



HAL
open science

Classificazione, analisi dei dati e modelli di conjoint analysis come strumenti integrati per le ricerche di mercato

Gilbert Saporta, Furio Camillo

► **To cite this version:**

Gilbert Saporta, Furio Camillo. Classificazione, analisi dei dati e modelli di conjoint analysis come strumenti integrati per le ricerche di mercato. *La Statistica per le imprese*, Societa Italiana di Statistica, Apr 1997, Torino, Italy. pp.15-32. hal-02514025

HAL Id: hal-02514025

<https://cnam.hal.science/hal-02514025v1>

Submitted on 24 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SOCIETÀ ITALIANA DI STATISTICA

Atti del convegno

La Statistica per le imprese

Sessioni plenarie
Sessioni specializzate
Tavole rotonde
volume 1

Convegno SIS
Torino 2-4 aprile 1997



Tirrenia
Stampatori



© Editrice Tirrenia Stampatori s.a.s., Torino, via Gaudenzio Ferrari 5

Stampa: Stampatre - Torino

Classificazione, analisi dei dati e modelli di *conjoint analysis* come strumenti integrati per le ricerche di mercato

Classification, Data Analysis and Conjoint Analysis Models : integrated tools for the marketing research

GILBERT SAPORTA

Conservatoire National des Arts et Métiers - Paris

FURIO CAMILLO

Dipartimento di Scienze Statistiche - Università di Bologna

1. Introduzione¹

Specie a partire dalla fine degli anni '70, le tecniche statistiche multivariate sono divenute uno strumento fondamentale per l'analisi dei dati relativi a studi di mercato. In molti casi proprio il particolare ambito di applicazione ha suggerito lo sviluppo di tecniche specifiche e di strategie particolari. La *conjoint analysis* rientra sicuramente in questa categoria di tecniche, visto che si ha traccia di uso di modelli di "misurazione congiunta" già nel 1964, negli Stati Uniti nell'ambito di problemi legati alla psicomètria (Luce e Tuckey, 1964). Negli anni immediatamente successivi gli aspetti applicativi maggiormente sviluppati furono quelli connessi alla ricerca delle condizioni sotto le quali stimare scale di misura per le variabili dipendenti da far corrispondere a scale di misura per quelle indipendenti, dato l'ordine degli effetti congiunti delle variabili indipendenti ed una specifica regola compositiva (Green e Srinivasan, 1978).

Il primo articolo dettagliato sulla *conjoint analysis* applicata allo studio del consumatore apparve nel 1971 (Green e Rao, 1971). In esso veniva discusso un possibile uso dei modelli multiattributivi per la quantificazione di giudizi di preferenza in termini di utilità ricevuta dal consumo di un prodotto. L'uso della *conjoint analysis* nell'ambito del marketing, specie negli Stati Uniti, si mostrò negli anni successivi di grande utilità, soprattutto per la realizzazione di nuovi prodotti o per simulare gli effetti sulle preferenze dei consumatori di cambiamenti di alcuni attributi di prodotti già esistenti. Lo sviluppo dell'apparato teorico ha però fortemente risentito degli aspetti e delle esigenze di tipo pratico e applicativo, il che ha portato i diversi ricercatori a proporre soluzioni molto personalizzate e difficilmente inquadrabili in un'opera completa. La stessa notazione formale presenta notevoli differenze così come è stata utilizzata nei diversi articoli apparsi nel corso degli anni su riviste specializzate.

Specialmente nella letteratura anglosassone è possibile individuare tre grandi classi di metodi usati negli studi di mercato per i dati di preferenza dei consumatori: i modelli di posizionamento, l'analisi esterna e interna delle preferenze, i modelli di scelta discreta del prodotto. Relativamente alla *conjoint analysis*, i modelli di posizionamento, che usano prevalentemente tecniche come il *multidimensional scaling* (metodi MDSCAL di J.B. Kruskal (1964) o INDSCAL di J.D. Carroll (1972)), permettono l'individuazione del posizionamento relativo dei prodotti, ma non di sapere qual è il prodotto ideale, né quali sono i prodotti preferiti. L'analisi interna ed esterna delle preferenze di fatto non riesce a spiegare le preferenze fra prodotti mediante le loro caratteristiche

¹ Il lavoro è il frutto comune degli autori. Tuttavia Saporta è l'autore dei par. 3.2, 3.3, 3.4, 3.6, 4.1 e Camillo dei par. 1, 2, 3.1, 3.5, 4.2. Il par. 5 è stato scritto da entrambi gli autori.

o i loro attributi (modello MDPREF (Carroll, 1972)). Infine i modelli di scelta discreta riescono ad analizzare la scelta o meno di un prodotto da parte dei consumatori sulla base di attributi qualitativi senza che gli intervistati debbano indicare un ordine fra differenti prodotti, i quali comunque è auspicabile che vengano costruiti con il ricorso a piani sperimentali (Sawtooth Software, 1995). Rispetto alla *conjoint analysis* la formazione del meccanismo di scelta è molto simile, tuttavia questi modelli sono di tipo globale e non riescono a trattare le preferenze espresse come scelte individuali.

Nel presente lavoro sono stati affrontati i principali aspetti metodologici legati alla conduzione di una *conjoint analysis* articolando la discussione secondo le diverse fasi di realizzazione sul campo di tali modelli e cercando di riacordare i diversi approcci presenti in letteratura. Allo stesso tempo sono state poste in evidenza le questioni applicative rilevanti e le relazioni di complementarietà della *conjoint analysis* con le strategie esplorative usualmente adottate per la segmentazione del mercato. In particolare la nostra proposta è relativa all'adozione di una strategia composta "in filiera" che consenta il completamento in senso "confirmatorio" della strategia esplorativa tipica della scuola francese dell'*Analyse des Données* nel trattamento di un'indagine diretta delle motivazioni, dei valori e delle preferenze dei consumatori.

2. L'*Analyse des Données* nelle inchieste per la segmentazione del mercato : una strategia composta

Secondo la classica definizione di W. Smith (1956) la segmentazione di qualunque tipo di mercato consiste nell'adeguare tanto i prodotti quanto le strategie di marketing alle differenze individuabili entro l'insieme delle esigenze manifestate dai consumatori. In questo contesto il compito della statistica è quindi quello di misurare e di rappresentare tali differenze in maniera che i prodotti e le attività di marketing possano adattarsi agli specifici bisogni individuati da ciascun segmento di consumatori. Più precisamente, gli studi di segmentazione hanno come "vocazione l'identificazione di gruppi omogenei d'individui sulla base di criteri comportamentali, attitudinali o percettivi. Tuttavia realizzare una segmentazione ha senso solo nel caso in cui i gruppi di individui possano essere identificati e se questa identificazione possa essere sfruttata in termini di marketing operativo" (Ladwein, 1996, pg. 14).

L'ultimo decennio ha visto il fiorire di numerose strategie di segmentazione del mercato mediante l'uso di dati provenienti da indagini dirette sui consumatori. La maggior parte di queste strategie, dovendo trattare una discreta quantità di variabili, ha fatto riferimento a metodologie statistiche di tipo multivariato, del tipo fattoriale e di *clustering*.

Nell'ambito operativo, la scelta del tipo di tecniche da usare si è dimostrata spesso dipendente dalla particolare formazione statistica dei ricercatori coinvolti nell'indagine, ma in effetti le diverse "filosofie" di ricerca riflettono differenze fondamentali già nelle ipotesi di partenza. La definizione di Ladwein sopra riportata sembra non lasciare dubbi sul fatto che solo le variabili comportamentali o motivazionali dovrebbero essere usate negli studi di segmentazione per definire gli algoritmi di *clustering*: infatti "i criteri socio-demografici, come l'età, il sesso, la categoria socio-professionale, il titolo di studio, sono molto utili ma solo perché permettono la descrizione operativa delle segmentazioni psicocognitive o comportamentali".

Nella tradizione anglosassone esiste invece una strategia di segmentazione che definisce *a priori* gli aspetti socio-demografici che determinano differenziali di comportamento o di valorizzazione dei prodotti. Questo procedimento implica che le variabili socio-demografiche siano usate non solo come "segnalistiche" per il riconoscimento operativo dei segmenti, ma anche come variabili strumentali alla suddivisione dell'insieme dei consumatori.

A partire dalla fine degli anni '70, soprattutto ad opera della scuola francese dell'*Analyse des données*, si è assistito ad un progressivo abbandono dell'idea di analizzare le potenzialità e la soddisfazione di un mercato mediante delle segmentazioni ipotizzate *a priori* per il prodotto. Tale impostazione ha proposto il

rovesciamento dello schema di analisi proprio in virtù della sua caratteristica descrittiva, ossia in maniera coerente alla sua peculiare funzione esplorativa, ricercando nei dati il modello comportamentale esistente, o meglio generato, dai soggetti indagati. Questi ultimi "generano" - e non "adottano" - uno o più modelli di comportamento o di valori, proprio come unità che singolarmente vi contribuiscono, cioè come "unità statistiche" che danno vita a più modelli collettivi sufficientemente omogenei al loro interno e altrettanto diversi fra loro (Evrard, Pras, Roux, 1993).

Dal punto di vista metodologico ciò ha implicato una particolare formulazione del questionario, in modo da consentire una segmentazione non ipotizzata *a priori* - in relazione ad un predefinito target socio-economico-demografico (variabili SED) -, bensì basata, dal punto di vista degli algoritmi di *clustering*, esclusivamente su domande relative a opinioni, motivazioni e percezioni circa il mercato o il prodotto in esame. Alcuni strumenti tipici della scuola francese dell'analisi multivariata, come quello della proiezione supplementare², sono risultati di fondamentale importanza e di fatto hanno consentito la messa a punto di strategie di analisi ben consolidate come la cosiddetta "analisi strutturale dei dati di inchiesta", definita usualmente come "*thémascopie*" (Morineau, 1994). L'impostazione tipologica francese propone quindi una classificazione dei soggetti intervistati innanzitutto per i loro comportamenti rispetto all'oggetto principale dell'indagine (nel caso del marketing, le motivazioni di consumo) e solo successivamente viene ricercata nelle variabili SED la spiegazione di tali comportamenti (Lebart 1986).

Una volta selezionate le unità statistiche costituenti il campione, le fasi principali di una ricerca possono essere schematizzate come segue: 1) redazione per sezioni, quindi per argomenti attivi e descrittivi (endogeni ed esogeni), del questionario di rilevazione; 2) applicazione di tecniche come l'Analisi delle Corrispondenze multiple per esaminare e rappresentare le interrelazioni fra le variabili indagate; 3) classificazione automatica dei soggetti del campione, eventualmente riponderati secondo il disegno campionario, mediante algoritmi di *cluster analysis*, condotti in genere solo sulle coordinate fattoriali rilevanti; 4) interpretazione dei comotati caratteristici delle tipologie individuate. Nelle ricerche di mercato le domande attive sono in genere costituite da valutazioni dei consumatori circa i fattori critici di successo del prodotto o i fattori potenzialmente desiderati.

Nell'ambito dell'impostazione tipologica francese è oggi usuale avvalersi di una o più domande aperte, le quali mirano a cogliere il senso evocativo e suggestivo del prodotto o del mercato studiato, sia per verificare l'eshaustività dell'insieme dei fattori critici di successo ipotizzati, sia, più genericamente, per completare in senso semantico l'atto del consumo da parte di ciascuna tipologia motivazionale individuata. Le domande aperte possono così costituire un supporto alle decisioni circa i contenuti creativi delle campagne di comunicazione promozionali e pubblicitarie sul prodotto, facendo emergere quelli che in genere vengono definiti come gli "schemi semiotici del consumo" (Camillo 1995; Grandi 1994).

Un questionario finalizzato all'approccio tipologico "alla francese" sarà pertanto articolato in quattro parti fondamentali: 1. insieme dei fattori di motivazione al consumo (fattori di successo), preferibilmente organizzati in variabili dicotomiche; 2. insieme delle domande sulle caratteristiche SED dei soggetti; 3. insieme delle domande che, mediante modelli statistici per lo più previsivi, consentano il *link* dell'indagine ad hoc con le informazioni di una o più grandi inchieste sulla fruizione dei media da parte dei cittadini; 4. una o più domande aperte d'ausilio al ruolo inventivo della pubblicità, ossia di analisi degli stereotipi derivati dall'evocazione.

In questo contesto di tipo esplorativo, o comunque di ricerca del modello comportamentale all'interno dei dati, l'introduzione di modelli di *conjoint analysis* può costituire un'integrazione analitica circa le preferenze dei soggetti intervistati, realizzando quello che in letteratura viene definito come "approccio composito" per la

² Nell'analisi delle corrispondenze multiple, la proiezione supplementare di una modalità *j* su un asse fattoriale è data da una particolare media ponderata delle coordinate degli individui che hanno scelto la modalità *j* di risposta. Un valore-test per la selezione delle sole modalità "rilevanti" per ogni singolo asse fattoriale è proposto in Lebart, Morineau e Piron (1995, pg. 123).

segmentazione di mercato. Classicamente, il modello di segmentazione composita sposta l'attenzione sull'individuazione dei profili dei consumatori che risultano maggiormente sensibili a particolari tipi di offerte di prodotto. In tale modo si cerca di misurare l'effetto congiunto delle caratteristiche dei consumatori e degli attributi del prodotto sulle preferenze espresse dagli intervistati (Brasini, F. Tassinari, G. Tassinari, 1995, pg. 179-184). Il modello di *conjoint analysis* decompone le utilità generali dei consumatori in specifiche utilità parziali, relative ai singoli attributi che costituiscono i connotati essenziali del prodotto o del servizio - già esistente sul mercato o anche solo potenziale - che è oggetto della ricerca. L'uso di una classificazione motivazionale o comportamentale consente la specificazione delle singole utilità per ogni segmento di consumatori e quindi offre un ventaglio di informazioni integrato per il perseguimento della soddisfazione del cliente. La decomposizione attuata dalla *conjoint analysis* presuppone, dal punto di vista assiologico, la possibilità di immaginare un prodotto o un servizio come una composizione soggettivamente percepita di vari attributi. La problematica del consumo può essere dunque spostata sulla "natura testuale" dell'oggetto consumato, ossia sul suo senso piuttosto che sulle sue caratteristiche oggettuali e concrete. Il prodotto, all'interno della teoria del marketing classico, si definisce essenzialmente per le sue proprietà fenomenologiche e tecniche: il volume, il peso, la forma. Ma tale concettualizzazione ha come conseguenza la rigida opposizione tra produzione e ricezione, la prima coincidente con la manifattura o la creazione, la seconda con l'atto fisico del consumo (Semprini, 1992).

In questo senso i modelli di *conjoint analysis*, innestati in studi di segmentazione, possono contribuire a evidenziare e quantificare anche le componenti immateriali del prodotto o del servizio, ossia quelle componenti capaci di incorporare un senso e di comunicare attraverso le più svariate materie espressive. Pertanto, uno dei momenti più delicati della conduzione sul campo di una *conjoint analysis* è proprio quello della definizione degli attributi rilevanti dei prodotti o servizi oggetto dell'indagine. In queste fasi l'analisi degli schemi semiotico-testuali del consumo derivati dal trattamento delle domande aperte del questionario motivazionale può contribuire in maniera determinante sia alla decodifica, sia alla comunicazione degli attributi chiave del prodotto all'interno di scenari di consumo sottoposti agli intervistati (Camillo, 1995; Lebart, 1996).

3. La Conjoint Analysis : definizioni, metodi e condizioni di realizzabilità

3.1 Il modello

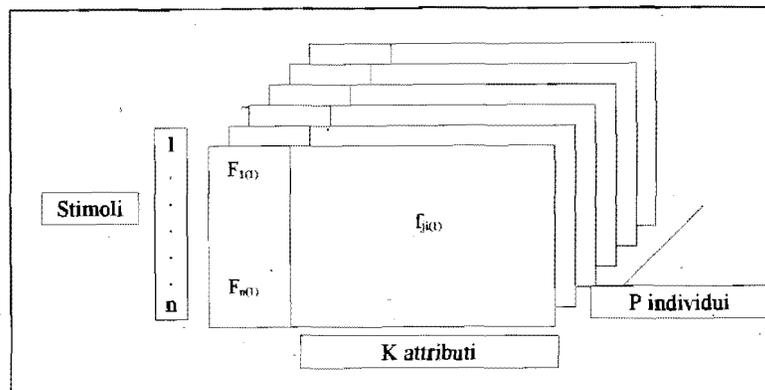
La valutazione dei giudizi di preferenza dei consumatori è stata affrontata attraverso la *conjoint analysis* già a partire dagli anni '70 facendo uso di quei modelli che esprimono le preferenze per un prodotto attraverso la specificazione di alternative multiattributive dello stesso. Una forma generale proposta di recente (van der Lans 1992) per tale tipo di modelli è quella per cui la variabile dipendente è una misura della preferenza che lo stimolo *j-esimo* assume per l'individuo generico *d-esimo* e che tale misura è funzione, per il singolo individuo, della modalità di percezione dell'attributo *i-esimo* dello stimolo *j-esimo* (Fig. 1).

Tali modelli sono costruiti sotto l'ipotesi che la preferenza di un individuo per uno stimolo o scenario - nel nostro caso un prodotto - sia funzione della posizione, così come è percepita dal singolo individuo, di quel particolare stimolo su un numero limitato di scale relative agli attributi usati.

L'accezione 'preferenza' assume diversi significati a seconda del campo d'indagine. Si può parlare di attitudine nel senso della propensione alla scelta, di valenza, ma anche di utilità soggettiva oppure, in casi particolari, di capacità o intenzione di acquisto di un prodotto.

Per 'stimolo', in via del tutto generale, si intende una qualunque entità che vuole essere sottoposta a giudizio. Nel caso specifico delle ricerche di mercato si tratterà di prodotti, servizi, ma anche di scenari rispetto ai quali si disegna un ambiente dove si svolge l'azione di acquisto, di consumo o di fruizione.

Fig 1 - Struttura dei modelli multiattributivi



Infine per 'attributo' si intende una caratteristica o un aspetto percepibile dello stimolo, quindi, per quanto riguarda un prodotto, l'attributo è uno dei possibili elementi che lo dota di utilità.

In termini formali, per il generico d -esimo individuo,

$$y_{j(d)} = F_{(d)}(f_{1(d)}, f_{2(d)}, \dots, f_{i(d)}, \dots, f_{K(d)})$$

con $y_{j(d)}$ = misura della preferenza che lo stimolo j assume per l'individuo d ; $F_{(d)}$ = funzione di preferenza dell'individuo d ; $f_{i(d)}$ = modalità percepita dall'individuo d , dell'attributo i , per lo stimolo j .

I valori di misura delle preferenze y sono numeri reali, vengono definiti sul prodotto cartesiano (stimoli - individui) e hanno una corrispondenza uno a uno con le preferenze. Assumendo che queste ultime rispettino i postulati d'ordine, non si perde di generalità nell'assegnare i valori in modo che a preferenze elevate corrispondano elevati valori di misura y . I modelli di preferenza multiattributivi, applicati al dominio microeconomico, possono essere interpretati come modelli di comportamento del consumatore, nei quali le singole funzioni di preferenza sono definite mediante l'utilità percepita nei singoli attributi di un prodotto. Tali modelli in ambito economico-aziendale hanno consentito lo studio dei comportamenti di scelta dei consumatori rispetto a una molteplicità di prodotti. In particolare la *conjoint analysis* consente di scomporre il giudizio globale che il consumatore dà a proposito di una serie di alternative in giudizi parziali circa i singoli attributi che compongono il prodotto stesso. Dal punto di vista metodologico la *conjoint analysis* include una serie importante di modelli e tecniche che consentono la trasformazione di risposte soggettive in stime di parametri: a partire da valutazioni globali si passa a scale di utilità separate, ma compatibili, per ogni attributo e con esse si possono misurare i benefici che il consumatore ricava dalle caratteristiche del prodotto. Tutto ciò nella *conjoint analysis* avviene senza che il consumatore esprima una valutazione diretta - e soprattutto non relativa a un prodotto specifico - dell'importanza dei singoli attributi. Inoltre, una volta stimati i parametri, si può costruire il giudizio globale relativo a prodotti "virtuali" (nuove combinazioni di attributi) che, in quanto tali, non sono stati oggetto diretto dell'espressione degli intervistati. I processi di simulazione delle scelte divengono pertanto utili alla definizione delle caratteristiche di un nuovo prodotto e stimano il suo successo relativo in termini di utilità arrecata ai singoli consumatori. Ovviamente, trattandosi di una tecnica applicata direttamente a dati individuali, l'affidabilità dei risultati di una *conjoint analysis* può risentire particolarmente dell'errore sia campionario che non campionario. In merito al primo è chiaro che occorrerà predisporre un piano di campionamento delle unità

statistiche (i consumatori, i clienti effettivi o potenziali, ecc.) che risponda il più possibile ai criteri di rappresentatività classici della teoria dei campioni. Nel presente lavoro non si è ritenuto opportuno trattare il problema del tipo di campionamento da adottare per una *conjoint analysis* visto che la questione può essere riportata al caso generale di predisposizione di un campione di unità rappresentative. Relativamente all'errore non campionario invece nel prosieguo del lavoro sono state discusse alcune problematiche connesse alla fase di raccolta dei dati, fase questa che per una *conjoint analysis* non è assolutamente indipendente né dal questionario adottato, né dal modello e dal metodo di stima prescelti.

A scopo illustrativo è possibile esemplificare il problema di una *conjoint analysis* con un semplice caso relativo a 12 tipi di autovetture, ciascuno caratterizzato da una combinazione particolare di 4 attributi: forma di pagamento (a1: sconto sul prezzo, a2: valutazione dell'usato, a3: rate a interesse zero), versione dell'auto (b1: monovolume, b2: due volumi, b3: station wagon), casa automobilistica (c1: Fiat, c2: Renault, c3: Ford, c4: Opel), vernice (d1: verde, d2: bordeaux, d3: nera, d4: grigia). Fra tutte le $3 \times 3 \times 4 \times 4 = 144$ combinazioni possibili delle modalità (tipi di auto) le 12 scelte per l'analisi consentono al disegno sperimentale di rispettare alcuni requisiti che verranno discussi nel successivo paragrafo 4.4.

A ciascun intervistato viene domandato di ordinare le 12 auto in ordine di preferenza e l'esempio riportato in fig.2 è relativo ad un soggetto che preferisce l'auto n°5 (graficamente evidenziata), composta da "valutazione dell'usato", "a due volumi", di marca Renault e di colore nero. Formalizzando, se y_j indica il rango d'ordine del prodotto n°5, si avrà che $f(y_j) = a_2 + b_2 + c_2 + d_3$, dove f sarà ipotizzata come una funzione monotona crescente. La stima dei parametri si avvarrà di una matrice X del disegno, che nel nostro esempio corrisponde a quella riportata nella parte destra della fig.2³. E' evidente che in tale matrice le colonne sono però legate da relazioni lineari che, rendendo indeterminati i coefficienti, impongono dei vincoli sui parametri, il più usato dei quali è quello di nullità della loro somma per il singolo attributo ($a_1 + a_2 + a_3 = 0$).

Il modello di stima deve essere quindi applicato a livello individuale, mediante tante regressioni, lineari o che utilizzino come variabile dipendente una trasformazione monotona dei giudizi di preferenza, quanti sono gli intervistati.

Fig.2 - Esempio di disegno, preferenze di un intervistato e matrice del disegno

auto	pagamento	versione	casa	vernice	rango	a1	a2	a3	b1	b2	b3	c1	c2	c3	c4	d1	d2	d3	d4
1	a1	b1	c1	d1	7	1	0	0	1	0	0	1	0	0	0	1	0	0	0
2	a2	b1	c1	d2	8	0	1	0	1	0	0	1	0	0	0	0	0	1	0
3	a3	b2	c1	d3	10	0	0	1	0	1	0	1	0	0	0	0	0	0	1
4	a1	b1	c2	d2	5	1	0	0	1	0	0	0	1	0	0	0	0	1	0
5	a2	b2	c2	d3	1	0	1	0	0	1	0	0	1	0	0	0	0	0	1
6	a3	b3	c2	d4	11	0	0	1	0	0	1	0	1	0	0	0	0	0	1
7	a1	b1	c3	d3	4	1	0	0	1	0	0	0	1	0	0	0	0	0	1
8	a2	b3	c3	d4	2	0	1	0	0	0	1	0	0	1	0	0	0	0	1
9	a3	b1	c4	d4	12	0	0	1	1	0	0	0	0	1	0	1	0	0	0
10	a1	b1	c4	d1	6	1	0	0	1	0	0	0	0	0	1	0	0	0	1
11	a2	b1	c4	d1	3	0	1	0	1	0	0	0	0	0	1	1	0	0	0
12	a3	b2	c4	d2	9	0	0	1	0	1	0	0	0	0	1	0	1	0	0

Le fasi di una *conjoint analysis*, pur essendo fortemente interrelate, possono essere, con buona approssimazione, organizzate nello schema seguente (Green e Srinivasan, 1978):

1. scelta del modello di preferenza rispetto al livello dell'attributo, ossia della funzione con la quale viene espresso il comportamento del consumatore all'atto di scegliere fra diverse alternative: massimizzazione della quantità di certi attributi oppure avvicinamento ad un proprio ideale modello di prodotto. Le possibili alternative sono fondamentalmente tre: il modello vettore (lineare), il modello punto ideale (quadratico), il modello *parth-worth* (discontinuo) (Green e Srinivasan, 1990);

³ In effetti nel nostro esempio, dato che la misura della preferenza avviene mediante il rango, la stima fornirà delle "anti-utilità", poiché i valori numerici più elevati di y sono associati ad automobili meno preferite.

2. scelta del metodo di raccolta dei dati, fra i due estremi di tipo "trade-off" fra due attributi alla volta e "full-profile", ossia di confronto dei profili completi del prodotto; a tale proposito nel seguito verrà presentata una particolare *conjoint analysis* definita "di tipo adattivo" (*Adaptive Conjoint Analysis*) (Johnson 1987) - completamente implementata in un software professionale di tipo interattivo - concepita proprio per risolvere i problemi derivanti da un questionario troppo complesso a causa dei molti attributi considerati;

3. costruzione dell'insieme di prodotti da sottoporre a giudizio mediante l'adozione di un disegno sperimentale che può passare da quello fattoriale completo ad uno definito con un processo di campionamento;

4. definizione delle modalità di presentazione dei prodotti agli intervistati, ossia mediante l'uso di descrizioni scritte o verbali del prodotto, opzione quest'ultima inevitabile per indagini somministrate telefonicamente, oppure utilizzando disegni, foto, altro materiale o addirittura i prodotti reali;

5. definizione della scala di misurazione delle preferenze dei prodotti dove le alternative sono rango d'ordine, *rating scale*, vari confronti a coppie o assegnazione di una modalità dal "non gradito" al "poco gradito", fino al "molto gradito";

6. scelta del metodo di stima fra tre grandi categorie: metodi metrici (regressione multipla), metodi non metrici (principalmente il MONANOVA o quello di Johnson di tipo adattivo) e, infine, metodi delle scelte in probabilità (logit, probit).

Nel prosieguo tali fasi non verranno discusse tutte in maniera puntuale, non solo per motivi di spazio ma anche per l'impossibilità di trattarle in maniera disgiunta l'una dall'altra.

3.2 La stima delle utilità

I dati di cui si dispone sono la descrizione di n prodotti sotto forma di scenari completi descritti da p fattori qualitativi in una tavola disgiuntiva completa X a n righe e la matrice Y a n righe e m colonne dei ranghi di classificazione fornita da ogni intervistato. I calcoli si effettuano separatamente per ognuna delle colonne y di Y . Per ottenere delle utilità e non delle anti-utilità si centra y e se ne cambia il segno.

Come accennato, si stimano le utilità individuali con una regressione lineare oppure con una regressione monotona. Per la regressione usuale, avremo che $y = X\beta + e$, stimando il vettore β attraverso b in maniera che

$\|y - X\beta\|^2$ sia minimo.

Tali stime però non sono definite se non si impone un vincolo sulle utilità di ogni attributo, il più usuale dei quali è quello di nullità della loro somma. Allora gli scarti fra utilità saranno stimabili univocamente.

Essendo y una variabile ordinale, appare più naturale ricercare la trasformazione monotona $T(y)$ che assicuri la migliore regressione multipla. Una trasformazione monotona di y può assumere la forma:

$$T(y_1) = a_1$$

$$T(y_2) = a_1 + a_2$$

$$T(y_3) = a_1 + a_2 + a_3$$

...

$$T(y_n) = a_1 + a_2 + \dots + a_n$$

dove a_1 è di qualunque segno mentre gli altri coefficienti sono positivi o nulli. $T(y)$ si esprime come una combinazione lineare di n variabili a coefficienti positivi (meno uno) e questo fornisce all'insieme delle trasformazioni monotone di y la struttura di un cono poliedrico convesso. La ricerca della minimizzazione di $\|T(y) - Xb\|^2$ è allora un problema di analisi canonica fra questo cono e il sottospazio vettoriale

W generato dalle colonne di X . Questo problema può essere risolto mediante un metodo di proiezioni successive (Agha, 1991; Tenenhaus, 1988). Conoscendo T , si tratta di una regressione multipla ordinaria intesa come proiezione sul sottospazio generato dalle colonne di X . Se Xb è fissato si ottiene T proiettandolo sul cono. Tale proiezione può essere ottenuta con un algoritmo dovuto a Kruskal e Carmone (1968) oppure applicando la tecnica proposta da Tenenhaus (1988), per cui proiettare un vettore z su un cono consiste

ogni fattore, e dall'altro queste somme non sono additive se non nel caso particolare dell'adozione di piani ortogonali. Occorrerebbe usare pertanto delle somme di quadrati del II tipo, con i relativi test F associati.

L'importanza relativa dei differenti fattori è calcolata tenendo conto del fatto che l'ampiezza massima per due prodotti (la differenza fra il miglior punteggio e il più debole) è uguale alla somma delle ampiezze delle utilità per ogni attributo.

3.4 La costruzione dei prodotti (scenari) mediante piani sperimentali

Se un prodotto è definito da p fattori con m_1, m_2, \dots, m_p categorie, esisteranno $(m_1 \times m_2 \times \dots \times m_p)$ prodotti. Presentarli tutti costituisce un piano fattoriale completo, ma in genere l'alto numero dei prodotti rende irrealizzabile l'uso di un tale piano. Per esempio, dovendo prendere in considerazione 5 attributi del servizio di una compagnia aerea (A, B, C, D, E), ciascuno dei quali è composto da due livelli, avremo $2^5=32$ scenari possibili da sottoporre agli intervistati⁵. Il numero di coefficienti (le utilità) da stimare per il modello a effetti additivi è uguale a $\sum_i(m_i - 1) = \sum_i m_i - p$; nel nostro esempio sarebbe uguale a 5. E' possibile pertanto ridurre il numero di scenari a condizione di rimanere comunque di alcune unità al di sopra del numero di parametri, specie se si effettua una trasformazione monotona sulle preferenze.

Come selezionare gli scenari da considerare? Si tratta di stimare al meglio i coefficienti di un modello lineare. E' noto che in tale modello la matrice varianze-covarianze di \mathbf{b} è $V(\mathbf{b}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$. Occorrerà scegliere la migliore matrice \mathbf{X} ed appare chiaro che ciò non dipende dall'uso o meno di una trasformazione monotona di y . La teoria dei piani sperimentali permette di rispondere a questo problema, utilizzando quando possibile delle frazioni ortogonali del piano completo, che permettano di ben separare le influenze dei fattori e che godano di proprietà di ottimalità universale. I piani *D-ottimali* sembrano costituire lo strumento adatto per la soluzione del problema, ma la generazione di questi impone l'uso di software specialistici, in genere progettati per il controllo di qualità industriale (Mitchell, 1974). La costruzione di piani ortogonali, per contro, può anche essere effettuata, nei casi semplici, senza l'ausilio dell'elaborazione elettronica.

	FRAZIONE %				
	A	B	C	D	E
1	1	1	1	1	2
2	2	1	1	1	1
3	1	2	1	1	1
4	2	2	1	1	2
5	1	1	2	1	1
6	2	1	2	1	2
7	1	2	2	1	2
8	2	2	2	1	1
9	1	1	1	2	1
10	2	1	1	2	2
11	1	2	1	2	2
12	2	2	1	2	1
13	1	1	2	2	2
14	2	1	2	2	1
15	1	2	2	2	1
16	2	2	2	2	2

Utilizzare dei piani frazionari, o a numero minimale di prove, non permette di stimare certe interazioni e conduce a limitarsi ai modelli puramente additivi, ossia al modello più tipico per la *conjoint analysis*. Si è inoltre visto in precedenza come le trasformazioni monotone sulla risposta hanno l'effetto di annullare certe interazioni o effetti moltiplicativi.

Poiché nel nostro esempio sul servizio aereo gli attributi hanno tutti 2 modalità (livelli), si userà preferibilmente i piani di tipo 2^{p-k} o piani di Placket e Burman, i quali sono ortogonali. Riprendendo l'esempio, il piano completo a 32 scenari sugli attributi A, B, C, D, E permette di stimare tutte le interazioni fino a quelle di ordine 4. E' relativamente facile ottenere una metà ortogonale del piano completo composto da 16 prove (Frazione₅). Si cercherà di avere delle colonne distinte e ortogonali per potere stimare gli effetti, ma siccome si incontreranno degli effetti di confondimento, si opererà in maniera che queste confusioni non compaiano se non a un livello elevato d'interazione. Si sceglierà nel nostro caso una "confusione" o generatore $E=ABCD$. Questo piano è di risoluzione V e la sua ortogonalità si verifica come segue: si rimpiazzano gli 1 con dei -1 e i 2 con degli 1. I prodotti scalari fra colonne del piano sono allora nulli.

⁵ Questi attributi potrebbero essere: A: classe (business, economica), B: durata del volo (1h, 2h), C: accesso al salone (si, no), D: pasto (caldo, freddo), E: servizio a bordo (ridotto, normale).

Si può ancora dividere per 2 il numero di prove usando lo stesso principio : si parte dal piano completo associato a A B C e si riaggiungono le colonne D e E con i generatori D=AB e E=AC.

FRAZIONE 1/4

	A	B	C	D	E		A	B	C	D	E	F	G
1	1	1	1	2	2	1	1	1	1	2	2	2	1
2	2	1	1	1	1	2	2	1	1	1	1	2	2
3	1	2	1	1	2	3	1	2	1	1	2	1	2
4	2	2	1	2	1	4	2	2	1	2	1	1	1
5	1	1	2	2	1	5	1	1	2	2	1	1	2
6	2	1	2	1	2	6	2	1	2	1	2	1	1
7	1	2	2	1	1	7	1	2	2	1	1	2	1
8	2	2	2	2	2	8	2	2	2	2	2	2	2

L'ultimo piano, di risoluzione III è il piano ortogonale in senso stretto, minimo per il nostro problema. Il primo piano è in effetti l'inizio di un piano a 8 scenari di risoluzione III, potendo trattare 7 fattori al massimo (piano Taguchi $L_8 2^7$). Dagli 8 fattori in poi bisognerebbe passare a 16 prove per avere un piano ortogonale. Tuttavia si userà un piano Plackett e Burman a 12 prove, piano molto classico, costruito con il ricorso alle matrici di Hadamard.

Quando gli attributi hanno tutti k modalità si è soliti usare i piani quadrati latini o greco-latini. Consideriamo ora un caso di detersivi caratterizzati da tre fattori a tre categorie: marca (m1, m2, m3), consistenza (classica, micro, liquida), prezzo (p1, p2, p3). Ci sono dunque 27 combinazioni possibili che si riducono a 9 usando un piano quadrato latino come quello riportato nella tabella di sinistra di Fig.3. I quadrati latini si costruiscono agevolmente per k qualunque, ma sono di scarsa utilità per la *conjoint analysis* se si procede con il metodo di rilevazione "full profile", ossia con un metodo di rilevazione per il quale occorre chiedere il giudizio su tutti i prodotti agli intervistati (vedi paragrafo seguente). Se k va oltre 4, k^2 potrebbe risultare troppo elevato.

Fig.3 - Quadrato latino e quadrato greco-latino

	Marca			Ammorbidente	Marca		
	m1	m2	m3		m1	m2	m3
Consistenza	p1	p2	p3	Consistenza	p1 α	p2 β	p3 γ
classica	p2	p3	p1	classica	p2 γ	p3 α	p1 β
micro	p3	p1	p2	liquida	p3 β	p1 γ	p2 α
liquida							

Questa costruzione può comunque essere generalizzata a 4 fattori (quadrato greco-latino): aggiungendo il fattore "ammorbidente", articolato anch'esso in tre livelli (α , β , γ) si otterrebbe lo schema riportato nella parte destra di Fig.3. Questo ultimo piano comporta solo uno scenario in più del numero dei parametri da stimare.

La costruzione di piani "asimmetrici", cioè per dei fattori aventi un numero diverso di modalità, è più delicato e non si può fornire alcun principio generale di costruzione. Tuttavia una tecnica molto vantaggiosa in tali situazioni si è mostrata quella del *collapsing* di livello, che partendo da un quadrato latino, mediante degli opportune sostituzioni di un livello con un altro, giunge a dei piani parsimoniosi e allo stesso tempo ortogonali (Saporta, 1992).

Quando non è possibile trovare dei piani ortogonali aventi un numero di scenari ragionevole, si adotta la soluzione di invertire il procedimento, fissando *a priori* il numero di scenari e scegliendo il miglior piano-risultato. Come accennato in apertura di paragrafo, il criterio usato è quello della *D-ottimalità*, che consiste nella minimizzazione del determinante della matrice $(X'X)^{-1}$, ossia nella massimizzazione di quello della $X'X$. La ricerca di tali piani si compie con l'aiuto di algoritmi, che in alcuni casi come la Proc OPTEX di SAS-QC, sono

nell'effettuare una regressione multipla ordinaria di z sulle variabili indicatrici dell'ordine, definito attraverso la matrice (n,n) sotto riportata.

Matrice (n,n)

```

100000000000
110000000000
111000000000
111100000000
.....
.....
111111111111

```

Se i coefficienti sono tutti positivi salvo il primo ciò che si ottiene è già la proiezione sul cono, altrimenti si eliminano le colonne corrispondenti a dei coefficienti negativi, o meglio si pongono tali coefficienti come nulli, ossia si danno degli ex-aequo, e si ricomincia la regressione multipla fino a quando tutti i coefficienti siano positivi o nulli. Un algoritmo iterativo, come quello contenuto nel software SPAD-MC, consiste quindi nella proiezione di y su W in y^* , quindi come nei minimi quadrati ordinati⁴. Poi y^* verrà proiettato sul cono, e questa proiezione su W , rinormalizzando le proiezioni a ogni passo. In effetti l'analisi della varianza con trasformazione monotona della risposta era stata inizialmente proposta da Kruskal al fine di eliminare gli effetti di

interazione dovuti alla non linearità della relazione fra risposta e fattori: il caso più classico è quello di un modello a effetto moltiplicativo $Y = X1 * X2$ che implica una dichiarazione di non nullità dell'interazione fra $X1$ e $X2$ nel modello lineare classico. Una trasformazione logaritmica di Y rende il modello additivo senza interazioni. Come avviene per questa trasformazione, che è una trasformazione monotona particolare, si riconosce che sono numerosi i casi in cui nella ricerca di trasformazioni monotone si giunge a eliminare l'eventuale non additività, ciò che costituisce uno degli argomenti a favore della limitazione nella *conjoint analysis* alla considerazione degli effetti principali, oltre che quello di non complicare l'interpretabilità dei risultati (Green, 1984).

Quale metodo di stima scegliere fra minimi quadrati ordinari e regressione monotona? E' evidente che le trasformazioni monotone meglio si adattano alla natura ordinale dei dati raccolti e al modello additivo postulato, ma possono condurre a degli aggiustamenti artificiali, specie se il numero di scenari è insufficiente (rischio di *overfitting*). Nelle applicazioni operative è quindi raccomandato di prendere in esame in primo luogo la soluzione dei minimi quadrati ordinari, prima di un'eventuale opzione per la regressione monotona: i primi sono in generale dotati di maggiore robustezza.

La bontà della ricostruzione delle preferenze è misurata sia con l' R^2 , per i minimi quadrati ordinari, sia con l'indice Tau di Kendall. Qualora si ottengano valori di tali indici troppo bassi si conclude per un'incoerenza delle risposte fornite dall'intervistato (stanchezza del rispondente, incomprensione del metodo o altri fattori che potrebbero riguardare l'errore non campionario), oppure per un'inadeguatezza del modello d'utilità additive. Occorre tenere presente che però quest'ultima spiegazione minerebbe la ricerca proprio nei presupposti originari dell'adozione della *conjoint analysis*. In generale, verranno quindi eliminate quelle osservazioni (rispondenti) per i quali l' R^2 o il Tau è inferiore ad una certa soglia fissata dal ricercatore. Operativamente non è esclusa la possibilità di rinunciare anche al 10-15% di risposte giudicate come incoerenti nella ricostruzione delle preferenze.

3.3 La gerarchia dei fattori

Per ogni intervistato, i fattori o variabili esplicative possono essere ordinati in una gerarchia di importanza. Poiché si tratta di un modello di analisi della varianza a effetti additivi potrebbe essere ragionevole utilizzare la somma dei quadrati spiegati da ciascun fattore. Questa soluzione in generale non viene però adottata in quanto da un lato le somme dei quadrati hanno tendenza ad aumentare con il numero di modalità di

⁴ Per un'esposizione di tale approccio geometrico al modello lineare e per la disamina dei legami con l'analisi canonica si rimanda a Lebart, Morineau, Piron (1995, pg. 225-230)

già implementati in pacchetti statistici. Si noti che se esiste un piano ortogonale, essendo questo ottimale, la procedura di SAS-System lo trova quando converge (SAS-QC Software, 1994). Abituamente nelle procedure informatiche è inoltre possibile imporre certi scenari e "lanciare" l'ottimizzazione del criterio solo su degli scenari ritenuti complementari.

3.5 A proposito della raccolta dei dati

I metodi per la raccolta dei dati di una *conjoint analysis* sono fondamentalmente due: il metodo *trade-off* e il metodo *full-profile*.

Il primo, denominato in italiano "due attributi per volta", consiste nel sottoporre agli intervistati delle tabelle in cui sono riportate tutte le possibili coppie di livelli (le modalità: non importante, poco importante, ecc.) di due soli attributi (le caratteristiche del prodotto: il prezzo, il *packaging*, ecc.) e domandare loro di ordinarle. Ci sarà quindi una tabella per ogni coppia di attributi possibile, tanto più complessa quanti saranno i livelli. Inoltre di tali tabelle ce ne saranno tante quante sono le coppie di attributi.

Il secondo metodo, il *full-profile*, consiste nel fare valutare agli intervistati un insieme di prodotti completi, descritti quindi facendo riferimento a tutti gli attributi considerati nei loro livelli.

La tecnica del *trade-off* venne sviluppata a partire dagli anni '70 da Johnson (1974), il quale valorizzò il maggiore potenziale informativo di tale metodo dato che l'intervistato valuta gli attributi singolarmente e in modo diretto. L'evoluzione di tale tecnica, integrata in algoritmi di calcolo che fanno ricorso a particolari ipotesi di ponderazione che si riadattano dinamicamente a seconda delle risposte fornite dall'intervistato, ha portato alla messa a punto del software "ACA-System" della *Sawtooth Software*, il quale esegue direttamente a computer sia la fase di raccolta dati mediante un'intervista interattiva, sia la successiva elaborazione (ACA System, 1994-95, app. G). Si tratta di un metodo di tipo "ibrido" che unisce approcci differenti: dapprima con un "*self-explicated model*" individua a livello di singolo rispondente quali sono gli attributi più importanti e in un secondo momento, solo sui più importanti, viene adottato un modello *part-worth* di confronto a coppie. Il software ACA è del tutto integrato e consente una gestione globale dell'indagine già a partire dalla costruzione del questionario, che avviene in maniera automatica dopo avere imputato i dati del problema circa gli attributi e le modalità di loro articolazione.

In letteratura sono numerosi gli studi di verifica delle procedure proposte globalmente da ACA System, sia in termini di efficienza del metodo che di accuratezza delle stime (Green, Krieger, Argawal, 1991).

Il problema fondamentale tuttavia resta quello, come tutti i metodi derivati dall'approccio "*trade-off*", che gli intervistati non valutano effettivamente i prodotti, bensì compiono una serie di scelte a catena circa livelli e attributi di prodotti selezionati secondo un criterio adattivo. Di fatto ciò determina da parte dell'intervistato un ragionamento circa i singoli attributi che compongono un prodotto, distogliendo la sua attenzione dal prodotto globalmente considerato, che procura utilità anche grazie all'addizione dei diversi attributi. Solo ad esempio, presentare le opzioni come una lista di marche, ognuna affiancata dai suoi attributi, non provoca le stesse valutazioni che la presentazione di una serie di attributi salienti ciascuno dei quali è affiancato da una o più marche (Rumiati 1990).

Per tale motivo ci sembra che nelle fasi di raccolta dei dati per una *conjoint analysis* un intervento che metta direttamente in evidenza gli attributi di un prodotto potrebbe disturbare l'apparato di valutazione del rispondente il quale, proprio per gli scopi della *conjoint analysis*, deve invece concentrarsi sull'utilità apportata globalmente da tutto l'insieme degli attributi che definiscono il prodotto. Ci sembra che solo in tale modo la filosofia di decomposizione analitica delle preferenze, tipica della *conjoint analysis*, venga rispettata e che possa permettere di stimare la misura latente delle utilità parziali, la qual cosa è effettivamente strategico rispetto agli obiettivi operativi dello studio di mercato. Per questo motivo il *wording* del questionario descrittivo dei prodotti dovrebbe riuscire a spiegare l'esistenza di un livello o modalità di un attributo senza farne diretto riferimento. E'

ovvio che se si vuole sfruttare a pieno le caratteristiche decompositive della *conjoint analysis* occorre adottare un metodo di raccolta del tipo *full-profile*, che però allo stesso tempo non sia troppo complesso da condurre i rispondenti ad un effetto di *information overload*, ossia ad un fenomeno di confusione fra le effettive utilità globali apportate dai prodotti proposti.

Il controllo della complessità del questionario dunque può essere effettuato operando sul numero di prodotti proposti, ossia sul disegno sperimentale, oppure cercando di definire in maniera essenziale il numero di livelli o modalità di un attributo, fino a spingersi dove possibile all'indicazione dicotomica della singola caratteristica (presenza/assenza).

Dal punto di vista pratico è ragionevole non presentare più di 16 scenari all'intervistato. Una tecnica particolare per la limitazione della complessità è costituita da una tecnica di "ponte" o "passerella", che realizza un collegamento fra più *conjoint analysis*. In un caso di studio non pubblicato e ancora non divulgabile condotto in Francia, ci si trovava di fronte a scenari definiti da ben 12 attributi. Si è fatto pertanto ricorso a tale tecnica, realizzando tre *conjoint analysis* con 4 attributi ciascuna. In un secondo momento è stata realizzata un'analisi di congiunzione considerando tre attributi, ciascuno proveniente dalle tre diverse analisi precedenti. Mediante delle regressioni lineari è possibile ricavare le ponderazioni da applicare a ciascun insieme di utilità, individuo per individuo, per giungere ad una formula globale.

3.6 L'uso nelle simulazioni

Uno dei principali interessi della *conjoint analysis* risiede nella possibilità di stimare delle preferenze rispetto a prodotti non proposti nell'inchiesta. Questo avviene grazie al modello delle utilità additive. Si faccia l'ipotesi che la scelta finale si riduca a tre scenari. E' possibile dunque calcolare per ogni intervistato i punteggi di questi tre prodotti e dedurre la proporzione di coloro che preferiscono l'uno piuttosto che l'altro. In altri termini, è possibile una stima delle quote di mercato di questi tre prodotti. Questo metodo è detto dell'utilità massimale.

Esistono altri due metodi più classici di stimare le quote di mercato in maniera meno "grossolana", i quali sono da tenere in considerazione specie se le utilità risultano molto simili: il modello logistico, dove la probabilità di scelta di un prodotto di utilità U_i è proporzionale a $\exp(U_i)$; il modello di Bradley Terry Luce (BTL) dove questa probabilità è proporzionale a U_i (può essere applicato solo quando U_i è positiva).

I due modelli hanno un certo interesse teorico perché probabilizzano la decisione del consumatore. Tuttavia il modello logistico non è invariante, se non ad una traslazione delle utilità, e quello BTL non lo è, se non per una moltiplicazione delle utilità con una stessa costante. Il modello di utilità massimale è invece invariante per tutte le trasformazioni monotone e fornisce, se il campione è di grandi dimensioni, una stima corretta delle probabilità. Una critica di fondo è comunque formulabile a tutti questi modelli: dal punto di vista metodologico non sembra sempre giustificato dedurre da una preferenza un'intenzione d'acquisto. Occorre inoltre pensare anche alla cosiddetta "opzione zero", dove nessuno dei prodotti proposti sarà acquistato.

Chiaramente la *conjoint analysis* non ha come obiettivo primario la stima delle quote di mercato, bensì quello dell'esplicitazione delle preferenze. Probabilmente se occorresse conoscere l'intenzione d'acquisto dei consumatori varrebbe la pena domandarglielo esplicitamente, in una ricerca progettata e condotta *ad hoc*.

Tuttavia, se si dispone per ogni prodotto della variabile sì/no, prodotto accettabile o no, si può studiare la percentuale d'accettazione o di rifiuto di ogni prodotto. Se inoltre le condizioni d'omogeneità sono rispettate, potrà essere tracciata una nuvola dei punti, con in ascissa il punteggio medio d'utilità s di ogni prodotto e in ordinata il tasso d'intenzione di acquisto p . In generale un aggiustamento mediante una funzione logistica $p = \exp(a + bs) / (1 + \exp(a + bs))$ fornisce dei buoni risultati e permette di posizionare i prodotti non effettivamente testati, quindi simulandoli (Darmon, 1992).

4. L'uso della *conjoint analysis* : due casi di studio

4.1 La telefonia cellulare in Francia : un caso classico

Uno degli operatori francesi desiderava studiare diversi scenari di servizi e tariffe per un segmento di clientela particolare. Gli attributi utilizzati nella relativa *conjoint analysis* sono stati i seguenti 7 :

il costo dell'apparecchio telefonico (gratis oppure 700F), i costi di allacciamento alla rete (gratis oppure 200F), la durata del contratto (2 modalità), il periodo di applicazione della tariffa ridotta (2 modalità), l'abbonamento mensile (3 modalità), la tariffa di comunicazione in ora di punta (3 modalità), la tariffa di comunicazione in ora non di punta (3 modalità). Si trattava di stimare 10 parametri. Il piano *D-ottimale* che segue, organizzato in 12 scenari, è quasi ortogonale : i fattori a due modalità sono ortogonali fra loro e ortogonali ai fattori a tre modalità.

scenario	apparecchio	allacciamento	durata	periodo ridotta	abbonamento	tariffa o. punta	tariffa n.o. punta	
1	2	2	2	2	3	1	1	1
2	2	2	1	1	3	3	2	2
3	2	2	1	1	1	2	3	3
4	2	1	2	2	1	2	2	2
5	2	1	2	1	2	1	3	3
6	2	1	1	2	2	3	1	1
7	1	2	2	2	1	3	3	3
8	1	2	2	1	2	2	1	1
9	1	2	1	2	2	1	2	2
10	1	1	2	1	3	3	2	2
11	1	1	1	2	3	2	3	3
12	1	1	1	1	1	1	1	1

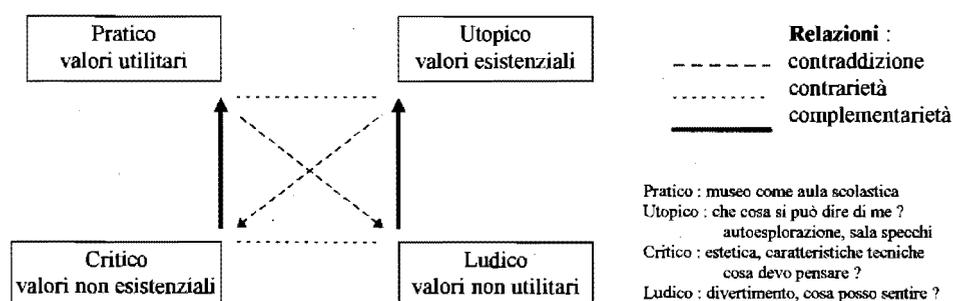
Un campione di 263 individui ha consentito di realizzare, mediante il calcolo dell'importanza media dei singoli attributi, una gerarchia dei fattori : 1. il costo dell'apparecchio telefonico, 2. la tariffa di comunicazione in ora di punta, 3. l'abbonamento mensile, 4. la tariffa di comunicazione in ora non di punta, 5. i costi di allacciamento alla rete, 6. la durata del contratto, 7. il periodo di applicazione della tariffa ridotta. La ricerca ha mostrato come il costo dell'apparecchio sia il fattore più importante fra quelli da considerati nel *marketing mix* della compagnia telefonica. Lo studio ha consentito l'individuazione di una certa sensibilità ancora molto forte a fattori globalmente indicatrici del "costo" del servizio, come il tipo di tariffa in ora di punta oppure le varie forme di abbonamento. Coerentemente, in Fig.4, le utilità medie mostrano l'insieme di attributi determinanti nella definizione del servizio di telefonia mobile. Il campione si è mostrato sufficientemente omogeneo e i risultati individuali hanno potuto essere rappresentati attraverso le loro medie.

Per ognuno dei 12 scenari si possedeva anche il giudizio circa l'intenzione di acquisto degli intervistati. L'adattamento di una funzione logistica che legava la percentuale di intenzione di acquisto e *score* d'utilità media è risultato eccellente ($R^2 = 0.96$). Nello studio, la stima di tale funzione ha permesso l'uso della *conjoint analysis* anche per l'ottenimento previsioni della probabilità di acquisto di vari tipi di servizi di telefonia mobile, anche se composti da attributi non combinati in uno specifico scenario sottoposto a valutazione degli intervistati.

4.2 "Gli italiani e l'arte" : una ricerca condotta con una strategia in filiera

Si fa qui riferimento ad una ricerca nazionale condotta mediante somministrazione di un questionario telefonico a proposito delle modalità di fruizione e di consumo di arte in Italia. La ricerca aveva come obiettivo

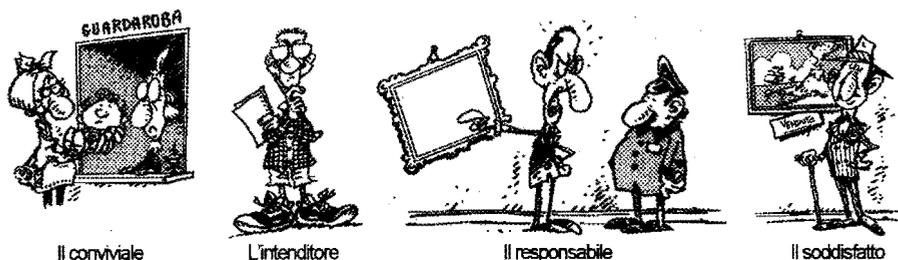
fondamentale quello di definire “gli stili di fruizione dell’arte” attraverso la valutazione di una serie di fattori potenzialmente determinanti per una buona riuscita di una mostra o di un’esposizione. L’indagine ha riguardato un campione di 3089 soggetti, selezionati mediante una stratificazione per area geografica, sesso e classe d’età⁶. Il questionario sottoposto agli intervistati era composto in maniera aderente ai “canoni” esplicitati nei paragrafi precedenti dell’approccio tipologico francese. La sua redazione si era però avvalsa di un metodo semiotico di modellizzazione della ricezione (il consumo) del “prodotto” museo o esposizione. Tale modello è detto quadrato semiotico ed è composto da posizioni logiche (e non tanto atteggiamenti individualizzabili) e soprattutto sono in relazione dinamica fra loro. Secondo Floch, il capostipite della scuola semiotica applicata al marketing, a seconda dei casi, la valorizzazione di un prodotto potrà essere pratica, utopica, critica o ludica (Floch, 1992). In tale modello interpretativo la cosa interessante è che il passaggio da un tipo di valorizzazione all’altro è solo questione di occasione o contesto e di un implicito (ignaro) uso degli operatori logici che conduce i soggetti a consumare in un modo piuttosto che in un altro, preferendo quindi un bene a scapito di un altro.



Il modello semiotico applicato nella costruzione del questionario de “Gli italiani e l’arte” è stato elaborato partendo dallo schema, sotto riportato, adottato in occasione di uno studio condotto sull’esposizione *Passport to the World* all’*Indianapolis Children’s Museum* (Umiker-Sebeok, 1992). Occorre qui sottolineare che nello schema semiotico non si fa riferimento ad una segmentazione di soggetti, quanto a diversi modi di valorizzazione del prodotto che mutano secondo percorsi logici dinamici. Sono stati pertanto individuati 14 possibili fattori di successo di una mostra-esposizione che avrebbero successivamente consentito anche un’ipotesi di posizionamento dinamico dei soggetti intervistati e non solo una segmentazione del tipo *clustering*. Quest’ultima ha prodotto 7 tipologie di italiani, 4 di visitatori più o meno abituali di mostre o musei e 3 di non visitatori. La strategia adottata ha fatto ricorso ad un’analisi delle corrispondenze multiple e all’algoritmo di *cluster analysis* non gerarchico-misto delle nuvole dinamiche di Diday, particolarmente adatto a questo tipo di problemi (Celeux e altri, 1989).

Volendo soffermarsi sulle 4 tipologie di visitatori - descritte sotto con le vignette ufficiali di Silver degli *Stili di Fruizione dei Musei*[®] (ArtMonitor, 1996) - nella ricerca si è poi proceduto all’analisi delle due risposte aperte contenute nel questionario, relative all’evocazione suscitata negli intervistati dalle parole *arte* e *museo*.

⁶ La ricerca “Gli italiani e l’arte” è stata realizzata nel 1996 nell’ambito dell’attività di ARTMONITOR SZ[®] - Osservatorio permanente per lo studio della fruizione e del consumo di arte in Italia - il cui comitato scientifico è composto da alcuni docenti dei dipartimenti di Comunicazione, Economia e Statistica dell’Università di Bologna, nonché dello IULM di Milano. La rilevazione è stata curata dalla società Agorà di Ferrara.



L'analisi delle corrispondenze della matrice lessicale, definita con sulle righe le forme testuali lemmatizzate semanticamente⁷ e sulle colonne le 4 tipologie di visitatori, insieme ad un'analisi delle specificità lessicali condotta sia col metodo della "proporzione", sia con quello del CHI2 (Lebart, Salem, 1994), ha consentito la messa a punto del sistema di ipotesi dinamica delle posizioni sul quadrato semiotico rispetto a ciascuna delle 4 tipologie. Tale procedura ha permesso la redazione di un questionario particolare per una successiva reintervista di un campione di 356 soggetti che facevano parte delle 4 tipologie di visitatori: l'obiettivo della reintervista era quello di realizzare una *conjoint analysis* sui fattori che determinano le preferenze circa il tipo di visita a musei, mostre o esposizioni. La particolarità del questionario somministrato è che ciascuno degli 8 scenari costruiti con un piano sperimentale di frazione 0.5 a 4 fattori dicotomici (comodità, fruibilità, accuratezza, prezzo/bene sociale), è stato proposto agli intervistati mediante la narrazione di storielle o la descrizione di situazioni. Ad esempio, il primo scenario, quello relativo ad un museo comodo, di facile visitabilità, ma non troppo curato nell'esposizione delle opere e interamente autofinanziato, è stato descritto in questo modo:

Lei entra, trova che il biglietto costa 25000 lire, ma non si sorprende, né si lamenta, dato che è noto che quel museo non riceve più fondi pubblici da tempo. Nonostante il discreto affollamento riesce a lasciare rapidamente le sue cose al guardaroba e si incammina verso il percorso indicato dai grandi cartelli luminosi. Dopo un'ora di visita ha gli occhi stanchi, dato che per vedere bene le opere si è dovuto sforzare sempre di trovare la giusta posizione. Si ferma al bar, dove trova un buffet fornitissimo, con dei prezzi che però sono più alti che fuori.

A questo punto all'intervistato veniva chiesto di esprimere un voto alla situazione descritta sulla usuale scala scolastica da 0 a 10 e l'uso di un sistema CATI consentiva alla fine la conferma della classifica che lo stesso intervistato aveva generato, chiedendo contemporaneamente l'eliminazione delle situazioni *ex-aequo*. I risultati di questo sistema sono stati confrontati con quelli di un campione di controllo intervistato proponendo lo stesso piano sperimentale a mezzo di un lungo, complicato e antieconomico questionario telefonico, nel quale si chiedeva di ordinare gli scenari in maniera standard⁸. Specie all'interno della segmentazione motivazionale degli Stili di Fruizione, le utilità sono risultate sorprendentemente simili fra i due modi di redazione del questionario, ma nel caso di narrazione dello scenario l'intervista durava circa 12 minuti, mentre nel secondo caso circa 25 minuti. Bisogna inoltre aggiungere che la sensibilità all'accuratezza nell'esposizione delle opere, se proposta come attributo non esemplificato concretamente, risultava nel caso standard sovrastimata dal fatto che è difficile immaginare solo a parole di aver voglia di visitare una mostra con delle opere mal disposte o poco illuminate. E proprio di questo fattore nell'indagine generale si era invece lamentato oltre il 60% dei visitatori abituali di musei o mostre.

⁷ Per una descrizione delle procedure di lemmatizzazione semantica adottate vedi Camillo, Filippucci (1996).

⁸ Il lavoro sul campione di controllo è stato oggetto di una tesi di laurea discussa presso la Facoltà di Scienze Statistiche di Bologna da Isabella Guaitoli, relatore Carlo Filippucci.

5. *Conjoint analysis* e impostazione tipologica : alcune conclusioni

L'integrazione di una tecnica come la *conjoint analysis* con quanto proposto negli ultimi anni dall'approccio tipologico "alla francese" si basa quindi su due punti principali. Il primo è che i soggetti indagati giocano un ruolo fondamentale come unità statistiche sia nell'approccio tipologico che nella *conjoint analysis*. Il secondo, forse più importante, è che la natura esplorativa della ricerca motivazionale della scuola francese riesce a fornire, grazie ad alcuni accorgimenti particolari come la "proiezione supplementare", una segmentazione la cui interpretazione è di fatto già contenuta nei risultati. La segmentazione del mercato costituisce uno scenario ideale per una *conjoint analysis* contemporanea o successiva, dato che quest'ultima fornisce degli strumenti operativi di quantificazione del comportamento del consumatore, come misure di elasticità parziali o simulazioni della soddisfazione potenzialmente ottenibile da un bene non ancora esistente sul mercato.

L'integrazione della *conjoint analysis* con una strategia di ricerca più generale di tipo esplorativo "alla francese" può apportare miglioramenti all'interpretabilità dei suoi risultati. Infatti, l'indagine preliminare esplorativa, che definisce le tipologie e che quindi segmenta il mercato secondo le motivazioni o i valori, potrà fornire un quadro informativo diverso a seconda della tipologia motivazionale, ossia a seconda del target a cui ci si riferisce. E' possibile che proprio le differenziazioni esistenti fra le motivazioni o i valori presenti sul mercato costituiscano la causa delle eventuali difficoltà di sintesi dei risultati di una *conjoint analysis*. Uno degli strumenti operativi di valutazione dei risultati è infatti fornito dal valore medio delle utilità individuali stimate. Molte volte la loro eccessiva variabilità pone grossi problemi di generalizzazione delle stime. Il calcolo dell'utilità media per ogni attributo a partire da quelle individuali può non essere un'operazione sempre consigliabile: per un motivo filosofico generale, una sintesi così grossolana va contro lo spirito microeconomico dei modelli di analisi congiunta, e questo a maggior ragione quando gli intervistati sono soggetti estremamente diversi, che danno vita, per segmenti, a indici di variabilità delle utilità per attributo molto differenti. In tali casi vale la pena studiare ex-post le cause di tali differenziazioni e nella migliore delle ipotesi modellare le differenze tramite le variabili "segnalistiche" di tipo SED disponibili (Saporta, 1992).

Un esempio interessante di applicazione di analisi tipologiche effettuate sui risultati di una *conjoint analysis* a proposito del *packaging* di prodotti per la casa è riportato in Morineau (1994bis). In tale studio l'approccio tipico dell'analisi dei dati viene applicato alla matrice delle utilità individuali per attributo e la segmentazione è stata effettuata ex-post, adottando una classica diagonalizzazione in componenti principali e un successivo *clustering*. Il mantenimento in illustrativo delle caratteristiche SED degli intervistati ha poi consentito di riconoscere operativamente i diversi target di consumatori.

In una recente proposta si adotta in maniera innovativa una classificazione di tipo *fuzzy* eseguita a catena su un'analisi fattoriale non simmetrica - analisi in componenti principali in rapporto a un sottospazio di riferimento (D'Ambra e Lauro, 1982) - di una matrice di preferenze. La classificazione avviene sulle variabili-preferenza ottimizzando un particolare criterio discriminante. L'interesse operativo per questa strategia risiede nella possibilità di ottenere rappresentazioni grafiche in cui ciascun punto, su un piano cartesiano, è individuato dal grado di appartenenza dei livelli degli attributi a delle classi di intervistati (target): punti posizionati lungo la direttrice degli assi indicano livelli "tipici" dei segmenti di mercato, mentre punti più accentrati possono identificare nuove potenziali nicchie di mercato (Giordano e Verde, 1995). In tale proposta si tiene conto in modo originale sia delle preferenze espresse, sia delle caratteristiche "segnalistiche" dei consumatori intervistati.

L'altra strategia - di conduzione di più *conjoint analysis*, ciascuna relativa ad ogni segmento motivazionale o valoriale - consente da un lato di evidenziare le caratteristiche importanti del prodotto studiato e le loro interrelazioni, ma dall'altro permette di impostare l'analisi congiunta già su segmenti omogenei al loro interno. L'effettuazione in filiera di una *conjoint analysis*, differenziata a seconda dei segmenti individuati, può consentire inoltre una migliore predisposizione del questionario, soprattutto relativamente alla definizione

semantico-valoriale degli attributi rilevanti per l'utilità globale del prodotto. L'adozione di una strategia "in filiera" consentirebbe di usare l'analisi delle eventuali risposte aperte di natura evocativa che, come ricordato, entrano sempre più spesso a far parte del questionario di base dell'approccio tipologico. Come visto nel caso della ricerca "Gli italiani e l'arte", un'analisi dei campi semantici e di senso dei testi, che mescoli opportunamente tecniche di statistica testuale e modelli interpretativi della semiotica moderna, giova all'impostazione dell'indagine, all'interpretazione dei risultati e alla loro comunicazione.

Riferimenti bibliografici

- ACA System, (1994-95), *app.G, version 4.0*, Sawtooth Software, Evanston, IL
- Agha A.K. (1991), *Régression et analyse canonique sous contraintes linéaires, algorithmes et applications*, Université Paris IX Dauphine, UER Sciences des Organisations
- ArtMonitor (1996), I sette stili di fruizione dei musei, in *SZ*, n.4, pg. 65-69, SanZanobi Editore, Bologna
- Brasini S., Tassinari F., Tassinari G. (1995), *Marketing e pubblicità*, Il Mulino, Bologna
- Camillo F. (1995), Risposte aperte, statistica testuale e approccio semiotico: una ricerca di integrazione per la comunicazione pubblicitaria; *Atti di Jact95* (eds. S.Bolasco, L.Lebart, A.Salem), vol.2, Cisu edizioni, Roma
- Camillo F., Filippucci C. (1996), Statistical Analysis of Textual Information Stemming from Focus Group, Atti dell'International Conference on Survey Measurement and Process Quality, Bristol (UK) 1-4 aprile 1995, ASA, 1996
- Carrol J.D. (1972), Individual differences and Multidimensional scaling, in *Multidimensional Scaling: Theory and Applications in the Behavioural Sciences* (vol.1), New York, Seminar Press, pg. 105-155
- D'Ambra L., Lauro N.C. (1982), Analisi in Componenti Principali in Rapporto ad un Sottospazio di Riferimento, in *Rivista di Statistica Applicata*, n.15
- Darmon R. (1992), L'établissement des quotas de vente par l'analyse conjointe, in *L'analyse conjointe et le produit idéal* (AA.VV.), Cisia, Paris
- Evrard Y., Pras B., Roux E. (1993), *Market, Études et recherches de marketing*, Nathan, Parigi
- Floch J.M. (1992), *Semiotica, marketing e comunicazione*, F. Angeli, Milano
- Giordano G., Verde R. (1995), Cluster Conjoint Analysis, in *Riassunti delle Giornate di Analisi dei Dati Multidimensionali*, Napoli, 30-31 ottobre 1995, DMS, Rocco Curto Editore
- Grandi R. (1994), *Semiotica al marketing*, F. Angeli, Milano
- Green P. (1984), Hybrid Models for Conjoint Analysis: an Expository Review, in *JMR*, 21, pg. 155-159
- Green P., Krieger Abba M., Argawal Manoj K. (1991), Adaptive Conjoint Analysis: some caveats and suggestions, in *JMR*, 28, pg. 215-222
- Green P., Rao R.V. (1971), Conjoint Measurement for Quantifying Judgmental Data, in *JMR*, n°5
- Green P., Srinivasan V. (1978), Conjoint analysis in Consumer Research: Issues and Outlook, in *Journal of Consumer Research*, 5, pg. 103-123
- Green P., Srinivasan V. (1990), Conjoint analysis in Marketing: New Developments with Implications for Research and Practice, *Journal of Marketing*, 54, October, pg. 3-19
- Johnson R.M. (1974), Trade off analysis of consumer values, in *JMR*, n°11, pg. 121-127
- Johnson R.M. (1987), Adaptive Conjoint Analysis, in *Sawtooth Conference on Perceptual Mapping, Conjoint Analysis and Computer Interviewing*, Ketchum, Sawtooth Software, pg. 253-265
- Kruskal J.B. (1964), Multidimensional scaling by Optimizing Goodness-of-Fit to a Nonmetric Hypothesis, in *Psychometrika*, 29, pg. 1-28 e 115-129
- Kruskal J.B., Carmone F.J. (1968), *MONANOVA: a Fortran IV program for Monotone Analysis of Variance*, Bell Telephone Laboratories, Murray-Hill, N.J.

- Ladwein R. (1996), *Les Études Marketing*, Economica, Parigi
- Lebart L. (1986), Qui pense quoi ? in *Consommation Revue de Socio-Economie*, Dunod, 4, pg. 3-22
- Lebart L. (1996), Questions ouvertes et questions fermées dans les enquêtes de satisfaction, Atti del "Colloque Isup : La satisfaction des produits et des services", ASU-ISUP, Paris
- Lebart L. Morineau A., Piron M. (1995), *Statistique exploratoire multidimensionnelle*, Dunod, Parigi
- Lebart L., Salem A. (1994), *Statistique Textuelle*, Dunod, Parigi
- Luce R. D., Tukey J. (1964), Simultaneous Conjoint Measurement : a new type of fundamental measurement, in *Journal of Mathematical Psychology*, 1, 1-27.
- Mitchell T.J. (1974), An Algorithm for the Construction of D-optimal Experimental Designs, in *Technometrics*, 20, pg. 203-210
- Morineau A. (1994), Le "themascope" ou analyse structurelle des données d'enquête, in *Traitements Statistique des Enquêtes*, Dunod, Parigi.
- Morineau A. (1994bis), Aplicacion del Analisis Conjunto al estudio del embalaje de un producto domestico, in *MARKSTAT '94* (ed. T. Aluja Banet), UPC, Barcellona
- Rumiati R. (1990), *Giudizio e decisione : teoria e applicazione della psicologia della decisione*, Il Mulino, Bologna
- Saporta G. (1992), Méthodologie de l'analyse conjointe et construction des produits par plans d'expérience, in *L'analyse conjointe et le produit idéal* (AA.VV.), Cisia, Paris
- SAS-QC Software (1994), *ADX Menu System for Design of Experiments*, Cary, Sas Institute
- Sawtooth Software (1995), *The CBC System for Choice-Based Conjoint Analysis*, Sawtooth Software, Washington
- Semprini A. (1992), Introduzione all'edizione italiana di *Semiotica, marketing e comunicazione*, di J.M. Floch, F. Angeli, Milano
- Smith W. (1956), Product Differentiation and Market Segmentation as Alternative Marketing Strategies, in *Journal of Marketing*, 21, pg. 3-8
- Tenenhaus M. (1988), Canonical Analysis of two Polyedral Convex cones, in *Psychometrica*, vol.53 n.4, pg. 503-524
- Torgerson, W.S. (1958), *Theory and method of scaling*, New York, Wiley
- van der Lans I. (1992), *Nonlinear Multivariate Analysis for Multiattribute Preference Data*, DSWO Leiden University, MT Series 22
- Umiker-Sebeok J. (1992), Meaning construction in a cultural gallery : a sociosemiotic study of consumption experiences in a museum, in *Advances in Consumer Research XIX*

Summary

Classificazione, analisi dei dati e modelli di *conjoint analysis* come strumenti integrati per le ricerche di mercato

Classification, Data Analysis and Conjoint Analysis Models : integrated tools for the marketing research

The work aim is the study of the main methodological topics related to the Conjoint Analysis (CA). This goal is reached by the survey of different literature contributes about this matter. In meantime hereby we propose the French explorative approach to multidimensional data integrated and completed with CA models for the marketing research.