

UNIVERSITA' DEGLI STUDI DI MESSINA  
FACOLTA' DI ECONOMIA

DOTTORATO DI RICERCA in  
STRUMENTI MATEMATICI PER  
L'ECONOMIA E LA FINANZA

XV CICLO

# OPERATORI DI AGGREGAZIONE NELLA CUSTOMER SATISFACTION

Massimo Riccardo Costanzo

TESI DI DOTTORATO

RELATORE:  
Chiar.mo Prof. Salvatore Greco

COORDINATORE  
Chiar.ma Prof.essa Maria T. Calapso

Anno Accademico 2001-2002

*Dedicata  
a mio padre e mia madre per l'amore,  
l'immensa comprensione  
e la grande fiducia.*

***Ringraziamenti:***

Nella preparazione della mia tesi ho beneficiato dell'aiuto di molte persone. In particolare, desidero esprimere particolare gratitudine al Prof. Salvatore Greco, mio relatore, per la Sua guida e i Suoi consigli. La Sua assistenza è stata indispensabile per il completamento di questa Tesi.

Ringrazio l'Università tutta per avermi dato la possibilità di effettuare questo affascinante corso di studi. Altresì, intendo ringraziare i miei familiari, per aver sopportato la mia distanza dalla "vita comune" nelle serate e nei fine-settimana in cui ho, anche, scritto la mia Tesi.

Massimo Riccardo Costanzo  
Ottobre 2002  
Catania, Italia

Massimo Riccardo Costanzo  
Università di Catania – Facoltà di Economia  
DIP. DI ECONOMIA E METODI QUANTITATIVI  
Corso Italia 55  
95128 – Catania  
Tel. 095-375344 int. 202 (università)  
Tel. 347- 0362008 (cellulare)  
e-mail:[costanzomassimo@ctonline.it](mailto:costanzomassimo@ctonline.it)  
e-mail:[costanzomassimo@hotmail.com](mailto:costanzomassimo@hotmail.com)

Titolo in inglese:

**AGGREGATION OPERATORS FOR CUSTOMER SATISFACTION © 2002**

## INDICE

### Introduzione

<b>1 CUSTOMER SATISFACTION</b>	<b>pag. 13</b>
1.1 DEFINIZIONE ED ASPETTI AD ESSA COLLEGATI	
1.2 PSICOLOGIA DELLA CONOSCENZA	
1.3 ASPETTI MOTIVAZIONALI	
1.4 SISTEMA DI MISURAZIONE E SISTEMA OPERATIVO MANAGERIALE	
1.5 OBIETTIVI DEL MANAGEMENT	
1.6 LEGAMI TRA CSE QUALITA'	
<b>2 INDICI DI CS</b>	<b>pag. 29</b>
2.1 ACSI, ECSI, SCS	
2.2 CUSTOMER RETENTION	
<b>3 CUSTOMER RELATIONSHIP MANAGEMENT</b>	<b>pag. 37</b>
3.1. I PRINCIPI DEL CRM: PROCESSI E COMPONENTI	
3.2. CUSTOMER BASED VIEW (CONOSCENZA)	
3.3. LEGAMI TRA CUSTOMER VALUE- CUSTOMER SATISFACTION- CUSTOMER LOYALTY – PROFITTO	
<b>4 MODELLI FORMATIVI (DESCRITTIVI)</b>	<b>pag. 63</b>
4.1 INTRODUZIONE	
4.2 IL COSTRUTTO CONCETTUALE E IL PROCESSO DI MISURAZIONE	
4.3 IL PROBLEMA DELLE SCALE ORDINALI	
4.4 LE TECNICHE “ATTRIBUTE BASED”: DESCRIZIONE METODOLOGICA DEL “MODELLO DEI GAP” DEL PROF. PARASURAMAN: VANTAGGI E SVANTAGGI (marketing relazionale, gap di valore, gap di percezione, gap di sintonia, gap di progettazione e di realizzazione, gap di allineamento, coinvolgimento, consonanza)	
4.5 MODELLI FORMATIVI O COMPOSITIVI SERVQUAL (QUALITA' DEL SERVIZIO) SERVPERF	
4.6 ANALISI DISCRIMINANTE LINEARE	
<b>5 MODELLI STRUTTURALI (DIPENDENZA CONOSCITIVA)</b>	<b>pag. 84</b>
5.1. ANALISI FATTORIALE E MODELLI STRUTTURALI LINEARI	
5.2. L'ANALISI LISREL	
5.3. L'ACIMO-PLSPER LA VALUTAZIONE DELLA CUSTOMER SATISFACTION	
5.4. L'UTILIZZO DELLE RETI NEURALI NELLA CS	
<b>6 MODELLI DI STATISTICHE MULTIVARIATE (INFERENZIALI)</b>	<b>pag. 98</b>
6.1. MODELLI DIRETTI ESPLICATIVI O DECOMPOSITIVI: MODELLI DI REGRESSIONE	
6.2. MODELLO DI REGRESSIONE LOGISTICA LOGIT	
6.3. REGRESSIONE LINEARE MULTIPLA A RISPOSTA POLITOMICA	
6.4. L'ANALISI IN COMPONENTI PRINCIPALI (ACP)	
6.5. L'ANALISI IN COMPONENTI PRINCIPALI NON LINEARE	
6.6. REGRESSIONE LINEARE MULTIVARIATA	
6.7. ANALISI CORRELAZIONI CANONICHE	
6.8. ALBERI DECISIONALI ED ANALISI A CAMPIONE CHAID	
6.9. STIMATORI: OLS, ANOVA	

<b>7. MODELLI DI STATISTICHE MULTIVARIATE CON TECNICHE DI AGGREGAZIONE</b>	<b>pag.120</b>
7.1. MULTICRITERIA SATISFACTION ANALYSIS(MUSA): DATA MINING	
7.2. SISTEMI DI DISAGGREGAZIONE INTERATTIVA	
7.3. LA CLUSTERANALYSIS	
7.4. METODI GERARCHICI DI CLASSIFICAZIONE SFOCATA	
7.5. METODI NON GERARCHICI DI CLASSIFICAZIONE SFOCATA	
7.6. FUZZY CLUSTERANALYSIS	
7.7. CONJOINT ANALYSIS(ANALISI CONGIUNTA)	
<b>8. L'UTILIZZO DI OPERATORI DI AGGREGAZIONE FUZZY</b>	<b>pag.177</b>
8.1. DEFINIZIONE DI OPERATORI DI AGGREGAZIONE	
8.2. PROPRIETA' MATEMATICHE	
8.3. OPERATORI MATEMATICI	
8.4. LA LOGICA FUZZY	
8.5. T-NORMSE T-CONORMS	
8.6. MEDIA, MEDIANA, MINIMO, MASSIMO	
8.7. AGGREGAZIONE ORDINATA PESATA (OWA)	
8.8. GLI INTEGRALI FUZZY DI CHOQUET E SUGENO	
<b>9. IL METODO ROUGH SETS PER LA CUSTOMER SATISFACTION</b>	<b>pag.205</b>
9.1. L'APPROCCIO ROUGH SETS PER LA STIMA DELL'IMPORTANZA RELATIVA DI CIASCUN ATTRIBUTO	
9.2. CONFRONTO CON ALTRE METODOLOGIE	
<b>10. LA CUSTOMER SATISFACTION E L'UTILIZZO DI FUNZIONI DI UTILITA' NON ADDITIVE</b>	<b>pag.222</b>
10.1. INTRODUZIONE	
10.2. AI METODI TRADIZIONALI AI METODI DECISIONALI MULTICRITERIO	
10.3. LA METODOLOGIA	
10.4. PROBLEMA MULTICRITERIALE	
10.5. LE FASI DELLA RICERCA – I – L'INTEGRALE DI CHOQUET E DI SUGENO	
10.6. LE FASI DELLA RICERCA – II – L'INTEGRALE GERARCHICO	
10.7. LE FASI DELLA RICERCA – III - IL CUSTOMER PROFILING COME PROBLEMA DI CLASSIFICAZIONE	
<b>11. L'APPLICAZIONE</b>	<b>pag.254</b>
11.1. METODOLOGIE E ATTIVITÀ DI REALIZZAZIONE DEL SONDAGGIO	
11.2. IL CAMPIONE	
11.3. IL METODO MONTECARLO	
11.4. L'APPLICAZIONE – I – L'INTEGRALE DI CHOQUET E DI SUGENO	
11.5. L'APPLICAZIONE – II – L'INTEGRALE GERARCHICO	
11.6. L'APPLICAZIONE – III - IL CUSTOMER PROFILING COME PROBLEMA DI CLASSIFICAZIONE	
11.7. ALGORITMI GENETICI	

<b>ALLEGATI</b>	<b>pag.276</b>
-----------------	----------------

<b>CONCLUSIONI</b>	<b>pag.285</b>
--------------------	----------------

<b>BIBLIOGRAFIA</b>	<b>pag.286</b>
---------------------	----------------

## **INTRODUZIONE**

La soddisfazione del cliente (Customer Satisfaction) è un processo difficile ma rappresenta un obiettivo ambito perché per suo tramite è possibile saldare e rinsaldare le relazioni, costruire percorsi di fiducia e fedeltà tra le parti, generando valore reciproco. La misurazione della soddisfazione dei clienti rappresenta uno strumento importante al fine di valutare il lavoro svolto e la competitività delle proprie offerte rispetto a quanto proposto dai diretti concorrenti: un cliente soddisfatto sarà portato a riacquistare il prodotto/servizio offerto, sarà fedele e, considerato l'elevato costo relativo all'acquisizione di nuovi clienti rispetto a quello sostenuto per trattenerli, è facile immaginare l'importanza che un cliente soddisfatto ricopre nella vita di un'azienda.

La ricerca proposta risulta di grande interesse non solo ai fini scientifici, ma anche a carattere operativo, costituendo un reale e nuovo modello di supporto alle decisioni in ambito di valutazione di Customer Satisfaction.. Infatti, le problematiche affrontate in azienda, le alternative che il management deve fronteggiare e la complessità dell' ambiente esterno (istituzionale, economico, del mercato di riferimento, ecc.) impongono un approccio alla risoluzione dei problemi saldamente ancorato alla razionalità, alla quantificazione e valutazione delle realtà interne ed esterne all' azienda, alla misurazione oggettiva dei diversi elementi che concorrono a creare una situazione di vantaggio o di svantaggio competitivo.

Il principale oggetto della dissertazione riguarda le procedure di aggregazione in ambito di Customer Satisfaction.

Gli operatori di aggregazione sono utilizzati per ottenere un valore unico di un'alternativa rispetto a più criteri.

Il più importante contributo di questa tesi è la considerazione dell'importanza dell'interazione tra criteri.

La regola di composizione secondo cui il singolo consumatore aggrega l'utilità associata a ciascun attributo per ottenere il valore del prodotto/servizio, costituisce l'aspetto di primario interesse per chi si avvale di questa metodologia. Tale regola è generalmente nota come modello di preferenza.

Le tecniche classiche di misurazione della soddisfazione del cliente suppongono che i vari attributi sono indipendenti. Un significativo aspetto di aggregazione è la differenza di importanza dei criteri, che è usualmente modellato usando differenti pesi. L'utilizzo di operatori di aggregazione ponderati non è appropriato quando si considerano criteri interattivi. Infatti si tende a costruire criteri indipendenti, o criteri supposti tali causando alcuni difetti nella valutazione. Al fine di avere una rappresentazione di complessi fenomeni di interazione tra criteri (positiva o negativa sinergia), è utile sostituire al vettore dei pesi *un'insieme di funzioni non additive* che permettono di definire un peso non solo per ogni criterio, ma anche su ogni sottoinsieme di criteri.

Si propone un metodo che, a partire da alcune preferenze espresse su un insieme di azioni e da altre indicazioni circa l'importanza e l'interazione dei criteri, fissa i parametri di una funzione di utilità nella forma di un particolare integrale fuzzy, l'integrale di Choquet. Per la costruzione di questa funzione di utilità occorre inferire: 1) un insieme di pesi non additivi che costituiscono una misura fuzzy sull'insieme dei criteri considerati e che

possono essere interpretati come "importanza" delle coalizioni di criteri; 2) le funzioni di utilità marginale relative a ciascun criterio, che consentono di esprimere valutazioni con riferimento a differenti criteri su un' unica scala di valutazione. Dal punto di vista formale, il problema affrontato si presenta come un problema di programmazione non lineare e non differenziabile, ove i vincoli e la funzione obiettivo hanno una formulazione particolarmente complessa, dipendente anche dai valori assegnati dalle funzioni di utilità marginale sulla scala comune a tutti i criteri.

Per raggiungere l'obiettivo, anziché utilizzare le classiche metodologie statistiche, si è preso in considerazione l'utilizzo di funzioni di utilità non additive, nell'ambito dei cosiddetti integrali fuzzy, che permettono di modellizzare strutture di preferenza anche in presenza di interazione tra attributi.

La metodologia utilizzata consente la decomposizione di valutazioni globali in scale di utilità, corrispondenti a ciascuno degli attributi considerati, separate e comparabili, in modo tale che le valutazioni globali originarie possono essere correttamente ricostruite. Le ipotesi di base su cui poggia tale metodologia sono essenzialmente due: che il consumatore scelga tra prodotti/servizi alternativi in base al valore soggettivo da lui stesso assegnato ad ognuna di essi, che il valore di ciascun prodotto/servizio sia dato dalle combinazioni dei valori associati a ciascun livello degli attributi caratterizzanti il prodotto/servizio stesso.

Tale metodologia consente di rilevare i punti di forza e di debolezza del servizio esaminato. Sulla base dei dati ottenuti si può pianificare un miglioramento del servizio.

Sfruttando le peculiarità della logica Fuzzy, si ottiene una rilevazione della CS svincolata dalle tipiche scale numeriche. Negli approcci tradizionali i clienti sono chiamati ad esprimere le loro valutazioni su una scala numerica; ciò costringe l'intervistato ad operare scelte che spesso alterano il suo reale giudizio.

Utilizzando i metodi tradizionali per rilevare la CS risulta difficile tradurre in dati oggettivi l'informazione espressa verbalmente.

La Fuzzy consente di trattare le valutazioni verbali senza eliminarne l'ambiguità e privilegiando la significatività del risultato piuttosto che la precisione. La Fuzzy utilizza operatori d'aggregazione (OWA) in grado di comporre giudizi espressi su scale differenti da diversi valutatori, senza alterarne la vaghezza e l'ambiguità. Nei metodi di rilevazione più diffusi, quali il SERVQUAL, i pesi degli elementi valutati sono attribuiti dal valutatore col rischio di avere una classificazione poco affidabile; la Fuzzy consente, invece, di ricavare i pesi da attribuire ai singoli elementi direttamente dalla coerenza dei giudizi espressi, garantendo una maggiore veridicità delle priorità emerse.

Nella Tesi si precisa che un usuale operatore aritmetico non può aggregare valori ordinali. E', quindi, necessario ricorrere ad un aggregatore tale da riflettere la valutazione qualitativa. In questo contesto l'integrale di Sugeno appare un potenziale candidato a risolvere il problema. Infatti l'integrale di Sugeno è riconosciuto come la naturale controparte dell'integrale di Choquet nel caso di insiemi ordinali, dove la somma è sostituita dal *max* (nel caso di integrale discreto) e il prodotto dal *min*



Nel lavoro, pertanto, viene proposta un'applicazione con l'utilizzo dell'integrale di Sugeno, nell'ipotesi in cui gli attributi di valutazione sono espressi in termini qualitativi.

Un altro risultato evidenziato nella Tesi si basa sulla possibilità di determinare una valutazione delle varie componenti, secondo l'approccio decompositivo, che spinge l'analisi di CS verso livelli sempre più disaggregati..

La logica della disaggregazione dei dati è utilizzata molto spesso in modelli di analisi multicriteriale.

Lo schema generale della filosofia della disaggregazione è altresì impiegato in altri approcci, come i rough sets, il machine learning e le reti neurali.

Nell'ambito in cui si formula la ricerca è sicuramente importante la conoscenza della percezione della soddisfazione di ogni singolo attributo. E' ormai universalmente noto che la qualità del prodotto/servizio, così come è intesa dai consumatori, si può definire come il grado di discrepanza tra le aspettative o i desideri dei clienti e le loro percezioni.

Si è proposta un'analisi multicriteriale per la valutazione di un insieme delle funzioni marginali di soddisfazione rappresentanti il livello di ogni criterio. E' possibile determinare degli indici che mostrano il livello di soddisfazione parziale dei clienti secondo ogni sub-criterio, similmente all' indice globale di soddisfazione. L'approccio seguito è quello degli integrali gerarchici di Murofushi, Sugeno, Fujimoto.

Infine nel lavoro viene dato particolare risalto alla comprensione del cliente e dei suoi comportamenti, basilare per il customer profiling.

Nella Tesi non si affrontano tutte le componenti del processo di analisi, si prendono solo in considerazione le tecniche di segmentazione. In particolare si utilizzano gli algoritmi degli aggregatori fuzzy per determinare una classificazione-segmentazione della clientela.

Il processo di segmentazione della clientela è un processo chiave in quanto, se svolto appropriatamente, consente di raggiungere una conoscenza reale della struttura del portafoglio clienti. Tale conoscenza è fondamentale, in quanto costituisce le fondamenta per l'identificazione dei target di clientela, la scelta dei prodotti da mettere sul mercato, l'impostazione del marketing mix: in breve, per l'intera azione commerciale.

Nell'applicazione si fa riferimento al caso di una banca che intende suddividere i clienti a seconda del valore (attuale/potenziabile) e differenziarli a seconda dei loro comportamenti (spesso indicatori anche dei loro bisogni).

L'approccio proposto prevede la costruzione di una funzione di utilità non additiva con la tecnica degli integrali bipolari fuzzy di Choquet e di Sugeno.

Il vantaggio degli integrali bipolari di Sugeno e di Choquet consiste nel prendere in considerazione i valori maggiori e minori rispetto ad un livello neutrale di riferimento per ciascun criterio.

Più precisamente, l'estensione degli integrali di Sugeno e di Choquet valuta che il peso attribuito ad un dato insieme di valutazioni dipende anche dall'insieme delle valutazioni simmetriche.

Infatti, nella recente letteratura viene sottolineato come sia interessante da un punto di vista decisionale oltre che considerare l'aspetto classico di confronto tra singole alternative anche l'esistenza di un livello neutrale per

ogni criterio rispetto al quale poter classificare un' azione come attrattiva o repulsiva, o meglio, come nel nostro caso, soddisfatto non soddisfatto.

Nel capitolo 1 si forniscono i primi concetti di Customer Satisfaction secondo l'approccio tipico aziendale.

Nel capitolo 2 si esaminano gli indici (SCSI, ACSI, ECSI) di rilevazione della CS per confrontare prodotti/servizi di diversi settori e/o Paesi e il concetto di customer retention.

Nel capitolo 3 si presentano i principi del Customer Relationship Management.

I capitoli 4, 5 e 6 richiamano, in modo pressoché esaustivo, i principali modelli e metodi statistici utilizzati per stabilire una corrispondente misura di “customer satisfaction”, o almeno precisarne il significato. Nei modelli compositivi la “customer satisfaction” è associata ad una variabile latente il cui valore è ottenuto convenzionalmente da quelli delle variabili manifeste associate al costrutto; nei modelli strutturali vale la stessa assunzione, però, il valore di “customer satisfaction” è stimato da quelli delle variabili manifeste tenendo conto della struttura del costrutto; nei modelli di regressione la “customer satisfaction” è, invece, una variabile manifesta di tipo solo ordinale che pone problemi di scala – peraltro presenti anche negli altri approcci – da affrontare con tecniche appropriate.

Nel capitolo 7 si esplicitano i modelli di statistiche multivariate che associano le tecniche di aggregazione. Si presentano in particolare i modelli più utilizzati in ambito di CS, come il metodo MUSA, il metodo UTADIS, la Cluster Analysis, la Fuzzy Cluster Analysis, la Classificazione sfocata, l'Analisi Congiunta.

Nel capitolo 8 si considerano alcune particolari famiglie di operatori di aggregazione. In particolare è definito il concetto di operatore di aggregazione. Si presentano, altresì, le proprietà matematiche. Si discute infine della necessità dell'uso del concetto di misura fuzzy. Due classi di fuzzy integrali vengono esaminati e caratterizzati: l'integrale di Choquet e l'integrale di Sugeno.

Nel capitolo 9 si presenta il metodo dei Rough Sets per la Customer Satisfaction, quale diverso approccio per la stima dell'importanza relativa di ciascun attributo valutativo.

Nel capitolo 10 si esaminano i problemi e le fasi della ricerca relativamente all'utilizzo delle funzioni di utilità non additive .

Nel capitolo 11 dopo aver dettagliato sulle metodologie si propongono alcune applicazioni relative all'utilizzo di funzioni di utilità non additive nell'ambito della Customer Satisfaction, presentando anche il metodo Montecarlo e gli Algoritmi Genetici.

## CAPITOLO 1 -

### **CUSTOMER SATISFACTION**

#### 1.1 DEFINIZIONE ED ASPETTI AD ESSA COLLEGATI

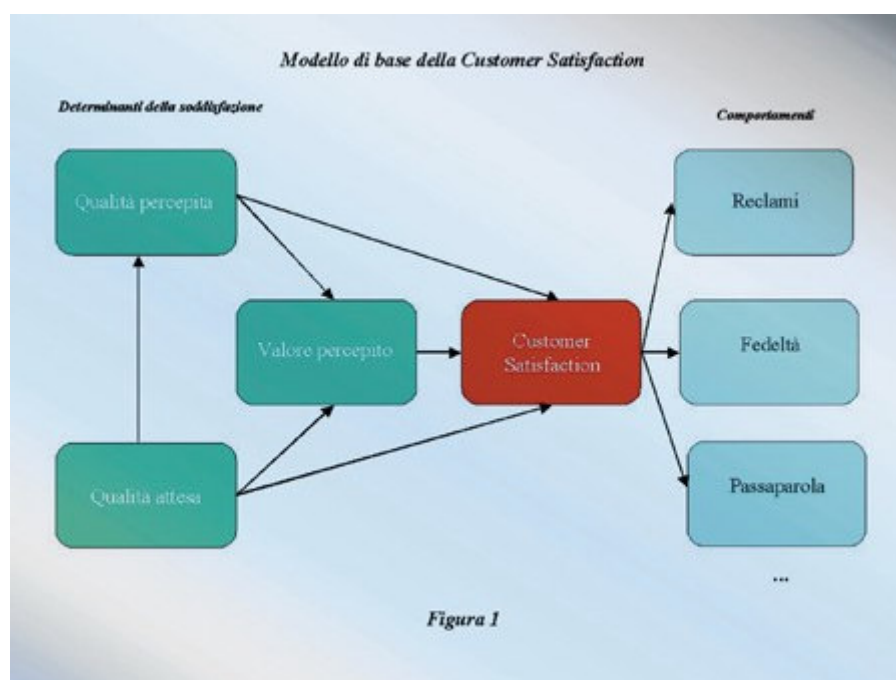
Fare la Customer Satisfaction è diventato ormai un tema di gran moda un po' in tutti i settori economici e a tutti i livelli delle organizzazioni aziendali. Anche nelle organizzazioni pubbliche l'interesse per questo tema è crescente. L'interesse nasce soprattutto dall'enfasi che è stata data nel marketing alla tematica della percezione della qualità da parte del cliente (intermedio o finale) come determinante principale delle decisioni di acquisto e, soprattutto, di riacquisto di determinati prodotti o servizi (AA.VV.GRAMMA, 1993).

L'ipotesi su cui è fondata la rilevanza pratica della Customer Satisfaction dal punto di vista manageriale, ovvero che è meglio avere clienti soddisfatti invece che insoddisfatti, può a ragione sembrare la scoperta dell'acqua calda, ma si tratta di un'impressione superficiale. Considerare la soddisfazione a livello individuale come la leva di marketing per la gestione attiva del portafoglio clienti non è una scelta ovvia. Non appena ci si comincia a chiedere in un caso concreto che cosa significhi soddisfare un cliente e soprattutto cosa e come fare per potere influire sul suo livello di soddisfazione, la saggezza istintiva di qualsiasi homo oeconomicus necessita di spiegazioni, a cominciare dal fatto se qualcosa come un livello di soddisfazione negli esseri umani esista o meno. Non si può certo dare per scontato che esso esista nello stesso senso in cui esiste un livello di adrenalina nel sangue umano. Non esistono cose come la soddisfazione o l'insoddisfazione in assoluto, ma tutta una serie di discrepanze, di distanze, di differenziali più o meno variabili tra ciò che un cliente si aspetta dall'acquisto di un certo prodotto o di un certo servizio e ciò che in realtà ottiene in cambio. Dato che tali discrepanze hanno a che fare con la percezione della qualità, esse sono certamente completamente soggettive ed individuali, molto probabilmente legate a fattori intangibili dell'utilizzazione e del consumo di prodotti e servizi e perciò molto difficili da ricondurre a fattori oggettivi. Soddisfazione e insoddisfazione non sono

concetti unidimensionali, ma multidimensionali e complessi, costituiti sia da tanti fattori distinti sia da fattori tra loro sovrapposti.

Ancora più che in passato, l'approccio della Customer Satisfaction sostenuto negli ultimi anni dal marketing mette in luce la necessità della congruenza tra domanda e offerta sui mercati della percezione della qualità (Swan e Mercer, 1981). Ove l'esperto di marketing riscontra incongruenze tra livello di qualità attesa e livello di qualità percepita, egli raccomanda al committente di "colmare il gap", il che significa quasi sempre di intraprendere dei programmi di cambiamento (di "miglioramento") che si devono confrontare con la percezione di qualità del cliente. L'orientamento al cliente richiede all'azienda una disponibilità illimitata a migliorare continuamente i propri prodotti e servizi, le proprie procedure operative e la qualità del proprio personale, almeno fino a che tutto ciò è economicamente giustificato. Al di là di quel punto occorre valutare se invece non sia il caso di attuare il Business Process Reengineering (bpr) invece della Customer Satisfaction.

I presupposti di buon senso dell'approccio della Customer Satisfaction sono corroborati da un corpus notevole di ricerche econometriche e statistiche che attestano legami causali strettissimi tra qualità, soddisfazione, fedeltà, e redditività delle imprese dei settori più disparati. [Figura 1: Modello di base della Customer Satisfaction](#)



## 1.2 PSICOLOGIA DELLA CONOSCENZA

Le ricerche sulla customer satisfaction trovano il loro fondamento teorico negli studi sulla psicologia della personalità condotti da Hoppe (1930) e da Lewin (1936) nella prima metà del XX secolo. Indagando il costrutto "autostima/autofiducia" e le sue determinanti, tali studiosi hanno di fatto posto le basi concettuali e metodologiche per lo sviluppo del cosiddetto *paradigma della conferma/disconferma delle aspettative*.

Applicando tale paradigma ai processi valutativi post-acquisto, infatti, la percezione di soddisfazione/insoddisfazione dipenderebbe dal confronto fra aspettative e percezione di performance (valore atteso *versus* valore percepito)

Le sperimentazioni sulla *customer satisfaction* hanno condotto all'elaborazione di diversi modelli, tutti genericamente definiti *gap models*, al fine di connotare la natura sottrattiva della soddisfazione. Oltre un centinaio di articoli apparsi sulle più importanti riviste scientifiche hanno, di volta in volta:

- avallato la validità del paradigma della conferma/disconferma, definendone configurazioni alternative in funzione della natura delle aspettative (Cadotte, Jenkins e Woodruff, 1987), e proponendo addirittura una "standardizzazione" delle dimensioni caratterizzanti il processo valutativo, con riferimento alla misurazione della "qualità percepita" nei servizi (Parasuraman, Zeithaml e Berry, 1988);
- sperimentato la maggiore articolazione del paradigma, che agirebbe sulla soddisfazione sia indirettamente, via disconferma, sia direttamente, mediante le aspettative e le percezioni di performance (Oliver, 1980; Churchill e Suprenant, 1982);
- sostenuto la prevalenza delle performance percepite, talvolta nel corso di esperimenti volti a misurare la qualità percepita (seppure definita in modo identico alla soddisfazione e con scale di dubbia validità), in altri casi nell'ambito di misure delle determinanti della *customer satisfaction* (Oliver e De Sarbo, 1988; Cronin e Taylor, 1992 e 1994);
- dimostrato, al contrario, il maggior peso delle aspettative (Oliver, 1980; Cadotte, Woodruff e Jenkins, 1987), specificamente in processi d'acquisto ambigui e caratterizzati da una scarsa conoscenza dei prodotti

Ne sono emerse configurazioni alternative del paradigma di riferimento.

Al riguardo, è stato proposto (Costabile 1996b) il ricorso alla distinzione fra beni ricerca, esperienza (Nelson, 1970) e fiducia (Darby e Karni, 1973), estendendo tale tipologia di analisi ai processi d' acquisto, per sostenere che:

- nel processo valutativo dei prodotti "esperienza", la customer satisfaction è spiegata dalla percezione di performance in misura maggiore rispetto allo scostamento fra aspettative e performance, oppure rispetto alle sole aspettative, proprio in quanto le specificità della valutazione non consentono la formazione di aspettative affidabili sull' offerta;
- nel processo valutativo dei prodotti "ricerca", la customer satisfaction è spiegata in misura prevalente dallo scostamento fra aspettative e performance, piuttosto che dalle sole performance o dalle sole aspettative;
- nel processo valutativo di prodotti "fiducia", la customer satisfaction è spiegata in misura maggiore dalle aspettative, rispetto allo scostamento fra aspettative e performance o alle sole percezioni di performance; e ciò in quanto la valutazione delle performance rimane incerta e la percezione di soddisfazione subisce un effetto di assimilazione delle aspettative (Hoch e Ha, 1986).

Qualora le suddette proposizioni venissero confermate dalla verifica empirica si potrebbero ipotizzare metodologie di misurazione differenziate in funzione della tipologia di processo valutativo.

Si fa sovente riferimento al ruolo rivestito dal cliente nella caratterizzazione del servizio e della *misura della qualità*.

Diversi studiosi si sono occupati di definire meglio i caratteri del legame tra azienda e clientela. Si citi l'Albrecht (Albrecht 1992), secondo il quale, riferendosi alle aziende orientate al servizio, il consumatore è sito al centro del sistema progettuale, produttivo, distributivo del servizio, predisposto questo in funzione, non solo del profitto, ma anche della domanda.

Così come l'Albrecht, anche il Carlzon ribadisce l'effetto della clientela sul successo o sull'insuccesso dell'organizzazione: essa si colloca all'apice della struttura aziendale



(“il cliente ha sempre ragione”), cosicché tutto sia finalizzato alla soddisfazione del cliente, o come si suol dire, alla *customer satisfaction* (Heider,1958).

Con riferimento al nostro Paese, la qualità intesa come capacità di soddisfare le esigenze del cliente ha cominciato a radicarsi grazie a tecniche gestionali e operative volte al mercato ed all'efficienza gestionale più che alla produzione.

Secondo questa visione, l'acquirente è il riferimento permanente verso cui allineare la gestione, fondandosi questa sull'adeguato impiego delle risorse finanziarie ed umane a disposizione.

### 1.3 ASPETTI MOTIVAZIONALI

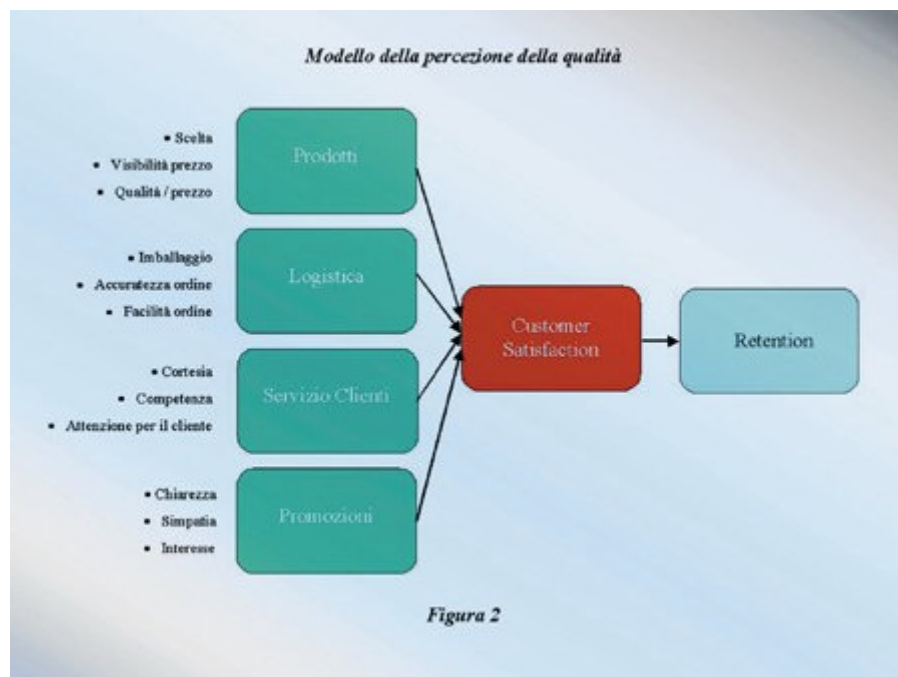
Come si fa a capire se un cliente è soddisfatto o no? Occorre – ovviamente – osservarne attentamente il comportamento. Si suppone che quando un cliente è soddisfatto tipicamente si comporta in modo diverso da quando è insoddisfatto. Tuttavia l'osservazione del comportamento non basta al manager che deve e che vuole orientare la sua azienda al cliente. Questo manager infatti deve conoscere le cause del comportamento del cliente in modo da poter predisporre in modo adeguato l'organizzazione della sua azienda, da poter stimare quante risorse gli sono necessarie per soddisfare tutte le variazioni di preferenze della clientela e da decidere come utilizzarle. La pura e semplice osservazione, per quanto acuta, a questo punto non basta più. Occorrono una comprensione profonda e un monitoraggio continuo delle motivazioni e delle scelte dei clienti (Castaldo, 1994 e 1995).

Ciò è possibile solamente per mezzo dell'ascolto attivo della voce del cliente, ma soprattutto dando al cliente risposte sulla sua stessa "lunghezza d'onda", cioè formulate nello stesso linguaggio utilizzato dal cliente. Ogni cliente ha in testa un modello mentale e semantico che lo guida nell'acquisto e nella fruizione di prodotti e servizi e ogni azienda farebbe bene a cercare di conoscere quelli dei propri clienti. Purtroppo anche ipotizzando di poter aprire la testa dei propri clienti per guardare cosa c'è dentro non si raggiungerebbe molto. Non c'è altra via che domandare loro che cosa si aspettano da un dato prodotto o da un dato servizio e in che misura ciò che hanno acquistato soddisfa o meno tali aspettative.

In realtà questo è molto meno semplice di quanto non possa sembrare in quanto le domande devono essere tali da consentire non solo di ricostruire il modello mentale e semantico della qualità percepita del cliente, ma anche di collegare uno ad uno ogni aspetto di questo con il sistema operativo manageriale dell'azienda. Rispetto a forme più tradizionali di ricerche di mercato, i progetti di misurazione della Customer Satisfaction hanno la peculiarità di richiedere come preconditione non solo la comprensione delle motivazioni di soddisfazione e di insoddisfazione dei clienti, ma anche del contesto organizzativo aziendale in cui queste informazioni vanno presentate

e soprattutto utilizzate in modo tale da ridurre o addirittura eliminare, dove esistono, le eventuali discrepanze di valori tra aziende e clienti.

*Figura 2: Modello della percezione della qualità*



Le aziende che si trovano di fronte a grandi numeri di clienti, magari anche molto dispersi geograficamente, incontrano problemi particolari a organizzare sistematicamente le loro procedure di ascolto della voce del cliente e a misurare periodicamente il livello di soddisfazione.

I clienti si possono ascoltare a uno a uno, in gruppi (o in focus groups), in giurie, tramite loro rappresentanti scelti in libere elezioni o anche tutti in una volta in uno stadio (se ci entrano) e più o meno periodicamente. Comunque sia, è consigliabile che quanto viene detto in qualche modo venga registrato per essere studiato e discusso accuratamente a distanza di qualche giorno dal manager e dal suo assistente.

Ancora più importante è che il manager abbia a disposizione una metodologia che gli garantisca con un grado ragionevole di affidabilità di fare acquisire effettivamente tutte le informazioni di cui ha bisogno per gestire attivamente il suo portafoglio clienti.

La rilevanza dell'ascolto della voce del cliente del punto di vista del marketing va ben al di là del semplice buon senso. Nell'approccio Customer Satisfaction l'opinione e le preferenze del cliente non sono solo dei dati da raccogliere e da interpretare, ma dei termini di comparazione, ovvero, come si suol dire, il benchmark fondamentale dal punto di vista operativo sia per l'esperto di marketing sia per l'uditore di turno in azienda. Questo ovviamente non vuol dire che altri approcci e altre tecniche di ascolto del cliente non abbiano ragione di esistere, ma che l'ascolto della voce del cliente, specialmente tramite una misurazione statistica della Customer Satisfaction, ha una funzione talmente preminente che certamente ha la priorità su tutte le altre, una funzione che in ultima analisi trova la sua radice nel principio economico della sovranità del consumatore.

Un progetto di Customer Satisfaction naturalmente può avere come esito anche che l'azienda in oggetto è sostanzialmente congruente con il sistema di valori dei suoi clienti. Una tale conferma naturalmente non dovrebbe solo indurre il management a pentirsi di aver impiegato tempo e denaro in un progetto di questo tipo, ma soprattutto a chiedersi se l'azienda sarà in grado di mantenere lo stesso livello di congruenza anche in futuro, magari quando il ciclo vitale dei suoi prodotti e/o dei servizi attuali volgerà alla fine o quando sul suo mercato di riferimento potrà cambiare completamente la situazione competitiva o avvenire un cambiamento tecnologico radicale.

#### 1.4 SISTEMA DI MISURAZIONE E SISTEMA OPERATIVO MANAGERIALE

Un sistema di Customer Satisfaction è costituito da almeno due componenti: un sistema di misurazione del livello di soddisfazione e un sistema operativo manageriale. Il sistema di misurazione può essere costituito da un semplice questionario su carta o da un sistema informativo a sé stante, interno o esterno rispetto all'azienda (Customer Satisfaction Council, 1995).

Il sistema operativo manageriale può essere allocato in una funzione o in un reparto ben specifici (come ad esempio nel reparto marketing, nel controllo della qualità o nella pianificazione strategica, direzione commerciale, alta direzione) o essere un processo aziendale autonomo legato agli altri processi aziendali nei modi più svariati. La competenza per la Customer Satisfaction normalmente non può essere allocata a priori in una sola funzione. Si tratta di un processo intrinsecamente trasversale che va a influire sulla percezione di qualità del cliente con innumerevoli ramificazioni. Benché gli sforzi volti a adottare uno standard di riferimento per la misurazione della Customer Satisfaction non siano ancora giunti a compimento, qualsiasi sistema di misurazione della Customer Satisfaction deve soddisfare gli stessi requisiti minimi (Gerson, 1993):

- Il sistema deve essere affidabile;
- I risultati devono essere ripetibili e confrontabili (anche da altri);
- I risultati devono essere discriminanti;
- Il sistema deve essere coordinato con i processi aziendali/organizzativi.

L'affidabilità è certamente la caratteristica più importante di qualsiasi sistema di misurazione della Customer Satisfaction. La misurazione per mezzo di scale quantitative multiple e ponderate è ormai uno standard accettato universalmente nelle misurazioni di Customer Satisfaction in quanto questo tipo di scala garantisce la migliore precisione statistica della misurazione di un concetto, quello della soddisfazione, che è intrinsecamente vago.

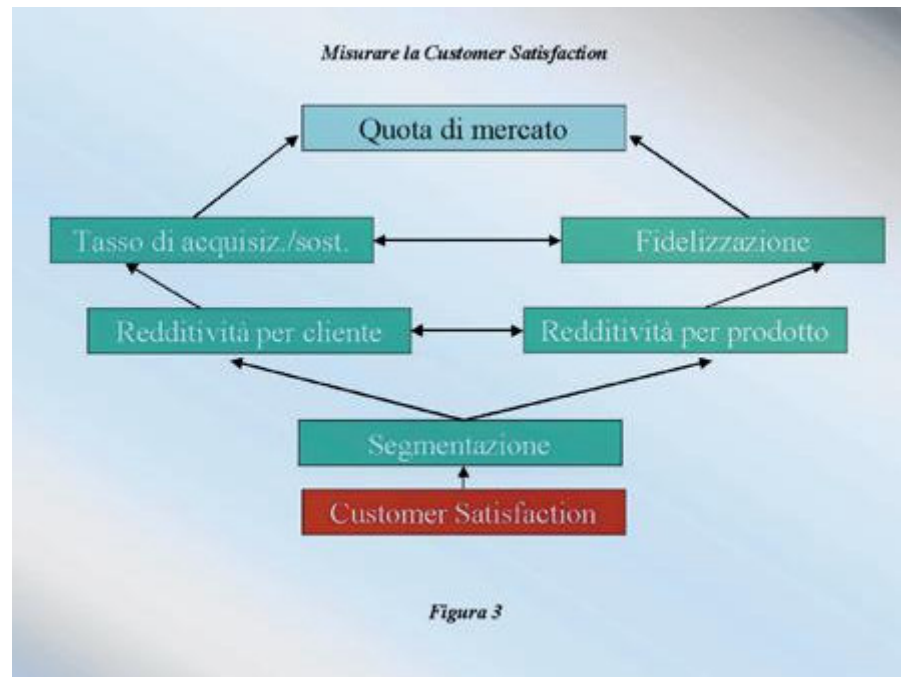
Le procedure con cui vengono pianificate le misurazioni e vengono elaborati i dati devono essere ripetibili e comparabili. Questo implica ovviamente non solo che una misurazione eseguita a distanza di pochi mesi da un'altra deve sostanzialmente misurare la stessa "cosa", ma soprattutto che non dovrebbero esistere motivi per imporre il "segreto professionale" sulle procedure di misurazione e di elaborazione dei dati. Questo problema oggi esiste in quanto esistono i più svariati "consulenti" che dicono tutti di avere scoperto il metodo di misurazione migliore del mondo per misurare la Customer Satisfaction e che vorrebbero mantenere il segreto professionale su di esso (Berry e Parasuraman, 1991).

I risultati del sistema di misurazione (i dati rilevati) devono essere discriminanti, cioè aiutare effettivamente dal punto di vista logico i loro utilizzatori a identificare le cause delle differenze di livello dell'indice di soddisfazione e le possibili azioni da intraprendere per ridurre o eliminare le discrepanze tra qualità attesa e qualità percepita.

Solo se la logica di un sistema di misurazione della Customer Satisfaction è collegata con un sistema operativo manageriale la misurazione della Customer Satisfaction va al di là dell'esercizio accademico e diventa una realtà operativa per la gestione strategica di un'azienda o di un'organizzazione.

La produzione di indici totali o parziali di (in-)soddisfazione è solo una delle funzioni di un sistema di Customer Satisfaction. L'indice di Customer Satisfaction è la base su cui si costruiscono misure della redditività per cliente e per prodotto, misure dell'efficienza del processo di acquisizione e di sostituzione di clienti, misure del grado di fidelizzazione, della segmentazione della clientela e della dimensione della quota di mercato. Le misure di redditività per cliente e per prodotto servono per allocare con efficienza gli sforzi di marketing e di vendita dell'azienda cliente per cliente, segmento per segmento e prodotto per prodotto. Insieme con un indice di soddisfazione queste misure consentono di misurare direttamente e/o indirettamente i tassi di acquisizione o di sostituzione dei clienti, il costo di queste operazioni e gli utili derivanti dalla permanenza di un cliente nel portafoglio clienti dell'azienda. Si può parlare di una "permanenza ottimale" di un cliente nel portafoglio clienti nel senso che la crescita

dell'indice di soddisfazione nel tempo e la fedeltà contribuiscono essenzialmente a ridurre i costi totali che l'azienda deve sostenere per poterlo servire, ma da un certo punto in poi tentare di incrementare ancora la soddisfazione può diventare controproducente. [Figura 3: Misurare la Customer Satisfaction](#)



Tutte insieme queste misure consentono di determinare la quota di mercato dell'azienda intesa sotto vari aspetti: come quota di profitti lordi presenti in un dato mercato, come percentuale del fatturato totale di quel mercato, come quota di spesa dei clienti, come numerosità di clienti sul totale del potenziale di mercato – e chi più ne ha più ne metta. La misura che interessa di più l'azienda è la quota di mercato: le altre misure servono per determinarla più o meno (in-)direttamente.

Nel caso di organizzazioni pubbliche questo modello va modificato notevolmente in quanto di solito si tratta di monopoli locali o nazionali creati dal diritto pubblico (che nondimeno vengono considerati "monopoli naturali"). La quota di "mercato politico" che garantisce la sopravvivenza e lo sviluppo dell'organizzazione in questi casi si pone in termini completamente diversi. Ciascuna misura presa singolarmente non ha un significato intrinseco. È la loro combinazione a agire potenzialmente come catalizzatore dello sviluppo strategico di un'azienda o di un'organizzazione pubblica che "pensa" orientandosi a un modello di comportamento economico.

## 1.5 OBIETTIVI DEL MANAGEMENT

Lo scopo operativo ultimo di qualsiasi misurazione di Customer Satisfaction non è di massimizzare, bensì di ottimizzare la soddisfazione dei clienti. La possibilità di soddisfare i propri clienti presenta senza dubbio dei limiti assoluti al di là dei quali non solo gli investimenti necessari non sono giustificati, ma si hanno anche effetti controproducenti. Infatti non è affatto inspiegabile che proprio i clienti più soddisfatti paradossalmente siano maggiormente predisposti a abbandonare un dato prodotto o servizio: il motivo principale è che sono stati soddisfatti a tal punto che non hanno più aspettative nei confronti di quel prodotto o servizio e che vogliono semplicemente sperimentare qualcos'altro (ad es. una Lexus dopo 20 anni di Mercedes). È un comportamento normale e compatibile con la dinamica delle economie di mercato.

Gli innumerevoli fattori che contribuiscono a formare la percezione della qualità non possono essere né di esclusiva competenza del reparto di controllo della qualità né del reparto marketing, ma coinvolgono tutte le funzioni e i processi che in qualche modo determinano direttamente e/o indirettamente la soddisfazione dei clienti. Dare una tipologia di un sistema operativo manageriale valida per tutti i settori economici e per tutte le aziende dello stesso settore perciò è piuttosto difficile. È il caso di mettere in rilievo due aspetti: il capitale umano e i sistemi di incentivazione.

Il trasferimento del valore di un prodotto o di un servizio a un prospect o a un cliente già noto nella maggior parte dei casi è esso stesso un servizio. Quindi non può essere immagazzinato, trasportato e trasferito se non per mezzo del capitale umano (ovvero di persone in carne e ossa, tanto per intenderci). Ciò significa che ogni programma di cambiamento e/o miglioramento dell'organizzazione di un'azienda o di un'organizzazione pubblica deve necessariamente passare attraverso la selezione e/o la formazione delle persone adatte per i propri clienti. Laddove il portafoglio clienti è estremamente variegato occorrerà formare il personale in modo tale che esso sia in grado di parlare il maggior numero di "linguaggi" possibili, di fargli conoscere dettagliatamente i modelli mentali della percezione della qualità delle persone con cui



ha a che fare e educarlo a fare le domande giuste ai clienti. Questo può avvenire solo se dall'azienda il personale viene trattato almeno altrettanto bene come i suoi clienti. La soddisfazione del cliente è sì "spiegabile" con il suo modello mentale di percezione della qualità, ma diventa gestibile soprattutto per mezzo del comportamento del personale che più è in grado di influire sulla percezione di qualità del cliente. A sua volta il comportamento del personale è determinato dal livello di soddisfazione per il lavoro che svolge. Questa è l'esatta controparte della capacità dell'azienda di selezionare i clienti "giusti" proprio per il livello di qualità che è in grado di fare percepire. Quanto detto sopra basta per relativizzare il peso che è stato dato ai sistemi di incentivazione della Customer Satisfaction, che tendono a essere basati quasi tutti esclusivamente sul principio della "durata della fedeltà" alla stessa azienda e che non sembrano essere molto efficaci. I sistemi di incentivazione basati su misurazioni quantitative di Customer Satisfaction inoltre hanno lo spiacevole effetto di indurre gli interessati a inventare modi per manipolare i dati e perciò vanno utilizzati con grande cautela.

Come abbiamo già avuto modo di vedere, avere clienti fedeli di per sé può non essere né una garanzia né un plus per l'azienda. Senza doverci addentrare troppo nei meandri dei principi di retribuzione del lavoro, possiamo dire che dall'approccio Customer Satisfaction al capitale umano emerge non solo l'importanza della fedeltà all'azienda, che, dove e quando opportuno, può certamente essere incentivata con certi sistemi retributivi, ma che emerge altrettanto forte anche l'importanza dell'allocazione delle risorse umane a disposizione dell'azienda nelle direzioni in cui la percezione della qualità da parte del cliente è destinata a avere un impatto più forte. Come la Customer Satisfaction raccomanda di investire di più sui fattori che determinano maggiormente la percezione della qualità da parte del cliente nel caso degli investimenti strumentali e organizzativi, nel caso delle risorse umane questo approccio non può che raccomandare di investire relativamente di più sulle persone preposte e in grado di trasferire "meglio" la percezione della qualità dell'azienda ai clienti. Naturalmente il significato di "meglio" va discusso (e misurato) caso per caso.

## 1.6 LEGAMI TRA CS E QUALITA'

Nell'ottica della piena soddisfazione del cliente, viene ad essere proprio il TQM (Total Quality Management) la dottrina gestionale che mira al beneficio dei compratori e dell'interocomparto aziendale (AA.VV.,1979, AA.VV.,1987, AA.VV.,1990, AA.VV.,1992).

Essa si esplicita attraverso la partecipazione di tutte le funzioni (ricerche di mercato, progettazione, approvvigionamenti, produzione,...) e quindi di ogni membro dell'organizzazione (Total): ogni singola parte deve esprimere la sua essenza in funzione dell'acquirente, e di conseguenza il lavoro e le tecniche delle singole unità devono avere ampio respiro e non essere fini a se stesse.

Si ricordi l'espressione "Il processo a valle è il tuo cliente" del prof. Ishikawa: così il sistema diventa una rete di rapporti cliente-fornitore focalizzati affinché l'azienda possa ottimizzare il servizio ai clienti effettivi. Il riferimento alla qualità viene ad esplicitarsi come la congruenza alle specifiche (*requirement*) decretate dal cliente. Infatti, la ignoranza di tali richieste è la causa degli errori e non permette di offrire prodotti e/o servizi fruibili.

L'alta dirigenza ha allora il compito della supervisione di ogni ambito del TQM e della customer satisfaction, che nella sua accezione più rilevante non può ridursi alla mera fornitura di prodotti conformi. Infatti il rispetto di qualsivoglia specifica o standard viene a perdere valore se vi è insoddisfazione da parte del cliente.

La politica del monitoraggio dei reclami del cliente mostra correntemente che essi trovano espressione in riferimento alle attività non-produttive più che ai prodotti non soddisfacenti.

Pertanto l'attenzione alla qualità di tutto il personale tende a migliorare l'immagine dell'azienda.

E' significativo a questo punto riallacciarsi al CWQC, o *Company-Wide Quality Control*. Esso è un sistema manageriale fondato in Giappone nell'ambito della rinascita industriale post-bellica e basato proprio sull'estensione dei concetti e delle tecniche del Controllo Qualità a tutti i settori dell'azienda.

Grazie a questa tecnica, il significato della parola qualità viene proiettato all'esterno, perdendo le soggettive connotazioni interne: l'unica definizione di qualità è quella del cliente.

In effetti, il profitto, originaria priorità aziendale, è appannaggio di pochi individui ed è legato a conoscenze in genere ignorate dai dipendenti, mentre la customer satisfaction è relativa ad ogni singolo membro dell'organizzazione in quanto cliente di un generico fornitore.

L'adempimento dei fattori di base è un presupposto irrinunciabile che deve essere corredato da altri fattori ai fini dell'accrescimento della soddisfazione.

Infatti, si ha anche la cosiddetta qualità "latente", nel senso che vi sono dei bisogni inespressi dai clienti, il rispetto dei quali è fondamentale perché non atteso. Questi sono definiti come "fattori di delightment": meravigliano il cliente e ne accentuano ampiamente il soddisfacimento.

Deve allora essere annoverata la necessità di recepire anche quei fattori, detti prestazionali che sono dati per scontati ed evidentemente inespressi. In questa accezione la customer satisfaction si delinea come una sensazione su cosa sia stato offerto dall'azienda al consumatore e sulle modalità attraverso le quali ciò è avvenuto.

E' necessario fornire sempre qualcosa di nuovo (capacità proattiva): non è importante quale sia il livello di gradimento acquisito perché è un fattore provvisorio destinato alla crescita verso un livello più elevato. C'è quindi una richiesta di prodotti sempre innovativi, il che spinge a politiche aziendali

estremamente dinamiche, dato che la vita del prodotto-tipo è in continuo calo, e questo procede di pari passo alla crescente complessità delle risorse da attuare per fidelizzare i clienti.

Ma le richieste riguardano anche la qualità del servizio offerto: dalla affidabilità nell'acquisizione degli ordini all'efficacia della rete di vendita e del servizio postvendita, e nell'odierna ottica ambientalista, anche la riciclabilità dei prodotti.

L'impegno deve essere continuo ed attento. Infatti, la eventuale penalizzazione dell'acquirente ha gravi effetti: si verifica spesso che, a causa di insoddisfacimento, non siano effettuati reclami e si verifichi il conseguente allontanamento del cliente.

E' da tenere in considerazione anche il cattivo ritorno pubblicitario derivante dalla perdita di un cliente: esso infatti tende ad esternare ad altri il suo punto di vista, con le reazioni a catena del caso.

Perciò l'utente insoddisfatto deve ricevere un servizio metodico e personalizzato, indipendentemente dalle richieste di altri: ha la possibilità di determinare il futuro dell'azienda e il veicolo prioritario per far questo è l'ottimizzazione della customer satisfaction.

## CAPITOLO 2 -

### INDICI DI CUSTOMER SATISFACTION

#### 2.1 ACSI, ECSI, SCSi

La misurazione della customer satisfaction è divenuta uno dei fondamentali processi di valutazione delle performance aziendali. Lo studio del legame fra soddisfazione, fiducia, immagine e fedeltà dei clienti infatti ha evidenziato che dalla customer satisfaction dipendono le performance economiche e competitive e, quindi, il valore stesso dell'impresa (Busacca e Costabile, 1995; Reichheld, 1996; Srivastasa, Shervani e Fahey, 1998; Naumann e Giel, 1995).

Da quanto visto emerge che, col passare del tempo, sono andati ad aumentare gli approcci scientifici tesi all'incontro delle necessità di qualità ed affidabilità espresse dalle varie utenze (Hill, 1996).

Così la qualità viene ad essere una sorta di compromesso tra le performance del prodotto e/o servizio fornito e le attese del consumatore:

$$\text{Aspettative} = \text{Performance} / \text{Qualità}$$

Allora ogni prestazione eccessiva comporta dei costi elevati, mentre ogni calo di standard porta ad un calo della qualità complessiva. Il giusto accordo si ha per valori prossimi all'unità, garantiti da rapporti diretti con gli acquirenti. Si deve ancora ribadire che un elevato grado di customer satisfaction è foriero di elevati profitti per l'organizzazione. Infatti, studi recentemente effettuati su imprese-campione statunitensi hanno mostrato esservi una correlazione diretta tra i profitti e la coppia quota di mercato-qualità percepita, essendo quest'ultima proprio la capacità di incontro dei bisogni dell'utente propria di un bene e/o servizio. Si sono registrati degli incrementi di profitto (pari anche al 60%) ogniquale volta il livello qualitativo percepito di un'azienda superava quello delle dirette concorrenti, anche nel caso di una ristretta quota di mercato della prima rispetto alle seconde.

Si è scritto degli effetti correlati al non soddisfacimento di determinate esigenze: la soddisfazione del cliente è il mezzo atto a garantirne la fidelizzazione. Questa è da ottenersi mediante la politica delle indagini di mercato verso utenti singoli o

associazioni di consumatori e la conseguente creazione di vere e proprie forme di collaborazione (*comakership*) tra le parti in questione.

Dall'esame dei dati forniti dalla corrente pratica aziendale, risulta che i clienti completamente soddisfatti abbandonano i fornitori nel 27% dei casi, nel 55% dei casi quelli "normalmente" soddisfatti e nell'83% quelli insoddisfatti.

In genere, le aziende hanno ottenuto di mantenere clienti insoddisfatti grazie a tempestive reazioni ai reclami. Proprio in virtù del rapido interessamento delle organizzazioni, questi stessi clienti hanno contribuito al profitto in maniera crescente negli anni successivi l'evento.

Grazie alla capacità di mantenere i clienti si sono registrati tendenzialmente degli aumenti del volume di acquisti anche pari al 50% in un periodo di mantenimento pari a 10 anni. Gli acquirenti soddisfatti, come già accennato in precedenza, rafforzano l'immagine delle compagnie e portano incrementi di utili non indifferenti.

La tendenza attuale degli Enti di Certificazione, sulla scia del ristrutturato approccio metodologico fornito dalle Vision 2000, non è più la semplice verifica dell'applicazione delle norme ISO, ma anche l'analisi di *indici di qualità*, come i *Customer Satisfaction Index (CSI)*, che forniscono indicazioni sul trend di miglioramento delle imprese.

L'affermazione di un'azienda è figlia del corretto management dei suoi asset, sebbene la odierna contabilità aziendale tenga considerazione solamente di quelli manifesti e non di altri, come quello rappresentato dalla clientela, che sono una voce rilevante del reale valore di un'organizzazione.

La misura del "fattore clientela" è di necessaria importanza ai fini del completamento degli indicatori gestionali e relativi alla contabilità.

In tal senso, la creazione di Indici di Customer Satisfaction comporta una valutazione quantitativa della qualità con riguardo al cliente. Essi forniscono utili indicazioni sugli interventi strategici imprenditoriali finalizzati al ritorno in termini di immagine, investimenti (ROI), vendite (ROS) e, in generale, sull'efficacia di tutte le tecniche orientate al servizio. Tra queste è necessario che si annoverino i seguenti: la realizzazione di complementi alle prestazioni di base, come quelli finanziari,

L'implementazione di sistemi di assistenza operativi in tempo reale e di strategie atte alla gestione dei reclami, e, *dulcis in fundo*, la messa in atto di Sistemi di Gestione della Qualità sul modello delle norme ISO 9000.

Un efficace CSI deve permettere di valutare quali siano i limiti da porre agli investimenti in qualità al fine di avere dei "return" positivi.

Per la valutazione delle dimensioni economiche del valore del cliente è possibile ricorrere alle misurazioni inerenti a: fatturato, margine e redditività, fruendo delle numerose matrici di analisi del portafoglio clienti e facendo riferimento a matrici della profittabilità che rappresentano indicatori del profiling del cliente.

Tra i rivelatori di customer satisfaction che si basano sugli esistenti processi informativi aziendali, detti sistemi di misurazione indiretti, è possibile ascrivere i vari LTV<sub>i</sub>, CRR, AMP.

Nell'ordine, l'impresa che non misura la soddisfazione non è, a conoscenza dei costi indiretti dell'insoddisfazione: un cliente perduto rappresenta una perdita di ricavi e profitti, che spesso è per sempre o difficilmente recuperabile, se non a fronte di ingenti e costosi investimenti. Tale perdita può essere stimata con il *Life time Value* attraverso il quale si determina il valore che un cliente rappresenta per l'impresa.

Tale valore, espresso in termini di ricavi, si calcola moltiplicando il ciclo di vita medio di un cliente per i suoi acquisti medi annui. In termini analitici:

$$LTV_i = VM \cdot FA \cdot CV_i$$

Dove LTV<sub>i</sub> = Life Time Value relativo al cliente i-esimo;

VM = Valore Medio della transazione;

FA = Frequenza annua di acquisto;

CV<sub>i</sub> = ciclo di vita atteso del cliente i-esimo (funzione anche del suo tasso di fedeltà), l'anzianità attesa per il cliente i-esimo.

Quest'ultimo indice è la valutazione finanziaria di un acquirente dal contatto iniziale con un'azienda al momento in cui non è più economicamente redditizio per l'impresa stessa. L'analisi LTV<sub>i</sub> permette di comprendere il valore corrente e potenziale dei

clienti attuali, valutare gli investimenti necessari all'ottenimento di nuovi, far fronte al cambiamento delle relazioni con gli acquirenti, effettuare un miglior *business planning* e, infine, predire il tempo residuo di mantenimento di un singolo utente.



## 2.2 CUSTOMER RETENTION

Il Customer Retention Rate (CRR) fornisce una indicazione sulla entità dei clienti fidelizzati in un determinato lasso di tempo rispetto a quelli che sono passati ad altri fornitori. Esso, pertanto, esprime il numero di clienti rimasti fedeli alla fine dell'anno rispetto a quelli che vi erano all' inizio ed ai nuovi acquisti:

*CRR istantaneo*

$$(\text{Customer Retention} = \frac{(\text{Clienti a fine anno} - \text{Nuovi clienti})}{(\text{Clienti ad inizio anno})} \times 100).$$

*CRR storico*

$$(\text{Customer Retention} = \frac{\text{CRR}_{2000} + \text{CRR}_{2001} + \text{CRR}_{2002} + \dots}{\text{Anni}} \times 100).$$

Mentre il CRR istantaneo stima il tasso di fedeltà della clientela per anno preso in considerazione, il CRR storico rappresenta la media aritmetica dei CRR istantanei ottenuti nel corso del periodo di osservazione.

L'AMP è correlato al CRR mediante la relazione

$$AMP = N / (1 - CRR)$$

ed esprime l'Anzianità Prospettica Media dell'utente, ovvero una stima del suo tempo di mantenimento. All'interno della formula dell'anzianità media prospettica “*N*” indica il numero di anno, o l'anno, o la frazione di un anno che definisce il ciclo di acquisto/riacquisto caratteristico in cui opera l'impresa. Per convenzione, adottando una visione sincrona rispetto ad altre misure di valore (reddito, flussi di cassa, fatturato), *N* viene sempre utilizzato quale valore pari ad 1. Tale convenzione non viene rispettata solo per quei prodotti che evidenziano un ciclo di riacquisto/acquisto di molto superiore; infatti in tali casi il tasso di fidelizzazione e il calcolo del LMP o AMP vengono misurati su base pluriennale e, in seguito, espressi in termini di media “annua”.

L'anzianità media prospettica della clientela viene quindi calcolata quale rapporto inverso al turnover di suddetta clientela e, in base a ciò, si intende stimare la durata della relazione con quest'ultima. Il calcolo di tale valore aiuta a misurare il LTV così

come illustrato in precedenza e, indi, a sostituire al ciclo di vita – inteso quale determinazione temporale della relazione con il cliente - l'AMP, che si sostanzia in un numero di anni o mesi a seconda della visuale temporale adottata.

Agli approcci appena evidenziati, devono esserne affiancati degli altri, più sfruttati, che si basano fondamentalmente su indagini e analisi di mercato, e quindi su un diretto rapporto con i consumatori (sistemi diretti).

Queste metodiche sono normalmente riferite alla clientela esterna, ma trovano anche espressione nella valutazione della rete di rapporti interfunzionali all'interno delle organizzazioni.

E' da sottolineare però che un criterio come quello appena visto, basato sulla realizzazione di un modello mediante inchieste a campione, non fornisce un approccio oggettivo alla valutazione della customer satisfaction.

Il Customer Satisfaction Index (CSI) è un indicatore che consente di monitorare e condensare le informazioni raccolte circa le aspettative e le percezioni degli utenti di un prodotto/servizio.

Esso è ottenuto con il metodo delle medie ponderate secondo la

$$CSI_i = \sum \frac{x_{ij} z_{ij}}{z_{ij}}$$

ove i pedici "i" rappresentano rispettivamente l'i-esimo individuo intervistato e il j-esimo attributo in esame,  $x_{ij}$  il parere sulla qualità percepita riguardo il j-esimo attributo da parte dell'i-esimo utente, e  $z_{ij}$  il giudizio sulla qualità attesa.

E' un indice individuale in cui i giudizi sulle percezioni vengono espressi in una scala tra 1 e 10 e si ponderano secondo una successione di aspettative crescenti di livello da 1 a 3.

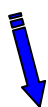
Mediante aggregazioni successive il grado di analisi si sposta dai singoli a classi di utenza più vaste.

Su questa metodologia si fondano l'ECSI (European Customer Satisfaction Index) e l'ACSI (American Customer Satisfaction Index).

Il primo è un indice, ancora in fase di sperimentazione, applicato in 12 paesi europei al fine di valutare la qualità di beni e servizi di comparti merceologici differenti acquistati in Europa e prodotti in qualunque paese, anche non appartenente all'UE. Il progetto pilota è stato promosso nel 1998 dalla Commissione Europea, sviluppato dalla EFQM (European Foundation for Quality Management), dall'EOQ (European Organization for Quality), dal CSI network (un'organizzazione di 8 università europee) e supportato dall'ESOMAR (European Society for Opinion and Marketing Research).

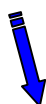
## Il modello ECSI

**Immagine**



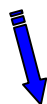
**Aspettative dei**

**clienti**



**Qualità percepita**

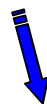
**(fattori tangibili) (fattori intangibili)**



**Valore percepito,**

**rapporto**

**qualità/prezzo**



**Fedeltà**

Altri compiti dell'ECSI sono il dar "Voce" ai consumatori e l'essere di complemento agli attuali indici prestazionali dell'economia europea: benessere, ritorno, stabilità, output economici, come anche competitività con l'estero.

L'American Customer Satisfaction Index rileva il livello di soddisfazione dei consumatori U.S.A. con scadenza trimestrale a partire dal 1994.

E' connesso al trend dei consumi, alla redditività ed alla concorrenza tra le imprese.

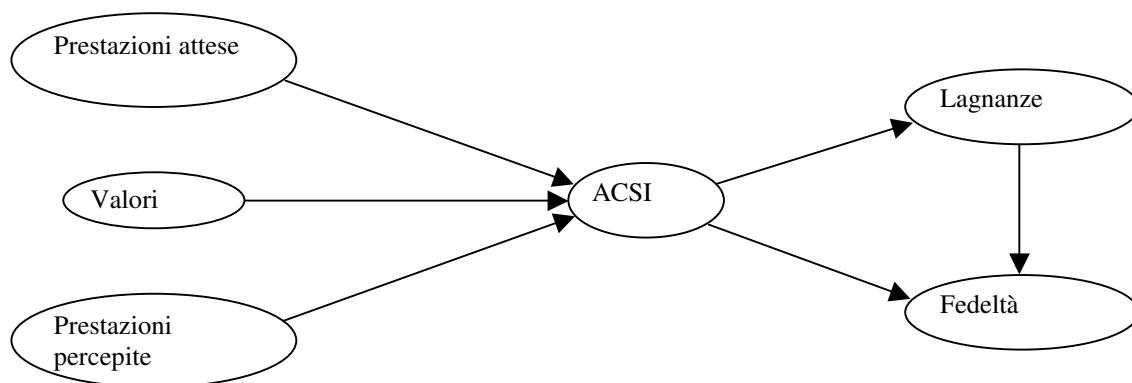
Anche in questo caso i prodotti e/o servizi possono essere anche realizzati fuori dagli U.S.A.

I settori di analisi, in comune con l'ECSI, vanno dall'elettronica di consumo ai cibi, dai trasporti ai supermercati, dai media alle automobili.

L'approccio ECSI sfrutta l'esperienza del metodo ACSI, ma comunque entrambi i modelli sfruttano svariati studi statistici e sono compatibili in virtù dei vigenti commerci internazionali e del diffuso benchmarking.

Le vendite delle aziende considerate dall'ACSI rappresentano il 30-40% del PIL (prodotto interno lordo) statunitense, ed inoltre sono considerate in esso le aziende estere con ampia rilevanza sul mercato.

### Il modello ACSI



Si ricordi che con l'espressione "valore percepito" si intende ancora il rapporto qualità/prezzo.

In conclusione, è possibile affermare che il compiacimento della clientela, al di là della presenza sul mercato di notevoli offerte concorrenziali, sia lo strumento cardine di ogni azienda volto all'espansione della propria quota di mercato, alla crescita dei profitti, al miglioramento dell'immagine, ovvero tutti i fini della moderno quality management nonché i canoni della moderna globalizzazione.

## CAPITOLO 3-

### CUSTOMER RELATIONSHIP MANAGEMENT

#### 3.1 I PRINCIPI DEL CRM: PROCESSI E COMPONENTI

Il Customer Relationship Management è l'insieme di **Processi e sistemi** aventi lo scopo di gestire e mantenere le **relazioni** con i clienti.

Osservando l'evoluzione della management science, si nota che si stanno affermando delle filosofie secondo le quali, partendo dall'assunto che molti aspetti importanti del business devono impiegare diversi anni per realizzarsi e svilupparsi, si conclude che, per valutare un'impresa, non basta misurare solo gli asset finanziari, ma è necessario prestare molta attenzione a ciò che vengono definiti gli human asset. In questo contesto viene ampiamente riconosciuta, per l'ennesima volta, l'importanza delle tecnologie dell'informazione come risabilitanti del Customer Relationship Management (CRM), una filosofia più che un semplice insieme di apparati e di persone, che porta alla "celebrazione" dell'individuo, del singolo cliente.

Dalla fertilizzazione incrociata di nuovi orientamenti culturali, originati dagli scenari competitivi e dalle pervasive potenzialità latenti del CRM, e dall'evoluzione delle tecnologie dell'informazione, l'attività direzionale regionale ne esce trasformata: diviene meno intuitiva e più "scientifica". Le decisioni e gli interventi utilizzano modelli sempre più accurati e metodi di indagine sempre più sofisticati, tanto che oggi si parla sempre più di Business Intelligence, un'area ancora in via di consolidamento ma oggetto di particolari attenzioni da parte degli studiosi.

I presupposti economici del rapporto azienda-cliente stanno cambiando radicalmente ed i modelli guida della produzione e del marketing di massa, creati dalla Società Industriale, stanno cedendo il posto a nuovi paradigmi nei quali la relazione con il cliente è sempre più l'elemento centrale del business. Innanzitutto è importante sottolineare come il "ciclo di vita del rapporto con il cliente" sia diventato un elemento chiave di qualsiasi strategia di business. Gli obiettivi delle imprese, nell'attuale quadro economico, puntano molto sulla fidelizzazione sulla massimizzazione del grado di soddisfazione del cliente come requisiti fondamentali per

migliorare la redditività dell' azienda e aumentare il valore dell' azienda per gli azionisti. L' obiettivo non è più solo quello di acquisire nuovi clienti, ma anche di ottimizzare l' interazione in tutti i suoi aspetti.

In secondo luogo, gli strumenti e le tecnologie rese generalmente disponibili grazie ad Internet, alle intranet e al data warehousing consentono una migliore gestione delle informazioni strategiche relative al cliente e, di conseguenza, di ogni singolo rapporto azienda-cliente. Mentre un tempo era impensabile personalizzare le strategie di marketing e di vendita prodotti nei confronti di ogni singolo cliente, oggi gli strumenti di CRM offrono in tal senso nuove opzioni e opportunità.

La value chain delle imprese che hanno adottato una strategia di business incentrata sull' attenzione al cliente è passata dal vecchio principio *design-build-sell* al nuovo ordine *sell-redesign-rebuild*. In altre parole, l' impresa oggi deve essere organizzata in base ad una visione focalizzata sul cliente piuttosto che sul prodotto.

Nell' ambito di questa nuova cultura di business, l' implementazione di un sistema CRM diventa essenziale, così come l' esigenza di capitalizzare il patrimonio di conoscenze aziendali e la fedeltà del rapporto con i clienti sta dando vita ad applicazioni aziendali strategiche.

La concentrazione dei processi di business sul rapporto con il cliente mira principalmente a creare ed estrarre valore dalle interazioni al fine di migliorare la redditività aziendale. Infatti, in termini di valore finanziario, la gestione del rapporto con il cliente consente un aumento delle entrate per cliente e della produttività; a livello organizzativo, le applicazioni CRM consentono di migliorare i processi aziendali associando i servizi e il supporto post-vendita alle operazioni di vendita e marketing.

Inoltre, in termini di offerta di prodotti e servizi, l' implementazione di un sistema CRM consente alle aziende di presentare ai propri clienti un' offerta personalizzata; infine, in termini di posizione di mercato, si fornisce alle aziende un significativo elemento di differenziazione ed un indiscusso vantaggio competitivo.

Una strategia CRM-oriented fornisce l' opportunità per integrare le applicazioni e la tecnologia

esistente ed ottimizzare l' utilizzo degli strumenti a disposizione in una prospettiva operativa. Come mostra un sondaggio condotto da IDC e da CAP Gemini nel corso del 1999, le aziende che stanno implementando un' applicazione CRM o intendono farlo nei prossimi due anni hanno già compiuto consistenti investimenti. Il 73% di tali aziende dispone infatti di call center, il 69% di canali Internet e il 61% di soluzioni di front-office, mentre tecnologie più sofisticate, quali siti web interattivi, data warehousing, CTI, etc, sono già in cantiere e saranno implementate nel corso dei prossimi due anni.

Le aziende intervistate danno la massima priorità agli investimenti nelle funzioni di supporto, dato che si tratta di funzioni immediatamente operative, mentre le tecnologie e gli strumenti più interattivi saranno implementati gradualmente in una fase successiva. Questo spiega perché i progetti CRM condotti finora non presentano necessariamente costi proibitivi. Secondo i risultati del sondaggio, l' investimento medio per un progetto CRM (strumenti, software, servizi, consulenze, etc) è pari a di 3,1 milioni di dollari. I budget relativamente contenuti rivelano la presenza di un gran numero di progetti di dimensioni ridotte, che coinvolgono l' integrazione di un call center e di una soluzione di frontoffice a livello di singolo reparto. Le previsioni di un ritorno dell' investimento non sono entusiasmanti. I vantaggi anticipati dall' aziende sembrano essere moderati e stimati in termini di conservazione del tasso di redditività del capitale investito. Tali risultati riguardano aziende operanti in settori industriali maturi, dove la crescita è costosa e difficile.

Il contatto con il cliente è chiaramente un elemento fondamentale per il successo di qualsiasi azienda, dato che, se non si interagisce con i propri clienti, non si può vendere loro alcun prodotto. Essere semplicemente in grado di portare a termine una transazione commerciale di base, quale può essere l' acquisizione di un ordine, non è più sufficiente per competere. La qualità e la completezza del contatto con il cliente sono diventati gli elementi distintivi fondamentali per qualsiasi azienda che operi in un mercato competitivo. L' infrastruttura che sta alla base del contatto con il cliente coinvolge tutti i sistemi, le informazioni e i processi necessari per assicurare il successo totale dell' interazione. La qualità della relazione può essere significativamente migliorata se tutto il personale addetto alle vendite, al marketing e al supporto, ovvero chiunque sia in relazione con il cliente, può accedere alle stesse informazioni

relative alle preferenze e ai contatti precedenti. Queste informazioni devono essere disponibili in maniera immediata ogni volta e in ogni luogo sia necessario accedervi e devono essere estremamente accurate e complete e sempre aggiornate.

Un ambiente di contatto con il cliente efficace deve presentare i seguenti requisiti fondamentali:

♣ **Attenzione alle esigenze**

L' ambiente deve fornire un unico punto di contatto per il cliente per tutti i prodotti e servizi deve aiutare ad applicare i processi di business aziendali in modo che qualsiasi aspetto venga affrontato correttamente fin dall' inizio della relazione.

Deve inoltre utilizzare tutti i dati noti relativi al cliente, in modo che quest' ultimo non si costretto ogni volta a fornire le stesse informazioni. Fatto ancora più importante, deve essere il cliente a decidere come interagire con l' azienda, scegliendo tempi e modi dell' interazione. Infine, i clienti più preziosi devono essere trattati con i guanti, ottenendo un elevato livello di servizio.

♣ **Miglioramento della conoscenza**

E' fondamentale memorizzare e capitalizzare ogni interazione con il cliente. L' ambiente deve facilitare il dialogo in modo che l' azienda sia in grado di accumulare informazioni preziose.

♣ **Sfruttamento della conoscenza**

La conoscenza del cliente da parte dell' azienda deve essere utilizzata in maniera intelligente ogni volta che si interagisce. Ogni contatto deve essere visto come un' opportunità di vendita non come un evento casuale, offrendo il prodotto più appropriato al prezzo più appropriato.

Le informazioni strategiche sul cliente dovrebbero inoltre essere utilizzate a vantaggio delle varie funzioni all' interno dell' azienda (Si consideri, per esempio, lo stato di solvibilità di cliente: deve essere immediatamente disponibile all' agente) .



### ♣ **Maggior efficienza**

In una realtà economica altamente concorrenziale, il contenimento dei costi rimane una delle principali leve d'azione per la redditività dell'impresa. Questo nuovo ambiente "cliente-centric" deve pertanto funzionare in modo estremamente efficiente, permettendo di ottimizzare l'utilizzo di personale grazie all'utilizzo dell'ICT.

### ♣ **Maggior flessibilità**

La flessibilità dell'ambiente di contatto con il cliente svolgerà un ruolo sempre più centrale. L'ambiente dovrebbe pertanto supportare la rapida introduzione di nuovi sistemi di produzione per ridurre il time-to-market e consentire all'azienda di sfruttare nuovi canali, quali per esempio Internet.

Sono necessarie capacità non indifferenti per integrare in maniera appropriata tutti i componenti di un sistema CRM nell'infrastruttura ICT e nelle strutture informative esistenti in un'azienda.

Affinché i dipendenti addetti al rapporto con il cliente siano produttivi, è fondamentale che possano accedere alle applicazioni CRM necessarie per servire il cliente in maniera adeguata. In tal senso, è necessario disporre di un'unica interfaccia utente, che consenta di accedere tutti i sistemi richiesti. Una soluzione che prevede modalità d'accesso diverse sarebbe più lenta, più soggetta ad errori e anche molto più costosa da far funzionare e gestire.

L'obiettivo di qualsiasi rapporto con un cliente dovrebbe essere quello di trarre un vantaggio in termini di aumento del fatturato (ovvero delle vendite), riduzione dei costi ed una gestione più efficiente dei processi aziendali. Il CRM, uno degli elementi chiave di una strategia di business orientata al cliente, è difficile da adottare senza ridisegnare i sistemi legacy (orientati al prodotto) attualmente esistenti all'interno delle imprese.

Un fornitore di soluzioni ICT può contribuire notevolmente alla soluzione del problema se è in grado di fornire una combinazione di capacità e strumenti, quali:

\* Personale in grado di sviluppare software e interfacce che integrino e completino i sistemi

legacy esistenti.

\* Processi che possano essere implementati utilizzando elementi software coerenti che si integrino con le applicazioni di business sottostanti sia da un punto di vista tecnico sia di logica di business.

\* Partner che consentano di utilizzare prodotti "best of breed" all' interno dell' ambiente assicurino la disponibilità delle competenze necessarie.

\* Piattaforme che forniscano il supporto tecnico necessario alla disponibilità e all' accessibilità utilizzando i sistemi operativi più diffusi

Non esiste una soluzione unica, pacchettizzata ed esportabile per i progetti di CRM. Le esigenze specifiche, le caratteristiche e la cultura dell' azienda rendono unico ogni progetto

I processi che devono essere attivati sono i seguenti:



L' acquisizione dei dati costituisce una fase imprescindibile di un progetto di CRM, mentre la presenza e le caratteristiche delle altre fasi dipendono dalle scelte strategiche e dall' avanzamento del progetto stesso. Si tratta di un processo ricorsivo che riguarda ogni momento dell' interazione con l' utente: registrazione iniziale, richieste esplicite nel corso dell' interazione, tracking del comportamento, ecc.. Conoscere il cliente nella sua globalità significa infatti disporre di informazioni progressive nel tempo e relative a sfere diverse, riguardanti sia il cliente in quanto persona sia il cliente in quanto cliente:

- dati socio-demografici, capacità reddituale, esposizione ai media, atteggiamenti valoriali, stile di vita, ecc.

- storia delle transazioni o dell' utilizzo del servizio, lamentele ed esiti, storico proposte statistiche di comportamento, costi relativi all' operatività del cliente, ecc.

I dati acquisiti potranno poi essere utilizzati in modo atomico per campagne specifiche o potranno entrare in gioco in fase di **analisi** per ottenere i dati sintetici necessari alla produzione di conoscenza sul cliente.

La *Business Intelligence* del sistema si colloca in questa fase. Si tratta di una fase importante in quanto consente di estrarre conoscenza dai dati acquisiti e di conseguenza di massimizzare le successive fasi di **profilazione, personalizzazione e ottimizzazione** del servizio. L' analisi, effettuata con *DataMining* (Berry et al., 1997; Michalski et al., 1998; Agrawal et al., 1993), TextMining, OLAP, sistemi a regole, ecc. deve produrre sia indicatori sintetici sia modelli predittivi. Es. individuazione delle correlazioni tra la chiusura di un contratto e specifiche tipologie di insoddisfazione o di caratteristiche dei clienti.

La **profilazione** dei clienti copre uno spettro che va dalla segmentazione al one-to-one mktg impiegando tecniche e metodologie diverse:

- Segmentazione basata su criteri e caratteristiche impostati a priori e supportati da sistemi OLAP. Es. criteri precisi per campagne mirate, criteri derivanti da ricerche e fonti esterne.
- Classificazioni sulla base di tecniche di DataMining: *cluster analysis*, pattern recognition, *reti neurali* (che permettono di individuare associazioni e segmenti imprevedibili).
- Modellazione utenti basata su conoscenza: conoscenza stereotipale, reti semantiche e conoscenza operativa (business rules).

**Interazione personalizzata** e integrata con l' utente, sulla base di quattro dimensioni:

- il modello o profilo dell' utente
- gli obiettivi e sotto-obiettivi specifici (prodotti da promuovere, budget da raggiungere, ecc.)
- le caratteristiche del canale utilizzato

- la sincronizzazione e la coerenza con le azioni condotto sugli altri media

**L'ottimizzazione** si realizza tramite feedback sul servizio e sulle linee strategiche di tutte le funzioni aziendali: marketing, customer care, logistica, produzione, R&S, ecc.

### 3.2 CUSTOMER BASED VIEW (CONOSCENZA)

Contestualmente all' evoluzione degli stati concorrenziali, che in un numero sempre maggiore di mercati vanno assumendo i caratteri dell' ipercompetizione (D' Aveni, 1994; Valdani, 1995 e 1997; Ancarani, 1999), è significativamente aumentata l' attenzione che le imprese dedicano al comportamento dei propri clienti, con l' obiettivo di sviluppare e consolidare le relazioni di mercato.

L' accresciuto interesse verso le interazioni fra domanda e offerta è stato spiegato ricorrendo all' evidenza della crescente complessità tecnologica, concorrenziale e relazionale (Busacca, 1994; Busacca, Grandinetti e Troilo, 1999). L' evoluzione delle forme di concorrenza, la progressiva saturazione di molti mercati, e le strutturali modificazioni dei processi di scambio, in parte indotte dall' emergere dell' economia digitale, infatti, stanno obbligando le imprese ad adottare una prospettiva di prioritario orientamento allo sviluppo e al consolidamento della relazione con i clienti, recentemente definita da Valdani e Busacca (Caldani e Busacca 1999) *customer-based view*.

A partire dai primi anni ottanta, peraltro, anche le ricerche accademiche sulla domanda hanno riguardato, con frequenza crescente, le fasi del comportamento d' acquisto successive alle iniziali manifestazioni di preferenza verso una marca o un' insegna, concentrandosi, seppure con differenti approcci, sui fenomeni che definiscono lo stato e la dinamica delle relazioni fra impresa e cliente - *postconsumption research* (Oliver, 1997) -. Parallelamente all' ormai consolidato filone di studi *subconsumer behavior* - sempre più ricco di contributi interdisciplinari - si sta quindi sviluppando un nuovo insieme di modelli specificamente riferibili al comportamento d' acquisto dei clienti, volti cioè a descrivere e interpretare il *customer buying behavior*.

La consapevolezza che la "Teoria del Consumatore" e i modelli sul comportamento della domanda non avessero adeguatamente approfondito i fenomeni conseguenti alla decisione d' acquisto è stata segnalata da diversi studiosi (Day e Wensley, 1983; Gronroos, 1994a e 1994b). Tutta la manualistica sul comportamento d' acquisto riserva una marginale attenzione al processo valutativo post-acquisto. E solo in tempi

relativamente recenti, alcuni autori, da sempre impegnati nello studio del comportamento della domanda, hanno evidenziato che l'estensione delle ricerche all'interazione prodotto-consumatore successiva all'acquisto è da considerarsi fondamentale, per generare nuova conoscenza su costrutti di importanza critica quali la *customer satisfaction* e la *customer loyalty* (Tse, Nicosia e Wilton, 1990).

Nei modelli più consolidati di analisi della domanda, quindi, non vengono approfondite le complesse dinamiche del comportamento d'acquisto del cliente, nel corso del ciclo di vita della sua relazione con l'impresa, ma ci si limita a descrivere il sistema valutativo post-acquisto come l'insieme delle percezioni da cui ha origine il flusso di retroazione sul sistema motivante e su quello percettivo, interpretando il processo di *feedback* esclusivamente sulla base del costrutto "soddisfazione/insoddisfazione" per l'esperienza d'uso del prodotto (Busacca, 1990 e 1994; Costabile, 1996a). Le carenze di modelli teorici condivisi, peraltro già evidenziate da Iacobucci et al. (1992), permangono, nonostante i primi consistenti sforzi di indagine sulle determinanti della *customer satisfaction* e del comportamento post-acquisto risalgano alla seconda metà degli anni settanta (Hunt, 1977; Oliver, 1977; Olson e Dover, 1979); e nonostante le numerose sperimentazioni e le diverse applicazioni manageriali registrate a partire dai primi anni ottanta e proseguite poi per tutto il decennio successivo.

E' necessario proporre modelli che sintetizzino concettualmente alcuni dei principali risultati conseguiti nell'ambito delle ricerche e delle sperimentazioni sul sistema delle valutazioni post-acquisto e sul comportamento dei clienti, facendo in prevalenza riferimento a due filoni di studio: quelli sulla *customer satisfaction* e quelli sul *marketing relazionale*.

Il primo ha avuto origine proprio dall'esigenza di approfondire la comprensione dei processi valutativi post-acquisto. Gli esperimenti volti a verificare le determinanti e le conseguenze della soddisfazione del consumatore, nonché il legame fra soddisfazione e comportamenti di riacquisto, hanno riguardato diversi ambiti settoriali e variegati contesti di consumo, pervenendo a risultati non sempre convergenti, ma in definitiva interpretabili alla luce della comparazione fra aspettative e percezioni di performance. Le dinamiche che caratterizzano tale processo di comparazione sono descritte dal

cosiddetto "paradigma della conferma/disconferma della aspettative" (Cardozo, 1965; Oliver, 1980; Iacobucci e al., 1992; Costabile, 1996a e 1998; Oliver, 1997; Fournier e Mick, 1999). Analogamente, gli studi sul marketing relazionale, e le ricerche volte a comprendere le dimensioni cognitive, emotive e comportamentali che definiscono il concetto di "relazione", e conseguentemente l'essenza paradigmatica del marketing relazionale, sono stati numerosi e decisamente pervasivi, partendo dagli scambi fra imprese industriali, ed estendendosi poi alle relazioni distributive, ai servizi e ai beni di consumo.

I due filoni di ricerca hanno in realtà concentrato la loro attenzione sul medesimo processo - lo sviluppo delle relazioni di mercato - seppure partendo dai due estremi opposti: lo studio delle percezioni di soddisfazione, interpretabili quali primo stadio di tale processo, e quello delle relazioni collaborative di lungo periodo, ritenute il traguardo verso il quale tutte le relazioni di mercato dell'impresa dovrebbero idealmente tendere.

Gli studi sulla customer satisfaction, infatti, sono stati sviluppati nell'ambito della ricerca sul consumatore, con l'obiettivo di indagare le conseguenze, cognitive, emotive e comportamentali della scelta d'acquisto. Tali studi hanno esaminato le determinanti della percezione di soddisfazione e le sue conseguenze, talvolta validando, altre volte falsificando - ma solo in parte - il paradigma della conferma/disconferma delle aspettative. E' ormai ampiamente condiviso, tuttavia, che la customer satisfaction sia il fondamentale antecedente della fiducia e della fedeltà, e pertanto all'origine di tutte le forme di relazione e di valore dell'impresa (Costabile, 1996b e 1998).

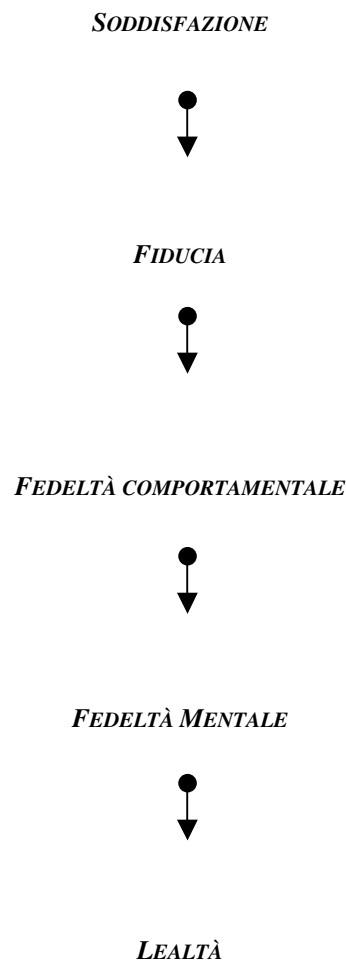
Gli studi sul marketing relazionale, invece, si sono concentrati sull'evidenza del contenuto prevalentemente sociale di alcune relazioni di mercato, ritenendo inadeguato il paradigma del marketing mix quale modello interpretativo degli scambi ripetuti nel tempo, e tentando di identificare i costrutti cognitivi e comportamentali ad essi sottostanti, al fine di isolarne le fondamentali determinanti (Hakansson, Johanson e Wootz, 1976; Dwyer, Schurr e Oh, 1987; Morgan e Hunt, 1994; Ganesan, 1994).

Analizzando le conseguenze della customer satisfaction, successive ai processi di scambio, e gli antecedenti della solidità delle relazioni longeve e collaborative, entrambi

i filoni di studio hanno sviluppato ricerche originali proprio sul comportamento d' acquisto dei clienti. E in particolare, sul legame fra soddisfazione, fiducia e fedeltà, nell' ambito delle ricerche sulla customer satisfaction; sulla fiducia, sulla fedeltà e sulla natura delle relazioni collaborative fra acquirente e venditore, nell' ambito del marketing relazionale.

I principali costrutti emersi dallo studio delle conseguenze della soddisfazione del cliente e degli antecedenti delle relazioni collaborative di lungo periodo sono: la soddisfazione; la fiducia; la fedeltà comportamentale (la ripetizione d' acquisto); la fedeltà mentale; la customer loyalty e la partnership collaborativa; anche se la maggior parte degli studi sul *continuum* relazionale hanno riguardato soddisfazione, fiducia e cooperazione.

*Figura - Il continuum relazionale*





Il tema del valore economico delle relazioni di mercato è oggetto di una crescente attenzione da parte degli studiosi dell'economia d'impresa e degli operatori aziendali, in gran parte riconducibile alla radicale trasformazione dei *business model* innescata dalle nuove tecnologie e alla conseguente focalizzazione degli investimenti nell'area della *internet economy* sullo sviluppo della *customer base*. Quale contributo all'approfondimento del tema indicato, in un recente lavoro (Valdani e Busacca, 1999) è stata proposta una prospettiva teorica, la *customer-based view* (CBV), le cui proposizioni centrali correlano il valore che l'impresa è capace di generare per i propri clienti, il valore di questi ultimi e il valore del capitale economico. Tale prospettiva si inserisce nel filone di studi volto ad approfondire il legame fra orientamento al mercato, soddisfazione del cliente e business performance (Buzzel e Gale, 1987; Narver e Slater, 1990; Jaworski e Kohli, 1993; Deshpande, Farley e Webster, 1993; Narver, Jacobson e Slater, 1993; Anderson, Fornell e Lehmann, 1994; Gale, 1994; Slater e Narver, 1995). Sul piano teorico, oltre che sui contributi indicati, essa si fonda sulla *resource-based view* (Penrose, 1959; Rumelt, 1984; Wernerfelt, 1984; Barney, 1986 e 1991; Itami, 1987; Dierickx e Cool, 1989; Grant, 1991 e 1996; Amit e Schoemaker, 1993; Peteraf, 1993) e sulla *teoria di creazione del valore*, nella versione proposta dalla Scuola italiana, (Guatri 1991 e 1992; Guatri e Massari, 1992), proponendo la focalizzazione di tali costruzioni concettuali sul ruolo centrale del cliente nel processo di sviluppo del capitale economico.

In particolare, la CBV riconduce i meccanismi causali che spiegano il legame fra le risorse dell'impresa e il suo successo economico/competitivo attraverso i nessi di causalità esistenti fra il valore generato per i clienti, il livello di customer satisfaction, la consistenza del portafoglio delle relazioni di mercato, il valore dello stesso, la dimensione del capitale economico e la dotazione di risorse immateriali dell'impresa.

Nella CBV la creazione di valore si pone quale requisito inderogabile per la sopravvivenza e il successo dell'impresa, requisito il cui soddisfacimento è assicurato dalla focalizzazione dei processi aziendali sul valore offerto ai clienti. Ciò infatti garantisce lo sviluppo del patrimonio aziendale, attraverso le relazioni circolari che legano gli intangibile customer-based alle risorse fondate sulla conoscenza e sulle relazioni (fiducia) con altre categorie di stakeholder. Peraltro, proprio alla diversa

capacità di presidio della sequenza “valore per il cliente - risorse intangibili - valore del capitale economico” la CBV riconduce le differenti performance delle imprese. Essendo la suddetta capacità il risultato dello stock iniziale di conoscenza e di fiducia, emerge con evidenza il collegamento diretto fra risorse intangibili e risultati aziendali. Da questo punto di vista, la CBV risulta coerente con le assunzioni di base della teoria di creazione di valore e della resource-based view. In particolare, ci si riferisce al riconoscimento che il valore del capitale economico “è legato non tanto ai flussi (di reddito, di cassa) che l’impresa produce nell’immediato, ma alle potenzialità accumulate di produrre in futuro, e per lungo tempo, tali flussi positivi” (Guatri, 1992: 6), nonché alla concezione del processo di creazione di nuovo valore economico quale condizione di esistenza dell’impresa (Vicari, 1991), e alla correlazione tra le differenze nella dotazione di risorse (scarse e difficili da imitare) e le differenze di performance. I principi sui quali si fonda la CBV sono rappresentati dalla considerazione del cliente quale fonte primaria della generazione di valore economico e dal riconoscimento della centralità della customer satisfaction quale linguaggio organizzativo indispensabile per presidiare adeguatamente tale fonte di valore.

Pertanto, il *customer-based view* è una prospettiva teorica, le cui proposizioni centrali correlano il valore che l’impresa è capace di generare per i propri clienti (value proposition), il valore di quest’ultimi (customer equity) e il valore del capitale economico (equity d’impresa). Tale prospettiva si inserisce nel filone di studi volto ad approfondire il legame tra orientamento al mercato, soddisfazione del cliente e business performance.

L’affermazione di standard universali e aperti, quali internet, unitamente alla diffusione e al perfezionamento delle tecnologie di comunicazione (fisse e mobili, wire e wireless, terrestri e satellitari) hanno determinato l’esplosione della connettività e la progressiva separazione delle cose fisiche dalla conoscenza, che in passato erano indissolubilmente legate. La divergenza delle economie delle cose da quella della conoscenza ed informazione, in questi ultimi anni, si è sviluppata nella successione di tre distinte fasi. Nella prima le imprese hanno investito significative risorse nei processi di generazione della conoscenza e nella sua trasformazione da tacita in esplicita per poterla rendere pervasiva, disponibile e appropriabile per misurare, controllare e coordinare i processi

aziendali. Il possesso dell'informazione era utile per migliorare l'efficacia e l'efficienza dei processi aziendali allo scopo di migliorare la soddisfazione della clientela e la redditività. La seconda fase ha impegnato l'impresa nel primo tentativo di parziale virtualizzazione della catena del valore. Per acquisire nuovi vantaggi competitivi, incrementando l'efficienza dei processi interni e migliorando la soddisfazione della clientela, l'impresa ha iniziato a sperimentare il trasferimento di alcune fasi della sua catena fisica del valore nel market space. Con tale sperimentazione l'impresa attivava il processo di separazione dell'economia delle cose fisiche da quella della conoscenza ed informazione, generando le prime ma frammentate catene virtuali del valore, che operavano comunque integrate in quelle fisiche, con lo scopo di migliorarne la performance e la sua competitività. La terza fase, oggi in atto, esprime la nuova traiettoria competitiva che mira a gestire contemporaneamente l'economia delle cose e l'economia della conoscenza, riconoscendone ed enfatizzandone gli elementi di distintività. Nella nuova economia, l'attuazione della CBV implicherà pertanto una profonda riconfigurazione dei processi che governano la creazione di valore, che incentiverà le cose fisiche e il contenuto della conoscenza e dell'informazione a seguire con maggiore indipendenza la propria economia, con lo scopo, comunque, di sostenersi reciprocamente ogni qual volta una sarà necessaria all'altra.

Se ogni impresa è predestinata a dover competere contemporaneamente nel mondo delle cose fisiche e in quello virtuale della conoscenza e dell'informazione (il market space), l'aspetto più critico risiede nella capacità di generare valore sia nel primo sia nel secondo, perché le economie dei due mondi sono profondamente diverse. Tale diversità, se da un lato accresce la rilevanza del cliente e quindi della CBV, dall'altro subordina l'attuazione di questa prospettiva al pieno sfruttamento del potenziale coevolutivo intrinseco al reticolo delle connessioni presenti ed operanti nel market space. La realizzazione di reti digitali del valore (digital value network) rappresenta la modalità per attivare il suddetto potenziale. I digital value network esprimono infatti comunità nelle quali imprese e clienti, attraverso l'interazione consentita dalle nuove tecnologie, possono ricercare nuove configurazioni e originali ambiti di integrazione delle proposizioni di valore. Le relazioni che si generano in un digital value network sono fluide, nascono e si dissolvono per riaggregarsi in base alle nuove dinamiche espresse

dall'incontro della domanda con l'offerta. Il fornitore di ieri può divenire il cliente di oggi e forse il concorrente di domani. Nella rete digitale del valore si esprimono con maggiore evidenza il significato e la strategicità del concetto di "coevoluzione".

In questa rete infatti imprese complementari e concorrenti attivano relazioni cooperative e competitive che stimolano tutti i soggetti coinvolti a utilizzare le rispettive competenze e risorse per inventare nuove modalità e soluzioni che favoriscano la generazione e il trasferimento di valore ai clienti. I partecipanti al network si attivano e operano gli uni a vantaggio degli altri, nella consapevolezza che il ruolo svolto da ciascun attore è fondamentale per il conseguimento degli obiettivi condivisi. La rete digitale del valore si esprime attraverso l'interconnessione e lo stretto rapporto tra le varie catene virtuali del valore che in essa operano per massimizzare la soddisfazione dei clienti. Questi rapporti sono resi possibili da piattaforme digitali funzionali, basate su standard più o meno universali, e da agenti facilitanti che operano nella rete quali infomediary, facilitando l'accesso alle informazioni, ai contenuti e alla conoscenza. In sintesi, nella nuova economia la realizzazione del circolo virtuoso della CBV dipenderà in misura crescente dalla capacità di attivare e gestire una rete del valore sempre più estesa e articolata. Su questo obiettivo dovranno presumibilmente concentrarsi i maggiori sforzi di competence building delle imprese. Infatti, a dispetto del generalizzato riconoscimento della superiorità della rete, quale soluzione organizzativa per il governo della complessità, ancora poche imprese, soprattutto nel nostro Paese, hanno affiancato alla propria catena fisica del valore una catena virtuale e un numero ancora inferiore è stato in grado di attivare un digital value network.

Le linee di azione di un CBV si riferiscono:

1. al sistema delle rilevazioni contabili ed extra contabili necessarie per la quantificazione del valore dei clienti;
2. alla verifica della distribuzione di tale valore;
3. all'articolazione della value proposition offerta dall'impresa in coerenza con la distribuzione suddetta.

In relazione al primo punto occorre anzitutto precisare che il valore economico di un cliente deve essere valutato considerando non soltanto il flusso di profitti derivante dai suoi acquisti, ma anche il contributo fornito in termini di attivazione di nuove relazioni e di sviluppo delle risorse aziendali di conoscenza. La misurazione del valore economico dei clienti presuppone pertanto la costruzione di un articolato sistema di rilevazioni, le cui unità di analisi sono costituite dai singoli clienti, concernenti:

- ♣ • • i consumi totali della categoria in cui rientra il prodotto identificato dalla marca;
- ♣ • • le quantità acquistate nel periodo;
- ♣ • • i margini prodotti da ciascun acquisto;
- ♣ • • la durata della relazione, espressa in termini di numero di periodi di acquisto;
- ♣ • • il tasso di sviluppo dei volumi di acquisto
- ♣ • • i costi sostenuti per l'acquisizione, lo sviluppo e la conservazione del cliente;
- ♣ • • la sua propensione ad attivare comunicazioni informali positive sull'impresa;
- ♣ • • il suo potere di influenza nell'ambito dei gruppi di riferimento;
- ♣ • • la sua disponibilità a collaborare con l'impresa, ad esempio ai fini della sperimentazione di innovazioni (di prodotto e di marketing) o della produzione di maggiore conoscenza sulle proprie esigenze funzionali e simboliche.

Com'è agevole notare si tratta di rilevazioni analoghe a quelle da tempo proposte dagli studiosi di marketing industriale per l'analisi delle relazioni *business to business*. La loro disponibilità nel caso di beni o servizi di basso valore unitario e ad acquisto ricorrente è evidentemente subordinata alla piena attivazione del potenziale di connessione delle nuove tecnologie che consente, a costi contenuti, di interagire con le singole unità di consumo. D'altro canto, la costruzione di ampi *customer data base* si rivela indispensabile non solo per determinare il valore delle relazioni di mercato (e orientare di conseguenza l'allocazione degli investimenti di marketing), ma anche per ricostruire i profili di comportamento dei clienti e sfruttare appieno sia il potenziale di apprendimento intrinseco alle relazioni di mercato sia le opportunità di creazione di valore dei sistemi produttivi flessibili. I nuovi ambienti comunicativi internet-based e le

tecniche di data mining già oggi permettono di perseguire concretamente tali obiettivi e nel prossimo futuro la capacità di realizzare un circolo virtuoso fra *customer knowledge management*, *customer connecting technologies* e *customer economics* rappresenterà una delle principali determinanti del vantaggio competitivo conseguibile.

Al di là della puntuale verifica della capacità dell'impresa di accrescere la dimensione del capitale economico, la misurazione della customer equity è essenziale ai fini della verifica delle opportunità connesse alla segmentazione del portafoglio clienti (sulla base, appunto, del loro valore) e alla conseguente articolazione della value proposition, il che introduce alla seconda e alla terza implicazione manageriale, in precedenza citate.

In presenza di una elevata varianza della customer equity, lo sviluppo del capitale economico implica la *focalizzazione* o la *personalizzazione* della value proposition, in funzione dell'uniformità o della eterogeneità dei bisogni espressi dai clienti chiave. In entrambe le situazioni indicate risultano ovviamente critiche la chiara identificazione di tali clienti e la conquista della loro fedeltà, attraverso approcci di marketing relazionale e di marketing one to one.

Qualora sia il valore economico dei clienti che le esigenze di acquisto/consumo risultino relativamente uniformi la standardizzazione della value proposition si rivela una scelta obbligata. In questo caso gli spazi per la creazione di valore si restringono sensibilmente ed assume grande rilevanza la progressiva ottimizzazione del rapporto qualità/prezzo, attraverso il presidio dell'integrità del prodotto e la ricerca di continui incrementi di efficienza. Infine, nel caso in cui l'eterogeneità della customer base sia elevata con riferimento ai bisogni e modesta in relazione al valore economico delle singole relazioni, si impone la differenziazione della value proposition in funzione delle esigenze specifiche del segmento obiettivo. In questa fattispecie assumono grande rilevanza le conoscenze di marketing che orientano le strategie di segmentazione della domanda. Per cogliere appieno i vantaggi connessi alla disomogeneità dei bisogni è infatti necessario sofisticare il tradizionale approccio alla segmentazione, procedendo all'aggregazione dei consumatori sulla base di variabili causali e quindi alla puntuale identificazione degli stessi sulla base di variabili descrittive.

### 3.3 LEGAMI TRA CUSTOMER VALUE- CUSTOMER SATISFACTION- CUSTOMER LOYALTY – PROFITTO

Sviluppare le relazioni con i clienti, dall' attivazione alla soddisfazione, sino alla *customer loyalty*, richiede la capacità di gestire l' interazione assicurando che la percezione di valore sia sempre in linea con le aspettative. Non si tratta tuttavia di un semplice problema di gestione della customer satisfaction, ma di una più complessa capacità di interpretare differenti configurazioni di valore, al quale il cliente mostra sensibilità nelle diverse fasi del ciclo di vita delle relazioni.

La percezione di valore da parte del cliente nel tempo, infatti, varia in funzione della variabile tempo e può riguardare:

- il valore atteso, in termini relativi, ossia il rapporto fra i benefici attesi e i costi che si ritiene di dover sostenere per l' acquisizione e il godimento dei predetti benefici, e la cui percezione è influenzata dal confronto con le alternative disponibili;
- il valore percepito dopo l' acquisto e l' uso, generalmente rapportato al valore atteso per derivarne la percezione di soddisfazione o insoddisfazione, e quindi adottato quale riferimento prevalente per valutare l' esperienza d' acquisto e di consumo, e dunque l' affidabilità dell' impresa;
- il valore percepito in termini comparativi dopo le prime esperienze d' uso, vale a dire confrontato con le alternative d' offerta prese in considerazione nel corso del ciclo di vita della relazione. Tale configurazione viene denominata "valore monadico" per enfatizzarne la componente di soggettivismo determinata dalla prospettiva strettamente individualistica del cliente (monade) che conduce la valutazione, comparando esperienze (certe), maturate nel corso della relazione con l' impresa e basate sulla percezione dei benefici ottenuti (Gardiel et al., 1994), e aspettative (aleatorie), che riguardano le eventuali alternative d' offerta. La percezione di valore monadico, pertanto, ha quale riferimento l' offerta dell' impresa verso la quale si dimostrano forme di ripetizione d' acquisto, ed emerge dalla comparazione del valore sperimentato con il valore atteso nell' ipotesi di transizione (*brand switching*);

- il valore equità, ossia il rapporto fra il valore che il cliente ritiene di aver ottenuto (benefici/costi) e quello che ritiene di aver generato per l'impresa (ricavi/ costi), nel corso della "storia" della relazione. Tale configurazione può essere definita "valore diadico", per evidenziarne la prevalente natura di comparazione interna alla diade "cliente-impresa", nella prospettiva dell'equità seriale che è stata sperimentata (percepita) nel corso di una specifica relazione.

La percezione e la rilevanza di tali differenti, seppur correlate, configurazioni di valore sono soggette a variazioni lungo il ciclo di vita delle relazioni. La loro dinamica definisce, pertanto, quattro principali fasi del processo di sviluppo della relazione, che conduce alla *customer loyalty*:

- I. la fase della soddisfazione e dell'accumulazione di fiducia;
- II. la fase della fiducia e della fedeltà comportamentale;
- III. la fase della fedeltà mentale;
- IV. la fase della lealtà.

#### *I. La fase della soddisfazione e dell'accumulazione di fiducia*

La preferenza da cui ha origine la scelta di beni e servizi è in genere fondata su una percezione di valore differenziale che il cliente ritiene di poter ottenere in seguito al loro acquisto. Come noto, il valore per il cliente è definito dal rapporto fra i benefici attesi e i diversi tipi di costo - sacrifici in senso lato - che devono essere sostenuti per acquisire e godere dei predetti benefici in associazione a una data offerta (prodotto, servizio, marca o impresa):

$$V = B/S$$

La scelta d'acquisto, pertanto, si fonda sulle aspettative di valore, e in particolare sulla percezione di capacità dell'impresa nell'offrire i benefici ricercati meglio dei concorrenti.

Come i numerosi studi sulla *customer satisfaction* hanno dimostrato, dalla congruenza fra valore atteso e valore percepito, in seguito all'acquisto e all'esperienza d'uso, ha



origine la percezione di soddisfazione. Tale percezione rappresenta un "flusso", prodotto a seguito di ogni interazione che il cliente ha con l'impresa, o con uno specifico prodotto dell' impresa (ad esempio ogni qualvolta viene utilizzato un personal computer o un servizio di telefonia cellulare). Tale "flusso" - consapevolmente o inconsapevolmente - alimenta uno "stock": la fiducia, intesa quale pregiudizio (atteggiamento) riguardante la capacità dell' impresa (o del prodotto, o della marca, o dell' insegna, ecc.) di offrire un valore congruente con quanto atteso (Costabile, 1996).

## *II. La fase della fiducia e della fedeltà comportamentale*

Le esperienze di acquisto e consumo caratterizzate da soddisfazione del cliente alimentano la tendenza al riacquisto (Boulding, Kalra, Staelin e Zeithmal, 1993); e se da tali riacquisti l' esperienza "soddisfacente" viene ulteriormente confermata, si raggiungono livelli sempre più consistenti di fiducia che determinano un' evoluzione della relazione verso la fedeltà (Bolton e Drew, 1991; Chang e Wildt, 1994; Morgan e Hunt, 1994).

Tale fenomeno è interpretabile alla luce del ruolo che la fiducia esercita sui costi di transazione, per cui al crescere dello stock di fiducia il riacquisto diventa economicamente più conveniente. Le principali categorie di economie generate dalla fiducia sono riconducibili:

- ai *costi cognitivi*, derivanti dallo sforzo di ricerca e di elaborazione delle informazioni, e che dovrebbero essere sostenuti nell' ipotesi in cui il cliente non riacquistasse dall' impresa che ha offerto performance soddisfacenti;
- ai *costi emotivi*, legati alla percezione di rischio e incertezza che, in genere, la fiducia contribuisce a ridurre in misura considerevole. Tali costi sono elevati in conseguenza della rilevanza delle differenti componenti di rischio percepito che di solito accompagnano l' acquisto e il consumo (fisico, economico, sociale, psicologico, funzionale o di performance - Kaplan, Szybillo, Jacoby, 1974), presenti in processi d' acquisto ad elevato coinvolgimento;

- ai *costi operativi*, e quindi al tempo, ai costi di trasferta, e a tutte le altre categorie di costo da sostenere per la valutazione delle alternative d' offerta;
- ai *costi strutturali* del cambiamento, derivanti da specificità tecnologiche del prodotto in uso (conversioni, interfacce, accessori, e così via) e da eventuali strategie di *lock in* (Shapiro e Varian, 1999) adottate dall' impresa fornitrice.

In una prospettiva dinamica, la seconda fase della relazione ha una durata che varia in funzione del livello di pressione competitiva, del livello di obsolescenza delle soluzioni tecnologiche percepito dal cliente, e del suo livello di coinvolgimento. Più specificatamente, è possibile sostenere che la durata della seconda fase è funzione del tipo di bene o servizio offerto dall' impresa e delle caratteristiche (individuali, sociali, economiche, e così via) degli attori coinvolti della relazione. Essa investe un intervallo definito da un numero variabile di riacquisti ( $t_1, \dots, t_m$ ), e si conclude, in genere, a seguito di nuovi stimoli, interni o esterni (situazionali e/o concorrenziali - Oliver, 1997) al sistema valutativo del cliente.

In sintesi, l' evoluzione della relazione consegue a una nuova valutazione comparativa fra il valore sperimentato nel periodo " $t_1, \dots, t_m$ " e il valore delle alternative disponibili sul mercato. La comparazione, che convenzionalmente potrà essere identificata con il momento " $t_m$ " nel ciclo di vita della relazione, avviene, in genere, sulla base del valore monadico, ossia del rapporto fra benefici e sacrifici attesi nell' offerta delle diverse alternative di mercato disponibili e il valore sperimentato in quella dell' impresa. La comparazione produce, in genere, tre risultati alternativi:

- il primo è quello che Hirschman (1970) avrebbe definito "*exit*". Il cliente, cioè, verificato che, nella sua prospettiva individualistica (monadica), vi sono imprese con offerte di valore significativamente superiori, decide di interrompere la relazione;
- il secondo risultato, invece, non conduce all' interruzione della relazione in considerazione delle "economie della fiducia" sperimentate nelle prime due fasi. In tal caso la relazione prosegue configurandosi come una forma di fedeltà "spuria" (Day, 1970) oppure coatta, ossia obbligata dalla convenienza rilevata su singole componenti di costo - ad esempio il costo di accesso all' offerta tipico della fedeltà comportamentale ad alcuni punti di vendita al dettaglio (Castaldo e Costabile, 1996).

- il terzo risultato, infine, è quello che rinforza la relazione. Qualora, infatti, la valutazione comparativa dimostri che il valore offerto dall' impresa è superiore rispetto a quello proposto dai concorrenti, il "conflitto" si risolve positivamente e la relazione si consolida entrando nella fase successiva.

### *III. La fase della fedeltà mentale*

La fedeltà mentale è una convinzione relativa alla capacità dell' impresa di mantenere nel tempo un differenziale di valore costante, o comunque positivo, rispetto ai concorrenti. Tale convinzione, rinforza anche il senso di autoefficacia del cliente, relativo cioè alla propria capacità di scelta dell' alternativa "migliore" fra quelle disponibili sul mercato.

Si tratta di uno stadio di sviluppo della relazione caratterizzato da elevata stabilità e disponibilità del cliente all' ampliamento della portata della relazione. E' con clienti in questo stadio della relazione, ad esempio, che sono frequenti - e di successo - le politiche di *cross selling* (Busacca e Costabile, 1995), oppure le estensioni della marca a nuove aree di *business*.

Il comportamento del cliente mentalmente fedele è caratterizzato da comportamenti di riacquisto durante i quali l' attenzione alle offerte dei concorrenti diventa selettiva o addirittura nulla. Ed è in questo stadio del ciclo vitale che si possono registrare i fenomeni di fedeltà proattiva descritti da Oliver (1997), ossia di riacquisto della marca o dell' offerta verso la quale si è fedeli anche in presenza di negative influenze situazionali o concorrenziali (evidenti vantaggi economici derivanti dal cambiamento di marca o di fornitore).

Tale forma di fedeltà, tuttavia, non rappresenta ancora lo stadio più evoluto che la relazione "cliente-impresa" può raggiungere. La fedeltà mentale, infatti, può assumere nel tempo due differenti configurazioni, in funzione del risultato di un ulteriore processo valutativo, generalmente, condotto dai clienti più anziani e longevi.

### *IV. La fase della lealtà*

Nella fase della fedeltà mentale i clienti hanno, in genere, maturato una lunga consuetudine di rapporti con l' impresa, acquisendo conoscenze approfondite sia

sull' offerta che sui suoi processi organizzativi. Ed è proprio tale maggiore conoscenza dell' impresa e della sua offerta, nonché la crescente autofiducia del cliente nelle proprie capacità di valutazione della categoria di prodotto, che provoca un' ulteriore momento di "verifica".

Anche in questo caso, si tratta di una comparazione del valore che tuttavia si concentra sull' equità della relazione con l' impresa. In sostanza, il cliente confronta il valore "storicamente" ottenuto dall' impresa con il valore che, nel corso del ciclo di vita della relazione, ritiene di aver generato per l' impresa.

La motivazione di tale comparazione può avere diverse origini, sovente riconducibili alla crescente capacità dello stesso cliente di valutare accuratamente l' offerta dell' impresa e lo sforzo economico e organizzativo della stessa profuso nella relazione. In altri casi, la verifica si fonda sulla maggiore consapevolezza del valore cumulato che i propri riacquisti hanno generato per l' impresa nel tempo. Si tratta, in sostanza, di ragioni che derivano dai fisiologici effetti di apprendimento sull' impresa, sui concorrenti e sul prodotto in senso lato.

Il valore diadico emerge dal confronto di benefici (B) e sacrifici (S) derivanti dall' acquisto e dal consumo dei beni e dei servizi dell' impresa (valore monadico), con costi (C) e ricavi (R) che si ritiene rappresentino la contropartita del valore per l' impresa:

$$B/S \cong R/C$$

Ovviamente, è solo nell' ipotesi che le ragioni di scambio vengano percepite come eque, e il valore offerto dall' impresa sia conseguentemente ritenuto corretto, che la relazione raggiunge la fase della *customer loyalty*.

La *customer loyalty* rappresenta lo stadio più alto del processo evolutivo di una relazione di mercato. Il cliente leale, infatti, è legato all' impresa da una relazione di fedeltà, mentale e comportamentale, ma anche da una convinzione di equità e correttezza che, sulla base del principio di reciprocità (Cialdini, 1984), conduce all' assunzione di atteggiamenti e comportamenti corretti e cooperativi.

Il cliente leale, quindi, è pronto a collaborare con l'impresa, sia sotto il profilo commerciale - ad esempio attivando spontaneamente flussi di passaparola positivi - sia sotto il profilo tecnico-produttivo - ad esempio fornendo suggerimenti su come migliorare i prodotti, i processi, e le forme di interazione cliente-impresa, fino a sperimentare nuove soluzioni organizzative o commerciali.

Il valore della *customer loyalty*, così definita, è riconducibile soprattutto alle opportunità di co-evoluzione della diade impresa-cliente (Busacca, 1997). I mercati ipercompetitivi - quali quello della telefonia mobile -, infatti, costringono le imprese alla continua innovazione, e di conseguenza all'incessante ricerca di nuove modalità per il miglioramento dei prodotti e dei processi aziendali. E' evidente, pertanto, che in tali mercati il valore economico e competitivo di relazioni con clienti leali sia particolarmente elevato, proprio in quanto è dallo stato delle suddette relazioni che dipende il valore delle opzioni di sviluppo delle imprese, e quindi il loro valore azionario.

### *La Customer Equity*

Adottando la prospettiva di Wayland e Cole (1997), con alcune rettifiche proposte da Costabile (2001), Grisaffe et al. (1998), si può sostenere che il valore del cliente – definito anche dalla locuzione *customer equity* – abbia tre componenti:

- il valore attuale;
- il valore delle opzioni di sviluppo della relazione;
- il valore delle risorse immateriali generate o co-generate dal cliente mediante interazione relazionale.

La formalizzazione analitica di tali componenti può essere espressa come segue:

$$CE = \left[ \sum_{t=1}^n (Q_t M_t) d^t - \sum_{t=1}^n (F_t d^t) \right] - A_1 + \left[ \sum_{z=1}^n (Q_z M_z) d^z * P_z - \sum_{z=1}^n (S_z + F_z) d^z \right] + K + R$$

dove:

- "Q" rappresenta la quantità acquisita dal cliente al tempo "t";

- “M” il valore dei margini;
- “ $\delta$ ” il tasso di attualizzazione dei flussi attesi e di capitalizzazione dei costi;
- “F” i costi di fidelizzazione;
- “A” i costi di acquisizione, che in linea di principio dovrebbero essere sostenuti solo al periodo “1”;
- “S” i costi di sviluppo della relazione, e in particolare “S<sub>z</sub>” quelli di esercizio delle opzioni dalle quali derivano gli incrementi dei margini “M<sub>z</sub>” e delle quantità “Q<sub>z</sub>”;
- “Z” un intervallo temporale  $< \tau$ , in quanto il numero di periodi (anni per convenzione) che compongono l’orizzonte temporale di manifestazione dei flussi incrementali parte dal momento dell’esercizio dell’opzione (o delle opzioni);
- “P<sub>z</sub>” la probabilità necessaria a stimare la quota parte di clienti che aderirà all’opzione esercitata dall’impresa al tempo “Z”, ossia che accetterà di ampliare la portata della relazione, mediante maggiori volumi acquisiti, *cross buying*, acquisti di prodotti di più alta gamma (*trading up*) e acquisti di nuovi prodotti dell’impresa, disponibili in seguito all’ingresso della stessa in nuove aree di *business*;
- “K” il valore della conoscenza generabile dall’interazione con il cliente;
- “R” il valore della reputazione che il medesimo cliente è in grado di diffondere a beneficio dell’impresa.

La prima parte dell’equazione, al netto dei costi di acquisizione, rappresenta il valore attuale; la seconda parte il valore delle opzioni di sviluppo della relazione, e quindi il valore-potenzialità; la terza il valore delle risorse immateriali generabili a partire da relazioni leali e collaborative.

## CAPITOLO 4–

### **MODELLI FORMATIVI (DESCRITTIVI)**

#### 4.1 INTRODUZIONE

Gli studi di *customer satisfaction*, concernenti, cioè, il soddisfacimento del cliente o consumatore o utilizzatore di un prodotto/servizio, si sono inizialmente sviluppati, sia dal punto di vista pratico che teorico, nell'ambito del Marketing. La conoscenza e l'interpretazione sistematica di come i consumatori percepiscono e valutano la qualità dei prodotti/servizi si è subito delineata essenziale per l'orientamento della gestione e delle strategie d'impresa. Da ciò trae motivazione l'attuale sviluppo degli studi rivolti al tema del "soddisfacimento del cliente", che segnala un allargamento degli ambiti disciplinari coinvolti, che ora includono, oltre al Marketing, l'Economia aziendale, la Sociologia, l'Economia, e, più decisamente, la Statistica (Vedaldi, 1997; Frosini, 1996), ma anche col nostro scritto la Teoria delle Decisioni. Sembra, pertanto, giunto il momento in cui appare utile fare il punto circa i molteplici contributi e le proposte disponibili sul tema in esame, con riferimento ai diversi settori disciplinari. Il presente contributo si muove in questa direzione con riferimento a quanto la metodologia statistica ha dato, e potrà ancora dare, allo sviluppo ed alla chiarificazione delle problematiche collegate alla "customer satisfaction". Il presente capitolo, e i due successivi, si riferiscono alle principali metodologie statistiche e di data-mining. La loro conoscenza rappresenta idoneo elemento di confronto con la metodologia da noi proposta relativa all'ambito della Teoria delle Decisioni.

#### 4.2 IL COSTRUTTO CONCETTUALE ED IL PROCESSO DI MISURAZIONE

Conviene porsi il quesito preliminare: cosa vuole indicare l'espressione "customer satisfaction"? In sostanza si tratta di una *nozione* o *concetto relazionale* che assume pieno significato sperimentale, vale a dire statistico, solo in base ai diversi aspetti del prodotto/servizio cui è associato attraverso un cosiddetto costrutto concettuale ed un corrispondente procedimento di misurazione. Si tratta, quindi, di una funzione di un insieme multidimensionale di attributi e variabili – vettore – che stabiliscono la struttura sottostante alla sintesi attuata da detta funzione.

Quando si vuole associare ad un concetto – inteso come tratto unificante di certe manifestazioni fenomeniche, nel caso in studio collegate ad una popolazione "sufficientemente omogenea" di individui – un *costrutto concettuale* ed un corrispondente processo di misurazione, si può adottare un procedimento riassunto nei punti seguenti, Bollen (1989), p. 179 e seg..

1. *Ricerca di un'accurata definizione verbale del concetto*. Ad esempio la definizione di *customer satisfaction* data da Hunt nel 1977 si riferisce ad un bene o servizio dopo l'acquisto ed afferma "*The evaluation rendered that the experience was at least as good as it was supposed to be*", si veda Evrard (1993).

2. *Individuazione degli aspetti essenziali, dimensioni, delle manifestazioni fenomeniche collegati al concetto di interesse*. Essi stessi sono tipicamente di natura concettuale. Si giunge ad una "rete" di concetti in cui è "incapsulato" quello in studio e, globalmente, al cosiddetto *costrutto concettuale*. Ciò può farsi, in primo luogo, attraverso l'analisi semantica della definizione del concetto di interesse. Ad esempio la precedente definizione di Hunt include un confronto cui corrispondono due dimensioni collegate rispettivamente alla percezione del prodotto/servizio acquistato ed alla "base di riferimento iniziale", anteriore all'acquisto. I successivi approfondimenti hanno portato al corrispondente costrutto della "discrepanza" con insiemi, tipicamente speculari, di dimensioni attinenti a: a) *prestazioni percepite*, b) *prestazioni attese*, collegate in modo specifico al prodotto/servizio in esame e tali che dal loro confronto originano le: c) *discrepanze*, che determinano la *customer satisfaction*. L'essenzialità della dimensione "discrepanza" non è condivisa, come si dirà, da tutti gli autori, si veda Cronin et al.



(1992). Altre indagini hanno portato ad aggiungere le dimensioni: a) “costo” percepito, in relazione alle prestazioni percepite ed attese (valore), b) motivi di lagnanze, c) fedeltà del cliente. La precedente rete di dimensioni ed il corrispondente costruito, sono quelli alla base dell’American Customer Satisfaction Index (ACSI), si veda ad esempio Anderson e Fornell (2000). Non molto diverso si presenta il costruito utilizzato nella definizione dell’European Customer Satisfaction Index (ECSI). Ma di questi abbiamo già discusso nel capitolo secondo.

3. *Definizione di un procedimento di misurazione a partire dalla scelta, per ogni dimensione, di uno o più corrispondenti indicatori o variabili osservabili o manifeste.*

Quando si considerano più indicatori  $X_i$  per una stessa dimensione si è soliti associare a quest’ultima un’unica “variabile latente non osservabile”  $\xi$  secondo un modello, cosiddetto di misurazione. In generale può adottarsi un modello di misurazione complessivo, attinente o a tutte le variabili latenti associate alle dimensioni del costruito, o a un loro sottoinsieme. Precisamente si può assumere il modello:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \mathbf{M} \\ \mathbf{X}_p \end{bmatrix} = \begin{bmatrix} \lambda_1 & \mathbf{0} & \Lambda & \mathbf{0} \\ \mathbf{0} & \lambda_2 & \Lambda & \mathbf{0} \\ & & \Lambda & \\ \mathbf{0} & \mathbf{0} & \Lambda & \lambda_p \end{bmatrix} \cdot \begin{bmatrix} \xi_1 \\ \xi_2 \\ \mathbf{M} \\ \xi_p \end{bmatrix} + \delta = \Lambda_x \xi + \delta \quad (1)$$

dove  $\mathbf{X}_j$  sono vettori colonna,  $q_j \times 1$ , riassuntivi delle variabili manifeste  $X_{ji}$ ,  $i = 1, 2, \dots, q_j$ , che corrispondono alle componenti del vettore casuale  $\xi$  delle variabili latenti  $\xi_j$  associate a  $p$  dimensioni del costruito,  $j=1, 2, \dots, p$ , e ritenute non correlate con le componenti aleatorie d’errore  $\delta_{ji}$ , riassunte dal vettore  $\delta$ , di media nulla, varianza finita fra loro non correlate;  $\lambda_j$  sono vettori colonna riassuntivi di incogniti parametri di scala;  $\mathbf{0}$  dei vettori colonna di componenti nulle,  $\Lambda_x$  è la matrice corrispondente al modello parametrico anzidetto. L’analisi, su base campionaria, del modello (1) è tipicamente condotta in due passi. a) Per ogni insieme  $X_{ji}$ ,  $i=1, 2, \dots, q_j$ , si verifica che “misurino” una stessa variabile latente  $\xi_j$ . Al riguardo è spesso utilizzato il coefficiente  $\alpha$  di Cronbach definito, per la generica variabile latente  $\xi_j$ , come il quadrato del coefficiente di correlazione fra la stessa ed  $H_j = \Sigma_i X_{ji}$ ,  $0 \leq \alpha_j \leq 1$ , si veda Bollen (1989). Sono

considerati coerenti all'ipotesi, in senso descrittivo, valori campionari  $\alpha_j \geq 0.7$ . b)  
Successivo studio del modello (1): si devono, precisamente, determinare i parametri di  
una struttura di covarianza del tipo:

$$\Sigma_x = E(\mathbf{X} \mathbf{X}') = \Lambda_x \Phi \Lambda_x' + \Psi$$

dove  $\Phi = E(\xi\xi')$ , e  $\Psi = E(\delta\delta')$  è matrice diagonale,  $E(\cdot)$  indica il valore atteso, l'apice  
l'operazione di trasposizione. Ciò è tipicamente fatto, su base campionaria, mediante le  
tecniche di analisi dei fattori che consentono di confermare o eventualmente modificare  
il numero ed il significato delle dimensioni e/o dei corrispondenti modelli di  
misurazione.

#### 4.3 IL PROBLEMA DELLE SCALE ORDINALI

##### 1. *Qualche riflessione preliminare.*

Per rendere più chiaro quanto segue si consideri, ad esempio, uno “sportello” di una determinata sede bancaria e l’insieme di  $N$  successive operazioni correnti: versamenti, prelievi, bonifici, ecc. espletate in un certo periodo di tempo. Si suppone di chiedere ad un “soggetto”, dopo che abbia fruito del servizio, di qualificarne l’attuazione scegliendo una voce della scala verbale, semantico-differenziale, stabilita dalle seguenti affermazioni: *estremamente inefficiente, inefficiente, normale, efficiente, estremamente efficiente*. Secondo la teoria assiomatica, attinente alla costruzione di scale di misurazione, una scala ordinale nel campo reale, cioè numerica, può ottenersi nelle condizioni e secondo le modalità seguenti (si veda ad esempio Thomas (1985)). Sia  $A$  un insieme finito di  $N$  elementi, tipicamente non numerici, per il quale è assegnata una relazione binaria  $\varphi$  fra le coppie di elementi (se  $a_1, a_2 \in A$ ,  $a_1 \varphi a_2$  si legge “ $a_1$  è in relazione con  $a_2$ ”), quindi, definita nell’insieme  $A \times A$ ;  $\varphi$  soddisfa, inoltre, le condizioni seguenti:  $\forall a, b, c \in A$ : 1) riflessività (identificabilità):  $a \varphi a$ ; 2) concatenamento:  $a \varphi b$  ovvero  $b \varphi a$ ; 3) transitività: se  $a \varphi b$  e  $b \varphi c$  allora  $a \varphi c$ . Sotto le precedenti condizioni si dimostra il seguente teorema, che stabilisce la rappresentabilità degli elementi di  $A$  e delle rispettive relazioni nel campo reale: “Esiste una funzione  $f: A \rightarrow \mathbb{R}$  tale che:  $a \varphi b$  se e solo se  $f(a) \geq f(b)$ ; la funzione  $f$  non è unica nel senso che la precedente relazione è garantita da una qualsiasi altra funzione reale  $h[f(x)]$ ,  $x \in \mathbb{R}$ , con  $h(\cdot)$  funzione strettamente monotona crescente”.

Si è in tale modo definito un procedimento di misurazione e si dice che gli elementi di  $A$ , con la struttura di relazione che li caratterizza, sono rappresentati attraverso le misurazioni  $f(a)$ ,  $a \in A$ , su una scala ordinale. Mediante la scala ordinale definita come si è detto possono classificarsi gli  $N$  elementi “materiali” dell’insieme  $A$ . Nell’esempio precedente  $A$  è costituito dall’insieme di  $N$  operazioni bancarie considerato: il precedente risultato può ritenersi applicabile qualora tutte le valutazioni siano eseguite da uno stesso soggetto. Se, però, consideriamo le valutazioni di  $M$  diversi soggetti, e questi, come è ragionevole, non possono ritenersi “identici”, diventa in generale impossibile ammettere l’esistenza di un sistema relazionale empirico fra gli elementi

materiali di  $A$  con le proprietà sopra specificate, che sono necessarie per giungere ad una scala ordinale. Quanto precede sembra giustificare la notevole cautela nel fare riferimento a più soggetti suggerita nell'impiego della tecnica cosiddetta di *conjoint analysis*, che si basa sulle valutazioni di elementi per costruzione diversi. È facile convincersi che non si hanno, invece, difficoltà quando più soggetti valutano una *stessa unità* sperimentale, o più unità ritenute identiche (è il caso della produzione di serie).

## 2. Trasformazioni per ottenere scale metriche.

Col ricorso alla nozione di variabile latente e ad un corrispondente modello probabilistico è possibile trasformare – sulla base di convenienti ipotesi – i valori di una scala ordinale in altri espressi su una scala metrica, tipicamente ad intervalli. Seguendo Zanella et al. (2000) si supponga di poter idealizzare la legge secondo cui si manifesta il risultato osservato da un'unità sperimentale – costituita, ad esempio, dal rispondente ad un questionario – mediante una variabile aleatoria categorica multidimensionale  $\mathbf{X} = (X_1, X_2, \dots, X_K)'$  con componenti che assumono, per semplicità, uno stesso numero  $I^* = I + 1$  di modalità qualitative ordinate  $x_{ki}$ ,  $i = 1, 2, \dots, I^*$ ,  $k = 1, 2, \dots, K$ , che possono sempre rappresentarsi su una scala ordinale mediante i valori  $x_{ki} = i$ ,  $i \in \{1, 2, \dots, I^*\} = S$ . Sia  $P(X_{ki} = i) = p_{ki}$  la probabilità, marginale,  $\sum_i p_{ki} = 1, \bullet \forall k$ , che la componente  $k$ -ma di  $\mathbf{X}$  assuma il valore  $i$  e si indichi con:

$$F_k(i) = \sum_{\{i \in S, j \leq i\}} p_{kj}$$

la probabilità che si verifichi una o l'altra delle prime  $i$  modalità. Si consideri l'approccio psicometrico di Thurstone, proposto originariamente nell'ambito dello studio delle valutazioni soggettive conseguenti agli stimoli provenienti da un oggetto e con riferimento alla distribuzione normale. Si supponga precisamente che alla variabile categorica  $\mathbf{X}$  corrisponda ordinatamente una variabile casuale multidimensionale di tipo continuo  $\mathbf{Z} = (Z_1, Z_2, \dots, Z_K)'$  non direttamente osservabile o latente le cui componenti, quindi, assumono per ipotesi valori su scale metriche, tipicamente ad intervalli;  $\mathbf{Z}$  sia collegata con la variabile categorica manifesta  $\mathbf{X}$  tramite le condizioni:

$$\Psi_k \left( \frac{\xi_{ki} - \mu_k}{\sigma_k}; \alpha_k \right) = F_k(i), \quad k = 1, 2, \dots, K, \quad (2)$$

dove  $\Psi_k(\cdot)$  indica la funzione di ripartizione marginale della generica componente  $Z_k$  di  $\mathbf{Z}$ ,  $\mu_k, \sigma_k > 0$  sono, rispettivamente, i relativi parametri di posizione e di scala,  $\alpha_k$  è un vettore riassuntivo di altri parametri. Nel caso di distribuzioni normali  $\mu_k, \sigma_k$  divengono rispettivamente la media aritmetica e lo scarto quadratico medio e non esiste  $\alpha_k$ .

La critica che può rivolgersi all'utilizzazione di una tecnica di cambiamento di scala del tipo espresso dalla (2), o alla sua estensione bivariata, pure utilizzata, si veda Bollen (1989), è che non è agevole la verifica delle assunzioni su cui si fonda. Vi sono situazioni, però, nelle quali una verifica diretta della validità delle trasformazioni (2) risulta possibile. Si supponga, infatti, che alle modalità ordinate  $x_{ki} \in \{1, 2, \dots, I\}$  delle variabili categoriche  $X_k$ ,  $k = 1, 2, \dots, K$ , per ciascuna delle variabili casuali latenti  $Z_k$ , corrisponda uno *stesso* insieme ordinato di valori  $z_{ki} = \xi_i, \forall k, \xi_i < \xi_{i+1}, i = 1, 2, \dots, I - 1$ . Si abbia, quindi, in corrispondenza alla (2), dove ora  $\xi_{ki} = \xi_i, \forall k$ :

$$(\xi_i - \mu_k) / \sigma_k = \Psi_k^{-1}[F_k(i)] = \zeta_{ki}(\alpha_k)$$

$i = 1, 2, \dots, I, k = 1, 2, \dots, K$ . Si tratta di  $K$  relazioni lineari simultanee delle quali è possibile la verifica su base campionaria, si veda Zanella et al. (2000). È questo un tema attuale di ricerca.

Con la metodologia che presenteremo nei prossimi capitoli cercheremo di affrontare, con approccio diverso, il problema delle scale ordinali.

Parleremo più avanti di altri approcci (il modello di Rasch).

#### 4.4 LE TECNICHE “*ATTRIBUTE BASED*”: DESCRIZIONE METODOLOGICA DEL “MODELLO DEI GAP” DEL PROF. PARASURAMAN, VANTAGGI E SVANTAGGI

Le ricerche di marketing per la C.S.M. (Customer Satisfaction Measurement), consistenti nella rilevazione della soddisfazione e misurazione delle performance d'alcune aree della qualità, si differenziano dalle normali ricerche di marketing per due aspetti fondamentali:

1. Ha una direzione d' indagine interna ed esterna all' organizzazione, perché il cliente è sollecitato ad esprimersi non solo sul risultato del processo, ma anche sull' efficienza ed efficacia dei suoi elementi.
2. E' un insieme di parti integrate ed interagenti, perché l' informazione di CS nasce dall' interpretazione di strumenti di misurazione diretta (indagini qualitative, quantitative, o integrate), ed indiretta (*customer retention rate*, tasso di riacquisto, volume delle vendite, ecc.), perciò la ricerca di marketing è strumento per la misurazione della CS.

Possiamo distinguere due sistemi di misurazione:

- **Misurazione indiretta** attraverso l' impiego di "filtri", o indicatori che consentono di correlare i diversi risultati operativi al livello di soddisfazione.
- **Misurazione diretta** fa riferimento a quelle tecniche di misurazione che prevedono un esplicito coinvolgimento del cliente, al quale è chiesto di esprimere un giudizio puntuale sul livello di soddisfazione per il servizio erogatogli. La sua caratteristica è la volontà dell' impresa di ascoltare direttamente "la viva voce del cliente".

Presso talune aziende che esercitano attività particolari, la verifica della soddisfazione del cliente è perfettamente confusa con la sua operatività (es. Grimaldi crociere). Per altre del tipo job-shop, con un contatto continuo col cliente, la CS è stata da sempre applicata.

La CS management costituisce una tipologia di gestione che per essere in sintonia col mercato implica un continuo processo di ristrutturazione dell' organizzazione (il c.d.

*business process reengineering*), compito del top management, è di gestire il cambiamento in modo graduale e senza strappi.

Per la misurazione della CS, vi sono due grandi famiglie di modelli:

- La prima nasce sulla scia degli studi del prof. Parasuraman e si fonda sulla logica dei gap (scostamenti tra dimensioni critiche per la qualità del servizio);
- La seconda più eterogenea si focalizza sull' indice cosiddetto dell' overall satisfaction e si concentra una volta sul processo, un' altra sulla qualità percepita, e così via.

Il primo modello esplica l' analisi della CS verificando la consonanza esistente tra:

1. Gli obiettivi di CS pianificati dal management aziendale (*soddisfazione pianificata*);
2. Il livello di soddisfazione desiderato dal cliente (*soddisfazione desiderata*);
3. Gli obiettivi di CS percepiti dal personale aziendale (*soddisfazione recepita*);
4. Il livello di soddisfazione percepito dal cliente (*soddisfazione percepita*);
5. Il livello di soddisfazione effettivamente offerto dall'impresa (*soddisfazione offerta*).

Eventuali scostamenti tra gli elementi indicati "**gap**", costituiscono altrettante possibili strategie d'accrescimento del livello di CS, identificando le cause di ciascuno scostamento e le azioni idonee a rimuoverli. Nell' intento di identificare e rimuovere le possibili cause di insoddisfazione del cliente – che nel tempo indebolisce le fonti dei vantaggi strategici conquistati dall' impresa- si può far ricorso a un modello utile a verificare la conoscenza esistente tra la qualità pianificata dal top management e quella desiderata dal consumatore, gli obiettivi di qualità percepiti dal personale aziendale, la qualità offerta al mercato e quella effettivamente percepita dall' acquirente. Gli eventuali scostamenti rilevabili tra questi elementi, possono essere ricondotti a otto tipologie fondamentali:

- *gap di sintonia*;
- *gap di valore*;
- *gap di percezione*;
- *gap di allineamento*;

- *gap di progettazione;*
- *gap di realizzazione;*
- *gap di coinvolgimento;*
- *gap di consonanza.*

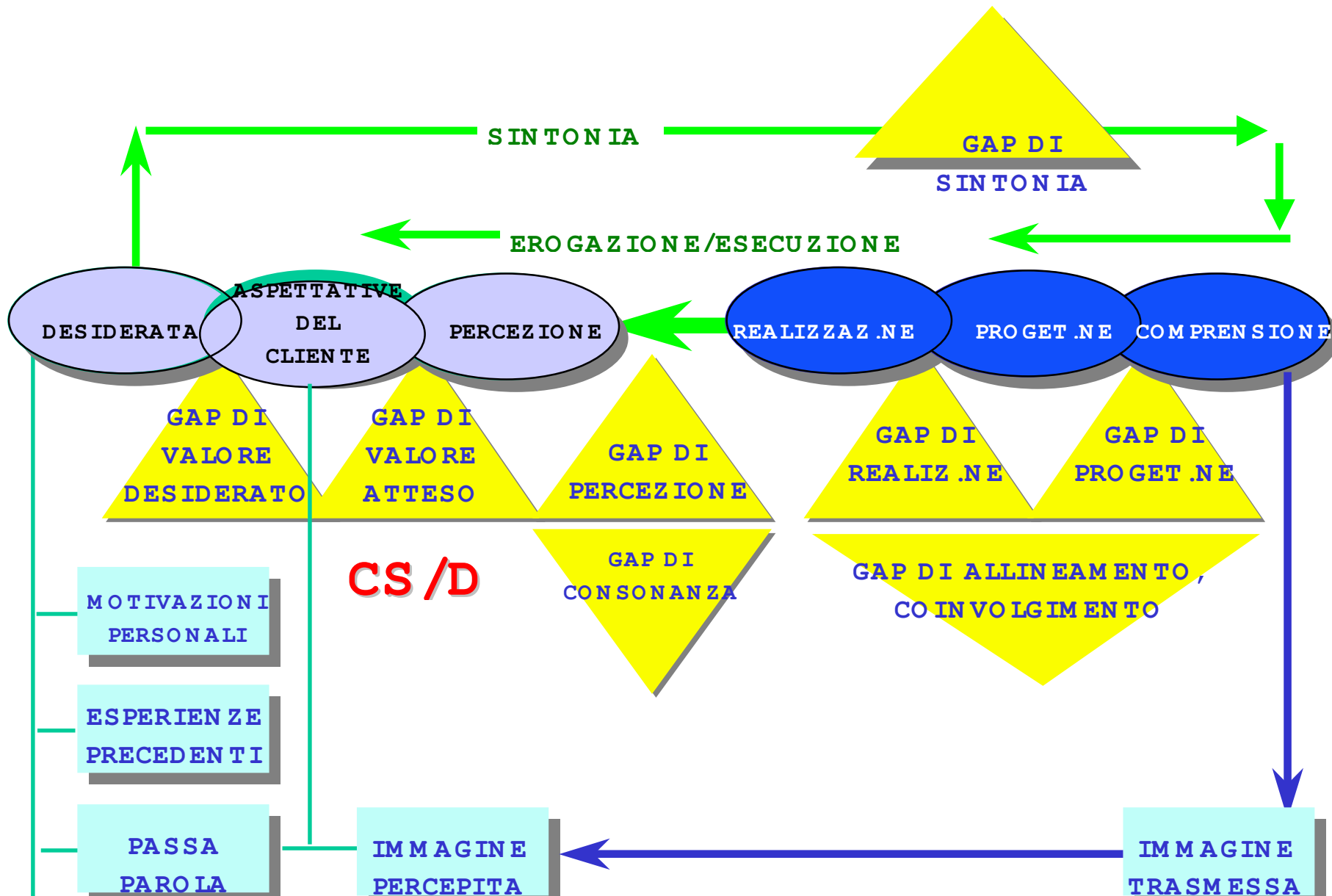
Il ***gap di sintonia*** si manifesta quando l' impresa definisce le caratteristiche d' uso e di immagine della propria offerta, prescindendo dalla ricerca di una sintonia tra la soddisfazione pianificata dal manager e quella desiderata dai consumatori. Ciò, inevitabilmente, si traduce nella progressiva erosione delle potenzialità competitive. Per rilevare un simile gap è indispensabile verificare l' allineamento della propria offerta rispetto ai desideri e alle aspettative dei consumatori, attraverso ricerche di marketing qualitative e quantitative.

Il gap di sintonia può scaturire da tre specifiche cause:

- i clienti esprimono esigenze altamente differenziate, dando vita ad una domanda complessa ed articolata, che impone all' impresa di sviluppare nuove capacità per segmentare più specificatamente e creativamente il mercato;
- i consumatori incontrano difficoltà ad esprimere e comunicare i propri bisogni e desideri, per cui è necessario dedicare notevoli risorse all' osservazione e all' ascolto dei clienti, coinvolgendoli nel processo di sviluppo dell' offerta e cooperando con essi alla ricerca di soluzioni tecnologiche reciprocamente vantaggiose;
- i clienti manifestano esigenze mutevoli e aspettative crescenti, alle quali l' azienda può positivamente rispondere solo accrescendo la propria flessibilità, predisponendo piani di azione "proattivi" e organizzando meccanismi operativi idonei a incrementare la velocità di risposta e il patrimonio di risorse naturali e immateriali a sua disposizione.



## IL PROCESSO DI CUSTOMER SATISFACTION



Il **gap di valore** si verifica nel caso in cui l' offerta dell' impresa genera al cliente una soddisfazione inferiore a quella desiderata. Si tratta di un gap difficile da rilevare, dal momento che il valore percepito è funzione di molteplici variabili, alle quali ciascun individuo attribuisce una diversa importanza. Esso scaturisce, in sintesi, da una ponderazione dell' inutilità derivante dalle caratteristiche d' uso<sup>ed</sup> immagine dei prodotti con i rispettivi prezzi e con eventuali costi non monetari connessi alle attività di acquisto e consumo.

Il gap di valore può derivare da:

- un' errata valutazione delle correlazioni tra le caratteristiche del prodotto e le esigenze dei consumatori;
- da una definizione non corretta della posizione occupata dal profilo di offerta ideale nello spazio percepito dei consumatori;
- dalla non corretta interpretazione dei meccanismi sottostanti alla percezione delle caratteristiche qualitative e di immagine dell' offerta.

Alla luce di queste considerazioni, il gap di valore può essere colmato ricorrendo a due fondamentali strategie, graduate secondo una scala crescente di complessità. La prima è costituita dalla modificazione della posizione occupata dal prodotto nello spazio percettivo del cliente, mediante interventi sulle caratteristiche dell' offerta e/o sul sistema di convinzioni e atteggiamenti; la seconda, dallo spostamento della localizzazione del profilo di offerta ideale, attraverso azioni sul sistema di preferenza e sui parametri di giudizio individuali.

Il **gap di percezione** nasce da una divergenza fra la soddisfazione teoricamente traibile dall' offerta aziendale e la soddisfazione percepita dai consumatori. Il gap di percezione è riconducibile a tre distinte cause esemplificative:

- incapacità dell' impresa di comunicare adeguatamente le caratteristiche qualitative della propria offerta;
- mancanza di sintonia tra le strutture cognitive dei consumatori e le comunicazioni da essi ricevute;

- mancanza di corrispondenza tra le dimensioni utilizzate dai consumatori per inferire il livello qualitativo dei prodotti e i parametri adottati dall' impresa per definire la qualità della propria offerta.

Per annullare il gap di percezione è quindi necessario modificare la posizione occupata da una determinata alternativa di offerta nello spazio percettivo dei consumatori, attraverso il cambiamento delle convinzioni concernenti i rapporti attributo-beneficio, prodotto-attributo e prodotto-beneficio, che, a propria volta, richiede la preliminare definizione

dell' organizzazione interna delle strutture cognitive.

Il *gap di allineamento* si manifesta qualora la soddisfazione pianificata dal vertice aziendale diverga dagli obiettivi di customer service recepiti dall' organizzazione. Tale scostamento può verificarsi quando il management non si preoccupa di suscitare una partecipazione diretta di tutte le componenti dell' impresa, oppure di generare una visione unitaria sulle conseguenze derivanti dall' orientamento aziendale alla soddisfazione del cliente. Alcune imprese si sono sforzate di definire i principi fondamentali che stanno alla base dell' orientamento della customer service elaborando una visione atta a descrivere e far condividere i valori ad essa sottostanti. Altre realtà aziendali hanno, invece, perfezionato modelli e meccanismi organizzativi per favorire, da un lato, la condivisione dei nuovi valori di customer service e, dall' altro, la definizione delle responsabilità secondo principi idonei a facilitare il coordinamento e l' integrazione dei compiti. La generazione di una visione unitaria necessita, comunque, di grande innovazione anche nei processi di comunicazione dell' impresa: maggiore è lo scambio di informazioni, più numerose sono le opportunità di miglioramento della customer service.

I *gap di progettazione* e di realizzazione si manifestano quando gli obiettivi di customer service condivisi dall' organizzazione non si riflettono nella soddisfazione effettivamente offerta al mercato (gap di progettazione), oppure qualora l' impresa sia incapace di comporre un profilo di offerta perfettamente rispondente alle specifiche del progetto con una affidabilità tecnico funzionale senza difetti (gap di realizzazione). Per ridimensionare questi scostamenti occorre sviluppare maggiore integrazione esterna,

allo scopo di generare concetti di prodotto che anticipino i bisogni e i desideri dei futuri clienti, e far penetrare tali concetti nei progetti di dettaglio. Il processo di sviluppo dell' offerta non dipende tuttavia soltanto dal flusso di informazioni acquisite dall' impresa, ma anche dalla circolazione di tali informazioni all' interno dell' impresa stessa. Ciò implica che l' integrazione esterna si estenda in profondità a ogni livello dell' organizzazione, traducendosi in una effettiva collaborazione tra le aree e le funzioni che svolgono un ruolo determinante nello sviluppo di un nuovo prodotto: marketing, ricerca & sviluppo e produzione. Ciò che distingue la capacità di introdurre prodotti di successo è infatti la coerenza tra la struttura formale e l' organizzazione informale caratterizzante l' attività critica dello sviluppo.

Il **gap di coinvolgimento** è evidenziato dallo spostamento fra gli obiettivi di customer service pianificati dal management e la soddisfazione traibile dall' offerta dall' impresa. La sua esistenza attesta una condivisione, all' interno dell' impresa, dei valori e dei principi sottostanti ad un effettivo orientamento alla customer service. In termini generali, il gap di coinvolgimento può essere ridotto intervenendo simultaneamente su quattro fattori reciprocamente interconnessi:

- la partecipazione diretta ai cambiamenti (stimolando la diagnosi congiunta sui problemi di customer service);
- il coordinamento (per integrare funzioni e attività nella realizzazione e nel perseguimento della visione della customer service);
- gli incentivi e i meccanismi di riconoscimento (finalizzati all' incremento dell' assegnazione e assunzione di responsabilità del personale).

Le esigenze dei clienti vengono infatti adeguatamente soddisfatte se tutti i componenti dell' impresa, pur appartenendo a funzioni diverse, si sforzano di partecipare all' avanzamento dei processi organizzativi secondo una logica relazionale, anziché preoccuparsi solo del successo dell' attività propria competenza.

Il **gap di consonanza** si manifesta qualora gli obiettivi di customer service recepiti e dichiarati dal personale divergano dalla soddisfazione percepita dai clienti. L' esistenza di tale scostamento costituisce, nel lungo periodo, una sostanziale minaccia alla

sopravvivenza stessa dell' impresa; il suo perdurare si traduce, infatti, in una crescente distanza tra le aspettative della domanda e il risultato dell' esperienza di acquisto e di consumo, minando la credibilità e l' immagine aziendale. L' individuazione del gap di consonanza implica un approccio analitico articolato in cinque fasi fondamentali:

- disaggregazione dell' offerta aziendale in una serie di attributi rilevanti suscettibili di orientare il processo di formazione delle preferenze della clientela;
- specificazione delle aree di responsabilità a livello funzionale con riferimento alle componenti del profilo di offerta precedentemente individuate;
- definizione degli obiettivi di customer service che ciascuna funzione si propone di conseguire in relazione alle proprie possibilità operative;
- determinazione degli effettivi criteri di scelta utilizzati dai consumatori, della loro importanza relativa e della correlazione esistente tra tali criteri e gli attributi di offerta aziendale;
- quantificazione dei giudizi formulati dalla clientela con riferimento ai diversi attributi e verifica degli scostamenti esistenti tra il profilo di offerta percepito dalla domanda e gli standard qualitativi perseguiti dall' impresa.

L' eliminazione del gap di consonanza è subordinata all' esplicitazione delle cause ad esso sottostanti. In primo luogo, queste appaiono riconducibili all' errata identificazione delle esigenze maturate dai consumatori (gap di sintonia) o all' incapacità del management di attivare efficaci processi di comunicazione degli obiettivi di customer service perseguiti (gap di allineamento). In secondo luogo, il gap di consonanza può derivare dal fallimento di strategie finalizzate all' incremento del patrimonio di risorse immateriali a disposizione dell' impresa. L' adozione di tali strategie può, infatti, tradursi nella fissazione di obiettivi di customer service che eccedono le attuali capacità, competenze e risorse a disposizione dell' azienda. Ciò, necessariamente, determina uno scostamento tra le finalità recepite dall' organizzazione e la soddisfazione dai clienti, scostamento comunque destinato ad annullarsi progressivamente al crescere del patrimonio di risorse immateriali.

**Le principali metodologie** adottate per la misurazione della soddisfazione del consumatore si possono distinguere in due categorie fondamentali:

- ⇒ *Attribute based*, presuppone che il beneficio fruibile dal prodotto o servizio sia scomponibile in un insieme (bundle) di caratteristiche (attributi); al fine di rilevare tali giudizi si ricorre soprattutto alle scale di valutazione, che possono essere di diversi tipi:
- *Scala unipolare*, finalizzata a definire la quantità di un certo attributo, che è offerto da una determinata azienda;
- *Differenziale semantico*, scala di tipo bipolare, mediante la cui è chiesto al consumatore di esprimere un giudizio tra due opposte possibilità con cui può manifestarsi un certo attributo;
- *Scala Likert*, regola indiretta, basata sulla richiesta di un giudizio d'accordo/ disaccordo su una serie d'informazioni predeterminate dal ricercatore.
- *Stapel scale*, si propone di definire il profilo di un prodotto, in base alla valutazione dei consumatori espresse in una scala numerata con un estremo negativo e l'altro positivo.

**Limiti delle analisi basate sulle scale:**

- Dato che, gli attributi che il cliente valuta, sono predeterminati dal ricercatore, c'è il rischio di trascurare attributi rilevanti; si può minimizzare tale rischio attraverso un'indagine esplorativa di tipo qualitativo finalizzata alla definizione degli attributi mediante focus group;
- Non è facile scomporre l'offerta in un insieme d'attributi, soprattutto nel caso dei servizi (caratterizzati dall'intangibilità).

*Non attribute based*, si basano su valutazioni sintetiche.

#### 4.5 MODELLI FORMATIVI O COMPOSITIVI

Caratteristica di questi modelli è la dominanza del carattere di “attributo multidimensionale” della “customer satisfaction”, nel quale ciascuna componente corrisponde ad una dimensione del costrutto concettuale, cioè, ad un aspetto di un prodotto/servizio ritenuto essenziale nella determinazione della “customer satisfaction”. Quest’ultima è resa variabile manifesta attribuendole un valore ottenuto *componendo*, ad esempio in modo additivo, le valutazioni osservabili in corrispondenza alle diverse dimensioni: la sintesi delle valutazioni sui singoli attributi “marginali” ha carattere definitorio e, quindi, convenzionale. Il modello in oggetto è stato sviluppato a partire dal lavoro fondamentale di Parasuraman et al. (1988) che sta alla base del notissimo modello SERVQUAL, relativo alla qualità complessiva dei servizi. In questo e nei successivi lavori, è stato accuratamente applicato in 8 casi (soprattutto Banche e Assicurazioni). Il metodo ha consentito di confermare, attraverso le corrispondenti variabili manifeste, il cui numero è indicato fra parentesi, 5 dimensioni: aspetto tangibile (4), affidabilità (5), capacità di risposta (4), capacità di rassicurazione (4), empatia (5). Ogni variabile manifesta è valutata dal rispondente al questionario su una scala di punteggio intero, da 1 a 7. Questa è ancorata, per le aspettative nei confronti dell’aspetto cui si riferisce, alle affermazioni estreme “del tutto non essenziale”, “del tutto essenziale”; per le percezioni “dissentito fortemente”, “concordo fortemente” nei confronti di una valutazione positiva di tale aspetto per il caso in studio. I due tipi di apprezzamento sono presentati in modo speculare ma distinto nel questionario. E’ anche richiesto al rispondente di assegnare ad ogni dimensione un “peso”, che ne esprime l’importanza, col vincolo che il totale dei pesi abbia valore 100.

Siano  $y_{jih}$ ,  $z_{jih}$ , i punteggi ottenuti dal rispondente  $h$ -mo,  $h = 1, 2, \dots, N$ , per la variabile manifesta  $i$ ,  $i = 1, 2, \dots, q_j$ , della dimensione  $j$ -ma,  $j = 1, 2, \dots, 5$ , e relativi rispettivamente alle percezioni attinenti al servizio reale ( $y_{jih}$ ) e alle aspettative ( $z_{jih}$ ). L’indicatore complessivo, che esprime il *paradigma della “discrepanza o gap”* ed è alla base del SERVQUAL, viene definito come:

$$CSI = \sum_{h=1}^N \left\{ \sum_{j=1}^5 w_{jh} \left[ \sum_{i=1}^{q_j} (y_{jih} - z_{jih}) / q_j \right] \right\} / N \quad (3)$$

dove  $w_{jh}$ ,  $\sum_j w_{jh} = 1$ , indicano, divisi per 100, i valori dei pesi.

Una critica profonda all'approccio di Parasuraman et al. è stata quella mossa da Cronin et al. (1992), che hanno messo in dubbio, su base sperimentale, che l'indicatore di Parasuraman et al. coincida anche con una misura della qualità di un servizio, come originariamente affermato da questi ultimi autori. Cronin et al. sostengono che la qualità di un servizio è meglio descritta dall'indicatore SQI ottenuto dalla (3) omettendo i valori  $z_{jih}$ , *paradigma SERVPERF*.

I modelli compositivi richiedono, però, di verificare che l'indicatore compositivo prescelto è effettivamente adatto a descrivere il concetto in studio: deve seguire, pertanto, una verifica di validità del modello, si veda Parasuraman et al. (1991) che utilizzano, in particolare, un modello di regressione lineare in cui la variabile dipendente è costituita da una valutazione complessiva di soddisfacimento, indicata, da parte dell'intervistato, su una scala di punteggio intero 1-10, mentre le variabili esplicative sono le dimensioni.

Si noti che un punto di debolezza del modello compositivo qui considerato risiede nel fatto che i punteggi convenzionali, che sono valori su scale ordinali, vengono assunti come valori di una stessa scala ad intervalli.



#### 4.6 ANALISI DISCRIMINANTE LINEARE

L'analisi discriminante è un metodo di classificazione che misura l'importanza dei fattori che determinano l'appartenenza di un'osservazione ad un gruppo. Per esempio, supponiamo di voler stabilire quali sono i fattori che inducono i clienti di una banca a non rinnovare i loro certificati di deposito. Se il modello riuscirà a identificare i fattori giusti, allora potremo usare lo stesso modello per "discriminare" coloro che probabilmente rinnoveranno i certificati da coloro che probabilmente non li rinnoveranno.

L'analisi discriminante viene condotta per definire una modalità di assegnazione dei casi a differenti gruppi, in funzione di una serie di variabili fra di loro correlate. I gruppi sono già definiti al momento dell'analisi, pertanto l'interesse è rivolto **definire un modello che consenta di assegnare un nuovo caso ad un gruppo predefinito**, in funzione di un certo numero di variabili. Questa analisi è molto usata in CS, come nel caso di una serie di clienti, studiati tramite una serie di variabili di tipo valutativo. Tramite l'analisi discriminante è possibile definire un modello matematico che ci consenta di collocare un eventuale nuovo cliente, una volta misurate le variabili, in uno dei due gruppi (soddisfatto o non soddisfatto), in modo tale da ottimizzare le politiche di marketing.

Nell'analisi discriminante si trova una combinazione lineare di variabili che consente di calcolare il coefficiente di discriminazione (D):

$$D = b_0 + b_1x_1 + \dots + b_nx_n$$

dove:

$D$  = coefficiente discriminante

$b_0$  = costante

$x_n$  = n - ma variabile indipendente

$b_n$  = n - mo coefficiente delle funzioni discr.

Il numero di funzioni discriminanti ottenibili è uguale a  $k-1$  dove  $k$  è il numero dei gruppi. Per semplicità considereremo solo il caso della discriminazione fra due gruppi, ma le considerazioni fatte sono estendibili alla discriminazione fra più gruppi.

Il metodo di calcolo della funzione discriminante e' quello dei **minimi quadrati**, analogo a quello studiato per la regressione lineare multipla, che consente di ottenere, per i valori di D, una variabilita' minima all' interno dei gruppi e massima fra i gruppi. Per due gruppi (A, B) essendo k=2 esistera' una sola funzione discriminante i cui coefficienti sono dati dalla risoluzione della seguente equazione:

$$b = \Delta x \cdot W$$

dove:

$$b = \begin{pmatrix} b_0 \\ b_1 \\ . \\ b_n \end{pmatrix}$$

$$\Delta x = \begin{pmatrix} \bar{x}_{1A} - \bar{x}_{1B} \\ \bar{x}_{2A} - \bar{x}_{2B} \\ \dots \\ \bar{x}_{nA} - \bar{x}_{nB} \end{pmatrix}$$

$W$  = matrice di dispersione comune

$\bar{x}_{nA}$  = media della n - ma variabile del gruppo A

$\bar{x}_{nB}$  = media della n - ma variabile del gruppo B

I **valori medi** dei coefficienti di discriminazione per i due gruppi sono calcolabili nel seguente modo:

$$\bar{D}_A = b_0 + b_1 \bar{x}_{1A} + \dots + b_n \bar{x}_{nA}$$

$$\bar{D}_B = b_0 + b_1 \bar{x}_{1B} + \dots + b_n \bar{x}_{nB}$$

con:

$\bar{x}_{nA}$  = media della n - ma variabile indep. del gruppo A

$\bar{x}_{nB}$  = media della n - ma variabile indep. del gruppo B

Le **varianze** del coefficiente di discriminazione per i due gruppi sono cosi' calcolabili:

$$s^2_{D_A} = b \cdot X'_A X_A \cdot b$$

$$s^2_{D_B} = b \cdot X'_B X_B \cdot b$$

dove:

$X'_A X_A$  = matrice di dispersione del gruppo A

$X'_B X_B$  = matrice di dispersione del gruppo B

La **soglia discriminante** deve essere tanto piu' vicina ad una delle due medie tanto minore e' la deviazione standard del gruppo corrispondente, pertanto viene calcolata nel seguente modo:

$$D_0 = \frac{\overline{D}_A s_{D_B} + \overline{D}_B s_{D_A}}{s_{D_A} + s_{D_B}}$$

Pertanto un nuovo elemento (ad esempio un nuovo paziente) viene assegnato al gruppo in funzione della soglia discriminante: se il valore di D e' superiore alla soglia viene assegnato al gruppo con la media di D piu' alta, viceversa se il valore e' piu' piccolo. All' assegnamento di un elemento ad un gruppo puo' essere assegnata una probabilita' , in funzione del valore dello scarto standardizzato fra il valore della media di gruppo ed il valore della soglia discriminante, come visto per gli scarti standardizzati della distribuzione di frequenza campionaria normale:

· per il gruppo A:

$$z_A = \frac{D_0 - \overline{D}_A}{s_{D_A}}$$

con:

$$P(z_A) \Rightarrow \text{Tavole } f(z)$$

· per il gruppo B:

$$z_B = \frac{\overline{D}_B - D_0}{s_{D_B}}$$

con:

$$P(z_B) \Rightarrow \text{Tavole } f(z)$$

## CAPITOLO 5 -

### **MODELLI STRUTTURALI (DIPENDENZA CONOSCITIVA) E RETI NEURALI**

#### 5. 1 ANALISI FATTORIALE E MODELLI STRUTTURALI LINEARI

L' analisi fattoriale è una tecnica di riduzione dei dati che costruisce un modello a partire dai dati grezzi. L' analisi fattoriale ricava fattori riassuntivi detti anche "variabili latenti" che concentrano le informazioni contenute originariamente in un numero elevato di variabili. Essa rientra nell'ambito dello studio dell'interdipendenza tra variabili di tipo quantitativo.

Lo scopo è quello di condensare l'informazione contenuta in un numero elevato di variabili originarie in un numero esiguo di nuove variabili, *fattori latenti*, ottenute come combinazione lineare delle variabili originarie, con una minima perdita di informazioni.

Per esempio, in un' indagine di mercato viene chiesto un giudizio su 9 caratteristiche di un prodotto. L' analisi fattoriale trova tre fattori latenti. Le variabili che hanno "pesato" di più nell' identificazione di questi fattori, forniscono informazioni su cosa i fattori potrebbero rappresentare. Per esempio, se tre attributi come supporto tecnico, servizio clienti e formazione pesano fortemente su un fattore, potremmo definirlo "livello di servizio". Questa tecnica aiuta a identificare variabili che non sono state misurate, ma che sono "manifestate" da variabili osservate.

Per costruire un modello fattoriale, il numero dei fattori dovrebbe essere noto a priori.

Ad esempio il comportamento dei clienti di una banca, che dal punto di vista quantitativo può essere rilevato attraverso le variabili di movimentazione, utilizzo plastic cards, rimborso prestiti, presenza assicurazioni, risparmio in titoli, etc..., può essere espresso in funzione di due fattori latenti quali la *ricchezza* e l'*utilizzo* non misurabili direttamente. Da tale trasformazione è possibile ricavare la matrice di varianza e covarianza delle nuove variabili, ossia  $V_Z = TV_X T'$  che sarà una matrice diagonale i cui valori sono le varianze delle nuove variabili che non sono altro che gli autovalori della matrice  $V_X$ .

Un' altra buona applicazione dell' analisi fattoriale è il raggruppamento di prodotti in base a similitudini rilevate nel comportamento d' acquisto. In questo modo si possono rilevare opportunità di vendite incrociate e determinare la convenienza di offerte cumulative di più prodotti. In base a questo raggruppamento sarà possibile progettare offerte cumulative, o tentare di vendere un prodotto ai clienti che ne hanno già comprato uno dello stesso gruppo.

Un tale modello consente di stabilire, su base campionaria, i legami delle variabili latenti, associate alle dimensioni, fra di loro e con quella – endogena – attinente alla “customer satisfaction”, che sta a fondamento del costrutto, e rende possibile, la *determinazione completa del costrutto concettuale* secondo metodi rigorosi di inferenza statistica. Il modello consta di due sistemi di equazioni:

a) equazioni strutturali

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}, \text{ da cui si ammetta segua: } \boldsymbol{\eta} = (\mathbf{I} - \mathbf{B})^{-1} (\boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}), \quad (1)$$

dove:  $\mathbf{B}$ ,  $\boldsymbol{\Gamma}$ : sono matrici  $m \times m$ ,  $m \times q$  di incogniti parametri;  $\boldsymbol{\eta}$ : variabile causale  $m$ -dimens.,  $E(\boldsymbol{\eta}) = \mathbf{0}$ , variabili latenti endogene;  $\boldsymbol{\xi}$ : variabile causale  $q$ -dimens.,  $E(\boldsymbol{\xi}) = \mathbf{0}$ , variabili latenti esogene  $E(\boldsymbol{\xi}\boldsymbol{\xi}') = \boldsymbol{\Phi}$ ;  $\boldsymbol{\zeta}$ : variabile causale  $m$ -dimens.,  $E(\boldsymbol{\zeta}) = \mathbf{0}$ ,  $E(\boldsymbol{\xi}\boldsymbol{\zeta}') = \mathbf{0}$ ,  $E(\boldsymbol{\zeta}\boldsymbol{\zeta}') = \boldsymbol{\Psi}$ , errori casuali.

b) modello di misurazione

$$\begin{aligned} \mathbf{Y} = \mathbf{\ddot{E}}_y \boldsymbol{\zeta} + \mathbf{\ddot{\alpha}} &= \begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{M} \\ \mathbf{Y}_m \end{bmatrix} = \begin{bmatrix} {}_y\lambda_1 \boldsymbol{\zeta}_1 \\ \mathbf{M} \\ {}_y\lambda_m \boldsymbol{\zeta}_m \end{bmatrix} + \mathbf{\ddot{\alpha}} \\ \mathbf{X} = \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta} &= \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{M} \\ \mathbf{X}_q \end{bmatrix} = \begin{bmatrix} {}_y\lambda_1 \boldsymbol{\xi}_1 \\ \mathbf{M} \\ {}_x\lambda_q \boldsymbol{\xi}_q \end{bmatrix} + \boldsymbol{\delta}, \quad (6) \end{aligned}$$

$\mathbf{Y}$ : vettore di variabili casuali **osservabili**, in genere almeno 2 per ogni  $\eta_i$ ;  ${}_y\lambda_i$ : vettori di incogniti parametri;  $\boldsymbol{\varepsilon}$ : errori casuali,  $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ ,  $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \boldsymbol{\Theta}_{\boldsymbol{\varepsilon}}$ ,  $E(\boldsymbol{\eta}\boldsymbol{\varepsilon}') = E(\boldsymbol{\xi}\boldsymbol{\varepsilon}') = \mathbf{0}$ ;  $\mathbf{X}$ ,  ${}_x\lambda_j$ ,

$\delta$ : analoga configurazione,  $E(\delta) = \mathbf{0}$ ,  $E(\delta\delta') = \Theta_\delta$ ,  $E(\eta\delta') = E(\xi\delta') = \mathbf{0}$ ,  $E(\epsilon\delta') = \mathbf{0}$ , in genere.

I modelli ad equazioni strutturali con variabili latenti, forniscono una metodologia per valutare gli elementi principali della verifica di una teoria, ovvero l'adeguatezza della misurazione delle variabili e l'analisi delle relazioni ipotizzate tra le variabili stesse. In tale ambito, particolare interesse ha sollevato, in letteratura, il problema della costruzione e della stima dei parametri e della validazione dei modelli strutturali. Due approcci sono seguiti, principalmente, in letteratura per la stima dei parametri del modello: la massima verosimiglianza (Joreskog, 1970) e il PLS (Wold, 1985, Tenenhaus, 1999) il cui approccio geometrico (soft modeling) è inquadrabile nell'ambito dell'Analisi dei Dati.

I metodi di stima dei parametri correntemente utilizzati sono: 1) il criterio della massima verosimiglianza che, nell'ipotesi di normalità congiunta delle componenti di  $\bullet Y', X' \bullet$ , assicura stime asintoticamente efficienti ed distribuite normalmente; 2) il criterio dei minimi quadrati generalizzato che assicura le stesse proprietà ma non richiede l'assunzione di normalità delle osservazioni; 3) il criterio dei minimi quadrati ordinari. Per ciascuno dei criteri anzidetti l'ottimizzazione è condotta per via numerica.

Esiste una copiosa disponibilità di software incluso nei "package" statistici di uso corrente: SPSS, SAS, Statistica, ecc.. Per le condizioni di identificabilità parametrica e le tecniche di validazione del modello (1), si veda Bollen (1989). Anche i valori delle variabili latenti, in primo luogo di quella descrittiva della "customer satisfaction", devono essere stimati ed esistono al riguardo, ad esempio, metodi basati su "tecniche di regressione", si veda ad esempio Bollen (1989).

Il metodo di stima cosiddetto del "Partial Least Squares" (PLS) (Metodo dei Minimi Quadrati Parziali) procede esattamente in modo "opposto" rispetto ai metodi visti in a): 1) si "stimano" in primo luogo i valori delle variabili latenti, quindi anche di quella che è associata alla "customer satisfaction", come combinazioni lineari delle corrispondenti variabili manifeste, con coefficienti ottenuti con un procedimento iterativo; 2) si stimano, quindi, i parametri del modello mediante il principio dei minimi quadrati ordinari applicato alle singole "equazioni di regressione" cui si è ricondotti, sia per la

parte strutturale che per le equazioni di misurazione (si veda l'articolo fondamentale di Wold (1985)). È il metodo utilizzato per il calcolo degli indici nazionali quali l'ACSI e l'ECSI.

Negli ultimi anni l'utilizzo dei modelli ad equazioni strutturali, nell'ambito della Customer Satisfaction, ha avuto un notevole incremento in termini di diffusione (ad es: Boari, 2000, Manaresi *et al.*, 2000). Ciò è dovuto alla possibilità, offerta da tali modelli, di rendere rigoroso il procedimento di definizione del concetto di customer satisfaction e, quindi, per le sue valutazioni.

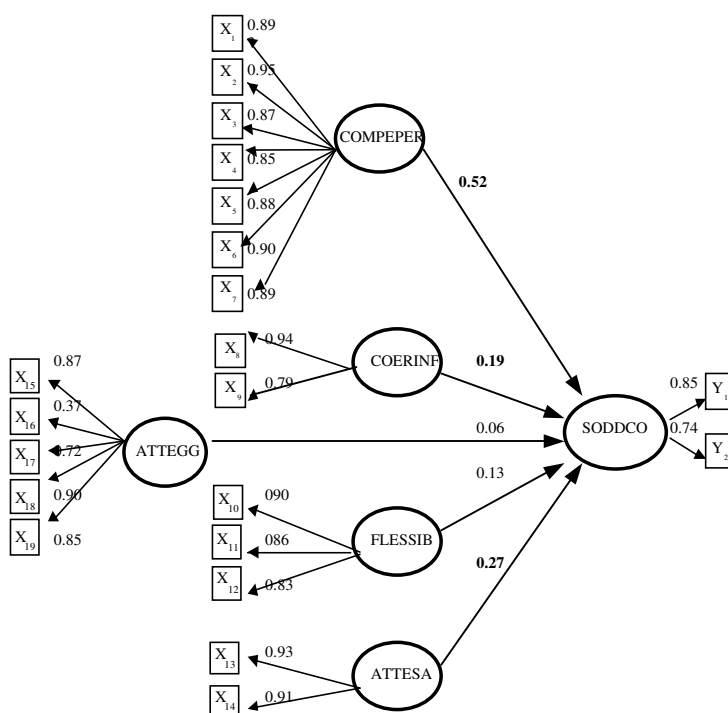
Sull'onda dell'esigenza di una maggior scientificità negli approcci, sono state sfruttate delle tecniche più avanzate e sofisticate, come gli algoritmi genetici (GA) o il metodo della regressione dei minimi quadrati parziali (PLS), il pioniere del quale fu Herman Wold (1908-1992), economista e studioso di statistica svedese.

## 5.2 L'ANALISI LISREL

L'approccio LISREL permette sia di misurare variabili che rappresentano costrutti teorici non osservabili, sia di trattare la causalità esistente tra variabili, latenti o osservate. La duplice natura di questo approccio è rilevabile nelle due parti che lo costituiscono: il modello di misurazione e il modello strutturale: il primo specifica come le variabili latenti siano misurate tramite le variabili osservate e serve per determinare i caratteri di tale misurazione; il secondo specifica le relazioni fra le variabili latenti e serve per determinare gli effetti causali e l'ammontare di varianza non spiegata (Jöreskog e Sörbom, 1993).

Per la fase di stima dei parametri del modello, poiché le variabili sono di tipo ordinale e presentano distribuzione con forte asimmetria negativa (come di norma accade per variabili che rilevano la soddisfazione) è necessario ricorrere al metodo dei minimi quadrati pesati (WLS) che richiede alcune elaborazioni preliminari per calcolare un' opportuna matrice di pesi per la funzione di adattamento, su cui si basa il processo di verifica e stima dei parametri. In figura riportiamo un esempio del modello con le relazioni di causalità e i parametri stimati.

	INDICATORE
X <sub>1</sub>	Personale gentile e cortese
X <sub>2</sub>	Conoscenze del personale
X <sub>3</sub>	Mancanza competenza del personale
X <sub>4</sub>	Informazioni fornitemi complete
X <sub>5</sub>	Personale incapace di consigliarmi
X <sub>6</sub>	Personale disponibile ad aiutarmi
X <sub>7</sub>	Chiarezza del personale
X <sub>8</sub>	Durata dell' attesa accettabile
X <sub>9</sub>	Condizioni dell' attesa accettabili
X <sub>10</sub>	Personale elastico
X <sub>11</sub>	Personale indifferente ai problemi
X <sub>12</sub>	Personale capisce le esigenze
X <sub>13</sub>	Informazioni coerenti con segreteria
X <sub>14</sub>	Informazioni coerenti con Università
X <sub>15</sub>	Competenza personale
X <sub>16</sub>	Inadeguatezza orari apertura
X <sub>17</sub>	Chiarezza informazioni (generale)
X <sub>18</sub>	Efficienza ed organizzazione
X <sub>19</sub>	Accessibilità del servizio
Y <sub>1</sub>	Soddisfazione complessiva
Y <sub>2</sub>	Aspettative





### 5.3 L'ACIMO-PLS PER LA VALUTAZIONE DELLA CUSTOMER SATISFACTION

Nel modello ad equazioni strutturale ipotizzato, nel caso in cui siamo in presenza di un solo modulo, figura un primo sistema lineare di equazioni con coefficienti incogniti che collega, fra loro, un insieme di variabili, non osservabili (endogene) (esempio: grado di soddisfazione) così come con un secondo insieme di variabili, pure non osservabili, (esogene) (aspetti organizzativi) (2.a). Questa struttura è completata da due altri insiemi di equazioni lineari che collegano le variabili endogene ed esogene ad altre invece osservabili (2.b) (2.c). Ogni insieme di equazioni è perturbato per la presenza di errori accidentali. Il grado di soddisfazione complessivo è identificato con una delle variabili latenti.

$$\eta = B\eta + \Gamma\xi + \zeta \quad (a); \quad X = \Lambda_X\xi + \delta \quad (b); \quad Y = \Lambda_Y\eta + \varepsilon \quad (c)$$

(2)

dove  $\eta$  è il vettore delle variabili latenti endogene,  $\xi$  è il vettore delle variabili latenti esogene con  $\zeta$  vettore degli errori;  $B$  e  $\Gamma$  sono le matrici di coefficienti strutturali e si riferiscono la prima ai legami tra endogene, la seconda ai legami tra endogene ed esogene;  $X$  e  $\delta$  sono i vettori riferiti, rispettivamente, alle variabili esogene osservate ed agli errori,  $\Lambda_X$  è la matrice dei coefficienti strutturali tra variabili esogene osservate e latenti;  $Y$  e  $\varepsilon$  sono i vettori delle variabili endogene osservate e degli errori,  $\Lambda_Y$  è la matrice dei coefficienti strutturali tra le variabili endogene osservate e le latenti. Diverse sono le motivazioni per la scelta dell'approccio da utilizzare per la stima dei parametri. In generale, l'approccio di massima verosimiglianza si basa sull'ipotesi di multinormalità dei dati ed inoltre l'algoritmo di risoluzione prende le mosse dalla modellizzazione della matrice di covarianza, che, in alcuni casi, può portare a problemi d'identificazione e di non convergenza. Viceversa l'approccio PLS non si basa su ipotesi probabilistiche forti; inoltre, andando ad operare direttamente sui dati, attraverso

delle regressioni semplici o multiple, non porta a problemi di identificazione. Nel nostro caso le stime dei parametri avverranno con il PLS (Tenenhaus, 1999).

	<b>PLS</b>	<b>Lisrel</b>
<i>Approccio</i>	Esplorativo	Confermativo
<i>Ipotesi distributiva</i>	Nessuna	Distribuzione Normale Multivariata
<i>Stime dei parametri</i>	Consistenti al crescere della dimensione campionaria	Consistenti
<i>Variabili latenti</i>	<b>Stimate esplicitamente</b>	Indeterminate
<i>Implicazioni</i>	Ottimale per l'accuratezza della predizione	Ottimale per l'accuratezza dei parametri
<i>Complessità del modello</i>	Elevata (< 1000 indicatori)	Bassa (< 100 indicatori)
<i>Dim. campionaria min.</i>	Piccola (30-100 casi)	Media (200-800 casi)

Il modello ad equazioni strutturali, in presenza di più moduli, diviene particolarmente complesso poiché ciascuna variabile esogena è rilevata in occasioni differenti. Una sintesi ottimale delle misure delle singole variabili nei vari moduli, da utilizzare successivamente nel modello strutturale, può essere ottenuta con una estensione multi-tabellare del PLS: l'ACIMO-PLS (Analyse de Co-Inertie Multiple Ortogonale – Partial Least Square; Vivien, Sabatier). L'ACIMO-PLS è un metodo lineare iterativo basato sulla massimizzazione vincolata del criterio soggiacente il PLS. Rispetto al nostro caso, siano  $X_k$  ( $k=1, \dots, K$ ) variabili esogene, rilevate rispettivamente, in  $p_1, K, p_K$  moduli ed un gruppo di  $q$  variabili endogene identificato con  $Y$  raccolte sugli stessi  $n$  individui (Fig. 1.a). Scopo dell'ACIMO-PLS è quello di ricercare successivamente  $s$  ( $s=1, K, S$ ) componenti  $u_{k,s}$  per ciascuna matrice  $X_k$  e  $t_s$  per la tabella  $Y$  tali che: a) le  $t_s$  componenti sono tra loro ortogonali e sono le più esplicative di  $Y$  e b) le componenti  $u_{k,s}$  sono quelle che ottimizzano i criteri di massima rappresentatività del proprio gruppo di variabili  $X_k$  e di massimo legame con le  $t_s$ . Una estensione dell'ACIMO-PLS in presenza di variabili qualitative con modalità ordinate è stata proposta da D'Ambra *et al.* (2000b, 2000c) e Vivien e Sabatin (2000).

In particolare l'ACIMO-PLS di ordine 1 massimizza la seguente funzione

$$\text{obiettivo } f(b_1, \dots, b_K, a) = \sum_{k=1}^K \text{cov}^2(X_k Q_k b_k, Y Q a), \quad \text{con } \|b_k\|_{Q_k}^2 = 1, \forall k \quad \text{e}$$

$$\|a\|_Q^2 = 1, \text{ determinando } K+1 \text{ componenti di ordine 1, di cui le } u_{k,1} = X_k Q_k b_{k,1}$$

sono associate alle  $K$  tabelle  $X_k$  e la  $t_1 = YQa_1$  è associata alla tabella  $Y$ . Tale tecnica viene, quindi, a determinare le combinazioni lineari delle variabili esogene più legate alle variabili endogene. Tale caratteristica ci porta a preferire l'ACIMO-PLS ad altre tecniche di sintesi di matrici a tre vie (es: Analisi delle matrici principali, Rizzi, 1988) in cui le componenti di sintesi vengono determinate senza tener conto delle variabili endogene.

Le  $K$  componenti  $u_{k,1}$  così individuate, sintesi ottimali delle variabili per i differenti moduli, vengono ad assumere il ruolo di variabili esogene nel modello strutturale rispetto alle variabili endogene  $Y$ , la cui stima dei parametri avverrà ancora con il PLS.

La scelta dell' ACIMOPLS viene, quindi, motivata sia per il criterio di determinazione delle componenti sia per la coerenza con il metodo di stima dei parametri nel modello strutturale.

#### 5.4 L'UTILIZZO DELLE RETI NEURALI NELLA CS

Benché le reti neurali abbiano una storia relativamente recente, esiste già una vasta letteratura, sulla loro costruzione e sul loro funzionamento.

Premesso che esistono vari tipi di reti neurali, può essere sufficiente sapere che il principio di fondo consiste nella capacità di certi algoritmi di calcolo di emulare il comportamento umano estraendo valutazioni e considerazioni da situazioni complesse, non sempre ben definite, talvolta anche contraddittorie.

Come gli esseri umani, le reti neurali hanno la capacità di apprendere dalle esperienze pregresse per poi applicare a circostanze nuove le conoscenze acquisite. A differenza degli esseri umani, hanno la capacità di crearsi un sistema rappresentativo delle molteplici relazioni esistenti tra le variabili causali di un sistema complesso manifestando, di conseguenza, elevata attitudine ad esprimere con regolarità valutazioni appropriate e comportamenti adeguati in presenza di situazioni apparentemente caotiche.

Il principio di funzionamento di una rete neurale è molto semplice.

Al programma viene somministrata una certa quantità di esempi rappresentativi del problema da affrontare, con (reti neurali supervisionate) o senza (reti neurali non supervisionate) le corrispondenti soluzioni.

Le reti neurali non supervisionate esaminano gli esempi e creano dei raggruppamenti (clusters) in base al reciproco livello di affinità.

La rete supervisionata esamina gli esempi proposti e crea una serie di valori, detti pesi, la cui interazione con le variabili del problema, sulla base di determinate funzioni matematiche, produce delle ipotesi di soluzione; confronta quindi queste soluzioni teoriche con quelle effettive, ne misura lo scostamento e, se tale scostamento supera una soglia massima predefinita, torna indietro, modifica i pesi e ripete il ciclo.

In sostanza, procede per approssimazioni successive finché non vengono individuate soluzioni accettabilmente simili a quelle reali.

Ultimata questa fase, detta di addestramento, si passa a quella di verifica con la somministrazione di alcuni esempi del problema, diversi da quelli già utilizzati, le cui

soluzioni, questa volta, non vengono comunicate. La rete, sulla base dei pesi elaborati durante l'addestramento, fornisce le proprie soluzioni la cui validità è strettamente dipendente dal livello di efficienza raggiunto.

Se l'esito è soddisfacente, la rete è pronta per essere utilizzata nel settore di applicazione per il quale è stata creata, al fine di individuare delle soluzioni non conosciute.

Le reti neurali sono composte da una serie di strati, ciascuno dei quali possiede una funzione specifica.

Essenziali sono lo strato di input e quello di output. E' facoltativa in alcuni casi, ma in effetti necessaria per la soluzione di problemi complessi, la presenza di almeno uno strato cosiddetto nascosto.

Lo strato di input è costituito da tante unità elaborative quante sono le variabili indipendenti del problema da esaminare.

Lo strato di output, invece, è costituito da tante unità elaborative quante sono le variabili dipendenti che costituiscono la soluzione del problema.

Lo strato nascosto, infine, ha lo scopo di creare la rappresentazione del problema internamente alla rete. Poiché non esiste alcuna regola per la determinazione del numero di unità elaborative che devono comporre tale strato, si procede di solito per tentativi sulla base dell'esperienza.

La particolare struttura delle reti richiede precise modalità di rappresentazione del problema da risolvere. Tutte le variabili, cioè, devono trovare rappresentazione in forma numerica per ciascuna unità sia di input che di output.

Nel caso di variabili dicotomiche i valori possibili sono 0 e 1 oppure -1 e 1 in relazione al tipo di rete. Nel caso di variabili continue, invece, devono essere compresi tra -1 e 1 oppure tra 0 e 1. Valori diversi da quelli consentiti devono essere ricondotti alla forma richiesta attraverso un trattamento preliminare di interpolazione che, comunque, viene effettuato automaticamente dai programmi più recenti.

Una volta formalizzata la rappresentazione delle variabili, l'addestramento può avere inizio.

Ogni ciclo completo di confronto delle soluzioni reali di tutti gli esempi con le soluzioni proposte dalla rete prende il nome di epoca.

L'addestramento ha termine o dopo un certo numero di epoche prefissato dallo sperimentatore o dopo che il margine di scostamento tra soluzioni calcolate e soluzioni effettive si sia progressivamente ridotto a un livello predefinito.

Poiché non è sempre possibile identificare con precisione i fattori che incidono su un determinato fenomeno, può risultare conveniente abbondare nella somministrazione delle variabili di input. Sarà la stessa rete a riconoscere autonomamente quelle maggiormente significative, alle quali assegnerà pesi più consistenti, e quelle marginali, alle quali assegnerà pesi ridotti.

Il calcolo neurale è una tecnica di analisi dell'informazione. Alla base delle reti neurali si pone lo studio dei meccanismi di apprendimento del cervello umano, e della possibilità di riprodurli artificialmente. In pratica, le reti neurali sono sistemi che apprendono dall'esperienza: osservando un insieme di dati, le reti neurali imparano a riconoscere relazioni e modelli esistenti. Sebbene le reti neurali siano un ramo della ricerca sull'intelligenza artificiale, non vanno confuse con i sistemi esperti. A differenza dei sistemi esperti, infatti, le reti neurali non applicano regole note ai dati, ma costruiscono modelli basati sui dati. Questo comportamento è più flessibile e più facilmente adattabile a situazioni mutevoli nel tempo: un sistema esperto va riprogrammato ogni volta che cambiano le regole. Una rete neurale si accorge e segnala che le regole sono cambiate al variare dei dati in input.

Finora abbiamo trattato delle Reti Neurali Artificiali (*Artificial Neural Network*, ANN) di tipo supervisionato, che si prestano a risolvere problemi predittivi.

In effetti, i modelli ANN si basano sull'esistenza di una variabile obiettivo, rappresentata nell'architettura della rete neurale nello strato di output. Questo tipo di ANN viene detto *supervisionato* poiché i vari algoritmi di apprendimento lavorano basandosi sul *confronto tra il valore della variabile obiettivo stimato dal modello e quello osservato*; la *performance* del modello può essere quindi facilmente controllata grazie alla valutazione di tale scarto o residuo. Nell'ambito delle reti neurali artificiali esiste, tuttavia, un'altra famiglia di modelli costituita dalle reti neurali **non**

**supervisionate**, anche note come *Self Organizing Maps (SOMs)*. Questo genere di reti neurali sono state introdotte da Teuvo Kohonen e consentono di individuare per un certo insieme di records, una partizione data da gruppi omogenei rispetto alle variabili considerate.

Questa problematica risulta di notevole interesse in numerosi ambiti ed era già stata affrontata in statistica con lo sviluppo delle tecniche di *cluster analysis*, di cui parleremo successivamente.

Una delle applicazioni di *Data Mining* più comuni di questi modelli è la segmentazione comportamentale della clientela, che viene condotta quando l'azienda ha interesse nell'individuare segmenti della propria clientela simili rispetto al comportamento d'acquisto, per poter poi intraprendere iniziative di marketing diversificate per ciascun segmento/cluster. Le SOMs si differenziano dalle reti supervisionate *Multi-Layer Perceptron (MLP)* sia nella topologia, che nel metodo di apprendimento.

Le *Self Organizing Maps di Kohonen*, come le reti neurali supervisionate già viste, richiedono che le eventuali variabili categoriche siano ricodificate in tante variabili binarie (presenza/assenza) quante sono le categorie. Proseguendo l'analogia con le reti MLP, le SOMs sono formate - nella loro versione più semplice - da uno strato di input ed uno di output; dove però quest'ultimo non identifica la variabile obiettivo.

Nell'architettura delle SOMs, quindi, lo strato di input è composto da tanti neuroni quante sono le variabili esplicative di tipo numerico, mentre lo strato di output è dato da una griglia bidimensionale – rappresentabile in un piano cartesiano - di neuroni che costituiscono i potenziali clusters. Questa è la prima differenza sostanziale tra reti neurali supervisionate e non. La struttura delle connessioni di una SOM, invece, è simile a quella delle reti MLP nel senso che ogni neurone della griglia di output è connesso a ciascuna unità dello strato di input; mentre i neuroni che formano lo strato di output non sono connessi tra di loro. Ciò nonostante la struttura a griglia dello strato di output gioca un ruolo fondamentale nella fase di apprendimento della SOM.

L'algoritmo di training delle *Self Organizing Maps* di Kohonen prevede che i records vengano considerati ad uno ad uno per essere assegnati al neurone della griglia più simile rispetto alle variabili introdotte nello strato di input. Nella prima iterazione si

inizializzano casualmente i pesi associati alle connessioni tra neuroni di input e di output. I neuroni della griglia assumono in questo modo un profilo iniziale, che sarà poi modificato nel corso dell' apprendimento ogni volta che questi conquisteranno un nuovo record.

A differenza degli algoritmi statistici di cluster analysis non gerarchica, le SOMs hanno un meccanismo per evitare la formazione di pochi cluster molto numerosi ed altri poco popolati. In particolare, quando un neurone della griglia acquisisce o "vince" un record, viene previsto anche un cambiamento nel profilo dei neuroni vicini per favorire la possibilità che anche questi possano "catturare" records simili a quello appena selezionato.

La definizione dell' ampiezza del vicinato avviene mediante un opportuno parametro che cambia nel corso dell' apprendimento, decadendo con trend lineare o esponenziale. Se in una prima fase, infatti, il condizionamento tra i neuroni della griglia deve essere forte, e quindi bisogna definire un intorno d' azione ampio per tali neuroni, nella seconda fase si richiede che questo vicinato sia più ristretto per consentire la convergenza dell' algoritmo. La procedura descritta viene iterata fino a quando il flusso di records tra i neuroni della griglia non si stabilizza. A quel punto le unità "non vuote" presenti nello strato di output costituiscono i clusters o segmenti finali.

La procedura di apprendimento fa in modo che all' interno di ogni cluster si riscontri un' omogeneità dei records rispetto alle variabili utilizzate in input, mentre si abbia una forte diversità tra i profili medi (o centroidi) dei clusters.

Le reti neurali non supervisionate hanno un costo computazionale generalmente maggiore degli algoritmi di cluster non gerarchica, come il Kmedie. Tuttavia, rispetto a quest' ultimo algoritmo, le SOMs non richiedono di fissare inizialmente il numero di clusters ed assicurano inoltre una più equa distribuzione dei records tra i vari segmenti. Il vantaggio in termini pratici appare evidente se riprendiamo l' esempio della segmentazione comportamentale della clientela. In tale contesto è chiaro che si ricercano sottogruppi di clienti, oltre che omogenei rispetto al comportamento d' acquisto, anche sufficientemente consistenti per motivare azioni di CS diversificate. Detto altrimenti, l' individuazione di un cluster di soli 100 clienti su una customer-base



di milioni di records non costituirebbe un' informazione spendibile e comprometterebbe l' intera analisi. Per questa ragione le *Self Organizing Maps* di Kohonen trovano un largo impiego nelle applicazioni di Data Mining.

Queste informazioni permettono di intraprendere azioni mirate, specifiche per ogni segmento, in grado di incrementare sia il business che la soddisfazione del cliente.

Ancora una volta si è visto come le reti neurali rappresentino uno strumento analitico in grado di generare conoscenza ad alto valore e fornire un supporto decisionale importante per la migliore gestione dell' azienda nel rapporto con la propria clientela.

Le R. N. costituiscono un valido sostegno al Data Mining in quanto hanno:

- capacità di valutare un gran numero di fattori
- tolleranza verso dati imperfetti (presenza di dati mancanti, problemi di qualità dei dati)

Benché esistano molte possibilità di utilizzo di questo sofisticato strumento, va notato che una imprecisa individuazione del tipo dei valori da fornire in input costituisce spesso impedimento al buon esito dell'addestramento. Questa difficoltà, i tempi lunghi normalmente richiesti dalla sperimentazione e la necessità del possesso di alcune conoscenze specifiche da parte dell'utente hanno finora costituito un ostacolo di rilievo all'utilizzo diffuso delle reti neurali.

## CAPITOLO 6-

### **MODELLI DI STATISTICHE MULTIVARIATE (INFERENZIALI)**

#### 6.1 MODELLI DIRETTI ESPLICATIVI O DECOMPOSITIVI: MODELLI DI REGRESSIONE

L'approccio compositivo prima illustrato, anche se pone il problema della scelta più appropriata dell'indicatore che riassume il contributo delle varie dimensioni, cfr. Lauro et al. (1997), presenta, però, chiare connotazioni di "scientificità". Nei casi con creti, però, pensiamo ai cosiddetti beni "di largo consumo" per i quali non sussistono riferimenti teorici consolidati, come per i servizi, i modelli di regressione consentono di trattare in modo semplificato, ma unitario, una misura globale di "customer satisfaction" e la verifica del relativo costrutto concettuale. Nel seguito per brevità possono solo elencarsi i più consueti modelli e tecniche di analisi di regressione. Si ammetterà che attraverso un questionario si sia anche ottenuta una *valutazione globale o complessiva* del prodotto/servizio in esame, diciamo  $V$ , espressa in una scala semantico-differenziale, quindi ordinata, con associati dei punteggi convenzionali (ad esempio scala a 5 o 7 punti); detta scala potrà anche essere, o venire ridotta, a dicotomica in corrispondenza a due valutazioni riconducibili a "soddisfatto", "non soddisfatto".

## 6.2 MODELLO DI REGRESSIONE ‘LOGISTICA’ LOGIT.

L’approccio della regressione logistica consente di rispettare le caratteristiche di scala ordinale della risposta  ${}_cV$ , facendo riferimento in modo diretto alla distribuzione di probabilità di quest’ultima. Si descriva il soddisfacimento complessivo come probabilità di ottenere il giudizio ‘soddisfatto’, valore convenzionale 1, condizionatamente alle modalità che certi fattori concomitanti presentano per il rispondente, riassunte nel vettore  $\mathbf{x}$ ,  $q \times 1$ ; sia  $P({}_cV = 1 \mid \mathbf{x}) = P(1 \mid \mathbf{x})$ , detta probabilità. Il modello di regressione logistica pone:

$$P(1 \mid \mathbf{x}) = \frac{\exp(\beta_0 + \beta'_1 \mathbf{x})}{1 + \exp(\beta_0 + \beta'_1 \mathbf{x})} = F(\beta_0 + \beta'_1 \mathbf{x}) \quad (1)$$

dove  $\beta' = (\beta_0, \beta'_1)$  è un vettore,  $1 \times (q + 1)$ , di incogniti parametri e, si ricordi che  $F(x) = \exp(x) / (1 + \exp(x))$  è la funzione di ripartizione della distribuzione logistica. Per la teoria statistica che consente, tipicamente con riferimento a campioni ‘sufficientemente grandi’, l’analisi del modello (4) si veda Amemiya (1985). L’analisi statistica del modello (1) può condursi in modo automatizzato, formalmente abbastanza simile a quello dell’analisi della regressione lineare, mediante programmi inclusi nei ‘package’ più diffusi, quale, ad esempio, l’SPSS.

### 6.3 REGRESSIONE LINEARE MULTIPLA MULTIVARIATA A RISPOSTA POLITOMICA

Nella situazione applicativa sia il giudizio di valutazione globale o *overall* (variabile dipendente politomica  $Y$ ), sia le valutazioni sui singoli attributi di un prodotto (variabili indipendenti politomiche o predittori  $X$ ), sono formulati su una scala ordinale a tre passi: ‘insoddisfatto’, ‘soddisfatto’, ‘molto soddisfatto’.

Per la valutazione della qualità percepita del prodotto si può utilizzare un modello basato sulla regressione lineare multipla multivariata.

Per individuare le relazioni intercorrenti tra la variabile dipendente politomica ( $Y$ ) ed i suoi predittori qualitativi ( $X$ ), fissati in numero di  $m = 2$  nel modello in argomento, si muove dalla codifica binaria disgiuntiva completa. Le  $K = 3$  categorie o classi di giudizio di *overall* vengono espresse in funzione di tre variabili indicatrici binarie ( $Y_k, k = 1, 2, 3$ ), che assumono valore 1 se il rispondente assegna una valutazione di classe  $k$ , valore 0 nel caso contrario; parimenti, le categorie di giudizio di valutazione di ciascun attributo sono codificate secondo tre variabili binarie ( $Z_i, i = 1, 2, 3$ ) (vedi Tabella sotto).

Nel modello generale di regressione multipla multivariata *senza intercetta*, le  $K$  equazioni di regressione lineare su variabili indicatrici, collegate tra di loro tramite i residui (Sadocchi, 1981, p. 81), vengono rappresentate con  $K$  modelli *univariati* di regressione:

$$\mathbf{y}_k = \mathbf{Z} \boldsymbol{\delta}_k + \mathbf{e}_k, \quad k = 1, 2, 3, \quad (1)$$

dove:  $\mathbf{y}_k$  è il vettore colonna, di dimensioni  $n \times 1$ , delle osservazioni  $y_{kj}, j = 1, 2, \dots, n$ , sulla variabile dipendente  $k$ -sima;  $\mathbf{Z}$  matrice *fissa* (di dimensioni  $n \times 2I$ ) delle variabili indicatrici delle valutazioni ( $x_{ij} = 1, 2, 3$ ) espresse sui singoli attributi, associata a  $\mathbf{y}_k$ ;  $\boldsymbol{\delta}_k$  vettore colonna (di dimensioni  $2I \times 1$ ) dei parametri di regressione incogniti delle categorie;  $\mathbf{e}_k$  vettore colonna  $n \times 1$  degli errori eteroschedastici.

**Tabella :** *Codifica binaria disgiuntiva completa delle modalità di valutazione globale e delle valutazioni di due attributi di un prodotto*

Classi ordinali di valutazione	Variabili indicatrici delle modalità di valutazione globale			Variabili indicatrici delle modalità di valutazione di due attributi					
				attributo $X_1$			attributo $X_2$		
	$Y_1$	$Y_2$	$Y_3$	$Z_1$	$Z_2$	$Z_3$	$Z_1$	$Z_2$	$Z_3$
insoddisfatto	1	0	0	1	0	0	1	0	0
soddisfatto	0	1	0	0	1	0	0	1	0
molto soddisfatto	0	0	1	0	0	1	0	0	1

La forma algebrica del modello completo *con intercetta*, riparametrizzato in seguito alla soppressione, per entrambi gli attributi, della colonna corrispondente alla prima classe ( $Z_1$ ), onde evitare la singolarità della matrice del piano sperimentale (De Luca, 2000), è:

$$y_{kj} = c_k + \sum_{m=1}^2 \sum_{i=2}^3 \tilde{\delta}_{ki}^{(m)} \tilde{z}_{mij} + \sum_{i=1}^3 \sum_{h=1}^3 \tilde{\delta}_{kih}^{(12)} \tilde{z}_{ihj} + e_{kj}, \quad k = 1, 2, 3, \quad (2)$$

dove:  $c_k$  corrisponde alla media condizionata  $M(Y_{\tilde{z}_{k11}, \tilde{z}_{k21}}) = \tilde{\delta}_{k1}$  (media dei casi con valore 1 nel vettore di  $Y$  attinenti alla classe di riferimento);  $\tilde{\delta}_{ki}^{(m)}$  è il parametro della categoria  $i$  del predittore  $m = 1, 2$ ;  $\tilde{\delta}_{kih}^{(12)}$  è l'effetto di interazione fra la modalità  $i$  del fattore 1 e la modalità  $h$  del fattore 2;  $\tilde{z}_{ihj} = 0$  per  $i \vee h = 1$ ,  $\tilde{z}_{ihj} = 1$  per  $i, h \neq 1, j = 1, 2, \dots, n$ . La matrice  $\tilde{\mathbf{Z}}$  del piano degli esperimenti (fattoriale completo con replicazioni), sottostante la (2) è di dimensioni  $n \times 9$ . Le  $K$  equazioni (2), in forma compatta e con simbolismo matriciale, possono essere così espresse:

$$\mathbf{y}^* = \tilde{\mathbf{Z}}^* \tilde{\boldsymbol{\delta}}^* + \mathbf{e}^* \quad (3)$$

con:  $\mathbf{y}^*$  vettore composto, costituito da  $K = 3$  vettori colonna  $n \times 1$ , ognuno dei quali contiene le osservazioni della variabile dipendente su ciascuno degli  $n$  rispondenti;  $\tilde{\mathbf{Z}}^*$  matrice diagonale composta quadrata, contenente  $K \times K$  sottomatrici, delle quali le  $K$  sottomatrici  $\tilde{\mathbf{Z}}$  che si trovano sulla diagonale principale (uguali tra di loro) presentano in colonna le variabili indicatrici indipendenti corrispondenti alle diverse equazioni, mentre le restanti sottomatrici sono composte da elementi nulli;  $\tilde{\boldsymbol{\delta}}^*$  vettore composto contenente i  $K$  vettori colonna dei coefficienti di regressione  $\tilde{\boldsymbol{\delta}}_k$ ;  $\mathbf{e}^*$  vettore composto contenente i  $K$  vettori colonna degli errori  $\mathbf{e}_k$ , ciascuno di dimensioni  $n \times 1$ .

Essendo le categorie di valutazione di  $Y$  tra loro esclusive ed esaustive (*prima informazione estranea* al campione), la somma degli effetti di una data categoria di giudizio – inerente un attributo – sulle diverse variabili dipendenti si impone uguale a zero (vincolo di *uguaglianza*); parimenti, si impongono a somma nulla gli effetti di interazione fra categorie di giudizio. Si hanno pertanto le seguenti relazioni:

$$0 = \sum_{k=1}^3 \tilde{\delta}_{ki}^{(m)}, \quad m = 1, 2, \quad i = 2, 3; \quad 0 = \sum_{k=1}^3 \tilde{\delta}_{kih}^{(12)}, \quad i, h = 2, 3, \quad (4)$$

(l'intercetta  $\tilde{\delta}_{k1}$  non è sottoposta ad alcun vincolo) le quali, in forma compatta e con simbolismo matriciale, si rappresentano come:  $\mathbf{c} = \mathbf{R}\tilde{\boldsymbol{\delta}}$  ( $\mathbf{c}$  è un vettore colonna  $8 \times 1$  di termini nulli;  $\mathbf{R}$  è la matrice  $8 \times 27$  di variabili indicatrici, esprime la struttura dell'informazione estranea;  $\tilde{\boldsymbol{\delta}}$  è il vettore colonna  $27 \times 1$  dei parametri incogniti).

Con riferimento alla (3), la condizione di *disuguaglianza*:  $0 \leq p_{kj} \leq 1$ , da imporre ai valori attesi inerenti alla generica equazione  $k$ , si traduce nel vincolo  $0 \leq \tilde{\mathbf{z}}_{kj}' \tilde{\boldsymbol{\delta}}_k \leq 1$  (*seconda informazione estranea*), che si formalizza nel modo seguente:

$$0 \leq \tilde{\delta}_{k1} + \tilde{\delta}_{ki}^{(1)} + \tilde{\delta}_{kh}^{(2)} + \tilde{\delta}_{kih}^{(12)} \leq 1, \quad \forall i, h \in \{2, 3\}, k = 1, 2, 3. \quad (5)$$

La (5) impone, per la stima dei parametri, il ricorso al metodo dei minimi quadrati vincolati, che comporta l'uso della programmazione quadratica (PQ).

Per stimare i parametri del modello (3), l'espressione da minimizzare (funzione obiettivo  $F$ ), sotto i due ordini di vincoli (4) e (5) è, quindi, la seguente:

$$F = (\mathbf{y}^* - \mathbf{Z}^* \tilde{\boldsymbol{\delta}}^*)' (\mathbf{S}^{-1} \otimes \mathbf{I}) (\mathbf{y}^* - \mathbf{Z}^* \tilde{\boldsymbol{\delta}}^*), \quad (6)$$

dove: il simbolo  $\otimes$  indica il prodotto di Kronecker;  $\mathbf{S}^{-1} \otimes \mathbf{I}$  è la matrice *composta diagonale*  $nK \times nK$  (formata da sottomatrici diagonali  $n \times n$ , del tipo  $s^{hq} \mathbf{I}$ ,  $h, q = 1, 2, 3$ ,

dove  $s^{hq}$  è il generico elemento della matrice  $\mathbf{S}^{-1}$ , inversa della matrice delle varianze-covarianze dei residui delle regressioni *univariate*), relativa alla *prima fase* della procedura di stima dei parametri, svolta con il metodo dei *minimi quadrati ordinari*. L'utilizzazione della  $\mathbf{S}^{-1}$  – che permette di stimare facilmente i parametri del modello proposto – , in luogo della omologa matrice di secondo livello, relativa al metodo dei *minimi quadrati generalizzati* (cui si fa ricorso data l'eteroschedasticità degli errori  $e_{kj}$ ), è dovuta all'invarianza degli stimatori (e della connessa matrice dei residui) nel passaggio alla *seconda fase* di stima, essendo le variabili indipendenti dicotomiche (De Luca, 2000).

È da rilevare che nel caso in cui i [soli] vincoli di *disuguaglianza* sui parametri siano spontaneamente soddisfatti, il vettore degli stimatori  $\tilde{\mathbf{d}}^v$  del modello *vincolato* è:

$$\tilde{\mathbf{d}}^v = \tilde{\mathbf{d}}^* + \mathbf{P}\mathbf{R}'(\mathbf{R}\mathbf{P}\mathbf{R}')^{-1}(\mathbf{c} - \mathbf{R}\tilde{\mathbf{d}}^*) \quad (7)$$

( $\tilde{\mathbf{d}}^*$  è il vettore degli stimatori del modello (3) non vincolato;  $\mathbf{P} = [\tilde{\mathbf{Z}}^*(\mathbf{S} \otimes \mathbf{I})\tilde{\mathbf{Z}}^*]^{-1}$ );

i relativi errori standard sono dati dalla seguente relazione:

$$\text{var}(\tilde{\mathbf{d}}^v) = \mathbf{P} - \mathbf{P}\mathbf{R}'(\mathbf{R}\mathbf{P}\mathbf{R}')^{-1}\mathbf{R}\mathbf{P}. \quad (8)$$

Ove i vincoli di disuguaglianza non siano soddisfatti spontaneamente, gli errori standard degli stimatori del modello doppiamente vincolato – allo stato delle attuali conoscenze – restano da formulare; tuttavia, è possibile ottenere delle loro stime con il metodo bootstrap, all'interno di una procedura automatica di PQ.

#### 6.4 L'ANALISI IN COMPONENTI PRINCIPALI (ACP)

L'analisi della soddisfazione del cliente, nell'ambito della Qualità Totale, viene generalmente intesa come l'analisi di una serie di scostamenti o *gap*, che riguardano, sia l'azienda in termini di pianificazione ed offerta che il cliente in termini di attese e percezioni (Zeithaml *et al.*, 1991; D'Ambra e Lauro, 1982; D'Ambra e Amenta 2000a; Kroonenberg *et al.*, 1980, 1989). Nel seguito, ci soffermeremo sulla visualizzazione ed analisi del *gap* di valore o meglio dello scostamento tra la qualità attesa *a priori* dai clienti di un prodotto/servizio e qualità percepita *a posteriori*. Si affronta, in particolare, in un'ottica multivariata, il problema di valutare e rappresentare la soddisfazione di  $n$  consumatori rispetto a  $p$  attributi di uno scenario (servizio/prodotto). In Lauro *et al.* (1997) il problema viene affrontato con riferimento alla matrice delle differenze tra le percezioni e le attese. In questo modo si tiene conto dell'ampiezza del *gap* ma non del livello a cui si realizza. Il tema è la ricerca di una visualizzazione significativa dei dati della *Customer Satisfaction* e si affronta ricorrendo alle potenzialità della rappresentazione simbolica. Si evidenzierà come l'analisi simbolica consente il confronto, l'interpretazione e la misura di tali dati. La struttura dei dati è, dunque, costituita da una particolare matrice di dati ad intervallo, nella quale vengono registrati i giudizi dei diversi consumatori. Le analisi proposte su tale matrice e, dunque, le rappresentazioni grafiche possibili sono differenti. L'Analisi in Componenti Principali (ACP) su dati ad intervallo è suggerita con l'obiettivo di rappresentare e visualizzare i consumatori rispetto alla soddisfazione o insoddisfazione e gli attributi che soddisfano o meno i diversi consumatori. Con l'obiettivo di misurare la soddisfazione/insoddisfazione globale dei consumatori, un opportuno indice viene calcolato attraverso criteri geometrici derivati dalla rappresentazione fattoriale. Come ulteriori sviluppi di ricerca, infine, si suggerisce di utilizzare opportune analisi comparative per confrontare la soddisfazione espressa dai consumatori in tempi diversi (*panel* di consumatori) o su servizi diversi della stessa azienda.

La struttura principale dei dati è rappresentata da una matrice  $Y$  in cui sono sintetizzati i giudizi di  $n$  consumatori (in riga) su  $p$  attributi (in colonna), o dimensioni, di uno



scenario. La peculiarità di tale matrice è quella di essere costituita da dati ad intervallo. Si chiede, a ciascun consumatore  $C_i$  (per  $i=1, \dots, n$ ), di assegnare a ciascuna dimensione  $M_j$  (per  $j=1, \dots, p$ ) un punteggio (su scala ordinale) sulla qualità attesa (A) e uno su quella percepita (P). Tali punteggi debbono variare in un intervallo definito da due estremi indicanti la qualità minima (limite inferiore) e la qualità massima (limite superiore).

Il termine generico di  $Y$  è, quindi, un intervallo di valori, che denominiamo *intervallo di soddisfazione*, definito da due estremi uno indicante la qualità percepita e uno la qualità attesa dal consumatore per quella dimensione. Tale intervallo contiene, dunque, informazioni sul *gap* di soddisfazione o di insoddisfazione del consumatore rispetto a quel particolare attributo. Il *gap*, infatti, è dato dalla lunghezza del segmento che congiunge i due estremi dell'*intervallo di soddisfazione*. Dati  $n$  consumatori, essendo 2 i giudizi forniti da ciascun consumatore, le colonne della matrice  $Y$  saranno  $2p$ .

Si noti che per ciascuna dimensione  $j$ -esima, i giudizi di un consumatore possono determinare un gap pari a zero se le sue percezioni coincidono con le attese, un gap di significato positivo se le sue percezioni sono maggiori delle attese, o un gap di significato negativo se le attese sono maggiori delle percezioni. Per tener conto e rappresentare successivamente la differenza esistente tra un gap positivo e uno negativo, gli estremi dell'*intervallo di soddisfazione* debbono essere trasformati tutte le volte che un giudizio determina un gap negativo. A ciascun estremo si sostituisce il corrispondente numero negativo (es. percezioni =5 ed attese =6, diventa percezioni =-5 attese=-6). Gli intervalli relativi a gap positivi saranno rappresentati sul semiasse positivo di ciascuna dimensione, mentre quelli relativi a gap negativi su quello negativo.

Ciascun consumatore nella struttura dei dati appena proposta viene, dunque, definito attraverso un insieme di dati ad intervallo che rappresentano i suoi giudizi sulle diverse dimensioni del servizio. Si può quindi guardare a ciascun consumatore come ad un particolare *oggetto simbolico* rappresentabile con un ipercubo in  $R^p$ .

Nella terminologia dei dati simbolici, un *oggetto simbolico* (Bock e Diday; 1999) è un'unità statistica complessa definita da una congiunzione di eventi elementari. Un oggetto simbolico viene descritto da  $p$  variabili (descrittori)  $y_j$  che assumono valori all'interno di insiemi  $D_j$  ( $j = 1, \dots, p$ ):

$[\psi_l \in \delta_l] \wedge \dots \wedge [\psi_\varphi \in \delta_\varphi] \dots \wedge [\psi_\pi \in \delta_\pi]$ ; dove  $\bullet \delta_\varphi \subseteq D_j$ , ( $j = 1, \dots, p$ ) può essere, secondo la natura di  $y_j$ , un intervallo di valori o un insieme di modalità assunte dalla variabile  $y_j$ . Nel seguito, limiteremo la trattazione al caso di intervalli di valori generati da giudizi espressi su scala numerica. Oggetti simbolici descritti da insiemi di modalità possono risultare utili nella prospettiva di rilevare giudizi di natura qualitativa.

Un oggetto simbolico può essere definito *a priori* da un punto di vista *intenzionale*, indipendentemente da  $\Omega$ , insieme degli individui osservati. Il generico individuo  $\omega_i$  ( $i = 1, \dots, \text{card}(\Omega)$ ) può essere assegnato all'oggetto in base ad una funzione di riconoscimento Booleana  $a(.)$  del tipo (*vero, falso*). Si realizza  $a(\omega_i) = \text{vero}$  se  $\omega_i$  assume valori coerenti con la descrizione dell'oggetto. L'insieme degli individui che soddisfano questa condizione viene indicato come *estensione*.

La matrice di dati di partenza  $Y (n, 2p)$ , dove ciascun consumatore viene descritto da 2 valori (*vertici*) per ciascuna dimensione, può dunque essere vista come una particolare matrice di dati ad intervallo. E' evidente che le tecniche di analisi dei dati usuali non sono in grado di trattare una simile struttura dei dati, che necessita di essere trasformata.

$Y$  si può trasformare in una matrice  $Z_Y$ , di dimensioni  $(n2^p, p)$ , ottenuta considerando tutte le possibili combinazioni dei *vertici*. Nel caso più semplice di  $p=2$ , ciascun oggetto è rappresentato da  $2^2$  combinazioni dei valori dei 2 descrittori. Da un punto di vista geometrico, tale soluzione corrisponde alla rappresentazione di un rettangolo i cui vertici sono le 4 combinazioni  $\{min, min; min, MAX; MAX, min; MAX, MAX\}$ .

Si fa notare che nella struttura dei dati proposta nel lavoro, per ciascuna dimensione, i lati del rettangolo verranno rappresentati su un asse positivo (se le percezioni sono maggiori delle attese) o su un asse negativo (percezioni inferiori alle attese).

E' possibile supporre, inoltre, di avere rilevato per ciascuna dimensione una serie di informazioni esplicative sulle preferenze dei consumatori rispetto a diversi livelli dei  $p$  attributi del servizio. Si definisce, dunque, una matrice  $X (n, K)$  del piano sperimentale, dove  $K = \sum_{j=1 \dots p} k_j$  rappresenta la somma dei livelli  $k_j$  ( $j=1, \dots, p$ ) per ciascun attributo. L'elemento generico è rappresentato dalla preferenza di ciascun consumatore per uno dei livelli dei  $p$  attributi, dando il valore 1 al livello preferito e 0 a tutti gli altri.

L'analisi più immediata che può essere sviluppata per descrivere i dati della matrice  $Y$  ( $n, 2p$ ) è evidentemente un'Analisi in Componenti Principali (ACP).

L'ACP su oggetti simbolici definiti da variabili ad intervallo è stata proposta da Chouakria, *et al.* (1998) come Analisi in Componenti Principali sui Vertici (ACP-V).

L'ACP-V consiste in una normale ACP sulla matrice  $Z_Y$  ( $n \times 2p$ ) dei vertici (par.2). In questo caso, gli elementi del sottospazio di  $R^{2p}$  saranno i vertici che costituiscono gli oggetti consumatori mentre nel sottospazio di  $R^n$  è possibile rappresentare i diversi attributi. L'equazione caratteristica dell'analisi ACP-V è la seguente:

$$\frac{1}{n} Z_Y' Z_Y v_q = \lambda_q v_q \quad 1 \leq q \leq 2p$$

(1)

dove  $\lambda_q$  e  $v_q$  rappresentano rispettivamente gli autovalori e gli autovettori della matrice  $1/n Z_Y' Z_Y$ .

Lauro e Palumbo (2000) introducono una variante a tale approccio e definiscono l'ACP per Oggetti Simbolici (ACP-OS). La (1) è trasformata in modo da tener conto della coerenza tra vertici ed oggetti e da massimizzare la differenza tra gli oggetti simbolici:

$$\frac{1}{n} W Z_Y' P_{A_Y} Z_Y W v_q = \lambda_q v_q \quad 1 \leq q \leq p$$

(2)

dove  $P_{A_Y} = A_Y (A_Y' A_Y)^{-1} A_Y'$  è un operatore di proiezione ortogonale utilizzato per rappresentare i dati nel sottospazio di riferimento generato dalle colonne della matrice Booleana  $A_Y$  che indica il legame di appartenenza (coesione) dei vertici agli oggetti. Inoltre, la matrice  $W(p, p)$  è una matrice diagonale con termine generico il punteggio medio normalizzato assegnato dai diversi consumatori a ciascuna dimensione. Nel caso della struttura dei dati proposta è infatti interessante tener conto della differente importanza degli attributi per i diversi consumatori.

In maniera analoga, sulla matrice  $P_{A_Y} Z_Y Z_Y' P_{A_Y}$ , è definita l'equazione in  $R^p$ . Tale equazione ha gli stessi autovalori della (2) ma differenti autovettori  $u_q$ . Tra gli

autovettori esiste comunque la seguente relazione:  $\mathbf{p}_q = \lambda_q^{-1/2} \mathbf{Z}_Y' \mathbf{P}_{A_Y} \mathbf{u}_q$ . Le

coordinate di ciascun oggetto simbolico sull'asse principale sono date dal vettore:

$$\varphi_{i,q} = \mathbf{P}_{A_Y} \mathbf{Z}_{Y,i} \mathbf{p}_q.$$

In uno spazio bidimensionale definito da due componenti, le proiezioni dei vertici estremi determinano un rettangolo chiamato *rettangolo di massima copertura* (RMAC).

E' importante, a questo punto, fornire le regole di interpretazione dei risultati grafici dell'analisi:

- la rappresentazione in un sottospazio fattoriale di  $R^p$  a  $q < p$  dimensioni ci consente di visualizzare i diversi consumatori come differenti oggetti. Gli assi del piano fattoriale sono combinazioni lineari degli attributi e, dunque, la posizione di ciascun oggetto dipende dalla correlazione dei diversi attributi con gli assi;
- le caratteristiche dei consumatori in termini di soddisfazione/insoddisfazione possono essere lette sui piani fattoriali ricorrendo all'interpretazione tipica delle componenti principali nell'analisi simbolica. Il primo asse fattoriale, quindi, spiega la taglia della soddisfazione e distingue gli individui più che (o prevalentemente) soddisfatti (sul versante positivo) da quelli più che (o prevalentemente) insoddisfatti (sul versante negativo);
- la lettura congiunta del primo asse con quelli di ordine superiore permette di individuare diverse forme di soddisfazione in relazione agli attributi. In particolare, nei diversi quadranti troveremo individui per i quali esiste un consenso (primo e quarto quadrante) per tutti gli attributi in termini di soddisfazione (primo quadrante) o insoddisfazione (quarto quadrante), e invece, individui per cui esiste solo una prevalenza di soddisfazione (secondo quadrante) o di insoddisfazione (terzo quadrante);
- l'interpretazione dei dati viene arricchita guardando all'ampiezza e alla posizione degli oggetti sui piani fattoriali. In particolare, minore è l'area dell'oggetto, minore sarà il gap positivo o negativo. Inoltre, a parità di ampiezza, un oggetto più lontano di un altro dall'origine indica uno stesso *gap* ma ad un livello più elevato, con attese e percezioni più alte. In definitiva, per un'azienda le situazioni migliori saranno quelle definite da oggetti con area piccola e posizionati il più lontano possibile dall'origine degli assi.

Si tenga presente, che nella realtà, spesso le percezioni di un consumatore coincidono con le attese e quindi, il segmento indicante il *gap* potrebbe ridursi ad un punto.

La rappresentazione nel sottospazio di  $R^p$  va arricchita con le informazioni che si ottengono con la rappresentazione degli attributi nel sottospazio di  $R^n$ :

- le colonne di una ACP simbolica si rappresentano come vettori che, nel caso in esame, saranno orientati rispetto alla prevalenza di soddisfazione/insoddisfazione dei consumatori verso ciascun attributo. Nelle situazioni reali la maggioranza dei soggetti esprime situazioni di insoddisfazione, per cui le variabili saranno orientate verso il semiasse negativo;
- l'analisi contemporanea delle due rappresentazioni, dunque, ci consente di comprendere la forma, la dimensione e il livello di soddisfazione dei diversi consumatori sulla base della soddisfazione o meno verso ciascun attributo del servizio.

Tutte queste informazioni possono essere utilizzate al fine di costruire un indice di misura della *Customer Satisfaction*. A tale costruzione sono legate numerose problematiche relative soprattutto al sistema di ponderazione delle diverse dimensioni e alla sua robustezza rispetto a situazioni estreme.

Per tutti gli individui che giacciono sul versante positivo del primo asse, è possibile calcolare un'area (o volume se a più di due dimensioni) di *soddisfazione*. Tale area è ottenuta eliminando (con una procedura di *peeling*) dal grafico una percentuale fissa, pari ad  $\alpha$ , di oggetti anomali (con un'area più grande) e calcolando il seguente rapporto:

$$\frac{\sum_{i=1}^n a_i}{n_a d_i g^2} \quad (3)$$

dove  $a_i$  rappresenta l'area dell'oggetto  $i$ -esimo,  $n_a$  il numero di oggetti che giacciono sul versante positivo del primo asse fattoriale (individui soddisfatti) e  $g^2$  l'area di *gap* massimo, avendo indicato con  $g$  la massima differenza possibile tra attese e percezioni.

Al fine di includere l'informazione sul livello a cui si realizza la soddisfazione/insoddisfazione, si è inserito il coefficiente di ponderazione  $d_i$ , distanza del baricentro dell' $i$ -esimo oggetto dall'origine. Il peso è definito come il reciproco

della distanza per tener conto del fatto che maggiore è la distanza, maggiore è il livello di soddisfazione.

In maniera analoga si può calcolare l'area di insoddisfazione ed il relativo rapporto:

$$b_i = \sum_{i=1}^n b_i / n_b d_i g^2$$

(4)

L'indice di *Customer Satisfaction*  $ICS \in [-1, +1]$  può essere costruito nel seguente modo:

$$ICS = \frac{a/b}{a/b} \quad (5)$$

Tale indice assume valore pari a -1 se l'area di soddisfazione è nulla, pari a zero se le due aree coincidono, pari a +1 se l'area di insoddisfazione è nulla. Nella realtà, tale indice assumerà quasi sempre valore negativo, essendo maggiore l'area di insoddisfazione. L'indice ha numerosi vantaggi ed è in particolare un indice sintetico ponderato (basato su componenti principali), calcolato tenendo conto della taglia, della forma e del livello della soddisfazione/insoddisfazione di ciascun individuo. L'indice in (5) è quindi un indice che sintetizza tutte le informazioni ottenute dall'analisi fattoriale.

## 6.5 ANALISI IN COMPONENTI PRINCIPALI NON LINEARE

La P.C.A. è un metodo di trasformazione di un insieme di variabili in un nuovo insieme di variabili composite (componenti principali) ortogonali tra loro che spiegano la totalità della variabilità del fenomeno.

Le coordinate originarie  $x_1, \dots, x_n$  vengono trasformate nelle nuove coordinate  $z_1, \dots, z_n$  la cui caratteristica è l'indipendenza reciproca.

La trasformazione è  $Z=TX$  dove  $T$  è una matrice ortogonale composta da  $n$  vettori ortogonali tra loro che sono gli autovettori.

L'Analisi in Componenti Principali Non Lineare (van Rijckevorsel & de Leeuw 1988, van Rijckevorsel 1987) include informazioni a priori sulle unità statistiche (Takane & Shibayama 1991). Tale strategia può rivelarsi particolarmente utile nello studio di Customer Satisfaction (CS) in quanto l'informazione esterna può consentire di conoscere il livello di soddisfazione per gruppi omogenei (scuole, classi, sesso, titolo di studio, età, etc) di clienti/utenti, consentendo una valutazione più oggettiva ed accurata della CS.

## 6.6 REGRESSIONE LINEARE MULTIVARIATA

Per valutare come le informazioni esterne, rappresentate dai livelli degli attributi, in  $X$ , influenzano le dimensioni della soddisfazione/insoddisfazione, in  $Y$ , è opportuno far riferimento ad analisi fattoriali di tipo non simmetrico. Le preferenze per i livelli incluse nella matrice  $X$  solitamente presentano le seguenti caratteristiche:

- sono fortemente correlate in quanto legate agli attributi dello stesso scenario;
- la matrice  $X$  può presentare dei valori mancanti nel caso in cui il consumatore non esprima alcuna preferenza per i livelli di uno o più attributi (mancate risposte).

Queste caratteristiche specifiche delle variabili in  $X$  suggeriscono l'adozione della regressione PLS multivariata (PLS2, Tenenhaus 1998).

Nel primo passo della procedura, il PLS2 consiste in un'analisi inter-batteria dei fattori (Tucker, 1958). Gli elementi degli autovettori costituiscono i coefficienti delle combinazioni lineari delle variabili esplicative che rappresentano le soluzioni di primo ordine del PLS2. Per ottenere le soluzioni di ordine successivo, il PLS2 reitera il primo passo della procedura sulle matrici dei residui delle regressioni delle variabili dipendenti effettuate sulle componenti del primo ordine.

Con riferimento alle matrici  $X$  e  $Y$  definite nel paragrafo 1, analogamente all'ACP-OS, il PLS2 si applica alla matrice dei vertici coesi  $Z_Y^0 = P_{A_Y} Z_Y$  così da considerare la natura intervallare delle variabili in  $Y$  e l'appartenenza dei vertici allo stesso oggetto simbolico.

Al primo passo dell'iterazione la matrice da diagonalizzare (Höskuldsson, 1988) per la soluzione di ordine  $h$  dell'analisi in  $R^n$  (spazio degli attributi) è la seguente:

$$w_h X_{h-1}' Z_Y^0 Z_Y^0 X_{h-1} w_h \text{ con } w_h' w_h = 1$$

(6)



dove gli elementi dell'autovettore  $\mathbf{w}_h$ , associato all'autovalore di massimo modulo, rappresentano i pesi della combinazione lineare  $\mathbf{t}_h = \mathbf{X}_{h-1}\mathbf{w}_h$ , soluzione di ordine  $h$  del problema. Nel caso della soluzione di ordine  $h=1$ , la matrice  $\mathbf{X}_0$  è quella relativa alla matrice  $\mathbf{X}$  originaria, mentre per il generico  $h$  si sostituiscono le matrici dei residui:

$$\mathbf{X}_h = \left( \mathbf{I}_n - \frac{\mathbf{t}_h \mathbf{t}_h'}{\mathbf{t}_h' \mathbf{t}_h} \right) \mathbf{X}_{h-1}.$$

(7)

Si noti che la matrice dei vertici  $\mathbf{Z}_Y^0$  ha  $n2^p$  righe mentre la matrice  $\mathbf{X}$  ha  $n$  righe. Al fine di superare questo inconveniente, è necessario replicare  $n2^p$  volte ciascuna riga della matrice  $\mathbf{X}$ . In maniera duale è definita l'equazione nello spazio  $R^{p+K}$  dei consumatori.

Ai fini della rappresentazione grafica su piani fattoriali, è opportuno considerare le seguenti decomposizioni PLS delle matrici:

$$\mathbf{X} = \mathbf{t}_1 \mathbf{p}_1' + \mathbf{t}_2 \mathbf{p}_2' + \mathbf{K} + \mathbf{t}_m \mathbf{p}_m' \quad \text{e} \quad \mathbf{Z}_Y^0 = \mathbf{t}_1 \mathbf{c}_1' + \mathbf{t}_2 \mathbf{c}_2' + \mathbf{K} + \mathbf{t}_m \mathbf{c}_m'$$

(8)

dove  $\mathbf{p}_h = \mathbf{X}_{h-1}' \mathbf{t}_h / \mathbf{t}_h' \mathbf{t}_h$  e  $\mathbf{c}_h = \mathbf{Z}_{Y_{h-1}}^0 \mathbf{t}_h / \mathbf{t}_h' \mathbf{t}_h$  sono, rispettivamente, le coordinate (*factor loadings*) dei livelli degli attributi e dei giudizi dei consumatori sugli assi fattoriali. I valori dei  $\mathbf{t}_h$  rappresentano invece i punteggi (*factor scores*) dei vertici degli oggetti relativi ai consumatori sugli stessi assi fattoriali.

E' opportuno sottolineare che i pesi  $\mathbf{w}_h$  delle combinazioni lineari  $\mathbf{t}_h$  fanno riferimento alle matrici dei residui  $\mathbf{X}_{h-1}$  e pertanto sono di difficile interpretazione. E' comunque possibile trasformare questi pesi per mezzo della relazione:

$$\mathbf{w}_h^* = \mathbf{w}_h \left( \mathbf{p}_h' \mathbf{w}_h \right)^{-1}$$

(9)

così da ottenere i pesi  $\mathbf{w}_h^*$  delle variabili originarie che sono di diretta interpretazione.

Sulla base dei pesi, dei punteggi e delle coordinate, si possono costruire molteplici rappresentazioni bidimensionali che, rispetto a quelle mostrate nel paragrafo 3, si arricchiscono dell'informazione esplicativa apportata dai livelli preferiti. Le rappresentazioni più interessanti da un punto di vista informativo sono:

- rappresentazione dei vertici ( $t_h$ ) al fine di visualizzare gli oggetti relativi alla soddisfazione dei consumatori. Le regole per l'interpretazione di questa visualizzazione sono simili a quelle enunciate per l'ACP-OS. Il significato della posizione degli oggetti è però differente in quanto gli assi sono costruiti con un obiettivo esplicativo piuttosto che descrittivo. Ciascun asse, infatti, è orientato a sintetizzare la relazione di dipendenza tra la soddisfazione dei consumatori e i livelli degli attributi da questi preferiti. Pertanto, al fine di interpretare in maniera opportuna la posizione degli oggetti sui piani fattoriali, è importante definire il significato degli assi in termini delle correlazioni con gli attributi e i livelli preferiti.
- rappresentazione congiunta (*biplot*) di  $w_h^*$  e  $c_h$  per visualizzare il modello di regressione (struttura di dipendenza), infatti  $Z_Y^0 = X \sum_h w_h^* c_h' = XB$ . Questa rappresentazione permette una visualizzazione dei vettori relativi agli attributi ( $c_h$ ) che è speculare a quella della ACP-OS e, congiuntamente, una visualizzazione dei vettori relativi ai livelli ( $w_h^*$ ) degli stessi attributi preferiti dai consumatori. In particolare, interpretando l'angolo compreso tra i vettori che congiungono l'origine degli assi rispettivamente ad un livello e ad un attributo, si analizza l'influenza (positiva nel caso di un angolo acuto e negativa nel caso di un angolo ottuso) dei livelli preferiti dai consumatori sul loro grado di soddisfazione per un attributo.

## 6.7 ANALISI CORRELAZIONI CANONICHE

Nelle reali applicazioni aziendali è possibile che i giudizi espressi dai consumatori si presentino sotto forma di tabelle multiple legate ad osservazioni ripetute (Balbi e Esposito, 2000; Van den Wollenberg, 1977). Si pensi, ad esempio, ad un *panel* di clienti nel tempo (esempio *panel* dell'Auditel per gli ascolti televisivi) di cui si voglia confrontare i giudizi espressi in tempi differenti, oppure ad uno stesso gruppo di clienti che esprime dei giudizi su diversi servizi forniti da un'azienda. Queste situazioni generano delle tabelle denominate totalmente appaiate in quanto si riferiscono agli stessi individui (consumatori) ed alle stesse variabili (dimensioni o attributi del servizio) osservate in occasioni/condizioni differenti. L'obiettivo dell'analisi è quello di confrontare le diverse occasioni rispetto ad una che, per motivi interni all'azienda o di mercato, è definita come obiettivo. Nel caso delle strutture di dati ad intervallo introdotte in questo lavoro, ciascun consumatore (oggetto simbolico) è in realtà una configurazione di vertici che può essere direttamente confrontata con le configurazioni di vertici generate dallo stesso consumatore in tempi diversi o su servizi diversi. Per il confronto di oggetti simbolici rispetto ad un oggetto obiettivo è possibile far riferimento ad un'analisi comparativa, di natura essenzialmente geometrica, basata sulle rotazioni di Procuste.

Nel nostro caso, lavorando con oggetti simbolici, le configurazioni da confrontare sono quelle dei vertici coesi. Pertanto, date le configurazioni  $Z_{Y_1}^0$  (configurazione *obiettivo*)

e  $Z_{Y_2}^0$ , centrate per convenienza, la migliore trasformazione ortogonale di  $Z_{Y_2}^0$  verso

$Z_{Y_1}^0$  è  $Z_{Y_2}^0 R$ , con  $R$  operatore di rotazione ortogonale ( $R'R=I_p$ ) definito come:

$$R = Z_{Y_2}^0 Z_{Y_1}^0 (Z_{Y_1}^0 Z_{Y_2}^0 Z_{Y_2}^0 Z_{Y_1}^0)^{-1/2}.$$

(10)

L'analisi più immediata che può essere sviluppata per sintetizzare le relazioni tra le due matrici  $Z_{Y_1}^0$  e  $Z_{Y_2}^0 R$  è l'Analisi delle Correlazioni Canoniche (ACC). Questa analisi (Balbi & Esposito, 2000), quando applicata su matrici precedentemente trasformate per

mezzo di una rotazione di Procuste (Lafosse, 1989), conduce a variabili canoniche associate ad un unico sistema di assi e pertanto di più facile interpretazione rispetto ad un'ACC classica. Se la modellizzazione lineare, imposta a  $Z_{Y_2}^0$  dalla matrice  $R$ , conduce ad una sovrapposizione completa dei vertici, l'adattamento tra le due configurazioni è perfetto. Eventuali distanze residue riscontrate tra i vertici relativi alle diverse occasioni di osservazioni vanno interpretate come situazioni atipiche che costituiscono elementi di sostanziale differenza tra le valutazioni che i consumatori esprimono nei diversi tempi o sui diversi servizi. Queste differenze possono essere globali se esistono delle distanze residue per tutti i vertici o parziali se alcuni vertici coincidono. Rispetto all'analisi sviluppata in precedenza, i piani fattoriali permettono di interpretare tali distanze in termini di differenze tra le soddisfazioni espresse nelle diverse occasioni di osservazione con riferimento alla posizione degli oggetti piuttosto che al *gap*.

Tale tecnica permette di far fronte ai problemi relativi alla distribuzione non gaussiana ma asimmetrica dei dati, al gran numero di parametri in ballo e alle loro reciproche influenze, nonché all'elevato numero di indagini necessarie con gli altri metodi, pur avendo alta affidabilità.

## 6.8 ALBERI DECISIONALI E ANALISI *CHAID*

Gli alberi decisionali realizzano un procedimento di ripartizione della popolazione analizzata in gruppi mediante una successione di spaccature di tipo gerarchico, al fine di selezionare, tra le variabili esplicative, quelle che maggiormente spiegano la variabilità della variabile target.

Le partizioni prodotte sono di tipo gerarchico, e pertanto rappresentabile mediante un diagramma ad albero (dendrogramma), i cui nodi rappresentano gruppi di unità ai diversi stadi del livello di segmentazione, i rami rappresentano le condizioni che hanno determinato le suddivisioni, e le foglie sono i nodi terminali per i quali non è ritenuta utile un'ulteriore suddivisione.

La partizione prodotta dall'albero è denominata in ragione del numero di sottoinsiemi che possono essere formati nel suddividere un gruppo di unità: se si considerano solo partizioni a due vie, si ha la segmentazione binaria; se si considerano partizioni ad un numero qualsiasi di vie si parla di segmentazione multipla a  $k$  vie.

Le tecniche di costruzione degli alberi decisionali permettono di analizzare dati sia di tipo quantitativo che qualitativo (alberi di classificazione) e sono caratterizzate dalla facilità di interpretazione dei risultati (Del Ciello et al., 2000).

Partendo dall'algoritmo AID (Automatic Interaction Detection) in grado di creare attraverso bisezioni, dei sottoinsiemi dell'insieme di partenza si sono sviluppati algoritmi più sofisticati come il metodo *CHAID*.

La tecnica *CHAID* viene tipicamente utilizzata nelle analisi di segmentazione, quando l'obiettivo finale è di ripartire una popolazione (campione) in segmenti molto differenziati rispetto a un criterio di riferimento.

Più specificamente, il metodo *CHAID* si basa su una procedura iterativa. Nella prima fase, la popolazione viene divisa in due o più gruppi sulla base della variabile esplicativa più significativamente legata alla variabile dipendente prescelta; ciascuno di questi gruppi ottenuti nella prima fase viene ulteriormente diviso in sottogruppi più piccoli sulla base delle rimanenti variabili esplicative. Il processo di divisione termina

quando non possono più essere individuati predittori statisticamente significativi; i sottogruppi finali (segmenti) vengono quindi rappresentati in un diagramma ad albero.

I segmenti individuati sono mutuamente esclusivi ed esaustivi. In altre parole, i segmenti non si sovrappongono e ciascuna unità della popolazione è contenuta in un unico segmento. Dal momento che ogni segmento è definito da combinazioni delle variabili esplicative, è possibile riclassificare ogni osservazione (azienda) nel segmento appropriato semplicemente conoscendo tali variabili; per questo motivo il metodo *CHAID* viene talvolta utilizzato in alternativa all'analisi discriminante e al metodo *LOGIT*.

## 6.9 STIMATORI

L'attività di modellazione prevede nelle sue fasi la stima.

La stima dei parametri del modello ha lo scopo di assegnare specifici valori ai parametri (sconosciuti) del problema di interesse. Evidentemente, la disponibilità di stime permette di quantificare la relazione di causalità fra le variabili esplicative (nel caso più semplice, una sola  $x$ ) e la variabile dipendente (definita con  $y$ ). Un metodo largamente utilizzato per la stima del modello parametrico è quello dei minimi quadrati ordinari (OLS), che attribuisce ai parametri della relazione quei valori che minimizzano il quadrato delle distanze fra le osservazioni disponibili e la corrispondente retta di regressione; tali distanze sono anche dette residui. Dall'imposizione delle condizioni (necessarie) per il minimo si ottiene il sistema delle equazioni normali, dalla cui soluzione si ottengono i valori stimati. E' importante notare che, nel caso semplificato di una sola esplicativa, le stime OLS sono ottenute dall'imposizione di due condizioni (vincoli):

- (i) somma dei residui pari a zero
- (ii) ortogonalità fra i residui e la variabile esplicativa.

L'ANOVA, ovvero analisi della varianza, è una tecnica statistica che verifica le differenze fra le medie di una variabile dipendente in gruppi diversi, identificati dai valori delle variabili esplicative. Per esempio ANOVA può essere usata per verificare se esiste una differenza significativa fra le vendite realizzate in aree geografiche diverse.

Si possono inoltre specificare interazioni, ovvero effetti combinati di più variabili esplicative.

## CAPITOLO 7-

### **MODELLI DI STATISTICHE MULTIVARIATE CON TECNICHE DI AGGREGAZIONE**

#### 7.1 MULTICRITERIA SATISFACTION ANALYSIS (MUSA): DATA MINING

Per rinforzare l' orientamento del cliente su una base giornaliera, un numero crescente di aziende sceglie la soddisfazione di cliente come loro indicatore principale di prestazione. Tuttavia, è quasi impossibile mantenere un' intera azienda permanentemente motivata da una nozione così astratta e intangibile quale la soddisfazione di cliente. Di conseguenza, la customer satisfaction deve essere tradotta in un certo numero di parametri misurabili, tali da determinare gli elementi che influenzano i consumatori. (Deschamps e Nayak, 1995).

Lo scopo di questo paragrafo è di presentare un approccio metodologico originale al problema della valutazione della soddisfazione del cliente, unendo le tecniche di data-mining per induzione e le tecniche di analisi multicriteriale di disaggregazione di preferenza. Gli obiettivi principali sono: 1) confrontare i risultati dei due metodi, 2) valutare l' omogeneità dell' insieme dei clienti, 3) superare il problema di ~~nessu~~ risposta (mancanza di dati) nell' insieme di dati.

La struttura metodologica prevede l'uso del metodo MUSA (ANALISI di SODDISFAZIONE MULTICRITERIO – Multicriteria Satisfaction Analysis), che è basato su un modello di disaggregazione di preferenza. L' ~~g~~gregazione di diverse preferenze in una funzione è l' obiettivo principale di questo metodo. Più specificamente, è presupposto che la soddisfazione globale dei clienti possa essere spiegata da un insieme di criteri o dalle variabili che rappresentano le relative dimensioni.

La metodologia di disaggregazione di preferenza è un metodo basato sulla regressione ordinale (Lagrèze e Siskos, 1982; Siskos e Yannacopoulos, 1985) nel campo di analisi multicriteriale: il metodo UTA. È usata per la valutazione di un insieme delle funzioni marginali di soddisfazione in modo tale che il criterio globale di soddisfazione diventi consistente il più possibile ai giudizi dei clienti. Secondo il modello, ad ogni cliente è chiesto di esprimere i suoi giudizi, vale a dire la sua soddisfazione globale e la sua



soddisfazione riguardo all' insieme dei criteri discreti. I dati raccolti sono analizzati con il modello di disaggregazione di preferenza, rispettando la forma ordinale e qualitativa dei giudizi e delle preferenze dei clienti. I risultati principali del metodo sono (Grigoroudis ed altri, 1998; Siskos ed altri, 1998; Mihelis ed altri, 1998): 1) la determinazione della soddisfazione globale e parziale, 2) la determinazione dei pesi sui criteri (importanza relativa), 3) la determinazione degli indici medi della soddisfazione.

Seguendo i principi di analisi di regressione lineare vincolata ed usando tecniche di programmazione lineare, la funzione che rappresenta il modello è la seguente:

$$\begin{cases} e(Y^*) = \sum_{i=1}^n b_i X_i^* - \sigma^+ + \sigma^- \\ \sum_{i=1}^n b_i = 1 \end{cases}$$

dove  $e(Y^*)$  è la stima della funzione di soddisfazione globale  $Y^*$ ,  $X_i^*$  sono le funzioni di soddisfazione parziale e  $\sigma^+$  e  $\sigma^-$  sono gli errori sottostimati e le soprastimati, rispettivamente, e  $b_i$  è il peso del criterio  $i$ .

Al fine di ridurre il numero dei vincoli è possibile ricorrere alla seguente trasformazione:

$$\begin{cases} z_m = y^{m+1} - y^m \\ w_{ik} = b_i x_i^{*k+1} - b_i x_i^{*k} \end{cases}$$

con  $m=1, \dots, \alpha-1$ , con  $k=1, \dots, \alpha_i-1$ , con  $m=1, \dots, n$  pertanto il modello può essere

riscritto in forma di P.L., come segue:

$$\begin{cases} \min F = \sum_{j=1}^M \sigma_j^+ + \sigma_j^- \\ \text{sottovincoli} \\ \sum_{i=1}^n \sum_{k=1}^{\alpha_i-1} w_{ik} \sum_{m=1}^{y^i-1} z_m - \sigma_j^+ + \sigma_j^- = 0 \\ \sum_{i=1}^n \sum_{k=1}^{\alpha_i-1} w_{ik} = 100 \\ z_m \geq 0, w_{ik} \geq 0 \forall m, i, k \\ \sigma_j^+ \geq 0, \sigma_j^- \geq 0 \text{ per } j=1, 2, \dots, M \\ \sum_{m=1}^{\alpha-1} z_m = 100 \end{cases}$$

dove  $M$  sono il numero dei clienti,  $n$  il numero dei

criteri, e  $x$  e  $y$  i vari  $j$ -livelli delle variabili  $X_i$  e  $Y$ .

La metodologia di disaggregazione consiste altresì nell'analisi post-ottimale della stabilità del modello. La soluzione finale è ottenuta esplorando il poliedro delle soluzioni quasi-ottimali, che è generato dai vincoli del suddetto programma lineare. Questa soluzione è calcolata attraverso  $n$  programmi lineari (quanti il numero dei

criteri) nella seguente forma:

$$\left\{ \begin{array}{l} \max F^i = \sum_{k=1}^{\alpha_i-1} w_{ik} \\ per\ i = 1, 2, \dots, n \\ sottovincoli \\ F \leq F^* + \varepsilon \\ tutti\ i\ vincoli\ precedenti \end{array} \right.$$

Dove  $\varepsilon$  è una piccola percentuale di  $F^*$ .

Il data-mining è il processo di estrazione di conoscenza da banche dati di grandi dimensioni tramite l' applicazione di algoritmi che individuano le associazioni nascoste tra le informazioni e le rendono visibili. In altre parole, col nome *data mining* si intende l' applicazione di una o più tecniche che consentono l' esplorazione di grandi quantità di dati, con l' obiettivo di individuare le informazioni più significative e di renderle disponibili e direttamente utilizzabili nell' ambito del decision making. L' estrazione di conoscenza (informazioni significative) avviene tramite individuazione delle associazioni, o "patterns", o sequenze ripetute, o regolarità, nascoste nei dati. In questo contesto un "pattern" indica una struttura, un modello, o, in generale, una rappresentazione sintetica dei dati. Il termine *data mining* è utilizzato come sinonimo di *knowledge discovery in databases* (KDD) (Ostillio, 2002), anche se sarebbe più preciso parlare di *knowledge discovery* quando ci si riferisce al processo di estrazione della conoscenza, e di *data mining* come di una particolare fase del suddetto processo (la fase di applicazione di uno specifico algoritmo per l' individuazione dei "patterns"). L' obiettivo di datamining è estrarre le informazioni importanti dai dati, per scoprire le componenti nascoste al loro interno. Nel ambito del Decision Support Management, il data-mining può essere definito come processo dell' ausilio decisionale in cui si fa una ricerca di modelli sulle informazioni nei dati. (Parsaye, 1997).

Le Tecniche di data-mining sono basate sulle regole di induzione dei dati e su tecniche di distillazione di dati. Queste tecniche estraggono i modelli dall'insieme di dati e li usano per vari scopi, quale la previsione del valore di un campo dipendente (campo da predire). Automaticamente esplorando l'insieme di dati, il sistema di induzione forma la soddisfazione globale. Questi modelli possono essere di tipo logico. La logica può occuparsi sia dei dati numerici che non numerici. L'operatore centrale in logica è solitamente una variazione sulla dichiarazione "if-then". Tali regole collegano un risultato di interesse ad un certo numero di attributi. Sono di seguente forma (Akeel, 1994): se attributo1 = a ed se attributo2 = b allora il risultato = c (con probabilità = 0,9). La probabilità di una regola è la probabilità che un record casuale soddisfi il "rule's condition(s)", e che la conclusione compiuta (Meidan, 1999) è soddisfacente. Le regole possono andare facilmente oltre le rappresentazioni del valore di un attributo. Qui, nella logica di attributo, confrontiamo i valori dei due campi, senza determinare i valori. Le regole presentano il vantaggio di potere essere trattate sia con dati numerici che con dati non numerici (campi categorici).

La metodologia MUSA, combina il modello di disaggregazione di preferenza con il processo di rule-induction.

Le fasi principali della metodologia sono:

- 1) Analisi preliminare del processo di indagine di soddisfazione di cliente: gli obiettivi di ricerca di soddisfazione di cliente dovrebbero essere specificati in questa fase, per valutare le dimensioni di soddisfazione (consistenza dei criteri).
- 2) Determinazione dei questionari e dell'indagine di condotta: deriva dal punto precedente, questa fase si riferisce allo sviluppo del questionario, della determinazione dei parametri di indagine.
- 3) Analisi: i due differenti approcci determinano la previsione.

Nel caso la previsione non è considerata soddisfacente, una nuova selezione delle serie di dati è fatta ed il processo riavvia l'analisi. Nel caso opposto (della previsione soddisfacente), il valore previsto è usato per riempire le caselle vuote nella tabella dei dati. Le cellule vuote corrispondono ai casi di nessuna risposta. L'insieme di dati riempito

derivante è usato con il metodo di disaggregazione di preferenza per effettuare l' analisi finale.

La metodologia originale, pertanto, combina la metodologia di disaggregazione di preferenza con data-mining alla regola-induzione. La metodologia è proposta come soluzione potenziale al problema di nessuna risposta nell' insieme di dati che può essere dovuto ai questionari insufficientemente compilati. Il metodo della MUSA valuta la curva di incremento di soddisfazione riguardo ai giudizi dei clienti. Questa curva normalizzata tra [ 0, 100 ] mostra il valore ricevuto dai clienti per ogni livello della scala qualitativa ordinale di soddisfazione.

## 7.2 SISTEMI DI DISAGGREGAZIONE INTERATTIVA

Scopo di questo paragrafo è quello di presentare metodi che valutano un insieme degli oggetti su una gerarchia di criteri qualitativi. La valutazione dell' oggetto su un criterio riguardo alle relative valutazioni sui sub-criteri è formulata come problema di classificazione di tipo multicriteriale.

La struttura di tipo gerarchico è stata dimostrata essere di grande aiuto all' analista ed ai decisori che strutturano problemi complessi di decisione (Keeney 1992, Rasmussen 1985). Un vantaggio della struttura gerarchica degli obiettivi è che è una struttura semplice, facile da usare per il decisore. Tale struttura gerarchica allora è usata comunemente per definire un insieme di criteri che è presupposto per essere operativo, coerente ed esauriente (Roy, 1996 Keeney e Raiffa, 1976, Edwards e Von Winterfeldt, 1986). I criteri sono di natura qualitativa e normalmente non preesistono. Questa operazione è difficile da trattare nelle situazioni pratiche di decisione non soltanto perché è problematica da comunicare in merito alle conseguenze usando un attributo costruito, ma anche perché la sua determinazione è inesatta ed incerta.

Saaty (1980) ha usato il confronto a coppie fra le alternative per valutare le loro priorità sui criteri e per valutare la loro importanza relativa.

In questo paragrafo proponiamo di parlare dei criteri ordinali in forma gerarchica. Poiché possiamo ammettere che la valutazione ordinale è un caso della valutazione assoluta e la valutazione su un nodo è un' aggregazione della valutazione sui sub-nodi, possiamo usare i modelli di classificazione per valutare le alternative sul nodo che considera le loro valutazioni sui sub-nodi. Proponiamo, quindi, i modelli di valutazione tramite un metodo di aggregazione-disaggregazione.

Il metodo propone, formalmente, una gerarchia di criteri che è definita da un albero in cui ogni nodo rappresenta un criterio a cui si è associata una scala di valutazione. In una tale struttura,  $C_{i1}, C_{i2}, \dots, C_{in}$  sono "sub-criteri" del criterio  $C_i$ . I livelli della scala ordinale su ogni criterio sono definiti sia linguisticamente che dalle alternative tipiche descritte dalle loro valutazioni sui sub-criteri. Il problema è quello di valutare

un' alternativa sul criterio  $C$  secondo la relativa valutazione su  $C_{i1}, C_{i2}, \dots, C_{in}$  ed è formulato attraverso un modello di assegnazione in cui le categorie rappresentano i livelli della scala ordinale.

Gli attributi determinati o i criteri qualitativi considerano, in generale, più di una dimensione della preferenza. Una tecnica molto utile per occuparsi della bontà della costruzione dei criteri qualitativi consiste nel : 1) suddividerli in attributi più specifici, 2) costruire funzioni per ogni attributo secondario, 3) aggregare le funzioni di attributo secondario in una funzione generale.

Quindi possiamo considerare questo problema di costruzione dei criteri qualitativi come un problema multiplo di classificazione di criteri. dove l' insieme dei criteri e l' insieme dei subcriteri sono i livelli ordinati su scala qualitativa. Queste considerazioni ci risulteranno utili nell'applicazione.

Diversi sono i metodi che possono essere usati per costruire il criterio qualitativo di valutazione sui relativi sub-criteri, si citano ad esempio ELECTRE TRI (Mousseau e Slowinski, 1998), UTADIS, ORCLASS, ecc.

Usare un metodo di aggregazione-disaggregazione per valutare i parametri di modello sui nodi qualitativi con procedura gerarchica significa affrontare un problema della classificazione. Si può scegliere tra: 1. scegliere una procedura di aggregazione che costruisca il relativo parametro in un metodo decomposto, 2. usare un metodo di aggregazione-disaggregazione per determinare i modelli di classificazione da un insieme delle alternative valutate. Le scale ordinali che costruiamo sono ottenute in primo luogo dalla valutazione di un insieme di riferimento degli oggetti. Questa valutazione permette di ottenere delle scale frammentate che sono determinate attraverso una procedura di interazione usando alternative fittizie costruite in base alle informazioni ordinali predeterminate sull' importanza relativa dei criterio.

Può essere utile, allora usare i metodi di outranking. Questi metodi usano le soglie di veto, di indifferenza e di preferenza. Per i modelli cumulativi, le funzioni di valore di criterio devono essere valutate.

Abbiamo già parlato del metodo UTA utilizzato che Lagreze et Siskos (1982), cioè di un metodo di aggregazione-disaggregazione per determinare i parametri del modello da un insieme degli oggetti valutati.

Esistono altre formulazioni, quali il metodo UTADIS che permettono di stabilire le funzioni lineari di valore da un insieme di riferimento attraverso gli esempi assegnati. Questo metodo è interattivo.

Il metodo UTADIS (Utilities Additives DIScriminantes) è una variante interattiva del metodo di UTA, basata sul metodo di disaggregazione di preferenza valutata su funzioni di utilità additive e su profili usando tecniche di programmazione lineari per minimizzare l'errore di errata classificazione fra i valori predefiniti. Inoltre, permette di ottenere i parametri che corrispondono all'interpretazione nel modello usato. Questo metodo di aggregazione-disaggregazione può essere determinato anche attraverso il metodo TRI di ELECTRE che punta sulla determinazione del modello dagli esempi di assegnazione.

Ci sono molti fattori da considerare per scegliere un modello particolare da applicarsi per la procedura gerarchica. Il confronto dei modelli differenti dovrebbe essere basato sull'accettabilità dei loro principi qualitativi che, all'essenza, sono collegati con il grado di compensazione, sull'accettabilità di indipendenza di preferenza, e sulla possibilità di considerare l'incertezza, l'arbitrarietà e la determinazione inesatta. Il confronto può anche essere basato sulla loro capacità di ristabilire l'insieme degli oggetti di riferimento alle loro valutazioni iniziali, di rilevare l'incoerenza potenziale nell'insieme di riferimento delle alternative valutate e di aiutare i decisori che valutano questa incoerenza, sulla loro capacità di generare le regole che coprono tutti i casi possibili, ed infine, sulla relativa facilità di uso e convenienza al contesto del problema di decisione.

### 7.3 LA CLUSTER ANALYSIS

Tra i metodi di classificazione usati nell'analisi di Customer Satisfaction sono da considerare i metodi di *cluster analysis* o *clustering*. Essi hanno lo scopo di classificare le unità statistiche attraverso l'uso di procedure che, di solito, sono applicabili quando su ogni unità statistica sono state rilevate le modalità di M caratteri. Tali metodi si sono sviluppati fin dalla fine del XIX secolo e si valuta che gli algoritmi che sono stati elaborati fino ad oggi siano circa un migliaio. I motivi principali di tanto interesse per questo tipo di algoritmi sono essenzialmente due: 1) le tecniche di analisi dei gruppi sono largamente usate nei più svariati campi di ricerca (fisica, scienze sociali, economia, medicina, ecc.), in cui la classificazione dei dati disponibili è un momento essenziale nella ricerca di modelli interpretativi della realtà; 2) l'evoluzione degli strumenti di calcolo automatico ha consentito di affrontare senza difficoltà la complessità computazionale che è insita in molti dei metodi di classificazione e che in precedenza aveva spinto i ricercatori ad orientarsi verso quelle tecniche di analisi dei gruppi che erano più facilmente applicabili. Si è resa così possibile la produzione di diversi algoritmi di classificazione, sempre più complessi dal punto di vista computazionale, ma anche sempre più efficienti nel trarre informazioni dai dati attraverso una loro opportuna classificazione.

Gli autori non sono concordi nel definire un processo di clustering: secondo Sokal, consiste nel ripartire un insieme di unità elementari in modo che la suddivisione risultante goda di alcune proprietà considerate desiderabili; per altri studiosi classificare delle unità statistiche significa formare dei gruppi di unità in modo che le unità che sono assegnate allo stesso gruppo siano simili tra loro e che i gruppi siano il più possibile distinti tra loro (Gordon, 1988; Calinski et al., 1974). Indipendentemente dalla definizione, in generale un metodo di classificazione è caratterizzato da due fattori:

- a) una misura del grado di diversità tra le coppie di unità;
- b) un algoritmo con cui procedere alla ricerca dei cluster.



Modificando uno o l'altro di questi fattori si possono produrre una gran quantità di metodi diversi dei quali sono state proposte diverse classificazioni alcune basate sul tipo di algoritmo adottato dal metodo, altre basate sul tipo di risultato da esso fornito.

La più diffusa è quella, basata sul tipo di algoritmo, che distingue tra metodi *gerarchici* e metodi *non gerarchici*.

I primi sono metodi che producono raggruppamenti successivi ordinabili secondo livelli crescenti o decrescenti della distanza (o, viceversa, della similarità). Si tratta di procedura iterative che considerano tutti i livelli di distanza e i gruppi che si ottengono ad un certo livello di distanza sono contenuti nei gruppi ottenuti ad un livello di distanza inferiore. I metodi gerarchici si possono ulteriormente dividere distinguendo tra metodi *agglomerativi* e *scissori*. Sono *agglomerative* quelle tecniche che, partendo da  $n$  elementi distinti, producono di volta in volta un numero decrescente di clusters di ampiezza crescente, fino ad associare in un unico gruppo tutte le  $n$  unità di partenza. Viceversa, i metodi *scissori* ripartiscono gli stessi  $n$  elementi, inizialmente compresi in un unico insieme, in gruppi sempre più piccoli e numerosi, finché il numero di clusters viene a coincidere con il numero delle unità. Tra i due approcci, quello agglomerativo è stato sicuramente privilegiato: queste tecniche sono infatti più semplici da programmare e, come è stato osservato, comportano un minor rischio di pervenire a suddivisioni delle unità che non rispecchiano l'effettiva struttura dei dati, al contrario, i metodi scissori possono più facilmente realizzare allocazioni sbagliate delle unità, che però non vengono corrette se in seguito non sono previste particolari procedure di aggiustamento.

Quando l'algoritmo produce un'unica suddivisione dell'insieme di partenza, considerata ottimale rispetto al criterio adottato, la classificazione risultante è *non gerarchica*. Appartengono a questa categoria tutte le classificazioni prodotte da un metodo di programmazione matematica o quelle che, tentando di migliorare una suddivisione provvisoria delle unità, effettuano una serie di riallocazioni finché non risulta soddisfatto un dato criterio di ottimalità.

I metodi non gerarchici dipendono in generale da due fattori:

a) presenza o assenza di centri;

b) esistenza o meno di una funzione obiettivo.

Queste suddivisioni in realtà non comprendono tutti i vari tipi di metodi, ma riescono comunque a classificare quelli più usati. Esistono poi infinite versioni di uno stesso metodo, quando, pur applicando una stessa procedura di clustering, vengono utilizzate differenti distanze. Ciò spiega come mai i pochi metodi proposti in principio si siano moltiplicati fino a costituire un campo molto vasto e complesso, a cui i diversi schemi logici di unificazione hanno tentato di dare un ordine.

Senza voler approfondire tale approccio, estremamente rigoroso ma scomodo a causa della complicata terminologia di cui si avvale, possiamo più semplicemente fare una prima distinzione tra gli algoritmi *esatti* e quelli *euristici*. Sono esatte quelle procedure che determinano una suddivisione delle  $n$  unità in  $c$  clusters, la quale risulta ottima rispetto alla misura di omogeneità dei gruppi, o a quella di similarità delle unità, ossia genera la migliore tra tutte le possibili partizioni di  $n$  elementi in  $c$  clusters. Gli algoritmi euristici, o non esatti, danno luogo ad una suddivisione buona o approssimativamente ottima, ma che tuttavia si discosterà in qualche misura dall'essere la migliore possibile.

Si comprende come le tecniche più diffuse appartengano a questa seconda categoria: esse sono infatti computazionalmente più efficienti di quelle esatte, le quali per esaminare tutte le possibili partizioni necessitano spesso di un numero di operazioni elementari che cresce in maniera esponenziale con  $n$ .

Oltre alle differenziazioni basate sul tipo di algoritmo, i vari metodi si possono distinguere anche in base alla classificazione che essi producono. I risultati di una classificazione si possono rappresentare attraverso una matrice con tante righe quante sono le unità e tante colonne quanti sono i gruppi: se abbiamo  $n$  unità e  $G$  gruppi la matrice è di dimensione  $(n \times G)$  e contiene i valori di una **funzione di appartenenza**. Tale funzione, indicata con  $\mu_{ig}$ , è una funzione a  $G$  valori (dove  $G$  è il numero di gruppi della partizione o del ricoprimento) che associa ad ogni unità  $G$  numeri ognuno dei quali esprime il grado di appartenenza dell'unità  $i$ -esima al  $j$ -esimo gruppo (con  $i=1,2,\dots,n$  e  $g=1,2,\dots,G$ ). L'intervallo di definizione di tale funzione permette di distinguere tra metodi di classici e metodi non classici detti anche *sforcati*: per i metodi

classici la funzione di appartenenza è definita nell'insieme  $\{0,1\}$ , cioè assume solo i due valori 1 e 0, che indicano, rispettivamente, se una unità appartiene o non all'insieme; per i metodi sfocati l'insieme di definizione è l'intervallo  $[0,1]$  e quindi la funzione di appartenenza esprime il *grado* con cui una unità appartiene ad un gruppo. Se si distinguono inoltre i raggruppamenti, ottenuti in base ai vari metodi, in *partizioni*, che sono i raggruppamenti con la caratteristica di rispettare il vincolo

$$S = \sum_{g=1}^G \mu_{ig} = 1$$

e *ricoprimenti*, che sono i raggruppamenti con la caratteristica di rispettare il vincolo

$$S = \sum_{g=1}^G \mu_{ig} \geq 1$$

si ottiene la suddivisione di cui alla seguente tabella (Ricolfi, 1992):

tabella	METODI CLASSICI	METODI SFOCATI
	GRADO DI APPARTENENZA (0,1)	[0,1]
PARTIZIONI	$S = \sum_{g=1}^G \mu_{ig} = 1$ Classificazione classica	Classificazione sfocata
RICOPRIMENTI	$S = \sum_{g=1}^G \mu_{ig} \geq 1$ Classificazione sovrapposta	Classificazione sovrapposta sfocata

Da questo schema si evince che con metodi di *classificazione classica*, si intendono tutti quei metodi che forniscono una partizione classica (cioè una suddivisione delle unità in gruppi tra loro disgiunti e tali che la loro unione fornisca l'insieme di tutte le unità). Con i metodi di *classificazione sovrapposta* si indicano i metodi che forniscono una suddivisione delle unità in gruppi non disgiunti, cioè tali che una medesima unità possa appartenere a più di un gruppo (ricoprimento classico dell'insieme delle unità). Con metodi di *classificazione sfocata* indicheremo quei metodi che suddividono l'insieme delle unità in modo che una unità può appartenere solo in parte ad un gruppo e

quindi per la parte rimanente appartiene ad altri gruppi. Infine si può pensare ad un'ultima categoria di metodi, che risulta dall'unione di queste due ultime, la quale fornisca dei *ricoprimenti sfocati*, cioè dei gruppi sfocati sovrapposti. A questa categoria potremmo dare nome di metodi di *classificazione sovrapposta sfocata*. Naturalmente, all'interno di queste distinzioni, valgono ancora quelle fatte in precedenza tra metodi gerarchici e non gerarchici per cui, ad esempio, esistono tecniche gerarchiche e tecniche non gerarchiche di classificazione classica, sovrapposta, sfocata e sovrapposta sfocata.

Le tecniche di classificazione di solito utilizzate sono quelle che forniscono partizioni classiche dell'insieme iniziale, mentre per gli altri tre tipi il numero di algoritmi a disposizione è piuttosto ridotto. Ci sono inoltre altri metodi (Ponsard, 1985, Fustier, 1980) che danno una iniziale sfocatura ai dati assegnando all'inizio del procedimento una funzione ad ogni unità. Tale funzione, che chiameremo *funzione caratteristica*, misura la quantità di carattere posseduta da una unità rispetto a quella posseduta dalle altre unità. Perciò tali metodi non usano i dati di partenza per classificare le unità, ma li sostituiscono con dei dati 'sfocati'. In tal modo si possono utilizzare procedure di classificazione, che sono applicabili solo a caratteri misurabili, anche se si dispone di dati di qualunque natura. Le classificazioni che ne derivano sono però delle classificazioni classiche o al più sovrapposte, in quanto, secondo le definizioni appena date, la funzione di appartenenza finale è a valori in  $\{0,1\}$ . Non bisogna quindi confondere la classificazione sfocata con le classificazioni che si ottengono a partire da dati sfocati ma che presentano una funzione di appartenenza a valori in  $\{0,1\}$  e pertanto sono metodi da assegnare alla categoria di quelli classici. In questa trattazione ci soffermeremo sulla teoria degli insiemi sfocati in altri capitoli. Molti sono gli autori che si sono occupati della logica sfocata in modo rigoroso: pensiamo a Zadeh, uno tra i promotori di questa logica, a Ruspini, a Kaufmann, a Leung, a Ponsard e Tran, (DeSarbo et al., 1991), qui solo per citarne alcuni tra i più autorevoli. Perciò faremo riferimento solo ai concetti di logica sfocata, che di volta in volta, sarà più opportuno chiarire e rimandiamo a questi autori per ulteriori approfondimenti.

In questo gruppo abbiamo compreso tutti quei metodi di analisi dei gruppi che producono una partizione classica delle unità, cioè suddividono le unità in gruppi

disgiunti e tali che la loro unione fornisca l'insieme di tutte le unità. In questo paragrafo ci limitiamo a dare qualche accenno solo sui metodi più usati.

Tra i metodi gerarchici, quelli più usati sono quelli di tipo agglomerativo. Ad esempio:

a) il **metodo del legame singolo (SLM)**, che si basa sulle distanze tra le unità: le unità che sono le une rispetto alle altre a distanza minima vengono assegnate ad un unico gruppo; si calcola poi la distanza tra questo gruppo appena formato e le rimanenti unità come la *minima distanza* tra le unità del gruppo e le altre unità. Se si sono già formati dei gruppi si calcola la distanza tra il gruppo appena formato e gli altri gruppi come la minima distanza tra le unità del gruppo appena formato e le unità degli altri gruppi. Si ripete il procedimento fino a che tutte le unità sono nello stesso gruppo.

b) il **metodo del legame completo (CLM)**, che si basa su un algoritmo del tutto simile a quello del legame singolo con la sola differenza che la distanza tra il gruppo appena formato e ognuna delle rimanenti unità (o gruppi) è calcolata come la *massima distanza* tra le unità del gruppo e le rimanenti unità.

c) il **metodo del legame medio (ALM)**, che procede come i precedenti, ma calcola la distanza tra un gruppo ed una unità come la distanza tra l'unità e una unità fittizia in cui ciascun carattere è presente con una media delle modalità presentate dalle unità comprese nel gruppo.

E ancora: metodo del centroide, metodo della mediana, metodo della varianza minima, metodo del legame flessibile e tanti altri, su cui non ci soffermiamo.

Tra i metodi non gerarchici, i più usati sono quelli detti **metodi di suddivisione iterativa** e i **metodi di programmazione matematica**. In genere, questi metodi partono da una iniziale suddivisione delle unità e procedono spostando le unità da un gruppo all'altro fino a che non si raggiunge una situazione ottimale che non consente altri spostamenti.

I metodi di suddivisione iterativa eseguono degli spostamenti effettivi delle unità: si calcolano i centroidi dei vari gruppi (oppure si scelgono dei nuclei o dei semi intorno ai quali si devono raggruppare le unità) e si assegna ogni unità al gruppo più vicino; poi si ricalcolano di nuovo i centroidi e si ripete il procedimento fino a che non si possono più spostare le unità. In alcuni casi i metodi sono dotati anche di una funzione obiettivo che

valuta la bontà di una determinata partizione in modo che si possa scegliere lo spostamento più conveniente tra quelli possibili, come nel caso del *metodo delle k-medie (HCM)* (De Soete e Carroll, 1994).

I metodi di programmazione matematica si basano invece su spostamenti virtuali delle unità, fatti secondo la soluzione di un problema minimo o massimo vincolato, e non contemplano il calcolo dei centroidi dei gruppi.

Il punto debole di questo tipo di metodi non gerarchici sta nelle scelte che devono essere compiute all'inizio: si deve fare una iniziale partizione delle unità, si devono scegliere i nuclei o semi e, inoltre, si deve scegliere la funzione obiettivo. E' chiaro che scelte iniziali diverse porteranno inevitabilmente a diverse partizioni finali e che questi procedimenti possono diventare inaffidabili a meno che non si tenti di superare l'inconveniente di dover fare delle scelte iniziali (L. Ricolfi, 1992).

Ma in generale si pone il problema della scelta del metodo da adottare, dal momento che i diversi metodi di classificazione portano in genere a soluzioni diverse.

Perciò è stato proposto da diversi autori di costruire un'unica classificazione aggregando le classificazioni che possono risultare dall'applicazione di diverse procedure ad uno stesso insieme di dati. Si tratta delle tecniche dette del *consenso* che producono una classificazione unica a partire da più classificazioni. Si tratta di tecniche che ultimamente hanno avuto nuovi sviluppi perché sono state applicate allo studio delle matrici di dati a *tre vie* (cioè formate da unità, variabili e occasioni) in quanto sono ottimi strumenti di sintesi dell'informazione (Gordon e Vichi, 1997).

Questo tipo di analisi dei dati, nato nel contesto della teoria dei grafi, rappresenta il tentativo di superare i problemi che nascono quando si presenta il caso di unità con caratteristiche intermedie a due o più gruppi (caso, in realtà, piuttosto frequente) e sarebbe opportuno assegnare l'unità ad entrambi i gruppi. In tal modo non si ottiene una partizione delle unità, ma un ricoprimento e i gruppi formati vengono chiamati di solito *clump* (Gordon, 1981) anziché cluster, e la tecnica di analisi dei dati viene detta *clumping* o classificazione sovrapposta (*overlapping clustering*). Ai metodi che generano classificazioni sovrapposte è stata dedicata un'attenzione molto minore rispetto ai metodi di classificazione classica anche se essi sono di notevole interesse

teorico. Ciò è dovuto in parte al fatto che molti autori fanno una certa confusione tra l'approccio sfocato e l'approccio sovrapposto, per cui alcuni algoritmi che si definiscono come algoritmi di classificazione sfocata, in realtà sono solo di classificazione sovrapposta. D'altra parte alcuni metodi di clumping finiscono per sovrapporre i gruppi in modo eccessivo e il ricoprimento che si ottiene diventa impossibile da interpretare.

Ma anche per questo problema sono state proposte delle soluzioni, come l'introduzione di vincoli per limitare le sovrapposizioni consentite imponendo, ad esempio, agli insiemi un numero massimo di sovrapposizioni pari a  $(k-1)$ .

Vale la pena di ricordare il *metodo delle piramidi* (Diday, 1986), che è un metodo gerarchico di classificazione sovrapposta. Scelto un indice di similarità tra clump (che è anche un indice di similarità tra unità in quanto una unità è vista come un clump formato da un solo oggetto), si aggregano ad ogni stadio i clump più vicini formando proprio una struttura a piramide che ha alla base le unità ancora non raggruppate e il cui vertice rappresenta il clump finale, cioè quello che comprende tutte le unità. Anche per questo metodo sono state proposte delle limitazioni che servono ad ottenere un numero ridotto di sovrapposizioni: l'unione di due gruppi non è sempre consentita (se  $i$  e  $j$  sono gli elementi estremi di un gruppo  $g$ , cioè gli elementi che sono alla base del clump e delimitano il clump stesso, non si può unire a  $g$  nessun gruppo che non contenga sia  $i$  che  $j$ ).

L'interesse teorico di questi metodi è dovuto al fatto che essi trattano bene l'imprecisione: le unità statistiche non sono sempre classificabili con esattezza perché non è raro il caso di unità che possono essere assegnate indifferentemente a più gruppi. Questi metodi assegnano ogni singola unità in parte a ciascun gruppo in modo che la classificazione che risulta non solo mostri come aggregano le unità, ma riesca anche a mostrare quanto una unità appartiene ad un gruppo. In tal modo l'assegnazione di una unità ad un gruppo non è mai una forzatura mentre la non assegnazione di una unità ad un gruppo indica con certezza che quella unità non appartiene a quel gruppo.

I metodi di classificazione sfocata non hanno quindi la pretesa di dare risposte precise su come si aggregano i dati, cosa che si può fare più agevolmente con un metodo di

analisi classica, ma, al contrario, tentano di rappresentare proprio l'imprecisione insita nei dati. Perché, come dice Ponsard (1985), probabilmente solo un modello *impreciso* della realtà è in grado di rappresentarla più esattamente di quanto possa fare un modello *preciso*.

I metodi di classificazione sfocata, come abbiamo detto, producono delle partizioni o dei ricoprimenti dell'insieme dei dati, ma la funzione di appartenenza di ogni unità ai vari gruppi non assume, come nei metodi di classificazione classica, solo i valori 0 e 1, ma assume un valore compreso tra 0 e 1 e misura il grado di appartenenza dell'unità ad un gruppo. La partizione (o ricoprimento) che si ottiene si chiama partizione sfocata (o ricoprimento sfocato).

Anche questi metodi possono essere distinti in *gerarchici*, e *non gerarchici*.



#### 7.4 METODI GERARCHICI DI CLASSIFICAZIONE SFOCATA

I metodi di classificazione sfocata che descriveremo in questo capitolo si caratterizzano per il fatto che si tratta di procedure gerarchiche che si compongono di due fasi: nella prima si calcola una misura della similarità tra le coppie di unità mentre nella seconda si applica alla matrice delle similarità ottenuta una procedura di classificazione delle unità che si conclude con l'assegnazione, a ciascuna unità, di una funzione di appartenenza ai gruppi che si sono formati. Naturalmente si potrebbero escogitare un gran numero di metodi di classificazione sfocata, analogamente a quanto è accaduto per i metodi di classificazione classica. Perciò i metodi descritti di seguito possono essere modificati per produrre tutte le possibili varianti di essi che meglio si adattano a situazioni particolari.

Il metodo della sintesi di più partizioni (Zani, 1989) E' un procedimento che, partendo da  $M$  partizioni iniziali delle unità, arriva ad una classificazione sfocata di esse.

Supponiamo di avere  $n$  unità spaziali su cui siano rilevati  $M$  caratteri di tipo misto (quantitativi e/o qualitativi). Per determinare la similarità tra una coppia di unità si fa anzitutto, per tutti i caratteri considerati, una iniziale partizione in gruppi delle unità e si assume come indice di similarità tra le due unità che costituiscono la coppia, la frequenza relativa delle partizioni, una per ogni carattere, in cui le due unità si trovano incluse in uno stesso gruppo. Agli indici di similarità così ottenuti si applica poi una procedura di classificazione, molto simile alle procedure classiche, che produce una successione gerarchica di partizioni, cioè, per ogni livello di similarità considerato, la procedura produce una partizione sfocata delle unità i cui gruppi contengono quelli formati ai livelli precedenti di similarità. Nel corso di tale procedura ad ogni livello di similarità si assegna a ciascuna unità la sua funzione di appartenenza alla partizione sotto il vincolo che la somma dei gradi di appartenenza di ogni unità a tutti i gruppi sia uguale a 1 (vincolo necessario per ottenere una partizione sfocata anziché un ricoprimento, Bezdek, 1981).

Si ha quindi il problema della scelta della partizione iniziale di ciascun carattere. Tale scelta è però abbastanza naturale per i caratteri sconnessi o ordinali o per quelli

quantitativi discreti con un numero ridotto di modalità in quanto sono le stesse modalità che comportano generalmente per ogni carattere la partizione migliore. Per i caratteri quantitativi le cui modalità non possono che essere classi (o perché il carattere è continuo o perché è discreto ma con un numero molto alto di modalità), la partizione in classi presenta indubbiamente un più ampio margine di soggettività.

E' quindi almeno consigliabile (secondo S. Zani) ripetere la procedura considerando più partizioni iniziali con riferimento allo stesso carattere confrontando di conseguenza più risultati tra loro.

Per i caratteri quantitativi esistono comunque diversi criteri per la ricerca della partizione iniziale; possiamo, ad esempio, citarne alcuni:

- i) suddivisione in base ai quartili;
- ii) metodo della minima varianza (Spath, 1985);
- iii) metodo delle classi naturali di Mineo (1978);
- iv) metodo di ottimizzazione di Butler (1988).

Se si adotta il primo di questi criteri, cioè quello della suddivisione in quartili, le unità vengono raggruppate assegnando al primo gruppo quelle per le quali la modalità del carattere è compresa tra il minimo e il primo quartile, al secondo gruppo le unità con modalità del carattere compresa tra il primo e il secondo quartile, e così via.

Si individuano così le M partizioni indotte dalle M variabili. Siano  $g_k$  i gruppi della partizione indotta dal k-esimo carattere (con  $2 \leq g_k \leq n-1$  e  $k = 1, 2, \dots, M$ ).

Siamo in grado di definire il **grado di appartenenza congiunto** (ovvero la similarità) di due unità statistiche,  $a_i$  e  $a_j$ : esso è dato dalla frequenza relativa delle partizioni in cui le due unità sono incluse in uno stesso gruppo:

$$z_{ij} = \frac{1}{M} \sum_{k=1}^M \delta_k(i, j)$$

dove  $\delta_k(i, j)$  vale 1 se  $a_i$  e  $a_j$  sono nello stesso gruppo nella partizione k-esima, e vale 0 altrimenti.

$z_{ij}$  è un indice di similarità che assume valori nell' intervallo  $[0,1]$ , è simmetrico (perché  $z_{ij}=z_{ji}$ ), gode della proprietà riflessiva (perché  $z_{ii}=z_{ji}$ ) ed il suo massimo corrisponde alla piena appartenenza delle due unità allo stesso gruppo in ciascuna delle  $M$  partizioni ( $z_{ij} \leq z_{ii}=1$ ).

Con i valori dell' indice  $z_{ij}$  si costruisce la **matrice delle similarità**. Per ottenere le classificazioni sfocate, si possono applicare alla matrice delle similarità diversi metodi di cluster per i quali occorre però definire il grado di appartenenza di una unità ad un gruppo sfocato. Esso può essere:

- a) il massimo delle similarità tra essa e ciascuna delle altre unità incluse nel gruppo;
- b) il minimo delle similarità tra essa e ciascuna delle altre unità incluse nel gruppo.

Mentre la definizione b) presenta analogie con il metodo di classificazione classica del legame completo, la definizione a) presenta analogie con il metodo di classificazione classica del legame singolo, in quanto una unità viene inclusa in un gruppo sfocato, ad un certo livello, quando essa presenta un valore della similarità con almeno un elemento del gruppo pari al valore del livello.

La procedura di classificazione si articola in tre fasi:

- i) Si riuniscono tra loro le eventuali coppie di unità con similarità uguale ad 1, ottenendo una partizione delle unità. Si dice, in tal caso, che queste due unità costituiscono un "*nucleo*" (core) (Rolland-May, 1985). I nuclei possono essere ovviamente più di uno.
- ii) Nella matrice delle similarità si prende in considerazione la similarità a uguale a  $(M-1)/M$  e si individuano le coppie che presentano tale grado di similarità. Se entrambe le unità costituenti una coppia con similarità  $a$  non sono già state assegnate ad un gruppo precedente, esse vengono a formare un nuovo gruppo, con grado di appartenenza uguale ad  $a$  (tali unità individuano un "*nucleo al livello  $a$* "); se una delle unità era già stata inserita in un gruppo, l' altra viene aggregata a quel gruppo, con grado di appartenenza al gruppo medesimo uguale ad  $a$ .
- iii) Si itera la fase sub ii) considerando livelli via via decrescenti del grado di appartenenza:

$(M-2)/M$ ,  $(M-3)/M$ , .... In questi passi successivi si può manifestare anche il caso di coppie di unità che presentano tra loro il grado di appartenenza, considerato a quel livello, ma che sono già state inserite in gruppi diversi. In questa circostanza S. Zani suggerisce di assegnare "in parte" ciascuna unità all' altro gruppo, con grado di appartenenza uguale al valore di  $a$  in oggetto, purché la somma per riga dei gradi di appartenenza risulti uguale a 1; in caso contrario si potrà attribuire convenzionalmente come grado di appartenenza il valore massimo che soddisfa tale vincolo. Se tale valore massimo è 0 l' unità non può più essere assegnata a nessun gruppo anche se presenta il necessario livello di similarità con altre unità.

L' algoritmo genera dunque una successione gerarchica di raggruppamenti sfocati, in corrispondenza di livelli decrescenti della similarità (Dimitrescu,1988).

*Il metodo dei ricoprimenti sfocati è di un metodo gerarchico di classificazione sfocata basato sulla costruzione, e successiva scomposizione, di una matrice delle similarità che fornisce ricoprimenti sfocati delle unità di partenza.*

Occorre dunque definire l'indice di similarità adottato in questo metodo. Tale indice si adatta al minimo livello di misura dei caratteri che si hanno a disposizione per cui non costringe né a considerare soltanto variabili quantitative, come spesso accade quando si utilizzano i metodi di cluster analysis, né ad effettuare manipolazioni dei dati che comportano generalmente variazioni degli indici di similarità. Supponiamo di avere un collettivo di  $n$  unità su cui si rilevano  $m$  caratteri:

a) se tutti i caratteri sono quantitativi, per ogni carattere si calcola la distanza relativa (cioè la distanza rapportata alla massima distanza assunta per quel carattere) tra le coppie di unità e se ne fa il complemento ad uno; si ha cioè, per il generico carattere  $k$ -esimo (con  $k=1,2,...,m$ ):

$$V_{ij}(k) = 1 - \frac{d_{ij}(k)}{d_{\max}(k)}$$

dove  $d_{ij}(k)$  è un indice di distanza detto *distanza di Hamming generalizzata* calcolato tra le unità  $i$ -esima e  $j$ -esima relativamente al carattere  $k$ -esimo; l'indice di similarità complessiva tra le unità  $i$  e  $j$  risulta essere:

$$S(i, j) = \sum_{k=1}^m V_{ij}(k) \cdot p(k)$$

dove  $p(k)$  è il peso non negativo che, nel caso in cui si possa ritenere di poter dare pesi diversi ai caratteri, si può assegnare al  $k$ -esimo carattere, con

$$\sum_{k=1}^m p(k) = 1$$

L'indice  $S(i,j)$  è anche noto come indice di Gower (Gordon, 1981).

b) se i caratteri sono ordinali o quantitativi si calcola la distanza tra le unità come la distanza relativa (nel senso chiarito prima) tra le loro posizioni in graduatoria e se ne fa il complemento ad 1; si adotta cioè la formula 1 con la differenza che ora la distanza è una distanza tra le posizioni in graduatoria anziché tra le modalità assunte dalle unità;

c) se fra i caratteri ve ne sono anche di sconnessi, poiché non è possibile calcolare la distanza di Hamming generalizzata tra le unità, si adotta come indice di similarità quello proposto nel *metodo della sintesi di più partizioni* (Zani, 1993).

Perciò per ogni carattere si individua una iniziale partizione delle unità e, relativamente al carattere  $k$ , si calcola l'indicatore  $d_k(i,j)$  che vale 1 se le unità  $i$  e  $j$  si trovano nello stesso gruppo e vale 0 altrimenti. Tale indice si sostituisce a  $V_{ij}(k)$  calcolato per i punti a) e b). In pratica, quindi, la similarità tra due unità risulta essere pari alla frequenza relativa delle partizioni in cui le due unità si trovano incluse in uno stesso gruppo.

Alla matrice delle similarità così ottenuta si applica una procedura di classificazione che è simile alla procedura classica del legame completo secondo la quale una unità può entrare a far parte di un gruppo se la similarità tra essa e tutte le unità di quel gruppo è almeno pari ad un certo livello fissato di similarità. Le fasi di tale procedura sono descritte di seguito.

i) Nella matrice  $S$  si cercano le eventuali coppie di unità con indice di similarità uguale ad 1: esse formano un gruppo non sfocato detto 'nucleo' (o 'core', Rolland - May, 1985). Il risultato di questa prima fase è quindi una partizione non sfocata perché la funzione di appartenenza non può che assumere i valori caratteristici delle partizioni classiche.

ii) Nella matrice  $S$  si cerca il massimo valore degli indici di similarità che sia minore di 1. Si individua, così, almeno un indice di similarità  $S(i,j)$  a cui corrispondono le unità  $i$  e  $j$  (le quali formano un nuovo gruppo) e a cui corrisponde il livello di aggregazione  $a$ . Si formano tanti gruppi quanti sono gli indici di similarità pari ad  $a$  trovati nella matrice  $S$ . Si determinano i gradi di appartenenza come il minimo delle similarità tra ciascuna unità e tutte le unità comprese nel gruppo.

iii) Si procede in modo analogo al passo precedente: sempre nella matrice  $S$  si ricerca il massimo valore tra quelli non ancora considerati. Detto  $b=S(h,l)$  tale valore, che corrisponde anche al livello di aggregazione, le unità  $h$  ed  $l$  si riuniscono in un gruppo. Se però una di esse, ad esempio  $h$ , risultasse già appartenente ad un gruppo formato in precedenza, si verifica se la similarità tra  $l$  e tutte le unità del gruppo in cui è compresa  $h$  non è inferiore pari a  $b$ : in tal caso  $l$  può essere assegnata al gruppo a cui già apparteneva  $h$ , altrimenti le due unità formano un nuovo gruppo. Quando si sono formati tutti i gruppi del livello  $b$ , si procede alla determinazione della funzione di appartenenza come nel passo precedente.

iv) Si itera il passo iii) per livelli decrescenti degli indici di similarità fino a che non si considerano tutti i possibili livelli.

La procedura di classificazione adottata nel metodo dei ricoprimenti sfocati differisce dalla procedura classica per due ordini di motivi:

- se una unità viene assegnata ad un gruppo essa può ancora essere assegnata ad un altro gruppo a condizione che presenti il necessario livello di similarità con le unità di quel gruppo;
- viene prodotta una *funzione di appartenenza* che fornisce l'insieme dei gradi di appartenenza di ciascuna unità a ciascun gruppo. Il grado di appartenenza di una unità ad un gruppo equivale alla similarità minima tra questa unità e ciascuna delle unità del gruppo.

Tutti i metodi qui proposti si basano sul calcolo di una matrice di indici di similarità tra le coppie di unità. Si può presentare quindi il problema di applicare tali metodi a dati di qualsiasi ordine di misura. In genere, infatti, è agevole calcolare distanze e similarità

soltanto se si considerano caratteri quantitativi o, al più, ordinali; in realtà sono state proposte anche diverse soluzioni di calcolo delle similarità per caratteri qualitativi, (ad esempio, A. Di Ciaccio, 1990). L'indice di similarità proposto da S. Zani risolve con estrema semplicità il problema del calcolo degli indici di similarità per caratteri di qualunque ordine di misura. Tale indice, può non risultare il migliore quando si dispone soltanto di caratteri quantitativi situazione che nelle applicazioni pratiche è forse la più frequente.

Infatti, se per i caratteri qualitativi la scelta della partizione iniziale non è un problema rilevante in quanto la partizione migliore è proprio quella individuata dalle stesse modalità assunte dalle unità, nel valutare la similarità per caratteri quantitativi c'è, invece, un più ampio margine di soggettività al momento di individuare la partizione iniziale delle unità. Questo modo di procedere ha comunque il vantaggio di poter calcolare un indice di similarità che sintetizzi l'informazione fornita da caratteri di tipo diverso. Ma se si hanno caratteri tutti di uno stesso tipo, ad esempio solo caratteri quantitativi, oppure solo caratteri rettilinei ordinati, con questo indice di similarità i dati vengono trattati nello stesso modo in cui si tratterebbero se si avessero solo caratteri sconnessi. Ciò può risultare poco conveniente nei casi in cui è possibile usare un indice di similarità che si adatti di più alla natura dei caratteri considerati.

Da qui nasce l'idea di considerare un indicatore di similarità che si adatti al minimo livello di misura dei dati poiché è ovvio che è meglio usare un indice che sfrutta al massimo l'informazione disponibile. Inoltre, buona parte delle situazioni reali fa riferimento a caratteri quantitativi o che si possono rendere tali con facilità (ad esempio è il caso delle frequenze o delle percentuali che possono far riferimento a fenomeni per loro natura qualitativa ma che di fatto possono essere analizzati con metodi adatti a dati quantitativi) per cui è buona norma prevedere sempre un indice di similarità adatto a caratteri quantitativi.

Per quanto riguarda la procedura di classificazione, occorre chiarire che il metodo usato da Zani differisce dal metodo classico del legame singolo per il fatto che per quest'ultimo, se una unità entra a far parte di un gruppo, essa non può più essere assegnata ad altri gruppi e quindi in seguito non verrà più presa in considerazione.

Il metodo di Zani, invece, è in metodo di classificazione sfocata per cui, se una unità entra a far parte di un gruppo con grado di appartenenza minore di 1, essa potrà ancora essere presa in considerazione perché può appartenere, per la parte rimanente, ad altri gruppi. In tal modo, però, si può intensificare il noto *effetto catena*, tipico del metodo del legame singolo, che produce la formazione di gruppi in cui le unità possono non essere tutte simili allo stesso livello, ma fanno ugualmente parte di quel gruppo perché sono simili ad almeno una delle unità del gruppo. Questo effetto si amplifica se si adotta un metodo di classificazione sfocata in quanto le unità vengono considerate più di una volta e quindi aumenta il rischio di associarle a gruppi in cui sono già presenti unità dissimili da esse.

Questo implica che è stata tacitamente adottata la proprietà di transitività maxmin secondo cui, una unità è simile ad un livello  $\alpha$  alle unità inserite in un gruppo, se il massimo dei gradi di similarità tra essa e ognuna delle altre unità è almeno uguale ad  $\alpha$ . Ma se vogliamo che valga una proprietà di transitività più forte che garantisca che la similarità delle unità appartenenti allo stesso gruppo sia almeno uguale ad un certo livello, dobbiamo adottare un metodo simile al metodo classico del legame completo: una unità è inserita in un gruppo al livello  $\alpha$  con un certo grado di appartenenza se essa presenta grado di similarità con tutte le unità già presenti nel gruppo, almeno pari ad  $\alpha$  (come nel *metodo dei ricoprimenti sfocati*). Cioè si adotta una transitività di tipo min-max, anche detta *proprietà di affinità sfocata* (L.A. Zadeh, 1975). Ciò evita che si formino gruppi con unità poco simili tra loro anche se il numero di gruppi che si formano può essere molto superiore al numero di gruppi che si ottenevano, allo stesso livello, con la procedura di classificazione proposta da Zani, ma in compenso è garantita una maggiore omogeneità tra le unità che fanno parte di uno stesso gruppo.

Naturalmente, una classificazione che rispetta la proprietà di affinità sfocata può risultare piuttosto complessa. Questa complessità, da un lato può essere limitata, come si è detto, con delle semplici scelte per selezionare i risultati ottenuti, e dall'altro comporta una quantità notevole di informazione che con gli altri metodi va quasi sempre persa.



Un'alternativa al *metodo della sintesi di più partizioni* e al *metodo dei ricoprimenti sfocati* può essere rappresentata dal *metodo del legame medio sfocato*, ancora una volta di ispirazione classica (nel senso che è simile al metodo del legame medio classico), e che fornisce gruppi formati da unità abbastanza omogenee tra loro in quanto perché una unità entri a far parte di un gruppo occorre che la similarità media tra essa e tutte le altre unità già presenti nel gruppo sia almeno pari ad una certa soglia  $a$ , ma il numero di gruppi che si forma è più o meno intermedio al numero di gruppi che si formano con gli altri due metodi e ai livelli più bassi della gerarchia la sfocatura diventa talmente alta che le funzioni di appartenenza di ciascuna unità tendono ad equidistribuirsi tra i gruppi. Osserviamo ora che il vincolo secondo cui la somma dei gradi di appartenenza di ogni unità ai gruppi deve valere 1 (che da ora in poi chiameremo ***vincolo della somma per riga***) può impedire la formazione di alcuni gruppi qualora sia imposto prima che sia stata completata la procedura di aggregazione. E opportuno introdurre tale vincolo a posteriori rispetto alla formazione dei gruppi, attraverso, ad esempio, una normalizzazione dei gradi di appartenenza, cioè dividendo i gradi di appartenenza di ciascuna unità per la somma dei gradi di appartenenza di quella unità a tutti i gruppi a cui essa appartiene. In realtà, tale normalizzazione può anche non avere luogo in quanto ha l'effetto di diminuire alcuni gradi di appartenenza e quindi di fornire una lettura dei risultati non del tutto realistica. Se non si effettua la normalizzazione anziché ottenere delle partizioni sfocate si ottengono dei ricoprimenti sfocati

## 7.5 METODI NON GERARCHICI DI CLASSIFICAZIONE SFOCATA

I metodi di classificazione sfocata esaminati in questo paragrafo sono caratterizzati dall'essere tutti di tipo non gerarchico, ossia ciascuno di essi fornisce una classificazione sfocata delle unità in un ben determinato numero di gruppi che viene stabilito a priori fin dall' inizio della procedura di classificazione. La classificazione viene poi ottenuta attraverso un processo iterativo tendente alla ottimizzazione di una funzione obiettivo che, in genere, rappresenta una misura della dispersione dei punti dai centri dei cluster. La differenza principale tra i metodi di questo tipo consiste di fatto nella diversa funzione obiettivo adottata e, dunque, nel differente processo iterativo utilizzato per calcolare i gradi di appartenenza delle unità ai vari gruppi.

Nel seguito saranno esposti i metodi di classificazione più noti ma, essendo senza dubbio il metodo delle k-medie sfocato quello più conosciuto ed utilizzato, ad esso è dedicata una maggiore attenzione rispetto agli altri.

*Il metodo delle k-medie sfocato è tra i metodi di classificazione sfocata di tipo non gerarchico, il più conosciuto ed utilizzato (Bezdek, 1981). Esso rappresenta una generalizzazione del metodo classico delle k-medie ed è particolarmente indicato per trattare grosse matrici di dati poiché la convergenza verso la classificazione finale viene generalmente raggiunta in breve tempo.*

Per utilizzare il metodo delle k-medie sfocato si procede nel seguente modo:

dopo aver scelto il numero  $c$  di cluster in cui si vogliono suddividere le  $n$  unità sulle quali sono state rilevate le modalità  $x$  di  $p$  caratteri, si fornisce una partizione iniziale delle unità nei  $c$  gruppi (che può essere casuale o basata su conoscenze a priori del ricercatore). Partendo da questa si ottiene, attraverso successive iterazioni tendenti alla minimizzazione di una funzione obiettivo, una classificazione sfocata nella quale per ogni unità viene determinato il grado di appartenenza ai  $c$  gruppi.

Il grado di appartenenza di un' unità ad un gruppo viene espresso per mezzo dei valori

$\mu_{ik}$

assunti dalla funzione di appartenenza i quali sottostanno ai seguenti vincoli:

$$1) \quad 0 \leq \mu_{ik} \leq 1 \quad i=1, \dots, n \quad k=1, \dots, c$$

$$2) \quad \sum_{k=1}^c \mu_{ik} = 1$$

dove il vincolo 1) definisce l' insieme di definizione della funzione di appartenenza ed il vincolo 2) indica che la somma dei gradi di appartenenza di ogni unità ai vari gruppi deve valere 1.

L' insieme dei valori della funzione di appartenenza può essere rappresentato in una matrice  $U = [\mu_{ik}]$  di dimensione  $(n \times c)$ .

La funzione obiettivo da minimizzare, detta *funzione di ottimizzazione*  $J_m$ , viene utilizzata per calcolare i valori ottimi del grado di appartenenza ed è funzione del quadrato della distanza  $d_{ik}$  tra l' unità  $i$ -esima ed il centroide del  $k$ -esimo cluster e dipende da un parametro  $m$  che può assumere qualunque valore reale  $\geq 1$ :

$$J_m(U, v) = \sum_{k=1}^c \sum_{i=1}^n (\mu_{ik})^m (d_{ik})^2 \quad (*)$$

dove  $(d_{ik})^2 = \|x_i - v_k\|^2$  è un'opportuna norma su  $R^p$

$v_k$  è la componente  $k$ -esima del vettore dei centroidi

$x_i$  è la componente  $i$ -esima del vettore dell'unità  $x$

$U$  è la matrice di dimensione  $(n \times c)$  dei gradi di appartenenza.

La funzione obiettivo  $J_m$  ha una chiara interpretazione: per ogni dato cluster, il suo centroide è la miglior rappresentazione delle unità che lo compongono poiché esso minimizza la somma dei quadrati degli errori  $x_i - v_k$ . Così,  $J_m$  misura l' errore quadratico totale in cui si incorre nel rappresentare le  $n$  unità con i  $c$  centroidi dei cluster. Il valore di  $J_m$  dipende allora da come le unità sono raggruppate nei cluster e rappresenta dunque una misura della dispersione delle unità intorno ai centri dei cluster; la partizione ottima è considerata quella che minimizza  $J_m$ . Tale partizione è anche chiamata partizione di minima varianza.

Il parametro  $m$  che compare nella (\*) riveste una particolare importanza poiché, a seconda del valore che si sceglie (valore che deve essere fornito all'inizio della procedura), la classificazione che si otterrà sarà più o meno sfocata. L' algoritmo che descrive il metodo delle  $k$ -medie sfocato è il seguente:

**Passo 1:** fissato il valore di  $m \in [1, \infty)$  e di  $\chi \in [2, v)$  e scelta la metrica da utilizzare, si sceglie una partizione iniziale delle unità in  $c$  gruppi che può essere rappresentata con la matrice  $U(0) = [\mu_{ik}]$  dove con l' esponente si indica il numero di iterazioni.

**Passo 2:** si calcolano i  $c$  centroidi dei cluster  $v_k$

usando la formula:

$$V_k^{(o)} = \frac{\sum_{i=1}^n (\mu_{ik}) x_i}{\sum_{i=1}^n (\mu_{ik})^m} 1$$

**Passo 3:** si calcola la nuova matrice  $U$ , che rappresenta il risultato della prima iterazione, secondo le seguenti regole:

- a) se per qualche gruppo  $r$  si ha che  $d_{ir} = 0$ , si pone  $\mu_{ir} = 1$  e  $\mu_{ik} = 0$  per tutti i  $k \neq r$ ;
- b) se la precedente condizione non è soddisfatta allora si utilizza la seguente formula:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^c \left( \frac{d_{ik}}{d_{jk}} \right)^{2/(m-1)}}$$

**Passo 4:** si calcola la differenza tra i risultati ottenuti all' ultima e alla penultima

iterazione usando un' opportuna norma: se

$$|U^{(1)} - U^{(0)}| \leq \delta \quad (**)$$

dove  $\delta$  è un parametro stabilito a priori, allora ci si ferma e si considera come classificazione finale quella ottenuta all' ultima iterazione, altrimenti si torna al passo 2 e si esegue una nuova iterazione continuando il procedimento fin quando la (\*\*) non è soddisfatta.

Uno dei principali motivi per cui il metodo delle  $k$ -medie sfocato è molto utilizzato, risiede nella rapidità con la quale tale metodo arriva alla classificazione finale. Notiamo

inoltre che studi recenti (Bezdek e Hathaway, 1988) hanno dimostrato che le soluzioni finali ottenute corrispondono sempre e solamente ad un punto di minimo locale o globale della funzione obiettivo (\*) o, al peggio, ad un suo punto sella. D' altra parte, ciò può non essere uno svantaggio, poiché, seppure un punto di minimo globale è senza dubbio da preferire ad uno di minimo locale o ad un punto sella, è stato anche fatto notare che, in molti casi, le classificazioni corrispondenti ai tre diversi punti sono praticamente identiche tra loro. Un' altra delle caratteristiche positive della convergenza del metodo delle k-medie sfocato, consiste nel fatto che ad ogni successiva iterazione il valore della funzione obiettivo decresce rispetto a quello dell' iterazione precedente cosa che invece non sempre si verifica negli altri metodi di questo tipo suscitando così non poche perplessità sulla loro convergenza.

Infine, resta da segnalare che osservazioni empiriche hanno messo in evidenza che il metodo risulta essere "relativamente indipendente dalla scelta della partizione iniziale" (Bezdek e Hathaway, 1988) convergendo comunque sempre allo stesso punto della funzione obiettivo.

Come già detto, nella maggior parte dei casi il metodo delle k-medie sfocato fornisce lo stesso risultato indipendentemente da quale sia la partizione iniziale che può essere così casuale o scelta dall' utilizzatore in base a precedenti conoscenze del fenomeno. È ovvio però che, quanto più la partizione di partenza si avvicina a quella finale, tanto più si accelera il processo di convergenza.

Dunque, per ridurre i tempi di elaborazione, sembra essere buona regola scegliere sempre come  $U(0)$  quella ottenuta mediante un altro metodo di classificazione (classico o sfocato).

Per quanto riguarda la scelta del numero  $c$  di cluster, questa è lasciata di fatto alla sensibilità del ricercatore che dovrà basarsi in generale sulle supposizioni o conoscenze che ha del fenomeno. Infatti, poiché da uno studio dei diversi test proposti per la scelta del numero di cluster (Milligan, 1985) risulta che nessuno di questi può essere considerato esente da difetti, non esiste un criterio oggettivo per scegliere il parametro  $c$ . L' unico procedimento che può essere adoperato per aiutarsi nella scelta, è quello di provare alcuni valori di  $c$  e di confrontare le classificazioni ottenute facendo però

attenzione a non scegliere un numero di cluster troppo grande poiché, essendo ogni unità ripartita in parte nei vari cluster, si potrebbe ottenere una classificazione troppo sfocata e dunque di difficile interpretazione.

All' inizio della procedura di classificazione si deve scegliere, tra gli altri parametri, il valore di  $m$ . Questa scelta riveste una particolare importanza poiché a seconda del valore di  $m$  dato, la classificazione che si otterrà sarà più o meno sfocata.

In generale, si possono verificare due casi:

a)  $m=1$ ; se viene scelto questo valore per il parametro  $m$ , si dimostra che la classificazione delle unità è di tipo classico, ossia il grado di appartenenza assume solamente i valori 0 o 1 eliminando così qualunque tipo di sfocatura. In tal caso il metodo delle  $k$ -medie sfocato coincide con il metodo delle  $k$ -medie classico, che risulta essere così un caso particolare. Ogni unità viene in questo caso attribuita totalmente al cluster da cui ha distanza minore.

b)  $m > 1$  ; in questo caso, quanto più il valore di  $m$  sarà maggiore di 1, tanto più il grado di appartenenza tenderà ad assumere, per ogni unità, valori sempre più lontani dagli estremi 0 e 1 fino ad ottenere, al limite, il valore  $1/c$  corrispondente al caso di massima sfocatura della classificazione, dove ogni unità è equamente distribuita tra tutti i  $c$  cluster presenti.

Come è evidente, per riuscire ad ottenere un buon risultato la scelta di  $m$  risulta così determinante. Spesso, però, in assenza di ulteriori informazioni sul fenomeno da analizzare, tale scelta è piuttosto complicata e, per risolvere tale problema, sono state proposte alcune soluzioni, una delle quali verrà esposta nel seguente paragrafo.

Le diverse applicazioni effettuate con il metodo delle  $k$ -medie sfocato hanno messo in evidenza che non esiste un valore ottimale per il parametro  $m$ . Esso infatti varia a seconda delle singole applicazioni e ciò rappresenta un notevole difetto del metodo poiché la scelta di  $m$  può essere fatta solamente osservando le classificazioni finali ottenute in corrispondenza dei diversi valori e, in ogni caso, se non si hanno conoscenze del fenomeno indagato risulta difficile scegliere il livello di sfocatura ideale delle classificazioni.

Per ovviare a questo inconveniente, si è pensato di introdurre un indice che misuri il grado di sfocatura delle diverse classificazioni tenendo conto che una classificazione si dice totalmente sfocata se, detto  $c$  il numero di cluster, per ogni unità tutti i valori della funzione di appartenenza assumono valore  $1/c$ , oppure si dice classica se ogni unità appartiene ad un unico cluster.

Si noti che, secondo queste definizioni, il concetto di sfocatura è equivalente a quello di eterogeneità, cosicché una classificazione può essere detta più o meno sfocata a seconda se essa sia più o meno eterogenea. Da quanto detto, segue che un qualunque indice di eterogeneità può anche essere considerato un indice di sfocatura per cui per misurare il grado di sfocatura delle classificazioni si propone di utilizzare l'indice relativo di eterogeneità di Gini:

$$I = \frac{c}{c-1} \left[ 1 - \frac{1}{n} \sum_{i=1}^n \sum_{K=1}^c \mu_{ik}^2 \right]$$

Tale indice varia nell' intervallo  $[0,1]$  ed assume valore 1 nel caso di massima sfocatura e valore 0 nel caso in cui la classificazione è di tipo classico. Con l' aiuto di questo indice è così possibile avere una misura sintetica dell' effetto dei diversi valori di  $m$  sulle classificazioni corrispondenti e la scelta di  $m$  risulta così agevolata.

Dalle diverse applicazioni effettuate il valore migliore di  $m$  sembra essere quello ottenuto in corrispondenza del valore dell' indice  $I$  compreso tra 0.4 e 0.5.

E' stato sottolineato come il metodo delle  $k$ -medie sfocato, grazie alle proprietà di cui gode, risulti uno dei migliori metodi di classificazione sfocata.

Studi recenti hanno però evidenziato che, alcune volte, questo metodo può fornire una classificazione eccessivamente sfocata; in particolare, si considerino i due seguenti casi:

a) se il numero  $c$  di cluster è abbastanza grande, il fatto che ogni unità debba appartenere (almeno in parte) ad ognuno dei cluster, fa sì che la classificazione stessa potrebbe risultare troppo "frastagliata" e, dunque, di difficile interpretazione; b) se le unità sono raggruppate in cluster ben netti e separati tra loro, qualunque sfocatura nella classificazione risulterebbe fornire un risultato distorto.

Per eliminare alcuni di questi inconvenienti, recentemente sono stati proposti dei suoi perfezionamenti che danno luogo ad una classificazione che chiameremo **classificazione semisfocata** per distinguerla da quella sfocata prodotta dal metodo delle k-medie sfocato.

Questi nuovi metodi sono: i metodi di Kamel, Selim e Ismail

Nel 1984 Selim e Ismail hanno proposto tre nuovi metodi di classificazione semisfocata corrispondenti ad altrettante variazioni del metodo delle k-medie sfocato che, a parte queste, rimane inalterato. Tali variazioni sono le seguenti:

1) oltre al numero dei c cluster iniziali, si sceglie anche un numero di cluster  $p < c$  e si impone che ogni unità possa appartenere al più a p cluster; ciò si ottiene ordinando, per ogni unità, in ordine decrescente i primi p valori della funzione di appartenenza e ponendo uguale a zero i restanti c-p valori.

Per rispettare il vincolo secondo il quale per ogni unità la somma dei gradi di appartenenza deve essere uguale a 1, si opera poi la seguente normalizzazione:

$$w_{ik} = \frac{\mu_{ik}}{\sum_{k=1}^c (\mu_{ik})} \quad (***)$$

dove con  $w_{ik}$  si è indicato il valore della funzione di appartenenza normalizzata dell' unità-èsima al k-esimo cluster;

2) si impone che se un' unità è molto distante da un certo cluster, il suo grado di appartenenza a quel cluster è nullo, dunque, detta  $d_{ij}$  la distanza tra l' unità-èsima e il centro del j-esimo cluster, se  $d_{ij} > g$  (dove g è un valore prefissato), si pone  $\mu_{ij} = 0$ . Anche in questo caso si procede poi alla normalizzazione dei valori  $\mu_{ij}$  mediante la (\*\*\*)).

3) per eliminare la presenza di valori molto bassi della funzione di appartenenza, e rendere così più chiara la classificazione finale, se  $\mu_{ij} < \beta$  (dove  $\beta$  è un valore scelto a priori), si pone  $\mu_{ij} = 0$ .

Anche qui, come nei due metodi precedenti, si procede poi alla normalizzazione.



I tre metodi proposti, come gli stessi autori ammettono, non riescono però a risolvere tutti i problemi connessi con il metodo delle k-medie sfocato:

il metodo 1 infatti, risulta utile applicarlo solamente quando si possiedono precise informazioni a priori per poter scegliere un idoneo numero  $p$  di cluster, altrimenti limitare le unità ad appartenere ad un numero massimo di cluster può portare a serie distorsioni nel calcolo globale dei valori del grado di appartenenza delle unità stesse.

Il metodo 2 è utile se si ha il sospetto che esistano diversi dati anomali, ma anche in questo caso bisogna fare attenzione alla scelta del valore  $g$  perché se tale parametro viene scelto troppo piccolo, alcune unità potrebbero non essere assegnate ad alcun cluster.

Il metodo 3, infine, è quello che ha riscosso più consensi, in quanto sembra essere di più generale utilità e, inoltre, facendo una scelta appropriata del parametro  $\beta$ , per ogni unità viene così determinato in modo naturale il numero di cluster al quale appartiene, superando così il problema della scelta a priori del numero di cluster  $p$  incontrato nel metodo 1. Anche la scelta del parametro  $\beta$  però, non sempre risulta semplice, in quanto se viene scelto un valore troppo alto, alcune unità potrebbero non appartenere ad alcun cluster o, addirittura, qualche cluster potrebbe risultare vuoto.

Per superare il problema della scelta del parametro  $\beta$ , nel 1991 Kamel e Selim hanno proposto un metodo che, rispetto al precedente ha un' importante novità:

4) il metodo di classificazione TFCM (Thresholded Fuzzy C-Means). Tale metodo richiede che il valore del parametro  $\beta$  sia scelto solamente quando per tutte le unità sono stati calcolati i valori definitivi del grado di appartenenza mediante il metodo delle k-medie sfocato senza alcuna variazione; a questo punto, detti  $\mu_{ij}$  tali valori, si potrà scegliere  $\beta$  in base all' analisi dei risultati ottenuti nonché tenendo conto che il massimo valore che si può assegnare a  $\beta$  è dato da:

$$\beta_{\max} = \min \left\{ \min_i \max_j \mu_{ij}, \min_j \max_i \mu_{ij} \right\}$$

Una volta scelto il valore di  $\beta$ , tramite la (\*\*\*) si procede alla normalizzazione ricavando i nuovi valori dei gradi di appartenenza. I vantaggi di questo metodo rispetto al precedente sono due: anzitutto, sapendo quale è il valore massimo consentito per  $\beta$ , si

eliminano gli inconvenienti causati dalla scelta di un valore troppo alto; inoltre, poiché  $\beta$  viene utilizzato solamente alla fine del procedimento, è possibile considerarne diversi valori per poter confrontare le classificazioni corrispondenti senza che ciò comporti di dover ripetere per ogni valore di  $\beta$  l'intera procedura come invece accadeva con il metodo precedente. Questo, ovviamente, consente di risparmiare una grande quantità di tempo computazionale.

I quattro metodi precedentemente esposti, pur se riescono a superare alcune problematiche del metodo delle k-medie sfocato, non risolvono appieno uno dei suoi maggiori difetti che consiste nel fatto che, una volta scelto il numero  $c$  di cluster nel quale le unità devono essere classificate, il procedimento per mezzo del quale per ciascuna unità viene determinato il grado di appartenenza  $m_{ik}$  ad ognuno dei cluster non contempla il caso in cui un'unità non appartenga affatto ad uno o più cluster oppure, al limite, appartenga solamente ad uno di essi. Per ogni unità, infatti, viene determinato un grado di appartenenza ad ognuno dei  $c$  cluster maggiore di zero non ammettendo in tal modo che il legame tra un'unità ed uno o più cluster possa essere nullo. L'unica eccezione è rappresentata dal caso (molto raro, in verità) in cui un'unità coincide con il centro di un cluster, poiché in questo caso essa viene attribuita interamente a tale cluster ed il grado di appartenenza per tutti gli altri cluster viene posto uguale a zero (per rispettare il vincolo secondo il quale per ogni unità la somma dei gradi di appartenenza ai vari cluster deve essere uguale ad uno).

Per eliminare o, comunque, ridurre questo inconveniente, si è pensato quindi di apportare una nuova modifica al metodo delle k-medie sfocato che risponde all'esigenza di tenere in considerazione il fatto che, oltre alle unità coincidenti con il centro del cluster, anche alcune altre unità potrebbero possedere i requisiti necessari per essere classificate come totalmente appartenenti ad un solo cluster.

Nel metodo delle k-medie semisfocato (Iacovacci, 1997) si propone dunque di assegnare totalmente l'unità- $i$ -esima al cluster  $k$ -esimo se, detta  $d_{ik}$  la loro distanza, si ha  $d_{ik} < (1/a) D$  dove  $a$  è un parametro ( $>1$ ) determinato a priori, e  $D$  indica la distanza tra il centro del  $k$ -esimo cluster e il centro del cluster ad esso più vicino.

Questa formulazione risponde all' esigenza intuitiva di classificare come totalmente appartenente ad un cluster qualunque unità che abbia, rispetto al centro del cluster stesso, non solo una distanza ragionevolmente piccola, ma che inoltre sia abbastanza lontana da tutti gli altri cluster. Infatti, nel caso in cui l' unità- $i$ -esima fosse molto vicina al centro del cluster  $k$ -esimo, il quale è a sua volta poco distante dal centro del cluster  $p$ -esimo, classificare detta unità come totalmente appartenente al cluster  $k$ -esimo sarebbe un errore essendo in tal caso più giusto attribuire l' unità- $i$ -esima parte all' uno e parte all' altro cluster.

Oltre al metodo delle  $k$ -medie sfocato ed ai suoi perfezionamenti presentati, esiste in letteratura un grande numero di metodi di classificazione sfocata di tipo non gerarchico. Descriverli tutti esula dalle finalità del presente lavoro.

## 7.6 FUZZY ANALYSIS - IL METODO FUNNY e MND2

Certamente tra i compiti che più frequentemente un individuo si trova a dover risolvere quotidianamente, si annoverano la classificazione e la predizione. In effetti anche nel semplice atto di spostare un oggetto è necessaria una preventiva pianificazione dei movimenti, insieme ad una stima delle caratteristiche dell' oggetto da spostare (come ad esempio il suo peso, la posizione del baricentro, la matrice di inerzia) in base alle molteplici informazioni che il sistema nervoso centrale riceve dagli organi di senso.

L' evoluzione ha prodotto tecniche estremamente efficienti per assolvere questi compiti, tanto da far sembrare banale la soluzione di questa classe di problemi. In effetti, chi si occupa di robotica conosce bene le molteplici difficoltà che si incontrano nella pianificazione di una traiettoria e nel riconoscimento degli oggetti (Xie e Beni, 1991; Heiser e Groenen, 1997; Michailidis et al. 2000).

La classificazione dei dati appare dunque come il primo passo da compiere per dotare le macchine di quelle capacità in grado di renderle "intelligenti".

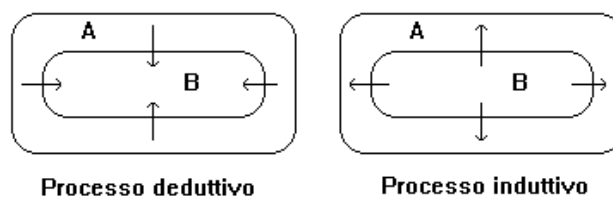
Si può dire che attualmente le macchine siano del tutto prive di capacità di classificazione? Certamente no . In effetti il "parser" di un compilatore implementato su un comune personal computer realizza un automa in grado di riconoscere un ben determinato linguaggio. Un automa riconoscente è in grado di partizionare l' insieme delle stringhe in due sottoinsiemi disgiunti, il primo dei quali comprende tutte le "frasi ben formate", mentre l' altro riunisce le stringhe che non sono compatibili con la grammatica del linguaggio.

Questo è un classico caso in cui viene applicato il principio aristotelico del terzo escluso: una stringa o appartiene al linguaggio, o non vi appartiene.

Nonostante i calcolatori sappiano risolvere brillantemente compiti di questo tipo, essi sono considerati sostanzialmente "stupidi". In effetti un parser non fa altro che compiere un procedimento deduttivo, in base alle regole che sono memorizzate in un opportuno

grafo (*diagramma sintattico*). La *deduzione*, attività logica per eccellenza è viziata da un' insuperabile sterilità, giacché si limita a trarre tutte le possibili conseguenze da alcune premesse, e dunque a esplicitare quanto implicitamente già contenuto in esse.

L' ~~ir~~irtevanza della deduzione nella ricerca e nella costruzione di nuova conoscenza dipende dal fatto che il processo deduttivo è un passaggio dal generale al particolare: una conseguenza della validità di una o più premesse non può che essere la validità di un caso meno generale. Da questo punto di vista, il ragionamento principe della deduzione è il seguente: *se supponiamo che tutti gli elementi di un certo insieme A soddisfino un determinato predicato, allora ne deduciamo che tutti gli elementi di un qualsiasi suo sottoinsieme B soddisfano sicuramente anch' essi quel predicato*(se supponiamo che gli italiani siano amanti della musica, allora i toscani sono amanti della musica).



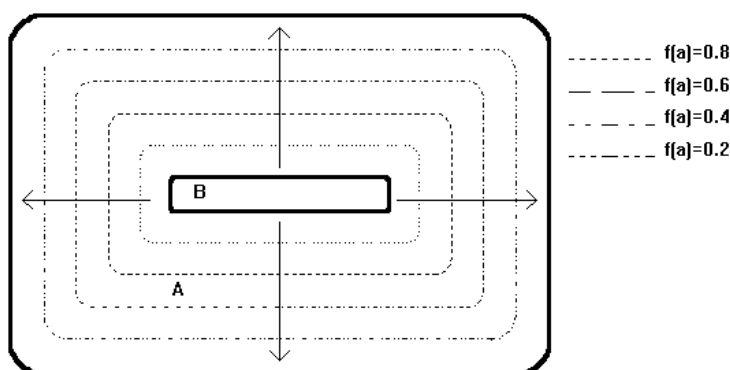
E' chiaro che un processo di costruzione di nuove conoscenze deve essere fondato proprio su un percorso in senso inverso (dal particolare al generale): *se supponiamo che tutti gli elementi di un certo sottoinsieme B di un insieme A soddisfino un determinato predicato, allora azzardiamo la tesi che tutti gli elementi dell' intero insieme A soddisfano anch' essi quel predicato*. Questo procedimento spericolato, questo ardito salto nel buio si chiama *induzione*. L' induzione, sia pure a rischio di commettere errori, ci consente di acquisire nuove informazioni.

La classificazione operata da un parser è di tipo deduttivo; se vogliamo indagare sui principi su cui si fonda l' intelligenza è necessario occuparsi di classificazione induttiva. Infatti, solamente l' ~~id~~induzione consente di progettare sistemi adattativi in grado di apprendere da esempi e di creare in base a questi l' informazione necessaria a classificare correttamente pattern mai elaborati, ossia di *generalizzare*.

Abbiamo detto, però, che l' induzione può essere causa di errori; ciò costituisce un effettivo pericolo qualora si utilizzi una logica che preveda la validità del principio del terzo escluso. Questo problema viene ridimensionato qualora si adotti la logica sfumata; per poter giustificare questa affermazione, si descriverà un processo induttivo su spazi normati che utilizza la logica fuzzy.

*Sia  $A$  un qualunque insieme in uno spazio normato e  $B$  un sottoinsieme di  $A$ ; supponiamo che ogni elemento di  $A$  soddisfi un predicato  $P$ ; all' insieme  $B$  viene associata una funzione di appartenenza  $f$  che soddisfi le proprietà di normalità, monotonicità e simmetria. Formuliamo quindi l' ipotesi che per ogni elemento  $aA$ , il postulato  $P$  sia vero in misura pari ad  $f(a)$ .*

Si noti che per la proprietà di monotonicità, all' aumentare della distanza dell' elemento  $aA$  dalla frontiera di  $B$ , diminuisce il "grado" di verità del predicato  $P$  (vedi figura ).



E' evidente che, in questo modo, la probabilità di commettere errori, si riduce: la generalizzazione operata in logica fuzzy non è più drastica e rischiosa, ma sfumata.

Del resto, tutti gli organismi a cui noi attribuiamo una certa intelligenza "ragionano" in logica fuzzy e per mezzo di questa rendono più sicure le decisioni derivanti da processi induttivi (ossia dall' aver generalizzato una determinata situazione).

Appare chiaro, allora, che la logica fuzzy consente di progettare algoritmi di classificazione flessibili e, allo stesso tempo, robusti.

FUNNY è il nome di un programma per personal computer il quale produce delle classificazioni sfocate utilizzando qualunque tipo di dati (siano essi numerici o misure

di dissimilarità). Sull'utilizzo di fuzzy architetture per la Customer Satisfaction si veda Temponi, Kuo, Corley 1999.

Il procedimento attraverso il quale si perviene alla classificazione delle unità è molto simile a quello utilizzato nel metodo delle k-medie sfocato: anche in questo caso, infatti, fissato il numero di gruppi nei quali si desidera suddividere le unità, si fornisce una partizione iniziale delle unità stesse assegnando a priori i valori della funzione di appartenenza; i valori finali si ottengono poi eseguendo una procedura di ottimizzazione di questi valori attraverso l'uso di una funzione obiettivo.

La differenza maggiore tra il metodo FUNNY (Kaufmann, 1990) ed il metodo delle k-medie sfocato consiste proprio nella funzione obiettivo che nel metodo

FUNNY assume la seguente forma:

$$C = \sum_{k=1}^c \left[ \sum_{i=1}^n \sum_{j=1}^n (\mu_{ik})^2 (\mu_{jk})^2 (d_{ij}) / 2 \sum_{j=1}^n (\mu_{jk})^2 \right]$$

dove  $(d_{ij})$  rappresenta la distanza (o la dissimilarità) tra l'unità- $i$ -esima e l'unità- $j$ -esima e  $\mu_{ik}$  esprime il grado di appartenenza dell'unità  $i$  al cluster  $k$ . La soluzione di minimo per la funzione si trova applicando il metodo dei moltiplicatori di Lagrange sotto le condizioni di Kuhn e Tucker alla funzione sotto i vincoli :

$$i) \mu_{ij} \geq 0 \quad i=1, \dots, n \quad k=1, \dots, c$$

$$ii) \sum_{k=1}^c \mu_{ik} = 1 \quad i=1, \dots, n$$

Tale soluzione è in genere una soluzione di ottimo locale e si ottiene per mezzo di una procedura di ottimizzazione che qui, per brevità di esposizione, non esporremo. Come detto, la differenza tra il metodo FUNNY e il metodo delle k-medie sfocato risiede solamente nella differente funzione obiettivo adottata, ed in particolare nel fatto che il metodo delle k-medie sfocato considera il quadrato della distanza, mentre il metodo FUNNY la utilizza con l'esponente uguale ad 1; inoltre, mentre nel metodo delle k-medie sfocato il parametro  $m$  può variare ( $m > 1$ ), nel metodo FUNNY esso è fisso e posto uguale a 2 poiché, secondo l'autore, è con tale valore che pare si ottengano i risultati migliori (notiamo comunque che, variando l'esponente, si ottengono delle classificazioni più o meno sfocate proprio come col metodo delle k-medie sfocato).

Un altro metodo di classificazione sfocata che presenta delle affinità con i metodi precedenti, è l' algoritmo *MND2* di Roubens (Roubens, 1978) che si differenzia dagli altri metodi a causa della funzione obiettivo adottata che è la seguente:

$$R = \sum_{k=1}^c \sum_{i=1}^n \sum_{j=1}^n (\mu_{ik})^2 (\mu_{jk})^2 (\mu_{ij})^2$$

Sebbene tale funzione appare molto simile a quella adottata dal metodo FUNNY (la differenza maggiore consiste nel non considerare il denominatore), la sua minimizzazione però tende a distorcere i risultati, in particolare se il numero di unità da classificare è molto elevato. In tal caso, infatti, il metodo *MND2* non sempre assicura una convergenza ottimale il che, abbinato all' altro difetto ora esposto, rende questo metodo decisamente meno preferibile rispetto a quelli considerati in precedenza.

I metodi che sono stati esposti in questo paragrafo presentano, come è stato visto, forti analogie tra di loro in quanto i procedimenti utilizzati per ottenere le classificazioni cercate sono tra loro molto simili. Poiché questi metodi sono tutti di tipo non gerarchico e alcuni di essi rappresentano una generalizzazione dei metodi classici, essi presentano in generale lo stesso tipo di problemi propri dei metodi non gerarchici di classificazione classica che consistono principalmente nella difficoltà di scelta delle condizioni iniziali.

Per quanto riguarda la scelta di alcuni parametri, nei paragrafi precedenti sono state proposte alcune soluzioni, ma diversi problemi restano di difficile soluzione come per esempio la scelta iniziale del numero *c* di gruppi la quale può essere fatta solamente dopo aver compiuto un' accurata analisi dei dati a disposizione e dopo aver ripetuto la procedura di classificazione per diversi valori di *c* in modo tale da poter valutare quale sembra essere il valore migliore.

Un altro inconveniente comune a questi metodi è poi quello relativo alla convergenza che non sempre risulta essere ottimale in quanto alle volte i risultati ottenuti corrispondono a soluzioni di minimo locale e non assoluto.

È infine da notare che, poiché il problema della convergenza è legato a quello della scelta della funzione obiettivo, quasi tutti i metodi di questo tipo utilizzano una funzione simile o riconducibile a quella utilizzata nel metodo delle *k*-medie sfocato giacché questo è quello che sembra convergere più rapidamente alla soluzione ottimale.



Nel complesso, comunque, i metodi precedentemente esaminati forniscono dei risultati che possono essere considerati molto soddisfacenti e, trattandosi di metodi tutti molto recenti, è facile prevedere che nell' immediato futuro verranno apportati ulteriori miglioramenti.

## 7.7 LA CONJOINT ANALYSIS

La **Conjoint analysis (analisi congiunta)**, è una tecnica d'analisi multivariata, che si basa sulla concezione di Lancaster secondo cui il valore di un prodotto può essere scomposto in un insieme di valori parziali, connessi alle utilità parziali dei singoli attributi dell'offerta, e alle sue componenti costo-sacrificio; consente di rilevare l'importanza relativa dei diversi attributi d'offerta, nonché il valore parziale che il cliente assegna a ciascun livello di prestazione.

L'applicazione dell'analisi congiunta consiste in una simulazione del processo di valutazione e scelta del cliente, con riferimento a diverse alternative d'offerta. Uno dei maggiori pregi di questa tecnica è rappresentato, dall'opportunità di rilevare le percezioni di valore in condizioni di cognizione e scelta molto simili alla realtà del processo d'acquisto. In sostanza il cliente non è costretto ad esprimere le sue valutazioni sui singoli attributi del prodotto, bensì su servizi alternativi verso i quali potrebbe verosimilmente indirizzare le sue preferenze. Ne derivano informazioni utili ed affidabili sui criteri di scelta e valutazione d'importanza delle specifiche prestazioni (livelli) dei prodotti, che emergono dall'elaborazione statistico - quantitativa dei giudizi rilevati.

Anche nel caso di misurazione di Customer Satisfaction mediante analisi congiunta, è necessario condurre una preliminare indagine qualitativa volta ad identificare la catena del valore del cliente, gli attributi rilevanti, e i livelli di prestazione attesi o desiderati da ciascun cliente, Inoltre, considerando che:

- ⌘ La misurazione delle preferenze è realizzata senza una valutazione diretta dell'importanza degli attributi;
- ⌘ Le preferenze, per i livelli d'attributi, sono misurate individualmente, anche se la forma del modello che descrive le preferenze il ricercatore può, pertanto, individuare facilmente eventuali omogeneità o disomogeneità;

⌘ La procedura, comporta la raccolta d'informazioni generali sugli intervistati, (caratteristiche demografiche, situazioni reddituali, d'acquisto e d'uso del prodotto), attraverso le quali si possono descrivere i segmenti;

⌘ Generalmente, è inclusa una fase di *simulazione*, in cui il ricercatore può: testare profili di prodotti nuovi o modificati, oppure, analisi di preferenza rispetto alla concorrenza.

Si evince chiaramente che l'analisi congiunta è, inoltre, particolarmente utile per affrontare problemi connessi alla segmentazione del mercato e allo sviluppo di nuovi prodotti.

Sul piano algoritmo utilizza la regressione multipla su variabili qualitative (c.d. variabili dummy), la cui applicazione presuppone:

- L'individuazione delle caratteristiche dell'offerta considerate rilevanti dal cli ente, ciò può avvenire, sia tramite ricerche qualitative sia interviste in profondità o focus group (di cui parleremo in seguito);
- Identificazione dei benefici, degli attributi e dei livelli di prestazione: minima, attesa, ottimale.
- La loro articolazione in livelli discreti, riconducibili alle diverse modalità in cui ciascuna caratteristica può concretamente manifestarsi;
- La definizione di un insieme di profili d'offerta alternativi;
- Raccolta, attraverso interviste personali e scale di valutazione proporzionali, dei giudizi formulati dal cliente sui profili d'offerta selezionati (Card).

Possiamo individuare sette fasi rilevanti nella strutturazione della conjoint analysis:

1. Scelta di un modello di preferenza, tra i seguenti: vettore, punto ideale, 'part-wort', modelli misti;
2. Metodo d'intervista (analisi 'trade off', 'full profile');
3. Costruzione di un insieme di stimoli, (metodo full profile), alternativamente mediante: disegni fattoriali frazionari, campionamento casuale da distribuzioni multivariate;
4. Presentazione degli stimoli mediante: descrizione verbale, rappresentazione pittorica o modello tridimensionale;

5. Livello di misurazione delle valutazioni, che possono essere espresse in termini di:
  - *Ranking* chiedendo al consumatore di ordinare i profili d'offerta da quello maggiormente preferito a quello meno gradito;
  - *Rating* esprimendo un voto per ciascun profilo d'offerta.
6. Metodo di stima: OLS, monanova, trade-off;
7. Modello di simulazione: First choice, BTL, LOGIT.

Non tutte le combinazioni alternative sono realizzabili da un punto di vista metodologico. L'obiettivo è quello di massimizzare la validità predittiva del modello, con i vincoli dati dai tempi d'intervista e dal budget di ricerca; la scelta dipenderà da fattori quali: il tipo di prodotto, mercato, il numero di attributi rilevanti e il tipo di intervistato.

1. **Scelta di un modello di preferenza**, sia  $t = 1, 2, \dots, k$  l'insieme dei  $k$  attributi (fattori) prescelti per descrivere il servizio; sia, inoltre,  $f_{jt}$  il livello del  $t$ -esimo attributo relativo allo stimolo  $j$ . Gli attributi considerati possono essere di natura *qualitativa* o *quantitativa*.

- ⊗ Il **modello vettore**, proposto da Srinivasan e Shocker, pone la preferenza (utilità)  $p_j$ , espressa per il  $j$ -esimo stimolo, pari a:

$$p_j = \sum_{t=1}^k w_t f_{jt}$$

dove i coefficienti  $w_t$  sono i pesi dei  $k$  attributi. Tali pesi saranno, in generale, diversi per i differenti individui nel campione.

- ⊗ Il **modello punto ideale**, la preferenza  $p_j$  sia negativamente correlata con la distanza al quadrato  $d_j^2$ , opportunamente ponderata, tra il livello del  $t$ -esimo attributo per il  $j$ -esimo stimolo ( $f_{jt}$ ) e il livello ideale per il  $t$ -esimo attributo ( $o_t$ ), con  $d_j^2$  definita nel modo seguente:

$$d_j^2 = \sum_{t=1}^k w_t (f_{jt} - o_t)^2$$

quindi, quanto più ridotta è la distanza di uno stimolo dal punto ideale, tanto più elevato è il gradimento  $p_j$  per lo stimolo stesso.

⊗ IL **modello Part-Worth**, esprime la preferenza per lo stimolo  $j$  attraverso una funzione discontinua  $s$ :

$$p_j = \sum_{t=1}^k s_t(f_{jt})$$

definita, per un insieme opportunamente selezionato di livelli (in genere, tre o quattro). Nel caso degli attributi quantitativi, la stima delle preferenze avviene attraverso un'interpolazione lineare.

*Il modello part- worth è il più flessibile*, permettendo alla funzione d'utilità di assumere diverse forme, a seconda della specificazione  $s_t(f_{jt})$ . Infatti definendo  $s_t = -w_t(f_{jt} - o_t)^2$ , la formula, (del modello part- worth),  $p_j = \sum s_t(f_{jt})$ , fornisce l'espressione del modello punto ideale; ponendo  $s_t(f_{jt}) = w_t f_{jt}$  si ottiene il modello vettore. Tuttavia, considerato il numero di parametri da stimare, il modello parth- worth è anche *il più complesso*: se vi sono  $q$  livelli per ciascuno dei  $k$  attributi, occorre stimare  $(q - 1)t^4$  parametri; il modello vettore, invece, comporta la stima dei  $k$  parametri  $w_t$ , nel caso del modello punto ideale, occorre stimare i  $2k$  parametri  $w_t$  e  $o_t$ . Risulta evidente, che esiste un trade- off tra flessibilità e semplicità dei modelli.

E' possibile, combinare le caratteristiche dei tre modelli, realizzando un **modello misto**, un attributo con  $q$  livelli può essere trasformato in  $(q-1)$  variabili *dummy* (a valori 0/1 – NO/SI), utilizzando queste il modello parth-worth è trasformato nel vettore. Il modello misto può essere riassunto dall'espressione seguente:

$$p_j = \sum_{t=1}^k \sum u_t z_{jt}$$

dove  $k$  è il numero totale di pseudo attributi (che coincide con il numero di parametri da stimare), e le quantità  $z_{jt}$  sono definite in relazione a  $f_{jt}$  nel modo seguente:

- ♣ Nel caso in cui ci si aspetti che le preferenze siano monotone, cioè con andamento lineare  $z_{jt} = f_{jt}$ ;
- ♣ Nel caso in cui si prevede che le preferenze siano non lineari, concave o convesse, o del tipo "punto ideale", per ogni attributo  $k$  sono definite due variabili  $z$ , una pari a  $f_{jt}$  e l'altra a  $f_{jt}^2$ ;

- ♣ Nel caso d'attributi qualitativi o quelli, la cui funzione d'utilità non sia quella del modello misto, per ogni attributo  $k$  con  $q$  livelli, sono definite  $(q-1)$  variabili dummy.

Se la preferenza complessiva, per i profili di servizio sottoposti a valutazione, è misurata a livello d'intervallo (è una variabile quantitativa), e per la stima dei parametri è usato il metodo di regressione multipla, gli usuali test sui parametri stimati forniscono utili indicazioni per la scelta dei diversi modelli di preferenza (per es. se il coefficiente del termine  $f^2$ , risultasse non significativo per l'attributo  $k$ , con il coefficiente del termine  $f$  significativo, il modello vettore risulterebbe più indicato rispetto al modello punto ideale).

**2. Il metodo d'intervista**, due sono le principali alternative di raccolta delle informazioni:

- ⊗ La procedura **trade off** (due fattori per volta), essa ha il **vantaggio** di rendere più agevole il compito dell'intervistato, evitando di richiederli troppe informazioni, e, non richiedendo particolari metodologie di somministrazione, si presta anche ad essere inserita in questionari postali. Più rilevanti gli **svantaggi**:

- Decomporre l'insieme globale degli attributi in una serie di coppie da esaminare separatamente, fa sì che l'intervistato, nel valutare il ranking delle coppie di livelli di due attributi, può dimenticare le ipotesi relative ai rimanenti  $k - 2$  attributi;
  - Il compito dell'intervistato può diventare lungo e noioso, infatti, nel caso di sei fattori con quattro livelli ciascuno, il numero totale di tavole da valutare è abbastanza alto, pari a  $(6 * 5) / 2 = 15$  con  $4 * 4$  celle ciascuna;
  - Gli intervistati hanno la tendenza a dimenticare il punto della tavola in cui sono arrivati, o ad adottano schemi di risposta standardizzati, (ad es. prendendo in considerazione sistematicamente variazioni di un fattore prima di considerare l'altro;
  - Infine, la procedura si presta maggiormente a descrizioni verbali degli stimoli, piuttosto che a rappresentazioni "visive", si pensi allo studio sull'aspetto esteriore del servizio.
- ⊗ L'approccio **Full profile**, le considerazioni fatte sul metodo precedente si rovesciano, i principali vantaggi, sono i seguenti:

- Minor numero di valutazioni, (seppure più complesse rispetto a quelle trade- off), da parte del rispondente.
- E' possibile richiedere valutazioni di tipo rating (punteggi su scale del tipo 1 -5, con 1 = poco gradito, 5 = molto gradito), o ranking (ordinamento) dei profili globali, dal più gradito al meno gradito.
- Consente una più realistica descrizione degli stimoli, definiti attraverso specifici livelli per ciascuno degli attributi selezionati; è quindi possibile tener conto della potenziale correlazione tra i fattori, ciò è tanto più vero quanto minori, sono le difficoltà di valutazione per il rispondente.
- *L'unico inconveniente* è il sovraccarico d'informazioni sull'intervistato con la conseguente tendenza a semplificare la sperimentazione, trascurando le variazioni nei fattori giudicati meno rilevanti o semplificando psicologicamente i livelli. Considerato che, in genere le scelte avvengono all'interno di un ridotto insieme d'alternative si comprende perché si tende a ridurre il numero d'attributi (cinque o sei al massimo), e di livelli considerati.

**3. La costruzione (metodo full profile) e la presentazione dell'insieme di stimoli**, il numero dei profili d'offerta che il consumatore considera quando opera la sua scelta, sono, nella realtà limitati.

⊗ Un primo problema, riguarda la *scelta del numero degli stimoli*. Nel caso in cui il metodo di stima sia il modello di regressione, si ha che l'errore quadratico medio di previsione è positivamente correlato, al rapporto tra il numero di parametri da stimare ed il numero d'osservazioni (in questo caso gli stimoli); da qui l'indicazione ad aumentare al massimo possibile il numero di stimoli da sottoporre a valutazione.

⊗ Un secondo problema è la *scelta dei campi di variazione e della correlazione tra gli attributi*, infatti:

- Se da un lato è preferibile utilizzare specifiche degli attributi simili a quelli reali, per dare al rispondente un'impressione di credibilità della procedura di valutazione;
- Dal punto di vista metodologico, esiste l'opportunità di scegliere disegni degli esperimenti, che riducano ad una dimensione ottimale il numero dei profili, che porta

verso la specificazione di profili d'offerta diversi da quelli reali. *I principali metodi di riduzione degli stimoli* sono:

- **Il quadrato latino**, particolare disegno fattoriale frazionato, con cui si ottiene un'elevata riduzione del numero di combinazioni, trascurando tutte le potenziali iterazioni; si tratta di un *disegno simmetrico*, in quanto, il quadrato latino prevede che ciascun fattore considerato abbia lo stesso numero di livelli;
- **Matrici ortogonali**, consentono di sviluppare disegni ancora più efficaci nella riduzione del numero di combinazioni necessarie per stimare gli effetti principali dei fattori. Nel caso di **disegni simmetrici**, in cui ciascun fattore ha  $q$  livelli, e se esso è un numero primo o una potenza di esso, si può costruire un disegno ortogonale; se  $q$  non soddisfa tali condizioni o si consideri un disegno asimmetrico, le matrici ortogonali sono costruite aggregando i livelli di opportuni disegni simmetrici.

Combinazione	fatt. 1	fatt. 2	fatt. 3	fatt. 4	fatt. 5	fatt. 6
1	1	1	1	1	1	1
2	2	2	2	2	1	1
3	2	1	2	1	2	1
4	1	2	1	2	2	1
5	1	2	2	1	1	2
6	2	1	1	2	1	2
7	2	2	1	1	2	2
8	1	1	2	2	2	2

Tab. # Esempificazione di disegno ortogonale simmetrico

La tabella #, riporta un disegno ortogonale simmetrico, comprendente 6 fattori con due livelli ciascuno, che consente una riduzione da  $2^6 = 64$  ad 8 combinazioni. Si noti che ciascun livello appare uno stesso numero di volte, (in questo caso 4), per ciascun fattore. Questa condizione non è necessaria nel caso di matrici ortogonali (simmetriche e asimmetriche). Condizione necessaria e sufficiente affinché gli effetti principali di ogni coppia di fattori siano non correlati, è che ogni fattore appaia con frequenza proporzionale a quella di ogni livello di un altro fattore.



- **Campionamento casuale di una distribuzione multivariata**, si tratta di una procedura alternativa ai disegni ortogonali, per creare descrizioni degli stimoli. Nel caso che, tutti gli attributi siano quantitativi e continui, è possibile individuare una distribuzione multivariata (normale) a partire dalle medie, dalle deviazioni standard e dalle correlazioni tra attributi; la descrizione degli stimoli può quindi essere estratta casualmente da tale distribuzione. *Metodologicamente*, per utilizzare questo metodo di riduzione del numero degli stimoli, è opportuno generare un certo numero di profili casuali, in eccesso rispetto al numero desiderato, in modo da eliminare i profili *dominati*, ovvero quelli che hanno livelli degli attributi meno desiderabili per ciascun fattore; attraverso una serie di estrazioni, è sempre possibile definire un insieme di  $n$  descrizioni, in modo che nessun profilo domini uno dei rimanenti  $n - 1$  profili. Questa procedura è più complessa e laboriosa rispetto alla costruzione di disegni ortogonali, ed è da preferirsi a questa solo in presenza di forti correlazioni tra i fattori considerati.
- ⊗ Per quanto concerne la **presentazione degli stimoli**, sono state individuate due differenti alternative:
  - **La descrizione verbale**, vi sono due di somministrazione degli stimoli al cliente:
    - Nel primo caso, a ciascun intervistato sono sottoposte  $n$  schede di prodotto, ciascuna delle quali descrive schematicamente i livelli di ciascuno dei  $k$  attributi; al rispondente è chiesto di ordinarle o di valutarle in una scala quantitativa. Il principale **vantaggio** di questa procedura è la sua semplicità e l'efficienza con cui i dati possono essere raccolti; uno **svantaggio possibile** è, invece, dato dal fatto che, (è stato dimostrato) la stima dell'importanza di un attributo è in qualche modo influenzata dall'ordine e dalla posizione dell'attributo sulla scheda; per ridurre la distorsione potenziale, l'ordine viene in genere variato casualmente tra gli intervistati, così come l'ordine di presentazione delle schede.
    - Il secondo tipo di descrizione verbale, si basa su una descrizione realistica e completa degli stimoli e ha il **vantaggio** di consentire, contestualmente, una valutazione dell'efficacia dell'annuncio pubblicitario. Lo **svantaggio** è che tale metodologia porta ad un'ulteriore riduzione del numero dei profili valutabili dal rispondente con la dovuta attenzione, con maggiore inaccuracy delle stime dei parametri a livello individuale.

- **La rappresentazione visiva** comporta notevoli **vantaggi rispetto ai profili verbali**: è risparmiato lo sforzo di traduzione dello stimolo dal piano verbale a quello visivo; inoltre, attraverso tale procedura si forza una maggiore omogeneità delle percezioni tra i rispondenti, rendendo gli stimoli più realistici, e il compito del valutatore meno faticoso. Lo **svantaggio**, è dato dal conseguente aggravio dei costi.

**4. Il livello di misurazione delle valutazioni, e il metodo OLS di stima dei parametri,**  
la misurazione può essere espressa in termini di *preferenza globale*, oppure come *propensione all'acquisto*, (nel caso in cui lo studio si riferisca a nuove tipologie di prodotti / servizi). Secondo lo scopo dell'indagine, individuiamo una delle due seguenti scale di misurazione:

- **Metodi non metrici (ad ordinamento degli stimoli)**, Metodologicamente la procedura richiede un'intervista personale e guidata, perché è preferibile, che l'ordinamento avvenga in tre fasi:
  - ♣ Dapprima, separando in due o più gruppi gli stimoli, dai meno preferiti ai maggiormente graditi;
  - ♣ Successivamente si procede all'ordinamento all'interno dei gruppi;
  - ♣ Ed infine, unire i gruppi di schede e controllare l'ordinamento ottenuto.

**Principali vantaggi:**

- ♦ *Semplicità*, è più facile per il rispondente, (anche se necessita di un maggiore dispendio di tempo), scegliere gli stimoli preferiti, piuttosto che esprimerne il grado di preferenza.
- ♦ *Possibilità d'uso sia di modelli additivi che moltiplicativi*, infatti, la trasformazione logaritmica, che preserva l'ordinamento della variabile dipendente, consente di ricondurre il modello moltiplicativo a quello additivo.

Nella scelta del livello di misurazione delle preferenze, si utilizza la procedura MONANOVA di Kruskal

- **Metodi metrici (a punteggio)**, il principale **vantaggio** rispetto al metodo precedente è dato dal contenuto informativo potenzialmente più elevato, e dal fatto che tale approccio può essere applicato anche in un'indagine postale.

Nella scelta del livello di misurazione delle preferenze, s'impiegano prevalentemente i metodi di regressione lineare (minimi quadrati ordinari – OLS). L'evidenza empirica dimostra che i risultati ottenuti applicando i metodi di regressione classici, (a dati interi, in cui la variabile dipendente è ridefinita come variabile misurata a un livello d'intervallo), differiscono solo leggermente dai risultati forniti metodi non metrici (più complessi).

Il metodo più frequentemente utilizzato per la stima dei parametri è quello dei minimi quadrati. Più precisamente, l'analisi di regressione è applicata ad un sistema lineare composto di un numero d'equazioni pari a quello dei profili selezionati; i giudizi espressi su ciascun profilo di servizio, costituiscono le variabili dipendenti, mentre le variabili indipendenti sono rappresentate dai livelli discreti, (codificati come variabili *dummy*), delle caratteristiche dell'offerta. Il metodo OLS, prevede che la variabile dipendente dovrebbe essere misurata a livello d'intervallo, è tuttavia, prassi diffusa utilizzare valutazioni di tipo rating o ranking.

L'elaborazione fornisce i coefficienti numerici relativi ad importanza, e utilità associate da ciascun cliente ai diversi livelli degli attributi, considerati nel processo di scelta. Si supponga di voler studiare la relazione tra una variabile dipendente  $Y$  e  $k$  variabili esplicative  $X_1, X_2, \dots, X_k$ ; il modello di regressione lineare multipla può essere descritto in forma matriciale, nel modo seguente:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$

dove,  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  è il vettore, formato dalle  $n$  osservazioni disponibili su  $Y$ , mentre  $\mathbf{e}$  è un vettore di  $n$  osservazioni relative ad una variabile casuale  $e$  (detta 'residuo') per ipotesi distribuita secondo una normale: media nulla e varianza  $\sigma^2$ ;  $\mathbf{X}$  è la matrice (detta 'matrice disegno') formata dalle  $n$  osservazioni sulle variabili  $X_1, X_2, \dots, X_k$ , e, nel caso si desideri stimare anche l'intercetta, da un vettore colonna unitario:

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix}$$

Obiettivo dell'analisi di regressione è fornire una stima  $\mathbf{b}^{\wedge}$  dei parametri  $\mathbf{b}$  del modello, sulla base delle  $n$  osservazioni disponibili. La stima dei minimi quadrati è il valore  $\mathbf{b}^{\wedge}$  che si ottiene minimizzando rispetto a  $\mathbf{b}$  l'espressione:

$$S(\mathbf{b}) = (\mathbf{y} - \mathbf{Xb})^T (\mathbf{y} - \mathbf{Xb}) = \mathbf{e}^T \mathbf{e}$$

che è la somma dei quadrati dei residui. E' possibile, una volta ottenuti tali valori, è possibile quantificare:

- I livelli di ciascuna caratteristica che massimizza l'utilità dei clienti;
- L'importanza relativa delle caratteristiche del servizio considerato, determinata dagli scarti d'utilità, ottenuti calcolando, per ciascun attributo, la differenza di valore - utilità esistente tra il livello migliore e quello peggiore. Rapportando poi, tale differenza alla sommatoria di tutti gli scarti d'utilità, si ottengono i valori dell'importanza relativa per il campione intervistato.
- I profili d'offerta che meglio si adattano a gusti ed esigenze della domanda.

Il modello, se utilizzato sistematicamente, consente di:

- I. Identificazione del potenziale competitivo e del suo grado di sfruttamento, connesso alle caratteristiche rilevanti dell'offerta;
- II. Determinare un indice sintetico del livello di CS raggiunto dall'impresa, idoneo a favorire la programmazione e il controllo delle azioni finalizzate all'incremento delle risorse di fiducia.
- III. Valutazione della solidità della posizione di mercato rispetto all'offerta dei concorrenti più significativi, a livello aggregato e disaggregato; possiamo reiterare il procedimento della tabella 2.3., rapportando l'indice sintetico di CS ottenuto dall'impresa a quello di specifici concorrenti. Si ottiene in tal modo un indicatore complessivo del vantaggio competitivo nell'ottica della domanda, disaggregabile con riferimento alle singole dimensioni di scelta.
- IV. Misurazione del valore di un determinato servizio: agli intervistati sono sottoposti servizi alternativi dei quali si chiede un giudizio di preferenza, che rappresentano i

prodotti alternativi, sui quali è effettuata la rilevazione delle percezioni degli acquirenti.

La rilevazione, si estrinseca nei seguenti passaggi:

- Raccogliere le valutazioni dei clienti, sotto forma di punteggi su una scala a intervalli (conjoint analysis metrica) o di classifica ranking delle preferenze (conjoint analysis non metrica), riguardo le “card” (profili di servizio) che descrivono le possibilità d’offerta;
- Disaggregazione del valore globalmente percepito (l’intenzione d’acquisto), nella fase d’elaborazione, in valori d’utilità parziale, ossia nella parte del valore complessivo generata per il cliente per ciascun livello di prestazione presentato nelle card, in modo che le valutazioni globali originarie possano essere correttamente ricostruite sommando i valori parziali dell’insieme di livelli che descrivono ciascun prodotto;

Il **processo d’elaborazione multivariato** consiste nello sviluppo di un sistema di regressioni multiple per ogni cliente intervistato, costituito dall’insieme delle valutazioni espresse da ciascun cliente, (ogni regressione rappresenta un giudizio formulato su un profilo di prodotto- card). Ogni valutazione (sotto forma di classifica o di punteggio) rappresenta una variabile dipendente il cui valore è spiegato da  $n$  variabili indipendenti, quanti sono i livelli degli attributi descritti in ciascuna card.

La logica del procedimento, considera la preferenza del cliente la variabile che dipende dalle singole caratteristiche del servizio, oggetto di valutazione; e, pertanto, definisce il valore complessivo come funzione dei valori parziali (attribuiti ai livelli di performance offerta su ciascuna caratteristica). Uno degli assunti fondamentali dell’analisi, riguarda il consumatore, che opererebbe una scelta fra gli attributi presenti nelle diverse scelte d’acquisto.

La soluzione del sistema di regressioni multiple è rappresentata dai coefficienti beta che, per ciascun profilo di prodotto misurano il valore, implicitamente percepito dal cliente, nei livelli di prestazione, e nei diversi attributi del servizio. Tali valori sono disponibili a livello individuale, oppure quali medie dei beta di tutti i clienti intervistati.

Ulteriori, elaborazioni utili, si hanno rapportando lo scarto d’utilità parziale dei livelli, minimo e massimo, di ciascun attributo al totale degli scarti d’utilità, si può cogliere in

che misura l'incremento di prestazione su una determinata caratteristica possa incidere sul valore percepito dal cliente. La misura dell'importanza relativa per l'attributo  $x$ ,  $I(x)$  è pari a:

$$I(x) = \frac{U_{\max}(x) - U_{\min}(x)}{\sum_{i=1}^n U_{\max}(i) - U_{\min}(i)}$$

Quanto maggiore è la variazione d'utilità complessiva conseguente al cambiamento nel livello di prestazione di un attributo, tanto maggiore sarà la sua rilevanza nelle preferenze e nella percezione di valore da parte del cliente. Inoltre, sempre dai dati delle utilità parziali, è possibile quantificare il valore monetario di ciascun attributo e dell'eventuale differenziale di performance: giacché i valori parziali sono espressi secondo la medesima scala di misura; alla presenza di variabili quantitative, è possibile convertire il loro valore nella scala delle utilità. Il valore dell'utilità unitaria in termini di prezzo è calcolabile nel modo seguente:

Valore monetario	$= \frac{P_{\max} - P_{\min}}{U_{\max}(P_{\min}) - U_{\min}(P_{\max})}$
dell' utilità Unitaria	(prezzo massimo- prezzo minimo) / (Utilità massima – Utilità minima)

Si potrà così, simulare, l'impatto che un miglioramento nelle prestazioni del servizio produce sulla percezione del valore da parte del cliente, e conseguentemente sul potenziale competitivo dell'impresa.

La potenzialità informativa della tecnica, può essere incrementata significativamente qualora alla sua applicazione sia combinata la cluster analysis. I valori che emergono dall'analisi congiunta, infatti, sono parametri di regressione valutati a livello di singolo cliente e sui quali il calcolo delle medie ha frequentemente scarso significato a causa d'elevati livelli di varianza. In tali casi, avendo preventivamente accertato l'esistenza di una significativa varianza tra le preferenze; sulla matrice conjoint dei risultati, (che riassume l'utilità assegnata da ciascun individuo alle varie caratteristiche dell'offerta), è

opportuno applicare la cluster analysis. Dopo aver preventivamente descritto i cluster attraverso l'analisi di *cross-tabulation* condotte sulle variabili d'interesse.

**I limiti dell'analisi congiunta** si possono considerare:

- Impossibilità di verificare sperimentalmente tutte le possibili combinazioni sui livelli di prestazione del servizio, col conseguente, rischio di tralasciare attributi indispensabili all'indagine, che provoca una riduzione del suo potenziale informativo;
- Difficoltà a descrivere gli attributi simbolici, e quindi scarsa efficacia della tecnica per la misurazione di questi;

Alcuni di questi limiti, sono superabili combinando i valori parziali rilevati mediante la *conjoint* con quelli delle misurazioni alla *Fishbein* (con riferimento alle percezioni di performance delle singole marche sui diversi attributi), in quanto consente di determinare quale valore differenziale è attribuito ad una determinata caratteristica, in funzione della fiducia verso l'impresa che lo offre sul mercato.

Tenuto conto che le tecniche di misurazione forniscono indicazioni parziali, e che la soddisfazione espressa è un giudizio complesso formato dall'operare congiunto di più determinanti; si comprende che per misurare efficacemente tutti gli aspetti della CS è necessario l'utilizzo contestuale di più strumenti con l'applicazione delle tecniche d'analisi multivariata. (Analisi sistematica ed integrata composta sia da studi qualitativi che quantitativi).

L'atteggiamento complessivo nei confronti di un prodotto viene pertanto a dipendere:

- Dall'importanza relativa di ciascun attributo;
- Dall'insieme dei giudizi formulati dal consumatore sulla presenza e posizione occupata dal prodotto stesso con riferimento agli attributi percepiti;

Da qui la necessità di ricorrere a metodologie d'analisi in grado di cogliere i *trade-off cognitivi*. Il cliente, definisce le proprie preferenze prendendo in considerazione le caratteristiche rilevanti del prodotto; la considerazione separata dei diversi parametri valutativi non consente all'impresa di determinare gli effetti che l'aumento o riduzione di un certo attributo è suscettibile di esercitare sul valore percepito del prodotto, ed in quale misura tali costi aggiuntivi possano essere assorbiti da aumenti di prezzo.

E' proprio per questo che il nostro studio propone un'applicazione metodologica in cui le caratteristiche delle preferenze sono considerate interattive. Ma di questo parleremo in seguito.



## CAPITOLO 8 -

### **L'UTILIZZO DI OPERATORI DI AGGREGAZIONE FUZZY**

#### 8.1 DEFINIZIONE DEGLI OPERATORI DI AGGREGAZIONE

Nella teoria delle decisioni multicriteriali, ed in particolar modo nell'oggetto della nostra dissertazione relativo alla Customer Satisfaction, la scelta di un appropriato operatore di aggregazione che sintetizzi un insieme di dati già convertiti in valori numerici in un unico valore, assume un ruolo molto importante. A tali operatori si richiedono usualmente alcune proprietà di base che caratterizzano in maniera stringente il tipo di aggregazione proposto. Storicamente, la classe più impiegata di operatori di tal sorta è costituita dalle medie, la cui conoscenza teorica è stata quasi completamente "coperta" dal celeberrimo teorema di rappresentazione di Kolmogorov. Successivamente, la letteratura ha evidenziato la necessità di definire operatori sempre più deboli dal punto di vista delle proprietà assiomatiche richieste. Recentemente, sono state studiate alcune nuove classi di aggregatori, tra cui, proposti da Yager, i cosiddetti ordered weighted averaging (OWA) operators, i quali hanno riscosso molto interesse, sia per il fatto di generalizzare diversi esempi di operatori classici, sia perché si sono naturalmente collegati all'integrale di Choquet. In questo capitolo si propone una classificazione teorica di aggregatori, di cui si investigano le proprietà generali, cercando di catturare quelli classici come casi particolari. Si cerca poi di giustificare la loro rilevanza nel collegarli alla logica fuzzy.

Gli operatori di aggregazione sono oggetti matematici che hanno la funzione di ridurre l'insieme dei numeri in un unico elemento rappresentativo (o espressivo). Qui presentiamo un loro elenco. Ritraiamo le loro caratteristiche e vantaggi, ma proviamo ad essere obiettivi dando anche i loro svantaggi.

Nel prossimo paragrafo ci occupiamo delle regole matematiche che definiscono un operatore di aggregazione. Ciò è essenziale, perché ogni operatore matematico che trasforma un insieme dei numeri in un unico valore, necessariamente non dà un valore finale rappresentativo o espressivo.

In un senso piuttosto informale, il problema di aggregazione è costituito nell' aggregamento di n-uple di oggetti appartenenti ad un dato insieme, in un singolo oggetto dello stesso insieme. Nel caso dell' operatore matematico di aggregazione questo insieme è tutti i numeri reali. In questo modo, un operatore di aggregazione è semplicemente una funzione  $\text{Aggreg} : \mathfrak{R}^n \rightarrow \mathfrak{R}$ , che assegna un numero reale  $y$  a una n-upla  $(x_1, x_2, \dots, x_n)$  di numeri reali:

$$y = \text{Aggreg} (x_1, x_2, \dots, x_n)$$

Naturalmente, dovremmo imporre determinate proprietà su  $\text{Aggreg}(\cdot)$  per giustificare il nome "dell' operatore di aggregazione". Parecchi autori hanno proposto un insieme delle circostanze fondamentali che definiscono gli operatori di aggregazione. E' da notare che queste definizioni di base non sono compatibili. Recentemente, Mesiar e Komorníková (Mesiar et al., 1997) hanno proposto un insieme delle proprietà fondamentali che raggruppano tutte le definizioni precedenti nelle circostanze più deboli.

Così, definiamo un operatore di aggregazione come funzione

$$\text{Aggreg}: \prod_{i \in N} [0,1] \rightarrow [0,1]$$

Che soddisfa:

$$\text{Aggreg} (\mathbf{x}) = x \quad \text{dove } \mathbf{x} = [x_1, x_2, \dots, x_n] \quad \forall x \in [0,1] \quad \textbf{idempotenza}$$

$$\text{Aggreg} (0, 0, \dots, 0) = 0 \text{ e } \text{Aggreg} (1, 1, \dots, 1) = 1 \quad \textbf{condizioni di frontiera}$$

$$\text{Aggreg} (x_1, x_2, \dots, x_n) \leq \text{Aggreg} (y_1, y_2, \dots, y_n) \quad \textbf{monotonia}$$

$$\text{se } (x_1, x_2, \dots, x_n) \leq (y_1, y_2, \dots, y_n)$$

Nel paragrafo seguente presentiamo una descrizione delle proprietà che possiamo attendersi da un operatore di aggregazione.

Per più particolari, si veda Fodor e Roubens (1994) e Grabisch (1995,1996) .

## 8.2 PROPRIETA' MATEMATICHE

Abbiamo definito un operatore di aggregazione come funzione

$$\text{Aggreg: } \prod_{i \in N} [0,1]^n \rightarrow [0,1] \quad (1)$$

Vediamo ora le proprietà matematiche (J.L. Marichal 1999).

♣ *Le condizioni a contorno.*

Prevediamo che un operatore di aggregazione soddisfi:

$$\text{Aggreg} (0, 0, \dots, 0) = 0 \quad (2)$$

$$\text{Aggreg} (1, 1, \dots, 1) = 1 \quad (3)$$

La condizione (2) significa che se abbiamo valutazioni soltanto completamente cattive, o false o i criteri sono non soddisfacenti l' aggregazione totale deve essere anche rispettivamente completamente cattiva, falsa o non soddisfacente. La (3) traduce che se osserviamo valutazioni soltanto vere o criteri completamente soddisfacenti allora l' aggregazione totale deve essere anche vera o soddisfacente. Come Mesiar e Komorníková (Mesiar et al., 1997) hanno precisato, questa proprietà sembra essere fondamentale nella definizione degli operatori di aggregazione. Un'estensione di questa condizione di base è stata proposta. Per esempio Mayor e Trillas (Mayor e Trillas, 1986, Trillas 1983) propongono come stato fondamentale per un operatore di aggregazione quanto segue:

$$\forall x \in [0,1] \text{ Aggreg} (x,0) = \text{Aggreg} (1,0) \cdot x \quad (4)$$

$$\forall x \in [0,1] \text{ Aggreg} (x,1) = (1 - \text{Aggreg} (1,0)) \cdot x + \text{Aggreg} (1,0) \quad (5)$$

Notiamo che la (4) richiede che il valore  $\text{Aggreg} (x,0)$  sia la media aritmetica ponderata della  $x$  e di 0; nello stesso senso,  $\text{Aggreg} (x,1)$  è la media aritmetica ponderata della  $x$  e di 1. Queste due circostanze restringono in uno più piccolo il gruppo degli operatori di aggregazione. In effetti (2) e (3) sono casi particolari per  $x=0$  e  $x=1$  rispettivamente di (4) e (5).

♣ *Monotonia (non decrescente)*

Ci occupiamo più precisamente della non decrescenza riguardo a ciascuna variabile.

Prevediamo che se un argomento aumenta allora l' aggregazione finale aumenti (o almeno non diminuisce, cioè rimane costante):

$$y_i \geq x_i \Rightarrow \text{Aggreg} (x_1, \dots, y_i, \dots, x_n) \geq \text{Aggreg} (x_1, \dots, x_i, \dots, x_n) \quad (6)$$

♣ *Continuità*

La funzione  $\text{Aggreg}(\cdot)$  è continua riguardo a ciascuna delle relative variabili. Questa proprietà è una garanzia per determinare la robustezza, la consistenza e per assicurare un comportamento non caotico.

♣ *Proprietà Associativa*

Questa proprietà assicura che la scelta di aggregazione a gruppi non abbia influenza sul risultato.

Per tre argomenti la proprietà può essere scritta:

$$\text{Aggreg} (x_1, x_2, x_3) = \text{Aggreg} (\text{Aggreg} (x_1, x_2), x_3) = \text{Aggreg} (x_1, \text{Aggreg} (x_2, x_3)) \quad (7)$$

Questa proprietà permette di definire completamente l'operatore quando questo è definito soltanto per due elementi.

♣ *Simmetria*

Tale proprietà è inoltre conosciuta come commutativa.

L' ordine delle argomenti non ha influenza sul risultato. Questa proprietà è obbligatoria quando l' aggregazione è fatta su argomenti che hanno la stessa importanza.

Per ogni permutazione  $\sigma$  di  $\{1, 2, \dots, n\}$  l' operatore deve soddisfare:

$$\text{Aggreg} (x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}) = \text{Aggreg} (x_1, x_2, \dots, x_n) \quad (8)$$

♣ *Bisimmetria*

Bisimmetria è una proprietà associata all' aggregazione di  $2n$  dati per un  $n$ -upla di operatori.

Se scriviamo questi dati in una matrice quadrata, allora la bisimmetria traduce il fatto che non importa se l'aggregazione avviene per vettori colonna o riga.

Per un operatore binario  $A$  questo significa che per ogni  $x_{11}, x_{12}, x_{21}, x_{22}$ , si ha:

$$A(A(x_{11}, x_{12}), A(x_{21}, x_{22})) = A(A(x_{11}, x_{21}), A(x_{12}, x_{22})) \quad (9)$$

Si noti che se un operatore è commutativo ed associativo allora è necessariamente bisimmetrico, tuttavia, non vale il contrario.

♣ *Elemento assorbente*

Si definisce elemento assorbente un valore  $a \in [0,1]$  tale che

$$\text{Aggreg}(x_1, \dots, a, \dots, x_n) = a \quad (10)$$

♣ *Elemento neutro*

Si definisce elemento neutro un valore  $e \in [0,1]$  tale che

$$\text{Aggreg}^{(n)}(x_1, \dots, e, \dots, x_n) = \text{Aggreg}^{(n-1)}(x_1, \dots, x_n) \quad (11)$$

Ove  $\text{Aggreg}^{(s)}(x_1, \dots, x_n)$  è l'operatore relativo all'argomento  $s$  con  $s=1,2,\dots, n$

♣ *Idempotenza*

Tale proprietà è conosciuta come unanimità o accordo. L'idempotenza rappresenta molto semplicemente l'idea che ogni operatore, di una serie di dati tutti uguali, restituisca lo stesso valore.

Formalmente essa si definisce:

$$\text{Aggreg}(x, x, \dots, x) = x \quad (12)$$

$$\forall x \in [0,1]$$

In realtà, anche se in rarissimi casi, in letteratura si trovano esempi di aggregatori non idempotenti (Zimmermann), la cui presenza è giustificata dal loro buon “funzionamento” in alcuni casi concreti, senza alcuna attenta analisi teorica.

#### ♣ *Compensazione*

Tale proprietà è conosciuta come proprietà di Pareto.

Tale proprietà impone che il risultato dell' aggregazione sia più basso del più alto elemento aggregato (il massimo) e più alto di quello più basso (il minimo):

$$\min_{i=1}^n (x_i) \leq \text{Aggreg}(x_1, x_2, \dots, x_n) \leq \max_{i=1}^n (x_i) \quad (13)$$

Questa proprietà non deve essere confusa con la proprietà del contrappeso

#### ♣ *Contrappeso*

Tale proprietà è denominata della compensazione da alcuni autori. Ciò significa che una certa confusione può determinarsi con la proprietà precedente. Denominiamo la proprietà del contrappeso, il comportamento di un operatore che fa diminuire il risultato finale se ci sono argomenti che entrano in un senso opposto.

$$\forall t \in ]0,1[ , \text{ e } \forall (x_1, x_2, \dots, x_n) \in [0,1]^n \quad \exists (y_1, y_2, \dots, y_m) \in [0,1]^m$$

$$\text{tale che} \quad \text{Aggreg}(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) = t \quad (14)$$

#### ♣ *Proprietà del “rinforzo”*

Una caratteristica di molti tipi di elaborazioni dell' informazione umane, che sono state precisate in maniera sconvolgente da Elkan (1994), Rybalov e di Yager (1998) è il “rinforzo completo”. Questa proprietà significa che un insieme di valori molto affermativi implica una risultante più affermativa, e di conseguenza, la presenza di numerosi valori bassi, rinforza nel dare un segno di "disconferma". Il primo concetto è denominato rinforzo ascendente ed il secondo concetto è denominato rinforzo

discendente. Yager mostra che le t-norme hanno soltanto un comportamento in discesa di rinforzo, mentre le t-conorms hanno soltanto un comportamento ascendente di rinforzo. Inoltre indica che le uninorme hanno un comportamento completo di rinforzo. Questa proprietà può essere molto interessante. Per esempio, nella diagnosi medica la presenza di un certo numero di sintomi indicativi di una malattia li renderà più sicuri nella diagnostica del paziente che soffre di una certa la malattia che tutti i sintomi da soli mentre l'assenza di questi sintomi renderà più sicuro diagnosticare un paziente non soffre più della malattia.

♣ *La stabilità per una funzione lineare*

Questa proprietà traduce una stabilità dell' operator nel caso di cambiamento della scala di misura:

$$\text{Aggreg} (r \cdot x_1 + t, r \cdot x_2 + t, \dots, r \cdot x_n + t) = r \cdot (\text{Aggreg} (x_1, x_2, \dots, x_n)) + t \quad (15)$$

♣ *L' invarianza*

Si ha quando per i valori aggregati  $(x_1, x_2, x_n)$  rappresentativi della misura di determinati criteri, dovremmo specificare una scala in cui queste misure sono state realizzate. Inoltre, possiamo considerare la funzione Aggreg di aggregazione rispettare un rapporto espressivo riguardo alla data scala.

La nozione della piena significatività è formalizzata nel teorema di rappresentazione di Kolmogorov come la proprietà di invarianza.

Per ogni trasformazione ammissibile  $f$ , abbiamo:

$$\text{Aggreg} (f(x_1), f(x_2), \dots, f(x_n)) = f(\text{Aggreg} (x_1, x_2, \dots, x_n)) \quad (16)$$

### 8.3 OPERATORI MATEMATICI

Nei prossimi paragrafi, presentiamo una descrizione dei principali operatori. Spieghiamo le loro proprietà e particolarità principali. Inoltre presentiamo alcuni casi particolari notevoli. Cominciamo presentando alcuni degli operatori di aggregazione più usati. Tra gli operatori di base troviamo il prototipo di un operatore di aggregazione, la media, ma inoltre troviamo la mediana, il minimo ed il massimo, così come alcune generalizzazioni classiche come la media pesata e le statistiche di K-ordine. Continuiamo il paragrafo presentando gli operatori di media quasi-aritmetica, una grande famiglia utile costruita su una trasformazione dell' operatore medio. Quindi presentiamo una generalizzazione della media pesata e la media pesata ordinata (OWA), che ha inoltre come caso particolare il minimo ed il massimo. Ciò ci conduce agli integrali fuzzy discreti: Choquet (Choquet, 1953) e Sugeno (Sugeno, 1974). L' integrale di Choquet generalizza gli operatori di OWA, mentre il Sugeno generalizza il massimo pesato e gli operatori minimi pesati. Tutti questi operatori danno un valore rappresentativo "di mezzo" dell' insieme aggregato. Ci occupiamo altresì di due famiglie specializzate sull' aggregazione sotto incertezza: le  $t$ -norme ed i  $t$ -conorms. Questi operatori non cercano di dare "un valore centrale", ma preferibilmente computano l' intersezione e l' unione (rispettivamente) degli insiemi sfocati. Questi operatori sono usati spesso, poiché possono inoltre essere visti come generalizzazione degli operatori logici di aggregazione: "coniunzione" ( $t$ -norme) ed "disgiunzione" ( $t$ -conorms). Un'altro genere di operatori è comparso quando si rilascia l' assioma che differenzia la  $t$ -norma ed la  $t$ -conorm: gli uninorms. Questi operatori risolvono un altro problema delle  $t$ -norme e dei  $t$ -conorms, che è la mancanza (verso il basso e verso l' alto) di rinforzo completo. Proviamo qui a presentare una descrizione obiettiva del dominio, presentando le caratteristiche, dei vantaggi e degli svantaggi di ogni operatore. Un'ottima descrizione è disponibile da Bouchon-Meunier (1997). Si vedano, altresì, Calvo e Mesiar (1999a;1999b); Calvo, De Baets e Fodor, Detyniecki e Yager (1999, 1998a,1998b,1998c); Detyniecki, Yager e Bouchon-Meunier (1999,2000); Detyniecki, Yager e Bouchon-Meunier (2000a,2000b); Dombi (1982); Dubois e Prade (1981,1986a,1986b,1988a,1988b,1988c,1988d,1985,1999); Dyckhoff e Pedrycz (1984); Frank (1979); Kelman e Yager (1999); Klement, Mesiar e Pap (1996,2000); Klement



(1982); Klir et al. (1988,1995); Komornikova (1999); Kolesarova e Komornikova (1999); Luhandjula (1982); March (1988); Marichal (1999); Mesiar e Komornikova (1997); Mesiar (1995); Ovchinnikov (1998); Prade (1985); Silvert (1979); Weber (1983); Murofushi e Sugeno (1991); Murofushi e Soneda (1993); Grabisch (1996), Roubens (1978,1982,1996).

#### 8.4 LA LOGICA FUZZY

L'abilità di prendere decisioni precise e significative diminuisce al crescere della complessità del sistema fino al raggiungimento di una soglia oltre la quale precisione e significatività sono concetti quasi mutuamente esclusivi.

La logica fuzzy, o logica sfumata, introdotta da Lofti A. Zadeh nei primi anni sessanta si basa sui principi formali del ragionamento approssimato, riuscendo così a modellizzare le capacità proprie della mente umana di prendere decisioni razionali in un ambiente di incertezza e imprecisione.

L'approccio alternativo di tipo qualitativo proposto dalla fuzzy logic si fonda sul presupposto che gli elementi chiave del pensiero umano non sono numeri ma valori linguistici che identificano **"fuzzy sets"**, cioè classi di oggetti per i quali il passaggio dalla appartenenza alla non appartenenza è graduale e non drastico come per il caso della logica booleana.

La Fuzzy Logic affronta il problema di assegnare valori di verità ad affermazioni per le quali non sia possibile stabilire inequivocabilmente la veridicità o falsità e per le quali risulti quindi insufficiente la logica classica a due valori.

La peculiarità di questa logica risiede nella possibilità che essa offre di lavorare su sistemi di cui non si ha descrizione di tipo analitico ma per i quali sia fornita una caratterizzazione di tipo linguistico, cioè qualitativa.

La logica sfumata ricopre un ruolo fondamentale in uno degli aspetti caratteristici del pensiero umano quale l'abilità di sintetizzare l'informazione, ossia di estrarre da una collezione di dati solamente quelli rilevanti ai fini del raggiungimento dell'obiettivo prefissato.

Per natura una sintesi rappresenta un'approssimazione di quanto si sta sintetizzando. In molti casi una caratterizzazione approssimata di una collezione di dati è sufficiente poiché molti obiettivi umani non richiedono un alto grado di precisione. La mente umana sfrutta il vantaggio che può derivare da questa tolleranza dell'imprecisione,

codificando le informazioni mediante fuzzy sets che costituiscono un' approssimazione dei dati reali.

L'abilità di manipolare fuzzy sets e la conseguente capacità di sintesi caratterizzano quindi il pensiero umano.

Le caratteristiche della fuzzy logic sono:

- L'uso di variabili linguistiche in luogo di variabili numeriche;
- caratterizzazione di relazioni tra variabili attraverso frasi condizionali fuzzy;

Sia  $U$  un insieme di oggetti, discreto o continuo. Questi d'ora in poi sarà chiamato *Universo del discorso*. Indicheremo inoltre con  $u$  un generico elemento appartenente a  $U$ .

Un sottoinsieme  $X$ , nel senso "classico" del termine, di  $U$  è definito dalla sua funzione caratteristica  $\chi_X$ , che vale 1 se  $u \in X$ , 0 se  $u \notin X$ .  $X$  è anche detto insieme Crisp.

Nel caso di un insieme fuzzy, la funzione caratteristica viene sostituita dalla cosiddetta *funzione di appartenenza* che può assumere valori compresi tra 0 e 1 per rappresentare il diverso grado di appartenenza degli elementi di un dato insieme. La funzione caratteristica può essere considerata un caso particolare della funzione di appartenenza.

Si definisce allora il fuzzy set nel modo seguente:

Un fuzzy set  $F$  in un universo del discorso  $U$  è un insieme di coppie ordinate:

$F = \{(u, \mu_F(u)) \mid u \in U\}$ , dove  $\mu_F: U \rightarrow [0,1]$  è la funzione di appartenenza di  $u$  in  $F$ .

Se  $U$  è discreto, un fuzzy set è rappresentato mediante la notazione  $\sum_{i=1}^n \mu_F(u_i) / u_i$ , se

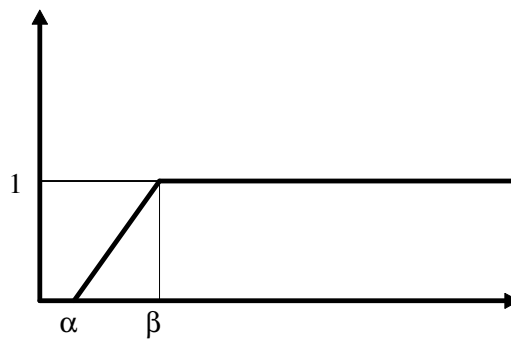
invece  $U$  è continuo si utilizza la notazione  $\int_U \mu_F(u) / u$ .

Se la funzione caratteristica assume valore 1,  $u$  appartiene completamente all' insieme, se viceversa assume valore 0,  $u$  non appartiene all' insieme. Per questo generalmente si tralasciano gli elementi con grado di appartenenza 0.

Di funzioni di appartenenza ce ne sono parecchie. Vediamo le più comuni:

1. Funzione  $\Gamma$ . E' una funzione a due parametri definita come segue:

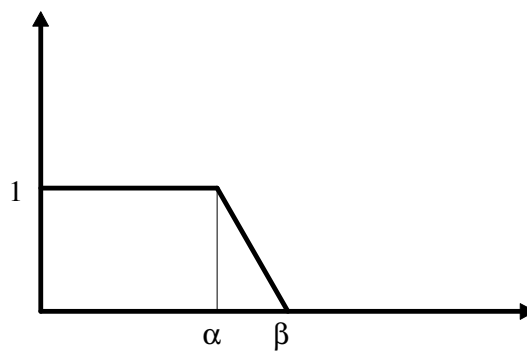
$$\Gamma = \begin{cases} 0 & \text{se } u < \alpha \\ \frac{u - \alpha}{\beta - \alpha} & \text{se } \alpha \leq u \leq \beta \\ 1 & \text{se } u > \beta \end{cases}$$



Funzione di appartenenza  $\Gamma$ .

2. Funzione L. Anche questa è una funzione a due parametri, definita come segue:

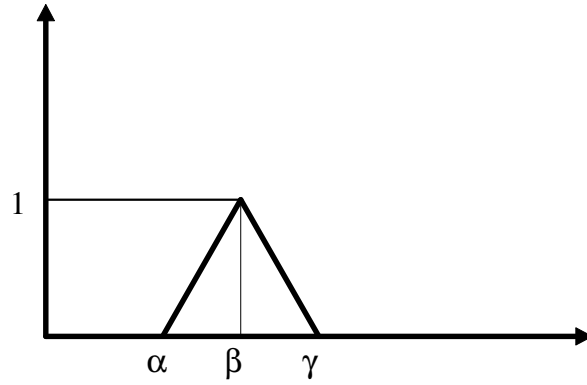
$$L(u, \alpha, \beta) = \begin{cases} 1 & \text{se } u < \alpha \\ \frac{\beta - u}{\beta - \alpha} & \text{se } \alpha \leq u \leq \beta \\ 0 & \text{se } u > \beta \end{cases}$$



Funzione di appartenenza L

3. Funzione  $\Lambda$ , o triangolare. In questo caso i parametri diventano tre e la definizione è:

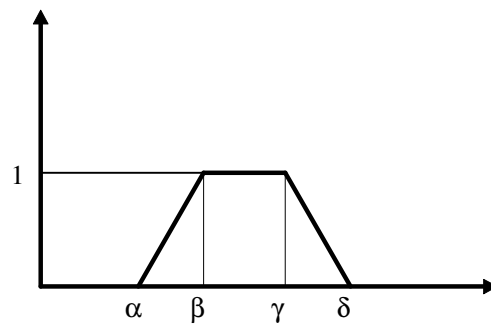
$$\Lambda(u, \alpha, \beta, \gamma) = \begin{cases} 0 & \text{se } u < \alpha \\ \frac{u - \alpha}{\beta - \alpha} & \text{se } \alpha \leq u \leq \beta \\ \frac{\alpha - u}{\beta - \alpha} & \text{se } \beta < u \leq \gamma \\ 0 & \text{se } u > \gamma \end{cases}$$



Funzione di appartenenza triangolare

4. Funzione  $\Pi$ . Questa funzione è dotata di quattro parametri, definita come:

$$\Pi(u, \alpha, \beta, \gamma, \delta) = \begin{cases} 0 & \text{se } u < \alpha \\ \frac{u - \alpha}{\beta - \alpha} & \text{se } \alpha \leq u \leq \beta \\ 1 & \text{se } \beta < u \leq \gamma \\ \frac{\gamma - u}{\delta - \gamma} & \text{se } \gamma < u \leq \delta \\ 0 & \text{se } u > \delta \end{cases}$$



Funzione di appartenenza  $\Pi$

Il **Supporto** di un fuzzy set  $F$  è l'insieme di tutti i punti tali che la funzione di appartenenza è strettamente positivo. Cioè

$$\text{Supp}(F) = \{u \in U \mid \mu_F(u) > 0\}$$

Più in generale si definisce  $\alpha$ -supporto di  $F$  l'insieme di tutti i punti  $u$  di  $U$  tali che la funzione di appartenenza applicata agli  $u$  è maggiore di  $\alpha$ . Cioè:

$$\text{Supp}(F)_\alpha = F_\alpha = \{u \in U \mid \mu_F(u) > \alpha\}$$

Un insieme fuzzy è **singolo** se e solo se il suo supporto è costituito da un solo elemento (se cioè  $\text{Supp}(F)$  è un singleton).

Un fuzzy set è **convesso** se:

$$\forall x, y \in U, \forall \lambda \in [0, 1] \Rightarrow \mu_F(\lambda x + (1 - \lambda)y) \geq \min(\mu_F(x), \mu_F(y))$$

Il **nucleo** (kernel) di un fuzzy set  $F$  è l'insieme crisp (crisp set) che contiene tutti e soli gli elemento con grado di appartenenza 1, vale a dire:

$$\text{Ker}(F) = \{u \in U \mid \mu_F(u) = 1\}$$

Un Fuzzy set è **normale** se e solo se

$$\text{Ker}(F) \neq \emptyset$$

Un **numero fuzzy**  $F$  in un universo continuo  $U$  è un fuzzy set  $F$  in  $U$  normale e convesso.

Siano  $A$  e  $B$  due fuzzy set di uno stesso universo  $U$ . Le operazioni su di essi sono definite tramite le loro funzioni di appartenenza. Operazioni del tipo *uguaglianza* e *inclusione* si derivano dalla teoria classica degli insiemi.

**Uguaglianza:** Due fuzzy set  $A$  e  $B$  sono uguali se e solo se

$$\forall u \in U \Rightarrow \mu_A(u) = \mu_B(u)$$

**Inclusione:** Il fuzzy set A è contenuto nel Fuzzy set B se e solo se

$$\forall u \in U \Rightarrow \mu_A(u) \leq \mu_B(u)$$

Più difficile risulta derivare dalla teoria classica le definizioni per le operazioni di *unione*, *intersezione* e *complemento* per i fuzzy set. Zadeh propose le seguenti:

**Unione:** La funzione di appartenenza dell' unione è definita da

$$\mu_{A \cup B}(u) = \max(\mu_A(u), \mu_B(u))$$

**Intersezione:** La funzione di appartenenza è definita da

$$\mu_{A \cap B}(u) = \min(\mu_A(u), \mu_B(u))$$

**Complemento:** La funzione di appartenenza è definita da  $\mu_{\neg A} = 1 - \mu_A$

Nelle tre precedenti definizioni gli operatori max, min, e 1-x possono considerarsi corrispondenti alle operazioni logiche “è”, “e”, “non”.

Le tre operazioni possono essere definite come *norme* o *conorme triangolari*.

**Norma Triangolare:** Una norma triangolare è una funzione  $T: [0,1] \times [0,1] \rightarrow [0,1]$  tale che valgano le seguenti proprietà:

- $T(a,b)=T(b,a)$  **commutativa**
- $T(a,T(b,c))=T(T(a,b),c)$  **associativa**
- $T(a,b) \geq T(c,d)$  se  $a \geq c$  e  $b \geq d$  **monotonia**
- $T(a,1)=a$  **identità**

Per esempio, T può essere una delle seguenti funzioni:

intersezione:  $x \wedge y = \min(x, y)$

prodotto algebrico:  $x * y = xy$

prodotto limitato:  $x \otimes y = \max(0, x + y - 1)$

prodotto drastico:  $x \times y = \begin{cases} x & \text{se } y = 1 \\ y & \text{se } x = 1 \\ 0 & \text{se } x, y < 1 \end{cases}$

**Conorma Triangolare:** Una conorma triangolare è una funzione  $S:[0,1] \times [0,1] \rightarrow [0,1]$

tale che valgano le seguenti proprietà:

- $S(a,b)=S(b,a)$  **commutativa**
- $S(a,S(b,c))=(S(S(a,b),c))$  **associativa**
- $S(a,b) \geq S(c,d)$  se  $a \geq c$  e  $b \geq d$  **monotonia**
- $S(a,0)=a$  **identità**

Per esempio  $S$  può essere una delle seguenti funzioni:

unione:  $x \vee y = \max(x, y)$

somma algebrica:  $x \bullet y = x + y - xy$

somma limitata:  $x \oplus y = \min(1, x + y)$

somma drastica:  $x \text{ Y } y = \begin{cases} x & \text{se } y = 0 \\ y & \text{se } x = 0 \\ 1 & \text{se } x, y > 0 \end{cases}$

Una relazione (in senso classico)  $n$ -aria fra  $X_1, \dots, X_n$  è un sottoinsieme del prodotto cartesiano  $X_1 \times \dots \times X_n$ , cioè un insieme  $R$  di  $n$ -uple ordinate  $(x_1, \dots, x_n)$ , con  $x_1 \in X_1, \dots, x_n \in X_n$ .

Proprio come gli insiemi classici, le relazioni possono essere descritte da una funzione

caratteristica  $\chi_R : X_1 \times \dots \times X_n \rightarrow \{0,1\}$ , con la definizione

$$\chi_R(x_1 \dots x_n) = \begin{cases} 1 & \text{se } (x_1 \dots x_n) \in R \\ 0 & \text{se } (x_1 \dots x_n) \notin R \end{cases}$$

Come già detto per i rapporti tra insiemi classici e fuzzy, nelle relazioni fuzzy  $\chi_R$  viene sostituita da una funzione caratteristica  $\mu_R$ , estesa all'intervallo  $[0,1]$ .

Se  $A_1, \dots, A_n$  sono fuzzy set in  $U_1, \dots, U_n$ , il loro **prodotto cartesiano (cross product)**

è un fuzzy set in  $U_1 \times \dots \times U_n$  definito da una delle seguenti funzioni di appartenenza:

$$\mu_{A_1 \dots A_n}(u_1 \dots u_n) = \min(\mu_{A_1}(u_1), \dots, \mu_{A_n}(u_n))$$

$$\mu_{A_1 \dots A_n}(u_1 \dots u_n) = \mu_{A_1}(u_1) \cdot \dots \cdot \mu_{A_n}(u_n)$$



Una **relazione fuzzy** n-aria  $R$  è un fuzzy set in  $U_1 \times \dots \times U_n$  definito dalla funzione di appartenenza  $\mu_R : U_1 \times \dots \times U_n \rightarrow [0,1]$ .

In particolare se  $U$  e  $V$  sono universi continui e  $\mu_R : U \times V \rightarrow [0,1]$ , allora

$R = \int_{U \times V} \mu_R(u, v) / (u, v)$  è una relazione fuzzy binaria in  $U \times V$ . Se  $U$  e  $V$  sono discreti, allora  $R = \sum_{U \times V} \mu_R(u, v) / (u, v)$

Siano  $R$  e  $S$  due relazioni binarie definite su  $X \times Y$ . La funzione di appartenenza  $\mu_{R \cup S}$  dell' **unione**  $R \cup S$  è definita da:

$$\mu_{R \cup S}(x, y) = \max(\mu_R(x, y), \mu_S(x, y)) \quad \forall (x, y) \in X \times Y \quad (R \cup S)$$

Nulla vieta però di definire altri tipi di unione attraverso altre conorme triangolari

Siano  $R$  e  $S$  due relazioni binarie definite su  $X \times Y$ . La funzione di appartenenza  $\mu_{R \cap S}$  dell' **intersezione**  $R \cap S$  è definita da:

$$\mu_{R \cap S}(x, y) = \min(\mu_R(x, y), \mu_S(x, y)) \quad \forall (x, y) \in X \times Y \quad (R \cap S)$$

Nulla vieta di definire altri tipi di intersezione attraverso altre norme triangolari

Due operazioni importantissime sui fuzzy set sono la *proiezione* ( $\text{Proj } R \text{ su } Y$ ) e l'*estensione cilindrica* (*cylindrical extension*,  $ce(X)$ )

La proiezione trasforma una relazione ternaria in una relazione binaria, o una relazione binaria in un fuzzy set, o un fuzzy set in un valore crisp. In generale la proiezione trasforma relazioni n-arie in relazioni (n-1)-arie. Daremo per semplicità la definizione solo nel caso binario

La **proiezione** di  $R$  su  $Y$ , dove  $R$  è una relazione e  $Y$  un fuzzy set è definita da

$$\text{proj } R \text{ su } Y = \int_Y \sup_x \mu_R(x, y) / y \quad \text{nel caso continuo} \quad \text{e} \quad \text{da}$$

$$\text{proj } R \text{ su } Y = \sum_Y \max_x \mu_R(x, y) / y \quad \text{nel caso discreto.}$$

L' *estensione cilindrica* è più o meno l' inverso della proiezione. Estende fuzzy set a relazioni binarie fuzzy, relazioni binarie fuzzy a relazioni ternarie fuzzy, ecc. In generale

trasforma relazioni n-arie in relazioni (n+1)-arie.

Come nel caso precedente, daremo solo una definizione binaria dell'estensione cilindrica.

L'estensione cilindrica del fuzzy set  $F$  su  $X \times Y$  è l'insieme di tutte le coppie  $(x, y) \in$

$X \times Y$  con grado di appartenenza  $\mu_F(y)$ , cioè:  $ce(F) = \int_{X \times Y} \mu_F(y) / (x, y)$ .

Si ha che  $proj(ce(S))$  su  $V=S$ , ma in generale  $ce(proj R \text{ su } V) \neq R$ .

Sia  $A$  un fuzzy set definito su  $X$  e  $R$  una relazione fuzzy definita su  $X \times Y$ . La composizione di  $A$  ed  $R$  risultante in un fuzzy set  $B$  definito su  $Y$  è data da:

$$B = A \circ R = proj(ce(A) \cap R) \text{ su } Y$$

Se l'intersezione è l'operazione di minimo e la proiezione quella di massimo si ha:

$$\mu_B(y) = \max_X \min(\mu_A(x), \mu_R(x, y))$$

ed è chiamata composizione *max-min*. Se l'intersezione è l'operazione di prodotto si ha:

$$\mu_B(y) = \max_X (\mu_A(x) \mu_R(x, y))$$

ed è chiamata composizione *max-dot*.

Il principio di estensione (extension principle), è uno dei più importanti principi della teoria dei fuzzy sets. Esso permette in sostanza di estendere qualsiasi operazione su  $U$  a fuzzy sets, e fornisce un metodo generale per combinare tra loro concetti fuzzy e non fuzzy (per esempio fuzzy sets e relazioni).

**Principio di estensione:** Siano  $X$  e  $Y$  due insiemi crisp, e sia  $f$  una mappa (funzione) da  $X$  a  $Y$ .

Sia ora  $A$  un sottoinsieme fuzzy di  $X$ . Allora  $f(A)$  è il sottoinsieme fuzzy di  $Y$  tale che  $f(A) = \bigcup_X \{A(x) / f(x)\}$ . Se indichiamo con  $B=f(A)$ , allora  $B$  è il sottoinsieme fuzzy

di  $Y$  tale che  $B(y) = \max_{\text{per tutti } x \text{ tali che } f(x)=y} [A(x)]$

In molte situazioni noi siamo in grado di caratterizzare informazioni numeriche in modo impreciso. Per esempio usiamo termini come ‘circa 5’, ‘vicino a 0’, ‘più o meno 10’: questi sono esempi di numeri fuzzy.

L'applicazione dell' "extension principle" permette di definire le operazioni aritmetiche sui numeri fuzzy.

Consideriamo due fuzzy set A e B appartenenti rispettivamente ad U e V.

La **coniunzione** dei due fuzzy set A e B è definita dalla funzione di appartenenza:

$$\mu_{A \text{ and } B}(u, v) = \mu_A(u) * \mu_B(v), \text{ dove } * \text{ è una norma triangolare}$$

La **disgiunzione** dei due Fuzzy set A e B è definita dalla funzione di appartenenza:

$$\mu_{A \text{ or } B}(u, v) = \mu_A(u) + \mu_B(v), \text{ dove } + \text{ rappresenta una conorma triangolare.}$$

Una funzione di **implicazione** fuzzy è una funzione dove l' antecedente e il conseguente contengono variabili fuzzy.

## 8.5 T-NORMS E T-CONORMS

Ogni aggregatore è definito da una funzione  $h:[0,1]^n \rightarrow [0,1], n \geq 2$ .

La funzione di appartenenza dell'insieme fuzzy aggregato è :

$$\mu_A(x) = h(\mu_{A_1}(x), \mu_{A_2}(x), \dots, \mu_{A_n}(x)), \forall x \in X.$$

I tipi di aggregatori possibili sono:

### 1. T-Norma (intersezione)

Rappresentano l'intersezione fuzzy, e sono estensioni dell'intersezione crisp.

### 2. S-Norma(unione) o T-Conorma

Rappresentano l'unione fuzzy e sono estensioni dell'unione crisp

### 3. Operatori di media (media generalizzata, OWA, quantificatori linguistici)

Le proprietà e le leggi crisp possono essere rappresentate negli insiemi fuzzy attraverso

T-norm e S-norm. Ad esempio:

- Negazione

$$\bar{a} = 1 - a$$

- Idempotenza

$$\forall a \ T(a, a) = a, \ S(a, a) = a$$

- Legge del terzo escluso

$$\forall a \ S(a, \bar{a}) = 1$$

- Principio di non contraddizione

$$\forall a \ T(a, \bar{a}) = 0$$

- Proprietà archimedeica

$$\forall a \in (0,1) \ T(a, \bar{a}) < a, \ S(a, a) > a$$

- Nilpotenza

$\forall$  sequenza  $a_i \in (0,1)$   $i \in \mathbb{N}$

$\exists n_0 \in \mathbb{N}$  tale che  $T(a_1, a_2, \dots, a_{n_0}) = 0$

$\exists n_0 \in \mathbb{N}$  tale che  $S(a_1, a_2, \dots, a_{n_0}) = 1$

Non tutte le T-norm e le S-norm hanno queste proprietà. Vediamo degli esempi particolari.

Sia T un arbitrario operatore T-norm. L'operatore T-conorm  $S(a,b) = 1 - T(\bar{a}, \bar{b})$  è detto il duale di T.

Abbiamo allora i seguenti esempi:

<b>t-norm</b>	<b>t-conorm</b>	<b>Nome</b>
$\min(a,b)$	$\max(a,b)$	Min/Max (Zadeh)
$ab$	$a+b-ab$	Product/Probabilistic Sum
$\max(0, a+b-1)$	$\min(1, a+b)$	Bold Union, Bounded Sum

Come scegliere tra questi gli operatori adatti per fare l'unione e l'intersezione ?

Semplicemente si controllano le loro proprietà e le si confrontano con le nostre aspettative.

In questo caso ad ogni numero fuzzy da aggregare viene associato un peso  $\alpha \in [0,1]$  che rappresenta l'importanza dell'insieme nell'aggregazione.

Il calcolo dell'aggregazione avviene tramite una trasformazione degli insiemi da aggregare (se  $A_i$  sono gli insiemi da aggregare e  $\alpha_i$  sono i pesi, allora si opera così:

$B_i(x) = f(A_i(x), \alpha_i)$ ), dopodiché si esegue l'aggregazione sugli insiemi aggregati ( $A = R(B_1, \dots, B_n)$ ).

L'intersezione si attua seguendo la formula:  $A(x) = \bigwedge_{i=1}^n (S(A_i(x), 1 - \alpha_i))$

L'unione invece utilizza la formula:  $A(x) = \bigvee_{i=1}^n (T(A_i(x), \alpha_i))$

In entrambi i casi S e T sono operatori duali.

## 8.6 MEDIA, MEDIANA, MINIMO, MASSIMO

In molti casi il tipo di aggregazione desiderata oscilla tra due estremi. Per questo si introducono gli operatori di media. Questo operatore è interessante perché dà un valore aggregato che è più piccolo dell'argomento più grande e più grande di quella più piccolo. Così, l' aggregazione risultante è "un valore centrale". Questa proprietà è conosciuta come la proprietà della compensazione. E' usato spesso poiché è semplice e soddisfa le proprietà del monotonia, della continuità, della simmetria, della proprietà associativa, dell'idempotenza e della stabilità per le trasformazioni lineari. Ma non ha né elemento assorbente né neutro.

**Definizione:** aggregatore medio n-dimensionale. Una funzione  $G:[0,1]^n \rightarrow [0,1]$  è detto aggregatore medio di dimensione n se soddisfa i seguenti assiomi:

1. Commutatività
2. Monotonicità
3. Idempotenza

La soddisfazione di questi assiomi implica che:  $\min_i(a_i) \leq G(a_i) \leq \max_i(a_i)$ , dove  $G$  è l'aggregatore.

In questa famiglia vi sono la media generalizzata, la media generalizzata pesata e l' OWA (di quest'ultimo parleremo nel prossimo paragrafo). La media Generalizzata è una famiglia parametrica di operatori medi che copre l' intero intervallo tra min e max:

$$G_{\alpha}(a_1, \dots, a_n) = \left( \frac{a_1^{\alpha} + \dots + a_n^{\alpha}}{n} \right)^{\frac{1}{\alpha}}, \text{ con } \alpha \in (-\infty, +\infty).$$

Si hanno i seguenti casi particolari:

$$G_{\infty} = \text{Max}$$

$$G_1 = \text{media aritmetica}$$

$$G_0 = \text{media geometrica}$$

$$G_{-1} = \text{media armonica}$$

$$G_{-\infty} = \text{Min}$$

Ad ogni insieme fuzzy da aggregare viene associato un peso  $w \in [0,1]$ , che rappresenta l'importanza dell'insieme nell'aggregazione. Allora la media generalizzata diventa, media generalizzata pesata:

$$G_{\alpha}(a_1, \dots, a_n; w_1, \dots, w_n) = \left( \sum_{i=1}^n w_i a_i^{\alpha} \right)^{\frac{1}{\alpha}}, \text{ con } \alpha \in (-\infty, +\infty), \quad e \quad \text{dove}$$

$$\sum_{i=1}^n w_i = 1, \text{ cioè i pesi sono normalizzati.}$$

Un altro operatore che segue l'idea di ottenere "un valore centrale" è la mediana. È costituita nell'ordinamento delle argomenti da quello più piccolo a quello più grande. Si prende, quindi, l'elemento di mezzo. Se la cardinalità dell'insieme delle argomenti non è dispari allora prendiamo la media dei valori centrali. Questo operatore di aggregazione soddisfa le condizioni di contorno, la monotonia, la simmetria, l'idempotenza ed evidentemente il comportamento della compensazione. Esiste una generalizzazione di questo operatore: la statistica di K-ordine, con cui possiamo scegliere l'elemento sulla posizione del k-esima sulla lista richiesta..

Due casi particolari notevoli massimi della statistica di K-ordine sono il minimo ed il massimo. Il minimo dà il più piccolo valore di un insieme, mentre il massimo dà quello più grande. Sono operatori di aggregazione poiché soddisfano gli assiomi della definizione. Le proprietà principali di questi operatori sono: monotonia, simmetria, proprietà associativa, idempotenza. Usando questi operatori non otterremo mai un valore aggregato "nella metà". Se consideriamo un intervallo limitato  $[a, b]$  il minimo ha per l'elemento assorbente  $a$  e per l'elemento neutro  $b$ , mentre per il massimo sarà l'opposto: la  $a$  sarà l'elemento neutro e la  $b$  quello assorbente. È importante notare che il minimo ha un comportamento congiuntivo ed è una t-norma particolare. Il massimo è una t-conorm ed ha un comportamento disgiuntivo. Poiché la possibilità di dare i pesi è importante, Yager (1980,1981,1988,1994,1996,1997) ha introdotto il minimo pesato e il massimo pesato.

Per media Quasi-aritmetica si intende l'estensione della media aritmetica semplice e cioè la media geometrica, la media armonica. In effetti tutti questi operatori appartengono alla famiglia delle media quasi-aritmetica. Questa famiglia è stata studiata dettagliatamente da Kolmogorov (1930) e da Aczel (1948,1966).

## 8.7 AGGREGAZIONE ORDINATA PESATA (OWA)

Gli operatori di aggregazione ordinata pesata (OWA) originalmente sono stati introdotti da Yager per fornire i mezzi per l' aggregazione di valori connessi con la soddisfazione di criteri multipli, tali da unificare in un operatore il comportamento congiuntivo e disgiuntivo.

Un operatore Owa di dimensione  $n$  è una funzione  $f:[0,1]^n \rightarrow [0,1]$  a cui è associato

un vettore di pesi  $W = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$ , tale che

$$(1) w_i \in [0,1]$$

$$(2) \sum_{i=1}^n w_i = 1.$$

dove  $f(a_1, \dots, a_n) = \sum_{j=1}^n w_j b_j$ , e  $b_j$  è il  $j$ -esimo elemento più grande nella collezione

$a_1, \dots, a_n$  ( in pratica, prima di fare l' OWA devo ordinare in senso non otono gli  $a_i$ ).

Un aggregatore OWA è rappresentato dal suo vettore di pesi. Un peso non è associato ad un argomento, ma piuttosto ad una particolare posizione ordinata. Il passo di ordinamento quindi introduce non linearità nel processo di aggregazione. Gli OWA sono commutativi, monotoni, idempotenti, sono stabili per le trasformazioni lineari positive ed hanno un comportamento compensativo. Questa ultima proprietà traduce il fatto che l' aggregazione fatta da un operatore di OWA è sempre fra il massimo ed il minimo.

Osserviamo ora alcuni casi particolari:

$$-W^* = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \Leftrightarrow F^* = \max$$



$$-W_A = \begin{bmatrix} 1/n \\ 1/n \\ . \\ . \\ 1/n \end{bmatrix} \Leftrightarrow F_A = \frac{1}{n} \sum_i a_i, \text{ cioè la media in senso classico}$$

$$-W_* = \begin{bmatrix} 0 \\ 0 \\ . \\ . \\ 1 \end{bmatrix} \Leftrightarrow F_* = \min.$$

Da questi esempi possiamo notare che un aggregatore di tipo OWA si muove in modo continuo tra il min e il max, dunque tra l'and e l'or. Per classificare gli operatori in base alla loro posizione su questo continuo, si può introdurre una misura di orness (Yager) associata al vettore dei pesi:

$$\text{Orness}(W) = \frac{1}{n-1} \sum_{i=1}^n ((n-i)w_i)$$

$$\text{orness} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \text{orness} \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} = 0.5$$

Questo grado descrive la dispersione dei pesi ed è basato sull'idea di entropia.

La **dispersione** o **entropia** indica il grado con cui usiamo tutte le informazioni contenute negli argomenti quando calcoliamo il valore aggregato.

$$\text{Disp}(W) = \sum_{i=1}^n (w_i \cdot \ln(w_i))$$

Per una data misura di orness ci sono più vettori di pesi, con diversa dispersione, che possono essere usati.

## 8.8 GLI INTEGRALI FUZZY DI CHOQUET E SUGENO

L'integrale sfocato riguardo ad una misura sfocata è stato principalmente studiato in ambito multicriteriale. È basato sulla nozione di una misura sfocata, che può essere osservata come il peso di importanza di un insieme (J.L. Marichal, 1999). Definiamo matematicamente la misura sfocata come segue:

Sia  $C = \{ c_1, \dots, c_n \}$  l'insieme dei criteri e  $P(C)$  l'insieme potenza di, cioè l'insieme di tutti i sottoinsiemi della misura sfocata. Una misura fuzzy è una funzione dell'insieme  $\mu : P(C) \rightarrow [0,1]$ , soddisfacente i seguenti assiomi.

- 1)  $\mu(\emptyset) = 0$  e  $\mu(C) = 1$  (condizioni di frontiera).
- 2)  $\forall A \subseteq B \subseteq P(C) \quad \mu(A) \leq \mu(B)$  (condizioni di monotonia).

Nell'ambito di un problema di decisioni multicriteriali, una misura fuzzy  $\mu(A)$  si interpreta come il peso che rappresenta l'importanza attribuita dal decisore a ciascun sottoinsieme di criteri  $A$ .

Una misura fuzzy si definisce additiva se risulta  $\mu(A \cup B) = \mu(A) + \mu(B) \quad \forall A, B$  tale che  $A \cap B = \emptyset$ . In questo caso particolare, la misura fuzzy  $\mu(A)$  risulta univocamente determinata una volta che si sono attribuiti i valori alle  $n$  misure fuzzy  $\mu(c_1), \mu(c_2), \dots, \mu(c_n)$ . In generale, tuttavia, le misure fuzzy da stimare sono  $2^{|C|}$ , essendo  $|C|$  il cardinale di  $C$ , cioè una per ogni specifico sottoinsieme  $C$  dei criteri.

Questo genere di misura è più flessibile di una probabilità, che deve essere dotata della relativa proprietà di additività. Infatti, l'importanza di due criteri nella struttura di probabilità non può essere niente altro che la somma delle diverse importanze, mentre le misure sfocate possono fornire (un valore più grande (misura superadditiva) o più basso (misura subadditiva)). Ciò permette di modellare l'interazione fra i criteri. Ora, usando una misura sfocata introduciamo gli integrali sfocati.

L'integrale discreto di Sugeno delle valutazioni  $x_1, x_n$  per i criteri  $c_1, \dots, c_n$  riguardo ad una misura sfocata  $\mu$ , è definito

$$Sugeno_{\mu}(x_1, x_2, \dots, x_n) = \max_{i=1}^n (\min(x_{\sigma(i)}, \mu(C_{\sigma(i)})))$$

dove  $\sigma$  è una permutazione ordinata degli elementi:  $x_{\sigma(1)} \leq x_{\sigma(2)} \leq \dots \leq x_{\sigma(n)}$  e dove  $C_{\sigma(i)} = \{c_{\sigma(i)}, \dots, c_{\sigma(n)}\}$ .

Un' espressione alternativa rappresentativa dell'Integrale di Sugeno è la seguente proposta da Kandel and Byatt (Marichal,1999):

$$\begin{aligned} \text{Sugeno}_{\mu}(x_1, x_2, \dots, x_n) &= \max_{i=1}^n (\min(x_{\sigma(i)}, \mu(C_{\sigma(i)}))) = \\ &= \min_{i=1}^n (\max(x_{\sigma(i)}, \mu(C_{\sigma(i+1)}))) = \max_{T \subseteq N} (\min(\mu_T, \min(x_{\sigma(i)}))) = \\ &= \min_{i=1}^n (\max(\mu_{(N-T)}, \max_{i \in T} (x_{\sigma(i)}))) = \text{median}(x_1, x_2, \dots, x_n, \mu_{\{(2), \dots, (n)\}}, \mu_{\{(3), \dots, (n)\}}, \dots, \mu_{\{(n)\}}) \end{aligned}$$

L' integrale discreto di Choquet delle valutazioni  $x_1, \dots, x_n$  per i criteri  $c_1, \dots, c_n$  riguardo ad una misura sfocata  $\mu$ , è definito

$$\text{Choquet}_{\mu}(x_1, x_2, \dots, x_n) = \sum_{i=1}^n [x_{\sigma(i)} - x_{\sigma(i-1)}] \mu(C_{\sigma(i)})$$

dove  $\sigma$  è una permutazione ordinata degli elementi:  $x_{\sigma(1)} \leq x_{\sigma(2)} \leq \dots \leq x_{\sigma(n)}$  e dove  $C_{\sigma(i)} = \{c_{\sigma(i)}, \dots, c_{\sigma(n)}\}$  e  $x_{\sigma(0)} = 0$ .

Un'espressione equivalente è la seguente:

$$\text{Choquet}_{\mu}(x_1, x_2, \dots, x_n) = \sum_{i=1}^n x_{\sigma(i)} (\mu(C_{\sigma(i)}) - \mu(C_{\sigma(i+1)}))$$

Con  $C_{\sigma(n+1)} = \emptyset$ .

Gli integrali di Sugeno e di Choquet sono interessanti poiché sono operatori monotoni, continui, idempotenti, con comportamento di compensazione. L' integrale di Choquet è stabile sotto trasformazione lineare positiva, mentre l' integrale di Sugeno è stabile sotto una trasformazione affine con minimo e massimo che sostituisce il prodotto e la somma rispettivamente. Questa ultima proprietà ci permette di affermare che l' integrale di Sugeno è più adatto ad un'aggregazione ordinale (dove soltanto l' ordine degli elementi è importante) mentre l' integrale di Choquet è adatto ad un'aggregazione cardinale (dove la distanza fra i numeri ha un significato). La commutatività è ottenuta soltanto quando la misura sfocata dipende dalla cardinalità degli insiemi, cioè  $\mu(A) = \mu(B)$  se  $\text{card}(A) = \text{card}(B)$ . La proprietà associativa non è solitamente soddisfatta. La possibilità di generalizzazione degli integrali di Sugeno e di Choquet è notevole. Entrambi contengono, come casi particolari, le statistiche di ordine ed in particolare il minimo ed il massimo. L' integrale di Choquet generalizza le medie pesate e l' operatore di OWA, mentre l' integrale di Sugeno generalizza il minimo pesato ed il massimo pesato.

Il problema principale nell' uso dell' integrale di Sugeno o di Choquet (oltre al fatto che

non sono associativi e commutativi) è il numero  $2^n$  di pesi da determinare, ad un' aggregazione semplice di criteri di  $n$ . Questi pesi sono niente altro che la descrizione della misura sfocata. Un'interessante ricerca è basata sui metodi per determinare o ridurre il numero di questi pesi. Con riferimento all'elevato numero di criteri presi in considerazione in un problema di decisione multicriteriale si dovrebbe ridurre il numero di misure fuzzy.

A questo proposito useremo il concetto introdotto da Grabisch (1996): misure fuzzy di ordine  $k$ .

Una misura fuzzy è definita  $k$ -additiva se  $a(T)=0 \quad \forall T \subseteq G$  tale che  $|T|>k$ . Si osservi che una misura 1-additiva è la consueta misura fuzzy additiva  $\mu_G=\mu_{(1)}+\mu_{(2)}+....\mu_{(G)}$ . La misura fuzzy 1-additiva è caratterizzata da  $n$  parametri (in termini di rappresentazione di Möbius, un valore  $a(\{g_i\})$  per ogni criterio  $g_i$ ,  $i=1,...,n$ ).

In concrete situazioni di decisione risultano interessanti le misure 2-additive. In questo caso si riescono a modellizzare interazioni positive e negative tra coppie di criteri, senza scendere a un grado di analiticità di rappresentazione delle preferenze che consideri anche interazioni tra terne, quaterne e in genere  $n$ -uple,  $n>2$ , di criteri. Dal punto di vista decisionale, l' utilizzo di misure 2-additive si giustifica osservando che sembrano concretamente interessanti informazioni che riguardano l'importanza di singoli criteri e l' interazione tra coppie di criteri. Dal punto di vista computazionale, risiede nel fatto che

una misura 2-additiva richiede di valutare un numero di parametri pari a  $n+\binom{n}{2}$  (in termini di rappresentazione di Möbius, un valore  $a(\{g_i\})$  per ogni criterio  $g_i$ ,  $i=1,...,n$ , e un valore  $a(\{g_i, g_j\})$  per ogni coppia di criteri  $g_i, g_j$ ,  $i,j=1,...,n$ ), laddove una generica misura fuzzy richiede di valutare un numero di parametri pari a  $2^n$ , (in termini di rappresentazione di Möbius, un valore  $a(T)$  per ogni insieme di criteri  $T \subseteq G$ ).

Con riferimento a una misura fuzzy 2-additiva la trasformazione inversa che consente di risalire dalla rappresentazione di Möbius  $a(S)$  alla misura fuzzy  $\mu(S)$  risulta definita :

$$\mu(S) = \sum_{i \in S} a(\{g_i\}) + \sum_{\{i,j\} \subseteq S} a(\{g_i, g_j\}) \quad \forall S \subseteq N.$$

$$1b) \mu(\emptyset)=0, \sum_{g_i \in G} a(\{g_i\}) + \sum_{\{g_i, g_j\} \subseteq G} a(\{g_i, g_j\}) = 1.$$

$$2b) a(\{g_i\}) \geq 0, \forall i \in G, a(\{g_i\}) + \sum_{g_j \in T} a(\{g_i, g_j\}) \geq 0, \forall g_i \in G \text{ e } \forall T \subseteq G - \{g_i\}$$

Nella nostra ricerca proprio la determinazione delle varie misure  $\mu$ , rappresenta particolare oggetto di indagine perché rappresentative delle importanze parziali relative alle singole componenti della Customer Satisfaction. Ma di questo parleremo nei capitoli seguenti.

## CAPITOLO 9-

### **IL METODO ROUGH SETS PER LA CUSTOMER SATISFACTION**

#### 9.1 L'APPROCCIO ROUGH SETS PER LA STIMA DELL'IMPORTANZA RELATIVA DI CIASCUN ATTRIBUTO

La teoria dei Rough Sets (insiemi approssimati), introdotta da Pawlak (1982, 1991), si è dimostrata spesso un eccellente strumento matematico per analizzare dati caratterizzati da imprecisione e vaghezza nella loro descrizione. Essa è fondata sull'assunzione che ad ogni oggetto dell'universo del discorso è associata qualche informazione (dati, conoscenza), espressa utilizzando opportuni attributi che descrivono gli oggetti considerati. Oggetti caratterizzati dalla stessa descrizione sono indiscernibili (similari) con riferimento alle informazioni disponibili. La relazione di indiscernibilità così generata costituisce il fondamento matematico della teoria dei rough sets.

In quest'ultimo caso, il sottoinsieme  $Y$  può essere caratterizzato da due insiemi ordinari, chiamati approssimazione inferiore e superiore. Un rough set è definito mediante queste due approssimazioni, che coincidono nel caso di un insieme ordinario. La teoria dei rough sets, che si propone di analizzare possibili relazioni di causa-effetto tra i dati imperfetti (caratterizzati da incertezza e vaghezza) disponibili, presenta talune intersezioni e si pone in alcuni casi come complementare a molte altre teorie e tecniche matematiche che trattano l'incertezza e l'imprecisione: teoria della probabilità, teoria dell'evidenza di Dempster-Shafer, teoria dei fuzzy sets, analisi discriminante, ecc.

Taluni importanti caratteristiche dell'approccio dei rough sets rendono tale strumento particolarmente interessante in numerose applicazioni a problemi concreti. Con riferimento all'output (informazioni ottenibili), è possibile avere a posteriori informazioni circa il ruolo (l'importanza) che taluni attributi o loro sottoinsiemi hanno nell'analisi del problema affrontato (senza dover predefinire trade-offs, ecc.) e si ottengono risultati facilmente comprensibili nella forma di regole decisionali "Se...allora", utilizzando gli attributi più rilevanti.

Il problema della discretizzazione dei dati quantitativi è abbastanza delicato, in quanto i risultati delle analisi possono dipendere dalla discretizzazione adottata.

Pertanto, ogni oggetto  $x$  di  $U$  sarà descritto da un vettore (stringa), ogni elemento del quale rappresenta il valore che il corrispondente attributo assume per  $x$ ; detto vettore è chiamato descrizione di  $x$  in termini delle valutazioni degli attributi di  $Q$  e denotato  $Des_Q(x)$ . Ovviamente, la relazione binaria di indiscernibilità così definita è una relazione di equivalenza (riflessiva, simmetrica e transitiva). La famiglia di tutte le classi di equivalenza della relazione  $I_p$  viene denotata con  $U/I_p$  e la classe di equivalenza contenente un elemento  $x \in U$  con  $I_p(x)$ . Le classi di equivalenza della relazione  $I_p$  sono chiamate insiemi  $P$ -elementari. Risulta  $0 \leq \alpha_p(X) \leq \gamma_p(X) \leq 1$  e la qualità rappresenta la frequenza relativa degli oggetti correttamente classificati usando gli attributi di  $P$ .

Se si considera un concetto vago, ossia allorché gli elementi dell' universo non possono essere classificati con certezza come appartenenti al concetto, l' incertezza è collegata al grado di appartenenza degli elementi all' insieme. Allora, per discutere il problema dell' incertezza dal punto di vista dei rough sets, occorre definire la funzione di appartenenza  $\mu_X^P(x)$  collegata al concetto di rough set (**rough membership function**).

Utilizzando la relazione di indiscernibilità, si ottiene

$$\mu_X^P(x) = \frac{|X \cap I_P(x)|}{|I_P(x)|}.$$

Nella teoria dei rough sets vi è, quindi, una stretta relazione tra vaghezza, insita negli insiemi e richiedente quindi le approssimazioni, ed incertezza, collegata agli elementi degli insiemi e per la quale è necessario introdurre il grado di appartenenza approssimativo. La peculiarità dei rough sets consiste nel trattare una imprecisa rappresentazione della realtà dovuta alla granularità della conoscenza, conseguenza della indiscernibilità tra oggetti aventi la stessa descrizione ("granuli").

Pertanto, un ridotto è un insieme di attributi che preserva le partizioni; cioè è un sottoinsieme minimale di attributi che consente di ottenere le stesse classificazioni, e quindi la stessa qualità dell'approssimazione, degli elementi di  $U$  ottenibili usando

l' intero insieme di attributi  $P$ . In altri termini, gli attributi che non appartengono ad un ridotto sono superflui rispetto alle classificazioni degli elementi dell' universo.

Possono esistere più ridotti di  $P$  in una tavola delle informazioni. Ai fini operativi, dunque, è sufficiente prendere in considerazione solamente i più importanti attributi (ridotti) per l' analisi della tavola delle informazioni considerata.

Se in una tavola delle informazioni gli attributi di  $Q$  vengono distinti in attributi condizionali (insieme  $C$ ) e attributi decisionali (insieme  $D$ ),  $C \cup D = Q$  e  $C \cap D = \emptyset$ , detta tavola è chiamata tavola delle decisioni. Gli attributi decisionali inducono delle partizioni di  $U$  dedotte dalle relazioni di indiscernibilità  $I_D$ , in maniera assolutamente indipendente dagli attributi condizionali di  $C$ . Nelle applicazioni operative, si tende a ridurre gli attributi condizionali preservando la dipendenza tra attributi condizionali e decisionali, il che consente di prendere decisioni adoperando minori informazioni. In altri termini, si vuole usare il minor numero possibile di attributi condizionali senza deteriorare la qualità dell' approssimazione della classificazione indotta dagli attributi decisionali. Poiché si tende a evidenziare la dipendenza funzionale tra gli attributi condizionali e quelli decisionali, una tavola delle decisioni può anche essere espressa come un insieme di regole decisionali. Queste sono delle proposizioni logiche (implicazioni) del tipo 'se...allora', ove l' antecedente riguarda valori assunti da uno o più attributi condizionali (descrizioni di insiemi  $C$ -elementari) ed il conseguente partizioni generate dagli (uno o più) attributi decisionali (descrizioni di insiemi  $D$ -elementari).

Le regole di decisione danno una sintesi, una rappresentazione comprensibile e generalizzata di conoscenza contenuta in un insieme di dati organizzato sotto forma d' una tabella delle informazioni. Le file della tabella sono identificate dagli oggetti, mentre le colonne sono identificate dagli attributi e le entrate della tabella sono attributi-valori. Se ci muoviamo in ambito di Customer Satisfaction, gli oggetti corrispondono ai clienti e gli attributi alle caratteristiche della valutazione compreso la valutazione completa di un prodotto, quindi le regole di decisione rappresentano l' atteggiamento preferenziale del cliente e spiegano i motivi delle sue preferenze. Secondo Slovic (1975), la gente prende le decisioni cercando le regole che forniscono la buona

giustificazione delle loro scelte. Una dichiarazione diretta delle regole di decisione richiede, tuttavia, uno sforzo conoscitivo grande dal decisore, essendo in genere più sicura quando prende le decisioni esemplari che spiegandole. Per questo motivo, l'idea di arguire i modelli di comportamento del cliente in termini di regole di decisione dalle decisioni esemplari è molto attraente. L'induzione delle regole dagli esempi è un metodo tipico di intelligenza artificiale. È concorde con il principio della razionalità e con logica di aggregazione-disaggregazione da Jacquet-Lagrèze (1981). Esistono molte applicazioni delle regole di decisione nel commercio e nella finanza. Per esempio:- i rivenditori usano le regole di associazione per capire le abitudini del cliente e le preferenze (analisi di mercato) e per applicare l'individuazione per lanciare le promozioni e la pubblicità efficaci, - le aziende di vendita diretta e di telemarketing usano le regole di decisione per ridurre il numero di chiamate fatte e per aumentare il rapporto delle chiamate riuscite. Altre applicazioni delle regole di decisione sono in tali settori come le linee aeree, il manufacturing, la telecomunicazione, le società di assicurazioni e così via. All'interno del campo del machine learning, del knowledge discovery e del data mining il concetto dell'insieme approssimativo, introdotto da Pawlak (1982), si è rivelato essere uno strumento efficace per l'analisi di una tabella delle informazioni (per esempio una tabella che rappresenta le risposte ad un questionario circa la valutazione di un prodotto o di un servizio) che descrive un insieme degli oggetti (risposte del cliente ad un questionario) da un insieme degli attributi a valori multipli (valutazioni del prodotto e dei servizi riguardo alle caratteristiche differenti, così come una valutazione completa).

Come precisato da Greco, da Matarazzo e da Slowinski (1996), il metodo approssimativo classico dell'insieme (CRSA) non considera, tuttavia, gli attributi con le scale preferenza-ordinate (dominii), cioè criteri. Tuttavia, in molti problemi reali le proprietà d'ordinamento degli attributi considerati svolgono un ruolo importante. Un caso tipico di presenza delle proprietà d'ordinamento è l'analisi di comportamento del cliente. Nell'ambito dell'analisi del comportamento deve valere il principio della dominanza che richiede che le valutazioni migliori sulle caratteristiche particolari devono non deteriorare (e possibilmente dovrebbe migliorare) la valutazione complessiva. Sulla base di suddette considerazioni, Greco, Matarazzo e Slowinski



hanno proposto un nuovo metodo dei rough sets a problemi dove le proprietà d'ordinamento sono importanti. Questo nuovo metodo sostituisce il rapporto di indiscernibilità, usato in CRSA, tramite un rapporto di dominanza (Greco, Matarazzo e Slowinski, 1998). Questo nuovo metodo, denominato metodo rough sets basato sulla Dominanza (DRSA), è generale e può essere usato a tutti i problemi di classificazione che coinvolge i criteri anziché gli attributi e, pertanto, rappresenta un modello specifico di analisi di decisione multicriteriale. Nell'utilizzo sui grandi insiemi di dati tabulati, come nell'analisi dei questionari dei clienti, un'altra modifica è necessaria: occorre modellare i rapporti dei dati in termini di distribuzione di frequenza piuttosto che come rapporto completo dell'inclusione. Questa modifica già è stata introdotta all'interno del metodo classico dei rough sets attraverso il cosiddetto modello a variabile di precisione (Ziarko 1993) che accetta un numero limitato di contro-esempi quando definisce i concetti principali dell'analisi dei rough sets.

In questo paragrafo consideriamo un'estensione recente del modello DRSA a variabile di precisione: cioè il Modello dei rough sets basato sulla dominanza a variabile di precisione (VPDRSA) (Greco, Matarazzo, Slowinski, Stefanowski 2000).

L'applicazione di VPDRSA proposta da Greco, Matarazzo, Platania e Slowinski è stata applicata ad analisi di soddisfazione di cliente e di comportamento del cliente. L'analisi di soddisfazione del cliente rappresenta uno strumento importante nell'ambito delle politiche di marketing delle imprese, rivolto a determinare rispettivamente che caratteristiche di un prodotto/servizio stanno mostrando la relativa resistenza e debolezza. Il risultato di questa analisi offre una conoscenza circa le preferenze e le aspettative dei clienti e costituisce un livello di valutazione delle prestazioni e del merito per tutta l'attività aziendale. La tecnica ampiamente usata dai ricercatori di analisi di mercato per misurare la soddisfazione di cliente è l'analisi congiunta (Krantz, Luce, Suppes, Tversky (1978), Kahneman e Tversky (1979); Cattin e Wittink (1982), Green e Srinivasan (1978)). Lo scopo dell'analisi congiunta, come già detto, è determinare una funzione che rappresenti le preferenze dei clienti riguardo alle caratteristiche differenti del prodotto o del servizio considerato. Ci sono parecchi metodi di valutazione comunemente usati. Accenniamo Monanova (Kruskal 1965), Linmap (Srinivisan e Shocker 1973), la regressione di variabile fittizia (Johnston 1972),

l' analisi di Logit (McFadden 1976), l' analisi di Probit (Goldberg 1964), il modello di disaggregazione di preferenza (Siskos, Grigoroudis, Zopounidis, Sauris 1998). In questo paragrafo ci proponiamo: è possibile usare un metodo di conoscenza-orientato, come l' analisi rough sets, per la soddisfazione di cliente e l'analisi del comportamento del cliente? In questo contesto l' uso di VPDRSA è imposto dal fatto che le preferenze del cliente, che costituiscono l' oggetto di analisi, sono tipicamente ordinate.

Sulla base di queste considerazioni, Greco, Matarazzo e Slowinski (1997a) hanno proposto un nuovo approccio dei rough sets per problemi di classificazione multicriteriale. Così come nell' analisi dei rough sets tradizionale, l' approccio proposto è basato su approssimazioni di una partizione degli oggetti dell'universo in alcune classi predefinite sulla base della tavola delle informazioni. Tuttavia, a differenza dell' approccio originario dei rough sets, le approssimazioni sono costruite usando relazioni di dominanza invece che di indiscernibilità. Questo permette di prendere esplicitamente in considerazione le proprietà ordinali degli attributi (criteri) considerati.

$\forall q \in C$  sia  $S_q$  una relazione di surclassamento (Roy, 1985) su  $U$  con riferimento all' attributo  $q$  tale che  $xS_q y$  significa "x è almeno tanto buono quanto y rispetto all' attributo  $q$ ". Si suppone che  $\mathcal{S}$  sia un preordine totale, cioè una relazione binaria fortemente completa e transitiva, definita su  $V_q$ . Inoltre sia  $Cl = \{Cl_t, t \in T\}$ ,  $T = \{1, \dots, n\}$ , un insieme di classi di  $U$ , tali che ogni  $x \in U$  appartenga a una e una sola classe  $Cl_t \in Cl$ . Si suppone che  $\forall r, s \in T$ , con  $r > s$ , gli elementi di  $Cl_r$  sono preferiti (strettamente o debolmente (Roy, 1985) agli elementi di  $Cl_s$ . Più formalmente, se  $S$  è una relazione di surclassamento globale su  $U$ , cioè se  $\forall x, y \in U$   $xSy$  significa "x è almeno tanto buono quanto y" si suppone che

$$[x \in Cl_r, y \in Cl_s, r > s] \Rightarrow [xSy \text{ e } y \mathcal{S} x].$$

Si considerano anche i seguenti insiemi

$$Cl_t^{\geq} = \bigcup_{s \geq t} Cl_s,$$

$$Cl_t^{\leq} = \bigcup_{s \leq t} Cl_s.$$

Si osservi che  $CI_1^{\geq} = CI_n^{\leq} = U$ ,  $CI_n^{\geq} = CI_1$  e  $CI_1^{\leq} = CI_1$ .

Si dice che  $x$  domina  $y$  con riferimento a  $P \subseteq C$ , indicata  $x D_P y$ , se  $x S_q y \quad \forall q \in P$ . Dati

$P \subseteq C$  e  $x \in U$  siano

$$D_P^+(x) = \{y \in U : y D_P x\},$$

$$D_P^-(x) = \{y \in U : x D_P y\}.$$

$\forall t \in T$  e  $\forall P \subseteq C$  definiamo approssimazione inferiore di  $CI_t^{\geq}$  con riferimento a  $P$ ,

indicata  $\underline{P}CI_t^{\geq}$ , e approssimazione superiore di  $CI_t^{\geq}$  con riferimento a  $P$ , indicata  $\overline{P}CI_t^{\geq}$ ,

rispettivamente:

$$\underline{P}CI_t^{\geq} = \{x \in U : D_P^+(x) \subseteq CI_t^{\geq}\},$$

$$\overline{P}CI_t^{\geq} = \bigcup_{x \in CI_t^{\geq}} D_P^+(x).$$

Analogamente,  $\forall t \in T$  e  $\forall P \subseteq C$  definiamo approssimazione inferiore di  $CI_t^{\leq}$  con

riferimento a  $P$ , indicata  $\underline{P}CI_t^{\leq}$ , e approssimazione superiore di  $CI_t^{\leq}$  con riferimento a  $P$ ,

indicata  $\overline{P}CI_t^{\leq}$ , rispettivamente:

$$\underline{P}CI_t^{\leq} = \{x \in U : D_P^-(x) \subseteq CI_t^{\leq}\},$$

$$\overline{P}CI_t^{\leq} = \bigcup_{x \in CI_t^{\leq}} D_P^-(x).$$

Le  $P$ -frontiere (regioni dubbie) di  $CI_t^{\geq}$  e  $CI_t^{\leq}$  sono definite come

$$B_{np}(CI_t^{\geq}) = \overline{P}CI_t^{\geq} - \underline{P}CI_t^{\geq},$$

$$B_{np}(CI_t^{\leq}) = \overline{P}CI_t^{\leq} - \underline{P}CI_t^{\leq}.$$

$\forall t \in T$  e  $\forall P \subseteq C$  si definiscono accuratezza dell'approssimazione rispettivamente di  $CI_t^{\geq}$

e  $CI_t^{\leq}$  i rapporti:

$$\alpha_P(CI_t^{\geq}) = \frac{|\underline{P}CI_t^{\geq}|}{|\overline{P}CI_t^{\geq}|},$$

$$\alpha_P(Cl_i^{\leq}) = \frac{|P(Cl_i^{\leq})|}{|\overline{P}(Cl_i^{\leq})|}.$$

Il rapporto

$$\gamma_P(Cl) = \frac{|U - (\bigcup_{t \in T} Bn_P(Cl_t^{\leq})) \cup (\bigcup_{t \in T} Bn_P(Cl_t^{\geq}))|}{|U|}$$

è definito qualità dell' approssimazione della partizione  $Cl$  per mezzo dell' insieme di attributi  $P$ , o in breve, qualità della classificazione. Essa esprime il rapporto tra tutti gli oggetti  $P$  - correttamente classificati e tutti gli oggetti della tavola.

Ogni sottoinsieme minimale  $P \subseteq C$  tale che  $\gamma_P(Cl) = \gamma_C(Cl)$  si chiama un ridotto di  $Cl$  ed è indicato  $RED_P$ . Si ricorda che una tavola delle informazioni può avere più di un ridotto. L' intersezione di tutti i ridotti si chiama nucleo (core) ed è indicato  $CORE$ .

Sulla base delle approssimazioni ottenute per mezzo della relazione di dominanza, si può ottenere una descrizione generalizzata delle informazioni preferenziali contenute in una certa tavola delle informazioni mediante un certo numero di regole di decisione.

L'approssimazione rough basata sulla dominanza delle classe di unione ascendenti e discendenti può servire ad indurre una descrizione generalizzata degli oggetti contenuti nella tabella delle informazioni in termini di regole di decisione se..., allora..... Per una data classe di unione ascendente o discendente,  $Cl_t^{\geq}$  o  $Cl_s^{\leq}$ , le regole di decisione implicano sotto certe ipotesi che gli oggetti appartenenti a  $\underline{P}(Cl_t^{\geq})$  o  $\underline{P}(Cl_s^{\leq})$  sono *positivi* e tutti gli altri *negativi*, e determinano un'assegnazione "almeno alla classe  $Cl_t$ " o "al più alla classe  $Cl_s$ ", rispettivamente. Sono denominate  $D_{\geq}$ -decision rules and  $D_{\leq}$ -decision rules, rispettivamente.

Inoltre, le regole di decisione sotto l'ipotesi che tutti gli oggetti che appartengono all'intersezione  $\overline{P}(Cl_s^{\leq}) \cap \overline{P}(Cl_t^{\geq})$  sono *positivi* e tutti gli altri *negativi*, determinano

un'assegnazione ad alcune classi tra  $Cl_s$  e  $Cl_t$  ( $s < t$ ). Queste regole sono chiamate  $D_{\geq}$ -*decision rules*. All'interno della VPDRSA, le regole decisionali conducono all'estensione delle approssimazioni ottenute per alcuni livelli di confidenza  $l$ . Per tale ragione, è necessario assegnare ad ogni regola decisionale un parametro aggiuntivo  $\alpha$  chiamato *confidenza della regola*. Ogni regola è caratterizzata inoltre da un secondo parametro, chiamato *supporto*, definito dal rapporto del numero degli oggetti soddisfacenti la parte di condizione della regola rispetto al numero totale degli oggetti.

Si assume che per ogni  $q \in C$ ,  $V_q \subseteq \mathbf{R}$  e che per ogni  $x, y \in U$ ,  $f(x, q) \geq f(y, q)$  implica  $x S_q y$ , i seguenti tre tipi di regole di decisione con confidenza  $\alpha$  possono considerarsi:

1) regole di decisione del tipo  $D_{\geq}$ , che si presentano nella forma:

$$[f(x, q_1) \geq r_{q_1} \text{ e } f(x, q_2) \geq r_{q_2} \text{ e } \dots f(x, q_p) \geq r_{q_p}] \Rightarrow x \in Cl_t^{\geq},$$

dove  $\{q_1, q_2, \dots, q_p\} \subseteq C$ ,  $r_{q_1} \in V_{q_1}$ ,  $r_{q_2} \in V_{q_2}, \dots, r_{q_p} \in V_{q_p}$  e  $t \in T$ ; tali regole sono ottenute sulla base delle approssimazioni inferiori delle classi  $Cl_t^{\geq}$ ;

2) regole di decisione del tipo  $D_{\leq}$ , che si presentano nella forma:

$$[f(x, q_1) \leq r_{q_1} \text{ e } f(x, q_2) \leq r_{q_2} \text{ e } \dots f(x, q_p) \leq r_{q_p}] \Rightarrow x \in Cl_t^{\leq},$$

dove  $\{q_1, q_2, \dots, q_p\} \subseteq C$ ,  $r_{q_1} \in V_{q_1}$ ,  $r_{q_2} \in V_{q_2}, \dots, r_{q_p} \in V_{q_p}$  e  $t \in T$ ; tali regole sono ottenute sulla base delle approssimazioni inferiori delle classi  $Cl_t^{\leq}$ ;

3) regole di decisione del tipo  $D_{\geq \leq}$ , che si presentano nella forma:

$$[f(x, q_1) \geq r_{q_1} \text{ e } f(x, q_2) \geq r_{q_2} \text{ e } \dots f(x, q_k) \geq r_{q_k} \text{ e } f(x, q_{k+1}) \leq r_{q_{k+1}} \text{ e } \dots f(x, q_p) \leq r_{q_p}] \Rightarrow x \in Cl_t^{\leq} \text{ o } x \in Cl_s^{\geq},$$

dove  $\{q_1, q_2, \dots, q_k\} \subseteq C$ ,  $\{q_{k+1}, q_{k+2}, \dots, q_p\} \subseteq C$ ,  $r_{q_1} \in V_{q_1}$ ,  $r_{q_2} \in V_{q_2}, \dots, r_{q_p} \in V_{q_p}$ ,  $s, t \in T$  tali che  $t < s$ ; tali regole sono ottenute sulla base delle frontiere delle classi  $Cl_t^{\leq}$  e  $Cl_s^{\geq}$ .

Un insieme di regole di decisione è *completo* se valgono le seguenti condizioni:

- 1) ogni  $y \in \underline{C}(Cl_t^{\geq})$  supporta almeno una decisione del tipo  $D_{\geq}$  "se  $f(x, q_1) \geq r_{q_1}$  e  $f(x, q_2) \geq r_{q_2}$  e  $\dots f(x, q_p) \geq r_{q_p}$ , allora  $x \in Cl_r^{\geq}$ ", con  $r, t \in \{2, \dots, n\}$  and  $r \geq t$ ,

2) ogni  $y \in \underline{C}(Cl_t^{\leq})$  supporta almeno una decisione del tipo  $D_{\leq}$  "se  $f(x, q_1) \leq r_{q_1}$  e  $f(x, q_2) \leq r_{q_2}$  e ...  $f(x, q_p) \leq r_{q_p}$ , allora  $x \in Cl_u^{\leq}$ ", con  $u, t \in \{1, \dots, n-1\}$  and  $u \leq t$ ,

3) ogni  $y \in \overline{C}(Cl_s^{\leq}) \cap \overline{C}(Cl_t^{\geq})$  supporta almeno una decisione del tipo  $D_{\geq}$  "se  $f(x, q_1) \geq r_{q_1}$  e  $f(x, q_2) \geq r_{q_2}$  e ...  $f(x, q_k) \geq r_{q_k}$  e  $f(x, q_{k+1}) \leq r_{q_{k+1}}$  e ...  $f(x, q_p) \leq r_{q_p}$ , allora  $x \in Cl_v \cup Cl_{v+1} \cup \dots \cup Cl_z$ ", con  $s, t, v, z \in T$  and  $s \leq v < z \leq t$ .

Rileviamo che l'applicazione dell'insieme completo delle regole decisionali sugli oggetti della tabella delle informazioni determina un'assegnazione esatta o approssimativa di questi oggetti alla classe  $Cl_t$ ,  $t \in T$ .

Per ogni oggetto  $x \in U$ , appartenente alla approssimazione inferiore dell'unione delle classi decisionali e non appartenente ad alcuna C-frontiera, la nuova assegnazione data dall'intersezione di tutte le unioni delle classi determinate per riassegnazione implica l'intersezione di tutte le unioni delle classi determinata per assegnazione nella conseguenza di  $D_{\geq}$ - e  $D_{\leq}$ -decision rules abbinata da  $x$ .

Dato un insieme completo di regole, e un oggetto  $y \in U$ , tale che  $y \notin Bn_C(Cl_s^{\leq})$  e  $y \notin Bn_C(Cl_s^{\geq})$  per ogni  $s \in T$ , possono verificarsi le seguenti condizioni:

- $y \in Cl_t$ ,  $t=2, \dots, n-1$ ; allora esiste almeno una  $D_{\geq}$ -decision rule la cui conseguenza è  $x \in Cl_t^{\geq}$ , e almeno una  $D_{\leq}$ -decision rule la cui conseguenza è  $x \in Cl_t^{\leq}$ ;
- $y \in Cl_1$ ; allora esiste almeno una  $D_{\leq}$ -decision rule la cui conseguenza è  $x \in Cl_1^{\leq}$ ;
- $y \in Cl_n$ ; allora esiste almeno una  $D_{\geq}$ -decision rule la cui conseguenza è  $x \in Cl_n^{\geq}$ .

In tutte le situazioni suddette, l'applicazione dell'insieme delle regole di oggetto  $y$ , provocherà l'assegnazione (esatta) di  $y$  alla classe  $Cl_t$ .

Similmente, per ogni oggetto  $y \in \overline{C}(Cl_s^{\leq}) \cap \overline{C}(Cl_t^{\geq})$ ,  $s < t$ , tale che  $y \notin \overline{C}(Cl_{s1}^{\leq}) \cap \overline{C}(Cl_{t1}^{\geq})$ ,

$s1 < [\leq] s$  and  $t \leq [\leq] t1$ , che significa che  $y$  appartiene esclusivamente alle frontiere

$Bn_C(Cl_v^{\geq})$ ,  $v=s+1, \dots, t$ , e  $Bn_C(Cl_z^{\leq})$ ,  $z=s, \dots, t-1$ , esiste almeno una  $D_{\geq}$ -decision rule la

cui conseguenza è  $x \in Cl_s \cup Cl_{s+1} \cup \dots \cup Cl_t$ . Allora, nel risultato dell'applicazione

dell' insieme delle regole completo ad oggetto ysarà rassegnata (approssimativamente) alle classi  $Cl_s \cup Cl_{s+1} \cup \dots \cup Cl_t$ .

Denominiamo minimo ogni insieme di regole minime di decisione che è completo non-ridondante. Una delle tre strategie di induzione può essere adottata per ottenere un insieme di decision rule (Stefanowski e Vanderpooten, 1994; Mienko, Stefanowski, Toumi e Vanderpooten, 1996):

generazione di una descrizione minima, cioè un insieme delle regole minimo,

generazione di una descrizione esauriente, cioè tutte le regole minime possibili per una data tabella delle informazioni,

generazione di una descrizione caratteristica, cioè un insieme delle regole minime che riguardano relativamente molti oggetti ciascuno, tuttavia, tutti insieme non necessariamente tutti gli oggetti dalla U.

Come accennato, l' analisi dei rough sets basata sulle approssimazioni mediante relazioni binarie di dominanza migliora, in generale, i risultati dei problemi di classificazione rispetto all' approccio classico basato sull' uso della relazione di indiscernibilità.

Nei problemi di classificazione, i vantaggi dell' approccio basato su relazioni di dominanza rispetto all' analisi dei rough sets originaria, basata sulla relazione di indiscernibilità, possono sintetizzarsi come segue:

- 1) Si ottiene spesso un minor numero di ridotti ed un nucleo più grande. Queste due caratteristiche sono generalmente riconosciute come delle proprietà desiderabili di una buona approssimazione (Pawlak, 1991, Slowinski e Stefanowski, 1996).
- 2) La qualità dell' approssimazione ottenuta usando le relazioni di dominanza può essere inferiore a quella ottenuta approssimando con relazioni d' indiscernibilità. Ma questo apparente inconveniente mostra, in verità, un altro notevole vantaggio dell' approccio considerato. Infatti, l' approccio mediante approssimazioni basate su relazioni di dominanza mette spesso in luce delle inconsistenze nei risultati della classificazione, che non possono essere colte dall' approssimazione tramite indiscernibilità. Quest' ultima, infatti, classifica gli oggetti dell' universo distinguendoli solamente in conseguenza di loro descrizioni differenti in termini degli attributi/criteri considerati, ma non coglie

assolutamente aspetti ordinali dei dati. Può pertanto accadere che due oggetti  $x$  e  $y$  siano stati classificati (da un esperto, in esperienze passate, ecc.) in maniera tale che la valutazione globale di  $x$  sia peggiore di quella di  $y$ , mentre dalla valutazione degli stessi, evidenziata nella corrispondente tavola delle decisioni, risulta che  $x$  domina  $y$ . L' approccio considerato, e solamente esso, consente di evidenziare questa inconsistenza, spiegando quindi anche la ragione dell' apparente peggioramento della qualità della classificazione.

Migliora la qualità dell' insieme delle regole decisionali ottenute dalle approssimazioni mediante relazioni di dominanza, che forniscono in generale una rappresentazione più sintetica della conoscenza contenuta nella tavola delle informazioni. Gli insiemi minimali di regole così ottenuti hanno un minor numero di regole ed usano un minor numero di attributi e descrittori rispetto all' algoritmo di classificazione basato sulla classica analisi dei rough sets. Inoltre, l' applicazione di tali regole a nuovi oggetti da classificare fornisce in generale risultati migliori; talvolta, infatti, utilizzando l' algoritmo originario non si è in grado di classificare qualche nuovo oggetto.



## 9.2 CONFRONTO CON ALTRE METODOLOGIE

Come già accennato, esistono numerose relazioni tra la teoria dei rough sets e altre teorie matematiche che si propongono di trattare particolari tipi di "incertezza" o di analizzare dati "imperfetti": per un accurato confronto della teoria dei rough sets con l'analisi discriminante, la teoria dei fuzzy sets e la teoria dell'evidenza si veda: Krusinska, Slowinski e Stefanowski (1992), Dubois e Prade (1992) e Skowron e Grzymala-Busse (1993)).

<b>Problema</b>	<b>Metodi statistici</b>	<b>Approccio dei Rough Sets</b>
Obiettivi	Identificazione e stima dei parametri delle equazioni strutturali	Riduzione degli attributi ridondanti, generazione di regole di decisione
Rappresentazione dei dati	Tavole a due entrate che rappresentano un campione	Tavola delle informazioni
Tipi di attributi	Attributi quantitativi (almeno nel caso classico)	Attributi qualitativi; gli attributi quantitativi sono trasformati in qualitativi per mezzo di una opportuna discretizzazione
Requisiti dei dati	Il campione deve essere statisticamente rappresentativo; distribuzione multivariata normale	Nessun requisito; possibilità di analizzare anche tavole delle informazioni di ridotte dimensioni
Operatori per l'aggregazione dei dati	Valori medi, matrice delle covarianze, test statistici	Nessun operatore; i dati vengono analizzati nella loro forma originaria
Riduzione dei dati	Selezione di attributi con il maggiore potere discriminante; tipico strumento: test statistici	Sottoinsiemi minimali di attributi che assicurano la stessa qualità di classificazione dell'intero insieme di attributi
Risultati finali	Rappresentazione funzionale	Regole di decisione nella forma di proposizioni logiche

Si riportano in precedenza e di seguito talune brevi considerazioni relative al confronto della teoria dei rough sets rispettivamente con quella dei fuzzy sets e la classica analisi statistica (analoga proposta da Stefanowski (1992)).

Spesso l' approccio dei rough sets non si pone come alternativo, ma come complementare ad altri approcci basati su teorie o tecniche differenti. Sono state effettuate diverse applicazioni concrete utilizzando differenti approcci; l' uso dei rough sets è risultato molto spesso particolarmente interessante, sia per le notevoli potenzialità applicative dovute alle sue peculiari proprietà (grande "povertà" di informazioni richieste) che per i peculiari risultati ottenuti (regole decisionali, rilevanza degli attributi).

<b>Problema</b>	<b>Fuzzy Sets</b>	<b>Rough Sets</b>
Incertezza (Conoscenza imperfetta)	Vaghezza delle informazioni dovuta alla non precisa definizione degli insiemi	Granularità della conoscenza dovuta alla indiscernibilità tra oggetti in un insieme
Informazioni aggiuntive	Grado di appartenenza	Nessuna
Elaborazione di dati	In modo "esatto", utilizzando la funzione di appartenenza	In modo "impreciso", usando le approssimazioni inferiori e superiori
Modellizzazione matematica	Generalizzazione al continuo della funzione caratteristica di un insieme, delle relazioni binarie e degli operatori logici	Calcolo di opportune partizioni, approssimazione di una data classificazione e utilizzazione di sottoinsiemi minimali di attributi
Un esempio dall'elaborazione di immagini	Livelli di grigio tra il nero e il bianco (gradi di funzioni di appartenenza)	Dimensione dei pixels usati per approssimare i contorni delle immagini (granularità)

Come accennato, le classiche definizioni di approssimazioni inferiore e superiore sono state introdotte con riferimento alla relazione binaria di indiscernibilità, che è una relazione binaria di equivalenza. In tale approccio, sia gli insiemi da approssimare che la relazione usata sono ordinari (crisp).

Una prima generalizzazione nella direzione dei fuzzy sets è stata proposta da Dubois e Prade (1990, 1992); essi hanno, infatti, considerato la possibilità che gli insiemi da approssimare e/o la relazione di indiscernibilità fossero fuzzy. Ma tale approccio è ancora basato sull' uso della relazione di indiscernibilità

Di particolare interesse sono le proposte di usare, anziché la relazione di indiscernibilità, delle relazioni binarie di similarità o di tolleranza, più deboli di quella di indiscernibilità, in quanto richiedono solo la riflessività, rilassando le assunzioni di simmetria e transitività (cfr. Slowinski, Vanderpooten, 1995, 1997).

L' indiscernibilità, come osservato, implica la assoluta impossibilità di distinguere due oggetti di  $U$  che hanno la stessa descrizione in termini degli attributi di  $Q$ . Tale relazione induce su  $U$  delle classi di equivalenza, che costituiscono i granuli fondamentali della conoscenza mediante l' indiscernibilità. Spesso, nella realtà, anche per l' imprecisione dei dati che descrivono gli oggetti, piccole differenze non sono considerate significative ai fini della distinzione degli oggetti corrispondenti.

In generale, le relazioni di similarità  $R$  non generano delle partizioni su  $U$ ; le informazioni sulla similarità possono rappresentarsi usando delle classi di similarità per ogni oggetto  $x \in U$ . Precisamente, la classe di similarità di  $x$ , denotata con  $R(x)$ , è costituita dall' insieme degli oggetti che sono simili ad  $x$ :

$$R(x) = \{y \in U: yRx\}.$$

La relazione di similarità è ovviamente riflessiva (ogni oggetto è simile a sè stesso). Slowinski e Vanderpooten (1995,1997) hanno proposto una relazione di similarità che sia solamente riflessiva, rilassando quindi le proprietà di simmetria e transitività. L' abbandono della transitività è facilmente giustificabile, ricordando ad esempio - il paradosso delle tazzine di caffè di Luce (1956). Pertanto, la simmetria della relazione di similarità non deve essere imposta. Inoltre, le definizioni proposte sono le uniche che caratterizzano propriamente l' insieme degli oggetti positivi (approssimazione inferiore) e l' insieme degli oggetti positivi o ambigui (approssimazione superiore) quando si usa una relazione di similarità riflessiva, ma non necessariamente simmetrica e transitiva.

Infine, un' ulteriore generalizzazione si è avuta con l' introduzione di una relazione binaria riflessiva fuzzy  $R(x,y)$  definita su  $U$ , ossia una relazione di similarità fuzzy, per definire le approssimazioni inferiore e superiore di un insieme, anch' esso fuzzy (Greco, Matarazzo, Slowinski, 1997c).

A tal fine vengono opportunamente usati la negazione ed i classici connettivi della logica fuzzy, in particolare delle t-norme  $T$  come congiunzioni, delle t-conorme  $S$  come disgiunzioni e delle opportune negazioni fuzzy  $N$ . L' insieme degli oggetti positivi e quello degli oggetti negativi sono dei fuzzy sets, le cui funzioni di appartenenza esprimono rispettivamente la credibilità che “per ogni  $y \in U$ ,  $x$  è non simile ad  $y$  e/o  $y$  appartiene ad  $X$ ” e che “per ogni  $y \in U$ ,  $x$  è non simile ad  $y$  e/o  $y$  non appartiene ad  $X$ ”.

Infine, per ogni attributo  $q \in Q$  si consideri una relazione binaria fuzzy  $R_q$ , cioè  $R_q: V_q \times V_q \rightarrow [0,1]$ , dove  $R_q(x,y)$  rappresenta il grado di similarità tra due valutazioni  $x,y$  rispetto all' attributo  $q$ ; precisamente  $\forall x,y,w,z \in V_q$ :

$R_q(x,y) = 0$  significa assenza di similarità tra  $x$  ed  $y$ ,

$R_q(x,y) = 1$  significa che  $x$  è assolutamente simile ad  $y$  ( $R_q(x,x)=1$ ),

$R_q(x,y) \leq R_q(w,z)$  significa che la similarità tra  $x$  ed  $y$  è almeno tanto credibile quanto la similarità tra  $w$  e  $z$ .

Modellizzando ed utilizzando opportunamente tale relazione di similarità fuzzy e le approssimazioni prima definite, è possibile ottenere delle regole decisionali, certe (ottenute dalle approssimazioni inferiori) e possibili (ottenute dalle approssimazioni superiori), ciascuna avente un suo grado di credibilità, che esprimono delle implicazioni logiche, i cui antecedenti sono, però, delle espressioni del tipo “se la valutazione sul criterio  $q_i$  è simile a  $f(x,q_i)$  e...”

Come accennato in precedenza, una tavola delle decisioni raccoglie tutte le informazioni relative ad un insieme di oggetti, descritti da un certo numero di

attributi. Più precisamente, gli attributi condizionali forniscono una descrizione di ogni oggetto in termini di valutazioni su ciascuno di essi; gli attributi decisionali, uno o più, rappresentano uno stato della conoscenza di ciascun oggetto, basata su esperienze pregresse, su opinioni di esperti, su preferenze di decisori, ecc. La tradizionale analisi di tale tavola mediante i rough sets consiste sostanzialmente nel confrontare le classificazioni degli oggetti di  $U$  indotte dagli attributi condizionali di  $C$  o di un sottoinsieme  $P \subseteq C$ , con quella dedotta dagli attributi decisionali  $D$ . Tali classificazioni sono, quindi, costruite l'una indipendentemente dall'altra, e non dedotte l'una dall'altra. Lo strumento che si utilizza per effettuare tali confronti sono le approssimazioni, inferiori e superiori, di ciascuna delle classi decisionali così ottenute, usualmente sulla base della classica relazione di indiscernibilità. Quest'ultima si basa fondamentalmente sulle regole di decisione ottenute dalla tavola analizzata; la fase della spiegazione, quindi, prepara quella della prescrizione, dandole utili informazioni per l'aiuto alle decisioni. Sotto tale aspetto, quindi, l'approccio dei rough sets è simile ad un processo di apprendimento induttivo. Ancora, le regole decisionali generate vengono "ottimizzate", sia con riferimento agli attributi effettivamente adoperati (ridotti), consentendo un grande risparmio nella gestione delle informazioni (eliminazione dei dati superflui), che con riferimento alle regole effettivamente utilizzate (generazione di insiemi di regole decisionali minimali), facilitano la comprensione delle stesse da parte del decisore mediante l'eliminazione di regole "ridondanti" (sulla problematica della generazione di insiemi appropriati di regole si veda comunque Mienko, Stefanowski, Toumi e Vanderpooten, 1996).

## CAPITOLO 10 -

### **LA CUSTOMER SATISFACTION E L'UTILIZZO DI FUNZIONI DI UTILITA' NON ADDITIVE.**

#### 10.1 INTRODUZIONE

Il tema della qualità correlato a quello della soddisfazione del cliente costituisce sicuramente uno degli argomenti più dibattuti in questi ultimi anni. Il concetto di qualità si è ampliato e non attiene esclusivamente agli aspetti strettamente inerenti il processo di produzione ma piuttosto la capacità globale di un'impresa di adattare il sistema aziendale ai bisogni reali dei suoi clienti. Se si intende la qualità dell'offerta aziendale come la capacità di garantire il completo soddisfacimento delle esigenze del cliente, l'analisi di C.S. dovrebbe quindi essere incentrata sulla valutazione della qualità ideale, della qualità attesa, di quella percepita ed infine delle eventuali discrepanze tra esse.

Queste ultime, infatti, si ritiene siano essenziali per il processo di valutazione della soddisfazione globale. In questo capitolo si tratta del problema della misurazione della qualità ideale e della soddisfazione globale in funzione, quest'ultima della distanza tra qualità attesa e percepita.

La misurazione della customer satisfaction è un problema importante per ogni organizzazione. La soddisfazione dei bisogni di cliente è l'obiettivo principale secondo i principi della scienza moderna di vendita. Le tecniche di indagine della soddisfazione di cliente possono essere calcolate tramite la valutazione di fedeltà di marca di un insieme di clienti.

L'errata convinzione dell'esistenza di una diretta correlazione tra qualità del servizio e soddisfazione del cliente deriva probabilmente dalla sottovalutazione della percezione della qualità da parte del cliente sia dalla non netta demarcazione tra i significati di qualità percepita e soddisfazione. Infatti, un'impresa può fornire benissimo un prodotto/servizio d'alta qualità e tuttavia non soddisfare il cliente, e ciò accade quando

gli aspetti su cui si concentrano gli interventi dell'impresa non sono significativi per l'utente finale. Pertanto, bisogna tenere conto del fatto che qualità e soddisfazione sono due aspetti che esprimono concetti diversi: la *qualità percepita* è una valutazione globale, indiretta e spesso comparativa, basata su ideali ed aspettative che non richiede alcun acquisto o esperienza del prodotto/servizio; mentre la *soddisfazione* è più specifica, basata sulla misurazione effettiva e ha luogo durante e/o dopo l'uso.

Quindi, la prima cosa da fare volendo mettere a punto un sistema per la valutazione della soddisfazione del cliente è individuare i cosiddetti **Satisfaction Drivers** effettivi, ossia i fattori che determinano realmente la soddisfazione o l'insoddisfazione degli utenti.

Sulla base di queste considerazioni nel presente capitolo si pone grande attenzione alla fase di identificazione delle premesse operative, ossia dei fattori che entrano a far parte del processo di definizione del grado di soddisfazione globale. Occorre considerare, infatti, che ad influire su questo non sono solo le aspettative, fondate su elementi di carattere cognitivo come le passate esperienze, ma anche i desideri risultanti dalle motivazioni personali dell'individuo. Da ciò si deduce chiaramente l'importanza dell'impiego di tecniche di tipo psicometrico per lo studio delle scelte del consumatore e, quindi per la valutazione della sua soddisfazione.

## 10.2 DAI METODI TRADIZIONALI AI METODI DECISIONALI MULTICRITERIO

Il paradigma della conferma/disconferma delle aspettative è un modello di consumer-based usato quale misura della soddisfazione di cliente. In particolare, Oliver (Oliver,1977,1980,1981,1997,1999) in diversi articoli descrive il processo di soddisfazione come segue:

- i. I compratori formano le aspettative delle prestazioni di prodotto prima dell' acquisto.
- ii. Il consumo rivela un livello di rendimento percepito che è confrontato ai livelli di aspettativa che sono confermati o annullati.
- iii. Se è negativo o positivo il livello della disconferma, allora la soddisfazione del cliente rispettivamente diminuisce o aumenta.

Secondo Yi (Yi,1990) la conferma o la disconferma di aspettativa sulla qualità del prodotto/servizio svolge un ruolo chiave nella determinazione se un individuo è soddisfatto o no.

Un altro metodo simile a quello sopracennato, esamina i collegamenti fra le misure customer-based delle prestazioni di un'impresa come qualità percepita e soddisfazione di cliente e misure tradizionali derivanti dai dati contabili dei volumi d'affari, come ad esempio la percentuale del mercato e del profitto. Il livello di soddisfazione allora è valutato per mezzo degli strumenti econometrici.

I metodi statistici classici, come analisi di regressione lineare multipla non possono essere applicabili nella misurazione della soddisfazione del cliente a causa della forma qualitativa delle preferenze dei clienti.

I dati di input nell' analisi di regressione lineare non seguono una scala ordinale.

Altri metodi statistici basati su analisi di dati, come i modelli loglineari e di analisi del logit sono stati sviluppati per superare il problema accennato prima. I modelli Loglineari descrivono i modelli di associazione fra le variabili esplicative che modellano i valori delle celle in una tabella di contingenza in termini di associazioni fra



queste variabili. Purtroppo, i modelli loglineari non distinguono fra la risposta e le variabili esplicative (cioè soddisfazione globale e parziale dei clienti rispettivamente). D' altra parte, i modelli generalizzati del logit per le variabili esplicative sono equivalenti ai modelli loglineari per le tabelle multiway di contingenza.

Le tecniche di analisi di dati, come conjoint analysis, inoltre, sono state applicate nella misurazione della soddisfazione di cliente. La conjoint analysis è un metodo basato sulla misurazione degli scambi di concessioni dei clienti fra il prodotto e gli attributi di servizio. L'approccio seguito è quello di tipo decompositivo, poiché cerca di decomporre una valutazione globale di gradimento nelle sue componenti. In altri termini mira a cogliere il meccanismo attraverso il quale valutazioni relative a caratteristiche separate concorrono alla formazione della valutazione globale.

Un altro metodo nella misurazione del problema di soddisfazione dei clienti è l' analisi dei dati del questionario per mezzo degli strumenti semplici di rappresentazione grafica, come gli istogrammi di differenza ed i diagrammi di probabilità. Questo metodo considera la differenza fra i segni per l' attributo "atteso" e l' attributo "percepito" come indicatore per soddisfazione di cliente.

La soddisfazione del cliente è in genere riconducibile a cause diverse, alcune delle quali possono essere quantificate e misurate in termini di aderenza a caratteristiche tecniche, altre, invece, riguardano comportamenti, sensazioni e benefici psicologici e sono difficilmente esprimibili con indici. Le scale di misura che si utilizzano per rilevare questo tipo di informazioni sono basate su punteggi o su ordinamenti. Il ricorso a questi strumenti consente di creare un sistema di riferimento nelle scelte di un prodotto/servizio, in funzione della maggiore o minore preferenza del cliente verso un attributo o una particolare combinazione di attributi.

Il presente lavoro si propone l'obiettivo della valutazione della soddisfazione del cliente, individuando gli attributi o la combinazione di attributi maggiormente discriminante, cioè quelli che influiscono sul processo di valutazione dell'utente, rispetto ai quali misurare gli scostamenti tra atteso e percepito.

Per raggiungere l'obiettivo, anziché utilizzare le classiche metodologie statistiche, di cui abbiamo trattato nei precedenti capitoli, prenderemo in considerazione l'utilizzo di

funzioni di utilità non additive, nell'ambito dei cosiddetti integrali fuzzy, che permettono di modellizzare strutture di preferenza anche in presenza di interazione tra attributi.

La metodologia utilizzata ci consente la decomposizione di valutazioni globali in scale di utilità, corrispondenti a ciascuno degli attributi considerati, separate e comparabili, in modo tale che le valutazioni globali originarie possono essere correttamente ricostruite. Le ipotesi di base su cui poggia tale metodologia sono essenzialmente due: che il consumatore scelga tra prodotti/servizi alternativi in base al valore soggettivo da lui stesso assegnato ad ognuna di essi, che il valore di ciascun prodotto/servizio sia dato dalle combinazioni dei valori associati a ciascun livello degli attributi caratterizzanti il prodotto/servizio stesso.

La regola di composizione secondo cui il singolo consumatore aggrega l'utilità associata a ciascun attributo per ottenere il valore del prodotto/servizio, costituisce l'aspetto di primario interesse per chi si avvale di questa metodologia. Tale regola è generalmente nota come *modello di preferenza*.

### 10.3 LA METODOLOGIA

Il nostro scopo è presentare una metodologia, basata sulla logica Fuzzy, per rilevare la Customer Satisfaction (CS).

Tale metodologia consente di rilevare i punti di forza e di debolezza del prodotto/servizio esaminato. Sulla base dei dati ottenuti si può pianificare un miglioramento del prodotto/servizio. Numerosi studiosi hanno elaborato strumenti per la rilevazione della CS, perché comprendere i desideri dei clienti consente di ottenere dati indispensabili per il progetto di strategie di mercato vincenti. L'interesse per la rilevazione della CS è testimoniato anche dal ruolo che essa gioca nella nuova ISO 9001:2000 e nel modello EFQM.

In tale contesto s' inserisce questo lavoro che, sfruttando le peculiarità della logica Fuzzy, consente una rilevazione della CS svincolata dalle tipiche scale numeriche (Zollo et al.,1999; Zollo et al.,1996). Negli approcci tradizionali i clienti sono chiamati ad esprimere le loro valutazioni su una scala numerica; ciò costringe l' intervistato ad operare scelte che spesso alterano il suo reale giudizio. La nostra ipotesi è che usando la logica Fuzzy è possibile conoscere le attese del cliente in modo più affidabile.

Utilizzando i metodi tradizionali per rilevare la CS risulta difficile tradurre in dati oggettivi l' informazione espressa verbalmente.

La Fuzzy consente di trattare le valutazioni verbali senza eliminarne l' ambiguità e privilegiando la significatività del risultato piuttosto che la precisione. La Fuzzy utilizza operatori d' aggregazione (OWA), di cui abbiamo già trattato, in grado di comporre giudizi espressi su scale differenti da diversi valutatori, senza alterarne la vaghezza e l' ambiguità.

Nei metodi di rilevazione più diffusi, quali il SERVQUAL, i pesi degli elementi valutati sono attribuiti dal valutatore col rischio di avere una classificazione poco affidabile; la logica Fuzzy consente, invece, di ricavare i pesi da attribuire ai singoli elementi direttamente dalla coerenza dei giudizi espressi, garantendo una maggiore veridicità delle priorità emerse.

L'analisi procede con l' aggregazione dei giudizi, in modo da ottenere, per ogni caratteristica del prodotto/servizio, una valutazione globale. I dati sono stati aggregati utilizzando gli operatori OWA e dei *quantificatori* che sono funzioni con cui si pesano i giudizi raccolti.

In dettaglio, la tecnica prevede diverse fasi che possono essere schematizzate come segue:

- Selezione degli attributi
- Selezione dei livelli di ciascun attributo
- Determinazione delle combinazioni di attributi
- Scelta del modello di preferenza
- Scelta del metodo di raccolta dei dati
- Definizione delle modalità di presentazione degli stimoli agli intervistati
- Scelta del metodo di stima

Terminato il processo di aggregazione, è possibile capire quali caratteristiche, fra quelle esaminate, soddisfano maggiormente i valutatori e quali, invece, sono più insoddisfacenti. Ciò completa il processo di elaborazione dei dati, fornendo gli elementi per valutare quali siano le caratteristiche del servizio da migliorare, per aumentare la soddisfazione dei clienti e quali di esse costituiscano già dei punti di forza.

La potenza delle metodologie Fuzzy consente di analizzare giudizi espressi col linguaggio naturale, senza alterarne il loro contenuto di vaghezza ed ambiguità, pur pervenendo ad indicazioni di carattere operativo. I risultati sperimentali ottenuti conferiscono validità alla procedura adottata, aprendo l' orizzonte verso nuovi ambiti d' indagine.

Nell'ambito della teoria dell'utilità multi -attributo (Multi Attribute Utility Theory - MAUT) sono state proposte diverse metodologie per modellizzare una funzione di utilità di un decisore al fine di rappresentare le sue preferenze. Molto spesso si adotta una funzione di utilità additiva, la quale, però, basandosi su ipotesi molto restrittive quale l' indipendenza delle preferenze, non consente di modellizzare con sufficiente

elasticità alcuni importanti aspetti della decisione come per esempio l' interazione tra i criteri.

Si presenta una metodologia per costruire funzioni di utilità non additive, nell'ambito dei cosiddetti integrali fuzzy, che permettono di modellizzare strutture di preferenza in presenza di interazione tra i criteri.

In particolare si propone un metodo che, a partire da alcune preferenze su un insieme di attributi e da altre indicazioni circa l' importanza e l' interazione dei criteri, fissa i parametri di una funzione di utilità nella forma di un particolare integrale fuzzy, l'integrale di Choquet. I parametri di questa funzione di utilità sono di due tipi: 1) pesi che possono essere interpretati come "importanza" delle coalizioni di criteri; 2) funzioni di utilità marginale relative a ciascun criterio considerato, che consentono di esprimere valutazioni con riferimento a differenti criteri su un' unica scala di valutazione.

#### 10.4 PROBLEMA MULTICRITERIALE

In un problema di decisione multicriteriale (Roy, 1985, 1990; Vincke, 1992) il decisore valuta un insieme di azioni  $A = \{a, b, c, \dots\}$  sulla base di una famiglia  $G$  di  $n$  punti di vista o criteri (attributi decisionali)  $G = \{g_1, g_2, \dots, g_i, \dots, g_n\}$  con  $g_i: A \rightarrow \mathfrak{R}$ . In particolare, sia  $S_i$ ,  $i=1, \dots, n$ , la relazione binaria definita in  $A$  tale che  $\forall a, b \in A$ ,  $a S_i b$  significa "a è almeno tanto buona quanto b con riferimento al criterio  $g_i$ ", o anche "a surclassa b con riferimento a  $g_i$ ". Per ogni  $g_i$  e  $\forall a, b \in A$ , si ha:

$$g_i(a) \geq g_i(b) \implies a S_i b.$$

Sia  $S$  la relazione binaria definita in  $A$  tale che  $\forall a, b \in A$ ,  $a S b$  significa "a è complessivamente almeno tanto buona quanto b", o anche "a surclassa b".

Una funzione di utilità  $U: \mathfrak{R}^n \rightarrow \mathfrak{R}$  rappresenta le preferenze del DM, espresse dalla relazione  $S$  se per  $\forall a, b \in A$

$$U(g(a)) \geq U(g(b)) \iff a S b,$$

$$\text{ove } U(g(\cdot)) = U(g_1(\cdot), g_2(\cdot), \dots, g_i(\cdot), \dots, g_n(\cdot)).$$

Il più comune modello di funzione di utilità è quello additivo cioè

$$U(a) = \lambda_1 u_1(g_1(x)) + \lambda_2 u_2(g_2(x)) + \dots + \lambda_n u_n(g_n(x))$$

ove  $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathfrak{R}^+$  e  $u_i$ ,  $i=1, \dots, n$ , sono funzioni non decrescenti.

La possibilità di adottare un modello di tipo additivo nell'aggregare le preferenze di un decisore è vincolata al verificarsi di alcuni assiomi, principalmente l'indipendenza preferenziale tra criteri. Molto spesso, invece, si osserva che le preferenze del DM non soddisfano questo assioma a causa dell'interazione esistente tra diversi criteri.

L'interazione tra criteri si può manifestare sotto forma di interazione negativa o ridondanza (due criteri si dicono ridondanti se hanno un'importanza complessiva minore della somma dell'importanza dei criteri considerati singolarmente) e di interazione positiva o sinergia (due criteri si dicono sinergici se hanno un'importanza complessiva maggiore della somma dell'importanza dei criteri considerati singolarmente).

Per rappresentare in maniera adeguata l' interazione tra i criteri sonostate proposte alcune specifiche formulazioni della funzione di utilità espressa in termini di integrale fuzzy (Costanzo et al., 2000). Gli integrali fuzzy furono presentati per la prima volta nella tesi di dottorato di Sugeno (1974).

Si prenderà in considerazione l' integrale di Choquet , la cui interpretazione in termini di integrale fuzzy venne proposta da Murofushi e Sugeno .

Nella teoria di aiuto alle decisioni si pongono essenzialmente due tipi di problemi:

- 1) Quali proprietà le preferenze del decisore devono soddisfare perché si possa rappresentarle mediante una funzione U con una data forma funzionale (additiva, moltiplicativa, mista etc.) ?
- 2) Quali metodologie adottare per costruire tali funzioni U e come attribuire i valori ai parametri che si riferiscono ad una data forma analitica di U ?

Con riferimento al secondo ordine di problemi, esistono due grandi famiglie di metodi per la determinazione di U in forma additiva: metodi diretti e metodi indiretti.

I metodi diretti consistono nella determinazione delle U ponendo delle opportune domande al decisore.

I metodi indiretti, di cui il più noto è il metodo UTA, consistono nella costruzione della U a partire dalle preferenze complessive del DM su un sottoinsieme A' di azioni di A.

Nel metodo UTA il modello di aggregazione dei criteri assume la forma di una funzione additiva:

$$u(g(.)) = \sum_i^n u_i(g_i)$$

$$u_i(g_{i^*}) = 0$$

$$\forall i = 1, 2, \dots, n$$

$$\sum_{i=1}^n u_i(g_i^*) = 1$$

#### 10.5 LE FASI DELLA RICERCA – I L'INTEGRALE DI CHOQUET E DI SUGENO

La procedura seguita per la realizzazione del presente lavoro si articola in più fasi. Non si ritiene, infatti, che per risolvere problemi connessi alla misurazione della qualità di un prodotto/servizio ad alto coinvolgimento socio-psicologico, la strategia più adatta possa essere rappresentata da una semplice rilevazione delle preferenze dei consumatori.

Inizialmente, si è proceduto all'individuazione di soggetti capaci di rappresentare i clienti potenziali e si è proceduto all'elaborazione di un questionario al fine di individuare quali sono gli attributi chiave su cui si sofferma il cliente per stabilire se il prodotto/servizio è di qualità o no.

Gli attributi così determinati sono utilizzati per stabilire, attraverso descrizione verbale dell'intervistato, eventuali profili che questi deve ordinare secondo proprie preferenze in modo crescente.

Quindi ci si è posti il problema della individuazione degli attributi maggiormente discriminanti, cioè di quelli che influiscono sul processo di valutazione dell'utente, rispetto ai quali misurare gli scostamenti tra atteso e percepito.

L'ordinamento in funzione del gradimento permette di individuare quello che in letteratura è noto come sistema della qualità attesa, mentre quello che viene indicato definisce la qualità percepita dagli utenti cioè la distanza tra i due ci dà un'indicazione utile ai fini della valutazione della soddisfazione del cliente.

Nel prossimo capitolo, presenteremo un metodo indiretto di fissazione dei parametri di una funzione di utilità espressa in termini di integrale di Choquet.

Il nostro approccio si ispira al lavoro di Marichal e Roubens (1999) che propongono un metodo indiretto di determinazione dei pesi interattivi per rappresentare le preferenze mediante una funzione di utilità espressa in termini di integrale di Choquet (Angilella et. Al., 1999). Tuttavia, nel loro approccio, Marichal e Roubens ipotizzano che esista come dato esogeno del problema una scala comune che permetta di confrontare valutazioni su criteri differenti. Il nostro approccio differisce in quanto si fa cadere



questa ipotesi e pertanto la scala comune per confrontare le valutazioni con riferimento a differenti criteri del problema viene determinata insieme ai pesi interattivi.

Dal punto di vista analitico il problema risulta ben più complesso del problema di Marichal e Roubens (1999) il quale si riduce a un problema di programmazione lineare. Analiticamente, invece, si presenta come un problema di programmazione non lineare e non differenziale ove i vincoli e la funzione obiettivo hanno una formulazione particolarmente complessa dipendente anche dall'ordinamento dei valori della scala comune a tutti i criteri.

Come vedremo, non si presenta una soluzione analitica e esatta del problema e si presenta invece una soluzione euristica e (talvolta) approssimata al problema. La metodologia proposta trova una sua giustificazione dal punto di vista decisionale, nell'idea di base del cosiddetto approccio costruttivo al supporto alle decisioni (Roy, 1993), secondo il quale non ha senso ipotizzare una qualche funzione di utilità preesistente al problema decisionale stesso e invece la funzione di utilità, come ogni altro strumento utilizzato come supporto alla decisione vengono costruite durante il processo stesso di supporto alla decisione. In quest'ottica non ha senso parlare della funzione di utilità ottima che esiste a priori anche se non si hanno gli strumenti analitici per calcolarla. D'altra parte, dal punto di vista analitico, anche la costruzione di una soluzione approssimata si basa su alcune raffinate tecniche di programmazione lineare (descrizione completa del semplice). E dal punto di vista decisionale la soluzione proposta, anche quando approssimata, possiede alcune caratteristiche di robustezza che la legittimano come ragionevole base per un proficuo processo di supporto alla decisione.

Gli input del problema sono

- un insieme  $A$  di azioni, che rappresentano le descrizioni verbali degli utenti, e un insieme  $G$  di criteri, che rappresentano gli attributi di cui si vuole individuare il valore discriminante,
- le valutazioni di ogni azione con riferimento a tutti i criteri, ovvero la valutazione globale della qualità percepita,

- un preordine parziale di  $G$  (cioè un ordinamento dei criteri in base alla loro importanza relativa)
- un preordine parziale sull'insieme delle coppie non ordinate di criteri (cioè l'ordinamento di coppie di criteri in base alla forza della loro interazione)
- il segno dell'interazione di alcune coppie di criteri

Gli output del problema sono:

- le misure associate ad ogni criterio rappresentanti l'importanza relativa e l'interazione,
- le funzioni di utilità marginale  $u_i$  associate.

Gli ordinamenti, trattati come variabile di risposta dell'esperimento, ci permettono di stimolare l'importanza delle caratteristiche individuate nello spiegare le attese del prodotto/servizio. Le modalità risultate più importanti sono quindi combinate insieme per definire il prodotto/servizio ideale.

Da un punto di vista formale il problema si può formulare come segue: trovare una misura 2-additiva e i valori delle funzioni di utilità marginale  $u_i(g_i(x)) \forall g_i \in G$  e  $x \in A$  che soddisfano le seguenti condizioni:

$$C_\mu(a) > C_\mu(b), \forall a, b \in A \text{ tale che } aPb,$$

$$C_\mu(a) = C_\mu(b), \forall a, b \in A \text{ tale che } aIb,$$

$$\phi(g_i) \geq \phi(g_j), \text{ se } g_i \text{ è almeno tanto importante quanto } g_j,$$

$$\phi(g_i) = \phi(g_j), \text{ se } g_i \text{ e } g_j \text{ hanno la stessa importanza,}$$

$$\phi(g_i, g_j) \geq \phi(g_k, g_l), \text{ se l'interazione tra } g_i \text{ e } g_j \text{ è almeno tanto forte quanto l'interazione tra } g_k \text{ e } g_l,$$

$$\phi(g_i, g_j) = \phi(g_k, g_l), \text{ se l'interazione tra } g_i \text{ e } g_j \text{ è uguale all'interazione tra } g_k \text{ e } g_l,$$

$$\phi(g_i, g_j) \geq 0, \text{ se l'interazione tra } g_i \text{ e } g_j \text{ è non negativa,}$$

$$\phi(g_i, g_j) \leq 0, \text{ se l'interazione tra } g_i \text{ e } g_j \text{ è non positiva,}$$

$$\phi(g_i, g_j) = 0, \text{ se non c'è alcuna interazione tra } g_i \text{ e } g_j.$$

Si osservi quanto segue:

- 1) i valori delle funzioni di utilità marginale  $u_i(g_i(x)) \quad \forall g_i \in G \text{ e } x \in A$  sono incognite del problema: questo differenzia l' approccio proposto dall' approccio di Roubens e Marichal (1999) i quali confrontano valutazioni di criteri differenti, supponendo perciò (almeno implicitamente) che le valutazioni sui vari criteri siano espressi su scale omogenee. Una disamina delle assunzioni teoriche sottostanti all' ipotesi di esistenza di un' unica scala omogenea per ogni criterio è stata condotta da Modave e Grabisch (1998).
- 2) formalmente il problema si presenta come un problema di ottimizzazione non-lineare: nell' approccio di Roubens e Marichal (1999), l' ipotesi di conoscere le funzioni di utilità marginale consente di impostare invece il problema in termini di programmazione lineare.
- 3) la formulazione delle funzioni di utilità in termini di integrale di Choquet dipende dall' ordinamento delle valutazioni delle azioni di  $A$  rispetto ai criteri di  $G$  nei termini delle funzioni di utilità marginale  $u_i(g_i(x))$ .

Data la complessità del problema risulta estremamente complesso, se non praticamente impossibile, pensare di risolvere con metodi analitici il problema. Si propone perciò una soluzione euristica del problema. In particolare, si propone la più semplice euristica: il metodo "Monte Carlo". In termini generali l' algoritmo di soluzione del problema procede mediante i seguenti passi:

- 1) generazione casuale di un certo numero di misure  $\mu$ , che rispettino i vincoli del problema (vincoli sulle importanze e sulle interazioni e vincoli sull' ordinamento delle valutazioni sui singoli criteri con riferimento alle utilità marginali)
- 2) calcolo delle utilità complessive dei vari clienti in termini di integrali di Choquet relativi alle misure  $\mu$  per ogni singolo criterio e alle utilità marginali considerate
- 3) calcolo del numero di inversioni tra l' ordinamento fornito dal decisore (qualità percepita) e l' ordinamento risultante dalle utilità calcolate al punto 2) (qualità attesa), ovvero calcolo della **somma dei quadrati delle differenze tra percepito ed atteso**

- 4) selezione delle misure e delle funzioni di utilità marginali in corrispondenza delle quali il numero di inversioni calcolate al punto 3) sia minimo, ovvero l'inversione **minima della somma dei quadrati delle differenze**.

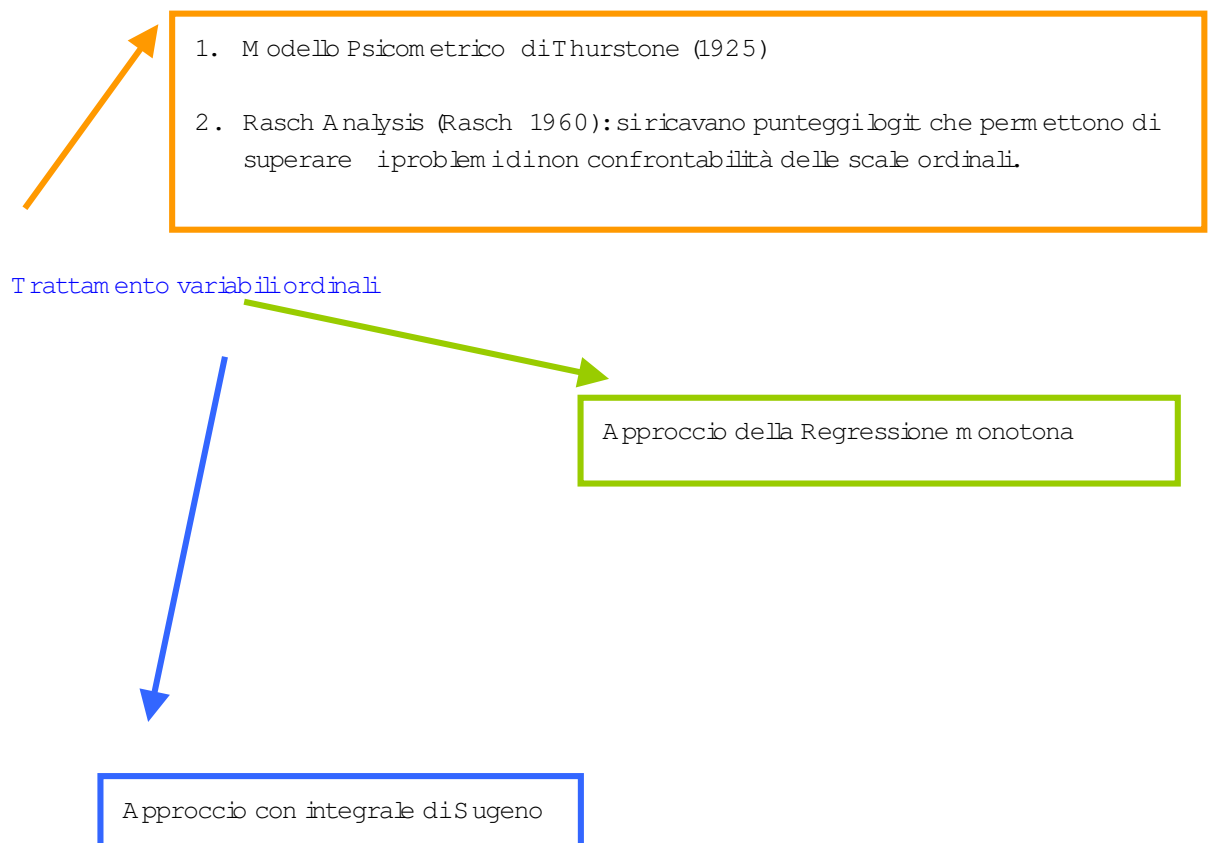
Ovviamente, essendo una procedura euristica, l' algoritmo proposto non dà alcuna garanzia circa l' effettiva esistenza di misure e di funzioni di utilità marginali le cui utilità complessive dei vari clienti, in termini di integrali di Choquet, rispettino le preferenze del decisore senza creare alcuna inversione rispetto al preordine totale fornito dal decisore. Tuttavia, nell' ambito di un approccio costruttivo al supporto alla decisione, si conviene che l' obiettivo non è di esplicitare una ipotetica funzione di utilità già preesistente al problema di decisione stessa, quanto piuttosto fornire al decisore degli strumenti che gli consentano di progredire verso una maggiore comprensione del problema di decisione. La funzione di utilità ottenuta rappresenta, perciò, uno strumento utile per chiarire alcuni aspetti del problema. In tale prospettiva la validità dello strumento non dipende tanto dal fatto di riuscire ad approssimare inequivocabilmente delle presunte preesistenti preferenze del decisore, quanto piuttosto nella capacità di aiutare il decisore a costruirsi delle preferenze robuste. Se a seguito della generazione di un gran numero valori della misura 2-additiva e delle funzioni di utilità marginali non si riesce a trovare dei parametri che permettano di rappresentare le preferenze del decisore, allora questo è quanto meno indice della scarsa robustezza della eventuale funzione di utilità che rappresenta correttamente le preferenze del decisore. Essa, infatti, se esiste si baserebbe su un insieme di possibili valori dei parametri considerati particolari: particolari nel senso che non si riescono a trovare possibili parametri neppure tra un gran numero di insiemi di parametri generati casualmente.

La metodologia proposta offre notevoli vantaggi rispetto agli approcci tradizionali, perché basata su una procedura di aggregazione delle preferenze che, oltre a non basarsi sull'ipotesi di indipendenza delle preferenze tipica dell'utilità additiva, permette di rappresentare interazioni positive e negative tra i criteri. E' utile rimarcare che la funzione di utilità espressa nei termini dell' integrale di Choquet presuppone l' esistenza di una scala comune che permetta di confrontare valutazioni su criteri differenti.

A differenza di altre proposte metodologiche, tale scala comune non viene supposta più o meno artificiosamente preesistente alla formulazione del problema di decisione, ma viene invece determinata insieme agli altri parametri che caratterizzano la funzione di utilità. Le informazioni richieste al decisore sono perciò minime: questi deve solamente fornire un preordine (eventualmente) parziale su un insieme di azioni. Opzionalmente il decisore può stabilire alcune relazioni di importanza tra i criteri o indicare interazioni tra questi.

La nostra ricerca in merito agli integrali fuzzy è stata estesa all'integrale di Sugeno.

In molti problemi di valutazione di Customer Satisfaction si ha un trattamento dei dati su scale ordinali. Abbiamo già discusso dell'importanza del problema nel capitolo 4.



Con l'approccio psicometrico di Thurstone consideriamo un questionario definito da  $K$  variabili ordinali  $\mathbf{X}_k$  ( $k = 1, K, K$ ) con egual numero di categorie  $I$  ( $i = 1, K, I$ ) di temine generale  $x_{ki}$ .

Sia  $\Pr(\mathbf{X}_{ki} = i) = p_{ki}$  la probabilità che la variabile  $k$  sia uguale a  $i$

$F_k(i) = \sum_{\{j \in S, j \leq i\}} p_{kj}$  è la probabilità che abbiamo almeno una delle prime  $i$  categorie, con

$$S = \{1, K, I\}.$$

A ciascuna variabile categorica ordinata (i giudizi)  $X$  corrisponde una variabile casuale latente di tipo normale  $Z$  ("true" scale).

In questo caso utilizziamo una scala ad intervallo.

A ciascun valore ordinato  $x_{ki}$  corrisponde un valore  $z_{ki} = \xi_i$  tale che  $\xi_i < \xi_{i+1} \quad \forall k$

( $i = 1, \dots, I$ ) definendo, per ciascuna distribuzione  $Z_k$ , il quantile associato alla categoria

$x_{ki} = i$  della variabile  $X_k$ . Da cui le seguenti identità:

$$\Phi\left(\frac{\xi_i - \mu_k}{\sigma_k}\right) = F_k(i)$$

$$\frac{\xi_i - \mu_k}{\sigma_k} = \Phi^{-1}[F_k(i)] = \zeta_{ki}$$

$$\begin{aligned} (i &= 1, K, I) \\ (k &= 1, K, K) \end{aligned}$$

Se  $\Phi(\cdot)$  è la funzione cumulativa di una variabile normale standardizzata allora,  $\mu_k$  e

$\sigma_k^2$  sono, rispettivamente, la media e la varianza della variabile normale latente  $Z_k$ .

$\zeta_{ki}$  è il valore osservabile (reiterando le  $X$  osservazioni) non superabile da una

variabile casuale normale con probabilità  $F_k(i)$ .

Consideriamo la media di  $\zeta_{ki}$  rispetto a K, abbiamo che  $\xi_i^* = (\bar{\sigma}\xi_i - \bar{\mu}) = \bar{\zeta}_i$ .

Se poniamo  $\bar{\sigma} = (1/K) \sum_{k=1}^K (1/\sigma_k)$  (reciproco della media armonica dei quadrati

medi delle differenze della variabile latente  $Z_k$ ),  $\bar{\mu} = (1/K) \sum_{k=1}^K \mu_k / \sigma_k$  e

$\bar{\zeta}_i = (1/k) \sum_{k=1}^K \zeta_{ki}$ , otteniamo i valori di  $\xi_i^*$  che sono compatibili con i “veri”

valori latenti  $\xi_i$ .

In tal modo, sostituiamo le originali variabili categoriche ordinali con i valori  $\xi_i^*$  definiti su una scala ad intervallo.

Con l’approccio della regressione monotona di Kruskal si ha che: Sia Y una variabile dipendente ordinale (convenzionalmente rappresentata da un insieme di punteggi)

$$y_1 \leq \dots \leq y_n \quad \text{e siano} \quad \hat{z}_1 \leq \dots \leq \hat{z}_n$$

le trasformazioni delle variabili di risposta secondo una funzione di trasformazione  $f(\cdot)$ , scelta in modo da rispettare l’ordinamento iniziale.

Si pone

$$z_j(\beta) = \sum_{s=1}^m g_{js} \beta_s$$

$g_{js}$  sono valori numerici noti che descrivono i livelli dei fattori esplicativi

$\beta_s$  i coefficienti

Indichiamo con  $\bar{z}_j(\beta)$  il valore di media aritmetica

Si procede in modo iterativo a determinare il “direct stress” (Kruskal 1965)

$$S(\hat{f}^*; \hat{\beta}^*) = \min_{\beta} \left\{ \min_f \left[ \frac{\sum_{j=1}^n (\hat{z}_j - z_j(\beta))^2}{\sum_{j=1}^n (z_j(\beta) - \bar{z}_j(\beta))^2} \right]^{1/2} \right\}$$

essa rappresenta una misura descrittiva della bontà di adattamento della trasformazione monotona  $f(\cdot)$

L'algoritmo di minimizzazione viene descritto in Kruskal (1965).

La Rasch Analysis (Rasch 1960) è una tecnica appartenente ai Latent Trait Models. Essa consiste nella trasformazione degli indicatori qualitativi in quantitativi, definiti su numeri naturali o reali.

La scelta dei questionari è tale che gli indicatori utilizzati sottendono un solo aspetto latente sulla base dell'abilità degli individui, delle difficoltà delle prove, dei diversi livelli di riuscita, (ad es. capacità motoria, capacità intellettiva dei disabili) e secondariamente attraverso l'utilizzo di un modello logistico che rende additive e quantitative le scale ordinali attraverso cui sono espressi i risultati dei questionari

La **Rasch Analysis** (Rasch 1960) consente di utilizzare i punteggi *logit* quali **misure oggettive**, indipendenti dal campione di individui e dalla tipologia del questionario, superando i problemi della non confrontabilità delle valutazioni espresse su scale ordinali.

Si è ribadito che un usuale operatore aritmetico non può aggregare valori ordinali. E', quindi, necessario ricorrere ad un aggregatore tale da riflettere la valutazione qualitativa. In questo contesto l'integrale di Sugeno appare un potenziale candidato a risolvere il problema. Infatti l'integrale di Sugeno è riconosciuto come la naturale controparte dell'integrale di Choquet nel caso di insiemi ordinali, dove la somma è sostituita dal *max* (nel caso di integrale discreto) e il prodotto dal *min*.

Abbiamo, pertanto, proposto un'applicazione con l'utilizzo dell'integrale di Sugeno, nell'ipotesi in cui gli attributi di valutazione sono espressi in termini qualitativi.



## 10.6 LE FASI DELLA RICERCA – II – L'INTEGRALE GERARCHICO

Un altro interessante approccio si basa sulla possibilità di determinare una valutazione delle varie componenti, secondo il già detto approccio decompositivo, che spinge l'analisi di CS verso livelli sempre più disaggregati..

La logica della disaggregazione dei dati è utilizzata molto spesso in modelli di analisi multicriteriale.

Si citano solo ad esempio i metodi ELECTRE III, ELECCALC, ELECTRE TRI, UTA method, MARKEX, MUSA system, MACBETH system, MIIDAS system.

Lo schema generale della filosofia della disaggregazione è altresì impiegato in altri approcci, come i rough sets, il machine learning e le reti neurali.

Nell'ambito in cui formuliamo la nostra ricerca è sicuramente importante la conoscenza della percezione della soddisfazione di ogni singolo attributo. E' ormai universalmente noto che la qualità del prodotto/servizio, così come è intesa dai consumatori, si può definire come il grado di discrepanza tra le aspettative o i desideri dei clienti e le loro percezioni.

Intendiamo proporre un'analisi multicriteriale per la valutazione di un insieme delle funzioni marginali di soddisfazione rappresentanti il livello di ogni criterio. E' possibile determinare degli indici mostrano livello di soddisfazione parziale dei clienti secondo ogni sub-criterio, similmente all' indice globale di soddisfazione.



E' possibile condurre questa fase di analisi al fine di definire gli attributi più interessanti per procedere alla determinazione dell'indice globale di soddisfazione.

La fase di decomposizione della soddisfazione globale ci consente di determinare i macro-attributi, rilevatori chiave delle aspettative.

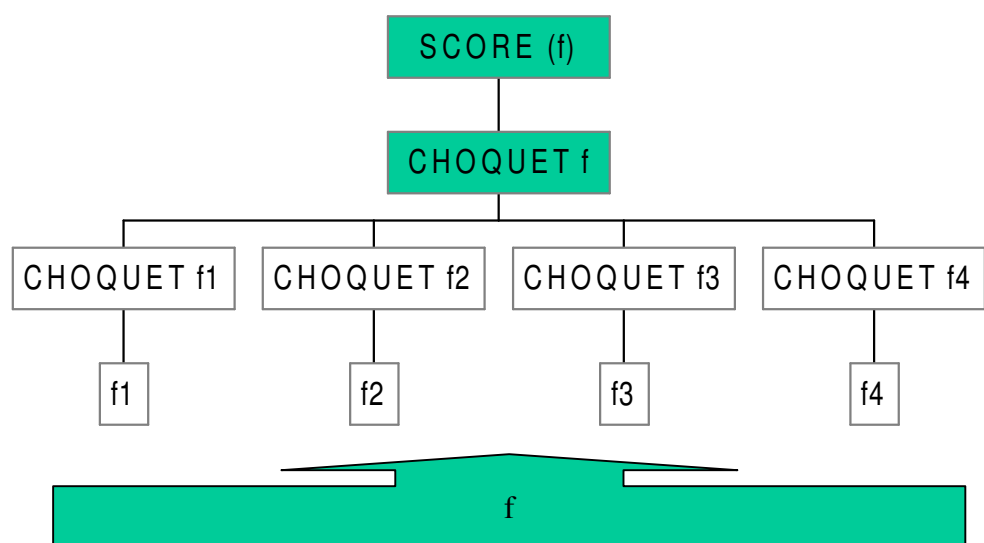
La necessità di ridurre al minimo il numero degli attributi (criteri) da utilizzare è dettata fondamentalmente dalla limitata capacità di valutazione comparativa dell'uomo, che, come dimostrato nel corso degli anni da vari psicologi (Miller 1964), non riesce a considerare simultaneamente più di sei/sette informazioni diverse.

Il numero degli stimoli ottenibili considerando tutte le dimensioni, sarebbe eccessivo, e rischierebbe di sovraccaricare gli intervistati, i quali non valuterebbero tutte le alternative proposte, ma solamente le prime o quelle che attirano maggiormente la loro attenzione, giudicando in modo frettoloso e superficiale tutte le altre.

L' obiettivo principale di analisi di decomposizione è identificare particolari aree di clienti con distinte preferenze e aspettative.

Anche per la procedura di decomposizione abbiamo utilizzato una funzione di utilità espressa in termini di integrale di Choquet.

### MODELLO INTEGRALI FUZZY STRUTTURA GERARCHICA



Siano  $x \in A$  e  $\mu$  una misura fuzzy sull'insieme  $G$ . Allora l'integrale di Choquet è definito da:

$$C(x, \mu) = \sum_{i=1}^n (x_{\pi(i)} - x_{\pi(i-1)}) \mu(A_{\pi(i)}) = (C) \int_X f d\mu$$

ove le valutazioni dell'utilità con riferimento a ciascun criterio sono riordinati in modo che:

$$x_{\pi(1)} \leq x_{\pi(2)} \leq \dots \leq x_{\pi(n)}$$

si pone  $A_{\pi(j)} = \{\pi(j), \dots, \pi(n)\}$  definito per ogni  $j=1, \dots, n$ .

#### TEOREMA SU DECOMPOSIZIONE DI INTEGRALI GERARCHICI

(Murofushi, Sugeno, Fujimoto, 1997)

Sia  $(X, \mathfrak{K})$  uno spazio misurabile, sia  $S$  un sottoinsieme non vuoto di  $X$ , sia  $x_S$  un punto non appartenente a  $X \setminus S$ , e sia

$$X_S = (X \setminus S) \cup \{x_S\},$$

$$\mathfrak{K}_{X_S} = \{ (A \setminus S) \cup B \mid A \in \mathfrak{K}, B = \emptyset \text{ o } \{x_S\} \},$$

$$\mathfrak{K} \cap S = \{ A \cap S \mid A \in \mathfrak{K} \}$$

Per una funzione  $f$  su  $X$  e una misura  $\nu$  su  $\mathfrak{K} \cap S$ , una funzione  $f_\nu$  su  $X_S$  è definita da:

$$f_\nu(x) = \begin{cases} f(x) & \text{se } x \in X \setminus S, \\ (C) \int_S (f \upharpoonright S) d\nu & \text{se } x = x_S, \end{cases}$$

dove  $f \upharpoonright S$  è la restrizione di  $f$  su  $S$  e  $(C) \int_S d\nu$  è l'integrale di Choquet rispetto a  $\nu$  su  $S$ .

Allora per ogni misura non-additiva  $\mu$  su  $\mathfrak{K}$ , esiste una misura non additiva  $\nu$  su  $\mathfrak{K} \cap S$  e una misura non additiva  $\lambda$  su  $\mathfrak{K}_{X_S}$  tale che per ogni funzione  $f$  su  $\mathfrak{K}$

$$(C) \int_X f d\mu = (C) \int_{X_S} f_\nu d\lambda$$

se e solo se  $\{S, X \setminus S\}$  è una partizione  $\mu$ -inter-additiva.

Una partizione di  $A$  è  $\mu$ -inter-additiva se è una partizione finita  $P$  di  $A$  tale che per ogni  $B \in A \cap P$  si ha che:

$$\mu(B) = \sum_{P \in P} \mu(P \cap B)$$

Si veda (Murofushi, Sugeno, Fujimoto 1997).

A quanto detto si associa il Teorema della indipendenza di preferenza dei modelli con integrali di Choquet, (PREFERENTIAL INDEPENDENCE THEOREM FOR CHOQUET INTEGRAL MODELS) che assicura la indipendenza nelle preferenze anche su sottoinsiemi  $S$  se  $\{S, X \setminus S\}$  è una partizione  $\mu$ -inter-additiva.

Si veda (Murofushi e Sugeno, 2000)

La possibilità di aggregazione di integrali gerarchici ci consente di esprimere un'analisi più compiuta perché essa può essere spinta fino ad esaminare i sub-attributi decisionali.

Il piano considerato permette di stimare anche i sub-effetti, in quanto contemplando le interazioni consente di avanzare delle ipotesi sul livello dei singoli sub-attributi.

L'approccio decompositivo, infatti, cerca di decomporre una valutazione globale di gradimento nelle sue componenti. In altri termini mira a cogliere il meccanismo attraverso il quale valutazioni relative a caratteristiche separate concorrono alla formazione della valutazione globale.

E' molto interessante qui citare un altro articolo di Murofushi, Sugeno, Fujimoto (1995), che sotto certe assunzioni determina una condizione necessaria e sufficiente per la decomposizione "sovrapposta" dell'integrale di Choquet. Si tratta del caso in cui i sub-integrali sono calcolati considerando l'intera azione tra le varie sub-funzioni.

### 10.7 LE FASI DELLA RICERCA III - IL CUSTOMER PROFILING COME PROBLEMA DI CLASSIFICAZIONE

Vista la centralità assunta dal cliente negli orientamenti d'impresa, si può affermare che di particolare interesse risulta la comprensione della clientela in termini di motivazioni, esigenze, atteggiamenti, attitudini, percezioni, preferenze, scelte e soddisfazione/insoddisfazione che quest'ultima esprime o può esprimere nei confronti dell'azienda. Diviene fondamentale comprendere, anche tramite processi analitici, come i comportamenti della domanda prendono forma, al fine di porre le basi per le decisioni e le scelte aziendali indirizzate e rivolte a sviluppo e mantenimento della relazione con la clientela.

La comprensione del cliente e dei suoi comportamenti prende le mosse dal sistema cognitivo del cliente, basilare per il customer profiling.

In questo scritto non affrontiamo tutte le componenti del processo di analisi, prendiamo solo in considerazione le tecniche di segmentazione. In particolare utilizziamo gli algoritmi degli aggregatori fuzzy per determinare una classificazione-segmentazione della clientela.

Il processo di segmentazione della clientela è un processo chiave in quanto, se svolto appropriatamente, consente di raggiungere una conoscenza reale della struttura del portafoglio clienti. Tale conoscenza è fondamentale, in quanto costituisce le fondamenta per l'identificazione dei target di clientela, la scelta dei prodotti da mettere sul mercato, l'impostazione del marketing mix: in breve, per l'intera azione commerciale.

Il processo di segmentazione ha inizio con l'individuazione, tra i possibili fattori di segmentazione, di quelli più significativi. I possibili fattori di segmentazione non si limitano al reddito e ai dati socio-demografici, ma si allargano ad includere i bisogni, le attitudini, i comportamenti, i canali preferiti e quant'altro differenzi tra loro i clienti. In via generale, una banca potrebbe suddividere i clienti a seconda del valore (attuale/potenziale) e differenziarli a seconda dei loro comportamenti (spesso indicatori anche dei loro bisogni). Il passo successivo è costituito dall'elaborazione del *database*

esistente e dall'analisi dei comportamenti storici, cui segue, in caso, lo svolgimento di ricerche di mercato ad hoc per completare la base informativa. Sulla base delle informazioni trovate, si procede ad una prima individuazione di gruppi, o segmenti, di clienti, che sarà poi perfezionata grazie ai *feedback* e all'apprendimento: nella prassi, infatti, i processi di tipo iterativo sembrano offrire i migliori risultati. La difficoltà sta nel trovare un equilibrio tra la raffinatezza della segmentazione e la sua utilizzabilità operativa; i due estremi da evitare sono, da una parte, di investire massicciamente in studi di segmentazione con modelli molto sofisticati e particolareggiati, che possono rivelarsi di difficile implementazione; dall'altra, di affidarsi ad approcci troppo semplicistici, che rischiano di produrre *cluster* di clienti non significativi dal punto di vista del marketing. I progressi tecnologici stanno avendo un profondo impatto sul modo di condurre e sulla stessa impostazione del processo di segmentazione, permettendo la costituzione di segmenti sempre più ristretti di utenti omogenei, fino al limite di segmenti costituiti dal singolo cliente (*"segment-of-one"*).

In realtà, proprio la disponibilità di strumenti tecnologici avanzati permette oggi alle aziende di svolgere analisi di segmentazione *ad hoc* per ogni singola azione di marketing. Nel caso dunque si intraprenda una campagna tattica (quindi il focus è sull'offerta di uno specifico prodotto/servizio, piuttosto che sull'offerta ad uno specifico segmento) si tende a privilegiare, ad una segmentazione di tipo statico, una segmentazione di tipo tattico.

Dal punto di vista statistico-matematico si individuano diversi metodi di segmentazione:

- **tecniche di segmentazione binaria**, che, per suddividere una popolazione in sub-popolazioni utilizzano un carattere "dipendente", di natura comportamentale (ad esempio: uso/non uso di un servizio). I segmenti individuati vengono poi descritti tramite i caratteri "indipendenti" o descrittori, che danno luogo ai profili dei segmenti. (*segmentazione di tipo "tattico"*, formulata per una determinata campagna)
- tecniche che ripartiscono una popolazione operando "direttamente" su descrittori (*cluster analysis*), senza far intervenire un carattere dipendente (*segmentazione di tipo "statico"*)

- l' approccio della funzione **discriminante**, che assegna a ogni azione (cliente) un valore che viene confrontato con opportune soglie che separano le classi considerate (analisi discriminante, nell' analisi PROBIT e LOGIT, nel metodo UTADIS, etc.)
- l' approccio dei **metodi di surclassamento**, che fissa alcune azioni (profili) soglia tra una categoria e l' altra e stima se le azioni (clienti) considerate sono preferite o meno a queste azioni soglia (ELECTRE TRI, MAPPAC SORT)
- l' approccio delle **regole di decisione**, che fissa una serie di regole "se ..., allora ..." che permettono di assegnare le azioni (clienti) nelle classi considerate. Quest' ultimo approccio è tipicamente considerato all' interno dell' intelligenza artificiale nei campi del Data Mining, del Machine Learning, del Knowledge Discovery. L' approccio dei rough sets basato sulla dominanza già presentato nei capitoli precedenti è l' unico tra questi metodi che considera l' ordine di preferenza stabilito dai criteri sull' insieme delle azioni. Il risultato principale di questo studio è che gli ultimi tre approcci, e per il quinto caso in particolare (l' approccio dei rough sets basato sulla dominanza), hanno una comune base assiomatica che sorprendentemente si fonda su un unico postulato che assicura che le classi preferenzialmente ordinate in cui sono ripartite le azioni inducono un ordine di preferenza su ciascun criterio.

La segmentazione dei clienti può considerarsi come un caso tipico di classificazione (sorting). Essa consiste nell'attribuire a classi (o categorie) un insieme, finito o evolutivo, di azioni di varia natura, intendendo per azione (Roy, 1996) la rappresentazione di un possibile contributo ad una decisione.

E' nostro scopo presentare l'utilizzo di tale problematica nel contesto di segmentazione.

Le classi possono indicare livelli diversi di accettabilità, adeguatezza o rischio, oppure differenti tipologie di 'utilizzo' delle azioni candidate. Possono derivare da una distinzione netta tra accettazione e rifiuto, oppure da una più sfumata in cui ad esempio l'accettazione è condizionata ad interventi migliorativi.

Se le azioni fossero in competizione diretta le une con le altre le si potrebbe confrontare ed ordinare in classi di preferenza decrescente. Nell'ambito di questa problematica decisionale (detta anche segmentazione, selezione o problematica  $\beta$ ) si vuole invece

aiutare il decisore a pronunciarsi sull'attitudine di ogni specifica azione a soddisfare determinati requisiti; è quindi il valore intrinseco delle azioni, rispetto al problema in esame ed al modello di riferimento adottato, che ne determina l'assegnazione ad una delle classi.

Passo essenziale e caratterizzante della problematica del sorting è la definizione delle *norme di assegnazione*, o *modello di riferimento* per il sorting, cioè della *procedura di assegnazione* e delle *azioni di riferimento*. Il modello di riferimento è sviluppato per poter rendere operative, nella procedura di sorting, normative, standard o protocolli che vincolano nella selezione di una azione candidata. Oltre a ciò il modello deve includere le esigenze, in termini di caratteristiche di 'qualità' e di assenza di particolari criticità, che un decisore, o più frequentemente una commissione giudicante, deve definire nella fase iniziale di una procedura di valutazione e selezione di azioni candidate.

La definizione della *procedura di assegnazione* comprende *struttura*, *numero* delle classi e *regole di attribuzione* dei candidati alle classi, definizione che deve avvenire a priori, al di fuori di ogni considerazione di merito circa le azioni candidate e coerentemente con lo specifico problema in esame. Molteplici possono essere gli elementi di conoscenza utilizzabili per la definizione formale delle *azioni di riferimento*. Possono derivare, ad esempio, da preesistenti modelli normativi e da standard, ad un buon livello di formalizzazione. Possono invece scaturire da un modello elaborato localmente che, anche se non particolarmente strutturato, potrebbe essere scaturito da un'esperienza magari pluriennale. Le azioni di riferimento, cioè quelle azioni con cui saranno confrontate le azioni candidate per essere selezionate, possono essere il risultato della modellizzazione di decisioni già prese in precedenza, cioè azioni analoghe per tipologia alle azioni candidate, ma di cui è possibile testare la reale corrispondenza alle attuali aspettative. Oppure le azioni di riferimento possono essere definite ad hoc dal decisore o dai soggetti coinvolti nella decisione, per rappresentare le caratteristiche e gli elementi significativi di ogni classe. Possono quindi essere *reali*, risultato di un'attenta analisi dei dati o di azioni di monitoraggio su precedenti situazioni reali oppure dedotte da altre situazioni simili, nella stessa o in altre organizzazioni. Possono essere *fittizie*, elaborate analiticamente sulla base di uno sviluppo iterativo di dimensioni, criteri, stati di qualificazione e valutazione ed essere integrate, convalidate o riformulate in un



processo di ‘negoiazione’ tra i soggetti che nel problema in esame assumono un ruolo decisionale.

I criteri con cui valutare le azioni candidate ed i coefficienti di importanza relativa da attribuire a questi criteri devono essere definiti e modellizzati congiuntamente con le azioni di riferimento, per caratterizzare il modello di riferimento da adottare nello specifico problema di sorting. Se le azioni di riferimento sono fittizie e devono essere elaborate, la definizione formale di criteri, coefficienti di importanza relativa e stati di qualificazione sui criteri dovrebbe essere supportata da specifiche *procedure di sviluppo*. Differenti approcci possono venire usati per formalizzare il modello di riferimento ed alcuni metodi sono stati sviluppati con queste finalità. Le condizioni di coerenza dell’insieme di riferimento permettono la definizione di categorie che sono sufficientemente distinte da impedire una attribuzione multipla di un candidato (cliente) a più di una classe.

Le classi possono essere ordinate e corrispondono a raccomandazioni come ‘ottimo’, ‘buono’, ‘discreto’, ‘sufficiente’, ‘scarso’, e così via. Spesso è necessario definire delle categorie aggiuntive corrispondenti all’esitazione dovuta alla difficoltà di scelta fra classi attigue, questo a causa dell’imprecisione dei dati e all’incertezza tra i criteri. In altri casi ci sono soltanto due classi (‘cliente buono’ e ‘cliente non buono’) o tre con quella degli ‘incerti’, a cui vengono assegnate azioni candidate che non possono venire classificate né come ‘accettabili’, né come ‘rifiutabili’. Se le azioni candidate sono tra di loro nettamente differenti (succede frequentemente se le caratteristiche del possibile candidato non sono definite in maniera restrittiva) e devono essere o accettate o rifiutate, le classi di attribuzione possono essere solo due o tre, ma il modello di riferimento deve essere sufficientemente articolato da permettere l’accettazione od il rifiuto del candidato, qualunque sia la sua tipologia. In questo caso il modello deve includere tante azioni di riferimento quante sono le differenti condizioni (o tipologie o modelli) di accettazione e di rifiuto.

Sia  $A = \{ x, y, z \dots \}$  l’insieme delle azioni con cardinalità  $|A|=m$ , descritte mediante un vettore di  $n$  criteri  $\mathbf{g} = \{g_1, g_2, \dots, g_n\}$ . Il problema della classificazione (*sorting*)

consiste nell' assegnare le  $m$  alternative in  $t$  predefinite classi ordinate  $\mathbf{CI} = \{CI_1, CI_2, \dots, CI_t\}$ .

L' assegnazione di un' alternativa  $a_i \in \mathbf{A}$  ad una specifica classe dipende dal confronto della valutazione della stessa sulla base di tutti i criteri con alcune azioni di riferimento che separano le classi decisionali contigue.

Definiamo le classi unioni *upward* e *downward* nel seguente modo:

$$CI_t^{\geq} = \bigcup_{s \geq t} CI_s \quad CI_t^{\leq} = \bigcup_{s \leq t} CI_s.$$

In un problema di classificazione multicriteriale vengono presi in considerazione due principali aspetti:

- 1) la forma del modello di aggregazione dei criteri;
- 2) La metodologia utilizzata per definire i parametri del modello.

Con riferimento alla prima delle principali problematiche del sorting, in letteratura si considerano due principali modelli di aggregazione basati rispettivamente, sulla relazione di surclassamento e sulla funzione di utilità. (Angilella et al., 2002; Costanzo et al., 2001; Costanzo et al., 2000).

La funzione di *utilità* assegna un valore reale  $f(x)$  ad ogni alternativa  $x \in \mathbf{A}$  ed attribuisce un' alternativa  $x$  alla classe  $CI_s^{\geq}$  se  $f(x) \geq \epsilon_s$ , dove  $\epsilon_s$ ,  $s=2, \dots, t$ , sono  $t-1$  soglie ordinate che soddisfano la seguente condizione:

$$\epsilon_2 \leq \epsilon_3 \leq \dots \leq \epsilon_t.$$

Una relazione binaria di surclassamento  $S$  è definita su  $\mathbf{A}$  nel seguente modo:

$xSy$  significa "  $x$  è complessivamente almeno tanto buona quanto  $y$ ".

Una relazione di surclassamento  $S$  su  $\mathbf{A}$ , assegna l' alternativa  $x$  alla classe unione *upward*  $CI_s^{\geq}$  se  $xSp^s$ , dove  $p^s$ ,  $s=2, \dots, t$ , sono  $t-1$  profili di riferimento, tali che  $p^{s+1}$  domina  $p^s$  (cioè  $p^{s+1}$  è almeno tanto buono quanto  $p^s$  con riferimento a ciascun criterio e  $c'$  è almeno un criterio per il quale  $p^{s+1}$  è strettamente preferito a  $p^s$ ),  $s=2, \dots, t-1$ .

Nel presente paragrafo prendiamo in considerazione l'utilizzo di aggregatori fuzzy. In particolare consideriamo gli integrali bipolari di Choquet e Sugeno, quali estensioni di quelli classici, per problematiche di classificazione.

**Definizione capacità bipolare.**

Dato  $N=\{1,\dots,n\}$ ,  $P(N)=\{(R,S): R \subseteq N, S \subseteq N \text{ con } R \cap S = \emptyset\}$ , si definisce capacità bipolare su  $N$  una funzione  $\mu: P(N) \rightarrow [0,1] \times [0,1]$ , tale che

- 1) Per ogni  $(R,S),(T,U) \in P(N)$  tale che  $R \supseteq T$  e  $S \subseteq U$ , con  $\mu(R,S)=(r,s)$  e  $\mu(T,U)=(t,u)$  con  $r \geq t$  e  $s \leq u$ ,
- 2)  $\mu(A,\emptyset)=(r,0)$  e  $\mu(\emptyset,B)=(0,s)$ ,  $r,s \in [0,1]$ ,
- 3)  $\mu(N,\emptyset)=(1,0)$  e  $\mu(\emptyset,N)=(0,1)$ .

**Integrali bipolari di Choquet e di Sugeno .**

Per ogni  $x \in \mathbb{R}$  si definisce  $x^+ = \max(x,0)$  la parte positiva e con  $x^- = \max(-x,0)$  la parte negativa (Greco et al.,2002; Greco et al., 2002b; Greco et al.2002; Sipos, 1979).

Per ogni  $\mathbf{x}=(x_1,\dots,x_n) \in \mathbf{R}^n$  denotiamo la parte positiva di  $\mathbf{x}$  con  $\mathbf{x}^+=(x_1^+, \dots, x_n^+)$  e con  $\mathbf{x}^-= (x_1^-, \dots, x_n^-)$  la parte negativa di  $\mathbf{x}$ .

Per ogni vettore  $\mathbf{x}=(x_1,\dots,x_n) \in \mathbf{R}^n$ , si consideri la funzione  $\pi: N \rightarrow N$  permutazione degli elementi di  $N$  tale che  $|x_{\pi(1)}| \leq \dots \leq |x_{\pi(n)}|$ . Per ogni  $i \in N$ , definiamo i seguenti insiemi:

- $A_{\pi(i)}^+ = \{j \in N: \pi(j) \geq \pi(i) \text{ e } x_{\pi(j)} \geq |x_{\pi(i)}|\}$  e
- $A_{\pi(i)}^- = \{j \in N: \pi(j) \geq \pi(i) \text{ e } -x_{\pi(j)} \geq |x_{\pi(i)}|\}$ .

L' integrale bipolare di Choquet di  $\mathbf{x}=(x_1,\dots,x_n) \in \mathbf{R}^n$  con riferimento alla capacità  $\mu(\cdot,\cdot)$  è definito nel seguente modo:

$$C^b(\mathbf{x},\mu) = C^+(\mathbf{x},\mu) - C^-(\mathbf{x},\mu).$$

dove  $C^+(\mathbf{x}, \mu) = \sum_{i \in N} (|x_{\pi(i)}| - |x_{\pi(i-1)}|) \mu^+(A_{\pi(i)}^+, A_{\pi(i)}^-)$  è l' integrale relativo alla parte

positiva e  $C^-(\mathbf{x}, \mu) = \sum_{i \in N} (|x_{\pi(i)}| - |x_{\pi(i-1)}|) \mu^-(A_{\pi(i)}^+, A_{\pi(i)}^-)$  è quello relativo alla

parte negativa, con  $x_{\pi(0)} = 0$ .

Mentre l' integrale bipolare di Sugeno  $\text{dia} = (a_1, \dots, a_n) \in [-1, 1]^n$  con riferimento alla capacità  $\mu(\cdot, \cdot)$  è definito nel seguente modo:

$$S^b(\mathbf{a}, \mu) = \begin{cases} -S^-(\mathbf{a}, \mu) & \text{if } S^+(\mathbf{a}, \mu) < S^-(\mathbf{a}, \mu) \\ 0 & \text{if } S^+(\mathbf{a}, \mu) = S^-(\mathbf{a}, \mu) \\ S^+(\mathbf{a}, \mu) & \text{if } S^+(\mathbf{a}, \mu) > S^-(\mathbf{a}, \mu) \end{cases}$$

dove  $S^+(\mathbf{a}, \mu) = \max_{i \in N} (\min(a_{\pi(i)}^+, \mu^+(A_{\pi(i)}^+, A_{\pi(i)}^-)))$  è l' integrale relativo alla parte

positiva e  $S^-(\mathbf{a}, \mu) = \max_{i \in N} (\min(a_{\pi(i)}^-, \mu^-(A_{\pi(i)}^+, A_{\pi(i)}^-)))$  è quello relativo alla parte

negativa.

### ***Rappresentazioni di problemi di sorting con integrali bipolari***

I problemi di sorting con funzione d' utilità può essere formulato, secondo l' approccio degli integrali bipolari, (Angilella et al., 2002) nel seguente modo:

$$C^b(\mathbf{x}, \mu) < \varepsilon_2 \quad \Rightarrow \mathbf{x} \in Cl_1$$

$$C^b(\mathbf{x}, \mu) < \varepsilon_3 \quad \Rightarrow \mathbf{x} \in Cl_2$$

....

$$C^b(\mathbf{x}, \mu) \geq \varepsilon_t \quad \Rightarrow \mathbf{x} \in Cl_t$$

La classificazione può essere fatta anche mediante una relazione di surclassamento  $S$  rispetto ad un' azione di riferimento  $a_r$ :

$xSa_r \Leftrightarrow u(x) \geq u(a_r)$  dove  $u: \mathfrak{R}^n \rightarrow \mathfrak{R}$  è una funzione d' utilità:

$$x \in Cl_r^{\geq} \Leftrightarrow S(x, a_r) \geq 0.$$

Allora,

$$xSy \Leftrightarrow C^b(g_1(x_1), \dots, g_n(x_n); \mu) \geq C^b(g_1(y_1), \dots, g_n(y_n); \mu),$$

dove  $C^b(x, \mu) = C^+(x, \mu) - C^-(x, \mu)$  è l'integrale bipolare di Choquet.

Con l' integrale bipolare di Sugeno e' possibile generalizzare la procedura di sorting anche con riferimento ad azioni con valutazioni ordinali.

In un problema di sorting, sia con un metodo di surclassamento che con una funzione d' utilità gli integrali bipolari di Choquet e di Sugeno per loro stessa definizione richiedono la normalizzazione delle valutazioni relative ai singoli criteri. In questo caso è possibile utilizzare un approccio empirico, dove si stimano le funzioni d' utilità mediante un algoritmo computazionale basato su alcune azioni di riferimento e sulle informazioni date dal decisore su dei sottoinsiemi di criteri.

Il vantaggio degli integrali bipolari di Sugeno e di Choquet consiste esplicitamente nel prendere in considerazione i valori maggiori e minori rispetto ad un livello neutrale di riferimento per ciascun criterio. Più precisamente, l' estensione degli integrali di Sugeno e di Choquet proposta in definisce il peso attribuito ad un dato insieme di valutazioni anche in funzione dell' insieme delle valutazioni simmetriche.

Infatti, nella recente letteratura (Costanzo et al.,2001; Angilella et al.,2002; Grabisch, Labreuche et Vansnick,2002; Grabisch, Greco et al.,2002; Grabisch e Labreuche,2000) viene sottolineato come sia interessante da un punto di vista decisionale, oltre a considerare l' aspetto classico di confronto tra singole alternative, anche l' esistenza di un livello neutrale per ogni criterio, rispetto al quale poter classificare un' azione come *attraattiva* o *repulsiva*.

## CAPITOLO 11 -

### **L'UTILIZZO DI FUNZIONI DI UTILITA' NON ADDITIVE - L'APPLICAZIONE**

#### 11.1 METODOLOGIE ED ATTIVITA' DI REALIZZAZIONE DEL SONDAGGIO

La fase più delicata di un'analisi di C.S. è sicuramente quella della scelta degli attributi o fattori e dei relativi livelli, in quanto la validità dei risultati ottenuti ne è condizionata fortemente; scelte sbagliate rischiano di inficiare le conclusioni ottenute (Hayes, 1992).

Domande tipiche di questa fase sono: quali sono i molteplici aspetti su cui i clienti basano la valutazione del servizio? Questi elementi variano a seconda dei diversi segmenti di clientela? Ammesso che le aspettative dei clienti svolgono un ruolo essenziale nella valutazione, quali fattori la determinano e influenzano?

Altro fattore da non trascurare in questa fase è la percezione che del servizio, e quindi dei suoi attributi, ha il consumatore. E' ormai universalmente noto che la qualità del servizio, così come è intesa dai consumatori, si può definire come il grado di discrepanze tra le aspettative o i desideri dei clienti e loro percezioni.

La nostra applicazione prevede l'esame della Customer Satisfaction di una Banca operante nella provincia di Catania con numerose agenzie.

Per la realizzazione della fase preliminare dell'indagine sono stati presi in considerazione dei *focus group*, composti da un numero minimo di membri, in possesso delle caratteristiche di età, status sociali e ruoli socio-professionali abbastanza omogenei e rappresentativi del target di clienti di una Banca.

Il *focus group* consiste nel riunire un numero limitato di persone che rappresentino, secondo certi criteri, la struttura della popolazione oggetto d'indagine. L'incontro viene gestito da un moderatore che stimola la discussione tra i partecipanti, coinvolgendoli, se necessario, su argomenti utili alla stesura del questionario della successiva fase quantitativa. Deve contenere, inoltre, gli eventuali conflitti. Un osservatore, presente all'incontro, annota gli interventi e gli elementi interessanti.

L'informazione ottenuta mediante queste interviste guidate ha permesso di individuare dei "macro-attributi".

Tutti gli indicatori sono dei "contenitori definiti dai testimoni in modo piuttosto omogeneo con qualche differenza semantica che comunque non distorce il significato che il senso comune attribuisce.

Il primo attributo attiene alla qualità *ambientale* - risponde alla domanda "dove" il cittadino riceve il servizio e si riferisce agli aspetti tangibili che consentono l'ottimale fruizione dei servizi, sia agli aspetti intangibili o di soddisfazione psicologica. Sono riconducibili a questa dimensione le strutture immobiliari e mobiliari nonché, a livello psicologico, l'impatto dell'ambiente sulle sensazioni e sulle emozioni.

Il secondo attributo è di qualità *tecnica* e risponde alla domanda "cosa fornisce" l'ufficio ai fruitori del servizio. Si misura in termini di tempi di risposta, alternative fornite, impatto...

Infine il terzo attributo riguarda la qualità *relazionale* e risponde alla domanda "come fornisce" il servizio l'ufficio e si riferisce agli aspetti comunicazionali, relazionali e di competenza. Sono riconducibili a questa dimensione la cortesia, la capacità tecnica nel risolvere i problemi, ...

Ciascuna delle dimensioni ha ricevuto, in termini di variabili poste, un'attenzione identica.

Sulla base della definizione di "grado di soddisfazione" come *scostamento* tra prestazione percepita e prestazione attesa, è stata intrapresa una fase di ricerca quantitativa volta a realizzare un sistema di misurazione delle attese e delle percezioni. Nell'elaborare una strategia di misurazione dello scostamento si è fatto riferimento al modello di Parasuraman. Il sistema prevede che lo strumento utilizzato per raccogliere informazioni (nel nostro caso il questionario strutturato) venga suddiviso in due sezioni: una dedicata alle aspettative, formata da un certo numero di proposizioni, e l'altra riservata alle percezioni, con un numero di proposizioni corrispondenti. Le proposizioni di ciascuna sezione prevedono una risposta obbligatoria espressa lungo una scala di valutazione numerica.

Nel nostro caso alla proposizione affermativa (verso cui esprimere il grado di accordo) è stata preferita la domanda. Si è scelta, come scala di valutazione, quella da cinque a uno: il valore cinque esprime la risposta più favorevole, il valore uno quella meno favorevole.

La differenza, per ciascuno degli aspetti indagati, tra il punteggio sulla percezione e quello (corrispondente) sull'attesa esprime il livello di soddisfazione dell'intervistato verso la prestazione.

E' bene, altresì, precisare che la natura del fenomeno indagato e la considerazione di una sola unità di osservazione (una sola banca) non consente di generalizzare i risultati ottenuti all'intero universo dei frequentatori di una banca, se non limitatamente ai clienti della particolare banca oggetto di studio.

Ciascuna sezione del questionario è formata da 12 domande; tre indagano la qualità ambientale, tre quella tecnica, tre quella relazionale, tre indagano gli aspetti complessivi.

Le domande del questionario sono introdotte dalla formula "quanto è soddisfatto per"; ed invitano ad esprimere un voto da 1 a 5.

*Per la realizzazione del questionario si sono tenute in debita considerazione le seguenti regole:*

- ♣ Esplicitare il problema di marketing
- ♣ Definire gli obiettivi informativi necessari a risolvere il problema di marketing
- ♣ Esplicitare attraverso un brainstorming tutte le domande a cui la ricerca deve rispondere
- ♣ Collegare ogni domanda ad ognuno degli obiettivi informativi
- ♣ Scartare quelle domande non direttamente associabili agli obiettivi informativi o ritenute poco significative
- ♣ Riorganizzare le domande per argomento
- ♣ Controllare l'ordine delle domande in modo che la risposta alla domanda precedente non influenzi le domande successive



- ♣ Controllare che non ci siano palesi ripetizioni, sovrapposizioni tra domande
- ♣ Identificare le domande di controllo (domande verbalizzate in modo diverso, ma con contenuto informativo identico utilizzate per verificare la veridicità della risposta)
- ♣ Attribuire ad ogni domanda un valore in termini di utilità dell'informazione di ritorno (in questo modo se l'intervistato mostrasse segni di noia, l'intervistatore si può focalizzare sulle domande più importanti)
- ♣ Prevedere il campo delle risposte possibili sia per le domande chiuse, che per quelle aperte
- ♣ Misurare il tempo di lettura e di compilazione delle risposte
- ♣ Definire ex ante una griglia di interpretazione delle risposte
- ♣ Testare il questionario su un campione del target in termini di
- ♣ Comprensibilità delle domande
- ♣ Durata della compilazione
- ♣ Predisporre la forma definitiva del questionario
- ♣ Predisporre in modo chiaro le istruzioni di compilazione o le istruzioni per l'intervistatore

I questionari sono riportati in allegato alla fine del capitolo.

E' bene precisare che diversamente da quanto accade in altre tecniche di valutazione di Customer Satisfaction, in cui non si tiene conto dell'interazione, nella nostra applicazione l'essenza del problema è proprio la valutazione dell'interazione dei fattori esaminati.

## 11.2 IL CAMPIONE

In generale un campione deve rispondere positivamente a due requisiti fondamentali: quello di rappresentatività (della realtà che si vuole indagare) e quello di significatività (in termini di numerosità e di attendibilità).

Se il modello non è corretto, infatti, il modellatore si trova con delle distorsioni che non è in grado di controllare e il descrittore dell'analisi, invece, stima il vettore dei parametri di popolazione, che se da una parte ha un qualche significato dall'altro perde di significato venendo meno il modello che ne giustifica l'interesse.

Relativamente al modello in esame ciò implica che i risultati ottenuti saranno validi unicamente in relazione alla situazione sperimentale considerata; qualora dovesse cambiare anche uno solo degli elementi che la caratterizza (attributi, livelli, tipologia dei clienti intervistati, ecc.) tali risultati non saranno più validi. Inoltre la natura individuale di questo tipo di analisi non permette di ottenere la distruzione di campionamento dei singoli parametri stimati, il che rende impossibile il calcolo di intervalli di confidenza e la verifica di ipotesi ad essi relative.

Questo approccio è sembrato il più opportuno, in quanto il campo di osservazione del presente lavoro è circoscritto agli utenti del servizio bancario di un particolare istituto, per cui è verosimile assimilare questi soggetti ad una sub-popolazione di dimensione finita, realizzazione casuale di una superpopolazione costituita da tutti i clienti di banche.

Nel predisporre il questionario si è preventivamente analizzato il tipo di popolazione da interessare al sondaggio, tenuto conto che nella Banca non era mai stata svolta un'indagine di Customer Satisfaction, né era stata realizzata un'approfondita segmentazione della clientela; si è quindi deciso di rivolgere l'analisi a tutta la clientela, per avere un quadro complessivo, sia pur generico, dell'immagine percepita della banca, (nei suoi aspetti tangibili, empatici, di tempestività nell'erogazione del servizio, e cura delle esigenze della clientela), dalla clientela, riservandosi nel caso in cui si intravedessero situazioni di insoddisfazione, di approfondire l'analisi nel segmento

individuato. Trattandosi dell'intera clientela della banca, ne deriva una forte eterogeneità interna del campione, riguardo sia gli aspetti relativi a reddito, età, sesso, titolo di studio, uso dei servizi, frequenza d'uso; nonché attese emozionali e tangibili.

Non è previsto un *numero massimo* di clienti da intervistare, considerata l'eterogeneità del campione, più ampia è la popolazione intervistata, maggiore sarà la rappresentatività del campione, e quindi più chiaro sarà il quadro d'analisi che ne risulta.

Per quanto riguarda, invece, un *numero minimo*, si ritiene indispensabile analizzare almeno 90 questionari.

Per quanto riguarda la frequenza prevista, non è stata programmata una periodicità d'analisi.

Gli obiettivi previsti dal sondaggio sono :quelli di avere una visione d'insieme del *grado di soddisfazione* generale della clientela, analizzandola dal punto di vista sintetico (complessivamente intesa, o suddivisa per Filiale), che analitico (per ciascun segmento di clientela del campione, es. i correntisti, i titolari di conto on line ...).

Tutto ciò, allo scopo di individuare le aree a maggiore redditività, che necessitano un intervento correttivo urgente.

Nello specifico il grado di soddisfazione per un servizio erogato può essere misurato solo su coloro che vi hanno avuto accesso e che, pertanto, possono esprimere un giudizio di percezione (oltre a quello di attesa). E' un aspetto di cui si è dovuto tener conto nella costruzione dei campioni.

Con riferimento alla popolazione da indagare si è deciso di riprodurre la popolazione secondo gli elementi di sesso e di età. L'altra condizione, di significatività, poneva il problema della consistenza numerica del campione. Per garantire conclusioni significative è necessario che l'errore massimo di campionamento sia contenuto al 10%. In questo modo viene garantito un livello di confidenza accettabile.

La fase successiva è stata stabilire le modalità di individuazione delle unità del campione, in modo da garantire una rilevazione corretta dei dati. La limitata disponibilità della risorse (economiche ed umane) ha favorito la scelta di formare il campione contattando la gente in azienda in modo casuale. Il questionario è stato

somministrato a quelle che presentavano i requisiti di conoscibilità e utilizzo dell'ufficio.

L'obiettivo delle 90 unità, stratificate secondo i criteri di sesso ed età introdotti, ha richiesto il contatto di almeno 600 persone.

La somministrazione del questionario è stata sempre preceduta da due domande *filtro*. L'una atta a testare se il cittadino conoscesse o meno la C.S.; l'altra a rilevare se l'avesse mai considerata. Solo nei casi in cui le risposte per entrambe le domande erano positive si è proceduto all'intervista.

Questo semplice criterio naturalmente presenta dei pregi e degli inconvenienti. Al primo aspetto si riferisce la possibilità di controllare la qualità delle risposte degli intervistati, maggiormente stimolati dal contatto diretto ad assumere la responsabilità delle proprie valutazioni. Al secondo appartiene la consapevolezza della difficoltà di completare le classi del campione.

### 11.3 IL METODO MONTECARLO

Nelle industrie, o in azienda, quando si deve riprodurre in scala ridotta la dinamica evolutiva di situazioni reali al fine di prendere decisioni operative o fare previsioni, se non è possibile rappresentare fisicamente i suddetti scenari si cerca di realizzarne un modello virtuale utilizzando quella che gli addetti ai lavori definiscono **tecnica di simulazione**. In altre parole, mentre un modello fisico rappresenta sia pure in scala ridotta la realtà, possiamo dire che la simulazione la imita. Così, se è possibile realizzare in scala il profilo di un'ala d'aereo, o il modello della carena di una nave, di cui testare rispettivamente l'efficacia in una galleria del vento o in una vasca idrodinamica, più difficile se non addirittura impossibile è modellizzare, per esempio, il rendimento di una linea di produzione, l'andamento di una campagna di vendita, o le richieste di un determinato servizio che si possono materializzare, per esempio, nell'afflusso di auto ad un casello, o di clienti ad uno sportello bancario. Anche l'utilizzo della simulazione implica la realizzazione di un modello, ma in questo caso la sua natura è squisitamente matematica. Potrebbe essere più complicato impostarlo concettualmente, ma sicuramente si rivela più economico di un modello fisico e, se è stato strutturato correttamente, la sua flessibilità è decisamente superiore. Non abbiamo certo la presunzione di riconoscere in queste tecniche la soluzione totale dei problemi che assillano i moderni manager, ma sta di fatto che tali strumenti, se interpretati dall'analista con creatività e buon senso, si rivelano preziosi quando è il momento di prendere decisioni che richiedono di gettare uno sguardo nel futuro più o meno prossimo senza affidarsi alla tradizionale sfera di cristallo. È il caso, tanto per fare un esempio che si riallaccia ai nostri preamboli, dell'interpretazione dell'evolversi di fenomeni legati alla casualità, che può essere simulata con una ottima verosimiglianza grazie alle speciali funzioni che operano nell'ambito di un foglio elettronico, a cominciare da Excel 2000. Le prime applicazioni delle tecniche di simulazione risalgono al secolo scorso, ed uno dei metodi che ha incontrato maggiore successo è quello cosiddetto di **Montecarlo**. Chi immagina una correlazione con la omonima e famosa casa da gioco ha visto giusto! Infatti, quale migliore laboratorio di un casinò per studiare la casualità? E proprio a Montecarlo, nella seconda metà del secolo scorso, registrando puntualmente le uscite dei numeri alla roulette, vennero creati i primi

elenchi di **numeri casuali** che furono, e sono tuttora, alla base delle moderne tecniche di simulazione. Sono le leggi della causalità, infatti, che regolano molti fenomeni della vita quotidiana, e non solo il comportamento della pallina protagonista del gioco delle “roulette”. Il metodo di Montecarlo, pertanto, si rivela particolarmente indicato per la simulazione di tutte le situazioni che sono regolate dal caso: il numero di vetture che transitano ai caselli di un’autostrada, quello degli ordini che pervengono ad una azienda, quello dei clienti che affluiscono ad una tavola calda, e così via. Infatti, per impostare una simulazione che descriva l’evolversi di una dei suddetti fenomeni si genera una serie di numeri casuali, e si associa ad ognuno dei suoi elementi l’evento che ha la medesima probabilità di essere generato.

Seguendo questi orientamenti e con il preciso scopo di valutare diverse ipotesi di scenario nella realizzazione della ricerca si è fatto uso del Metodo Montecarlo.

Presentiamo di seguito gli aspetti più significativi dell’analisi.

#### 11.4 L'APPLICAZIONE I – L'INTEGRALE DI CHOQUET E DI SUGENO

Come già detto nel precedente capitolo per l'applicazione ci si è posti il problema della individuazione degli attributi maggiormente discriminanti, cioè di quelli che influiscono sul processo di valutazione dell'utente, rispetto ai quali misurare gli scostamenti tra atteso e percepito.

L'ordinamento in funzione del gradimento permette di individuare quello che in letteratura è noto come sistema della qualità attesa, mentre quello che viene indicato definisce la qualità percepita dagli utenti cioè la distanza tra i due ci dà un'indicazione utile ai fini della valutazione della soddisfazione del cliente.

Per fare questo, utilizzando il metodo Montecarlo abbiamo determinato in metodo indiretto i parametri di una funzione di utilità espressa in termini di integrale di Choquet.

1. In particolare, si è proceduto alla simulazione di 300 vettori di misure associate ad ogni criterio rappresentanti l'importanza relativa e l'interazione, attraverso la generazione casuale. Le misure  $\mu$ , rispettano i vincoli del problema e cioè vincoli sulle importanze e sulle interazioni e vincoli sull'ordinamento delle valutazioni sui singoli criteri con riferimento alle utilità marginali.
2. Quindi si sono calcolate le utilità complessive dei vari clienti in termini di integrali di Choquet relativi alle misure  $\mu$  per ogni singolo criterio e alle utilità marginali considerate.
3. Si sono determinati gli scostamenti tra l'ordinamento fornito dal decisore (qualità percepita) e l'ordinamento risultante dalle utilità calcolate al punto 2) (qualità attesa), ovvero calcolo dei quadrati delle differenze tra percepito ed atteso (onde evitare determinazioni negative)
4. Si è poi, proceduto alla selezione delle misure e delle funzioni di utilità marginali in corrispondenza delle quali il numero di inversioni (scostamenti) calcolate al punto 3) sia minimo, ovvero calcolo dei minimi quadrati.

Dall'esame dei dati si evince con chiarezza che nessuno dei criteri di qualità (ambientale = aspetto esteriore ( $\mu(1)=0,21$ ), tecnica = servizio ( $\mu(2)=0,15$ ), relazionale = personale ( $\mu(3)=0,20$ )) è fortemente importante, mentre assume particolare importanza la loro combinazione a coppie, ovvero la interazione tra aspetto esteriore e servizio è fortemente interattiva ( $\mu(1,2)=0,80 > \mu(1) + \mu(2)$ ), così come quella tra aspetto esteriore e personale ( $\mu(2,3)=0,44 > \mu(3) + \mu(2)$ ).

Tale considerazione è verificata anche dall'esame di altri scenari che determinano scostamenti leggermente più elevati.

La simulazione ci permette di osservare che l'attributo personale è il meno importante rispetto agli altri. Infatti a valori di  $\mu(3)$  elevati servizio è più importante di personale, corrispondendo maggiori scostamenti tra atteso e percepito.

L'applicazione si è quindi spostata all'analisi condotta attraverso l'integrale di Sugeno.

L'elaborazione è stata condotta prendendo in considerazione quanto detto nel capitolo precedente sulle valutazioni qualitative.

Si è esaminato il numero delle inversioni tra atteso e percepito determinato utilizzando l'integrale di Sugeno sulle risposte di tipo qualitativo espresse dal questionario, relative al giudizio sui tempi di servizio, sugli orari e sulle informazioni ottenute.

1. In particolare, si è proceduto alla simulazione di 81 vettori rappresentanti i casi possibili di elaborazione a 3 livelli ( $3^3$ ) di misure associate ad ogni criterio rappresentanti l'importanza relativa e l'interazione, attraverso la generazione casuale. Le misure  $\mu$ , rispettano i vincoli del problema e cioè vincoli sulle importanza e sulle interazioni e vincoli sull'ordinamento delle valutazioni sui singoli criteri con riferimento alle utilità marginali.
2. Quindi si sono calcolate le utilità complessive dei vari clienti in termini di integrali di Sugeno relativi alle misure  $\mu$  per ogni singolo criterio e alle utilità marginali considerate.
3. Si sono determinati gli scostamenti tra l'ordinamento fornito dal decisore (qualità percepita) e l'ordinamento risultante dalle utilità calcolate al punto 2) (qualità attesa), ovvero calcolo delle inversioni tra i due livelli (percepito ed atteso)



4. Si e poi, proceduto alla selezione delle misure e delle funzioni di utilità marginali in corrispondenza delle quali il numero di inversioni (scostamenti) calcolate al punto 3) sia minore.

Si presenta il risultato il quale evidenzia il miglior livello e ciò si ha quando si presentano le misure indicate in tabella:

$\mu(1)$	$\mu(2)$	$\mu(3)$	$\mu(1,2)$	$\mu(1,3)$	$\mu(2,3)$	INVERSIONI
2,00	<b>2,00</b>	1,00	2,00	2,00	2,00	15
1,00	<b>2,00</b>	1,00	2,00	1,00	2,00	15
2,00	<b>2,00</b>	2,00	2,00	2,00	2,00	15
1,00	<b>2,00</b>	2,00	2,00	2,00	2,00	15

Con

$\mu(1)$	<b>tempi servizio</b>
$\mu(2)$	<b>Orario</b>
$\mu(3)$	<b>informazioni</b>
$\mu(1,2)$	<b>servizio+orario</b>
$\mu(1,3)$	<b>servizio+informazioni</b>
$\mu(2,3)$	<b>Orario+informazioni</b>

$\mu=1$	Insoddisfatto
$\mu=2$	soddisfatto
$\mu=3$	Pienamente soddisfatto

Dalla tabella si evince che se l'orario  $\mu(2)$  è almeno soddisfacente allora il livello qualitativo globale è almeno soddisfacente. Si veda in particolare anche l'interazione con gli altri attributi. Mentre i tempi di servizio e le informazioni considerate da sole possono anche essere non soddisfacenti e garantire un livello qualitativo globale soddisfacente.

### 11.5 L'APPLICAZIONE II – L'INTEGRALE GERARCHICO

Per la nostra ricerca è sicuramente importante la conoscenza della percezione della soddisfazione di ogni singolo attributo.

Pertanto, si è proceduto all'applicazione con analisi multicriteriale per la valutazione di un insieme delle funzioni marginali di soddisfazione rappresentanti il livello di ogni criterio. E' possibile determinare degli indici mostrano livello di soddisfazione parziale dei clienti secondo ogni sub-criterio, similmente all' indice globale di soddisfazione.

In modo simile a quanto indicato nel paragrafo precedente ed utilizzando il modello di Murofushi, Sugeno, Fujimoto, sulla decomposizione di integrali gerarchici si è proceduto alla determinazione in metodo indiretto dei parametri di una funzione di utilità espressa in termini di integrale di Choquet. Anche qui si è proceduto elaborando i dati utilizzando il metodo Montecarlo.

1. In particolare, si è proceduto alla simulazione di 300 vettori di misure associate ad ogni sub-criterio rappresentanti l'importanza relativa e l'interazione, attraverso la generazione casuale. Le misure  $\mu$ , rispettano i vincoli del problema e cioè vincoli sulle importanza e sulle interazioni e vincoli sull' ordinamento delle valutazioni sui singoli criteri con riferimento alle utilità marginali.
2. Quindi si sono calcolate le utilità parziali dei vari clienti in termini di integrali parziali di Choquet relativi alle misure  $\mu$  per ogni singolo sub-criterio e alle utilità marginali considerate.
3. Poi si è proceduto al calcolo delle utilità complessive dei vari clienti in termini di integrali di Choquet relativi alle misure  $\mu$  per ogni singolo macro-criterio e alle utilità marginali considerate, secondo il metodo di aggregazione proposto da Murofushi, Sugeno, Fujimoto. Si è esclusa la possibilità di sub-interazione tra i sub-criteri.
4. Si sono determinati gli scostamenti tra l' ordinamento fornito dal decisore (qualità percepita) e l' ordinamento risultante dalle utilità calcolate al punto 3)

(qualità attesa), ovvero calcolo dei quadrati delle differenze tra percepito ed atteso (onde evitare determinazioni negative)

5. Si è poi, proceduto alla selezione delle misure e delle funzioni di utilità marginali in corrispondenza delle quali il numero di inversioni (scostamenti) calcolate al punto 4) sia minimo, ovvero calcolo dei minimi quadrati.

Il risultato che ha prodotto il minore scostamento è quello presentato in tabella:

<b>aspetto esteriore</b>					
$\mu(1)$	$\mu(2)$	$\mu(3)$	$\mu(1,2)$	$\mu(1,3)$	$\mu(2,3)$
0,14	0,09	0,04	0,31	0,73	0,34
<b>servizio</b>					
$\mu(1)$	$\mu(2)$	$\mu(3)$	$\mu(1,2)$	$\mu(1,3)$	$\mu(2,3)$
0,49	0,27	0,35	0,72	0,68	0,61
<b>personale</b>					
$\mu(1)$	$\mu(2)$	$\mu(3)$	$\mu(1,2)$	$\mu(1,3)$	$\mu(2,3)$
0,33	0,23	0,33	0,94	0,84	0,88
<b>globali</b>					
$\mu(1)$	$\mu(2)$	$\mu(3)$	$\mu(1,2)$	$\mu(1,3)$	$\mu(2,3)$
0,20	0,16	0,02	0,45	0,48	0,61

Con misure associate relative agli attributi corrispondenti alle seguenti:

<i>Valutazione complessiva sull' aspetto esteriore</i>	
Arredamento elegante ed accogliente	$\mu(1)$
Impiegati curati nell' aspetto e nel vestire	$\mu(2)$
I materiali associati al servizio sono bene esposti?	$\mu(3)$
<b>Valutazione complessiva servizio</b>	
Interesse a risolvere i problemi dei clienti	$\mu(1)$
servizio erogato in breve tempo	$\mu(2)$
Orari d' apertura comodi	$\mu(3)$
<b>Valutazione complessiva personale</b>	
Gli impiegati con il loro comportamento Le infondono fiducia?	$\mu(1)$
Impiegati gentili	$\mu(2)$
Preparazione tecnica impiegati	$\mu(3)$

I valori globali confermano quanto precisato, in merito all'analisi, nel paragrafo precedente, e cioè che l'attributo "personale" è il meno importante rispetto agli altri, ma che esiste forte interazione tra tutte le sue sub-componenti.

Che in merito ai servizi risulta essere più importante "l'interesse a risolvere i problemi"  $\mu(1)=0,49$  ed, inoltre, che gli "operatori di apertura comodi" ( $\mu(3)$ ) rendono sub-additive le misure con le altre componenti  $\mu(1,3)<\mu(1)+\mu(3)$  e  $\mu(2,3)<\mu(2)+\mu(3)$

Che in merito all'aspetto esteriore l'arredamento accogliente deve essere accompagnato da materiali associati al servizio ben esposti.  $\mu(1,3)=0,73$ .

### 11.6 L'APPLICAZIONE III - IL CUSTOMER PROFILING COME PROBLEMA DI CLASSIFICAZIONE

Il terzo aspetto applicativo che si è evidenziato nel capitolo precedente è quello dell'utilizzo delle misure non additive nella segmentazione dei clienti come un caso tipico di classificazione (sorting). L'ulteriore applicazione si basa sull'indagine su campione, già descritta, dei clienti dell'istituto finanziario ai quali è stato chiesto, oltre che, un giudizio complessivo sull'istituto ed una valutazione dell'importanza da loro attribuita ad alcuni elementi che caratterizzano il processo di scelta.

Più in dettaglio si voleva mettere in relazione l'interesse verso alcune caratteristiche del servizio ed il giudizio complessivo sull'istituto con la propensione a consigliarlo a parenti ed amici. Il problema, quindi, può porsi in termini di classificazione di clienti fra coloro che sono propensi a consigliare (PC) e coloro che non sono propensi a consigliare (NPC).

Nell'applicazione abbiamo utilizzato gli integrali bipolari di Sugeno e di Choquet. Il vantaggio della formulazione bipolare consiste esplicitamente nel prendere in considerazione le valutazioni positive e negative simmetriche rispetto ad un livello neutrale di riferimento per ciascun criterio. Più precisamente, l'estensione degli integrali di Sugeno e di Choquet utilizzata definisce il peso attribuito ad un dato insieme di valutazioni anche in funzione dell'insieme delle valutazioni simmetriche.

Nel nostro caso, infatti, è interessante, da un punto di vista decisionale, oltre che considerare l'aspetto classico di confronto tra singole alternative, anche l'esistenza di un livello neutrale per ogni criterio, rispetto al quale poter classificare un'azione come *attrattiva* o *repulsiva*.

L'analisi condotta ci permette di interpretare il costrutto anche nell'ambito della *customer loyalty* (fedeltà del cliente) (Costabile, 2001; Dick e Basu, 1994) perché il consigliare ad altri può essere inteso come possibilità di riacquisto del prodotto.

Si è proceduto come segue:

1. In particolare, si è normalizzata la matrice dei valori relativi ai “vari attributi (aspetto esteriore, servizio, personale) + valutazione globale” calcolando le differenze rispetto al valore neutrale (soddisfatto) pari a 3.
2. Si è proceduto all'analisi degli altri clienti calcolando l'integrale bipolare di Choquet con le misure associate, positive e negative, ad ogni sottoinsieme di criteri.
3. Quindi si sono calcolate le utilità complessive dei vari clienti in termini di integrali di Choquet relativi alla misura bipolare  $\mu$ , tenendo conto della diversa importanza tra i criteri e in particolare considerando il servizio più importante del personale e questo più importante dell' aspetto esteriore. Le misure sono state predeterminate, ma è possibile calcolare le misure utilizzando le tecniche già proposte nei precedenti paragrafi.
4. Si è determinato se il valore così trovato era superiore o inferiore al cliente ("neutro") che dà tutte le valutazioni uguali all' elemento "neutro".
5. Si è effettuata la classificazione tra clienti propensi a consigliare (PC) e clienti non sono propensi a consigliare (NPC).

A titolo di esempio si riportano i calcoli relativi alle valutazioni di alcuni clienti:

Siano **A = aspetto esteriore**, **B = servizio**, **C = personale**, gli insiemi dei tre criteri.

Si è proceduto ai seguenti calcoli:

$$\text{Cliente 71} = 1\mu^+(\{C\}, \{A\}) - [1\mu^-(\{C\}, \{A\}) + 1\mu^-(\{\emptyset\}, \{A\})]$$

$$\text{Cliente 1} = -[1\mu^-(\{\emptyset\}, \{A, B, C\}) + 1\mu^-(\{\emptyset\}, \{A, B\})]$$

$$\text{Cliente 26 e 43} = -1\mu^-(\{\{\emptyset\}, \{A, B\}\})$$

$$\text{Cliente 28 e 55} = 1\mu^+(\{B, C\}, \{A\}) - 1\mu^-(\{B, C\}, \{A\})$$

$$\text{Cliente 57} = [1\mu^+(\{B, C\}, \{A\}) + 1\mu^+(\{C\}, \{\emptyset\})] - 1\mu^-(\{B, C\}, \{A\})$$

$$\text{Cliente 84} = +1\mu^+(\{C\}, \{A\}) - 1\mu^-(\{C\}, \{A\})$$

$$\text{Cliente 2} = -1\mu^-(\{\emptyset\}, \{B\})$$

$$\text{Cliente } 29 = 1\mu^+(\{A\}, \{C\}) - 1\mu^-(\{A\}, \{C\})$$

$$\text{Cliente } 80 = 1\mu^+(\{A, C\}, \{\emptyset\}) + 1\mu^-(\{A\}, \{\emptyset\})$$

Con  $\mu^+$

parte positiva	parte negativa	valore
C	A	0,6
B,C	A	0,8
A	C	0,2
A,C	$\emptyset$	0,7
A	$\emptyset$	0,1

$\mu^-$

parte positiva	parte negativa	valore
C	A	0,2
$\emptyset$	A,B,C	0,9
$\emptyset$	A	0,4
$\emptyset$	A,B	0,7
B,C	A	0,2
$\emptyset$	B	0,6
A	C	0,4

rappresentanti i valori delle misure positive e negative, dei sottoinsiemi bipolari (parte positiva e parte negativa).

Il risultato è quello presentato nella seguente tabella:

clienti	aspetto esteriore	servizio	personale	globale	A = aspetto esteriore	B = servizio	C = personale	D = globale	valutazione choquet bipolare	classificazione
71	1	3	4	3	-2	0	1	0	0,0	NPC
1	2	1	1	1	-1	-2	-2	-2	-1,6	NPC
26 e 43	2	2	3	2	-1	-1	0	-1	-0,8	NPC
28 e 55	2	4	4	4	-1	1	1	1	0,6	PC
57	2	4	5	4	-1	1	2	1	1,0	PC
84	2	3	4	3	-1	0	1	0	0,4	PC
2	3	2	3	2	0	-1	0	-1	-0,6	NPC
29	4	3	2	3	1	0	-1	0	-0,2	NPC
80	5	3	4	4	2	0	1	1	0,8	PC

Da tale prospetto si evince come la maggiore attenzione ai criteri personale e servizio determina risultati diversi dalla valutazione globale per i clienti 71,84,29, in termini di diversa classificazione.

## 11.7 GLI ALGORITMI GENETICI

Solo al fine di completezza espositiva si precisa che il problema da noi trattato in termini applicativi, anziché con l'utilizzo del Metodo Montecarlo, può essere implementato con l'utilizzo di algoritmi genetici.

Gli algoritmi genetici (GA), sono algoritmi basati sui processi genetici degli organismi biologici, e possono essere usati per risolvere problemi di ricerca e ottimizzazione.

I GA riprendono la teoria dell'evoluzione e simulano su computer il fenomeno della selezione naturale, applicandoli alle soluzioni di un problema anziché ad esseri viventi.

L'idea di fondo è la seguente. Supponiamo di avere un certo insieme di soluzioni di un problema di ottimizzazione. Tra queste, ce ne saranno di più buone o meno buone. La "bontà" di una soluzione è misurata da una funzione di merito, detta *fitness function*, che in genere coinciderà con la funzione obiettivo, ma non sempre. A questo punto, vogliamo generare nuove soluzioni, con la speranza ovviamente che fra queste ve ne siano di sempre migliori. L'idea è allora quella di ricavare nuove soluzioni facendo accoppiare le soluzioni tra di loro, oppure modificandone alcune di loro. Allora, da una certa popolazione di individui, se ne ricava un'altra, che costituisce una nuova generazione, ossia quella dei figli della popolazione di partenza. Questa nuova generazione potrebbe essere molto più numerosa della precedente, e allora quello che si fa è effettuare una selezione, cioè si escludono dalla popolazione tutte le soluzioni che hanno un valore di fitness function inferiore a una certa soglia. Con la popolazione così selezionata, si ricomincia, generando quindi i nipoti delle soluzioni di partenza e così via per numero fissato di generazioni.

I GA rientrano nell'insieme delle meta-euristiche.

Un algoritmo è di tipo *euristico* quando è concepito in modo da produrre soluzioni che si sperano buone, ma senza garanzia a priori sulla vicinanza all'ottimo.

Una meta-euristica è un approccio di tipo generale, la struttura e l'idea di fondo di ciascuna meta-euristica sono sostanzialmente fissate, ma la realizzazione delle varie componenti dell'algoritmo dipende dai singoli problemi. Gli approcci metaeuristici



possono vedersi in realtà in modo omogeneo, come generalizzazioni di un unico approccio fondamentale, che è quello della *ricerca locale*.

I GA al pari delle altre meta-euristiche, per la loro efficacia è cruciale il modo in cui vengono effettuate diverse scelte parametriche.

L'impiego di un algoritmo genetico richiede la definizione di:

- Codifica o rappresentazione di una soluzione;
- Funzione fitness;
- Definizione dei processi riproduttivi (mutazione, crossover, ...);
- Dimensione della popolazione.

Negli algoritmi genetici gioca un ruolo fondamentale il modo in cui vengono rappresentate le soluzioni ammissibili. Tipicamente, una soluzione sarà rappresentata da un insieme di stringhe di interi o binarie, che sono i cromosomi. Ciascun cromosoma a sua volta è composto da geni.

Esempi di casi particolari:

- In alcuni problemi di ordinamento, (come il problema del commesso viaggiatore), è opportuno usare la codifica con permutazione;
- Una codifica col valore diretto può essere usata in problemi dove si usano valori complicati come i numeri reali. L'uso di una codifica binaria (De Luca 2000a, 2000b, 2001) per questo tipo di problemi sarebbe molto difficile. In questa codifica, ogni cromosoma è una stringa di alcuni valori. I valori possono essere qualsiasi cosa correlata al problema: da interi, numeri reali o caratteri fino ad alcuni oggetti complicati.

Per ciascun problema da risolvere deve essere costruita una specifica funzione fitness. Dato un particolare cromosoma, la funzione fitness restituisce un singolo valore numerico "fitness" o una "figura di merito", che si suppone sia proporzionale alla utilità o abilità dell'individuo che il cromosoma rappresenta. Per molti problemi, in particolari funzioni di ottimizzazione, è ovvio che la funzione fitness deve misurare il valore stesso della funzione. Ma non è sempre questo il caso.

Durante la fase di riproduzione di un GA, gli individui sono selezionati tra la popolazione e ricombinati, producendo la discendenza che sarà compresa nella generazione successiva. I genitori sono selezionati a caso usando uno schema che favorisce gli individui migliori. Gli individui buoni saranno probabilmente selezionati più volte per la riproduzione, mentre quelli peggiori potrebbero non essere mai scelti. Avendo selezionato due individui, i loro cromosomi sono ricombinati, tipicamente usando il meccanismo del *crossover* e la *mutazione*.

I geni nei cromosomi del nuovo individuo dovranno provenire da quelli dei due genitori.

In genere può avvenire in modo casuale, ossia se indichiamo con  $x_i$  e  $y_i$  il gene  $i$ -esimo di un certo cromosoma dei due genitori rispettivamente, il corrispondente gene nel cromosoma-figlio sarà uguale all'uno o all'altro con probabilità  $p$  e  $1-p$  rispettivamente.

Oppure si tagliano in  $n$  parti le stringhe dei genitori e si ricombinano per ottenere varie combinazioni di figli (Tipicamente  $n$  vale 2 o 3).

In ogni caso, il crossover dipende soprattutto dal tipo di codifica adottata, e dal problema.

La mutazione è applicata ad alcuni individui singolarmente dopo il crossover. Viene alterato a caso ogni gene con una probabilità bassa. La mutazione aiuta ad assicurarci che nessun punto nello spazio abbia probabilità nulla di essere esaminato.

Un'altra scelta molto importante è quella relativa alla dimensione della popolazione: a ogni generazione, infatti, verranno escluse dall'evoluzione dell'algoritmo tutte le soluzioni con basso valore di funzione obiettivo. È una scelta comunque quella di mantenere costante nei vari passi il numero di individui della popolazione.

A titolo di esemplificazione, si consideri il problema NP-completo che va sotto il nome di multiprocessor scheduling. Questo problema consiste nell'assegnare  $n$  oggetti di peso  $w_1, w_2, \dots, w_n$  a  $m$  contenitori, con l'obiettivo di minimizzare il peso del contenitore più carico. In questo caso possiamo scegliere di rappresentare la generica soluzione (assegnamento di oggetti a contenitori) per mezzo di una stringa (un unico cromosoma) di  $n$  interi in  $\{1, \dots, m\}$ , che indichiamo il contenitore cui è assegnato ciascun oggetto.

Così ad esempio si consideri un' istanza costituita da 8 oggetti, di peso 2, 3, 4, 6, 7, 9, 10, 14. Consideriamo, tra le altre, le due soluzioni ammissibili:  $x = (1\ 3\ 3\ 2\ 3\ 3\ 2\ 1)$   $y = (3\ 2\ 1\ 2\ 1\ 2\ 3)$  di valore 23 e 22 di funzione obiettivo rispettivamente. Possiamo dunque combinarle per ottenere nuove soluzioni, ciascun gene delle quali coinciderà con quello di uno dei due genitori, e in particolare, sarà identico a quello di ambedue i genitori qualora esso sia uguale nei due cromosomi (come il penultimo gene in questo esempio). Generando allora un certo numero di figli da  $x$  a  $y$ , tra di essi vi potrebbe essere:  $z = (3\ 2\ 1\ 2\ 3\ 3\ 2\ 1)$  che è la soluzione ottima del problema, di valore pari a 19.

Si noti che in questo caso il meccanismo dell' algoritmo è particolarmente semplice. Infatti, la rappresentazione delle soluzioni è tale che scegliendo in qualunque modo il valore di un gene tra quelli di due genitori, si ottiene una soluzione ammissibile. Questo non è sempre vero, almeno con rappresentazioni semplici delle soluzioni ammissibili. Infine, è evidente che in questo caso una mutazione può consistere semplicemente nell' assegnare un oggetto a un contenitore diverso.





















## ***Conclusioni***

Malgrado non esistano ancora standard di misurazione della Customer Satisfaction la strada della standardizzazione dei processi di lavoro in questo campo ormai è stata intrapresa. Le tendenze nell'ambito dei sistemi di misurazione sono abbastanza definite. Essendo questa un'area prettamente tecnico-statistica è anche molto più facile da standardizzare che non il sottosistema operativo manageriale della Customer Satisfaction.

L'interesse diffuso per le indagini di Customer Satisfaction tuttavia fa pensare che nei prossimi anni i sistemi informativi aziendali presenteranno sia procedure per il rilevamento periodico di dati di Customer Satisfaction sia procedure standardizzate per la gestione attiva del portafoglio clienti in cui entreranno notevoli moli di dati provenienti da indagini di Customer Satisfaction. I sistemi di Customer Satisfaction sono destinati a diventare o parte dei sistemi informativi integrati o sistemi informativi stand-alone di tipo DSS (Decision Support Systems).

Tutto questo servirà per vendere di più? A questa domanda ovviamente non si può rispondere in generale con un semplice sì o con un no: ci sarà chi riuscirà a vendere di più anche grazie a questi supporti e chi non ci riuscirà neanche con questi supporti. Ma ci saranno certamente sempre più aziende di piccole e medie dimensioni che riusciranno a vendere meglio i loro prodotti e servizi gestendo più efficientemente il loro lavoro di marketing e di vendita e valutando più accuratamente il roi sulla base della percezione della qualità dei loro clienti.

Questo lavoro ha presentato un metodo per misurare la CS. La potenza delle metodologie Fuzzy consente di analizzare giudizi espressi col linguaggio naturale, senza alterarne il loro contenuto di vaghezza ed ambiguità, pur pervenendo ad indicazioni di carattere operativo. I risultati sperimentali ottenuti conferiscono validità alla procedura adottata, aprendo l'orizzonte verso nuovi ambiti d'indagine.