'inferenza con logica fuzzy. Fonte: Cosma, S., Facchinetti, G., Mastroleo, G., Ferretti, R., "I sistemi di credit scoring: un'applicazione della logica fuzzy nella valutazione delle piccole e medie imprese" (2001).

Utilizzando un sistema esperto *fuzzy* per un problema decisionale legato al rischio di credito possiamo tenere conto degli elementi soggettivi importanti ai fini della valutazione ma anche degli elementi oggettivi derivanti da dati economici e finanziari legati al debitore. Il sistema necessita comunque l'utilizzo di esperti del settore. Per tale motivo, il processo non è mai del tutto standardizzato ma deve, al contrario, essere personalizzato. Si inizia quindi con l'individuazione dell'istituto di credito che vuole implementare il sistema. Il passo successivo è la formazione dei blocchi di regole che traducono in un linguaggio "astratto" la volontà degli esperti relativa al loro giudizio. Nello *step* successivo bisogna "pesare" le regole tramite la produzione di casi

non reali (quindi non provenienti dai dati storici). Il programma dovrebbe a questo punto essere pronto per il controllo di efficacia finale. L'ultimo passaggio sarà quindi quello di analizzare le risposte del programma ai dati storici, quindi, di situazioni iniziate e già concluse di cui si dispongono i dati di *input* e *output*, con la verifica definita dell'efficacia del modello che potrà essere utilizzato, se ritenuto idoneo, dall'istituto di credito.

2.5 Il modello di Merton e il modello KMV

In questo paragrafo verrà approfondito il modello di Merton³³, appartenente alla famiglia di modelli che adottano un approccio di natura "strutturale" (o *firm's value models*), da distinguere dai modelli in forma ridotta (*intensity based models*) che analizzeremo nel successivo paragrafo, nonché dai modelli ibridi, ossia quelli che presentano caratteristiche intermedie tra i due approcci sopra elencati.

Come accennato nei paragrafi precedenti, i modelli con approccio strutturale cercano di descrivere la struttura della società analizzata con l'obiettivo di ricondurre in modo diretto l'eventuale evento di *default* a un avvenimento interno. Invece, i modelli in forma ridotta non precisano le motivazioni del *default* in senso stretto ma vogliono focalizzare l'attenzione sulla probabilità che esso si verifichi. Relativamente ai modelli ibridi si potrebbe obiettare, invece, che non esista un vero e proprio approccio "ibrido", per cui non potremmo classificarli in maniera autonoma. Essi, infatti, potrebbero essere pensati come modelli strutturali o in forma ridotta a cui, tuttavia, vengono aggiunte caratteristiche dei modelli dell'altra tipologia.

2.5.1 L'approccio del modello di Merton

Il modello di Merton prende avvio dai risultati ottenuti da Black e Scholes nel 1973 relativamente al modello del pricing delle opzioni. L'intuizione di Merton si basa su alcune premesse fondamentali. In particolare, il modello parte dal presupposto che una società risulti insolvente se il valore delle attività sia inferiore a quello delle passività. L'insolvenza ha luogo al momento del primo pagamento dovuto dall'impresa dopo che il valore delle passività abbia superato quello delle attività, rendendo più conveniente per gli azionisti lasciare la società ai creditori. Alla luce di quanto detto, evidenziamo anche le due ipotesi fondamentali alla base del modello:

1. Si fa sempre riferimento al rischio di insolvenza di una società il cui debito sia costituito integralmente da titoli obbligazionari zero-coupon;
2. il valore complessivo delle attività della società segue un processo diffusivo Browniano geometrico³⁴;

³³ La pubblicazione del lavoro originale si deve al Journal of Finance (1974).

³⁴ $V_t = V_0 \exp\{(\mu - \sigma^2/2)t + \sigma\sqrt{t}Z_t\}$, dove μ indica la *drift* percentuale istantanea, σ la volatilità percentuale istantanea, t il tempo e Z un processo di *Wiener*, o moto browniano standard.

Le due ipotesi conducono chiaramente a una semplificazione che, tuttavia, rappresenta anche i limiti del modello stesso. In primo luogo, considerando sempre debiti obbligazionari del tipo zero-coupon, assumiamo che l'insolvenza possa verificarsi solamente alla scadenza del debito. Inoltre, assumiamo di poter prevedere lo sviluppo delle attività della società (tramite il moto Browniano geometrico), potendo facilmente stimare la probabilità di insolvenza in relazione a un orizzonte temporale predefinito. Non è sempre verosimile che l'andamento delle attività sia approssimato correttamente tramite questo metodo. Proprio in risposta a tali limiti, negli anni successivi alla pubblicazione dei risultati raggiunti da Merton, alcuni studiosi hanno cercato di generalizzare il modello³⁵.

2.5.2 Il calcolo della *probability of default*

Il modello di Merton, appartenendo alla famiglia dei modelli con approccio strutturale, cerca di analizzare la struttura della società per ricondurne in modo diretto l'evento di *default*. Il calcolo della probabilità di insolvenza dipende allora dalle tre variabili qui descritte:

1. valore delle attività: esso è rappresentato dal valore di mercato degli *asset* della società. Può essere considerato come il valore attuale di tutti i flussi di cassa futuri che verranno generati dalla società (al lordo degli oneri finanziari) attualizzato secondo un tasso opportuno;
2. valore delle passività: può essere calcolato come il valore contabile delle passività rilevato dall'azienda. È, quindi, il valore nominale dei debiti che l'azienda è tenuta a rimborsare ai propri creditori;
3. rischio di attività (o *asset risk*): indica la volatilità del rendimento del valore di mercato delle attività e ne rappresenta, quindi, il livello di incertezza.

Quanto indicato sopra può essere visualizzato tramite la figura 5. È facile notare come l'insolvenza sopraggiunga quando il valore dell'attivo diventi inferiore a quello del debito. La probabilità di insolvenza è evidenziata nell'area in grigio nella coda sinistra della distribuzione. Intuitivamente, si può notare che la probabilità di insolvenza risulta maggiore all'aumentare del valore del debito D , al diminuire del valore dell'attivo A o all'aumentare della volatilità del rendimento del valore dell'attività σ_A .

³⁵ Tra gli studi più noti segnaliamo quello di Black e Cox del 1976, di Madan e Unal del 1994 e, infine, di Duffie e Singleton del 1994. Il primo ammette l'insolvenza anche prima della scadenza del titolo obbligazionario. Ciò è possibile considerando la presenza delle cedole oppure fissando una soglia nel valore dell'attivo al di sotto della quale l'emittente sia da considerarsi in default. Resta presente, in ogni caso, l'assunzione sull'andamento del valore dell'attivo secondo il moto Browniano geometrico. Invece. Negli approfondimenti di Madan e Unal e di Duffie e Singleton abbiamo un nuovo modo per stimare l'andamento del valore delle attività nel tempo. Non più il moto Browniano geometrico ma un processo evolutivo "a salti" (*jump process*).

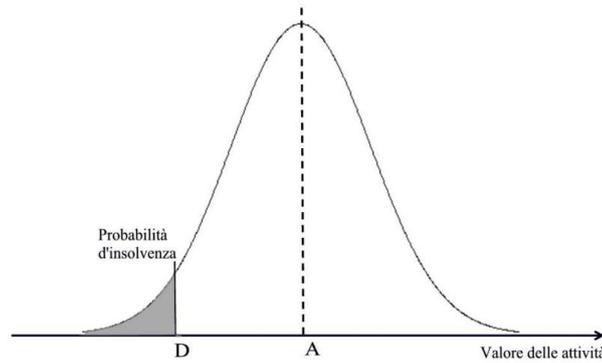


Figura 5 Distribuzione del valore di mercato dell'attivo e probabilità di insolvenza. Fonte: "Introduzione al modello di Merton/KMV per la stima delle probabilità di insolvenza di società quotate", a cura di C. Epis

Delle variabili menzionate, ossia A , D e σ_A , solo il valore del debito D è noto, poiché è semplice stimarlo dai dati di bilancio. Per quanto riguarda A e σ_A , Merton assimila il *payoff* dell'azionista della società a quello di un investitore acquirente di un'opzione *call* di tipo europeo sul valore dell'attivo della società con scadenza e prezzo di esercizio pari alla scadenza e al valore contabile del debito della società stessa:

1. se il valore dell'attivo A dell'impresa alla scadenza T è maggiore del valore del debito emesso D , allora i titoli valgono il loro valore nominale e la società risulta essere solvente;
2. al contrario, se il valore dell'attivo A dell'impresa alla scadenza T è minore del debito emesso, la società viene dichiarata insolvente e i creditori hanno il diritto di rivalersi sul suo attivo per la quota di credito loro spettante.

Per quanto riguarda i creditori, la loro posizione è assimilabile alla vendita di una ipotetica opzione *put* di tipo europeo. Essi hanno diritto, infatti, a ricevere alla scadenza T un ammontare pari al valore nominale dei titoli di debito. Pertanto, per i debitori il *payoff* alla scadenza T sarà calcolato come $\min\{A(T), K\}$, dove K è lo *strike price* dell'opzione. Per gli azionisti è possibile ottenere la parte dell'attivo che residui dopo la soddisfazione dei creditori. Possiamo allora calcolare il loro *payoff* come $\max\{A(T) - K, 0\}$, dove al posto dello *strike price* possiamo considerare il valore nominale del debito emesso.

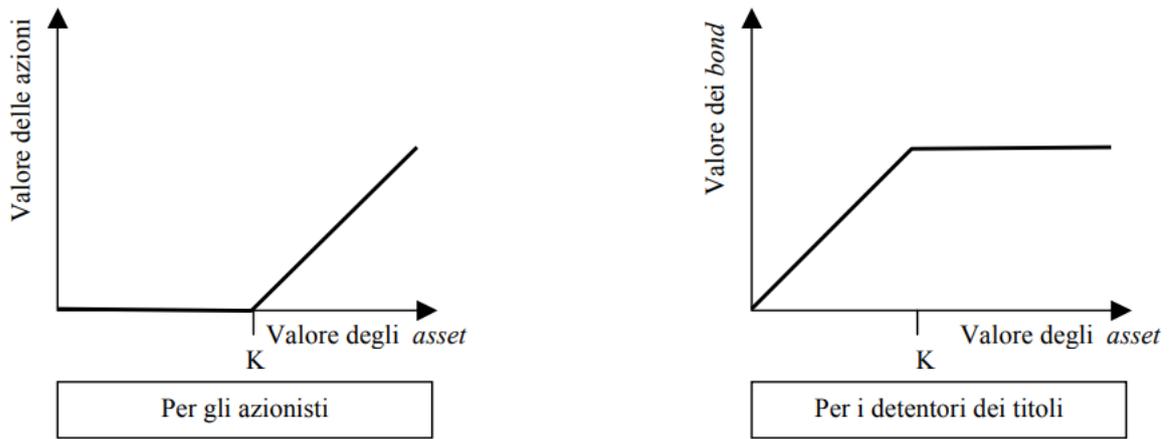


Figura 6 Payoff per gli azioni e per i creditori. Fonte: "Approcci matematico-statistici alla valutazione del rischio di credito". A cura di A. Matta

Un altro esempio può essere quello proposto in figura 7.

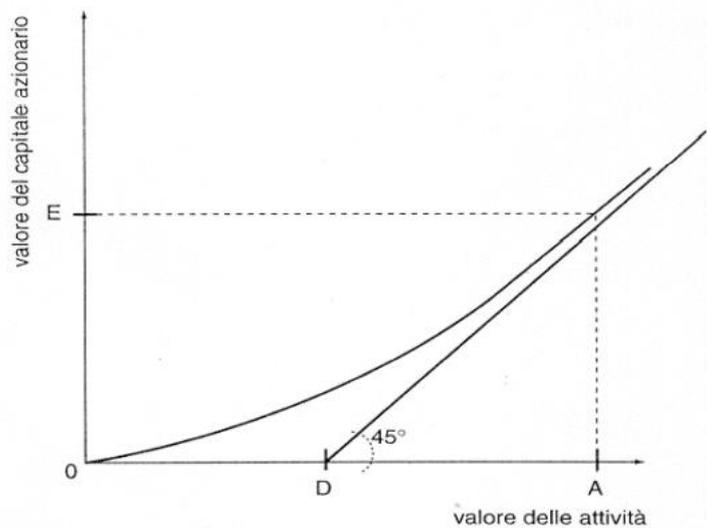


Figura 7 Valore dell'attivo e valore dell'equity. Fonte: "Sironi e Marsella (1999)"

Sull'asse delle ascisse troviamo il valore delle attività A mentre sull'asse delle ordinate trovate il valore dell'equity E . La retta inclinata a 45° indica il valore dell'attivo alla data di scadenza del debito T mentre l'altra curva indica lo stesso valore ad una data antecedente. Notiamo che a parità di valore dell'attivo A le due curve indicano valore della società differenti. Questo si spiega poiché prima della scadenza T sono possibili aggiustamenti strategici che portino a un incremento di valore della società. Per lo stesso motivo, per valore dell'attivo inferiori al debito il valore dell'equity sia positivo al posto di essere nullo (come invece accadrebbe alla scadenza).

L'incremento di valore dell'equity è, di fatto, dovuto a un trasferimento di valore dai debitori agli azionisti poiché gli azionisti della società potrebbero essere tentati, col fine di aumentare il valore della società, di accettare un livello di rischio maggiore implementando strategie più rischiose che aumentano il rischio di insolvenza e, conseguentemente, una riduzione del valore di mercato dei debiti della società. La differenza

tra i due valori dell'*equity* indicati dalle due curve è tanto più lontana quanto più grande è la volatilità del rendimento dell'attivo σ_A .

Precedentemente abbiamo paragonato il valore dell'*equity* al valore di un'opzione *call*. Possiamo affermare che il valore di un'opzione è una funzione di cinque variabili: prezzo di esercizio, valore di mercato dell'attività sottostante, volatilità dell'attività sottostante, tempo rimanente fino alla scadenza e tasso di interesse. Ne segue che possiamo impostare la seguente relazione:

$$E = f(D, A, \sigma_A, T, i)$$

dove T e i sono il tempo rimanente fino alla scadenza e il tasso di interesse. Delle sei presenti nella relazione precedente, notiamo che soltanto due non sono note: il valore di mercato A e la sua volatilità σ_A . Le altre sono tutte stimabili a patto di avere il prezzo di mercato delle azioni per la stima dell'*equity*. Per la stima delle due incognite possiamo impostare un'altra relazione:

$$\sigma_E = g(D, A, \sigma_A, T, i)$$

ossia un'equazione che lega la volatilità del valore di mercato dell'*equity* e la volatilità dell'attivo dell'impresa. Anche in questa equazione ritroviamo le stesse due variabili ignote ma possiamo risolvere un sistema formato dalle precedenti due equazioni e ricavare i valori di tutte le variabili.

Nonostante le intuizioni interessanti alla base del modello di Merton, non sono stati moltissimi i tentativi di applicazione pratica. La carenza di lavori empirici è forse riconducibile a un paio di fattori. Il primo è la difficoltà pratica dell'individuazione della soglia oltre la quale una diminuzione del valore dell'attivo conduce all'insolvenza. Un secondo fattore è dato dal fatto che non è verosimile applicare una distinzione netta tra *equity* e obbligazioni ma generalmente il finanziamento dell'impresa è più complesso e prevede la presenza di strumenti ibridi. A prescindere da ciò, un'applicazione molto interessante è stata quella proposta dalla società statunitense KMV.

2.5.3 Il modello KMV

In generale, ai fini della determinazione della *probability of default* necessitiamo di tre tipi di informazioni: dati di carattere finanziario, prezzi di mercato delle azioni e dei debiti e, infine, stime soggettive sulle prospettive future dell'impresa. I dati finanziari del bilancio sono dati che possiamo ottenere *ex post* e fanno riferimento al passato mentre i prezzi di mercato riflettono le aspettative che gli investitori hanno per il futuro. È facile comprendere, quindi, che il modello di cui abbiamo bisogno deve essere in grado di coniugare informazioni derivanti da dati storici e prospettici. Il modello sviluppato da KMV, in accordo con quanto detto, calcola la PD facendo riferimento sia ai dati di carattere finanziario sia ai prezzi di mercato delle azioni. Il primo passo è quello di stimare l'*asset value* e l'*asset volatility* della società. Successivamente, si calcola la

distance to default, grandezza che sarà introdotta e analizzata con maggior grado di dettaglio nei paragrafi successivi. Infine, otteniamo una stima della *PD*.

2.5.3.1 Stima dell'*asset value A* e dell'*asset volatility* σ_A

La probabilità che un'impresa superi il "punto di insolvenza", ossia il valore dell'attivo oltre il quale si presume si verifichi il fallimento della società, è funzione di importanti variabili: il valore dell'attivo A della società e la volatilità del rendimento dell'attivo σ_A . Il metodo attraverso cui vengono stimati tali valori è analogo a quanto descritto in precedenza per il modello di Merton. Le azioni della società, quindi, vengono pensate come delle opzioni *call* scritte sul valore dell'attivo aziendale, con *strike price* pari al valore dell'importo nominale del debito. In questo modo possiamo scrivere il valore delle azioni in funzione dell'attivo e della volatilità del suo rendimento. Similmente, possiamo immaginare il legame tra la volatilità delle azioni σ_E e la volatilità del rendimento dell'attivo σ_A . Possiamo allora scrivere il seguente sistema di due equazioni in due incognite:

$$\begin{cases} E = f(A, D, \sigma_A, T, i) \\ \sigma_E = g(A, D, \sigma_A, T, i) \end{cases}$$

Se σ_E fosse direttamente osservabile potremmo risolvere il sistema di equazioni e ottenere il valore di A e σ_A . Tuttavia, spesso, a causa della sensibilità di σ_E alle variazioni dell'attivo, la volatilità istantanea delle azioni risulta instabile³⁶, rendendo difficile la sua stima a partire dai dati di mercato. In questi casi, in modello KMV prevede l'utilizzo di un processo iterativo³⁷ per il calcolo del valore σ_E . Una volta ottenute queste stime è possibile procedere con il calcolo della *distance to default*.

2.5.3.2 Stima della *distance to default*

Le variabili che possiamo prendere in considerazione per la stima della *PD* in un dato orizzonte temporale sono:

1. Il valore corrente delle attività di una società;
2. La distribuzione di valore delle attività in un tempo t futuro all'interno dell'intervallo temporale selezionato H ;
3. La volatilità futura delle attività all'interno dell'intervallo temporale H ;
4. Il livello del punto di insolvenza (o "*default point*") e, di conseguenza, il valore delle passività iscritto in bilancio;
5. Il tasso di crescita delle attività all'interno dell'intervallo temporale H ;

³⁶ Ciò significa che la relazione tra σ_A e σ_E vale solamente all'istante.

³⁷ Di tale processo, tuttavia, KMV non fornisce documentazione specifica.

6. L'intervallo temporale H.

Considerando che l'intervallo H è fissato dagli analisti e che il tasso di crescita dell'attivo all'interno di tale intervallo generalmente impatta in maniera trascurabile sulla PD, le variabili su cui possiamo focalizzarci sono le prime quattro indicate nell'elenco precedente.

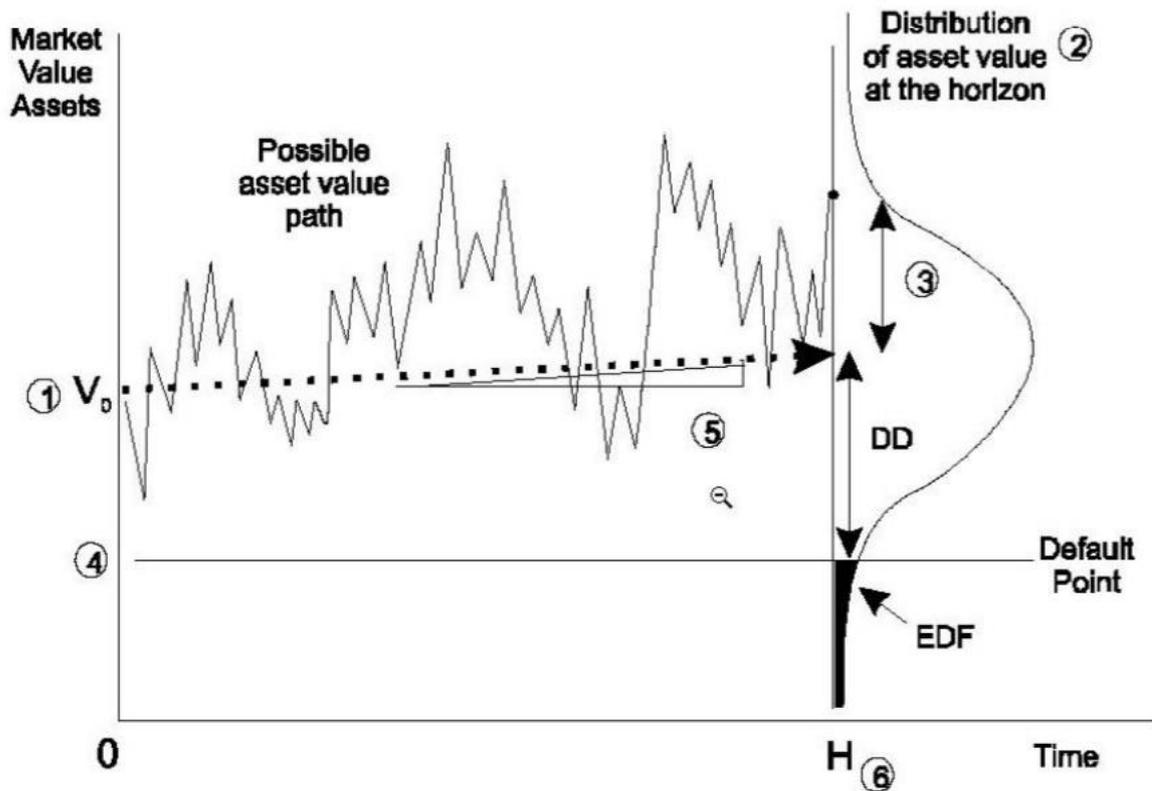


Figura 8 Indicazione delle variabili rilevanti all'interno del modello KMV. Fonte "Crosbie (1997)"

È facile verificare che maggiore è il valore delle attività rispetto al valore dei debiti minore sarà la PD. Allo stesso modo, minore è la volatilità del rendimento dell'attivo all'interno dell'intervallo in analisi e minore sarà, ancora una volta, la PD.

Se la distribuzione del valore delle attività fosse nota potremmo indicare la PD come la probabilità che il valore dell'attivo scenda al di sotto del default point. Tuttavia, tale distribuzione non è nota e non è sempre verosimile assumerne la normalità o la lognormalità, come invece avviene nel modello di Merton. Per tali motivi, la società KMV ha preferito optare per una distribuzione di tipo empirico.

Prima del calcolo della PD è stato allora introdotto un passaggio intermedio. Possiamo adesso introdurre una nuova variabile, la *distance to default* (DD). Essa identifica la "distanza", espressa in numero di deviazioni standard, dal valore di mercato delle attività dal default point. Un valore di DD pari a tre indica, quindi, che la distanza tra il valore di mercato dell'attivo dal punto di insolvenza è pari a tre deviazioni standard dell'attivo. In altri termini, una DD pari a tre indica che occorre che il valore di mercato delle attività si riduca di almeno tre volte la propria volatilità affinché si verifichi il default.

Possiamo calcolare la DD con l'equazione seguente:

$$DD = \frac{A - Dp}{A\sigma_A}$$

Considerando che:

$A - Dp = MNW$, dove MNW indica il *market net worth*, abbiamo che:

$$DD = \frac{MNW}{A\sigma_A}$$

Osservando le formule precedenti, possiamo notare che la DD è tanto minore quanto maggiore è volatilità del rendimento dell'attivo. Al contrario, la DD è tanto maggiore quanto maggiore è la differenza tra il valore dell'attivo e il punto di insolvenza.

Inoltre, avendo diviso il MNW per la volatilità dell'attivo $A\sigma_A$, abbiamo reso confrontabili i valori della DD di imprese con valori di mercato anche molto diversi tra loro. Quest'ultima misura, infatti, da sola non basta come indicatore di confronto relativamente alla PD poiché non tiene in conto degli effetti sulla rischiosità legati a elementi come la dimensione, il settore di appartenenza o la posizione geografica. Tali elementi sono invece presenti nella volatilità dell'attività. Potrebbe capitare che una società, pur avendo un MNW più basso rispetto a un'altra, abbia una PD uguale o inferiore.

La *distance to default* rappresenta l'ultimo elemento necessario da stimare prima di poter procedere con il calcolo della PD.

2.5.3.3 Stima della *probability of default*

La stima della PD è basata sulla cosiddetta "tabella di frequenza" ossia una tabella da cui possiamo osservare la relazione tra la PD e il livello di DD, ottenuta tramite un'analisi di tipo empirico. Il campione utilizzato da KMV per la stesura della tabella di frequenza comprende oltre centomila società quotate sul mercato statunitense, di cui circa 3.400 fallite, monitorate a partire dal 1973. Successivamente, il *database* è stato aggiornato con società non statunitensi per cercare di rendere la tabella meno legata al mercato americano.

Per ottenere la PD relativa a una società di cui è nota la DD è sufficiente consultare correttamente la tabella di cui si riporta in basso una versione semplificata:

Numero società	Fallimenti	DD	PD (in bp)
9000	720	1	800
15000	450	2	300
30000	200	3	100

35000	150	4	43
40000	20	5	7
42000	17	6	4

Tabella 6 Tabella di frequenza. Fonte "Estrella (2000)"

Se immaginiamo che su 5.000 imprese aventi tutte la stessa *distance to default* (es. DD = 4), si sia registrato il fallimento, entro l'intervallo di tempo definito dagli analisti, di 20 di esse, si può concludere che il tasso di insolvenza sia pari a³⁸:

$$PD = \frac{\text{società fallite}}{\text{totale società}} = \frac{20}{5000} = 40 \text{ bp}$$

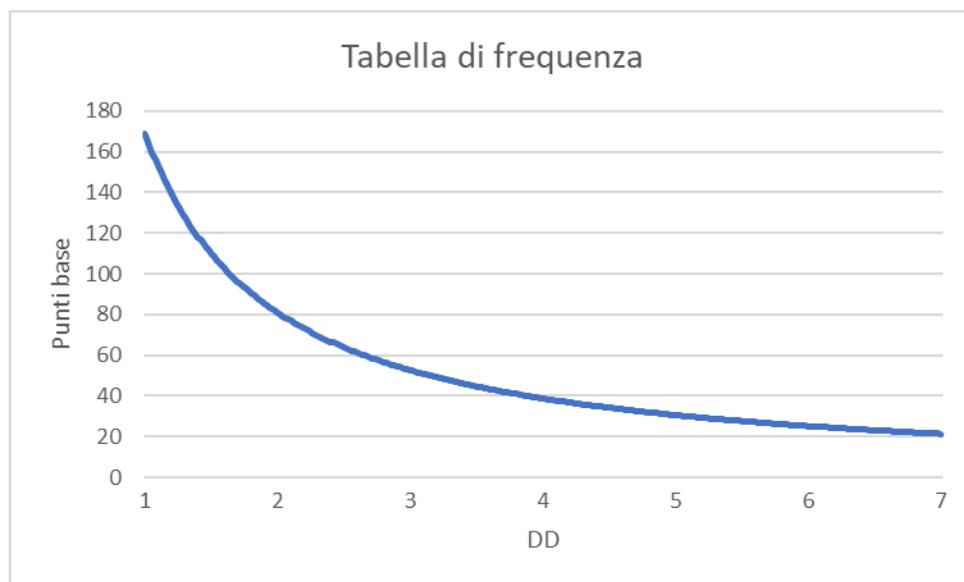


Figura 9 Rappresentazione grafica di una tabella di frequenza. Fonte "Introduzione al modello di Merton/KMV per la stima delle probabilità di insolvenza di società quotate", a cura di C. Epis

È opportuno sottolineare che la relazione tra PD e DD rimane verificata a prescindere dal tipo di società e dalla sua dimensione. Questo è dovuto al fatto che, come anticipato nel paragrafo precedente, le differenze dovute a tali fattori sono già rese confrontabili tramite il calcolo della DD.

Alcuni studi³⁹ hanno evidenziato che la relazione utilizzata da KMV per il calcolo della PD è utilizzabile, se necessario con alcuni correttivi, anche per aziende non statunitensi⁴⁰. Uno dei motivi per cui ciò accade è la costruzione della DD. Essa, infatti, viene calcolata come indicatore unicamente economico del rischio di insolvenza, non venendo presi in considerazione altri aspetti come, ad esempio, la tutela fornita dalla politica

³⁸ L'esempio è stato ripreso da "Introduzione al modello di Merton/KMV per la stima delle probabilità di insolvenza di società quotate", a cura di "C. Epis".

³⁹ Si veda, per esempio, lo studio di Sellers et al. (2000)

⁴⁰ Si ricordi che il campione utilizzato da KMV considerava aziende statunitensi. Una conferma della possibile estensione del modello è data dal mercato asiatico e, in particolare, alla crisi che colpì alcuni paesi asiatici negli anni Novanta che KMV fu in grado di prevedere con largo anticipo rispetto agli altri istituti specializzati nella stima dei tassi di insolvenza.

tramite aiuti statali (se così non fosse, le tabelle di frequenza costruite per uno stato sarebbero difficilmente applicabili ad altri paesi). Resta lasciato all'abilità e al giudizio professionale degli analisti la possibilità di integrare l'impatto di tali fattori, qualora considerati rilevanti, all'interno del processo di stima della PD.

Per confrontare i risultati ottenuti da KMV con quelli delle più importanti agenzie di *rating* è possibile consultare la seguente tabella indicante le corrispondenza tra i valori della PD e le classi di rating:

PD	S&P	Moody's
2-4 bps	AA	Aa2
4-10 bps	AA/A	A1
10-19 bps	A/BBB+	Baa1
19-40 bps	BBB+/BBB-	Baa3
40-72 bps	BBB-/BB	Ba1
72-101 bps	BB/BB-	Ba3
101-143 bps	BB-/B+	B1
143-202 bps	B+/B	B2
202-345 bps	B/B-	B2

Tabella 7 Corrispondenza tra PD e classi di rating. Fonte: "Bellucci et al. (1997)"

2.5.3.4 Calcolo della PD tramite la formula di Black e Scholes

La fase conclusiva del procedimento, ossia la stima della PD, può essere effettuata tramite analisi empirica e redazione della tabella di frequenza oppure tramite applicazione della formula di Black e Scholes, partendo dall'ipotesi secondo la quale i valori della distribuzione dell'attivo si distribuiscono secondo una log-normale.

Tramite il modello di Black e Scholes, la variazione nel tempo del valore di mercato delle attività di un'impresa può essere descritta tramite l'equazione differenziale:

$$dA_t = \mu A_t dt + \sigma_A A_t dW_t$$

dove:

1. A_t è il valore dell'attivo al tempo t ;
2. μ rappresenta il tasso di crescita istantaneo dell'attivo;
3. W è un moto Browniano standard.

Immaginando che esistano solo una categoria di debiti e una di azioni e indicando con D il valore di bilancio dei debiti in scadenza al tempo T , allora possiamo indicare il valore di mercato delle azioni tramite la seguente equazione:

$$E = AN(d1) - e^{-rT}DN(d2)$$

dove:

1. E indica il valore di mercato delle azioni;
2. r il tasso di interesse *risk free*;

$$3. d1 = \frac{\ln\left(\frac{A}{D}\right) + \left(r + \frac{\sigma_A^2}{2}\right)T}{\sigma_A\sqrt{T}};$$

$$4. d2 = d1 - \sigma_A\sqrt{T} = \frac{\ln\left(\frac{A}{D}\right) + \left(r - \frac{\sigma_A^2}{2}\right)T}{\sigma_A\sqrt{T}}.$$

In base al lemma di Ito⁴¹, possiamo osservare che deve essere soddisfatta la seguente relazione tra la volatilità delle azioni e la volatilità delle attività:

$$\sigma_E = \frac{A}{E}N(d1)\sigma_A$$

Risolvendo il seguente sistema possiamo ricavare il valore corrente delle attività e il valore del loro rendimento:

$$\begin{cases} E = AN(d1) - e^{-rT}DN(d2) \\ \sigma_E = \frac{A}{E}N(d1)\sigma_A \end{cases}$$

Ricordando che la PD per il modello KMV è la probabilità che il valore di mercato delle attività scenda al di sotto del punto di *default*, possiamo scrivere che:

$$P_t = Pr[A_t < Dp_t | A_0 = A] = Pr[\ln A_t \leq \ln Dp_t | A_0 = A]$$

dove P_t è la PD, A_t rappresenta il valore di mercato dell'attivo e Dp_t il punto di *default*, tutti relativamente all'istante t .

⁴¹ Il lemma di Ito è usato nel calcolo stocastico per calcolare il differenziale di una funzione di un tipo particolare di processo stocastico. Se ne propone una breve descrizione. Sia $x(t)$ un processo di Ito. In altri termini, $x(t)$ soddisfa l'equazione differenziale stocastica: $dx(t) = a(x(t))dt + b(x(t))dW_t$. Sia f una funzione di due variabili (x, t), differenziabili con continuità due volte rispetto alla prima variabile e una volta rispetto alla seconda. Allora, abbiamo che $f(x(t), t)$ è ancora un processo di Ito. Inoltre: $df(x(t), t) = \left(a(x, t)\frac{\partial f}{\partial x} + \frac{\partial f}{\partial t} + \frac{1}{2}(b(x, t))^2\frac{\partial^2 f}{\partial x^2}\right)dt + b(x, t)\frac{\partial f}{\partial x}dW_t$.

Considerando che l'evoluzione del valore delle attività, come sopra riportato, segue la seguente relazione $dA_t = \mu A_t dt + \sigma_A A_t dW_t$, possiamo affermare che il valore dell'attivo in t , nell'ipotesi che $A_0 = A$, sia data dall'equazione:

$$\ln A_t = \ln A + \left(\mu - \frac{\sigma_A^2}{2} \right) t + \sigma_A \sqrt{t} \varepsilon,$$

dove μ rappresenta il ricavo atteso dalle attività dell'impresa ε indica una componente aleatoria su tale guadagno.

Combinando la seguente equazione

$$P_t = Pr[A_t < Dp_t | A_0 = A] = Pr[\ln A_t \leq \ln Dp_t | A_0 = A]$$

con

$$\ln A_t = \ln A + \left(\mu - \frac{\sigma_A^2}{2} \right) t + \sigma_A \sqrt{t} \varepsilon$$

possiamo indicare la PD come:

$$P_t = Pr \left[\ln A + \left(\mu - \frac{\sigma_A^2}{2} \right) t + \sigma_A \sqrt{t} \varepsilon \leq \ln Dp_t \right]$$

È possibile riscrivere la precedente equazione nel seguente modo:

$$P_t = Pr \left[- \frac{\ln \frac{A}{Dp_t} + \left(\mu - \frac{\sigma_A^2}{2} \right) t}{\sigma_A \sqrt{t}} \leq \varepsilon \right]$$

Se assumiamo che la componente aleatoria della funzione di ripartizione N di una variabile aleatoria normale standard ε sia normalmente distribuita, possiamo definire la PD come:

$$P_t = N \left[- \frac{\ln \frac{A}{Dp_t} + \left(\mu - \frac{\sigma_A^2}{2} \right) t}{\sigma_A \sqrt{t}} \right]$$

La DD, ricordando che indica quanto una società sia "lontana" dal fallimento in termini di deviazioni standard, può essere scritta come:

$$DD = \frac{\ln \frac{A}{Dp_t} + \left(\mu - \frac{\sigma_A^2}{2} \right) t}{\sigma_A \sqrt{t}}$$

Considerando le ultime due formule congiuntamente, abbiamo che

$PD = N(-DD)$. Nell'ipotesi di neutralità verso il rischio, il guadagno atteso dell'attivo μ è uguale al tasso *risk free* per qualsiasi orizzonte temporale. In questo caso, possiamo scrivere che:

$$P_t = N \left[- \frac{\ln \frac{A}{Dp_t} + \left(r - \frac{\sigma_A^2}{2} \right) t}{\sigma_A \sqrt{t}} \right]$$

L'utilizzo del modello di Black e Scholes presenta il vantaggio di stimare la PD sulla base di un sistema matematico-finanziario, senza un necessario riferimento a delle serie storiche di dati. Tuttavia, vi sono alcuni aspetti negativi da tenere in considerazione. La DD tiene in considerazione solo gli incrementi dell'attivo A della società prodotti dal tasso $\left(\mu - \frac{\sigma_A^2}{2} \right)$, non considerando in questo modo le uscite di cassa prodotte, ad esempio, da interessi e dividendi.

Il secondo limite consiste nell'approssimazione legata all'assunzione di Normalità. In particolare, evidenziamo che la distribuzione empirica di KMV presenta code molto più ampie rispetto a quelle della Normale.

Infine, una semplificazione non verosimile è il ricorso alla *call* europea, legando quindi il fallimento dell'impresa a uno specifico intervallo temporale, in corrispondenza della scadenza del debito.

2.6 I modelli in forma ridotta

I limiti dei modelli strutturali hanno portato numerosi studiosi, a partire dalla seconda metà degli anni Novanta, a voler creare una nuova categoria di modelli, ossia i modelli in forma ridotta.

Come anticipato nell'introduzione al paragrafo 2.5, i modelli in forma ridotta non vogliono analizzare l'impresa per cercare di correlare dei valori di bilancio all'evento *default* ma si prefiggono come scopo quello di calcolare la probabilità di *default* direttamente dai dati di mercato. Si cerca quindi di porre tutta l'attenzione sull'evento *default* in quanto tale, considerandolo come un dato esogeno al modello.

2.6.1 L'intensità di default e il valore di recupero

Per realizzare tale tipo di analisi introduciamo una nuova grandezza detta "intensità di *default*", indicata con la lettera λ . Essa rappresenta la probabilità condizionale che l'impresa osservata non riesca ad adempiere alle proprie obbligazioni in un dato intervallo temporale successivo all'osservazione e sotto l'ipotesi che non si sia ancora verificata l'insolvenza.

Possiamo inoltre aggiungere che per lo studio degli eventi di *default* delle società che emettono titoli di debito è possibile sfruttare le proprietà dei processi di Poisson, che ben si adattano a tali scenari.

Un processo di Poisson $N(t)$ è definito come un processo di conteggio crescente che può assumere solo valori interi non negativi (0, 1, 2, ...) e assume valore iniziale uguale a zero.

Indicando con $\tau_1, \tau_2, \tau_3, \dots$ i tempi in cui avvengono i salti del processo, abbiamo che:

$$N(t) = \sum_{i=1}^{\infty} \mathbf{1}_{\{\tau_i \leq t\}}$$

Dove:

$$\mathbf{1}_{\{\tau_i \leq t\}} \begin{cases} \text{vale 1 se si verifica l'evento} \\ \text{vale 0 altrimenti} \end{cases}$$

Possiamo allora esprimere il *default* matematicamente tramite l'introduzione di una variabile aleatoria che chiameremo *time to default*, indicata con τ . Possiamo interpretarla come l'istante in cui avviene il primo salto in un processo di Poisson $N(t)$. Possiamo quindi scrivere che:

$$\tau = \inf\{t \geq 0 : N(t) > 0\}$$

Possiamo adesso legare il processo di Poisson omogeneo⁴² $N(t)$ all'intensità di *default* λ introdotta precedentemente. Si può provare che, per un processo di Poisson omogeneo, la probabilità che si verifichi un salto in un intervallo di tempo piccolo Δt , successivo a t , è data dalla seguente relazione:

$$P[N(t + \Delta t) - N(t) = 1] = \lambda \Delta t + o(\Delta t), \Delta t \rightarrow 0^+$$

Ciò implica che la probabilità che si verifichi di più di un salto in un singolo intervallo di tempo è trascurabile in relazione a $\lambda \Delta t$.

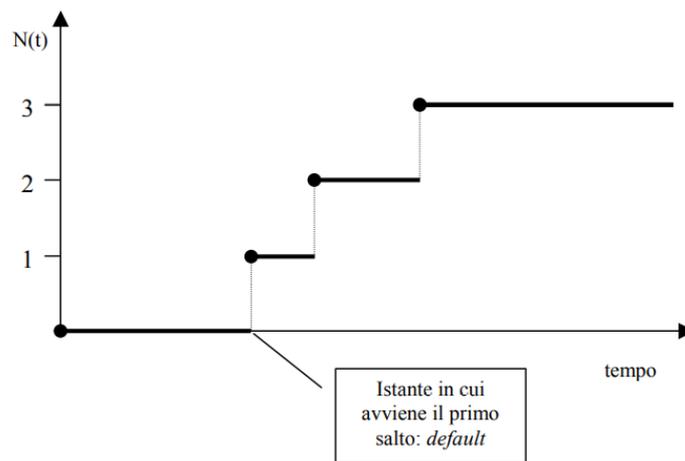


Figura 10 Illustrazione di un processo di Poisson con indicazione dell'istante in cui si verifica l'insolvenza. Fonte: "Approcci matematico-statistici alla valutazione del rischio di credito", a cura di A. Matta.

⁴² Il modo in cui definiamo l'intensità di *default* caratterizza il modello che utilizza i processi di Poisson. Possiamo allora presentare le seguenti definizioni:

- Processo di Poisson omogeneo: l'intensità di *default* è costante nel tempo e deterministica;
- Processo di Poisson inomogeneo: l'intensità di *default* varia in funzione del tempo;
- Processo di Cox: l'intensità di *default* viene descritta da una relazione di tipo stocastico.

Dall'equazione precedente si può ricavare la funzione di ripartizione $F(t)$ della variabile aleatoria τ , definita come:

$$F(t) = P[\tau \leq t] = 1 - \exp\{-\lambda t\} \text{ con } t \geq 0$$

che indica la probabilità che si verifichi il *default* entro il tempo t . Tramite il suo complemento a uno, di conseguenza, è agevole ricavare la funzione di "sopravvivenza", detta *survival probability*, che ovviamente indica la probabilità che l'insolvenza non si sia verificata fino al tempo t .

$$S(t) = 1 - F(t) = P[\tau > t] = \exp\{-\lambda t\} \text{ con } t \geq 0$$

Se ipotizziamo un tasso di interesse r costante e un *recovery rate* pari a zero, possiamo calcolare il valore in t di uno *zero coupon bond* rischioso con scadenza T per unità di nominale $D(t,T)$ paragonandolo a un titolo analogo ma privo di rischio $B(t,T)$ il cui rendimento venga aumentato dall'intensità di *default* con tasso di recupero nullo.

$$D(t,T) = P[\tau > t]B[t,T] = e^{-(r+\lambda)(T-t)}$$

In questo modo, λ può essere interpretato come lo *spread* creditizio di un'attività rischiosa rispetto al suo analogo *risk-free* nell'intervallo temporale di riferimento.

Tuttavia, di solito il fallimento di un soggetto non comporta la perdita totale dell'investimento. Rimane quindi una parte del valore dei titoli di debito emessi. Per la definizione dei tassi di recupero si fa spesso riferimento a delle convenzioni. Tra le più utilizzate vi sono quelle che prendono in considerazione, per unità di valore nominale:

1. un "*constant recovery*";
2. un "*equivalent recovery*";
3. un "*fractional recovery*".

La prima convenzione suppone possibile definire *ex ante* il valore di recupero in caso di insolvenza, senza dover considerare né quali siano state le cause che hanno portato al *default* né il tasso di interesse *risk free* (per attività prive di rischio) al momento dell'insolvenza.

Nonostante queste premesse siano molto stringenti, la teoria del *constant recovery* offre il vantaggio di poter semplificare notevolmente il modello e snellire i calcoli ed è particolarmente utile nel caso in cui si possa usufruire di dati medi di *recovery* forniti dalle agenzie di *rating*.

La seconda convenzione, il *recovery* di tipo "equivalente", poggia le sue basi su un'intuizione di Jarrow e Turnbull del 1995. Secondo loro, infatti, un'attività in *default* può essere pensata come una frazione R , $R \in [0,1]$, di titoli *risk free* con valore nominale pari al valore del debito a scadenza.

Infine, l'ultima convenzione è relativa al *fractional recovery*. Secondo questo modello, l'insolvenza non porta alla liquidazione complessiva dell'emittente ma solamente a una ristrutturazione del debito. Per tale motivo, in seguito al verificarsi dell'insolvenza, i titoli perdono solamente una porzione q del loro valore. La quota rimanente $(1-q)$ conserva il suo valore e può nuovamente essere negoziato tramite nuovi contratti. Tramite questa ipotesi possiamo calcolare il valore di un titolo senza cedola con valore nominale unitario tramite la seguente equazione:

$$D(t, T) = e^{-[r+(1-R)\lambda](T-t)}$$

Dove $R, R \in [0,1]$ rappresenta il *recovery rate*.

Nonostante le alternative più utilizzate circa la stima del *recovery rate* siano quelle sopra esposte, i risultati ottenuti una volta implementato il modello in forma ridotta non divergono significativamente. Per tale motivo, si preferisce spesso l'utilizzo del *constant recovery* in virtù delle semplificazioni di cui beneficia il modello.

Chiarito il tema del valore di recupero dobbiamo adesso concentrarci sulla stima dell'intensità di *default*. In un modello in forma ridotta, infatti, questa stima risulta particolarmente delicata. Generalmente, vengono impiegati due metodi:

1. il primo prevede l'utilizzo di dati storici relativi ai *default* forniti dalle agenzie di *rating*. Tali dati vengono ordinati nelle matrici di transizione, ossia matrici che indicano la probabilità di migrazione, empiricamente determinata, da una classe di *rating* di partenza a una di arrivo, in un intervallo di tempo determinato. Tramite questo metodo è possibile definire la probabilità storica di *default*, entro i successivi t periodi, di un soggetto appartenente a una determinata classe di *rating*;
2. il secondo metodo prevede, invece, l'utilizzo di dati direttamente osservabili sui mercati finanziari, facendo riferimento in particolare agli *spread* creditizi.

Per stimare l'intensità di *default* dai dati di mercato occorre calcolare il rendimento a scadenza di un titolo rischioso e confrontarlo con il tasso *risk-free* r alla stessa scadenza, ottenendo così una prima indicazione sul rischio di credito percepito dal mercato relativamente a uno specifico soggetto, in particolare in termini di *credit spread*.

Se consideriamo un tasso di *recovery* pari a zero, lo *spread* creditizio $S(t, T)$ è uguale all'intensità di *default*. Per cui abbiamo che $S(t, T) = \lambda$.

Diverso è il caso in cui ipotizziamo un tasso di *recovery* maggiore di zero. Se immaginiamo che il valore di recupero di un titolo rischioso senza cedole sia costante e uguale a R , possiamo calcolare il valore dell'intensità di *default* tramite la seguente equazione:

$$\lambda = -\frac{1}{T-t} \ln \left[\frac{D(t, T) - R \cdot e^{-r(T-t)}}{(1-R) \cdot e^{-r(T-t)}} \right]$$

Un esempio può chiarire maggiormente quanto esposto sopra⁴³. Consideriamo due titoli senza cedola con scadenza a un anno, uno rischioso che al momento dell'osservazione ha un valore di mercato di 0,95 mentre l'altro, non rischioso, con quotazione di 0,98. Il rendimento a scadenza del primo è il 5,13% mentre quello del secondo è inferiore, il 2,02%. Possiamo agevolmente calcolare lo *spread* creditizio, pari in questo caso a 3,11%. Questo valore coincide, sotto l'ipotesi del tasso di *recovery* nullo, con l'intensità di *default* λ . Nel caso in cui, invece, il tasso di *recovery* sia maggiore di zero, ad esempio un tasso costante e pari al 30% del valore nominale, l'intensità di *default* viene così stimata:

$$\lambda_{R=30\%} = -\frac{1}{T-t} \ln \left[\frac{D(t, T) - R \cdot e^{-r(T-t)}}{(1-R) \cdot e^{-r(T-t)}} \right] = -\frac{1}{1} \ln \left[\frac{0,98 - 0,30 \cdot 0,98}{(1-0,30) \cdot 0,98} \right] = 4,47\%$$

La differenza rilevata tra $\lambda_{R=0} = 3,11\%$ e $\lambda_{R=30\%} = 4,47\%$ può essere spiegata considerando che nel primo caso non si recupera nulla nel caso in cui si verifichi l'evento del *default* mentre nel secondo caso si recupererebbe una frazione pari al 30%.

2.6.2 Estensione al caso multivariato

In questo paragrafo analizziamo come si comportano i modelli in forma ridotta nel caso multivariato. Per operare questa estensione possiamo considerare un numero m di *shock*, ossia eventi che conducono al *default* degli emittenti, relativamente a un numero k di titoli appartenenti a un insieme n di sottostanti. Possiamo definire le relazioni all'interno di un portafoglio titoli modellando ciascuno *shock* j attraverso un Processo di Poisson $N_j(t)$, e ammettendo che i processi N_1, \dots, N_m siano indipendenti⁴⁴. Considerando queste ipotesi, possiamo scrivere la probabilità di sopravvivenza della i -esima società come:

$$S_i(t) = P[\tau_i > t] = \exp \left\{ - \sum_{k=1}^m a_{ik} \lambda_k t \right\}$$

Dove il termine a_{ik} può assumere un valore unitario oppure nullo a seconda del coinvolgimento o meno dell'attività di riferimento nello shock. Estendendo la formula al caso della sopravvivenza congiunta di n soggetti rischiosi, possiamo calcolare la probabilità di insolvenza come:

$$S_i(t_1, t_2, \dots, t_n) = P[\tau_1 > t_1, \tau_2 > t_2, \dots, \tau_n > t_n] = \exp \left\{ - \sum_{k=1}^m \lambda_k \max\{a_{1k}t_1, a_{2k}t_2, \dots, a_{nk}t_n\} \right\}$$

Consideriamo, per esempio, due società α e β . L'intensità della prima è stimata in 0,06 e quella della seconda in 0,04 per cui abbiamo che $\lambda_\alpha = 0,06$ e $\lambda_\beta = 0,04$. L'intensità di *default* congiunto vale 0,015, per cui $\lambda_{\alpha\beta} =$

⁴³ L'esempio proposto è tratto da "Approcci matematico-statistici alla valutazione del rischio di credito", a cura di A. Matta

⁴⁴ Gli *shock* possono essere analizzati come tendenza di una singola azienda ad andare incontro all'evento *default*, oppure come evento relativo a tutte le imprese appartenenti a uno specifico settore o a una specifica nazione oppure, ancora, come evento relativo all'intera economia.

0,015. Possiamo, quindi, stimare la probabilità di sopravvivenza a un anno delle due società prese singolarmente, come:

$$S_{\alpha}(1) = P[\tau_{\alpha} > 1] = \exp\{-(1 \cdot 0,06 + 0 \cdot 0,04 + 1 \cdot 0,015) \cdot 1\} = 92,77\%$$

$$S_{\beta}(1) = P[\tau_{\beta} > 1] = \exp\{-(0 \cdot 0,06 + 1 \cdot 0,04 + 1 \cdot 0,015) \cdot 1\} = 94,65\%$$

Tramite il calcolo della probabilità congiunta ricaviamo, invece, la stima della probabilità relativa al mancato verificarsi dell'evento *default* per entrambe le società entro un anno:

$$S_{\alpha\beta}(1) = P[\tau_{\alpha} > 1, \tau_{\beta} > 1] = \exp\{-(1 \cdot 0,06 + 1 \cdot 0,04 + 1 \cdot 0,015) \cdot 1\} = 89,14\%$$

Come ci aspettavamo, a una intensità di *default* λ superiore corrisponde una minore probabilità di sopravvivenza $S(t)$.

2.7 Accordi internazionali e modelli sviluppati da banche

Considerando la delicatezza del tema in analisi, la complessità delle valutazioni sul rischio di credito e gli impatti sull'intero sistema economico, non sorprende che negli anni siano nati accordi internazionali per regolamentare e armonizzare i comportamenti degli istituti di per minimizzare i potenziali riflessi negativi conseguenti a eventi connessi al *default* di alcune soggetti.

2.7.1 Il modello di Vasicek

I metodi IRB relativi ai requisiti patrimoniali per limitare il rischio di credito, voluti con gli accordi di Basilea II, si basano sul modello della copula Gaussiana proposto da Vasicek. Tramite tale modello è possibile calcolare i percentili della distribuzione del tasso di insolvenza su un portafoglio di prestiti.

Considerando un portafoglio di prestiti relativo a N società, sia $T_i (1 \leq i \leq N)$ il tempo in cui si verifica il fallimento della i -esima società. Ciò implica che, prima o poi, tutte le società andranno incontro all'evento *default*, tuttavia, il tempo rimanente prima del manifestarsi del *default* potenzialmente può essere estremamente lungo, anche più di un secolo.

Definiamo Q_i come la funzione di distribuzione delle T_i , per cui abbiamo che:

$$Q_i = \text{Prob}(T_i < T)$$

Possiamo adesso trasformare, percentile per percentile, le T_i nelle variabili U_i , aventi distribuzione normale standard, supponendo che la struttura delle correlazioni tra le U_i sia descrivibile tramite l'equazione:

$$U_i = a_i F + \sqrt{1 - a_i^2} Z_i$$

dove anche F e Z_i hanno distribuzioni normali standard indipendenti tra loro.

Ipotizziamo che la distribuzione Q_i del *time to default* T_i rimanga invariata per ogni i e sia uguale alla *probability of default* e che la *copula correlation* sia la stessa per ogni coppia di società i e j , con valore pari a ρ . Possiamo quindi scrivere che $a_i = \sqrt{\rho}$, per cui l'equazione che descrive le variabili U_i può essere riscritta come:

$$U_i = \sqrt{\rho}F + \sqrt{1 - \rho}Z_i$$

Possiamo adesso introdurre il *worst case default rate* $WCDR(T, X)$, ossia il "tasso di insolvenza relativo allo scenario peggiore" che, entro il tempo T e con probabilità X , non verrà sorpassato. Considerando quanto detto sopra, possiamo stimare tale tasso tramite la seguente equazione:

$$WCDR(T, X) = N \left\{ \frac{N^{-1}(PD) + \sqrt{\rho}N^{-1}(X)}{\sqrt{1 - \rho}} \right\}$$

Data la precedente equazione, notiamo che quando la *copula correlation* ρ è pari a zero, le insolvenze sui prestiti sono indipendenti l'una dall'altra e $WCDR(T, R) = \text{probability of default}$. Inoltre, un valore ρ più elevato comporterà anche un aumento del $WCDR(T, R)$.

Nel caso di un unico prestito, possiamo calcolare il percentile X della distribuzione delle perdite come:

$$WCDR(T, R) \times EAD \times LGD$$

Nel caso di un portafoglio di n prestiti, ciascuno dei quali è piccolo rispetto al totale, sfruttando il risultato ottenuto da Gordy⁴⁵ possiamo scrivere che il percentile X della distribuzione delle perdite è approssimabile a:

$$\sum_{i=1}^n WCDR_i(T, X) \times EAD_i \times LGD_i$$

dove $WCDR_i(T, X)$, EAD_i e LGD_i sono i valori $WCDR$, EAD e LGD relativi all' i -esimo prestito.

L'equazione appena riportata, con $T=1$ e $X=99,9\%$, indica il valore del capitale regolamentare per il *banking book*⁴⁶. È possibile che l'espressione indicata sotto il simbolo di sommatoria venga moltiplicata per il *maturity adjustment* MA_i relativo all' i -esimo prestito. Questa modifica viene integrata per tenere in considerazione del fatto che, se la durata del prestito è maggiore di un anno, la sua qualità creditizia potrebbe andare incontro a un peggioramento senza che questo conduca all'insolvenza.

⁴⁵ Si veda Gordy, M. B., "A Risk-factor Model Foundation for Ratings-Based Capital Rules", *Journal of Financial Intermediation* 12, no. 3 (luglio 2003): 199,233

⁴⁶ Il *banking book* (o portafoglio bancario) è una delle due categorie in cui rientrano le attività degli enti finanziari. Le attività non presenti nel *banking book* vengono riferite al *trading book* (o portafoglio di negoziazione). In particolare, il *banking book* consiste in un portafoglio di proprietà in cui sono detenute partecipazioni di natura strategica o verso controparti con le quali vi è una relazione di lungo periodo. Si veda anche https://www.bankpedia.org/index_voce.php?lingua=it&i_id=88&i_alias=b&c_id=18684-banking-book.

Nel *foundation IRB* (o “IRB di base”) la stima della PD viene effettuata dalla banca ma i valori di EAD, LGD e MA sono già fissati esternamente (secondo quanto previsto dagli accordi di Basilea). Nell’*advanced IRB* (o “IRB avanzato”), invece, le banche stimano anche gli altri parametri.

Per la stima della PD e di ρ si può utilizzare il metodo della massima verosimiglianza, sfruttando le serie storiche dei tassi d’insolvenza. Abbiamo che:

$$DR = N \left\{ \frac{N^{-1}(PD) + \sqrt{\rho}N^{-1}(G(DR))}{\sqrt{1-\rho}} \right\}$$

dove DR è il *default rate* e G(DR) è la funzione di distribuzione di DR. Dalla precedente equazione si può ottenere la seguente:

$$G(DR) = N \left\{ \frac{\sqrt{1-\rho}N^{-1}(DR) - N^{-1}(PD)}{\sqrt{\rho}} \right\}$$

Differenziando la funzione precedente otteniamo la funzione di densità del *default rate*:

$$g(DR) = \sqrt{\frac{1-\rho}{\rho} \exp \left\{ \frac{1}{2} \left[[N^{-1}(DR)]^2 - \left[\frac{\sqrt{1-\rho}N^{-1}(DR) - N^{-1}(PD)}{\sqrt{\rho}} \right]^2 \right] \right\}}$$

2.7.2 Credit Risk Plus

Alla fine degli anni '90, Credit Suisse propose una metodologia per il calcolo del VaR chiamata Credit Risk Plus⁴⁷.

Immaginiamo un’istituzione finanziaria con n prestiti di una determinata tipologia e una probabilità q di insolvenza a un anno per ciascun prestito. Il numero di insolvenze atteso sull’intero portafoglio è stimabile come $\mu = nq$. Supponendo che gli eventi di insolvenza siano indipendenti tra loro, la probabilità di verifica di m insolvenze contemporaneamente viene data dalla seguente distribuzione binomiale:

$$\frac{n!}{m!(n-m)!} q^m (1-q)^{n-m}$$

Tuttavia, nel caso in cui q sia un valore piccolo e n uno grande è noto che la distribuzione binomiale può essere approssimata dalla seguente distribuzione di Poisson:

$$\frac{e^{-\mu} \mu^m}{m!}$$

⁴⁷ Per un approfondimento si veda Credit Suisse Financial Products, “Credit Risk Management Framework”, ottobre 1997 e Resti A., “La gestione del rischio di credito con modelli di derivazione attuariale: il caso di CreditRisk+”, Working Papers, n. 4, Fondo Interbancario di Tutela dei Depositi, Roma, 2000.

La precedente espressione è utilizzabile anche qualora la probabilità di insolvenza non sia la stessa per tutti i prestiti purché tali probabilità siano piccole e finché q sia uguale alla probabilità di insolvenza media dei prestiti in portafoglio.

Nella realtà, tuttavia, è presente un certo grado di incertezza su quello che sarà il tasso di insolvenza durante l'anno successivo alla valutazione. Come possiamo notare dalla tabella 8, i tassi di insolvenza possono cambiare in modo significativo di anno in anno. Ipotizzando che il numero atteso delle insolvenze μ abbia una distribuzione gamma con media μ e deviazione standard σ , allora la distribuzione di Poisson sopra indicata si trasforma nella binomiale negativa e possiamo descrivere la probabilità che si verifichino m insolvenze come:

$$p^m(1-p)^\alpha \frac{\Gamma(m+\alpha)}{\Gamma(m+1)\Gamma(\alpha)}$$

dove:

1. $\alpha = \frac{\mu^2}{\sigma^2}$;
2. $p = \frac{\sigma^2}{(\mu + \sigma^2)}$;
3. $\Gamma(x)$ è la funzione gamma.

<i>Anno</i>	<i>Tasso d'insolvenza</i>	<i>Anno</i>	<i>Tasso d'insolvenza</i>	<i>Anno</i>	<i>Tasso d'insolvenza</i>
1970	2,631	1986	1,830	2002	2,924
1971	0,286	1987	1,423	2003	1,828
1972	0,453	1988	1,393	2004	0,834
1973	0,456	1989	2,226	2005	0,647
1974	0,275	1990	3,572	2006	0,593
1975	0,361	1991	2,803	2007	0,349
1976	0,176	1992	1,337	2008	2,507
1977	0,354	1993	0,889	2009	4,996
1978	0,354	1994	0,651	2010	1,232
1979	0,088	1995	0,899	2011	0,906
1980	0,344	1996	0,506	2012	1,230
1981	0,162	1997	0,616	2013	1,232
1982	1,040	1998	1,137	2014	0,939
1983	0,900	1999	2,213	2015	1,732

1984	0,869	2000	2,455	2016	2,149
1985	0,952	2001	3,679		

Tabella 8 Tassi di insolvenza annui. Fonte: Moody's

Invece, nella tabella 9 possiamo vedere la distribuzione probabilistica del numero delle insolvenze m in funzione della deviazione standard σ , posto che $\mu = 4$. Per il calcolo della colonna $\sigma = 0,0$ è stata utilizzata la distribuzione di Poisson: $\frac{e^{-\mu}\mu^m}{m!}$. Invece, le colonne successive tramite la binomiale negativa $p^m(1-p)^\alpha \frac{\Gamma(m+\alpha)}{\Gamma(m+1)\Gamma(\alpha)}$. Come notiamo dalla tabella, il valore restituito da quest'ultima formula tende alla distribuzione di Poisson quando σ tende a zero. Invece, al crescere di σ aumenta anche la probabilità di riscontrare un numero elevato di insolvenze.

Analizzando la tabella 9 e considerando che la perdita in caso di insolvenza sia costante, abbiamo che:

1. Se $\sigma = 0$ (per cui non vi è incertezza relativa al numero di insolvenze), il VaR al 99,9% è stimabile nella perdita conseguente a undici insolvenze.

Considerando invece il caso estremo opposto abbiamo che:

2. Se $\sigma = 10$ (quindi considerando un'elevata incertezza relativa al numero di insolvenze), il VaR al 99,9% è riconducibile alla perdita conseguente a 98 insolvenze.

N. di insolvenze (m)	Deviazione standard del numero delle insolvenze (σ)						
	0,0	0,1	0,5	1,0	2,0	5,0	10,0
0	0,0183	0,0184	0,0207	0,0281	0,0625	0,2814	0,5938
1	0,0733	0,0734	0,777	0,0901	0,1250	0,1553	0,00913
2	0,1465	0,1466	0,1486	0,1531	0,1563	0,1098	0,0509
3	0,1954	0,1952	0,1923	0,1837	0,1563	0,0833	0,0353
4	0,1954	0,1951	0,1895	0,1746	0,1367	0,0653	0,0268
5	0,1563	0,1561	0,1516	0,1396	0,1094	0,0523	0,0214
6	0,1042	0,1041	0,1026	0,0978	0,0820	0,0423	0,0177
7	0,595	0,0596	0,0603	0,0614	0,0586	0,0346	0,0150
8	0,298	0,0298	0,0315	0,0353	0,0403	0,0285	0,0129
9	0,132	0,0133	0,0148	0,0188	0,0296	0,0236	0,0113
10	0,053	0,0053	0,0064	0,0094	0,0175	0,0196	0,0099
11	0,019	0,0019	0,0025	0,0045	0,0111	0,0163	0,0088
12	0,006	0,0007	0,0009	0,0020	0,0069	0,0137	0,0079
13	0,0002	0,0002	0,0003	0,0009	0,0043	0,0115	0,0071
14	0,0001	0,0001	0,0001	0,0004	0,0026	0,0096	0,0064

N. di insolvenze (m)	Deviazione standard del numero delle insolvenze (σ)						
	0,0	0,1	0,5	1,0	2,0	5,0	10,0
15	0,0000	0,0000	0,0000	0,0001	0,0016	0,0081	0,0058
16	0,0000	0,0000	0,0000	0,0001	0,0009	0,0068	0,0053
17	0,0000	0,0000	0,0000	0,0000	0,0005	0,0058	0,0048
18	0,0000	0,0000	0,0000	0,0000	0,0003	0,0049	0,0044
19	0,0000	0,0000	0,0000	0,0000	0,0002	0,0041	0,0041
20	0,0000	0,0000	0,0000	0,0000	0,0001	0,0035	0,0038
Quantile 0,999	11	11	12	13	17	39	98

Tabella 9 Distribuzione probabilistica del numero di insolvenze per $\mu=4$. Fonte: J. C. Hull, "Risk Management e Istituzioni Finanziarie", 2018

Nella realtà, le banche interagiscono con varie categorie di controparti a cui sono legati diversi tassi di insolvenza. I passaggi da seguire possono essere sintetizzati come segue:

1. Estrazione del tasso di insolvenza complessivo (ad esempio, utilizzando una tabella simile alla tabella 8);
2. stima del tasso di insolvenza atteso per categoria di controparte;
3. estrazione del numero di insolvenze per categoria di controparte;
4. estrazione della LGD per ciascuna insolvenza;
5. calcolo della perdita complessiva;
6. calcolo della distribuzione delle perdite complessive (tramite la ripetizione dei passaggi precedenti);
7. calcolo del VaR sulla base della distribuzione delle perdite complessive.

Tuttavia, controllando attentamente i dati della tabella 8 possiamo notare che i tassi di insolvenza di anni consecutivi non sono completamente indipendenti rispetto a tassi rilevati l'anno precedente. Pertanto, estrarre casualmente un campione di dati da una tabella analoga alla tabella 8 potrebbe non essere la soluzione migliore. Potrebbe essere più opportuno stimare un modello autoregressivo sulla base dei dati storici relativi ai tassi di insolvenza oppure cercare una correlazione tra i dati e alcuni fattori macroeconomici osservati nell'anno precedente.

Possiamo in conclusione affermare che un ruolo cruciale è giocato dalla *default correlation*. Con un tasso di insolvenza costante, la *default correlation* tra le controparti sarebbe nulla e la probabilità di numerose insolvenze sarebbe estremamente bassa, come evidenziato nella tabella 9. Quando invece cresce l'incertezza sul tasso di insolvenza aumenta anche la *default correlation* e, di conseguenza, la probabilità di riscontrare numerose insolvenze. Inoltre, la *default correlation* è alta quando il tasso di insolvenza è uguale per tutte le

controparti, a prescindere dal fatto che quest'ultimo sia elevato o meno. Se la *default correlation* fosse nulla, la distribuzione probabilistica delle perdite sarebbe all'incirca simmetrica. Invece, nel caso di una *default correlation* diversa da zero, la distribuzione delle perdite per insolvenza presenta un'asimmetria positiva.

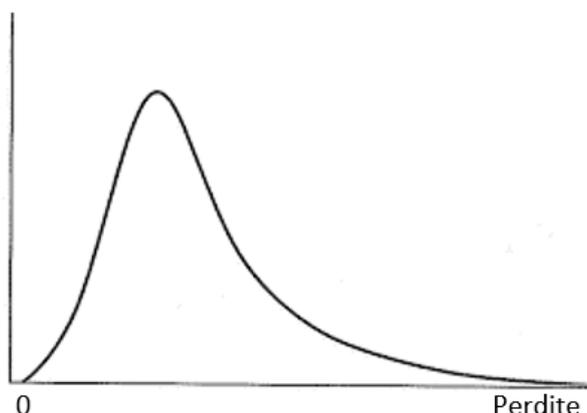


Figura 11 Distribuzione delle perdite per insolvenza. Asimmetria positiva (*default correlation* $\neq 0$). Fonte: Hull, J. C., "Risk Management e Financial Institution". John Wiley & Sons (2018).

2.7.3 Creditmetrics

Nei modelli analizzati nei paragrafi precedenti abbiamo preso in considerazione le perdite relative alle insolvenze ma abbiamo trascurato gli effetti legati al peggioramento della qualità creditizia dei prestiti⁴⁸.

Nel modello proposto da J. P. Morgan nel 1997, CreditMetrics, è stato invece espressamente previsto un metodo per tenere in considerazione sia le insolvenze sia i *downgrades*. Questo modello si basa, infatti, sulle matrici di transizioni dei *rating* (di cui un esempio è proposto nella tabella 10). I *rating* vengono forniti dalle agenzie di *rating* o calcolati internamente dalle banche stesse.

Per il calcolo del VaR creditizio a un anno relativo a un portafoglio di contratti contenenti numerose controparti si utilizza il metodo Monte Carlo. Per ciascuna simulazione si stima il livello di *rating* a un anno di ogni controparte, calcolando anche la perdita creditizia connessa. Qualora il *rating* della controparte non sia catalogato come "default", si stima il valore a un anno di tutti i contratti relativi a quella specifica controparte. Al contrario, se la controparte risulta essere insolvente si calcola il prodotto tra EAD e LGD. Generalmente, il *recovery rate* viene estratto da una distribuzione empirica.

Se i contratti con una controparte includono i derivati, il valore attuale delle perdite attese derivanti dall'insolvenza della controparte negli intervalli di tempo superiori all'anno è stimabile con la seguente espressione:

⁴⁸ È da notare, tuttavia, che il modello di Vasicek permette di aggiungere i *maturity adjustments* per tener conto del possibile deterioramento del merito creditizio.

$$\sum_{i=j}^n (1 - R)q_i v_i$$

dove:

1. il j -esimo intervallo è da considerarsi tra un anno;
2. q_i è la probabilità neutrale di fallimento della controparte nell'intervallo i -esimo;
3. v_i è il valore attuale dell'esposizione netta nei confronti della controparte;
4. R è il *recovery rate* che la banca stima nel caso di insolvenza della controparte e in assenza di garanzie accessorie.

Se indichiamo con q_i^* la probabilità di insolvenza nell' i -esimo intervallo di una specifica simulazione (per $i \geq j$), possiamo calcolare il valore attuale della perdita creditizia della medesima simulazione tramite la seguente formula:

$$\sum_{i=j}^n (1 - R)(q_i^* - q_i)v_i$$

Possiamo osservare che se il *rating* della controparte durante l'anno va incontro a un miglioramento allora è probabile ottenere una perdita creditizia negativa.

Se in una determinata simulazione la controparte va in *default* durante il primo anno allora si determina l'esposizione alla data dell'insolvenza e la si moltiplica per la LGD per ottenere la perdita creditizia.

Quando abbiamo introdotto il tema delle simulazioni con il metodo Monte Carlo non abbiamo supposto che le variazioni dei *ratings* delle controparti fossero indipendenti tra di loro. Mantenendo inalterata la situazione di partenza, possiamo stimare la distribuzione congiunta delle variazioni dei *rating* utilizzando il modello della copula gaussiana.

Infine, è interessante notare che basando Credit Risk Plus e CreditMetrics sulle stesse ipotesi entrambi forniscono la medesima distribuzione delle perdite nel lungo periodo pur differendo nella tempistica di tali perdite. A tal proposito, immaginiamo un prestito con scadenza a tre anni. Durante il primo anno il *rating* del debitore passa da A a BBB, durante il secondo da BBB a B e, infine, nel terzo si verifica il *default*. Secondo il Credit Risk Plus non si verificano perdite nei primi due anni ma solo al verificarsi dell'insolvenza, pertanto al terzo anno. Invece, secondo l'approccio di CreditMetrics si rilevano delle perdite in ognuno dei tre anni. La somma delle perdite stimate secondo quest'ultimo modello dovrebbe appunto restituire l'importo calcolato con il primo.

In conclusione, si riporta un esempio di tabella di transizione.

A	$x_A < -3,2771$	Aaa	B	$x_B < -3,6903$	Aaa
	$-3,2771 \leq x_A < -1,9378$	Aa		$-3,6903 \leq x_B < -3,3213$	Aa
	$-1,9378 \leq x_A < 1,5248$	A		$-3,3213 \leq x_B < -2,8757$	A
	$1,5248 \leq x_A < 2,4401$	Baa		$-2,8757 \leq x_B < -2,4545$	Baa
	$2,4401 \leq x_A < -2,8328$	Ba		$-2,4545 \leq x_B < -1,5499$	Ba
	$2,8328 \leq x_A < 3,0476$	B		$-1,5499 \leq x_B < 1,1929$	B
	$3,0476 \leq x_A < 3,1809$	Caa		$1,1929 \leq x_B < 1,7243$	Caa
	$3,1809 \leq x_A < 3,2384$	Ca-C		$1,7243 \leq x_B < 1,7928$	Ca-C
	$x_A \leq 3,2384$	Default		$x_B \leq 7928$	Default

Tabella 10 CreditMetrics: esempio di tabella di transizione. Fonte: J. C. Hull, "Risk Management e Istituzioni Finanziarie", 2018

CAPITOLO 3. Analisi della stima della *Probability of Default*

Il lavoro di analisi si pone come obiettivo quello di costruire un modello in grado di stimare la probabilità di *default* delle aziende partendo da dati di bilancio ritenuti significativi. Inoltre, consapevoli dell'impatto avuto dalla pandemia sull'economia globale, vogliamo testare se il modello continui ad avere una capacità previsionale elevata oppure se, al contrario, il nuovo contesto economico possa influenzare la significatività dei regressori selezionati. Tutti i dati dei campioni sono stati raccolti utilizzando il *database Refinitiv*.

Nel primo paragrafo presentiamo più nel dettaglio il funzionamento del modello logit, spiegandone anche le logiche di scelta del campione nonché di selezione dei parametri e della loro convalida. Nei successivi viene applicato il modello e sono quindi esposti i risultati dell'analisi.

3.1 Caratteristiche del modello e presentazione dei campioni utilizzati

Il modello logit viene utilizzato per analizzare la relazione tra una variabile indipendente y (di natura qualitativa) e n variabili x che possono essere di natura sia qualitativa sia quantitativa. Relativamente alla variabile y è bene notare che le modalità che può assumere sono mutuamente esclusive. Questo modello presenta diverse occasioni di applicazione come la scelta di acquisto di un prodotto (acquisto/non acquisto), il grado di soddisfazione di un cliente (soddisfatto/non soddisfatto) o, come nel caso in questione, il livello di salute di un'azienda (sana/in crisi oppure, relativamente all'evento *default*, verificato/non verificato).

Gli obiettivi per cui si implementano tali tipi di modelli possono essere diversi, in particolare:

1. individuare quale tra le variabili indipendenti x sia maggiormente in grado di influenzare la variabile dipendente y . Parliamo di "fattori di protezione" o di "fattori di rischio" a seconda del tipo di correlazione (positiva o negativa);
2. ottenere la combinazione lineare delle variabili indipendenti x che riesca meglio a discriminare il gruppo delle unità che possiedono o meno l'attributo;
3. ottenere una stima della probabilità che una nuova unità statistica possieda l'attributo.

Le tipologie di modelli logit applicabili possono essere sintetizzate come segue: modello logit per variabili binarie (dicotomiche), modello logit multinomiale (applicabile quando la variabile dipendente y ha più di due categorie) e, infine, modello logit ordinale. Quest'ultimo rappresenta un'estensione del modello logit semplice e si applica quando non siamo in presenza di una variabile indipendente y dicotomica ed essa può assumere valori appartenenti ad almeno due categorie di risultati. La variabile dipendente deve allora presentare due caratteristiche: deve essere "categorica" (ossia deve poter assumere valori compresi in almeno due classi di valori) e "ordinata" (ossia deve esistere un ordine nella classificazione delle categorie).

Nel resto del capitolo ci concentreremo sul modello logit per variabili dipendenti dicotomiche. In particolare, la variabile dipendente y potrà assumere solo due valori che indicano la presenza o meno di un attributo, ossia il verificarsi dell'evento *default*. Abbiamo allora che:

$$y = \begin{cases} 1, & \text{se si è verificato l'evento default} \\ 0, & \text{altrimenti} \end{cases}$$

Inoltre, y_i è la realizzazione di una variabile casuale Y_i che può assumere solo valori 0 e 1, con probabilità π_i e $1 - \pi_i$. La distribuzione di Y_i è nota come distribuzione di Bernoulli e può essere scritta come segue:

$$P(Y_i = y_i) = \pi_i^{y_i}(1 - \pi_i)^{1-y_i}$$

Possiamo quindi dire che per $y_i = 0,1$:

1. se $y_i = 1$: $P(Y_i = 1) = \pi_i$;
2. se $y_i = 0$: $P(Y_i = 0) = 1 - \pi_i$.

Relativamente alla variabile Y_i , si può facilmente verificare che il valore atteso $E(Y_i)$ e la varianza $V(Y_i)$ sono dati da:

1. $E(Y_i) = \pi_i$;
2. $V(Y_i) = \pi_i(1 - \pi_i)$.

Sia la media $E(Y_i)$ sia la varianza $V(Y_i)$ dipendono dalla probabilità π_i , pertanto ciascun fattore che abbia un impatto sulla probabilità di successo altera anche tali valori. Da qui possiamo affermare che ciascun modello lineare che assume che le variabili indipendenti influenzino la media senza modificare la varianza risulta inadatto allo studio di dati di tipo binario.

Nel modello che vogliamo costruire, le probabilità π_i dipendono da un vettore di covariate x_i . Possiamo allora immaginare un'equazione in cui la probabilità π_i dipenda linearmente dalle covariate, come indicato di seguito:

$$\pi_i = x_i' \beta$$

dove x_i' indica il vettore dei regressori e β indica il vettore di coefficienti che deve essere stimato successivamente.

Quando utilizziamo questo modello, tuttavia, notiamo l'esistenza di una problematica. Infatti, π_i , essendo una probabilità, può assumere valori tra 0 e 1. Invece, $x_i' \beta$ potrà assumere qualsiasi valore sull'asse reale. Per aggirare questo problema è possibile utilizzare il logaritmo dell'*odds* al posto della probabilità π_i . In

particolare, l'*odds* può essere definito come il rapporto tra probabilità di successo e il suo complemento a 1.

Possiamo scrivere che:

$$odds_i = \frac{\pi_i}{1 - \pi_i} = \frac{P(Y_i = 1)}{1 - P(Y_i = 1)} = \frac{P(Y_i = 1)}{P(Y_i = 0)}$$

Ad esempio, l'*odds* varrà 1 quando i due eventi rapportati sono equiprobabili, varrà 3/2 se $\pi_i = 3/5$.

Possiamo adesso considerare il logaritmo dell'*odds*, $\eta_i = \text{logit}(\pi_i) = \log \frac{\pi_i}{1 - \pi_i}$.

Esso permette di rimuovere le limitazioni relative ai valori assunti dalla probabilità. Infatti, per $\pi_i \rightarrow 0$, $odds \rightarrow 0$ e $\text{logit}(\pi_i) \rightarrow -\infty$; invece, per $\pi_i \rightarrow 1$, $odds \rightarrow +\infty$ e $\text{logit}(\pi_i) \rightarrow +\infty$.

Il modello logit che utilizziamo per l'analisi è allora definito assumendo che il logit della probabilità sia una funzione lineare delle variabili x_i , per cui possiamo scrivere che:

$$\text{logit}(\pi_i) = x' \beta$$

dove x_i è il vettore delle covariate e β è il vettore dei coefficienti, entrambi di dimensioni $h \times 1$. Il cambiamento dei coefficienti va interpretato allo stesso modo dell'analogo modello lineare ricordando che in questo caso la variazione del valore di ciascun coefficiente non impatterà direttamente la probabilità ma il logit. Per la stima dei parametri viene utilizzato il metodo della massima verosimiglianza. La soluzione ottimale può essere ottenuta iniziando con dei parametri di prova arbitrari che vengono successivamente modificati tramite software statistici con un processo iterativo. I tentativi andranno avanti fino a quando la funzione avrà dei margini di miglioramento. Quando tale miglioramento è infinitesimale otteniamo la migliore approssimazione possibile dei parametri.

Esplicitando l'equazione precedente per la probabilità π_i , possiamo riscrivere il modello come segue:

$$\pi_i = P(Y_i = 1|x_i) = \frac{\exp(x'_i \beta)}{1 + \exp(x'_i \beta)}$$

Relativamente all'interpretazione dell'*odds ratio* (OR), supponiamo di avere due variabili indipendenti x_1 e x_2 , con x_2 dicotomica. Possiamo scrivere che:

$$odds(x_2 = 1) = \frac{P(y = 1|x_1, x_2 = 1)}{1 - P(y = 1|x_1, x_2 = 1)} = \exp(\beta_0 + \beta_1 x_1 + \beta_2)$$

e

$$odds(x_2 = 0) = \frac{P(y = 1|x_1, x_2 = 0)}{1 - P(y = 1|x_1, x_2 = 0)} = \exp(\beta_0 + \beta_1 x_1)$$

quindi:

$$OR = \frac{odds(x_2 = 1)}{odds(x_2 = 0)} = \exp(\beta_2)$$

Se, ad esempio, $\exp(\beta_2) = 3$, possiamo interpretare questo valore come segue: le osservazioni caratterizzate da $x_2 = 1$ hanno il triplo della possibilità di restituire il valore $Y = 1$ rispetto a quelle con $x_2 = 0$ ⁴⁹. Se invece x_1 è una variabile continua la variazione dell'OR in corrispondenza di un incremento unitario di x_1 è analogo al caso visto in precedenza per variabile dicotomica. Si avrà allora che $OR = \exp(\beta_1)$. Se, invece, consideriamo un incremento di n unità, con $n \neq 1$, allora il logaritmo dell'OR sarà uguale a $\exp(n\beta_1)$.

3.1.1 Significatività del modello

Una volta strutturato il modello è necessario verificare quanto le stime da esso prodotte si avvicinino alle osservazioni effettuate. Il nostro obiettivo è chiaramente individuare un modello che minimizzi le differenze tra i valori stimati e quelli osservati.

Esistono diversi metodi per valutare la significatività di un modello. Tra i più noti ricordiamo:

1. Valutazione tramite il confronto con il modello saturo. Quest'ultimo è un modello che contiene un parametro per ciascuna osservazione, garantendo il massimo adattamento possibile. Introduciamo la grandezza D (*deviance*), definita come segue:

$$D = -2 \log \left\{ \frac{L_k}{L_{saturo}} \right\}$$

Dove L indica il livello di verosimiglianza del modello. Notiamo che:

- se $L_k < L_{saturo} \Rightarrow \frac{L_k}{L_{saturo}} < 1 \Rightarrow -2 \log \left(\frac{L_k}{L_{saturo}} \right) > 0$
- altrimenti $L_k \rightarrow L_{saturo} \Rightarrow D \rightarrow 0$

Quando il valore di L_k è molto vicino a quello di L_{saturo} il valore di D è basso e l'adattamento è buono. Tuttavia, questo può indicare l'utilizzo di un numero eccessivo di variabili che contravviene al principio di parsimonia;

2. Valutazione tramite confronto tra un modello completo A e un modello ridotto B.

Il modello completo presenta $k+r$ variabili indipendenti mentre il modello ridotto utilizza solo il sottoinsieme k . L'ipotesi che sottoponiamo a verifica è la seguente:

$$H_0: \beta_{k+1}, \beta_{k+2}, \dots, \beta_{k+r} = 0$$

⁴⁹ Ricordiamo che l'*odds* rappresenta la propensione al successo e l'OR è, in questo caso, il rapporto tra gli *odds* dei due valori che x_2 può assumere.

H_1 : almeno un coefficiente è diverso da zero

In questo modo abbiamo la possibilità di capire se almeno una delle r variabili aggiuntive del modello α sia significativa. Per effettuare questo confronto introduciamo la variabile G definita tramite la relazione seguente:

$$G = D_1(\text{modello } B) - D_2(\text{modello } A)$$

Questa variabile, nel caso dell'ipotesi H_0 , si distribuisce come un χ^2 con r gradi di libertà. Dato un valore α , possiamo individuare un valore di soglia $\chi_{\alpha,r}^2$ che permette di individuare una regione di rifiuto definita come $R = \{G > \chi_{\alpha,r}^2\}$.

Se l'ipotesi nulla viene rifiutata possiamo affermare che almeno uno dei parametri aggiuntivi del modello B è diverso da zero e significativo;

3. Valutazione tramite un parametro B_j usando la statistica di Wald. Consideriamo la seguente verifica di ipotesi:

$$H_0: B_j = 0$$

$$H_1: B_j \neq 0$$

Utilizzando la statistica test seguente:

$$W = \frac{B_j}{\hat{\sigma}(\hat{B}_j)}$$

dove $\hat{\sigma}(\hat{B}_j)$ è l'errore standard. Sotto l'ipotesi nulla la statistica di Wald si distribuisce come una normale standard. Dato un livello di significatività α , abbiamo che:

$$\text{se } |W_{obs}| < z_{\alpha/2} \Rightarrow \text{rifiuto } H_0;$$

4. Valutazione tramite lo Pseudo- R^2 di McFadden.

Tale valore è compreso tra 0 e 1 dove zero indica un modello con nessuna capacità predittiva mentre 1 indica la massima capacità di previsione possibile. Indichiamo con L_0 la verosimiglianza di un modello con la sola intercetta. Analogamente, con L_k indichiamo la verosimiglianza di un modello con k parametri. Possiamo calcolare la statistica Pseudo- R^2 come segue:

$$\text{Pseudo} - R^2 = 1 - \frac{\log L_k}{\log L_0}$$

Se tutte le stime dei coefficienti sono uguali a zero il valore dello Pseudo-R² sarà pari a zero. Se, invece, le previsioni del modello coincidono perfettamente con i valori osservati il valore restituito dallo Pseudo-R² sarà 1. Quest'ultimo è un valore teorico in quanto nella realtà non otteniamo modelli che si adatto così tanto al campione di osservazioni utilizzato;

5. Valutazione tramite una tabella di confronto tra i dati previsti e quelli osservati. Sommando il numero dei valori presenti nella diagonale principale $\eta_{00} + \eta_{11}$ possiamo ottenere il numero di casi correttamente previsti. Al contrario, la somma dei valori presenti nella diagonale secondaria $\eta_{01} + \eta_{10}$ indica il numero di casi non previsti correttamente dal modello. Il rapporto tra il numero di casi previsti correttamente e il numero di casi totale potrà fornire una semplice stima di accuratezza del modello.

		Previsto	
		0	1
Effettivo	0	η_{00}	η_{01}
	1	η_{10}	η_{11}

Tabella 11 Tabella di confronto tra risultati effettivi e previsti. Fonte: propria elaborazione.

6. Selezione delle variabili più utili tramite l'utilizzo del *P-value*. Tramite tale valore è possibile stimare l'influenza di ciascuna variabile indipendente x_i sulla variabile dipendente y_i . Il *P-value* può essere utilizzato per verificare che il coefficiente di una variabile sia diverso da zero. Per rigettare l'ipotesi, il *P-value* deve essere minore di un dato valore di α scelto. Nel caso in cui questo si verifichi, possiamo supporre che la variabile sia significativa all'interno della nostra analisi. Tramite questo criterio possiamo modificare il modello e selezionare le variabili che più abbiano un impatto sulla regressione.

3.1.2 Presentazione dei campioni costruiti

Presentiamo adesso i campioni raccolti che saranno analizzati nei prossimi paragrafi. Per la costruzione dei campioni si è cercato di raccogliere dati sufficientemente numerosi in modo tale, da un lato, di fare in modo che il modello risulti il più generalizzabile possibile e, dall'altro, di ricreare le condizioni di mercato. Per realizzare quest'ultimo punto è stato deciso di includere nel campione aziende appartenente a settori differenti. Più il campione è ampio e i settori diversi, più è verosimile supporre che esso si comporterà in maniera simile al mercato.

All'interno della nostra analisi, consideriamo la variabile dipendente dicotomica Y che assume valore 1 se si verifica l'evento *default*, 0 altrimenti. Per analisi su variabili dipendenti dicotomiche uno dei modelli più

utilizzati è il modello logit, approfondito nel paragrafo 3.1. La scelta ricade proprio su tale modello in quanto capace di coniugare ottimi risultati di stima e semplicità dei calcoli.

Il primo campione illustrato è quello statunitense, presentato nella tabella seguente, formato da 58 aziende di cui 29 presentano $Y=1^{50}$, rappresentanti quindi la metà dello stesso.

Nome Società	Y	Nome società	Y
AAON, INC.'	0	BRIGGS & STRATTON CORP	1
APPLE INC.	0	ASCENA RETAIL GROUP INC	1
ABCELLERA BIOLOGICS INC.	0	AKORN INC	1
ABIOMED, INC.	0	ENVISION HEALTHCARE CORP (DELAWARE)	1
ACADIA HEALTHCARE COMPANY, INC.	0	CEC ENTERTAINMENT INC	1
AXCELIS TECHNOLOGIES, INC.	0	COMMERCIAL BARGE LINE CO	1
ACM RESEARCH, INC.	0	COVIA HOLDINGS LLC	1
ALLEGRO MICROSYSTEMS, INC	0	DIAMOND OFFSHORE DRILLING INC	1
ALTAIR ENGINEERING INC.	0	FERRELLGAS PARTNERS LP	1
AMC NETWORKS INC.	0	EXTRACTION OIL & GAS INC	1
ADVANCED MICRO DEVICES, INC.	0	EQUINOX HOLDINGS INC	1
APOLLO MEDICAL HOLDINGS, INC.	0	ENVISION HEALTHCARE CORP	1
ALLIED MOTION TECHNOLOGIES INC.	0	FRONTIER COMMUNICATIONS PARENT INC	1
AMPHASTAR PHARMACEUTICALS INC	0	FORUM ENERGY TECHNOLOGIES INC	1
A-MARK PRECIOUS METALS, INC.	0	GUITAR CENTER HOLDINGS INC	1
THE ANDERSONS, INC.	0	HI-CRUSH INC	1
AMERICAN OUTDOOR BRANDS, INC.	0	JASON INDUSTRIES INC	1
AMERICAN PUBLIC EDUCATION, INC.	0	LIBBEY INC	1
APPFOLIO, INC.	0	NEIMAN MARCUS GROUP LTD LLC	1
ASIA PACIFIC WIRE & CABLE CORPORATION LIMITED	0	LSC COMMUNICATIONS INC	1
AERSALE CORPORATION	0	PIONEER ENERGY SERVICES CORP	1
ACADEMY SPORTS & OUTDOORS INC	0	INTERNAP HOLDING LLC	1
ASTEC INDUSTRIES, INC.	0	PIER 1 IMPORTS INC	1
ATRION CORPORATION	0	REVLON INC	1
AVNET, INC.	0	RENTPATH LLC	1
AXT, INC.	0	QUORUM HEALTH CORP	1
BARRETT BUSINESS SERVICES, INC.	0	WHITING PETROLEUM CORP	1
BIODELIVERY SCIENCES INTERNATIONAL, INC.	0	ULTRA PETROLEUM CORP	1
BLUE HAT INTERACTIVE ENTERTAINMENT TECHNOLOGY	0	SM ENERGY CO	1

Tabella 12 Presentazione del campione statunitense. Fonte: Database Refinitiv.

Di seguito si riporta il campione di verifica statunitense, formato da 52 società di cui 22 presentano $Y=1$.

Nome società	Y	Nome società	Y
AAON, INC.'	0	17 EDUCATION & TECHNOLOGY GROUP INC.	1
ABCAM PLC	0	BBQ HOLDINGS, INC.	1
ABIOMED, INC.	0	BOQII HOLDING LTD	1
ABSOLUTE SOFTWARE CORPORATION	0	CANADIAN NATURAL RESOURCES LIMITED	1

⁵⁰ In questo campione, consideriamo con $Y=1$ le aziende in cui si sia già verificato il *default* al 31.12.2019 oppure che abbiano avuto procedure fallimentari aperte in corso alla stessa data.

Nome società	Y	Nome società	Y
ACTIVISION BLIZZARD, INC.	0	CANADIAN SOLAR INC.	1
ADOBE INC.	0	DAQO NEW ENERGY CORP.	1
ADVANCED EMISSIONS SOLUTIONS, INC.	0	DRIL-QUIP, INC.	1
AESTHETIC MEDICAL INTERNATIONAL HOLDINGS GROUP LIMITED	0	ELITE EDUCATION GROUP INTERNATIONAL LTD	1
AKAMAI TECHNOLOGIES, INC.	0	FREQUENCY THERAPEUTICS, INC.	1
ALARM.COM HOLDINGS, INC.	0	GLBL X AI & TECHNOLOGY	1
ALIGN TECHNOLOGY, INC.	0	GREENPRO CAPITAL CORP.	1
ALLEGRO MICROSYSTEMS, INC	0	HP INC.	1
ALLIED MOTION TECHNOLOGIES INC.	0	IF BANCORP, INC.	1
AMC NETWORKS INC.	0	IQIYI, INC.	1
AMEDISYS, INC.	0	KELSO TECHNOLOGIES INC.	1
AMERICAN WOODMARK CORPORATION	0	LKQ CORPORATION	1
AMGEN INC.	0	MOVING IMAGE TECHNOLOGIES INC	1
AMPHASTAR PHARMACEUTICALS INC	0	NASDAQ, INC.	1
ANSYS, INC.	0	PETIQ, INC.	1
APPLE INC.	0	SQUARE, INC.	1
APPLIED MATERIALS, INC.	0	TOWNSQUARE MEDIA, INC.	1
APPLOVIN CORPORATION	0	TURQUOISE HILL RESOURCES LTD.	1
ARENA PHARMACEUTICALS, INC.	0		
ARRAY TECHNOLOGIES, INC.	0		
ASML HOLDING NV	0		
ASPEN TECHNOLOGY, INC.	0		
ATRION CORPORATION	0		
AUTOSCOPE TECHNOLOGIES CORPORATION	0		
BALCHEM CORPORATION	0		
MISSION PRODUCE, INC.	0		

Tabella 13 Presentazione del campione di verifica statunitense. Fonte: Database Refinitiv.

Infine, si presenta il campione europeo, formato da 61 aziende di cui 22 presentano Y=1.

Nome società	Y	Nome società	Y
AMPLIFON	0	CHL	1
ASKOLL EVA	0	CINEWORLD GROUP PLC	1
ATLANTIA	0	CODERE SA	1
BHB BRAU HOLDING BAYERN MITTE AG	0	DEOLEO SA	1
BIALETTI	0	DISTRIBUIDORA INTERNACIONAL DE ALIMENTACION SA	1
BIOMASS ENERGY PROJECT SA	0	ENERGY LAB	1
CNH INDUSTRIAL	0	EUROPCAR MOBILITY GROUP	1
CRANSWICK PLC	0	HAYA REAL ESTATE SAU	1
EMMI AG	0	INTELSAT SA	1
ENEL	0	LECTA SA	1
ENI	0	MALLINCKRODT PLC	1
EPRICE	0	MOBI SPA	1
FERRARI	0	NOBLE HOLDING CORPORATION PLC	1
FROSTA AG	0	NOSTRUM OIL & GAS PLC	1
GEL SPA	0	PACIFIC DRILLING SA	1
GENERAL ENERGY PLC	0	PGS ASA	1

GIGLIO GROUP	0	PIZZAEXPRESS FINANCING 1 PLC	1
KI GROUP	0	SAS AB	1
KRI KRI MILK INDUSTRY SA	0	TECHNICOLOR SA	1
LOTUS BAKERIES NV	0	TRANSOCEAN LTD	1
NATIONAL MILK RECORDS PLC	0	VESTEL ELEKTRONIK SANAYI VE TICARET AS	1
NEW NORDIC HEALTHBRANDS AB	0	ZELLIS HOLDINGS LTD	1
NWF GROUP PLC	0		
PHAROS ENERGY PLC	0		
PURCARI WINERIES PCL	0		
RAK PETROLEUM PLC	0		
SIRIO	0		
STMICROELECTRONICS	0		
STOCK SPIRITS GROUP PLC	0		
SUBSEA 7 SA	0		
TATE & LYLE PLC	0		
TECHNIPFMC PLC	0		
TEQNION AB	0		
TERNA	0		
TIM	0		
TREATT PLC	0		
VIMI FASTENERS	0		
VISIBILIA	0		
WYNNSTAY GROUP PLC	0		

Tabella 14 Presentazione del campione europeo. Fonte: Database Refinitiv.

3.2 Presentazione dei regressori e analisi del campione statunitense

Relativamente ai regressori, è necessario individuare delle variabili X_k che forniscano informazioni rilevanti per tutte le imprese e che siano ovviamente correlate con la probabilità di *default*. Generalmente, si opta per la scelta di alcuni indici di bilancio. Questi possono essere divisi in:

- indici di redditività: hanno come obiettivo generale la valutazione della capacità dell'impresa di generare un flusso di ricavi sufficiente alla copertura dei costi e alla produzione di utili per la crescita futura dell'azienda;
- indici di solidità: hanno come scopo quello di misurare l'equilibrio finanziario di medio e lungo periodo;
- indici di liquidità: hanno come scopo quello di valutare la capacità dell'impresa di far fronte alle obbligazioni con scadenza nel breve periodo tramite risorse che saranno liquidate nello stesso orizzonte temporale.

Nella presente analisi, la scelta degli indici è ricaduta sui seguenti:

1. X_1 : *quick ratio*⁵¹.

Questo indice segnala il livello di liquidità dell'azienda in relazione alle sue passività correnti, indicando quindi se l'azienda sarà in grado di far fronte ai debiti con scadenza breve. Il *quick ratio* considera solo la parte più "liquida" delle attività a breve, volutamente escludendo il magazzino che, pur essendo inserito in bilancio tra le attività correnti, non sempre può essere considerato fonte attendibile di liquidità nel breve periodo, soprattutto al valore inserito a bilancio. Aziende sane presentano spesso un valore di 1 o poco superiore. Un indice minore rispetto a tale soglia può indicare aziende non in grado di ripagare i debiti a breve mentre, al contrario, un valore troppo elevato può indicare eccessiva liquidità non investita, quindi carenza di efficienza nella gestione delle risorse. Il valore ottimale può comunque variare in relazione al settore di appartenenza;

2. X_2 : EBIT/Totale Attivo

Tale indice evidenzia quanto l'azienda sia in grado di sfruttare le sue risorse per la generazione dell'EBIT. La scelta dell'EBIT al posto dell'utile netto è riconducibile alla volontà di non influenzare tale valore dagli aspetti fiscali e finanziari, in modo tale da poter valutare efficacemente la sola gestione operativa;

3. X_3 : Totale Attivo a lungo/Totale Attivo

Tale valore indica quando l'azienda sia rigida. Tanto più questo valore è alto tanto più meno l'azienda sarà reattiva nella risposta a stimoli esterni. Tuttavia, non è sempre vero che una azienda più elastica risponda sempre bene a stimoli esterni. Se l'attivo circolante è costituito in prevalenza da crediti verso clienti e magazzino si corre il rischio di non poter liquidare nel breve periodo il valore del circolante;

4. X_4 : EBIT / Ricavi

Questo valore, detto anche *EBIT margin*, indica la percentuale di fatturato che si è trasformata in EBIT, indicando l'efficienza nella gestione operativa. Anche in questo caso, come nella variabile X_2 , si è voluto evitare l'impatto della componente finanziaria e di quella fiscale;

5. X_5 : Ricavi / Totale Attivo⁵²

Questo valore indica quanto l'azienda riesca a sfruttare i propri impieghi per generare ricavi. In altri termini, l'indice evidenzia il ricavo medio per unità di impieghi.

⁵¹ Conosciuto anche come *acid test*, viene calcolato come segue $QR = \frac{\text{Totale Attivo a breve} - \text{Magazzino}}{\text{Totale Passivo a breve}}$

⁵² Ossia l'indice di rotazione del capitale investito, una delle due componenti del ROI insieme al ROS.

Occorre ricordare che nella costruzione del modello è bene non aumentare inutilmente il numero dei regressori. Il modello sarà in questo modo più facile da generalizzare. Infatti, con un numero eccessivo di variabili aumentano anche gli errori standard delle loro stime e, soprattutto, si rischia di incorrere in problemi di *overfitting*. In altri termini, vogliamo evitare che il modello si adatti eccessivamente bene alle caratteristiche specifiche *training set*, diventando non utilizzabile a livello generale.

Un altro aspetto preliminarmente valutato all'interno dell'analisi è la cosiddetta multicollinearità. Essa si verifica quando dei regressori risultano essere linearmente correlati tra di loro. Tale dipendenza può pregiudicare la rilevanza statistica dei coefficienti stimati. Per tale motivo è stato eliminato un sesto regressore, $\frac{Cash\ Flow^{53}}{Totale\ attivo}$, poiché appare fortemente correlato con X_2 , come mostrato nel grafico in basso. Tra i due è stato preferito, come anticipato, X_2 poiché l'EBIT è stato ritenuto maggiormente significativo in quanto evidenzia l'andamento della sola gestione operativa, prima dell'influenza della gestione finanziaria e della componente fiscale.

Nome società	EBIT / Totale attivo	Cash Flow Totale Attivo
Amplifon	0,07	0,10
Askoll Eva	-0,26	-0,30
Atlantia	0,02	0,06
BHB Brauholding Bayern Mitte AG	0,03	0,13
Bialetti	0,00	0,20
Biomass Energy Project SA	0,02	0,02
CNH Industrial	0,04	0,05
Cranswick PLC	0,20	0,22
Emmi AG	0,09	0,13
ENEL	0,07	0,03
ENI	0,07	0,06
ePrice	-0,52	-0,69
Ferrari	0,17	0,18
FRoSTA AG	0,06	0,09
Gel SpA	0,03	0,06
General Energy PLC	0,08	0,09
Giglio Group	0,02	-0,06
Ki Group	-0,23	-0,16
Kri Kri Milk Industry SA	0,17	0,18
Lotus Bakeries NV	0,13	0,12
National Milk Records PLC	0,22	0,24
New Nordic Healthbrands AB	0,18	0,14
NWF Group PLC	0,07	0,08
Pharos Energy PLC	0,04	-0,03
Purcari Wineries PCL	0,16	0,14

⁵³ Dove il Cash Flow viene calcolato come *Utile netto + DA*

Nome società	EBIT / Totale attivo	Cash Flow Totale Attivo
RAK Petroleum PLC	0,06	0,11
Sirio	-0,01	0,10
STMicroelectronics	0,10	0,13
Stock Spirits Group PLC	0,08	0,06
Subsea 7 SA	0,03	0,07
Tate & Lyle PLC	0,17	0,17
TechnipFMC PLC	0,05	-0,08
Teqnon AB	0,09	0,11
Terna	0,06	0,07
TIM	0,05	0,09
Treant PLC	0,10	0,09
Vimi Fasteners	0,00	0,08
Visibilia	-0,14	-0,12
Wynnstay Group PLC	0,16	0,22
CHL	-0,11	-0,09
Cineworld Group PLC	0,06	0,04
Codere sa	0,05	0,08
Deoleo sa	0,02	0,00
Distribuidora Internacional de Alimentacion SA	-0,14	-0,08
Energy Lab	-0,19	-0,54
Europcar mobility group	0,04	0,02
Haya Real Estate SAU	-0,05	0,12
Intelsat SA	0,07	-0,02
Lecta SA	0,03	0,01
Mallinckrodt PLC	0,02	-0,01
Mobi Spa	-0,27	-0,04
Noble Holding Corporation PLC	-0,03	-0,05
Pacific Drilling SA	-0,10	-0,16
PGS ASA	0,03	0,20
Pizzaexpress Financing 1 PLC	0,03	-0,02
SAS AB	0,04	0,07
Technicolor SA	0,00	0,05
Transocean Ltd	0,00	-0,05
Vestel Elektronik Sanayi ve Ticaret AS	0,08	0,05
Zellis Holdings Ltd	-0,03	-0,02

Tabella 15 Dati per calcolo correlazione lineare tra regressori. Fonte: propria elaborazione

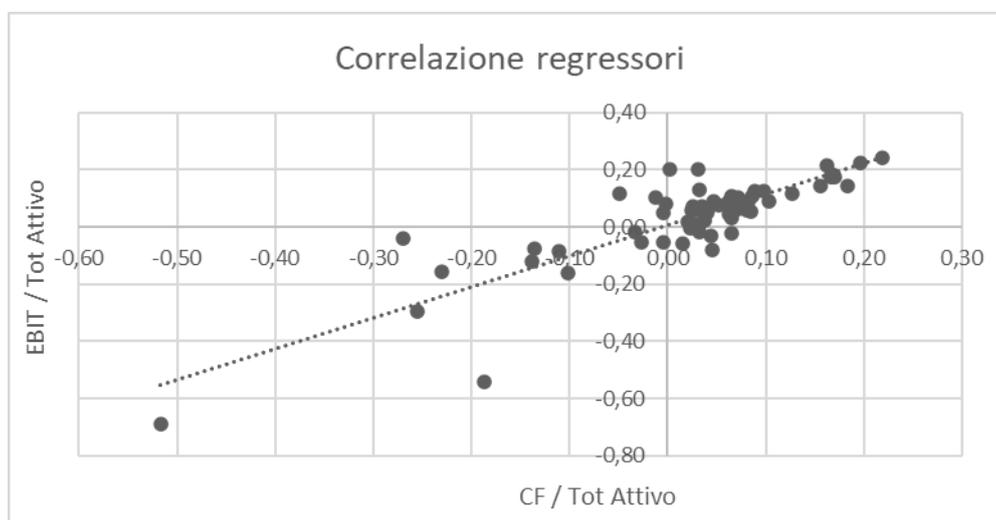


Figura 12 Rappresentazione grafica della correlazione lineare tra regressori. Fonte: propria elaborazione

Conclusa questa premessa possiamo ora affrontare il tema della costruzione del campione. Per l'analisi sono stati costruiti due campioni. Un primo campione di aziende quotato sul mercato statunitense e un secondo relativo ad aziende quotate su mercati europei. Avendo la possibilità di confrontare due database diversi, un ulteriore spunto di analisi potrà essere dato dal confronto tra i due modelli e, in particolare, dai regressori che risulteranno significativi. Nei prossimi paragrafi ci occuperemo di costruire il modello e testarne l'efficacia per entrambi i campioni, procedendo successivamente all'analisi dei risultati ottenuti.

Possiamo di seguito osservare l'intero campione selezionato insieme ai valori dei regressori per ciascuna azienda. I valori di bilancio fanno riferimento a dati relativi a bilanci chiusi entro l'anno fiscale 2019.

Nome Società	Quick ratio	$\frac{EBIT}{Tot\ Att}$	$\frac{Att\ a\ lungo}{Tot\ att}$	$\frac{EBIT}{Ricavi}$	$\frac{Ricavi}{Tot\ Att}$	Y
AAON, INC.	2,034	0,180	0,495	0,143	1,264	0
APPLE INC.	1,501	0,205	0,519	0,266	0,769	0
ABCELLERA BIOLOGICS INC.	1,601	- 0,100	0,462	- 0,202	0,494	0
ABIOMED, INC.	5,625	0,250	0,358	0,342	0,730	0
ACADIA HEALTHCARE COMPANY, INC.	1,159	0,047	0,921	0,104	0,452	0
AXCELIS TECHNOLOGIES, INC.	3,427	0,048	0,315	0,076	0,626	0
ACM RESEARCH, INC.	2,871	0,088	0,089	0,178	0,494	0
ALLEGRO MICROSYSTEMS, INC	2,274	0,134	0,512	0,139	0,963	0
ALTAIR ENGINEERING INC.	2,306	0,013	0,526	0,021	0,618	0
AMC NETWORKS INC.	2,897	0,110	0,584	0,202	0,547	0
ADVANCED MICRO DEVICES, INC.	1,532	0,079	0,237	0,071	1,117	0

Nome Società	Quick ratio	$\frac{EBIT}{Tot\ Att}$	$\frac{Att\ a\ lungo}{Tot\ att}$	$\frac{EBIT}{Ricavi}$	$\frac{Ricavi}{Tot\ Att}$	Y
APOLLO MEDICAL HOLDINGS, INC.	3,121	0,037	0,548	0,049	0,762	0
ALLIED MOTION TECHNOLOGIES INC.	1,335	0,095	0,622	0,078	1,213	0
AMPHASTAR PHARMACEUTICALS INC	1,608	0,107	0,565	0,195	0,549	0
A-MARK PRECIOUS METALS, INC.	0,695	0,029	0,052	0,004	6,781	0
THE ANDERSONS, INC.	0,538	0,021	0,501	0,010	2,095	0
AMERICAN OUTDOOR BRANDS, INC.	1,495	- 0,010	0,743	- 0,021	0,500	0
AMERICAN PUBLIC EDUCATION, INC.	4,923	0,058	0,372	0,072	0,807	0
APPFOLIO, INC.	1,302	0,025	0,762	0,025	0,984	0
ASIA PACIFIC WIRE & CABLE CORPORATION LIMITED	2,876	- 0,001	0,199	- 0,001	1,131	0
AERSALE CORPORATION	2,036	0,064	0,588	0,073	0,884	0
ACADEMY SPORTS & OUTDOORS INC	0,194	0,040	0,609	0,027	1,477	0
ASTEC INDUSTRIES, INC.	1,226	0,035	0,368	0,024	1,461	0
ATRION CORPORATION	8,005	0,155	0,495	0,261	0,592	0
AVNET, INC.	1,500	0,043	0,197	0,019	2,279	0
AXT, INC.	2,327	0,002	0,493	0,006	0,373	0
BARRETT BUSINESS SERVICES, INC.	1,239	0,073	0,508	0,067	1,095	0
BIODELIVERY SCIENCES INTERNATIONAL, INC.	1,971	0,020	0,356	0,034	0,609	0
BLUE HAT INTERACTIVE ENTERTAINMENT TECHNOLOGY	3,113	0,165	0,274	0,407	0,405	0
BRIGGS & STRATTON CORP	0,451	- 0,031	0,509	- 0,026	1,184	1
ASCENA RETAIL GROUP INC	0,761	- 0,041	0,572	- 0,024	1,754	1
AKORN INC	0,315	- 0,076	0,627	- 0,144	0,530	1
ENVISION HEALTHCARE CORP (DELAWARE)	3,384	0,046	0,714	0,097	0,472	1
CEC ENTERTAINMENT INC	0,450	0,024	0,951	0,055	0,431	1
COMMERCIAL BARGE LINE CO	0,718	0,048	0,878	0,075	0,643	1
COVIA HOLDINGS LLC	1,766	- 0,038	0,723	- 0,058	0,652	1
DIAMOND OFFSHORE DRILLING INC	1,576	- 0,048	0,918	- 0,301	0,161	1

Nome Società	Quick ratio	$\frac{EBIT}{Tot\ Att}$	$\frac{Att\ a\ lungo}{Tot\ att}$	$\frac{EBIT}{Ricavi}$	$\frac{Ricavi}{Tot\ Att}$	Y
FERRELLGAS PARTNERS LP	0,177	0,098	0,809	0,074	1,334	1
EXTRACTION OIL & GAS INC	0,437	- 0,009	0,924	- 0,029	0,310	1
EQUINOX HOLDINGS INC	1,104	0,089	0,772	0,134	0,666	1
ENVISION HEALTHCARE CORP	3,384	0,046	0,714	0,097	0,472	1
FRONTIER COMMUNICATIONS PARENT INC	1,083	0,084	0,826	0,181	0,464	1
FORUM ENERGY TECHNOLOGIES INC	1,277	- 0,006	0,426	- 0,007	0,825	1
GUITAR CENTER HOLDINGS INC	0,346	0,053	0,609	0,045	1,178	1
HI-CRUSH INC	1,098	0,003	0,839	0,005	0,573	1
JASON INDUSTRIES INC	2,257	- 0,003	0,548	- 0,004	0,873	1
LIBBEY INC	0,756	0,040	0,543	0,036	1,114	1
NEIMAN MARCUS GROUP LTD LLC	0,228	0,018	0,826	0,027	0,649	1
LSC COMMUNICATIONS INC	0,692	- 0,021	0,522	- 0,011	2,017	1
PIONEER ENERGY SERVICES CORP	1,752	- 0,056	0,729	- 0,065	0,854	1
INTERNAP HOLDING LLC	0,090	0,002	0,927	0,004	0,486	1
PIER 1 IMPORTS INC	0,467	- 0,285	0,278	- 0,120	2,366	1
REVLON INC	0,693	0,026	0,627	0,032	0,812	1
RENTPATH LLC	0,966	0,239	0,810	0,219	1,092	1
QUORUM HEALTH CORP	0,239	0,035	0,726	0,031	1,133	1
WHITING PETROLEUM CORP	0,601	0,013	0,957	0,062	0,206	1
ULTRA PETROLEUM CORP	0,054	0,125	0,931	0,306	0,409	1
SM ENERGY CO	0,535	0,010	0,960	0,038	0,253	1

Tabella 16 Composizione del campione statunitense e valori regressori. Fonte: propria elaborazione.

3.2.1 Analisi e risultati

La procedura di analisi effettuata inizia determinando il valore dell'intercetta e dei coefficienti β del modello tramite regressione logistica. Una volta ottenuti questi valori sarà possibile applicare il modello e determinare la probabilità di *default* π associata a ciascuna società del campione. Tuttavia, sarà necessario verificare la significatività dei regressori prima del calcolo di π . Per fare ciò, seguiamo una logica di tipo

“backward”, ossia partiamo dal modello con tutti i cinque regressori, calcoliamo il *p-value* per ogni coefficiente ed eliminiamo il regressore che presenta il *p-value* più elevato, costruendo così un nuovo modello. Tale procedura si arresta solamente quando tutti i coefficienti presentano un valore associato al *p-value* inferiore alla soglia prefissata $\alpha = 5\%$.

Un’analisi preliminare può comunque consistere nel controllo dei valori medi assunti dai regressori nel campione. Ci aspettiamo infatti che le società sane abbiano un valore del *quick ratio* più elevato e, in ogni caso, non inferiore a 1. Un ragionamento analogo può essere portato avanti anche per altri regressori come $\frac{EBIT}{Totale\ Attivo}$, $\frac{Ricavi}{Totale\ Attivo}$ e $\frac{EBIT}{Ricavi}$. Diverso è invece il caso dell’indice che misura la rigidità, $\frac{Totale\ Attivo\ a\ lungo}{Totale\ Attivo}$, in quanto è difficile prefissare una soglia sopra o sotto la quale una società sia da considerarsi sana. Questo valore dipende in larga misura da scelte manageriali, poiché un alto grado di rigidità permette di sfruttare maggiormente la leva operativa e ottenere risultati più che proporzionali all’aumentare dei volumi. Si rischia di andare però incontro a forti perdite nel caso in cui si verifichi una riduzione dei volumi. Inoltre, un’azienda più flessibile è in grado di adattarsi a nuove condizioni di mercato molto più rapidamente.

Analizzando i valori medi dei regressori osserviamo che effettivamente abbiamo ottenuto quanto ci aspettavamo, ossia tutti i valori medi dei regressori sono più alti per le aziende che presentano $Y=0$ ad eccezione dell’indice di rigidità che presenta valori maggiori (quindi maggiore rigidità) per le aziende contraddistinte da $Y=1$. Da questa prima analisi ci aspettiamo allora che tali valori di rigidità giocano a sfavore delle imprese e immaginiamo che i segni del coefficiente di questo indice sarà opposto rispetto agli altri.

Y	Quick Ratio	EBIT/Tot Att	Tot Att a lungo/Tot Attivo	EBIT/Ricavi	Ricavi/Tot Att
1	0,95	0,01	0,73	0,03	0,82
0	2,30	0,07	0,46	0,09	1,11

Tabella 17 Valori medi dei regressori nel campione statunitense. Fonte: propria elaborazione.

Presentiamo adesso i risultati delle regressioni logistiche partendo dal primo dello, ossia quello con tutti i regressori selezionati.

MODELLO 1

Regressori	Coefficienti	P-value
Intercetta	-4,239	0,14003
Quick ratio	-0,650	0,19408
EBIT/Totale Attivo	-18,528	0,09348
Ricavi/Totale Attivo	0,193	0,81549
EBIT/Ricavi	6,093	0,43756
Totale Attivo a lungo/Totale Attivo	8,958	0,00629

Tabella 18 Calcolo dei coefficienti e del *p-value* del modello 1 sul campione statunitense. Fonte: propria elaborazione.

Secondo la procedura sopra evidenziata, rimuoviamo il regressore che risulta meno significativo all’interno dell’analisi, ossia $\frac{Ricavi}{Totale\ Attivo}$. Procediamo quindi alla stima di un secondo modello con quattro regressori.

MODELLO 2

Regressori	Coefficienti	P-value
Intercetta	-3,749	0,04381
Quick ratio	-0,713	0,09205
EBIT/Totale Attivo	-18,488	0,09066
EBIT/Ricavi	6,209	0,42462
Totale Attivo a lungo/Totale Attivo	8,594	0,00242

Tabella 19 Calcolo dei coefficienti e del p-value del modello 2 sul campione statunitense. Fonte: propria elaborazione.

Continuiamo ulteriormente fino a quando il *p-value* di ogni regressore avrà un valore al di sotto della soglia α . Procediamo allora a rimuovere il regressore $\frac{EBIT}{Ricavi}$.

MODELLO 3

Regressori	Coefficienti	P-value
Intercetta	-3,769	0,04120
Quick ratio	-0,657	0,11467
EBIT/Totale Attivo	-11,683	0,02674
Totale Attivo a lungo/Totale Attivo	8,535	0,00248

Tabella 20 Calcolo dei coefficienti e del p-value del modello 3 sul campione statunitense. Fonte: propria elaborazione.

Concludiamo eliminando un ultimo regressore, il *quick ratio*.

MODELLO 4

Regressori	Coefficienti	P-value
Intercetta	-5,296	0,00098
EBIT/Totale Attivo	-13,671	0,00652
Totale Attivo a lungo/Totale Attivo	9,636	0,00035

Tabella 21 Calcolo dei coefficienti e del p-value del modello 4 sul campione statunitense. Fonte: propria elaborazione.

Dall'ultima regressione logistica possiamo ottenere dei risultati soddisfacenti in termini di significatività. Osserviamo infatti che tutti i *p-value* sono molto al di sotto della soglia prefissata $\alpha = 5\%$, potendo ipoteticamente considerare anche una soglia più bassa, ad esempio, $\alpha = 1\%$, senza bisogno di eliminare alcuna variabile. Come ci attendevamo prima di effettuare le regressioni, il segno del coefficiente relativo al regressore $\frac{\text{Totale Attivo a lungo}}{\text{Totale Attivo}}$ è di segno opposto rispetto all'ultimo altro regressore rimasto.

Come anticipato nel capitolo 3, ricordando che il modello logit che vogliamo costruire, se esplicitiamo il valore della probabilità di *default* π , può essere scritto come:

$$\pi_i = P(Y_i = 1|x_i) = \frac{\exp(x_i'\beta)}{1 + \exp(x_i'\beta)}$$

allora il nostro modello sarà il seguente:

$$\pi_i = \frac{\exp(-5,296 - 13,671x_1 + 9,636x_2)}{1 + \exp(-5,296 - 13,671x_1 + 9,636x_2)}$$

Dove x_1 indica il regressore $\frac{EBIT}{Totale\ Attivo}$ e x_2 rappresenta, invece, il regressore $\frac{Totale\ Attivo\ a\ lungo}{Totale\ Attivo}$.

Testiamo adesso l'efficacia del modello verificando quante aziende riesce a classificare correttamente. Tale verifica verrà eseguita confrontando il valore della variabile indipendente Y per ciascuna azienda (valore ovviamente noto *a priori*) con la probabilità di *default* π stimata dal modello. In seguito, viene stabilita una soglia di *cut off* che divide i valori delle probabilità in due parti. Per farlo, consideriamo gli errori di classificazione (ossia aziende con Y=0 classificate come in *default* e, viceversa, aziende con Y=1 classificate come sane) rispetto al totale delle osservazioni e minimizziamo tale valore. Otteniamo una soglia di *cut off* pari a circa il 44%.

Abbiamo che:

Nome azienda	Y	P	Classificazione	Classificazione corretta
AAON, INC.'	0	0,05	<i>non default</i>	sì
APPLE INC.	0	0,04	<i>non default</i>	sì
ABCELLERA BIOLOGICS INC.	0	0,63	<i>default</i>	no
ABIOMED, INC.	0	0,01	<i>non default</i>	sì
ACADIA HEALTHCARE COMPANY, INC.	0	0,95	<i>default</i>	no
AXCELIS TECHNOLOGIES, INC.	0	0,05	<i>non default</i>	sì
ACM RESEARCH, INC.	0	0	<i>non default</i>	sì
ALLEGRO MICROSYSTEMS, INC	0	0,1	<i>non default</i>	sì
ALTAIR ENGINEERING INC.	0	0,4	<i>non default</i>	sì
AMC NETWORKS INC.	0	0,24	<i>non default</i>	sì
ADVANCED MICRO DEVICES, INC.	0	0,02	<i>non default</i>	sì
APOLLO MEDICAL HOLDINGS, INC.	0	0,37	<i>non default</i>	sì
ALLIED MOTION TECHNOLOGIES INC.	0	0,35	<i>non default</i>	sì
AMPHASTAR PHARMACEUTICALS INC	0	0,21	<i>non default</i>	sì
A-MARK PRECIOUS METALS, INC.	0	0,01	<i>non default</i>	sì
THE ANDERSONS, INC.	0	0,32	<i>non default</i>	sì
AMERICAN OUTDOOR BRANDS, INC.	0	0,88	<i>default</i>	no
AMERICAN PUBLIC EDUCATION, INC.	0	0,08	<i>non default</i>	sì
APPFOLIO, INC.	0	0,85	<i>default</i>	no
ASIA PACIFIC WIRE & CABLE CORPORATION LIMITED	0	0,03	<i>non default</i>	sì
AERSALE CORPORATION	0	0,38	<i>non default</i>	sì
ACADEMY SPORTS & OUTDOORS INC	0	0,51	<i>default</i>	no
ASTEC INDUSTRIES, INC.	0	0,1	<i>non default</i>	sì
ATRION CORPORATION	0	0,07	<i>non default</i>	sì
AVNET, INC.	0	0,02	<i>non default</i>	sì
AXT, INC.	0	0,36	<i>non default</i>	sì
BARRETT BUSINESS SERVICES, INC.	0	0,2	<i>non default</i>	sì
BIODELIVERY SCIENCES INTERNATIONAL, INC.	0	0,1	<i>non default</i>	sì
BLUE HAT INTERACTIVE ENTERTAINMENT TECHNOLOGY	0	0,01	<i>non default</i>	sì
Briggs & Stratton Corp	1	0,51	<i>default</i>	sì
Ascena Retail Group Inc	1	0,69	<i>default</i>	sì
Akorn Inc	1	0,86	<i>default</i>	sì

Nome azienda	Y	P	Classificazione	Classificazione corretta
Envision Healthcare Corp (Delaware)	1	0,72	default	sì
Cec Entertainment Inc	1	0,97	default	sì
Commercial Barge Line Co	1	0,92	default	sì
Covia Holdings LLC	1	0,9	default	sì
Diamond Offshore Drilling Inc	1	0,99	default	sì
Ferrellgas Partners LP	1	0,76	default	sì
Extraction Oil & Gas Inc	1	0,98	default	sì
Equinox Holdings Inc	1	0,72	default	sì
Envision Healthcare Corp	1	0,72	default	sì
Frontier Communications Parent Inc	1	0,82	default	sì
Forum Energy Technologies Inc	1	0,25	non default	no
Guitar Center Holdings Inc	1	0,46	default	sì
Hi-Crush Inc	1	0,94	default	sì
Jason Industries Inc	1	0,51	default	sì
Libbey Inc	1	0,35	non default	no
Neiman Marcus Group Ltd LLC	1	0,92	default	sì
LSC Communications Inc	1	0,51	default	sì
Pioneer Energy Services Corp	1	0,92	default	sì
Internap Holding LLC	1	0,97	default	sì
Pier 1 Imports Inc	1	0,78	default	sì
Revlon Inc	1	0,6	default	sì
RentPath LLC	1	0,32	non default	no
Quorum Health Corp	1	0,77	default	sì
Whiting Petroleum Corp	1	0,98	default	sì
Ultra Petroleum Corp	1	0,88	default	sì
SM Energy Co	1	0,98	default	sì

Tabella 22 Classificazione delle aziende secondo il modello costruito. Fonte: propria elaborazione.

Analizzando anche il tipo di errore di classificazione possiamo vedere che:

		Previsto		Totale
		Y=0	Y=1	
Effettivo	Y=0	24	5	29
	Y=1	3	26	29
Totale		27	31	58

Tabella 23 Analisi degli errori per tipologia. Fonte: propria elaborazione.

Notiamo che 50 aziende sono correttamente classificate. Invece, 5 aziende sane sono classificate come insolventi mentre 3 insolventi, al contrario, sono classificate come sane. Utilizzando quindi la soglia del *cut off* che minimizza gli errori otteniamo un'accuratezza del modello dell'86,21%⁵⁴.

⁵⁴ Le soglie che minimizzano gli errori, presenti nell'analisi di ciascun campione presentato, sono calcolate utilizzando il componente aggiuntivo di Excel "Solver", impostando il vincolo di risultato tra 0 e 1 e chiedendo la minimizzazione della cella obiettivo, ossia la cella indicante la percentuale di errore.

Per ulteriore conferma dei risultati ottenuti, il modello è stato testato anche su un campione di verifica composto da 52 società, di cui 22 presentano $Y=1$. Anche in questo caso il campione è ampio e riferito a diversi settori in modo tale da cercare di imitare il più possibile l'andamento del mercato.

Tuttavia, una differenza importante consiste nella data di riferimento dei bilanci utilizzati nell'analisi. Infatti, mentre nel campione precedente i bilanci facevano riferimento a dati antecedenti al covid, non superando quindi la data del 31.12.2019, il campione di verifica vuole testare la significatività del modello dopo gli effetti della pandemia. Sono stati infatti selezionati bilancio riferiti relativi all'anno fiscale 2020. Di seguito si riporta il campione di verifica insieme all'applicazione del modello precedentemente illustrato. La soglia di *cut off* usata in questa prima fase è la stessa calcolata per il *training set*, ossia il 44%, in modo tale da mantenere coerenza per il confronto della correttezza di classificazione nei due campioni.

Nome dell'azienda	Y	P	Classificata correttamente
BBQ HOLDINGS, INC.	1	0,670	sì
CANADIAN SOLAR INC.	1	0,602	Sì
FREQUENCY THERAPEUTICS, INC.	1	0,987	sì
GREENPRO CAPITAL CORP.	1	0,173	no
iQIYI, Inc.	1	0,706	sì
IF BANCORP, INC.	1	0,008	no
LKQ CORPORATION	1	0,035	no
NASDAQ, INC.	1	0,101	NO
PETIQ, INC.	1	0,099	no
17 EDUCATION & TECHNOLOGY GROUP INC.	1	1,000	sì
GLBL X AI & TECHNOLOGY	1	0,025	no
BOQII HOLDING LTD	1	0,993	sì
CANADIAN NATURAL RESOURCES LIMITED	1	0,003	no
DAQO NEW ENERGY CORP.	1	0,003	no
DRIL-QUIP, INC.	1	0,896	sì
HP INC.	1	0,263	no
SQUARE, INC.	1	0,910	sì
TURQUOISE HILL RESOURCES LTD.	1	0,012	no
TOWNSQUARE MEDIA, INC.	1	0,017	no
ELITE EDUCATION GROUP INTERNATIONAL LTD	1	0,662	sì
KELSO TECHNOLOGIES INC.	1	0,833	sì
MOVING IMAGE TECHNOLOGIES INC	1	0,897	sì
AMERICAN WOODMARK CORPORATION	0	0,751	no
BALCHEM CORPORATION	0	0,732	no
AESTHETIC MEDICAL INTERNATIONAL HOLDINGS GROUP LIMITED	0	0,116	sì
ADVANCED EMISSIONS SOLUTIONS, INC.	0	0,071	sì
AMEDISYS, INC.	0	0,419	sì
MISSION PRODUCE, INC.	0	0,343	sì
AMGEN INC.	0	0,264	sì
ADOBE INC.	0	0,292	sì
AKAMAI TECHNOLOGIES, INC.	0	0,534	no

Nome dell'azienda	Y	P	Classificata correttamente
ANSYS, INC.	0	0,447	no
ACTIVISION BLIZZARD, INC.	0	0,402	sì
ALLIED MOTION TECHNOLOGIES INC.	0	0,355	sì
ABCAM PLC	0	0,249	sì
AMC NETWORKS INC.	0	0,235	sì
AMPHASTAR PHARMACEUTICALS INC	0	0,212	sì
ALARM.COM HOLDINGS, INC.	0	0,199	sì
ASPEN TECHNOLOGY, INC.	0	0,007	sì
APPLE INC.	0	0,043	sì
APPLOVIN CORPORATION	0	0,071	sì
ALLEGRO MICROSYSTEMS, INC	0	0,101	sì
AUTOSCOPE TECHNOLOGIES CORPORATION	0	0,151	sì
AAON, INC.'	0	0,048	sì
ATRION CORPORATION	0	0,067	sì
ASML HOLDING NV	0	0,075	sì
APPLIED MATERIALS, INC.	0	0,033	sì
ABIOMED, INC.	0	0,005	sì
ARENA PHARMACEUTICALS, INC.	0	0,001	sì
ALIGN TECHNOLOGY, INC.	0	0,007	sì
ABSOLUTE SOFTWARE CORPORATION	0	0,029	sì
ARRAY TECHNOLOGIES, INC.	0	0,034	sì

Tabella 24 Classificazione delle aziende del campione di controllo secondo il modello costruito. Fonte: propria elaborazione.

Utilizzando la soglia di *cut off* selezionata tramite il modello precedente abbiamo 4 aziende sane classificate con insolventi e, al contrario, 11 aziende insolventi classificate come sane. Questa classificazione ci porta a una percentuale di correttezza della classificazione pari a circa il 71,15%. Tale percentuale risulta inferiore a quella calcolata nel caso precedente ma è comunque significativa e accettabile.

Possiamo rappresentare graficamente tale risultato come segue:

		Previsto		Totale
		Y=0	Y=1	
Effettivo	Y=0	26	4	30
	Y=1	11	11	22
Totale		37	15	52

Tabella 25 Analisi degli errori per tipologia sul campione di verifica. Fonte: propria elaborazione.

Possiamo anche calcolare la soglia di *cut off* che minimizzi gli errori per il campione di verifica. Essa è uguale al 55% e permette di ottenere una corretta classificazione per il 75% delle imprese. Quindi, senza eccessiva influenza da parte della soglia di *cut off* selezionata, il modello presenta un buon adattamento sia per il *training set* che considerava valori di bilancio pre-covid sia per il campione di verifica che testa valori di bilancio post-covid.

Potremmo chiederci perché il modello continui a funzionare in maniera soddisfacente anche dopo la pandemia. Il motivo proposto nella presente analisi è il seguente: gli indici che sono risultati maggiormente

significativi colgono degli aspetti che sono particolarmente rilevanti anche e soprattutto durante il contesto post covid. Infatti, il *lockdown* imposto dalla politica in risposta al covid può essere visto come uno *stress test* per le aziende. Solo quelle più efficienti avevano buone possibilità di sopravvivenza. Inoltre, tale *stress test* ha riguardato anche le vendite che hanno subito una grandissima perdita a causa della chiusura forzata delle attività. L'EBIT/Totale Attivo rappresenta quindi l'efficienza nella produzione del reddito relativamente alle risorse patrimoniali a disposizione delle aziende. Invece, la rigidità delle può essere vista come un'integrazione verticale delle attività produttive, creando maggiori costi fissi e uno sfruttamento significativo della leva operativa. Questo porta a un aumento dei profitti migliore rispetto alle aziende più flessibili ma una esposizione maggiore nei confronti della variazione dei ricavi. Così le aziende più rigide sono state quelle maggiormente in difficoltà in seguito alla notevole riduzione degli utili derivante dal *lockdown*.

3.3 Considerazioni ulteriori tramite la costruzione di un campione europeo

Ultimata la costruzione del modello tramite il campione statunitense e il campione di verifica, è stato eseguito un ulteriore passaggio costruendo un altro campione di aziende, questa volta quotate sui mercati europei. I dati utilizzati sono relativi a bilanci chiusi entro l'anno fiscale 2019. Tale campione comprende 61 aziende di cui 22 in *default*. È stata seguita la stessa logica precedente per la costruzione del modello. Alla fine, confronteremo i regressori rimasti con quelli significativi per il campione statunitense.

Anche questa volta analizziamo i valori medi dei regressori e osserviamo che essi presentano nuovamente quanto ci aspettavamo, ossia tutti i valori medi dei regressori sono più alti per le aziende che presentano $Y=0$ ad eccezione dell'indice di rigidità che presenta valori maggiori (quindi maggiore rigidità) per le aziende contraddistinte da $Y=1$.

Y	Current Ratio	EBIT/Tot Att	Tot Att a lungo/Tot Attivo	EBIT/Ricavi	Ricavi/Tot Att
1	1,22	-0,02	0,73	-0,08	0,60
0	1,66	0,04	0,50	0,09	1,27

Tabella 26 Valori medi dei regressori nel campione europeo. Fonte: propria elaborazione.

Si riporta di seguito il campione europeo utilizzato per la stima del modello insieme ai valori dei regressori.

Nome Società	Quick ratio	$\frac{EBIT}{Tot Att}$	$\frac{Att a lungo}{Tot att}$	$\frac{EBIT}{Ricavi}$	$\frac{Ricavi}{Tot Att}$	Y
AMPLIFON	0,674	0,071	0,829	0,117	0,607	0,000
ASKOLL EVA	2,307	-0,255	0,477	-0,518	0,493	0,000
ATLANTIA	0,964	0,025	0,849	0,139	0,178	0,000
BHB BRAU HOLDING BAYERN MITTE AG	3,942	0,033	0,670	0,027	1,197	0,000
BIALETTI	1,420	0,002	0,550	0,003	0,859	0,000
BIOMASS ENERGY PROJECT SA	2,766	0,021	0,357	0,286	0,072	0,000
CNH INDUSTRIAL	1,610	0,039	0,299	0,069	0,571	0,000
CRANSWICK PLC	1,717	0,196	0,388	0,062	3,150	0,000
EMMI AG	2,309	0,089	0,457	0,062	1,431	0,000
ENEL	0,907	0,065	0,786	0,145	0,451	0,000

Nome Società	Quick ratio	$\frac{EBIT}{Tot Att}$	$\frac{Att a lungo}{Tot att}$	$\frac{EBIT}{Ricavi}$	$\frac{Ricavi}{Tot Att}$	Y
ENI	1,178	0,071	0,717	0,126	0,566	0,000
EPRICE	0,691	-0,517	0,490	-0,193	2,673	0,000
FERRARI	1,391	0,168	0,515	0,243	0,692	0,000
FROSTA AG	1,883	0,062	0,407	0,038	1,628	0,000
GEL SPA	1,990	0,027	0,354	0,032	0,833	0,000
GENERAL ENERGY PLC	5,699	0,083	0,719	0,430	0,192	0,000
GIGLIO GROUP	0,607	0,015	0,433	0,021	0,734	0,000
KI GROUP	0,443	-0,231	0,364	-0,100	2,298	0,000
KRI KRI MILK INDUSTRY SA	2,375	0,166	0,545	0,158	1,052	0,000
LOTUS BAKERIES NV	1,006	0,127	0,789	0,168	0,754	0,000
NATIONAL MILK RECORDS PLC	1,436	0,219	0,450	0,099	2,202	0,000
NEW NORDIC HEALTHBRANDS AB	2,094	0,183	0,038	0,080	2,277	0,000
NWF GROUP PLC	0,972	0,067	0,492	0,015	4,522	0,000
PHAROS ENERGY PLC	1,638	0,044	0,864	0,200	0,221	0,000
PURCARI WINERIES PCL	1,473	0,156	0,583	0,277	0,562	0,000
RAK PETROLEUM PLC	1,880	0,065	0,669	0,239	0,271	0,000
SIRIO	0,569	-0,012	0,821	-0,021	0,555	0,000
STMICROELECTRONICS	2,019	0,097	0,506	0,129	0,753	0,000
STOCK SPIRITS GROUP PLC	1,499	0,081	0,677	0,176	0,456	0,000
SUBSEA 7 SA	1,272	0,026	0,714	0,040	0,652	0,000
TATE & LYLE PLC	1,609	0,170	0,354	0,107	1,585	0,000
TECHNIPFMC PLC	1,172	0,045	0,495	0,079	0,571	0,000
TEQNION AB	1,859	0,085	0,415	0,067	1,269	0,000
TERNA	0,988	0,063	0,834	0,498	0,127	0,000
TIM	0,851	0,047	0,865	0,184	0,256	0,000
TRETT PLC	3,396	0,103	0,244	0,118	0,871	0,000
VIMI FASTENERS	1,222	-0,002	0,589	-0,003	0,687	0,000
VISIBILIA	1,235	-0,138	0,475	-0,285	0,483	0,000
WYNNSTAY GROUP PLC	1,817	0,162	-1,754	0,015	10,884	0,000
CHL	0,809	-0,112	0,573	-0,137	0,815	1,000
CINEWORLD GROUP PLC	0,301	0,063	0,964	0,178	0,352	1,000
CODERE SA	0,811	0,051	0,810	0,062	0,830	1,000
DEOLEO SA	1,731	0,025	0,686	0,037	0,663	1,000
DISTRIBUIDORA INTERNACIONAL DE ALIMENTACION SA	0,504	-0,136	0,738	-0,066	2,070	1,000
ENERGY LAB	1,182	-0,187	0,258	-0,637	0,294	1,000
EUROPCAR MOBILITY GROUP	1,099	0,037	0,351	0,102	0,364	1,000
HAYA REAL ESTATE SAU	2,392	-0,049	0,634	-0,101	0,488	1,000
INTELSAT SA	1,699	0,065	0,901	0,374	0,175	1,000
LECTA SA	1,097	0,031	0,581	0,026	1,197	1,000
MALLINCKRODT PLC	1,283	0,023	0,823	0,076	0,305	1,000
MOBI SPA	0,167	-0,269	0,834	-0,736	0,366	1,000
NOBLE HOLDING CORPORATION PLC	0,817	-0,026	0,949	-0,175	0,151	1,000
NOSTRUM OIL & GAS PLC	1,699	0,064	0,795	0,168	0,381	1,000
PACIFIC DRILLING SA	5,179	-0,102	0,827	-1,000	0,102	1,000
PGS ASA	0,518	0,032	0,815	0,078	0,405	1,000

Nome Società	Quick ratio	$\frac{EBIT}{Tot Att}$	$\frac{Att a lungo}{Tot att}$	$\frac{EBIT}{Ricavi}$	$\frac{Ricavi}{Tot Att}$	Y
PIZZAEXPRESS FINANCING 1 PLC	0,822	0,032	0,921	0,070	0,451	1,000
SAS AB	0,776	0,035	0,655	0,025	1,385	1,000
TECHNICOLOR SA	0,731	-0,005	0,649	-0,004	1,184	1,000
TRANSOCEAN LTD	2,111	-0,004	0,849	-0,032	0,128	1,000
VESTEL ELEKTRONIK SANAYI VE TICARET AS	0,656	0,085	0,506	0,091	0,928	1,000
ZELLIS HOLDINGS LTD	0,423	-0,034	0,919	-0,156	0,218	1,000

Tabella 27 Composizione del campione europeo e valori regressori. Fonte: propria elaborazione.

Continuiamo presentando i risultati delle regressioni che permettono di stimare i coefficienti dei regressori selezionati⁵⁵.

MODELLO 1

Regressori	Coefficienti	P-value
Intercetta	-2,583	0,22618
Current ratio	-0,787	0,16771
EBIT/Totale attivo	3,055	0,50782
Ricavi/Totale Attivo	-0,348	0,53397
EBIT/RICAVI	-6,179	0,02969
Totale Attivo a Lungo /Totale Attivo	5,416	0,02644

Tabella 28 Calcolo dei coefficienti e del p-value del modello 1 sul campione europeo. Fonte: propria elaborazione.

Seguendo la stessa logica utilizzata per il campione americano, rimuoviamo il regressore che presenta il *p-value* maggiore, ossia $\frac{Ricavi}{Totale Attivo}$. Continuiamo con la stima del modello successivo.

MODELLO 2

Regressori	Coefficienti	P-value
Intercetta	-3,495	0,03734
Current ratio	-0,702	0,20027
EBIT/Totale attivo	3,351	0,43083
EBIT/RICAVI	-6,477	0,02145
Totale Attivo a Lungo /Totale Attivo	6,218	0,00463

Tabella 29 Calcolo dei coefficienti e del p-value del modello 2 sul campione europeo. Fonte: propria elaborazione.

Nel secondo modello ritroviamo un altro regressore con *p-value* oltre la soglia fissata, ossia $\frac{EBIT}{Totale Attivo}$. Costruiamo un nuovo modello senza tale regressore.

MODELLO 3

Regressori	Coefficienti	P-value
Intercetta	-3,600	0,02541
Current ratio	-0,588	0,23457
EBIT/RICAVI	-4,941	0,00770
Totale Attivo a Lungo /Totale Attivo	6,121	0,00425

⁵⁵ Nel campione europeo il *quick ratio* è stato sostituito dal *current ratio* per una questione di semplicità di reperimento dei dati. Tuttavia, l'interpretazione del valore del regressore è analoga.

Tabella 30 Calcolo dei coefficienti e del p-value del modello 3 sul campione europeo. Fonte: propria elaborazione.

Otteniamo, infine, un ultimo modello eliminando il *current ratio*, poiché non significativo secondo il criterio di scelta adottato.

MODELLO 4

Regressori	Coefficienti	P-value
Intercetta	-4,700	0,00106
EBIT/RICAVI	-5,184	0,00743
Totale Attivo a Lungo /Totale Attivo	6,658	0,00186

Tabella 31 Calcolo dei coefficienti e del p-value del modello 4 sul campione europeo. Fonte: propria elaborazione.

L'ultimo modello risulta essere soddisfacente in termini di significatività dei regressori. Come possiamo notare, tutte i *p-value* sono ampiamente ricompresi nella soglia precedentemente scelta pari a $\alpha = 5\%$. Come nel caso del campione statunitense, i regressori, come ci attendevamo, si presentano con coefficienti di segno opposto. Inoltre, per entrambi i campioni il regressore $\frac{\text{Totale Attivo a lungo}}{\text{Totale Attivo}}$ rimane significativo mentre il secondo regressore è differente: $\frac{\text{EBIT}}{\text{Ricavi}}$ per il campione europeo, $\frac{\text{EBIT}}{\text{Totale Attivo}}$ per quello americano. Tuttavia, la logica sottostante e l'interpretazione risulta essere simile in quanto sono legati a un concetto di efficienza nella produzione dell'EBIT, dal punto di vista economico il primo, patrimoniale il secondo.

Analizziamo adesso la capacità di classificazione del modello. Calcoliamo allora le probabilità di *default* tramite la seguente equazione:

$$\pi_i = \frac{\exp(-4,700 - 5,184x_1 + 6,658x_2)}{1 + \exp(-4,700 - 5,184x_1 + 6,658x_2)}$$

Con la soglia di *cut off* che minimizza gli errori⁵⁶, pari al 57,82%, otteniamo che:

Nome azienda	Y	P	Classificazione corretta
Amplifon	0	0,553899415	Sì
Askoll Eva	0	0,762416042	No
Atlantia	0	0,557925568	Sì
BHB Brauholding Bayern Mitte AG	0	0,405628406	Sì
Bialetti	0	0,258885374	Sì
Biomass Energy Project SA	0	0,021843025	Sì
CNH Industrial	0	0,044432573	Sì
Cranswick PLC	0	0,080281709	Sì
Emmi AG	0	0,121630444	Sì
ENEL	0	0,445527315	Sì
ENI	0	0,359286258	Sì
ePrice	0	0,393706847	Sì
Ferrari	0	0,073886802	Sì
FRoSTA AG	0	0,101086363	Sì

⁵⁶ Utilizzando la soglia stabilita per il campione statunitense, pari a circa il 44%, otteniamo una corretta classificazione per circa il 72% delle imprese selezionate.

Nome azienda	Y	P	Classificazione corretta
Gel SpA	0	0,075017538	Sì
General Energy PLC	0	0,105438065	Sì
Giglio Group	0	0,127472441	Sì
Ki Group	0	0,14764583	Sì
Kri Kri Milk Industry SA	0	0,131499865	Sì
Lotus Bakeries NV	0	0,421291158	Sì
National Milk Records PLC	0	0,098204593	Sì
New Nordic Healthbrands AB	0	0,007687989	Sì
NWF Group PLC	0	0,181841472	Sì
Pharos Energy PLC	0	0,502821218	Sì
Purcari Wineries PCL	0	0,094826327	Sì
RAK Petroleum PLC	0	0,185314415	Sì
Sirio	0	0,706556382	No
STMicroelectronics	0	0,119266694	Sì
Stock Spirits Group PLC	0	0,248559637	Sì
Subsea 7 SA	0	0,46059168	Sì
Tate & Lyle PLC	0	0,052301363	Sì
TechnipFMC PLC	0	0,139799935	Sì
Teqnon AB	0	0,092019142	Sì
Terna	0	0,150835949	Sì
TIM	0	0,526522256	Sì
Treant PLC	0	0,024497807	Sì
Vimi Fasteners	0	0,318965479	Sì
Visibilia	0	0,485133966	Sì
Wynnstay Group PLC	0	7,15359E-08	Sì
CHL	1	0,456606351	No
Cineworld Group PLC	1	0,688820407	Sì
Codere sa	1	0,592802485	Sì
Deoleo sa	1	0,420179721	No
Distribuidora Internacional de Alimentacion SA	1	0,634668905	Sì
Energy Lab	1	0,578569982	Sì
Europcar mobility group	1	0,052571991	No
Haya Real Estate SAU	1	0,511104758	No
Intelsat SA	1	0,34515255	No
Lecta SA	1	0,275484278	No
Mallinckrodt PLC	1	0,594901178	Sì
Mobi Spa	1	0,990706736	Sì
Noble Holding Corporation PLC	1	0,925909508	Sì
Nostrum Oil & Gas PLC	1	0,431759897	No
Pacific Drilling SA	1	0,997502387	Sì
PGS ASA	1	0,579786168	Sì
Pizzaexpress Financing 1 PLC	1	0,744753595	Sì
SAS AB	1	0,384731623	No
Technicolor SA	1	0,411227422	No
Transocean Ltd	1	0,7541231	Sì
Vestel Elektronik Sanayi ve Ticaret AS	1	0,141474992	No

Nome azienda	Y	P	Classificazione corretta
Zellis Holdings Ltd	1	0,902614555	Sì

Tabella 32 Classificazione delle aziende secondo il modello costruito. Fonte: propria elaborazione.

Andando maggiormente nel dettaglio, possiamo classificare il tipo di errore come segue:

		Previsto		Totale
		Y=0	Y=1	
Effettivo	Y=0	37	2	39
	Y=1	10	12	22
Totale		47	14	61

Tabella 33 Analisi degli errori per tipologia. Fonte: propria elaborazione.

Possiamo osservare che 49 aziende sono state classificate correttamente. Invece, 2 aziende sane sono state classificate come insolventi e 10 insolventi sono state considerate sane. Possiamo concludere che anche il tentativo realizzato su un campione europeo conferma i risultati ottenuti tramite il modello statunitense.

3.4 Appendice: valori utilizzati per il calcolo dei regressori

3.4.1 Campione statunitense

I dati sono presentati nella tabella seguente, in migliaia di dollari.

Società	TCA - Inventory	TNCA	TA	TCL	Revenue	EBIT
AAON, INC. ¹	113.948	183.875	371.424	56.028	469.333	67.031
APPLE INC.	158.713.000	175.697.000	338.516.000	105.718.000	260.174.000	69.313.000
ABCELLERA BIOLOGICS INC.	12.640	10.848	23.488	7.895	11.612	-2.343
ABIOMED, INC.	596.255	377.149	1.054.346	105.998	769.432	263.360
ACADIA HEALTHCARE COMPANY, INC.	537.452	6.336.931	6.879.142	463.576	3.107.462	321.883
AXCELIS TECHNOLOGIES, INC.	235.315	172.415	548.094	68.664	342.958	26.077
ACM RESEARCH, INC.	153.600	19.307	217.703	53.502	107.524	19.121
ALLEGRO MICROSYSTEMS, INC	236.131	385.213	752.261	103.862	724.311	100.535
ALTAIR ENGINEERING INC.	352.457	390.688	743.145	152.866	458.915	9.759
AMC NETWORKS INC.	2.330.297	3.266.389	5.596.686	804.342	3.060.321	616.754
ADVANCED MICRO DEVICES, INC.	3.615.000	1.431.000	6.028.000	2.359.000	6.731.000	479.000
APOLLO MEDICAL HOLDINGS, INC.	329.068	399.645	728.713	105.424	555.409	27.016

Società	TCA - Inventory	TNCA	TA	TCL	Revenue	EBIT
ALLIED MOTION TECHNOLOGIES INC.	62.258	190.185	305.828	46.641	371.084	28.975
AMPHASTAR PHARMACEUTICALS INC	144.740	331.730	586.971	90.009	322.357	62.781
A-MARK PRECIOUS METALS, INC.	375.989	36.512	705.362	540.937	4.783.157	20.386
THE ANDERSONS, INC.	774.806	1.955.399	3.900.741	1.439.919	8.170.191	81.050
AMERICAN OUTDOOR BRANDS, INC.	30.177	263.655	354.773	20.189	177.363	-3.669
AMERICAN PUBLIC EDUCATION, INC.	222.909	131.988	354.897	45.278	286.270	20.648
APPFOLIO, INC.	61.791	198.311	260.102	47.451	256.012	6.461
ASIA PACIFIC WIRE & CABLE CORPORATION LIMITED	154.317	59.407	298.911	53.649	338.160	-191
AERSALE CORPORATION	83.867	202.198	343.984	41.194	304.201	22.058
ACADEMY SPORTS & OUTDOORS INC	133.116	1.971.685	3.238.957	684.483	4.783.893	128.950
ASTEC INDUSTRIES, INC.	211.768	294.194	800.498	172.767	1.169.613	28.260
ATRION CORPORATION	90.245	129.693	262.031	11.274	155.066	40.529
AVNET, INC.	3.867.912	1.688.220	8.564.556	2.578.576	19.518.592	365.911
AXT, INC.	64.053	110.144	223.349	27.526	83.256	494
BARRETT BUSINESS SERVICES, INC.	423.848	437.081	860.929	342.122	942.311	62.927
BIODELIVERY SCIENCES INTERNATIONAL, INC.	106.447	65.146	182.905	53.993	111.389	3.740
BLUE HAT INTERACTIVE ENTERTAINMENT TECHNOLOGY	42.592	16.098	58.816	13.680	23.834	9.700
Briggs & Stratton Corp	260.470	788.950	1.551.430	577.750	1.836.610	-48.220
Ascena Retail Group Inc	607.300	1.544.800	2.699.800	798.200	4.734.700	-111.600

Società	TCA - Inventory	TNCA	TA	TCL	Revenue	EBIT
Akorn Inc	310.000	808.590	1.288.640	985.590	682.430	-98.430
Envision Healthcare Corp (Delaware)	4.721.600	11.828.300	16.572.600	1.395.100	7.819.300	758.400
Cec Entertainment Inc	76.800	2.014.770	2.119.550	170.810	912.870	50.200
Commercial Barge Line Co	103.710	1.107.920	1.262.560	144.470	811.560	61.140
Covia Holdings LLC	555.230	1.769.500	2.446.520	314.430	1.595.450	-93.320
Diamond Offshore Drilling Inc	476.795	5.357.248	5.834.043	302.592	936.483	-281.730
Ferrellgas Partners LP	160.930	1.021.580	1.262.960	908.360	1.684.390	124.000
Extraction Oil & Gas Inc	202.790	2.703.210	2.926.960	464.530	906.640	-26.370
Equinox Holdings Inc	47.810	167.190	216.470	43.300	144.170	19.280
Envision Healthcare Corp	4.721.600	11.828.300	16.572.600	1.395.100	7.819.300	758.400
Frontier Communications Parent Inc	3.037.000	14.451.000	17.488.000	2.804.000	8.107.000	1.466.000
Forum Energy Technologies Inc	251.280	494.080	1.160.000	196.850	956.530	-6.880
Guitar Center Holdings Inc	145.300	1.106.300	1.816.560	419.720	2.139.190	95.670
Hi-Crush Inc	139.210	931.960	1.111.140	126.830	636.370	3.180
Jason Industries Inc	125.050	212.110	387.100	55.400	337.900	-1.200
Libbey Inc	131.940	342.530	630.400	174.610	701.970	25.460
Neiman Marcus Group Ltd LLC	196.020	6.234.040	7.545.900	859.680	4.900.440	134.230
LSC Communications Inc	618.000	861.000	1.649.000	893.000	3.326.000	-35.000
Pioneer Energy Services Corp	160.460	491.040	673.950	91.580	575.790	-37.550
Internap Holding LLC	43.630	555.730	599.360	485.580	291.510	1.040
Pier 1 Imports Inc	125.950	182.760	656.290	269.460	1.552.940	-186.820
Revlon Inc	663.000	1.869.200	2.980.600	956.900	2.419.600	76.700
RentPath LLC	39.950	172.420	212.740	41.340	232.220	50.870
Quorum Health Corp	370.290	1.082.840	1.491.880	1.550.790	1.689.630	52.080
Whiting Petroleum Corp	330.980	7.305.740	7.636.720	550.410	1.572.250	97.870
Ultra Petroleum Corp	125.490	1.689.790	1.815.280	2.324.220	742.030	227.070
SM Energy Co	252.630	6.039.600	6.292.230	472.040	1.589.240	60.550

Tabella 34 Valori di bilancio utilizzati per il calcolo dei regressori. Campione statunitense. Fonte: Database Refinitiv.

3.4.2 Campione di verifica statunitense

I dati sono presentati nella tabella seguente, in migliaia di dollari.

Società	TCA - Inventory	TNCA	TA	Revenue	EBIT
BBQ HOLDINGS, INC.	309.715	42.957	71.976	121.437	58.041
Canadian Solar Inc.	172.017	2.351.032	6.536.854	3.476.495	21.492
FREQUENCY THERAPEUTICS, INC.	343.066	39.658	264.722	36.984	16.777
GREENPRO CAPITAL CORP.	386.935	13.184	14.796	2.255	11.713
iQIYI, Inc.	1.938.766	25.895.005	48.185.429	29.707.215	599.939
IF BANCORP, INC.	4.744.000	537.748	735.517	23.071	1.369.000
LKQ CORPORATION	338.577	8.326.501	12.360.533	11.628.830	80.503
Nasdaq, Inc.	79.236	10.319.000	17.979.000	5.627.000	22.994
PETIQ, INC.	182.851	529.296	771.592	780.051	10.954
17 EDUCATION & TECHNOLOGY GROUP INC.	392.738	343.162	3.389.742	1.294.371	30.217
Globl X AI & Technology	1.115.716	546.706	659.864	505.549	5.715
Boqii Holding Ltd	105.535	185.782	880.574	1.010.985	23.495
CANADIAN NATURAL RESOURCES LIMITED	251.810	70.915.000	75.276.000	17.491.000	25.608
Daqo New Energy Corp.	199.353	1.058.872	1.239.143	675.602	9.783
DRIL-QUIP, INC.	167.328	287.979	1.151.172	364.973	8.096
HP INC.	178.964	14.033.000	34.681.000	56.639.000	11.283
SQUARE, INC.	636.251	2.107.921	9.869.550	9.497.578	420.398
TURQUOISE HILL RESOURCES LTD.	316.100	11.859.892	13.368.761	1.078.192	51.100
TOWNSQUARE MEDIA, INC.	62.021	614.211	768.502	371.338	35.668
Elite Education Group International Ltd	3.926.584	3.560	14.202	9.063	118.001
Kelso Technologies Inc.	113.003	4.807	12.017	11.149	7.740
Moving Image Technologies Inc	407.875	2.166	5.704	16.367	36.147
AMERICAN WOODMARK CORPORATION	197.507	1.223.896	1.529.931	1.645.319	146.540
BALCHEM CORPORATION	171.053	900.736	1.155.682	643.705	102.437
AESTHETIC MEDICAL INTERNATIONAL HOLDINGS GROUP LIMITED	238.574	782.441	1.047.135	869.050	301.740
ADVANCED EMISSIONS SOLUTIONS, INC.	31.588	126.751	173.799	70.086	54.710

Società	TCA - Inventory	TNCA	TA	Revenue	EBIT
AMEDISYS, INC.	350.554	912.179	1.262.745	1.955.633	183.851
MISSION PRODUCE, INC.	161.137	483.410	689.449	883.301	106.507
AMGEN INC.	14.856.000	41.267.000	59.707.000	23.362.000	10.427.000
ADOBE INC.	6.494.925	14.267.475	20.762.400	11.171.297	3.359.000
AKAMAI TECHNOLOGIES, INC.	2.231.613	4.775.273	7.006.886	2.893.617	580.749
ANSYS, INC.	1.555.480	3.283.407	4.838.887	1.515.892	515.040
ACTIVISION BLIZZARD, INC.	7.260.000	12.553.000	19.845.000	6.489.000	1.739.000
ALLIED MOTION TECHNOLOGIES INC.	62.258	190.185	305.828	371.084	28.975
ABCAM PLC	135.800	274.900	446.700	259.900	56.700
AMC NETWORKS INC.	2.330.297	3.266.389	5.596.686	3.060.321	616.754
AMPHASTAR PHARMACEUTICALS INC	144.740	331.730	586.971	322.357	62.781
ALARM.COM HOLDINGS, INC.	209.506	314.125	557.799	502.363	62.071
ASPEN TECHNOLOGY, INC.	429.040	474.025	903.065	598.345	311.923
APPLE INC.	158.713.000	175.697.000	338.516.000	260.174.000	69.313.000
APPLOVIN CORPORATION	586.221	616.264	1.202.485	994.104	194.371
ALLEGRO MICROSYSTEMS, INC	236.131	385.213	752.261	724.311	100.535
Autoscope Technologies Corporation	8.707	9.695	19.183	14.732	1.830
AAON, INC. ¹	113.948	183.875	371.424	469.333	67.031
ATRION CORPORATION	90.245	129.693	262.031	155.066	40.529
ASML Holding NV	8.321.900	10.498.500	22.629.600	11.820.000	2.790.800
APPLIED MATERIALS, INC.	6.732.000	8.818.000	19.024.000	14.608.000	3.530.000
ABIOMED, INC.	596.255	377.149	1.054.346	769.432	263.360
ARENA PHARMACEUTICALS, INC.	759.934	414.189	1.174.123	806.431	482.762
ALIGN TECHNOLOGY, INC.	1.521.368	867.283	2.500.702	2.406.796	555.123
ABSOLUTE SOFTWARE CORPORATION	68.381	34.929	103.311	98.909	11.110
ARRAY TECHNOLOGIES, INC.	471.660	303.897	923.581	647.899	83.409

Tabella 35 Valori di bilancio utilizzati per il calcolo dei regressori. Campione di verifica statunitense. Fonte: Database Refinitiv.

3.4.3 Campione europeo

I dati sono presentati nella tabella seguente, in migliaia di euro.

Società	TCA	TA	TNCA	Revenues	EBIT
Amplifon	486.560	2.851.340	2.364.780	1.732.060	202.360
Askoll Eva	18.550	35.500	16.950	17.500	-9.070
Atlantia	10.690.000	70.928.000	60.238.000	12.615.000	1.757.000
BHB Brauholding Bayern Mitte AG	4.770	14.440	9.670	17.290	470
Bialetti	73.340	162.970	89.630	139.960	370
Biomass Energy Project SA	4.370	6.800	2.430	490	140
CHL	11.520	26.990	15.470	21.990	-3.010
Cineworld Group PLC	400.090	11.106.590	10.706.500	3.904.510	695.800
CNH Industrial	30.760.900	43.873.290	13.112.390	25.040.650	1.730.790
Codere sa	312.960	1.650.520	1.337.560	1.370.510	84.660
Cranswick PLC	316.570	517.430	200.860	1.629.730	101.610
Deoleo sa	265.650	847.080	581.430	561.950	20.840
Distrib. Internacional de Alimentacion SA	871.270	3.319.430	2.448.160	6.870.440	-452.770
Emmi AG	1.192.030	2.195.850	1.003.820	3.142.930	194.550
ENEL	36.703.000	171.426.000	134.723.000	77.366.000	11.215.000
Energy Lab	13.190	17.770	4.580	5.230	-3.330
ENI	34.909.000	123.440.000	88.531.000	69.881.000	8.813.000
ePrice	24.970	49.000	24.030	130.980	-25.340
Europcar mobility group	5.386.160	8.295.280	2.909.120	3.022.390	307.030
Ferrari	2.641.200	5.446.370	2.805.170	3.766.620	913.920
FRoSTA AG	190.440	321.410	130.970	523.390	20.030
Gel SpA	13.270	20.530	7.260	17.110	550
General Energy PLC	491.610	1.751.650	1.260.040	337.040	144.840
Giglio Group	29.140	51.420	22.280	37.740	790
Haya Real Estate SAU	217.720	595.000	377.280	290.610	-29.290
Intelsat SA	1.046.870	10.530.220	9.483.350	1.842.010	688.170
Ki Group	7.690	12.100	4.410	27.810	-2.790
Kri Kri Milk Industry SA	48.790	107.340	58.550	112.900	17.850
Lecta SA	484.320	1.154.960	670.640	1.382.840	35.530
Lotus Bakeries NV	171.510	812.630	641.120	612.740	102.890
Mallinckrodt PLC	1.638.200	9.251.760	7.613.560	2.825.140	215.290
Mobi Spa	88.570	534.730	446.160	195.610	-143.900
National Milk Records PLC	6.460	11.750	5.290	25.870	2.570
New Nordic Healthbrands AB	18.070	18.790	720	42.790	3.430
Noble Holding Corporation PLC	376.540	7.390.270	7.013.730	1.113.410	-194.680
Nostrum Oil & Gas PLC	154.470	754.950	600.480	287.840	48.340
NWF Group PLC	85.740	168.650	82.910	762.620	11.360
Pacific Drilling SA	348.430	2.012.990	1.664.560	205.320	-205.320
PGS ASA	380.110	2.053.250	1.673.140	831.710	64.780
Pharos Energy PLC	104.460	765.390	660.930	169.510	33.950
Pizzaexpress Financing 1 PLC	106.930	1.361.410	1.254.480	613.840	43.160
Purcari Wineries PCL	30.880	74.040	43.160	41.620	11.540
RAK Petroleum PLC	1.059.410	3.202.580	2.143.170	867.990	207.030
SAS AB	1.090.280	3.161.070	2.070.790	4.378.540	110.910
Sirio	23.270	130.110	106.840	72.230	-1.550
STMicroelectronics	5.596.780	11.332.730	5.735.950	8.538.700	1.102.630

Società	TCA	TA	TNCA	Revenues	EBIT
Stock Spirits Group PLC	221.120	684.840	463.720	312.420	55.130
Subsea 7 SA	1.434.970	5.009.450	3.574.480	3.267.330	131.710
Tate & Lyle PLC	1.273.240	1.971.430	698.190	3.124.280	334.540
Technicolor SA	1.127.000	3.210.000	2.083.000	3.800.000	-15.000
TechnipFMC PLC	10.603.380	20.980.180	10.376.800	11.982.500	948.050
Teqnon AB	25.270	43.160	17.890	54.790	3.690
Terna	2.986.200	18.003.700	15.017.500	2.287.900	1.139.500
TIM	9.461.000	70.104.000	60.643.000	17.974.000	3.306.000
Transocean Ltd	3.247.100	21.503.110	18.256.010	2.759.260	-89.350
Treant PLC	110.670	146.440	35.770	127.570	15.030
Vestel Elektronik Sanayi ve Ticaret AS	1.441.090	2.917.450	1.476.360	2.706.570	246.690
Vimi Fasteners	27.680	67.420	39.740	46.300	-150
Visibilia	4.990	9.500	4.510	4.590	-1.310
Wynnstay Group PLC	140.600	51.060	-89.540	555.740	8.260
Zellis Holdings Ltd	59.630	736.690	677.060	160.930	-25.060

Tabella 36 Valori di bilancio utilizzati per il calcolo dei regressori. Campione di europeo. Fonte: Database Refinitiv.

Conclusioni

Tramite questo elaborato si è cercato di fare chiarezza su uno degli elementi fondamentali per la stabilità di un sistema economico, ossia il rischio di credito.

Siamo partiti dalla definizione di tale rischio illustrandone anche le componenti e la relazione con il *rating*. Successivamente, abbiamo esposto i criteri di corretta costruzione di un modello per lo studio di fenomeni economici, analizzando anche i modelli più conosciuti e utilizzati. In particolare, sono stati trattati i modelli *rating system*, *credit scoring*, i modelli strutturali (come il Modello di Merton e il KMV) e i modelli in forma ridotta, per concludere poi con alcuni modelli sviluppati da istituti di credito come Credit Risk Plus di CreditSuisse e Creditmetrics di J. P. Morgan.

Infine, è stata proposta un'analisi su due campioni di aziende (oltre a uno di verifica) tramite il modello logit. Tale modello, come spiegato nel capitolo 3, si adatta molto bene alle variabili dicotomiche. Nel nostro caso, ricordiamo, la variabile dipendente può rappresentare solo due eventi: *default* aziendale verificatosi o meno. Sono allora state descritte le fasi di costruzione del modello e, dopo aver presentato i campioni, sono stati stimati i parametri e, infine, le probabilità di *default* associate a ciascuna società. È stato quindi evidenziato come il modello logit possa essere particolarmente di ausilio agli istituti finanziari per l'allocazione delle proprie risorse, permettendo delle stime con buona affidabilità nonostante la semplicità di calcoli. Bisogna tuttavia osservare anche i limiti di questo approccio. La costruzione di un *training set* adeguato risulta essere essenziale ai fini della corretta stima della *probability of default*, la scelta delle imprese da includere può influenzare pesantemente i risultati ottenuti, fino a renderli inutilizzabili.

Un ulteriore limite del modello proposto va ricercato nella mancata considerazione della variabile temporale. I dati utilizzati per la stima del modello fanno riferimento a uno specifico istante, ad esempio la chiusura del bilancio per l'anno fiscale 2019 o 2020, e non considerano la loro evoluzione nel tempo. Alcuni dei modelli presentati nel capitolo 2, invece, tengono conto anche di questo aspetto. Ad esempio, nei modelli *Rating system* che utilizzano l'approccio detto *credit migration* considerano le probabilità che ad una società con un determinato *rating* ne vengano assegnati altri diversi in un intervallo di tempo t , spesso rappresentando in una matrice tali probabilità. Un ulteriore esempio è fornito dal modello di Merton che inserisce anche la variabile temporale t tra le diverse considerate, così come il modello KMV.

L'analisi presentata vuole cogliere un importante aspetto, ossia verificare l'impatto del cambiamento del contesto economico dovuto alla diffusione del Covid-19. Infatti, il modello è stato stimato con dati di bilancio antecedenti agli effetti della pandemia, prendendo dati relativi a bilanci chiusi entro il 31.12.2019. Invece, il campione di controllo è stato costruito tramite dati di bilancio successivi ad essi ossia con bilanci relativi all'anno fiscale 2020. Notiamo che il modello continua ad operare con un buon grado di adattamento e proviamo a fornirne una spiegazione. Gli indici che sono risultati maggiormente correlati con l'evento *default*

sono due: $\frac{\text{Totale Attivo a lungo}}{\text{Totale Attivo}}$ e $\frac{\text{EBIT}}{\text{Totale Attivo}}$. Questi indici colgono degli aspetti essenziali sia prima sia dopo la pandemia. È chiaro che quest'ultima ne ha amplificato notevolmente l'importanza. Il primo, indicando il grado di rigidità di una impresa, ci induce a pensare che le aziende più rigide abbiano seguito la scelta di internalizzare le attività produttive aumentando il livello delle immobilizzazioni. In questo modo, hanno puntato maggiormente sullo sfruttamento della leva operativa. Durante il *lockdown*, tuttavia, hanno subito pesanti riduzioni dei ricavi e, proprio a causa della leva operativa, sono andati incontro a perdite maggiori rispetto alle aziende più flessibili. Il secondo indice, invece, mostra l'efficienza nella produzione del reddito rispetto agli impieghi della società. In un contesto di *stress* come quello pandemico, l'efficienza nella produzione del reddito è sicuramente un fattore che ha giocato un ruolo chiave nella possibilità di sopravvivenza delle imprese.

Bibliografia

- Altman E.I., *"Financial ratios, discriminant analysis and the prediction of corporate bankruptcy"*, Journal of Finance 23, 589 – 609 (1968);
- Amman M., *"Credit Risk Valuation, Methods, Models, and Applications"*, Springer Verlag Berlin (2001);
- Banca d'Italia, *"Modelli per la gestione del rischio di credito, I ratings interni"* (2000);
- Banca d'Italia, *"Rapporto sulla stabilità finanziaria"* (2011);
- Banca d'Italia, *"Metodo dei rating interni per il calcolo del requisito patrimoniale a fronte del rischio di credito"* (2006);
- Belli G., *"Valutazione delle imprese e merito creditizio alla luce del Nuovo Accordo di Basilea"*, Università degli studi La Sapienza (2003-2004);
- Bielecki, T., Rutkowski, M., *"Credit Risk"* (2010);
- Boffelli G., *"Valutare l'impresa"*, Maggioli, Rimini (2002);
- Caramiello C., Di Lazzaro F., Fiori G., *"Indici di bilancio, strumenti per l'analisi della gestione aziendale"*, Giuffrè editore, Milano (2003);
- Credit Suisse Financial Products, *"Credit Risk Management Framework"* (1997);
- De Laurentis G., *"Il credito alle imprese dopo la crisi"*, Bancaria Editrice, Roma (2011);
- Gordy, M. B., *"A Risk-factor Model Foundation for Ratings-Based Capital Rules"*, Journal of Financial Intermediation 12, no. 3 (2003);
- Gai Lorenzo, *"Lineamenti di gestione bancaria"*, Franco Angeli Editore (2011);
- Hull, J. C., *"Risk Management e Financial Institution"*, John Wiley & Sons (2018);
- Knight, F., *"Risk Uncertainty and Profit"* (1921)
- Kulkarni A., Mishra A. K. & Thakker J., *"How good is Merton Model at Assessing Credit Risk? Evidence From India"*, Second Singapore International Conference on Finance (2008);
- McNeil, A., Frey, R., & Embrechts, P. *"Quantitative Risk Management"*. Princeton: Princeton University Press (2015);
- Monti A. C., *"Introduzione alla statistica"*, McGraw-Hill Education, Napoli (2008);

- Montrone A., *“Il sistema delle analisi di bilancio per la valutazione dell’impresa”*, Franco Angeli, Milano (2005);
- Nardon M., *“Un’introduzione al rischio di credito”*, Università Ca’ Foscari, Venezia (2004);
- Palomba G., *“Modelli a variabili dipendenti qualitative”* (2008);
- Polidoro M. F., *“Analisi di bilancio e decisioni strategiche”*, editore Maggioli, Rimini (2011);
- Resti A., *“Misurare e gestire il rischio di credito nelle banche: una guida metodologica”*, Fondo interbancario di Tutela dei depositi, Roma (2001);
- Resti A. & Sironi A., *“Rischio e valore nelle banche”*, Egea, Milano (2008);
- Resti A., *“La gestione del rischio di credito con modelli di derivazione attuariale: il caso di CreditRisk+”*, Working Papers, n. 4, Fondo Interbancario di Tutela dei Depositi, Roma (2000);
- Rossi M., *“Perdita attesa, perdita inattesa e diversificazione”* (2003);
- Senese, V., *“Regressione multipla e regressione logistica: concetti introduttivi ed esempi”*, I edizione, (2016).

RIASSUNTO

Il rischio di credito è un tema sempre attuale. Un'esposizione eccessiva nei confronti di debitori che versino in situazioni di emergenza finanziaria può portare gli istituti finanziari verso il *default* e generare ricadute sull'intera economia. È facile allora immaginare quanto sia delicato il tema della stima del rischio di credito delle controparti.

Negli ultimi venti anni abbiamo avuto modo di assistere a due crisi rilevanti. La prima, iniziata nel 2007, ha dimostrato quali possano essere i danni causati dall'eccessiva erogazione di prestiti senza prestare adeguata attenzione all'analisi della probabilità di *default* della controparte. La seconda, tuttora in corso, è stata causata dalla pandemia e dal conseguente *lockdown* imposto da molti governi per arginare e combattere la diffusione del virus. Ciò ha tuttavia portato a una riduzione dei ricavi delle imprese e a una maggiore necessità di approfondite analisi prima della concessione di eventuali crediti da parte degli intermediari finanziari.

Per rispondere a queste esigenze, è necessario dotarsi di strumenti sempre più raffinati, modelli in grado di stimare efficacemente il livello di rischio di una impresa, lasciando poi ovviamente al *management* il compito di stabilire il livello di rischio ritenuto accettabile.

Lo scopo del presente lavoro è quello di evidenziare l'importanza del rischio di credito per l'intera economia e costruire un modello che possa provare a stimare la probabilità di *default* di un campione di imprese.

Il lavoro si articola in tre capitoli. Il capitolo 1 si apre con la descrizione del concetto di rischio, fondamentale e ineludibile componente di qualsiasi attività d'impresa. Possiamo pensare il rischio di credito come il rischio che una variazione non attesa del merito creditizio di una controparte generi una variazione del valore della posizione creditoria.

Il perimetro del rischio di credito è più ampio di quanto possa apparire a una prima analisi. Non ci si limita infatti alla sola possibilità di insolvenza della controparte ma anche il solo deterioramento del merito creditizio della stessa deve essere considerato a tutti gli effetti una manifestazione di tale categoria di rischio.

La presenza del rischio di credito espone gli investitori al pericolo che essi non possano trarre i benefici economici desiderati. Per valutare attentamente il rischio e poter prendere le dovute cautele, oltre all'adempimento delle normative correlate, i prestatori, come ad esempio le banche, devono tenere in considerazione le componenti che costituiscono il rischio di credito. Esse sono la *probability of default* (PD), la *loss given default* (LGD) e l'*exposure at default* (EAD). Queste grandezze, opportunamente studiate, permettono di preventivare quali potranno essere le perdite collegate all'investimento che possono essere di due tipi: perdite attese o *Expected Loss* (EL) e perdite inattese o *Unexpected Loss* (UL).

Si prosegue con la definizione del *rating*, sintetica valutazione della solvibilità dei titoli e delle imprese elaborata da parte di agenzie indipendenti dette, appunto, agenzie di *rating*. Tale valutazione costituisce un

aiuto di facile reperibilità e comprensibilità relativo al livello di rischio che potenzialmente si decide di sostenere. Infine, si evidenziano le conseguenze economiche che una mancata attenzione nei confronti del rischio di credito può generare sia dal punto di vista degli *stakeholders* delle aziende sia da quello della società in senso lato.

Il capitolo 2 si apre con una panoramica delle *best practices* in tema di sviluppo e convalida dei modelli per lo studio di fenomeni economici. Essendo la realtà eccessivamente complessa non è possibile, o talvolta non è conveniente, cercare di studiare e analizzare i fenomeni economici e i dati a nostra disposizione così come essi si presentano. Dobbiamo quindi trovare un modo per semplificare la realtà senza però perdere di vista gli aspetti chiave dei fenomeni analizzati. Per questo motivo matematici, fisici ed economisti si impegnano nella costruzione di modelli, ossia semplificazioni della realtà, che riescano a cogliere gli aspetti salienti dei fenomeni studiati ma che allo stesso tempo non siano eccessivamente complicati in maniera tale che non risultino di difficile comprensione e applicazione.

In seguito, ci si concentra sui modelli utilizzabili per la stima della *probability of default*. In particolare, si evidenziano differenti approcci all'analisi come il *credit rating migration*, il *credit scoring*, il modello di Merton, il modello KMV e i modelli in forma ridotta. Tali modelli sono brevemente descritti di seguito:

- a. i modelli di *rating system* sono dei modelli utilizzati dalle agenzie di *rating* con lo scopo di analizzare la qualità del credito di specifiche aziende basandosi su dati sia quantitativi sia qualitativi. Il giudizio espresso dalle agenzie di *rating* non è un giudizio che viene modificato a ogni minimo cambio di condizione dell'azienda esaminata. L'obiettivo è più quello di riconoscere il *trend* generale. Tuttavia, è possibile che delle banche commissionino un giudizio esterno e indipendente sul credito relativo a una specifica azienda ottenendo un parere relativo alla stima della LGD e della PD;
- b. i modelli *credit scoring* sono invece basati sull'idea dell'attribuzione di uno *score* alle aziende, permettendone l'associazione a una determinata affidabilità creditizia. Uno dei modelli più conosciuti è il modello di Altman, tramite cui è possibile valutare il livello di solvibilità di un'azienda;
- c. il modello di Merton appartiene alla famiglia di modelli che adottano un approccio di natura "strutturale" (o *firm's value models*), cercando di descrivere la struttura della società analizzata con l'obiettivo di ricondurre in modo diretto l'eventuale evento di *default* a un avvenimento interno. Questo modello si basa sull'intuizione di considerare una società insolvente se il valore delle attività sia inferiore a quello delle passività. L'insolvenza ha luogo al momento del primo pagamento dovuto dall'impresa dopo che il valore delle passività abbia superato quello delle attività, rendendo più conveniente per gli azionisti lasciare la società ai creditori. Il *payoff* dell'azionista della società viene assimilato a quello di un investitore acquirente di un'opzione *call* di tipo europeo sul valore dell'attivo

della società con scadenza e prezzo di esercizio pari alla scadenza e al valore contabile del debito della società stessa:

1. se il valore dell'attivo A dell'impresa alla scadenza T è maggiore del valore del debito emesso D , allora i titoli valgono il loro valore nominale e la società risulta essere solvente;
2. al contrario, se il valore dell'attivo A dell'impresa alla scadenza T è minore del debito emesso, la società viene dichiarata insolvente e i creditori hanno il diritto di rivalersi sul suo attivo per la quota di credito loro spettante.

Per quanto riguarda i creditori, la loro posizione è assimilabile alla vendita di una ipotetica opzione *put* di tipo europeo;

- d. il modello KMV parte dal presupposto che, in generale, ai fini della determinazione della *probability of default* necessitiamo di tre tipi di informazioni: dati di carattere finanziario, prezzi di mercato delle azioni e dei debiti e, infine, stime soggettive sulle prospettive future dell'impresa. È facile comprendere, quindi, che il modello di cui abbiamo bisogno deve essere in grado di coniugare informazioni derivanti da dati storici e prospettici. Il modello sviluppato da KMV, in accordo con quanto detto, calcola la PD facendo riferimento sia ai dati di carattere finanziario sia ai prezzi di mercato delle azioni. Il primo passo è quello di stimare l'*asset value* e l'*asset volatility* della società. Successivamente, si calcola la *distance to default*, grandezza che identifica la "distanza" espressa in numero di deviazioni standard dal valore di mercato delle attività dal default point. Infine, otteniamo una stima della PD basata sulla cosiddetta "tabella di frequenza" ossia una tabella che assegna a degli intervalli di valori della DD un *rating* delle società S&P, ottenuta da KMV tramite un'analisi di tipo empirico;
- e. i modelli in forma ridotta si prefiggono come scopo quello di calcolare la probabilità di *default* direttamente dai dati di mercato. Si cerca quindi di porre tutta l'attenzione sull'evento *default* in quanto tale, considerandolo come un dato esogeno al modello. Per realizzare tale tipo di analisi viene utilizzata la cosiddetta "intensità di *default*", indicata con la lettera λ . Essa rappresenta la probabilità condizionale che l'impresa osservata non riesca ad adempiere alle proprie obbligazioni in un dato intervallo temporale successivo all'osservazione e sotto l'ipotesi che non si sia ancora verificata l'insolvenza.

Infine, nel capitolo 3 si presenta il lavoro di analisi svolto, ponendosi come obiettivo quello di costruire un modello in grado di stimare la probabilità di *default* delle aziende partendo da dati di bilancio ritenuti significativi. Inoltre, consapevoli dell'impatto avuto dalla pandemia sull'economia globale, vogliamo testare se il modello continui ad avere una capacità previsionale elevata oppure se, al contrario, il nuovo contesto

economico possa influenzare la significatività dei regressori selezionati. Tutti i dati dei campioni sono stati raccolti utilizzando il *database Refinitiv*.

Tramite delle regressioni logistiche costruiamo allora il modello logit, tipologia di modello particolarmente utilizzata per analizzare la relazione tra una variabile dipendente Y (di natura qualitativa) e n variabili X_1, \dots, X_n che possono essere di natura sia qualitativa sia quantitativa. Relativamente alla variabile Y è bene notare che le modalità che può assumere sono mutuamente esclusive. Nel nostro caso, la variabile dipendente Y potrà assumere solo due valori che indicano il verificarsi dell'evento *default* o meno. Abbiamo allora che:

$$Y = \begin{cases} 1, & \text{se si è verificato l'evento default} \\ 0, & \text{altrimenti} \end{cases}$$

Il primo campione raccolto è quello statunitense, formato da 58 aziende di cui 29 presentano $Y=1$, rappresentanti quindi la metà dello stesso. Consideriamo con $Y=1$ le aziende in cui si sia già verificato il *default* al 31.12.2019 oppure che abbiano avuto procedure fallimentari aperte in corso alla stessa data. Il campione di verifica è invece formato da 52 società di cui 22 presentano $Y=1$. Infine, il campione europeo è formato da 61 aziende di cui 22 presentano $Y=1$.

Riguardo i regressori relativi alle regressioni logistiche, è necessario individuare delle variabili X_k che forniscano informazioni rilevanti per tutte le imprese e che siano ovviamente correlate con la probabilità di *default*.

Nella presente analisi, la scelta degli indici è ricaduta sui seguenti:

1. X_1 : *quick ratio*: questo indice segnala il livello di liquidità dell'azienda in relazione alle sue passività correnti, indicando quindi se l'azienda sarà in grado di far fronte ai debiti con scadenza breve;
2. X_2 : EBIT/Totale Attivo: tale indice evidenzia quanto l'azienda sia in grado di sfruttare le sue risorse per la generazione dell'EBIT;
3. X_3 : Totale Attivo a lungo/Totale Attivo: tale valore indica quando l'azienda sia rigida. Tanto più questo valore è alto tanto più meno l'azienda sarà reattiva nella risposta a stimoli esterni. Tuttavia, non è sempre vero che una azienda più elastica risponda sempre bene a stimoli esterni. Se l'attivo circolante è costituito in prevalenza da crediti verso clienti e magazzino si corre il rischio di non poter liquidare nel breve periodo il valore del circolante;
4. X_4 : EBIT / Ricavi: questo valore, detto anche *EBIT margin*, indica la percentuale di fatturato che si è trasformata in EBIT, indicando l'efficienza nella gestione operativa;

5. X_5 : Ricavi / Totale Attivo: questo valore indica quanto l'azienda riesca a sfruttare i propri impieghi per generare ricavi. In altri termini, l'indice evidenzia il ricavo medio per unità di impieghi.

Possiamo adesso descrivere la fase di costruzione del modello. La procedura di analisi effettuata inizia determinando il valore dell'intercetta e dei coefficienti del modello tramite regressione logistica. Una volta ottenuti questi valori è possibile applicare il modello e determinare la probabilità di *default* π associata a ciascuna società del campione. Tuttavia, sarà necessario verificare la significatività dei regressori prima del calcolo di π . Per fare ciò, seguiamo una logica di tipo "backward", ossia partiamo dal modello con tutti i cinque i regressori, calcoliamo il *p-value* per ogni coefficiente ed eliminiamo il regressore che presenta il *p-value* più elevato, costruendo così un nuovo modello. Tale procedura si arresta solamente quando tutti i coefficienti presentano un valore associato al *p-value* inferiore alla soglia prefissata $\alpha = 5\%$.

Seguendo la procedura sopra descritta otteniamo il seguente modello:

$$\pi_i = \frac{\exp(-5,296 - 13,671X_1 + 9,636X_2)}{1 + \exp(-5,296 - 13,671X_1 + 9,636X_2)}$$

Dove X_1 indica il regressore $\frac{EBIT}{Totale\ Attivo}$ e X_2 rappresenta, invece, il regressore $\frac{Totale\ Attivo\ a\ lungo}{Totale\ Attivo}$.

Testiamo adesso l'efficacia del modello verificando quante aziende riesce a classificare correttamente. Tale verifica verrà eseguita confrontando il valore della variabile indipendente Y per ciascuna azienda (valore ovviamente noto *a priori*) con la probabilità di *default* π stimata dal modello. In seguito, viene stabilita una soglia di *cut off* che divide i valori delle probabilità in due parti. Per farlo, consideriamo gli errori di classificazione (ossia aziende con Y=0 classificate come in *default* e, viceversa, aziende con Y=1 classificate come sane) rispetto al totale delle osservazioni e minimizziamo tale valore. Otteniamo una soglia di *cut off* pari a circa il 44%. Analizzando gli errori e la loro tipologia di classificazione possiamo vedere che:

		Previsto		Totale
		Y=0	Y=1	
Effettivo	Y=0	24	5	29
	Y=1	3	26	29
Totale		27	31	58

Tabella 1 Analisi degli errori per tipologia. Fonte: propria elaborazione.

Notiamo che 50 aziende sono correttamente classificate. Invece, 5 aziende sane sono classificate come insolventi mentre 3 insolventi, al contrario, sono classificate come sane. Utilizzando quindi la soglia del *cut off* che minimizza gli errori otteniamo un'accuratezza del modello dell'86,21%, che è la frazione delle aziende correttamente classificate rispetto al totale.

Per ulteriore conferma dei risultati ottenuti, il modello è stato testato anche su un campione di verifica composto da 52 società, di cui 22 presentano $Y=1$. Anche in questo caso il campione è ampio e riferito a diversi settori in modo tale da cercare di imitare il più possibile l'andamento del mercato.

Ricordiamo che il campione di verifica vuole testare la significatività del modello dopo gli effetti della pandemia. Sono stati infatti selezionati bilancio riferiti relativi all'anno fiscale 2020. Di seguito si riporta il campione di verifica insieme all'applicazione del modello precedentemente illustrato. La soglia di *cut off* usata in questa prima fase è la stessa calcolata per il *training set*, ossia il 44%, in modo tale da mantenere coerenza per il confronto della correttezza di classificazione nei due campioni.

Utilizzando la soglia di *cut off* selezionata tramite il modello precedente abbiamo 4 aziende sane classificate con insolventi e, al contrario, 11 aziende insolventi classificate come sane. Questa classificazione ci porta a una percentuale di correttezza della classificazione pari a circa il 71,15%. Tale percentuale risulta inferiore a quella calcolata nel caso precedente ma è comunque significativa e accettabile.

Possiamo rappresentare graficamente tale risultato come segue:

		Previsto		Totale
		Y=0	Y=1	
Effettivo	Y=0	26	4	30
	Y=1	11	11	22
Totale		37	15	52

Tabella 2 Analisi degli errori per tipologia sul campione di verifica. Fonte: propria elaborazione.

Possiamo anche calcolare la soglia di *cut off* che minimizzi gli errori per il campione di verifica. Essa è uguale al 55% e permette di ottenere una corretta classificazione per il 75% delle imprese. Quindi, senza eccessiva influenza da parte della soglia di *cut off* selezionata, il modello presenta un buon adattamento sia per il *training set* che considerava valori di bilancio pre-covid sia per il campione di verifica che testa valori di bilancio post-covid.

In quest'ultima sezione dell'analisi è stato eseguito un ulteriore passaggio costruendo un campione di aziende quotate sui mercati europei. I dati utilizzati sono relativi a bilanci chiusi entro l'anno fiscale 2019. Tale campione comprende 61 aziende di cui 22 in *default*. Seguendo la medesima logica precedente per la costruzione del modello otteniamo la seguente equazione:

$$\pi_i = \frac{\exp(-4,700 - 5,184X_1 + 6,658X_2)}{1 + \exp(-4,700 - 5,184X_1 + 6,658X_2)}$$

Con la soglia di *cut off* che minimizza gli errori, pari al 57,82%, otteniamo che:

		Previsto		Totale
		Y=0	Y=1	
Effettivo	Y=0	37	2	39
	Y=1	10	12	22

	Previsto		Totale
	Y=0	Y=1	
Totale	47	14	61

Tabella 3 Analisi degli errori per tipologia. Fonte: propria elaborazione.

Possiamo osservare che 49 aziende sono state classificate correttamente. Invece, 2 aziende sane sono state classificate come insolventi e 10 insolventi sono state considerate sane. Possiamo concludere che anche il tentativo realizzato su un campione europeo conferma i risultati ottenuti tramite il modello statunitense.

Tramite questo elaborato si è cercato di fare chiarezza su uno degli elementi fondamentali per la stabilità di un sistema economico, ossia il rischio di credito, anche proponendo un modello per la stima della *probability of default*. È bene notare alcuni aspetti che possono costituire dei limiti all'affidabilità del modello stesso. In particolare, si fa riferimento alla costruzione di un *training set* adeguato, essenziale ai fini della corretta stima della *probability of default*, la scelta delle imprese da includere può influenzare pesantemente i risultati ottenuti, fino a renderli inutilizzabili.

Un ulteriore limite del modello proposto va ricercato nella mancata considerazione della variabile temporale. I dati utilizzati per la stima del modello fanno riferimento a uno specifico istante, ad esempio la chiusura del bilancio per l'anno fiscale 2019 o 2020, e non considerano la loro evoluzione nel tempo. Alcuni dei modelli presentati nel capitolo 2, invece, tengono conto anche di questo aspetto. Ad esempio, nei modelli *Rating system* che utilizzano l'approccio detto *credit migration* considerano le probabilità che ad una società con un determinato *rating* ne vengano assegnati altri diversi in un intervallo di tempo t , spesso rappresentando in una matrice tali probabilità. Un ulteriore esempio è fornito dal modello di Merton che inserisce anche la variabile temporale t tra le diverse considerate, così come il modello KMV.

L'analisi presentata vuole cogliere un importante aspetto, ossia verificare l'impatto del cambiamento del contesto economico dovuto alla diffusione del Covid-19. Infatti, il modello è stato stimato con dati di bilancio antecedenti agli effetti della pandemia, prendendo dati relativi a bilanci chiusi entro il 31.12.2019. Invece, il campione di controllo è stato costruito tramite dati di bilancio successivi ad essi ossia con bilanci relativi all'anno fiscale 2020. Notiamo che il modello continua ad operare con un buon grado di adattamento e proviamo a fornirne una spiegazione. Gli indici che sono risultati maggiormente correlati con l'evento *default* sono due: $\frac{\text{Totale Attivo a lungo}}{\text{Totale Attivo}}$ e $\frac{\text{EBIT}}{\text{Totale Attivo}}$. Questi indici colgono degli aspetti essenziali sia prima sia dopo la pandemia. È chiaro che quest'ultima ne ha amplificato notevolmente l'importanza. Il primo, indicando il grado di rigidità di una impresa, ci induce a pensare che le aziende più rigide abbiano seguito la scelta di internalizzare le attività produttive aumentando il livello delle immobilizzazioni. In questo modo, hanno puntato maggiormente sullo sfruttamento della leva operativa. Durante il *lockdown*, tuttavia, hanno subito pesanti riduzioni dei ricavi e, proprio a causa della leva operativa, sono andati incontro a perdite maggiori rispetto alle aziende più flessibili. Il secondo indice, invece, mostra l'efficienza nella produzione del reddito

rispetto agli impieghi della società. In un contesto di *stress* come quello pandemico, l'efficienza nella produzione del reddito è sicuramente un fattore che ha giocato un ruolo chiave nella possibilità di sopravvivenza delle imprese.