INTRODUZIONE AI METODI STATISTICI PER IL MARKETING:

LA CONJOINT ANALYSIS

A cura:

del Centro per la Modellistica, il Calcolo e la Statistica

(CMCS)

Arboretti
Boari
Salmaso

Facoltà di Economia

Università degli Studi di Ferrara

Documento preliminare - Non divulgabile

1 INTRODUZIONE

L'impiego di metodi statistici per la valutazione e il miglioramento della qualità dei sistemi di produzione di beni e servizi ha visto negli ultimi anni un consolidamento crescente come parte integrante dei sistemi di gestione aziendali. In un mercato sempre più complesso caratterizzato da elevata competitività, alto sviluppo tecnologico, volumi di produzione più elevati, maggiori e più differenziate esigenze individuali dei consumatori, l'orientamento aziendale alla qualità diventa sempre più una strategia importante per aumentare la produttività e la penetrazione nel mercato, per raggiungere un forte vantaggio competitivo.

In tale contesto, il concetto di qualità, da un significato iniziale di conformità del prodotto a specifiche prettamente di tipo tecnico imposte internamente dall'azienda, ha gradatamente assunto un significato più ampio investendo tutti gli aspetti dell'attività aziendale (formazione del personale, progettazione, scelte economiche), fino a coinvolgere aspetti esterni all'azienda quali le esigenze e i bisogni dei clienti, assumendo quindi anche un significato di capacità di soddisfare le aspettative dei consumatori.

L'evoluzione del concetto di qualità di un bene o servizio ha posto quindi l'esigenza di affiancare ai metodi tradizionali del controllo statistico della qualità, altre metodologie statistiche tra cui i metodi per la valutazione e l'interpretazione di come i consumatori percepiscono e valutano la qualità di prodotti o servizi indicata con il termine di "customer satisfaction". Tra i contributi metodologici più innovativi in questo ambito, si situa la Conjoint Analysis (CA) finalizzata allo studio dei modelli di scelta dei consumatori a partire dai giudizi di preferenza espressi da questi ultimi relativamente a diverse possibili configurazioni dei prodotti o servizi.

2 LA CONJOINT ANALYSIS: LA SUA STORIA ED I SUOI SCOPI

Il nome "Conjoint Analysis" fu usato per la prima volta da Green e Srinivasan nel 1978 per riferirsi ad un gruppo di paradigmi che descrivevano quantitativamente le preferenze dei consumatori o le scelte di valore ("modelli e tecniche che enfatizzano la trasformazione di risposte soggettive in parametri stimati"). In realtà già nel '64 Luce e Tukey (come riferito da Vriens, '94) svilupparono una metodologia di "misura congiunta" per analizzare le preferenze dei consumatori, introdotta nell'ambito del marketing (e quindi delle preferenze di prodotto) nel '71 da Green e Rao (come riferito da Green e Srinivasan, '78).

Green e Srinivasan ('90) definiscono conjoint analysis "ogni metodo decompositivo che stimi la struttura di preferenza di un consumatore (ossia i parametri di preferenza come i pesi relativi, i pesi di importanza, i punti ideali), dato il giudizio globale del cliente su un insieme di alternative".

Nella letteratura accademica la Conjoint Analysis è vista sia come supporto per problemi metodologici (es. metodi di stima, procedure per conoscere le

preferenze...), sia come metodo applicativo (statistico). C'è poi un'attenzione "commerciale": Wittink e Cattin (1982) hanno studiato la conjoint analysis e i suoi usi commerciali negli USA, Wittink, Vriens e Burhenne ('94) in Europa (Vriens, '94). Questi ultimi, in particolare, hanno sottolineato l'importanza crescente del metodo come supporto nel processo di sviluppo di nuovi prodotti: in una conferenza sul ruolo delle ricerche per migliorare lo sviluppo del prodotto (1990), la conjoint analysis era considerata seconda solo ai focus group. Nella loro analisi, Wittink, Vriens e Burhenne hanno trovato che gli scopi principali di una conjoint analysis (in Europa, anni '86-'91) erano relativi a decisioni di prezzo e allo sviluppo di nuovi prodotti (concept testing in particolare); altro scopo importante era quello della segmentazione del mercato. In genere l'analisi veniva svolta su beni di consumo (54%), mentre un 17% delle ricerche riguardava beni industriali (le percentuali erano molto simili a quelle riscontrate per le applicazioni negli USA tra l'81 e l'85).

Gli usi della conjoint analysis possono essere classificati sia basandosi sullo scopo da perseguire con l'analisi stessa, sia considerando i tipi di prodotti cui ci si riferisce (Vriens, '94).

Rispetto allo scopo della ricerca si distinguono:

1. Segmentazione del mercato. Ci sono forme diverse di segmentazione a seconda della base usata per suddividere il mercato: in genere si guarda ai benefici che il prodotto offre ai clienti e a come questi ultimi li colgano (approccio diretto e orientato al consumatore). I giudizi (pesi) ottenuti con la conjoint analysis possono servire come rappresentazioni numeriche dei benefici che il prodotto offre secondo i clienti.

In genere si procede in due fasi: prima si calcolano i pesi per ogni intervistato e poi si formano dei cluster (gruppi) usando degli algoritmi specifici. Ogni gruppo di clienti individuato sulla base della similarità dei benefici (pesi) è un segmento del mercato: si può così osservare se eventualmente la linea di prodotti di un'azienda copra o meno tutti i segmenti; può nascere l'idea di un nuovo prodotto che soddisfi specifiche necessità; si può formulare una campagna pubblicitaria più mirata per un prodotto esistente; etc.

2. Decisioni relative al prodotto. La conjoint analysis è utile per ottenere informazioni di supporto a decisioni del tipo:
   - modifiche ad un prodotto (es. cambiamento di 1 o 2 attributi)
   - progettazione ottimale di un prodotto (si considerano tutti gli attributi)
   - riprogettazione della linea
   - valutazione dei concetti di un nuovo prodotto.

3. Analisi competitiva. E' un aspetto molto legato al precedente: si utilizzano modelli di simulazione basati su dati ottenuti grazie ad una conjoint analysis, che permettono di prevedere per esempio i cambiamenti nella quota di mercato derivanti da variazioni apportate a prodotti o a linee di prodotto.

4. Decisioni sul prezzo. Grazie alla conjoint analysis, in particolare usando metodi non tradizionali, nuovi, è possibile ottenere informazioni sull'elasticità della domanda rispetto al prezzo, soprattutto per supportare decisioni sul mar-

keting mix migliore per un nuovo prodotto (laddove non ci sono dati storici che permettano per esempio un'analisi di regressione).

5. Decisioni sulla promozione e sulla distribuzione. Non ci sono molti esempi di utilizzo della conjoint analysis direttamente in queste aree. Eppure (magari indirettamente, o a margine C.A. mirate ad investigare magari altri aspetti) si possono ricavare informazioni molto utili anche rispetto a questi tipi di decisioni: si può condurre un'analisi che utilizzi come attributi variabili relative agli aspetti promozionali (pubblicità, livello di servizio ai distributori, condizioni di pagamento...), rilevando i giudizi che i distributori danno a certe combinazioni; si possono introdurre elementi relativi alla distribuzione, come per esempio "tempo di trasporto", o la variabile "disponibilità del prodotto".

Riferendosi ai prodotti su cui si può svolgere l'analisi, si può dire che siano state effettuate applicazioni di Conjoint Analysis praticamente per ogni tipo di prodotto:

Tab. 1: Tipologie di prodotti su cui sono state svolte applicazioni di C.A.

Tipo di prodotto	Esempi
beni di consumo "veloce"	alimentari, shampoo...
beni di consumo di acquisto frequente	scarpe, vestiario...
beni di consumo durevole	automobile, televisore..
prodotti industriali	sistemi informatici, ascensori...

3 SEGMENTAZIONE DEL MERCATO COME OBIETTIVO DI CONJOINT ANALYSIS

3.1 LA SEGMENTAZIONE DEL MERCATO

Per definire gli obiettivi e le modalità di una strategia commerciale un'azienda deve essere in grado di valutare le caratteristiche, i bisogni ed i comportamenti probabili degli acquirenti potenziale, di conseguenza, di delimitare il mercato in cui intende agire. Questa decisione comporta una suddivisione, o segmentazione del mercato in gruppi omogenei e distinti di consumatori. la segmentazione di un mercato consiste, quindi, nell'adeguare sia i prodotti che le strategie di marketing alle differenze individuabili entro l'insieme delle esigenze manifestate dai consumatori e/o utilizzatori.

La teoria economica del comportamento del consumatore considera il consumo come un'attività attraverso la quale i prodotti scelti sono utilizzati, da soli o in combinazione fra loro, al fine di produrre una serie di servizi che determinano una utilità per l'individuo stesso. In questo senso i prodotti sono configurabili come insiemi di caratteristiche o panieri di attributi, ed i consumatori come produttori di soddisfazioni.

Adottare una strategia di segmentazione del mercato significa, in altri termini, riconoscere esplicitamente l'esistenza di elementi di eterogeneità entro un mercato. Un'azienda non sempre si trova nella condizione di poter produrre e vendere un solo prodotto/servizio con un'offerta in grado di raggiungere il maggior numero possibile di consumatori, ma deve, più frequentemente, individuare e soddisfare domande diverse, una per ogni segmento di mercato servito. D'altra parte il consumatore singolarmente considerato rappresenta un segmento di mercato potenziale a se stante, in quanto portatore di una combinazione di bisogni e di desideri unica. Nel mercato dei beni di largo consumo esistono, da questo punto di vista, limiti oltre i quali il processo di segmentazione non può spingersi, al contrario di quanto accade nei mercati dei beni industriali e di investimento dove per una azienda può essere talvolta conveniente esaminare a fondo la situazione dei singoli clienti potenziali e mettere a punto per ciascuno di essi un'offerta specifica.

3.1.1 Requisiti della Segmentazione

L'efficacie attuazione delle politiche di marketing esige che ogni segmento di mercato cui l'azienda si rivolge presenti particolari connotati. I requisiti essenziali riguardano:

1) La tendenziale uniformità delle risposte degli acquirenti potenziali rispetto agli strumenti tipici del marketing-mix, come la qualità del prodotto/servizio, la pubblicità e le attività promozionali, i prezzi, il tipo di confezione o packaging, la distribuzione, i servizi e l'assistenza post-vendita, ecc...;

2) La profittabilità, nel senso che la dimensione del segmento o il suo livello di assorbimento devono essere tali da garantire un profitto;

3) L'accessibilità, ossia la possibilità per l'azienda di raggiungere il segmento di interesse tramite gli strumenti di marketing attualmente disponibili, con costi addizionali o perdite minimi.

3.1.2 Le fasi operative

Quando l'azienda desidera conoscere in profondità ol mercato nel quale è attiva o intende iniziare a competere può ricorrere utilmente ad uno studio di segmentazione. Ciò implica che si faccia riferimento sempre, ad ogni livello strategico, al concetto di segmento; così, ad esempio, conviene verificare la perfomance dei prodotti dell'azienda, in termini di quota del mercato, di tasso di crescita del volume delle vendite, di redditività, ecc..., non solo per l'intero mercato, ma anche per i singoli segmenti.

Da un punto di vista operativo, la segmentazione del mercato passa attraverso l'attuazione di alcune fasi che si possono compendiare nei seguenti punti:

1) definizione del problema e selezione della procedura di segmentazione;
2) messa a punto del programma dell'indagine sul campo;
3) elaborazioni, interpretazione de impiego dei risultati.

1) La selezione della procedura di segmentazione

Per la prima fase vi sono diverse soluzioni alternative che sono rappresentate dal uno dei seguenti modelli di segmentazione:

- a priori;
- a posteriori;
- flessibile.

- Nei modelli di segmentazioni a priori si procede alla suddivisione del collettivo in esame a seconda delle modalità presentate da una o più variabili o basi (dove per base si intende una variabile a disposizione che genera direttamente il processo di classificazione in gruppi delle unità statistiche in osservazione. Possono esserci una o più basi), specificate, appunto, a priori. Un esempio di questo tipo è la tradizionale segmentazione geografica dei consumatori presenti all'interno di un mercato nazionale, quando si assume quale base la regione di residenza, si determinano automaticamente venti segmenti, dei quali dovrà essere valutata esclusivamente l'ampiezza relativa. In questo caso, quindi, l'attività di segmentazione si riduce ad una semplice classificazione delle unità statistiche in classi preventivamente definite, sia per numero sia per tipologia.

- I modelli di segmentazione a posteriori, che si basano su algoritmi di raggruppamento (clustering), si differenziano dai precedenti solo in relazione al modo in cui viene selezionata la base di segmentazione. In questo caso manca una scala a priori, e non sono prefissati né il numero né le tipologie dei gruppi da formare. Il termine a posteriori evidenzia il fatto che i segmenti sono determinati attraverso la classificazione delle unità statistiche a seguito di una Cluster Analysis, cioè a partire dal grado di dissomiglianza rispetto ad un insieme prescelto di variabili (per dissomiglianza si intende il grado di discordanza tra profili individuali definiti sulla base dell'osservazione di più caratteri qualitativi e/o quantitativi. Ad esempio, la dissomiglianza tra due consumatori rispetto a p variabili quantitative può essere valutata mediante una misura della distanza tra i due punti che li rappresentano nello spazio a p dimensioni generate dalle p variabili.

- Con i modelli di segmentazione flessibile è possibile individuare un numero anche elevato di segmenti alternativi, ciascuno costituito dall'insieme dei consumatori con profili di risposta simili in termini di preferenza per prodotti o marche già esistenti, oppure potenzialmente introducibili sul mercato (realizzabili quindi in via ipotetica, sotto particolari condizioni ambientali e competitive). Tali prodotti sono definiti in base a specifici profili derivanti dalla combinazione di caratteristiche distintive quali design, prezzo, condizioni di garanzia e di assistenza post-vendita, ecc...L'approccio della segmentazione flessibile si fonda sull'integrazione dei risultati di uno studio che si avvale dell'applicazione di una tecnica di analisi statistica multivariata nota come Conjoint Analysis e di una simulazione sul comportamento di scelta dei consumatori.

2-3) Organizzazione del piano di indagine e lettura dei risultati

4 LA METODOLOGIA

I modelli di segmentazione flessibile si propongono di identificare segmenti di mercato costituiti da insiemi di clienti con profili di risposta simili in termini di gradimento e preferenza nei confronti di prodotti/servizi già esistenti sul mercato oppure in fase di progettazione. I prodotti/servizi sono descritti in base a specifiche combinazioni di modalità o di livelli di caratteristiche distintive (chiamate anche attributi), quali il design, il prezzo, le condizioni di garanzia, ecc... Lo schema della segmentazione flessibile si fonda sull'integrazione dei risultati della Conjoint Analysis e di quelli relativi ad una simulazione del comportamento di scelta dei consumatori.

I fondamenti concettuali della C.A. risiedono nelle considerazioni sviluppate da Lancaster in margine alla teoria del consumatore secondo le quali l'utilità d'uso di un bene deriva dalle singole caratteristiche che lo compongono. Dunque, risulta possibile scomporre l'utilità che un consumatore ricava da un prodotto/servizio in varie utilità separate che traggono origine dalle diverse caratteristiche di quel bene. Proprio in quest'ottica, che può essere denominata decompositiva, andrebbero allora a collocarsi sia il processo di formazione delle preferenze sia le vere e proprie decisioni di acquisto.

In questi ambiti, di norma, le caratteristiche di un bene sono giudicate non una alla volta, ma congiuntamente, nelle diverse modalità e/o intensità con le quali si presentano, nel senso che le possibili alternative sono prese in considerazione in base ai vantaggi ed agli svantaggi complessivi che il loro acquisto prospetta. La C.A. è stata messa a punto proprio per comprendere e misurare i diversi compromessi (trade-offs) che specifiche tipologie di consumatori compiono nel confronto e nella scelta di tali alternative.

Tramite questa tecnica è possibile valutare:

1) l'importanza che ogni individuo attribuisce a ciascuna caratteristica di un prodotto/servizio;

2) il grado di utilità corrispondente ad ogni livello o modalità di ciascuna caratteristica.

In termini operativi la sua applicazione richiede la selezione di un campione di consumatori potenziali ai quali sottoporre un elenco di versioni alternative del prodotto/servizio, dette anche stimoli, descritte sulla base delle modalità o intensità presentate da alcuni attributi ritenuti rilevanti. (tale elenco costituisce spesso solo un sottoinsieme di tutte le possibili combinazioni che le modalità di questi attributi possono generare, selezionato nel rispetto delle regole di un piano degli esperimenti). Ai consumatori contattati si chiede in genere di graduare in termini di preferenza (dal più al meno gradito) i profili di prodotto/servizio, oppure di esprimere un punteggio di gradimento per ogni alternativa su una scala prefissata, in modo tale da riflettere la probabilità di acquisto.

In sostanza, la C.A. mette in corrispondenza biunivoca le due nozioni di preferenza e di utilità. E'intuitivo pensare che quanto più uno specifico profilo incontrerà il gradimento di un consumatore tanto più la sua fruizione fornirà utilità. La preferenza può essere interpretata, a sua volta , come una funzione

delle modalità o dei livelli delle caratteristiche rilevanti del prodotto/servizio. Si tratta allora di stabilire la regola di composizione che caratterizza la formazione delle preferenze di un individuo. Sono disponibili a questo riguardo diverse alternative, ma il modello sicuramente più utilizzato è quello additivo, nel quale le utilità parziali dei singoli livelli di ogni attributo vengono sommate per ottenere l'utilità complessiva di un profilo.

Quindi, ciò che differenzia la C.A. dalle tradizionali procedure di analisi della Costumer satisfaction è che mentre queste ultime generalmente prevedono un'indagine campionaria basata su un questionario in cui ciascuna domanda mira a valutare la soddisfazione del consumatore rispetto ad una specifica caratteristica del prodotto/servizio ed a partire da queste valutazioni parziali si procede successivamente in fase di analisi dei dati raccolti alla stima della valutazione globale sull'insieme delle caratteristiche del prodotto/servizio. Al contrario la C.A. prevede un'indagine campionaria in cui si sottopongono agli intervistati diversi profili del prodotto/servizio ovvero diverse versioni di alternative di prodotto, definite in base a diverse combinazioni delle modalità di una serie di caratteristiche rilevanti del prodotto. Compito dei rispondenti è quello di assegnare a ciascun profilo il punteggio globale di gradimento o di redigere una graduatoria di preferenza degli stessi. Sulla base del punteggio globale assegnato da ciascun intervistato ai vari profili di prodotto, in fase di analisi dei dati sono stimate le preferenze parziali associate a ciascuna modalità di ogni caratteristica.

La C.A. consente quindi di affrontare lo studio della soddisfazione e del comportamento di scelta del consumatore con un'ottica in cui si ritiene che il processo di formazione individuale delle preferenze sia di tipo "multiattributivo", ossia il consumatore crea e quindi esprime una preferenza sul prodotto considerandolo nella sua interezza cioè dal punto di vista di tutte le sue caratteristiche (attributi) congiuntamente considerati.

A questo proposito sono tre i principali schemi ai quali si è soliti fare riferimento in letteratura:

-il modello vettore; -il modello punto-ideale; -il modello part-worth.	} Modelli di utilità

Il Modello Vettore

Indichiamo con il simbolo f_{jk} il livello d'intensità che l'attributo quantitativo k-esimo ($k = 1, 2, ..., K$) presenta nello stimolo j-esimo ($j = 1, 2, ..., J$) e con y_j la preferenza/utilità espressa nei confronti del medesimo stimolo, il modello vettore ipotizza che sussista la seguente relazione:

$$y_j = \sum_{k=1}^{K} w_k f_{jk} \qquad (1)$$

Nella 9 i coefficienti w_k vanno intesi come i pesi d'importanza assegnati a ciascun attributo k. Tali pesi possono essere diversi da consumatore a consumatore per tener conto della diversa struttura delle preferenze.

Il Modello Punto-Ideale

Il modello punto ideale suggerisce invece che esista per ogni consumatore un profilo ideale del bene costituito da un livello i_k per ogni attributo $(i_1, i_2, ..., i_K)$. Obiettivo del modello è quello di definire una misura di utilità che cresce al diminuire della distanza del profilo considerato da quello ideale: tanto più il prodotto/servizio è simile a quello ideale, tanto più elevata è l'utilità del consumatore. Anche in questo caso si mantiene una struttura di ponderazione per ogni attributo; una misura della distanza tra profilo considerato e profilo ideale è data da:

$$d_j^2 = \sum_{k=1}^{K} w_k \left(f_{jk} - i_k\right)^2 \tag{2}$$

Ovviamente l'utilità y_j risulta correlata negativamente con la distanza d_j^2.

Il Modello Part-worth

Questo modello esprime la preferenza/utilità y_j per lo stimolo j-esimo attraverso una funzione discontinua s definita per un insieme opportunamente selezionato di livelli degli attributi quantitativi o per le modalità degli attributi qualitativi:

$$y_j = \sum_{k=1}^{K} s_k \left(f_{jk}\right) \tag{3}$$

Nel caso degli attributi quantitativi la valutazione delle preferenze/utilità per i livelli compresi tra quelli selezionati avviene attraverso un'interpolazione lineare. Quest'ultimo modello è il più flessibile tra i tre presentati, poichè permette alla funzione di preferenza/utilità di assumere diverse forme a seconda della specificazione di s.

Per ottenere stime delle utilità parziali (part-worth utilities) riferibili alle modalità o ai livelli di ciascun attributo, determinandone di conseguenza l'utilità relativa, si assume che i consumatori, nell'esprimere i loro giudizi di preferenza, si attengano al criterio di massimizzazione dell'utilità attesa dalle varie alternative, a partire dai rispettivi vincoli di bilancio.

Questo modello presuppone un semplice schema additivo che tiene conto solo degli effetti principali (main effects) di ciascuna modalità e non anche delle possibili interazioni tra queste (cioè si ipotizza che i vari attributi siano tra loro indipendenti nella percezione dei consumatori).

5 FASI DELLA CONJOINT ANALYSIS

Il disegno e l'esecuzione di un esperimento di C.A. si articolano secondo una serie di fasi che partendo dal problema di ricerca possono essere sinteticamente schematizzate nel seguente modo:

1- Definizione dei fattori sperimentali e dei livelli:

in questa fase vengono individuate le caratteristiche del prodotto/servizio dette "attributi" o "fattori sperimentali" e per ciascun fattore vengono specificate le modalità o "livelli" di interesse. Le possibili combinazioni di livelli dei fattori definiscono i "profili" del prodotto/servizio che saranno sottoposti agli intervistati.

2- Definizione del piano fattoriale e dei profili di prodotto:

sulla base dei fattori e dei livelli individuati al punto precedente viene definito il piano fattoriale che potrà includere solo gli effetti principali dei fattori selezionati (modello additivo) o anche le interazioni tra i fattori (modello con interazione). La procedura di C.A. prevede usualmente la realizzazione di uno studio pilota con tutti i fattori su due livelli considerando un piano fattoriale completo che include le interazioni di ogni ordine. Successivamente l'analisi può essere ulteriormente riproposta considerando gli effetti principali veramente influenti sulla risposta sperimentale e le eventuali interazioni significative. Si consideri l'esempio di un'azienda interessata ad effettuare un esperimento di C.A. su un personal computer portatile. Vengono individuati tre attributi o fattori, ciascuno su due livelli:

Tab. 2: Esempi di attributi e relativi livelli

FATTORE A	LIVELLO 1: 1,5 Kg
	LIVELLO 2: 2 Kg
FATTORE B	LIVELLO 1: 2 ore
	LIVELLO 2: 4 ore
FATTORE C	LIVELLO 1: 1000 Euro
	LIVELLO 2: 2000 Euro

Indicando i livelli bassi ed i livelli alti dei tre fattori rispettivamente con "-" e "+", possiamo illustrare graficamente il piano fattoriale che è di tipo 2^3 (come base poniamo il numero dei livelli ed all'esponente il numero dei fattori) ipotizzando di essere interessati ai soli effetti principali dei tre fattori:

Tab. 3: Esempio di piano fattoriale

	FATTORI		
PROFILI	A	B	C
1) abc	+	+	+
2) ab	+	+	-
3) ac	+	-	+
4) a	+	-	-
5) bc	-	+	+
6) b	-	+	-
7) c	-	-	+
8) (1)	-	-	-

Ogni riga della matrice del piano fattoriale definisce un "profilo", ossia una combinazione dei livelli dei fattori selezionati. In corrispondenza del piano 2^3, risultano definiti otto profili, dove ciascuno di questi è stato indicato nella figura con lettere minuscole in modo tale che se una lettera è presente, allora il corrispondente valore risulta a sua volta presente in quel profilo al valore alto se, al contrario, è assente il fattore è presente al livello basso.

Il profilo in cui tutti i fattori sono presenti al livello basso è indicato con (1). Ciascun profilo corrisponde ad una diversa configurazione del prodotto/servizio. Ad esempio, al profilo 3 a "c" corrisponde un personal computer portatile con peso pari a 2 Kg, con durata della batteria pari a 2 ore e di costo pari a 2000 Euro.

In generale, dati K fattori, ciascuno caratterizzato da $P_1, P_2, ...P_k$ livelli, il numero totale di profili del prodotto/servizio sarà pari a $M = \prod P_k$.

3 - Definizione della modalità di presentazione dei profili agli intervistati:

i profili definiti al punto precedente sono sottoposti al giudizio di un campione di rispondenti detti anche "valutatori". La somministrazione dei profili ai rispondenti può essere effettuata in modo diverso. In particolare di parla di TRADITIONAL CONJOINT ANALYSIS o FULL PROFILE CONJOINT ANALYSIS quando ai valutatori sono sottoposti tutti i profili completi delle varie configurazioni del prodotto sotto forma generalmente di cartellini.

Riprendendo l'esempio precedente, i cartellini possono essere rappresentati nel modo seguente:

Tab. 4 e 5: Esempi di cartellini

CARTELLINO 1
PESO: 2 Kg
DURATA: 4 ore
PREZZO: 2000 Euro

CARTELLINO 8
PESO: 1,5 Kg
DURATA: 2 ore
PREZZO: 1000 Euro

I rispondenti daranno un giudizio in forma di punteggio o graduatoria relativamente all'insieme completo dei profili. Si parla invece di ADAPTIVE CONJOINT ANALYSIS quando, a differenza dell'approccio tradizionale, nella presentazione a coppie di profili successivi ai rispondenti, si tiene conto delle preferenze precedentemente ottenute, "adattando" di conseguenza i confronti successivi.

Nel seguito si farà riferimento esclusivamente alla tipologia FULL PROFILE.

4 - Definizione della risposta sperimentale:

la preferenza dei rispondenti sui vari profili può essere espressa mediante punteggio (RATING) su una scala metrica, ad esempio da 1 a 10, oppure mediante una graduatoria di preferenza dei profili (RANKING). Nel primo caso la variabile risposta è considerata continua in quanto rappresenta il punteggio espresso su scala metrica e si parla in questo caso di CONJOINT ANALYSIS METRICA, nel secondo caso la variabile risposta è di tipo ordinale in quanto rappresenta la posizione di graduatoria e si parla di CONJOINT ANALYSIS NON METRICA. Nel seguito si farà riferimento alla C.A.METRICA.

5 - Stima delle preferenze parziali dei livelli dei fattori e dell'importanza relativa di ciascun fattore:

dopo aver somministrato i profili di prodotto alle unità sperimentali (soggetti rispondenti), vengono stimate le preferenze parziali di ciascun livello dei fattori del prodotto e una misura dell'importanza di ciascun fattore per ogni rispondente. Nella C.A.Metrica la stima può avvenire ipotizzando per la variabile risposta un modello di regressione lineare multipla e stimando i parametri di tale modello con il metodo dei minimi quadrati.

6 - Verifica di ipotesi sugli effetti dei fattori del prodotto/servizio:

l'ultima fase della C.A. è relativa ai metodi di determinazione degli effetti maggiormente influenti sulla risposta sperimentale. Le procedure utilizzate in questa fase possono essere classificate principalmente secondo due diverse tipologie:

- Nel caso in cui il piano degli esperimenti che ha prodotto l'indagine di C.A. sia di tipo non replicato (ovvero nel caso di un singolo rispondente o nel caso di aggregazioni di giudizi di preferenza tramite la preferenza media) non essendovi gradi di libertà nella stima della varianza dell'errore sperimentale non è possibile condurre test di tipo parametrico. Quindi un metodo proposto in letteratura è quello introdotto da Daniel nel 1959 denominato NORMAL PROBABILITY PLOT che è un metodo grafico per la determinazione degli effetti che si discostano in modo sufficientemente evidente dal grafico di una distribuzione normale di media nulla.

- Nel caso in cui il piano sperimentale sia di tipo replicato, ovvero i rispondenti possono essere ragionevolmente identificati come appartenenti ad un gruppo omogeneo dal punto di vista delle variabili confondenti che potrebbero influire sul giudizio di preferenza, si può adattare la usuale procedura di ANALISI DELLA VARIANZA per piani di tipo multifattoriale.

6 SPECIFICAZIONE E STIME DEL MODELLO DI UTILITA′ INDIVIDUALE

La C.A. pone in corrispondenza biunivoca il concetto di "preferenza" con quello di "utilità", nel senso che il profilo di prodotto preferito è anche quello dal quale il rispondente ricava maggiore utilità. La preferenza di un acquirente è considerata funzione delle modalità (o livelli) degli attributi (o fattori) rilevanti del prodotto/servizio in studio. Tale funzione, come visto in precedenza, viene denominata "funzione di utilità individuale". Tra i vari modelli di utilità individuale proposti in letteratura vi è il modello "additivo a coefficienti separati" secondo il quale l′utilità globale di un generico profilo di prodotto è data dalla somma delle utilità parziali di ogni livello dei diversi attributi che definisco il prodotto.

Dati K fattori, ciascuno caratterizzato da $P_1, P_2, ...P_k$ livelli sia $M = \prod P_k$ il numero dei profili del prodotto. Per ciascun profilo costruiamo una variabile binaria come segue:

$$d_{mkp} = \begin{cases} 1 & \text{se il profilo } m \text{ presenta l'attributo } k \text{ con il livello } p \\ 0 & \text{altrimenti} \end{cases}$$

dove m indica il generico profilo, con $m = 1, ..., M$; k indica il generico fattore, con $k = 1, ..., K$; p indica il generico livello del fattore k, con $p = 1, ..., P_k$.

La preferenza o utilità globale per il profilo m, con riferimento ad un generico rispondente, può essere espressa secondo il seguente modello:

$$Y_m = \sum_{k=1}^{K} \sum_{p=1}^{P_k} w_{kp} d_{mkp} + e_m \qquad (4)$$

dove w_{kp} è un coefficiente che esprime l′importanza attribuita da un generico individuo al fattore k considerato al livello p. Tale coefficiente viene detto "utilità parziale" o "part-worth", come indicato in precedenza. L′insieme dei coefficienti w_{kp}, con $p = 1, ..., P_k$, $k = 1, ..., K$, rappresenta quindi il sistema individuale di preferenze per ilprofilo m. Le componenti casuali e_m, con $m = 1, ..., M$, rappresentano le componenti di errore del modello, dovute all′eventuale influenza di fattori noti o non noti controllati in sede di esperimento, che possono influire sulla preferenza y_m espressa dal generico rispondente del profilo m. Le componenti casuali e_m sono poste indipendenti de identicamente distribuite con media nulla e varianza costante σ^2. Il modello di utilità individuale si configura quindi come unmodello di regressione lineare multipla con covariate d_{mkp} binarie. Si asserisce che nel modello 4 non sono presenti termini di interazione corrispondenti a combinazioni di livelli; è possibile comunque prevedere modelli che comprendono anche interazioni.

Possiamo esprimere il modello 4 in forma matriciale nel seguente modo:

$$Y = ZW + e \tag{5}$$

con:

$E(e) = 0$
$Var(e) = \sigma_e^2 I_M$ con I_M: matrice identità di ordine M

Dove Y è il vettore colonna di dimensione $M \times 1$ dei punteggi di valutazione osservati per un dato rispondente; Z è la matrice $M \times P$, dove $P = \sum_{k=1}^{K} P_k$ è il numero totale dei livelli dei fattori, ed è la matrice delle variabili d_{mkp}; W è il vettore colonna $P \times 1$ dei coefficienti incogniti che rappresenta le utilità parziali ed e è il vettore colonna $M \times 1$ degli errori e_m.

Tornando all'esempio introdotto precedentemente e considerando i punteggi di preferenza espressi da un dato rispondente il modello 5 sarà del tipo:

$$\begin{array}{lcccccccc}
y_1 = 5,2 & & 10 & 10 & 10 & & & & e_1 \\
y_2 = 7,3 & & 10 & 10 & 01 & & w_{11} & & e_2 \\
y_3 & & 10 & 01 & 10 & & w_{12} & & e_3 \\
y_4 & = & 10 & 01 & 01 & \times & w_{21} & + & e_4 \\
y_5 & & 01 & 10 & 10 & & w_{22} & & e_5 \\
y_6 & & 01 & 10 & 01 & & w_{31} & & e_6 \\
y_7 & & 01 & 01 & 10 & & w_{32} & & e_7 \\
y_8 = 5,7 & & 01 & 01 & 01 & & & & e_8
\end{array}$$

La stima dei coefficienti incogniti W può essere ottenuta applicando il metodo dei minimi quadrati ordinari. Risultando però la matrice Z singolare, per poter ottenere le soluzioni ai minimi quadrati si deve porre una condizione complementare che può consistere, ad esempio, nel sopprimere, con scelta arbitraria, una colonna in ciascun blocco di variabili binarie della matrice Z, ponendo di conseguenza pari a 0 il coefficiente corrispondente. Nella matrice Z viene quindi soppresso un numero di colonne pari al numero di fattori.

Tornando all'esempio precedente il modello diventa:

$$\begin{array}{lcccccccc}
y_1 = 5,2 & & 1 & 1 & 1 & & & & e_1 \\
y_2 = 7,3 & & 1 & 1 & 0 & & & & e_2 \\
y_3 & & 1 & 0 & 1 & & w_{11} & & e_3 \\
y_4 & = & 1 & 0 & 0 & \times & w_{21} & + & e_4 \\
y_5 & & 0 & 1 & 1 & & w_{31} & & e_5 \\
y_6 & & 0 & 1 & 0 & & & & e_6 \\
y_7 & & 0 & 0 & 1 & & & & e_7 \\
y_8 = 5,7 & & 0 & 0 & 0 & & & & e_8
\end{array}$$

Verranno quindi stimati i coefficienti w_{11}, w_{21}, w_{31} mentre sono pari a 0 i coefficienti di w_{12}, w_{22}, w_{32}. In generale quindi con tale operazione il modello di regressione diventa:

$$Y = Z_0 W_0 + e \tag{6}$$

dove Y è il vettore colonna $M \times 1$, Z_0 è la nuova matrice non singolare $M \times (P - K)$; W_0 è il vettore colonna dei coefficienti $(P - K) \times 1$ ed e è il vettore colonna degli errori, $M \times 1$.

Il vettore delle stime dei coefficienti incogniti secondo il metodo dei minimi quadrati sarà:

$$W_0 = (Z_0' Z_0)^{-1} Z_0' Y \tag{7}$$

Si osservi che il modello 5 di utilità individuale viene stimato per ciascun rispondente. Gli elementi della matrice Z_0 sono fissi per ciascun rispondente mentre varierà il vettore Y dei punteggi di valutazione. Per ciascun rispondente verrà quindi stimato il vettore W_0 delle utilità parziali di ciascun livello dei diversi fattori. Tornando all'esempio precedente supponiamo che per un dato rispondente siano state stimate le seguenti utilità parziali per il fattore peso:

Tab. 6: Esempio di utilità parziali stimate.

FATTORE	LIVELLI	COEFFICIENTE DI UTILITA' PARZIALE
PESO	1 Kg	2,37
	2 Kg	0,00

Si osservi che avendo soppresso nella matrice Z l'ultimo livello dei fattori, questo rappresenta il livello di riferimento rispetto al quale sono misurati gli effetti dei restanti livelli.

Possiamo rappresentare graficamente i coefficienti di utilità parziale stimati per un rispondente nel modo seguente:

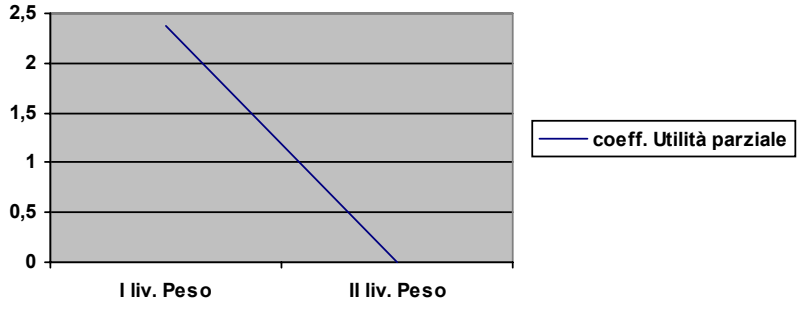


Fig. 1:

Nelle applicazioni di C.A. è inoltre consuetudine valutare l'importanza relativa dei fattori (allo scopo di rendere compatibili le utilità parziali e di pervenire a valori di importanza dei fattori compresi tra 0 ed 1) mediante il seguente indice:

$$I_{rk} = \Delta_k / \sum_{k=1}^{K} \Delta_k \qquad con \;\; k = 1, 2, ..., K \tag{8}$$

dove Δ_k è la differenza o campo di variazione tra l'utilità del fattore k. L'indice I_{rk} è tale che $0 \leq I_{rk} \leq 1$. L'importanza di un fattore è indicata quindi dalla capacità del fattore di determinare il maggior effetto differenziale nelle utilità parziali.

7 ANALISI GRAFICA DELL'INFLUENZA DEGLI EFFETTI DEI FATTORI SULA PREFERENZA DEI RISPONDENTI

Abbiamo visto precedentemente che il modello di utilità individuale viene stimato per ciascun rispondente dando quindi luogo ai vettori W0 delle utilità parziali dei livelli dei vari fattori inclusi nell'esperimento quanti sono i rispondenti. Supponiamo che sia d'interesse, in particolare, una valutazione degli effetti dei vari fattori sui giudizi medi espressi dai rispondenti. Riprendiamo l'esempio introdotto precedentemente e supponiamo di considerare oltre agli effetti principali dei tre fattori A, B, C anche le interazioni fino all'ordine massimo, per cui la matrice del piano fattoriale completo 23 è del tipo:

Tab. 7: Esempio di piano fattoriale completo

	F	A	T	T	O	R	I
Profili	A	B	C	AB	AC	BC	ABC
1) abc	+	+	+	+	+	+	+
2)ab	+	+	-	+	-	-	-
3)ac	+	-	+	-	+	-	-
4)a	+	-	-	-	-	+	+
5)bc	-	+	+	-	-	+	+
6)b	-	+	-	-	+	-	+
7)c	-	-	+	+	-	-	+
8)(1)	-	-	-	+	+	+	-

Supponiamo di aver aggregato i rispondenti secondo la media dei punteggi di preferenza rispetto a ciascun profilo di risposta per cui il vettore Y ora contiene le medie di tali punteggi.

$$\underline{Y} \quad \begin{matrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{matrix}$$

Supponiamo ora di aver stimato le utilità parziali di ciascun livello di ciascun fattore considerando come risposta i punteggi medi dei rispondenti, seguendo la procedura descritta precedentemente. L'interesse a questo punto è di valutare l'influenza dei fattori A,B e C e delle interazioni di tali fattori sui giudizi medi dei profili di prodotto espressi da tutti i rispondenti. Non essendovi gradi di libertà per la stima della varianza dell'errore, dato che la matrice dei profili è non replicata, è possibile evidenziare in modo grafico se un effetto influenzerà o meno la risposta sulla base della sua distanza dalla linea retta ideale tracciata sul NORMAL PROBABILITY PLOT. Dopo aver quindi stimato gli effetti dei fattori e delle interazioni si procede con la costruzione di un grafico probabilistico delle stime degli effetti. Se nessuno influenza la risposta, le stime si comportano come un componente casuale estratto da una distribuzione normale con media nulla, e gli effetti riportati nel grafico giaceranno approssimativamente su di una linea retta. Quegli effetti che si discostano dalla retta sono fattori che influenzano la risposta, dato che si potrebbe ipotizzare che si distribuiscano come .

Per disegnare il Normal Probability Plot si ordinano in senso crescente i 7 effetti dei fattori e delle interazioni che rappresenteranno le ascisse del grafico e si calcolano in corrispondenza i valori delle ordinate dati da:

$$Pi = 100(i - 1/2)/7 \quad \text{per} \quad i = 1, \ldots.., 7$$

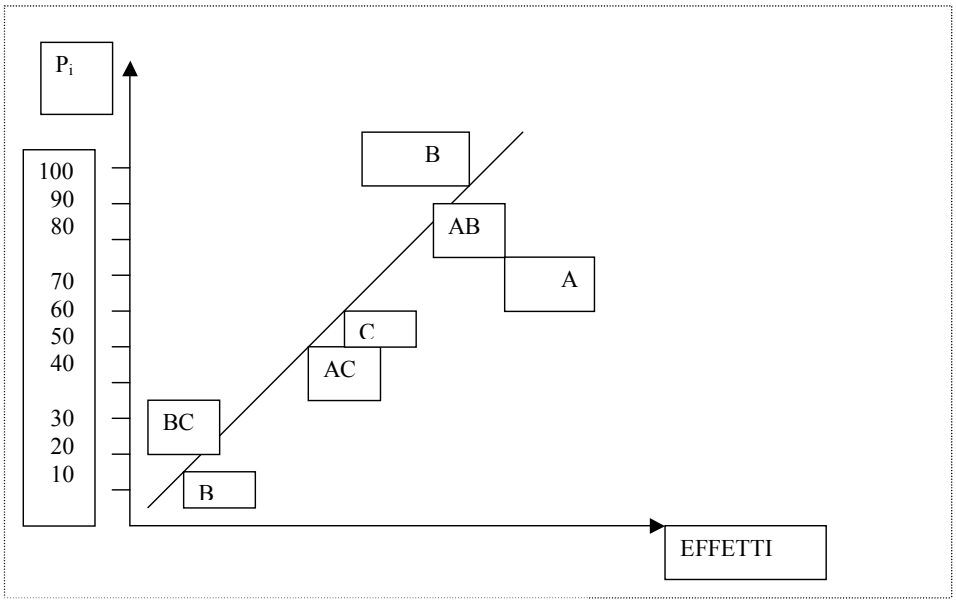


Fig. 2: Normal Probability Plot

Dall'ipotetico grafico riportato sopra si può evincere l'unico fattore che sembra influenzare la variabile risposta è il fattore A.

8 VERIFICA DI IPOTESI SUGLI EFFETTI DEGLI ATTRIBUTI MEDIANTE ANOVA

Supponiamo ora che i rispondenti siano stati preventivamente raggruppati in modo da costituire dei gruppi omogenei mediante ad esempio tecniche di CLUSTER ANALYSIS.

In tale situazione è possibile considerare i singoli rispondenti di ogni gruppo omogeneo come replicazioni. Quindi ad ogni singolo profilo di risposta è possibile attribuire una numerosità nm di unità sperimentali (soggetti rispondenti) con m = 1,. . . .,M, possibilmente bilanciata, ovvero nm = n per ogni profilo. Per esemplificare consideriamo il caso di un esperimento 22 replicato 2 volte (due rispondenti). La matrice dei profili e il vettore delle risposte sono i seguenti

Fatt.			Rispo.
A	B	AB	Y_{MJ}
+	+	+	Y_{11}
+	+	+	Y_{12}
+	-	-	Y_{21}
+	-	-	Y_{22}
-	+	-	Y_{31}
-	+	-	Y_{32}
-	-	+	Y_{41}
-	-	+	Y_{42}

dove m indicizza il numero dei profili $m = 1, \ldots.., 4$ e j indicizza le repliche, $j = 1, \ldots.., n_m$ e nel caso specifico$j = 1, 2$.

La corrispondente matrice 2 di variabili binarie per la stima delle utilità parziali dei singoli livelli dei fattori e delle combinazioni di livelli sarà quindi:

$$
Z= \begin{array}{cccccc}
1 & 0 & 1 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 \\
1 & 0 & 0 & 1 & 0 & 1 \\
0 & 1 & 1 & 0 & 0 & 1 \\
0 & 1 & 1 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 1 & 1 & 0 \\
\multicolumn{2}{c}{A} & \multicolumn{2}{c}{B} & \multicolumn{2}{c}{C}
\end{array}
$$

Da qui si può procedere secondo la procedura descritta precedentemente per stimare le utilità parziali dei singoli livelli e delle interazioni e per calcolare l'indice Irk di importanza relativa dei vari fattori.

Nel caso in cui si fanno delle replicazioni dei profili di prodotto, possiamo inoltre interpretare le y_{mj} secondo un modello di ANOVA multifattoriale. Nel caso specifico dell'esempio le y_{mj} possono essere espresse secondo un modello ANOVA a due vie:

$$
\begin{array}{l}
Y_{ilj} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + E_{ilj} \\
i = 1, \ldots.., a; \;\; l = 1, 2, \ldots.., b; \;\; j = 1, 2, \ldots.., n
\end{array}
$$

dove i indicizza i numero di livelli del fattore A, l indicizza il numero di livelli del fattore B e j il numero di replicazioni. Mediante l'applicazione dell'analisi della varianza sul modello ANOVA a due vie si può verificare tramite i test F parametrici se gli effetti dei fattori A e B e dalla loro intererazione influenzano in modo significativo le risposte sperimentali. L'analisi della varianza per la verifica di ipotesi sugli effetti dei fattori e dell'interazione è schematizzata nella tabella seguente:

Tab. 8: Tavola ANOVA

FONTE DI VARIABILITA	SOMMA DEI QUADRATI	G.d.l.	MEDIE DEI QUADRATI	TEST F
A	SS_A	a-1	$MS_A = SS_A/(a-1)$	$F_{OA} = MS_A/MS_E$
B	SS_B	b-1	$MS_B = SS_B/(b-1)$	$F_{OB} = MS_B/MS_E$
AB	SS_{AB}	(a-1) (b-1)	$MS_{AB}=SS_{AB}(a-1)(b-1)$	$F_{OAB} = MS_{AB}/MS_E$
ERRORE	SS_E	ab(n-1)	$MS_E = SS_E/ab(n-1)$	
TOTALE	SS_T	abn-1		

Dove SS_A è la devianza del fattore A e SS_E è la devianza dell'errore.
Ad esempio per verificare il seguente sistema di ipotesi:

$$M_0 = \alpha_1 = \alpha_2 = \ldots\ldots\ldots = \alpha_0 = 0$$
$$M_1 = \alpha_i \neq 0 \quad \textit{per almeno un } i$$

il rapporto $F_{OA} = MS_A/MS_E$ si distribuisce sotto l'ipotesi nella come una $Fa-1; ab(n-1)$ centrale.

Se il valore asservato di F_{OA} calcolato per il componente osservato F_{OSS} è tale che $F_{OSS} > F_{a-1,((ab(n-1):1-\alpha)}$.

si rifiuterà M_0 al livello di significatività α. Il grafico seguente mostra l'individuazione della regione critica del test.

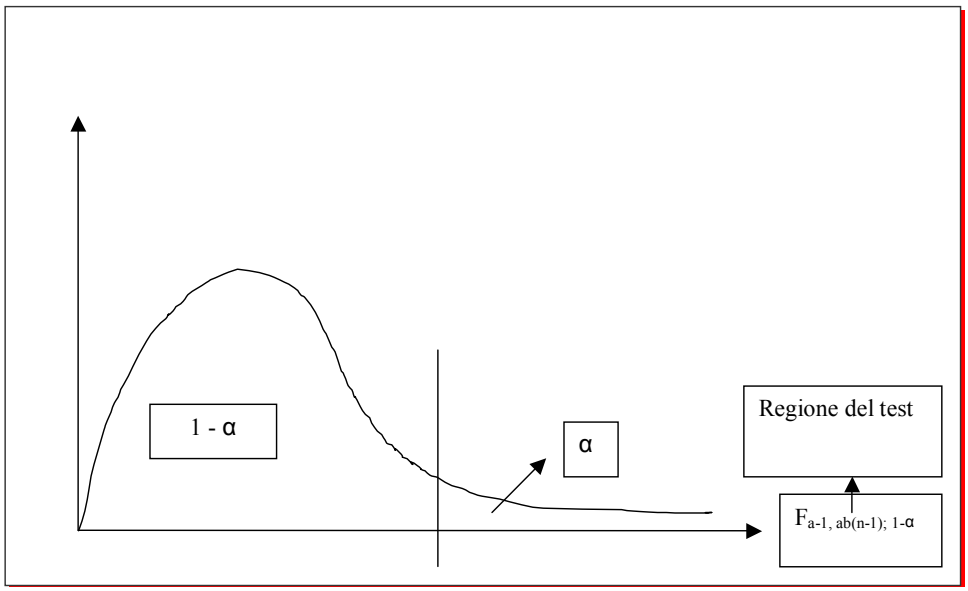


Fig. 3: Regione critica del test F

Similmente si deriva una verifica di ipotesi per F_{OB} e F_{OAB}.

9 L'USO DI PIANI FATTORIALI NELLA C.A.

Quando il numero dei fattori o il numero dei livelli aumenta, aumenta anche i numero dei profili da sottoporre ai rispondenti. Se ad esempio abbiamo cinque fattori su due livelli oppure tre fattori con rispettivamente 2, 3 e 4 livelli, nel primo caso avremmo un numero di profili pari a 32, nel secondo caso il numero di profili sarà 24, rendendo difficoltosa la somministrazione di tutti i profili. Per ovviare a tale inconveniente occorre fare riferimento alle tecniche di riduzione del piano sperimentale, sviluppate nell'ambito della pianificazione degli esperimenti. In particolare nel seguito consideriamo il caso in cui si vuole condurre uno studio pilota con fattori su due livelli dove il numero dei fattori è pari a 3 o a 4.

Se ad esempio il numero dei fattori è pari a 4, per realizzare un piano fattoriale completo che consenta la stima anche di tutte le interazioni fino a quelle di ordine massimo, bisogna disporre di un numero di profili pari a 16. Tale alto numero di profili può rendere difficile la realizzazione pratica dell'esperimento di CA, in quanto il rispondente potrebbe fornire giudizi di preferenza per lui significative in così tanti profili ed inoltre un numero elevato di profili può rendere l'indagine di CA troppo lunga e stancante per il rispondente.

Si consideri il seguente piano fattoriale 23:

Tab. 9: Esempio di piano fattoriale completo

	E	f	f.	F	a	t	t.
Profili	A	B	C	AB	AC	BC	ABC
abc	+	+	+	+	+	+	+
ab	+	+	-	+	-	-	-
ac	+	-	+	-	+	-	-
a	+	-	-	-	-	+	+
bc	-	+	+	-	-	+	+
b	-	+	-	-	+	-	+
c	-	-	+	+	-	-	+
(1)	-	-	-	+	+	+	-

dove ciascuna prova è indicata con lettera minuscola: se una lettera è presente, allora il corrispondente fattore è fissato al valore più alto se è assente, il fattore è presente al livello basso. La prova con tutti i fattori a livello basso è indicata con (1).

In generale una frazione, $\frac{1}{2}$ di un piano 2^K contiene 2^{K-1} prove e viene chiamato piano fattoriale frazionario 2^{K-1}. Ad esempio consideriamo il piano 2^{3-1}, cioè la frazione $\frac{1}{2}$ del piano 2^3. Supponiamo di scegliere le quattro prove a, b, c, abc come frazione $\frac{1}{2}$ del piano 2^3. Abbiamo selezionato le prove che forniscono un segno positivo per l'effetto ABC. Dunque ABC è detto generatore

di questa particolare frazione. Considerando I = ABC relazione di definizione del piano. Indicando in generale con l_A la stima dell'effetto del fattore A, si può dimostrare che $l_A = l_{BC}$, $l_B = l_{AC}$, $l_C = l_{AB}$. Conseguentemente non possiamo differenziare tra A e BC, B e AC e C e AB. Due o più effetti che hanno questa proprietà si dicono "alias". Nel piano considerato 2^{3-1}, A e BC, B e AC, C e AB sono alias.

In molte situazioni pratiche, sarà possibile scegliere la frazione in modo tale che gli effetti principali e le interazioni di basso ordine siano alias di interazioni di ordini alto (che saranno presumibilmente trascurabili). La struttura degli alias di questo piano, si determina a partire dalla relazione di definizione I = ABC. Moltiplicando ogni effetto per la relazione di definizione, si determinano gli alias di quell'effetto. Ad esempio l'alias di A è:

$$A = A \text{ X } ABC = A^2BC = BC$$

Si noti che $A^2 = I$ e quindi A^2BC restituisce la colonna BC. Consideriamo ora una frazione $\frac{1}{2}$ del piano 2^4. Supponiamo di utilizzare i piano 2^{4-1} con I = ABCD per analizzare i 4 fattori A, B, C, D. In questo piano gli effetti principali sono alias di interazioni fra i tre fattori:

$$A \text{ x } I = A \text{ x } ABCD$$
$$A = A^2BCD$$
$$A = BCD$$

Analogamente:

$$B = ACD$$
$$C = ABD$$
$$D = ABD$$

Le interazioni fra due fattori sono alias fra di loro:

$$AB = CD$$
$$AC = BD$$
$$AD = BC$$

Quindi in questo piano i fattori principali stimabili in modo non distorto se si assume che l'effetto delle interazioni di ordine tre è trascurabile, mentre gli effetti delle interazioni di ordine due non saranno stimabili in modo non distorto poichè le interazioni sono confuse a coppie.

Supponendo quindi di dover condurre un esperimento di CA con quattro fattori principali a due livelli, se è possibile ipotizzare che l'effetto delle interazioni di ordine superiore o uguale al terzo sia trascurabile, si possono ottenere le stime delle utilità parziali dei livelli dei fattori, applicando un piano frazionato 2^{4-1} con relazione definente I = ABCD. Questo permette quindi di sottoporre a ciascun rispondente un numero di profili pari ad 8 in luogo dei 16 che necessiterebbe un piano completo 2^4. Si noti che tutti i piani frazionati o completi

finora considerati hanno l'importante proprietà di essere piani di tipo ortogonale, ovvero sia X la matrice dei profili, si ha che $(X^TX)^{-1}= 2^K I_K$ dove I_K è la matrice identità di ordine K, con K numero di attributi. L'utilizzo di piani completi o frazionati ortogonali per condurre una Full Profile CA permette di rendere minima la varianza degli stimatori ai minimi quadrati utilizzati per la stima delle utilità parziali. Sia $Y = X\beta + \varepsilon$ il modello lineare di riferimento, con $E(\varepsilon) = 0$ e $Var(\varepsilon) = \sigma^2 I_m$, dove m è il numero dei profili, si ha che $Var(\text{ cov }(\beta)) = \alpha^2(X^TX)^{-1}$ dove β è la stima ai minimi quadrati di β. Nella formula della matrice di varianza e covarianza dei β compaiono due quantità σ^2 e $(X^TX)^{-1}$. La soluzione di minimo di tale quantità è realizzata solo nel caso in cui la matrice dei profili sia ortogonale.

Per concludere, nell'adattare una Full Profile CA diventa cruciale la scelta della matrice dei profili che nel caso in cui il numero di attributi sia relativamente elevato, dovrebbe essere scelta secondo un piano frazionato di tipo ortogonale.

10 Il pacchetto SPSS-CONJOINT: alcuni esempi applicativi

Questo software si basa su uno sviluppo di analisi congiunta di tipo tradizionale, in cui si prevede l'utilizzo di un piano fattoriale.

In particolare, l'opzione Conjoint, in appoggio all'SPSS Base System risulta completamente integrata in questo sistema. Questa opzione include:

1) Le procedure per lo sviluppo e la determinazione di profili usati negli studi di Conjoint.
2) Quattro differenti modelli per la Conjoint Analysis.
3) Un sommario grafico per i risultati dell'analisi congiunta.

10.1 Il metodo "full profile"

Questo metodo, prevede che il rispondente assegni un ordine ai profili presentati, in accordo con le sue preferenze.

Come già accennato visto in precedenza, in ciascuno dei profili presentati, sono presenti tutti i fattori attraverso diverse combinazioni di livelli. In questo modo, in ciascun profilo è descritto un servizio o un prodotto completo, ossia un full concept. Compito dell'intervistato è quello di classificare ciascun profilo secondo un criterio di preferenza basato, ad esempio, sul gradimento o sulla possibilità di acquisto o su altre scale di preferenza. Da questi rank o score la Conjoint Analysis determina le utilità (parth worth) per ciascun livello.

10.2 L'Orthogonal Array

Con un approccio di questo tipo, ben presto ci si accorge di un potenziale problema; al rispondente, infatti, viene presentato un numero troppo elevato di

combinazioni di livelli, affinché egli possa classificare i vari profili in maniera utile e significativa.

Per questa ragione, frequentemente, si ricorre ad un sottoinsieme di tutti i possibili profili. Questo sottoinsieme, chiamato Orthogonal Array, considera solo attributi presi singolarmente, mentre le interazioni tra i singoli fattori sono considerate trascurabili.

10.3 I profili di risposta

Una volta che l'Orthogonal Array è determinato, ciascun profilo deve essere presentato al rispondente in modo distinto dagli altri, al fine di permettere all'intervistato di focalizzarsi solo sul prodotto presentato. Gli "stimoli" dovrebbero inoltre essere standardizzati, in modo da assicurarsi che i diversi profili siano simili.

10.4 Raccolta ed analisi dei dati

Come accennato in precedenza, durante lo studio, a ciascun rispondente è presentato un set di profili e gli viene chiesto di esprimere una preferenza per il prodotto. L'analista può chiedere al rispondente di indicare una preferenza in molti modi. L'intervistato può, infatti, assegnare uno score a ciascun profilo, dove il più alto indica la preferenza più grande; alternativamente, il rispondente può usare la modalità rank da 1 a n per ciascun rispondente, dove n indica il numero di profili presentati e il rank più basso indica la preferenza migliore.

Qualsiasi sia il metodo usato, i dati vengono poi registrati per ciascun intervistato. La procedura di Conjoint è poi usata per stimare le utilità per ciascun individuo e per tutto il campione. I risultati evidenziano quale combinazione sia la preferita e quali siano le componenti del prodotto che maggiormente indirizzano la preferenza dell'intero prodotto e la relativa importanza di ciascun fattore. La presenza di un parth worth per ciascun livello di ciascun fattore permette anche di determinare gli effetti delle combinazioni di livelli non presenti attualmente nell'esperimento.

Tutte le informazioni ottenute da un'analisi congiunta, possono essere poi applicate ad una grande varietà di questioni riguardanti il market research. Come si è già approfondito in precedenza i risultati possono essere utili per analizzare aree quali, la progettazione di prodotto, l'analisi costi benefici, la segmentazione di mercato, strategie pubblicitarie e soprattutto la definizione del Concept Test.

10.5 Progettare un esperimento di analisi : generazione di un disegno ortogonale

Negli studi di analisi congiunta, l'analista assume che il prodotto in esame possa essere definito in funzione di poche importanti caratteristiche. Risulta inoltre evidente che, quando un consumatore compie una decisione su un determinato

prodotto, questa è determinata dai trade off tra le varie caratteristiche. In questo modo nessun prodotto presenterà le migliori caratteristiche, come nessuno le peggiori; sarà il consumatore a decidere quali caratteristiche saranno importanti e quali no. Lo scopo della Conjoint Analysis è quello di determinare l'utilità (parth worth) di queste caratteristiche; l'utilità misura, infatti, quanto importante è ciascuna caratteristica rispetto al sistema di preferenze totale del consumatore.

Come visto più volte, le caratteristiche di un prodotto sono descritte in termini di attributi e livelli e, considerando tutte le possibili combinazioni di attributi e livelli, ne risulta un numero di profili difficile da gestire per i rispondenti. Ecco il motivo per cui molti analisti utilizzano l'Orthogonal Array, che consente una stima efficace dei part worths per tutti gli effetti principali. In un Orthogonal Array, ciascun livello di ogni fattore si presenta con una frequenza proporzionale a quella degli altri assicurando in questo modo l'indipendenza degli effetti principali.

L'Orthogonal Array rappresenta il metodo più veloce per la stima di tutti gli effetti principali. Pur essendo vero che le stime dei part worth si incrementano notevolmente all'aumentare del numero di profili, si può dire che le informazioni relative ai profili omessi non sono perse, in quanto, una volta a disposizione delle ciascun livello di ogni fattore si presenta con una frequenza proporzionale a quella degli altri assicurando in questo modo l'indipendenza degli effetti principali.

In aggiunta ai casi in analisi, si possono specificare anche gli holdout cases, che risultano profili giudicati dai rispondenti, ma non sono usati per stimare le utilità, ma utilizzati come strumento di controllo per la validità delle "utilities" stimate. Gli holdout cases sono generati in un altro random plan e non all'interno del piano sperimentale ortogonale.

Da sottolineare che il piano ortogonale viene salvato come qualsiasi altro file di dati, in modo da poterlo utilizzare successivamente, sia per la presentazione dei profili ai rispondenti, sia per l'analisi congiunta. Da sottolineare che il piano ortogonale viene salvato come qualsiasi altro file di dati, in modo da poterlo utilizzare successivamente, sia per la presentazione dei profili ai rispondenti, sia per l'analisi congiunta.

Nel data file compaiono due variabili, denominate rispettivamente card e status.

La variabile card assegna un valore sequenziale ad ogni profilo, in relazione alle preferenze espresse, mentre la variabile status indica se un caso è parte del disegno sperimentale, oppure un holdout case, piuttosto che un simulation case.

10.6 Simulation Cases

Dopo aver generato un disegno ortogonale, si può creare anche un simulation case.

Queste sono combinazioni che non vengono prese in considerazione dagli intervistati, ma sono incluse nella Conjoint Analysis, basata sulle preferenze

dei profili sperimentali. Il modo più semplice per generare i simulation cases è quello di usare i value labels generati, una volta creato il disegno sperimentale.

11 Esempio applicativo (1) – Personal Computer

Di seguito viene descritto un esempio applicativo condotto col software, al fine di comprendere in maniera più realistica i concetti descritti in precedenza.

L'esempio in questione riguarda un personal computer portatile, in cui vengono individuati tre attributi, quali: peso, durata batteria, prezzo. Nel dettaglio ho quindi (vedi Fig -):

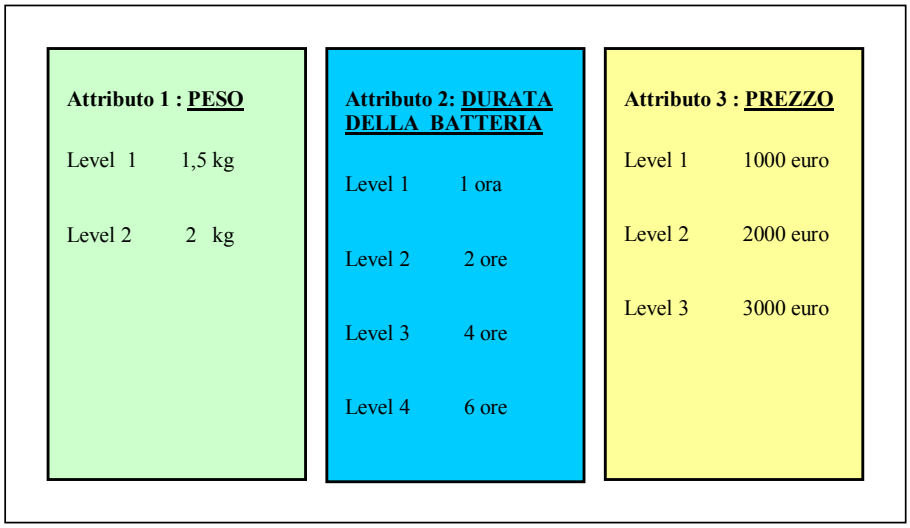


Fig.4: In figura sono riassunti gli attributi, con i rispettivi livelli.

11.1 Come si ottiene un disegno ortogonale

Ora, una volta definiti i livelli e i rispettivi attributi, è necessario generare un disegno ortogonale. Nell'esempio in questione risulta un piano ortogonale con 16 profili, come evidenziato nella tabella seguente, su un totale di 24 (2x4x3):

Tab. 10: Piano ortogonale relativo all'esempio

Peso (in kg)	Durata batteria (in ore)	Prezzo (in euro)
2	6	2000
1,5	2	2000
2	6	1000
1,5	1	3000
2	4	3000
1,5	6	1000
1,5	1	1000
2	2	1000
1,5	6	3000
2	4	1000
1,5	2	1000
1,5	4	1000
2	1	2000
2	2	3000
2	1	1000
1,5	4	2000

Nell'esempio trattato, il file contenente l'esempio in questione viene salvato con il seguente nome: computer_plan.sav.

La procedura che consente la generazione di un piano ortogonale, prevede la creazione di un file di dati contenente i principali effetti, che soddisfano la proprietà di ortogonalità senza testare ogni combinazione di valori.

- Per ottenere un disegno ortogonale, si sceglie dai menu:

Data
Orthogonal Design
Generate...

Questo permette di aprire il "*Generate Orthogonal Design dialog Box*".

Si deve poi definire almeno un fattore. Si digita un nome per il Factor Name, che può essere qualsiasi nome di variabile, eccetto Status o Card.

- Nell'esempio in questione i nomi delle tre variabili saranno: Peso, Durata della Batteria e Prezzo.
- Con il comando *Add* si aggiunge un factor name. Per cancellare un fattore lo si deve selezionare dalla lista e poi usare il comando *Remove*.
- Successivamente si devono assegnare i valori a ciascun livello del fattore o dei fattori precedentemente definiti. Per fare questo, dopo aver selezionato il numero di fattori desiderato, si utilizza il comando *Define Values*.Questo in

riferimento all'esempio trattato, permetterà di definire i livelli dei tre attributi trattati, che sono il peso del pc, la durata della batteria e il prezzo.

11.2 Raccolta dei dati

Una volta scelto l'esempio, compito dell'analista è quello di gestire i diversi profili per ciascun rispondente. La procedura di Conjoint consente tre metodologie di raccolta dati. Nel primo metodo, ai rispondenti è chiesto di assegnare un punteggio di preferenza dei vari profili; questo è tipicamente usato con una scala Likert, oppure quando viene chiesto all'intervistato di assegnare un numero da 0 a 100, per indicare la sua preferenza.

Nel secondo metodo, all'intervistato è chiesto di determinare un ranking, che va da 1 al numero totale di profili considerati.

Nel terzo e ultimo metodo, ai soggetti è chiesto di ordinare i profili in termini di preferenza. In questo ultimo caso, l'analista registra i numeri dei profili secondo l'ordine dato da ciascun rispondente.

Nell'esempio in questione si è scelto di raccogliere i dati in modalità rank, per un solo intervistato.

In riferimento al piano ortogonale generato, il file presenta il seguente ranking di profili (14, 11, 1, 3, 10, 8, 6, 16, 2, 4, 7, 12, 15, 9, 13, 5). Questo evidenzia, ad esempio, come il profilo 14, occupando la prima posizione nel ranking, risulti il preferito, al termine della fase di interviste, mentre il profilo 5, occupando l'ultima posizione, risulti il meno preferito. I dati di questo ordinamento vengono poi salvati in un file denominato *computer_ rank.sav*.

11.3 Come ottenere una Conjoint Analysis

Per ottenere una procedura di Conjoint Analysis si procede in questo modo. Dal menu si sceglie:

File
New $\Longrightarrow$
SpssSyntax...

Si digitano poi i comandi nella finestra di sintassi. Nell'esempio in questione questi sono i comandi utilizzati:

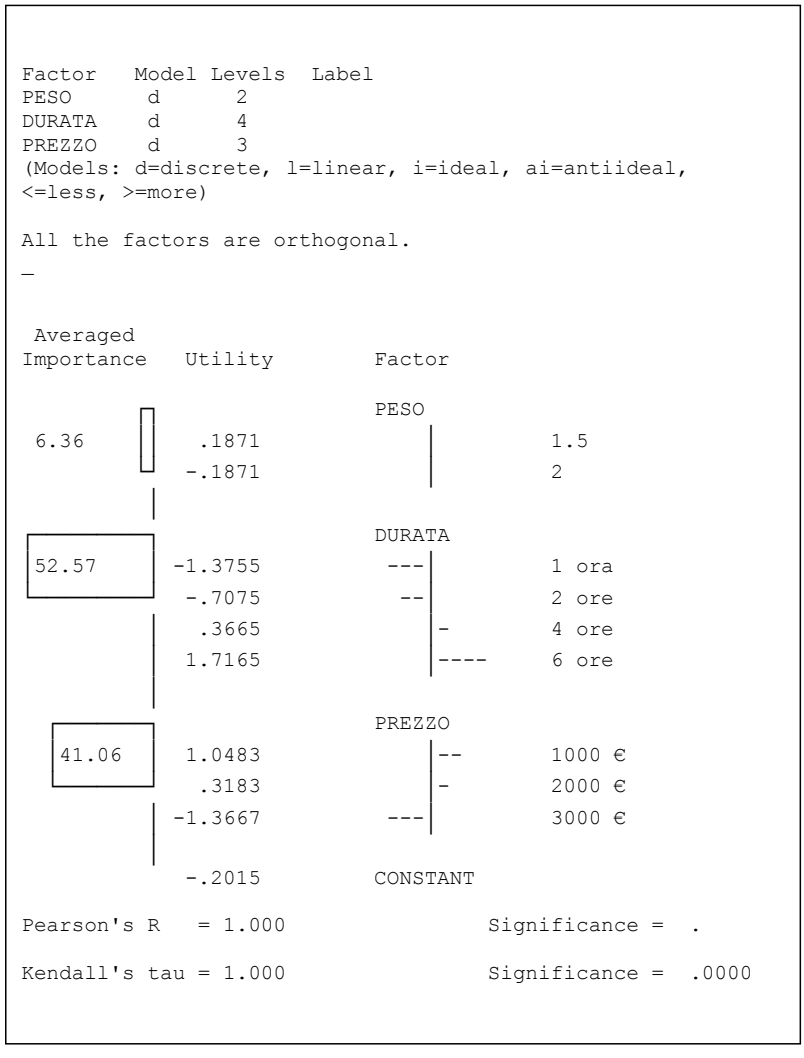
```
conjoint plan='C:\Documenti\ computer_ plan.sav'
 /data='C:\Documenti\computer_ score_ hold_ out.sav'
 /score=score_ 1 to score_ 16
 /plot summary.
```

Questa corrisponde, sostanzialmente, alla specificazione minima per la Conjoint; come si nota, la Conjoint Analysis richiede la definizione di due files, un data file e un plan file, creato usando la procedura Orthoplan, già descritta in precedenza e la specificazione di come i files sono ordinati.

I dati, infatti, possono essere ordinati in uno di questi tre modi:

- Sequenzialmente, attraverso il comando sequence, con il quale, ogni valore nel data file corrisponde ad un numero, che contraddistingue un profilo; si parte dal migliore e si termina con il peggiore, in termini di preferenza.
- Con modalità ranking ,attraverso il comando *rank*, si ha che una posizione bassa corrisponde ad un profilo con una buona preferenza.
- Con un punteggio di preferenza, utilizzando il comando *score*, si assegna un punteggio di preferenza ad ogni profilo, partendo dal punteggio del profilo 1, poi quello del profilo 2 e così via.

Ritornando all'esempio in questione, l'output della Conjoint Analysis è evidenziato nella figura seguente:



```
Factor   Model Levels  Label
PESO      d      2
DURATA    d      4
PREZZO    d      3
(Models: d=discrete, l=linear, i=ideal, ai=antiideal,
<=less, >=more)

All the factors are orthogonal.
_


 Averaged
Importance   Utility        Factor

                            PESO
 6.36          .1871                 1.5
              -.1871                 2

                            DURATA
52.57       -1.3755          ---     1 ora
             -.7075            --     2 ore
              .3665             -     4 ore
             1.7165                   6 ore

                            PREZZO
41.06        1.0483                  1000 €
              .3183             -     2000 €
             -1.3667          ---     3000 €

              -.2015        CONSTANT

Pearson's R   = 1.000                Significance =  .

Kendall's tau = 1.000                Significance =  .0000
```

Fig.5: Output dell'esempio di conjoint condotto con Spss

Come si può notare, oltre ad essere messi in evidenza i valori delle utilità per ogni livello di ogni attributo, si evidenziano i valori di alcuni indici.

In particolare se *Significance* < 0.05 e gli *indici di Pearson's R* e *Kendall's Tau* sono vicini a 1, allora il modello lineare stimato ben si adatta ai dati rilevati.

Di seguito sono riportati in modalità grafica i valori delle utilità per ogni attributo e successivamente confrontati, in termini di importanza, nell'ultimo grafico riassuntivo.

La modalità grafica è resa disponibile attraverso il comando *"plot"*, che consente di ottenere grafici ad alta risoluzione, in aggiunta a quello che è l'output usuale prodotto dal pacchetto CONJOINT.

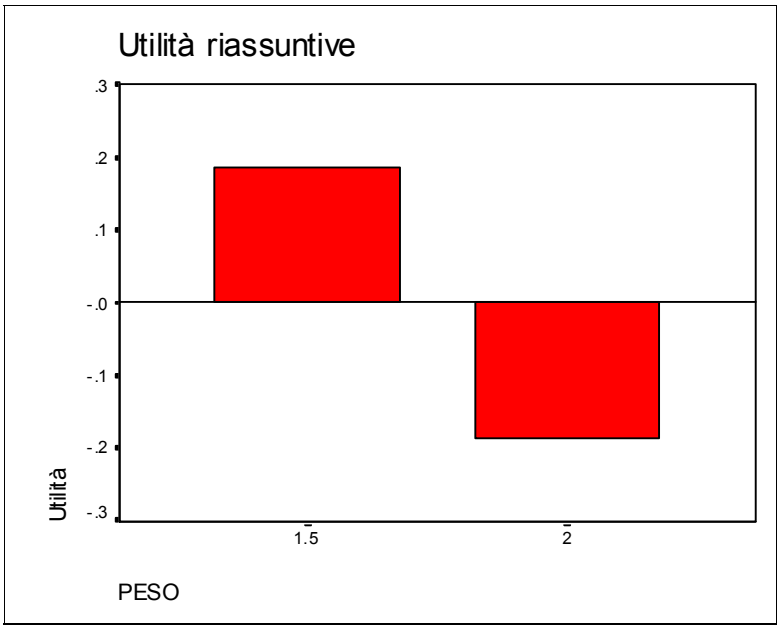


Fig. 6: Utilità dell'attributo "peso", espressa in modalità grafica.

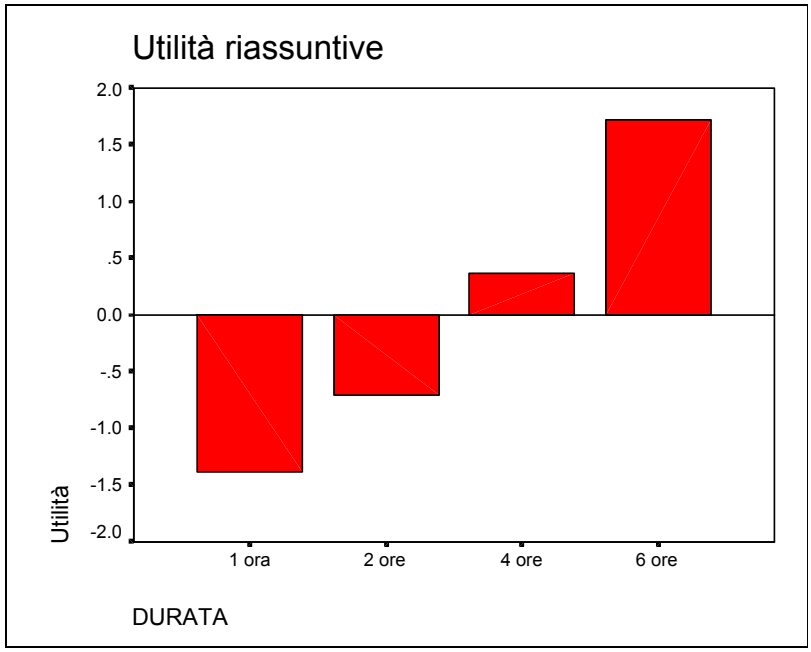


Fig. 7: Utilità dell'attributo "durata batteria", espresso in modalità grafica.

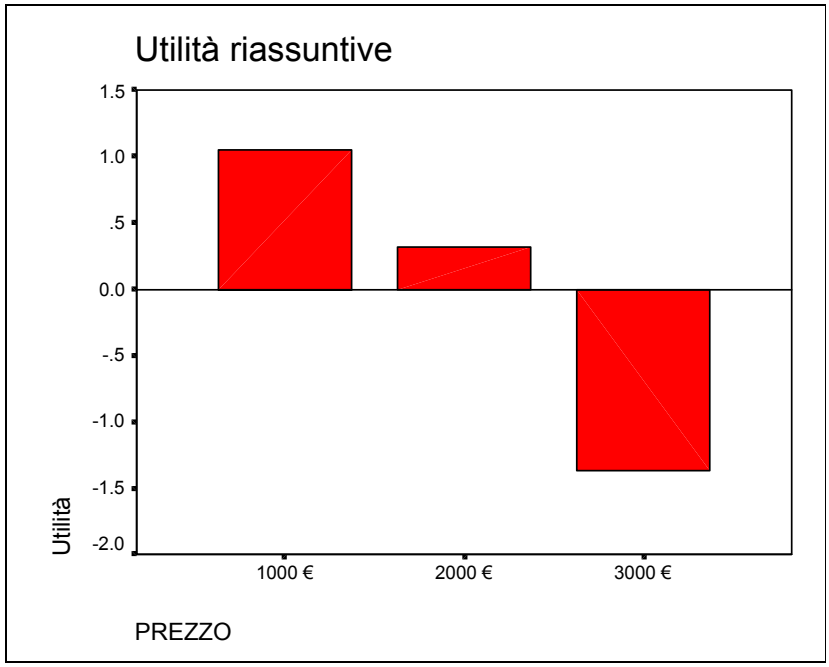


Fig. 8: Utilità dell'attributo "prezzo", espresso in modalità grafica.

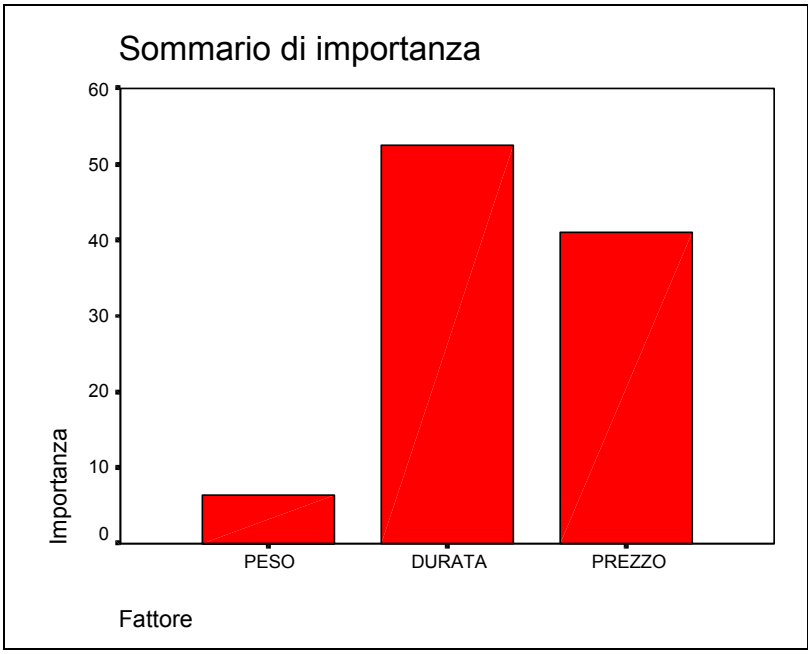


Fig. 9: Vengono messe a confronto le importanze dei tre attributi.

Di seguito, sempre dallo stesso esempio, si analizzano altri due esperimenti di Conjoint, che prevedono dieci rispondenti; nel primo caso si ha un'analisi aggregata dei dati, mentre nel secondo si presenta un'analisi di tipo disaggregato.

11.4 Esempio con 10 rispondenti (aggregato)

La sintassi di questa parte di esempio è la seguente:

```
conjoint plan='C:\Documenti\computer_plan.sav'
 /data='C:\Documenti\ computer_score.sav'
 /utility='C:\Documenti\ computer_utility.sav'
 /subject=risp
 /score=score_1 to score_16
 /plot summary.
```

L'output corrispondente è evidenziato nella figura seguente a pagina seguente:

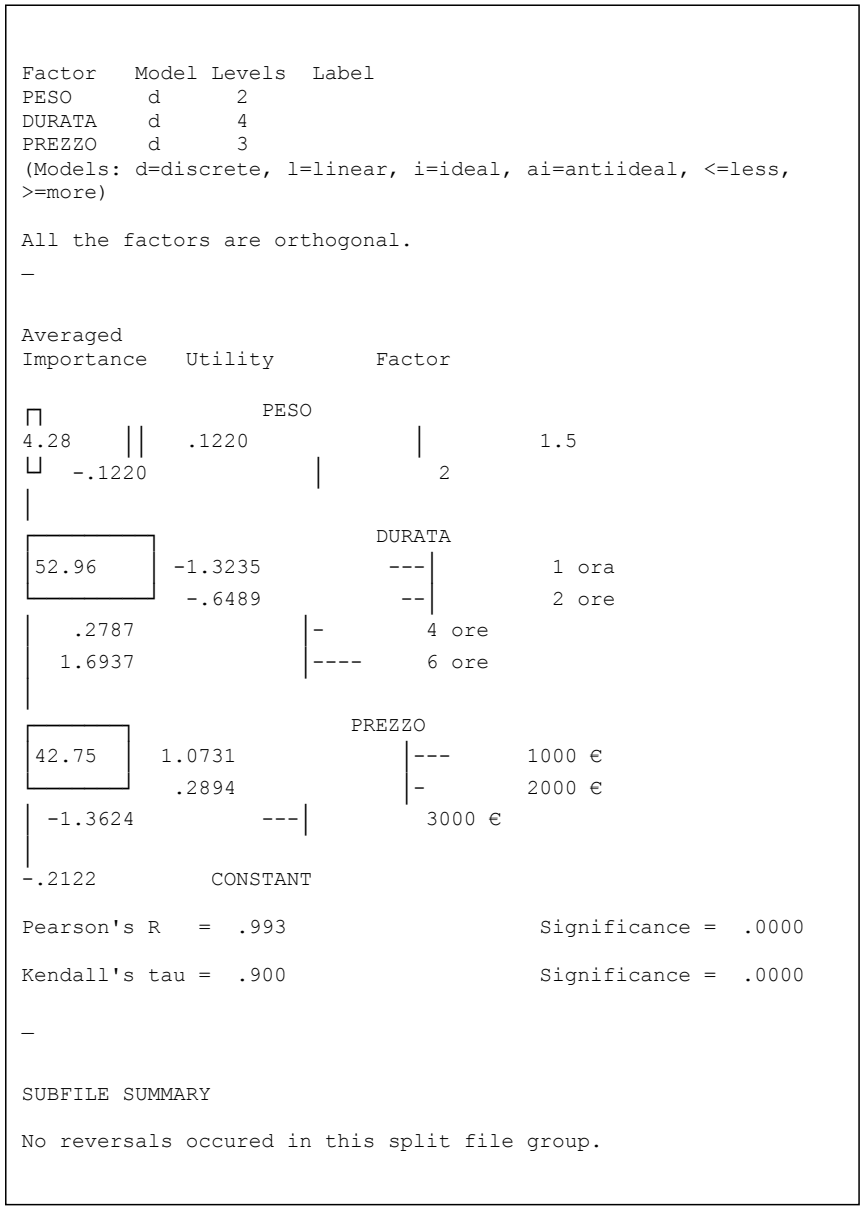


Fig. 10: Output del caso con 10 rispondenti e piano aggregato

In modalità grafica l'output è evidenziato nelle figure nelle pagine successive:

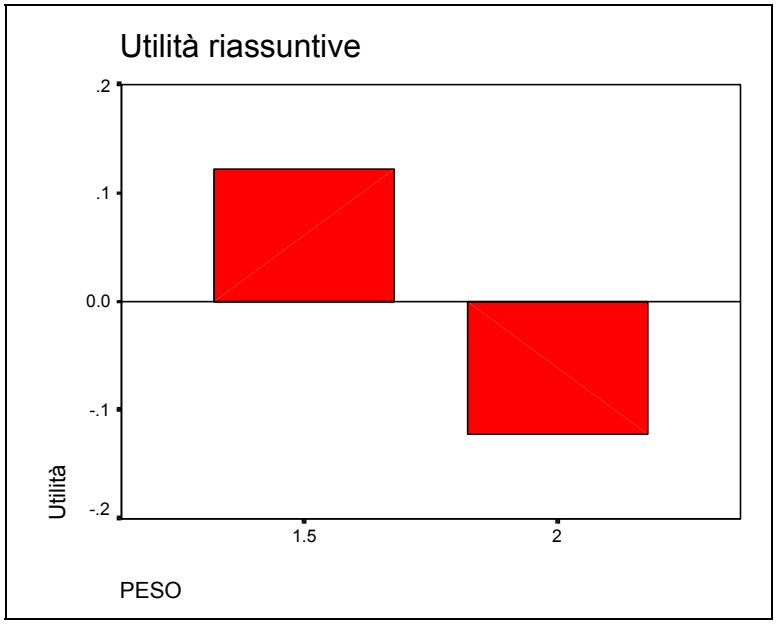


Fig. 11: Utilità del primo attributo

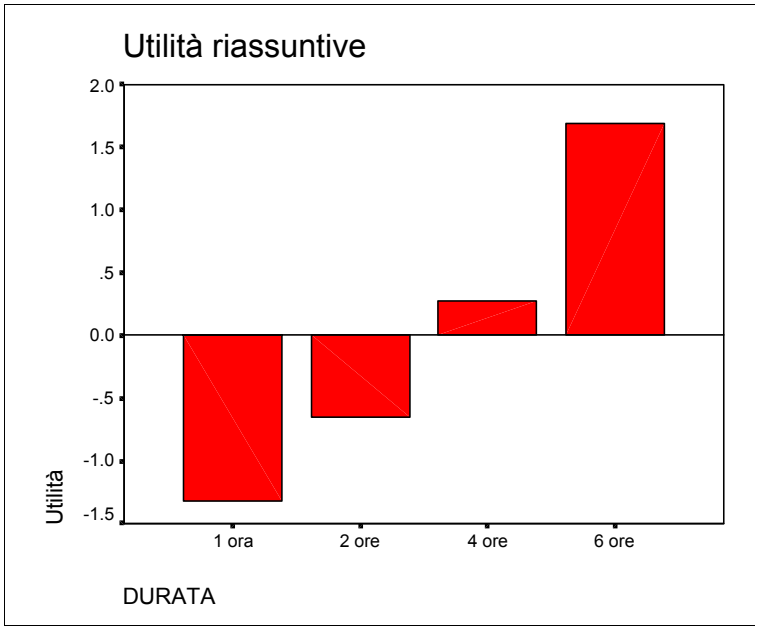


Fig. 12: Utilità del secondo attributo

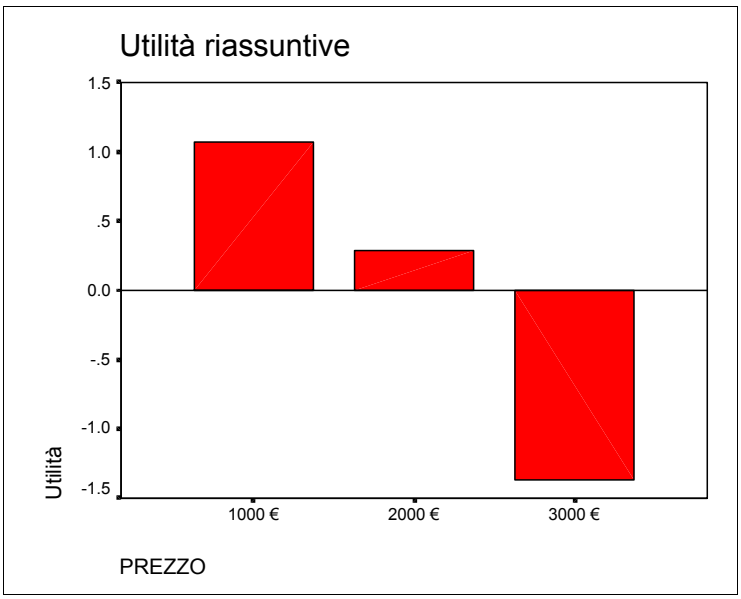


Fig. 13: Utilità delterzo attributo

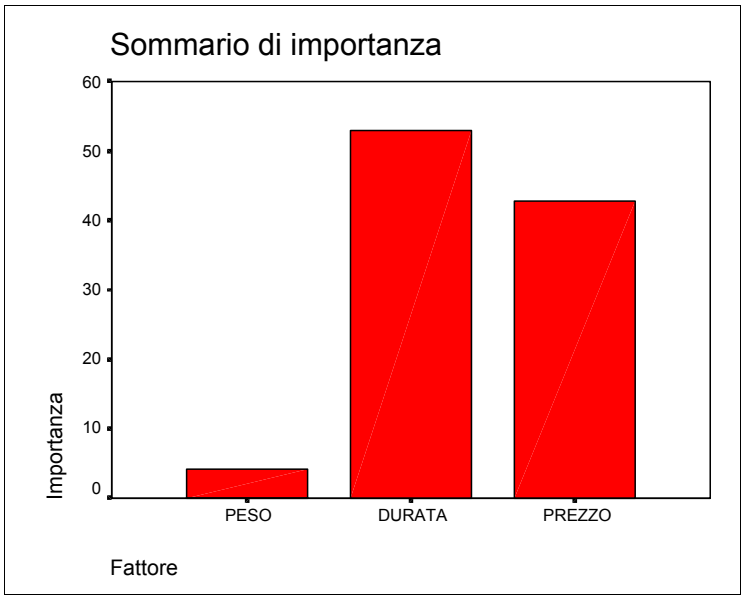


Fig. 14: Confronto complessivo delle utilità dei tre fattori con 10 rispondenti e con piano di tipo aggregato

11.5 Esempio con 10 rispondenti (disaggregato)

La sintassi di questa parte di esempio è la seguente:

```
conjoint plan='C:\Documenti\ computer_plan.sav'
 /data='C:\Documenti\ computer_score.sav'
 /utility='C:\Documenti\ computer_utility.sav'
 /subject=risp
 /score=score_1 to score_16
 /plot subject summary.
```

Nelle figure seguenti sono messi evidenza solamente le modalità grafiche di questa parte di esempio:

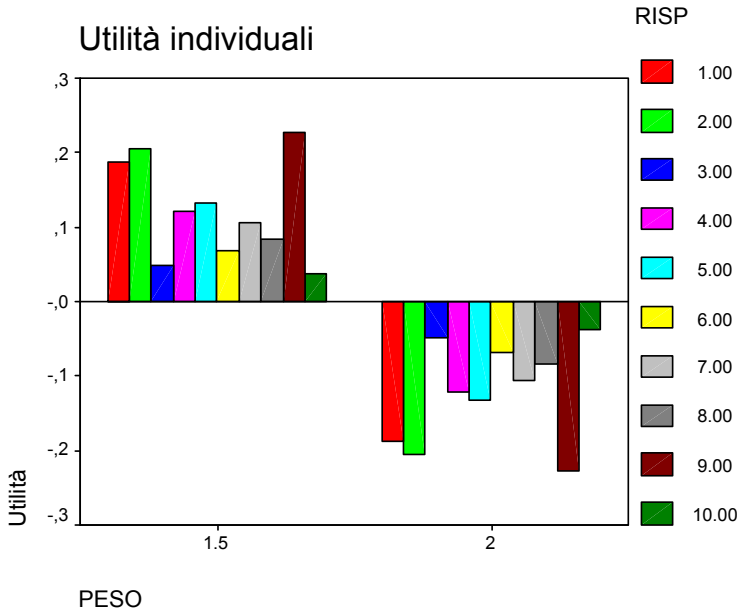


Fig.15: In figura, sono messi in evidenza i valori di utilità, in modalità disaggregata, per i
dieci rispondenti.

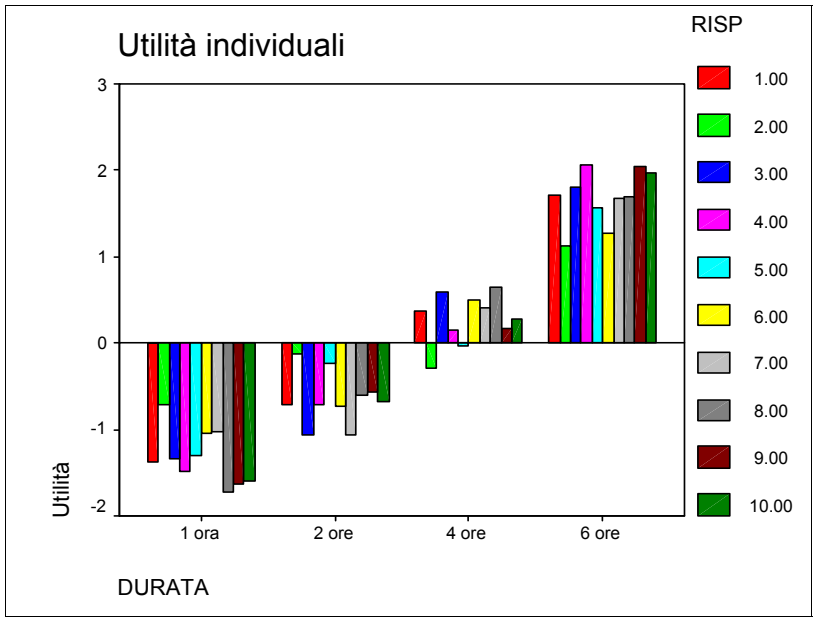


Fig. 16: Sono evidenziati i valori di utilità, in modalità disaggregata per i dieci rispondenti, per l'attributo "durata della batteria".

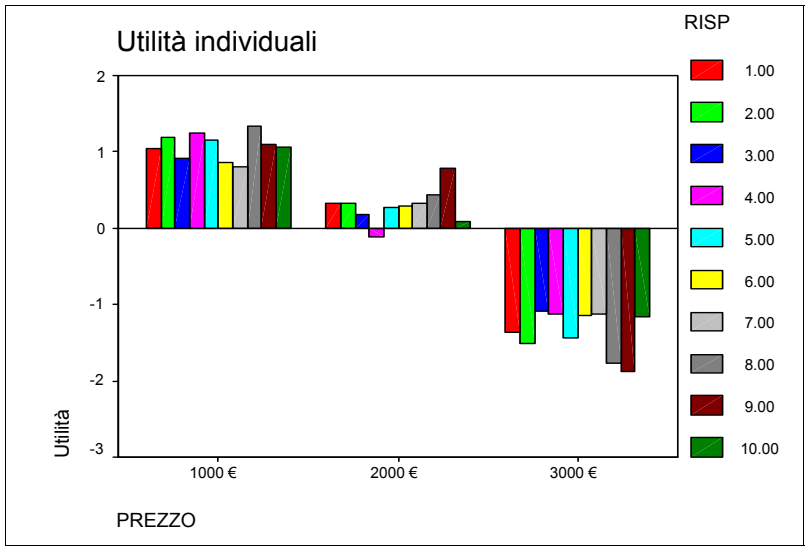


Fig.17: I valori di utilità dell'attributo "prezzo", per i dieci rispondenti, con piano disaggregato.

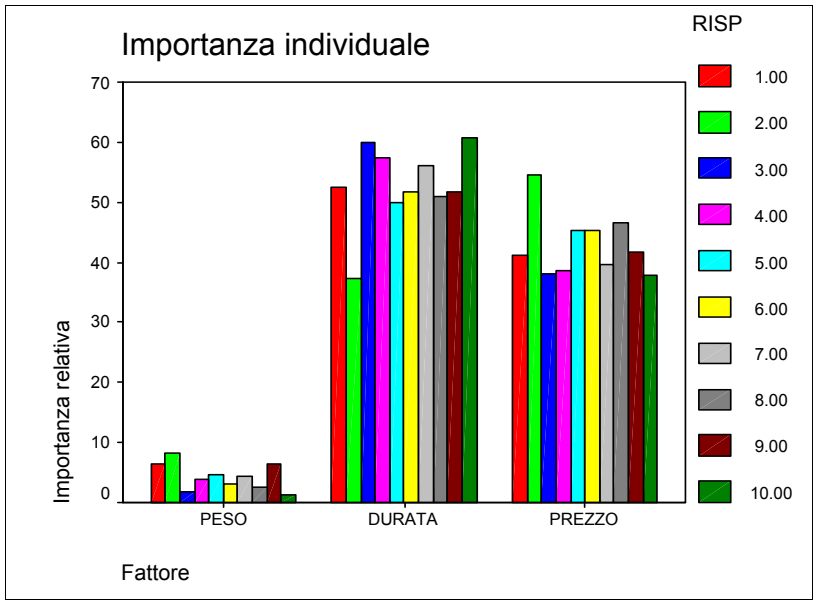


Fig.18: Grafico riassuntivo, che evidenzia l'importanza relativa per i dieci rispondenti.

12 Esempio applicativo (2) – Alimentari per animali domestici

(stima del mercato potenziale nel lancio di una marca di alimenti per animali domestici)

Un'impresa operante nel mercato degli alimenti per animali domestici (pet food) è interessata ad inserire nella sua linea di produzione una nuova marca di cibo per cani e gatti con piccoli appena nati. A giudizio del management dell'azienda sono cinque gli attributi che più di ogni altro possono influire sulla propensione dei proprietari di cani e gatti ad acquistare il prodotto:

1. il nome (sono state ipotizzate tre opzioni: Mamacat, Formula9 e Plus);
2. il tipo di confezione (sono contemplate tre varianti: scatoletta, vasetto di vetro e tetrapack);
3. il prezzo (sono allo studio tre alternative: 0.55€, 0.80€, 1€ a confezione);
4. l'approvazione da parte dell'ordine dei medici veterinari (due possibili modalità: sì o no);
5. la garanzia della restituzione dell'importo speso in caso di insoddisfazione (di nuovo due sole modalità: sì o no).

L'impresa dispone di un archivio di 1.500 clienti che hanno preso parte in passato ad iniziative di fidelizzazione, quali raccolte di punti e buoni omaggio, e decide quindi di selezionarne un campione con criterio probabilistico. Lo stanziamento a disposizione per la realizzazione dell'indagine sul campo, da effettuarsi mediante intervista diretta, consente di entrare in contatto con 150 consumatori. A partire da valutazioni di preferenza di questi ultimi, si tratta di capire quali attributi siano da ritenere più importanti e quale combinazione possa determinare il più elevato tasso di acquisto.

Tutte le possibili combinazioni delle 13 modalità dei 5 attributi, e quindi tutte le possibili versioni distinte della nuova marca, ammontano a $3 \times 3 \times 3 \times 2 \times 2 = 108$. Si potrebbe pensare, in prima istanza, di chiedere a ciascun consumatore del campione di esprimere un giudizio su ognuna di queste. E' del tutto evidente che in un caso simile l'onere di collaborazione a carico degli intervistati risulterebbe però eccessivo. Ai fini dell'indagine sul campo è necessario utilizzare un sottoinsieme assai più ristretto dei possibili profili. E' indispensabile però garantire che i risultati ottenuti possano essere estesi a tutti i profili, anche a quelli che non sono oggetto di valutazione diretta. Tale requisito viene soddisfatto, nelle applicazioni di C.A., grazie all'impiego di disegni fattoriali frazionati di tipo ortogonale (fractional factorial orthogonal arrays), cioè di piani degli esperimenti che consentono di stimare in modo indipendente, con il minor numero possibile di combinazioni, le utilità parziali associate a ciascuna modalità di ogni caratteristica. Ciò avviene valutando i soli effetti principali di tali modalità, coerentemente con il tipo di schema additivo ipotizzato per l'utilità totale. Nel caso considerato gli stimoli concretamente destinati agli intervistati si riducono in questo modo a 18, come descritto nella tabella seguente.

Tab. 11: Evidenzia il disegno comprendente gli stimoli concretamente destinati agli intervistati.

Stimoli	Nome	Confezione	Prezzo in euro	Approvazione dell'ordine	Garanzia del rimborso	Ordina-mento di preferenza	Punteggio di gradimento
1	Mamacat	Scatoletta	1	No	Sì	12	50
2	Formula 9	Scatoletta	0.55	Sì	No	7	78
3	Plus	Scatoletta	0.80	No	Sì	9	66
4	Mamacat	Vasetto	1	Sì	No	18	15
5	Formula 9	Vasetto	0.55	No	Sì	8	72
6	Plus	Vasetto	0.80	No	No	15	32
7	Mamacat	Tetrapack	0.55	No	No	4	90
8	Formula 9	Tetrapack	0.80	Sì	No	6	80
9	Plus	Tetrapak	1	No	Sì	5	84
10	Mamacat	Scatoletta	0.80	No	No	10	52
11	Formula 9	Scatoletta	1	No	No	16	28
12	Plus	Scatoletta	0.55	Si	Si	1	95
13	Mamacat	Vasetto	0.80	No	No	13	42
14	Formula 9	Vasetto	0.80	No	Si	11	50
15	Plus	Vasetto	1	Si	No	17	20
16	Mamacat	Tetrapak	0.80	Si	Si	2	92
17	Formula 9	Tetrapak	1	No	No	14	38
18	Plus	Tetrapak	0.55	No	No	3	90

Disegni di questo tipo non sono facili da progettare e, nella prassi corrente, sono il risultato dell'applicazione della procedura "Orthoplan" del pacchetto statistico SPSS. La presentazione delle 18 alternative è stata realizzata, da parte degli intervistatori, mediante l'impiego di appositi cartellini che riproducono per intero i profili. Tale soluzione è preferibile ad una descrizione puramente verbale, in quanto più efficace. Ciascun intervistato ha proceduto sia a fornire una graduatoria delle alternative in conformità alle proprie preferenze (fornendo, in questo modo, risposte su scala di tipo ordinale), sia a valutare le stesse in termini di gradimento, con un punteggio tra 0 e 100 (proponendo quindi giudizi che soddisfano i requisiti di una scala ad intervallo). Nella tabella sono riportati, a titolo di esempio, sia l'ordinamento di preferenza sia i punteggi di gradimento forniti per i 18 stimoli da una unità del campione.

Una volta disponibili i dati di base, il successivo compito da assolvere, sotto il profilo analitico, è quello di stimare le utilità parziali associate a ciascuna modalità di ogni attributo. Il metodo di stima da applicare deve essere tale da garantire che la somma delle utilità parth worth, riferite alle modalità contenute

nelle 18 alternative, dia luogo a valori di preferenza/utilità totale che riproducano il più fedelmente possibile l'ordinamento di preferenza o i punteggi di gradimento espressi. Ad esempio, sempre relativamente alle informazioni della tabella precedente, la somma delle utilità parziali associate alle modalità Plus del carattere nome, scatoletta del carattere tipo di confezione, 0.50 euro del carattere prezzo, sì del carattere di approvazione da parte dell'ordine dei medici veterinari e sì del carattere garanzia di rimborso, cioè quelle che designano la combinazione maggiormente preferita dall'intervistato, deve fornire il valore più elevato rispetto a tutti quelli corrispondenti alle altre combinazioni sottoposte a giudizio diretto.

12.1 Modelli di analisi

In relazione a quanto appena detto, le tecniche statistiche utilizzabili sono varie. Le esperienze di applicazione concordano nell'evidenziare che esse forniscono risultati simili e pertanto si può ritenere che abbiano prestazioni equivalenti. La regressione lineare multipla con variabili di tipo dummy costituisce, come visto nei capitoli precedenti, la tecnica maggiormente applicata quando le valutazioni fornite dagli intervistati sono misurate su una scala che può essere ragionevolmente ritenuta quantomeno ad intervallo. In ragione della sua ampia diffusione è il metodo illustrato dettagliatamente nel seguito. Un'altra tecnica, nota come LINMAP, si basa invece su un approccio di programmazione lineare ed assume che le valutazioni di preferenza o di gradimento abbiano esclusivamente la valenza di una scala di misura di tipo ordinale. Una terza, denominata MONANOVA, consiste in una procedura di analisi della varianza applicabile quando la variabile oggetto di studio (in questo caso la preferenza) è espressa su una scala ordinale.

12.2 Codifica degli attributi

Con riferimento al modello di regressione lineare multipla, l'espressione codifica degli attributi comporta che ogni modalità di ciascuna caratteristica sia espressa sotto forma di una variabile di tipo dummy; è possibile tuttavia omettere dalla specificazione della relazione lineare una modalità per ogni attributo, selezionabile arbitrariamente.

Pertanto il modello di regressione da utilizzare in questo caso è:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 \quad (9)$$

dove:

Y: indica la valutazione dell'intervistato su ciascun stimolo;

a : è la costante del modello;

X_1: è la variabile di tipo dummy generata dalla modalità Mamacat dell'attributo nome;

X_2: è la variabile di tipo dummy generata dalla modalità Plus dell'attributo nome;

X_3: è la variabile di tipo dummy generata dalla modalità tetrapak dell'attributo tipo di confezione;

X_4: è la variabile di tipo dummy generata dalla modalità scatoletta dell'attributo tipo di confezione;

X_5: è la variabile di tipo dummy generata dalla modalità 0.55 dell'attributo prezzo;

X_6: è la variabile di tipo dummy generata dalla modalità 0.80 dell'attributo prezzo;

X_7: è la variabile di tipo dummy generata dalla modalità sì dell'attributo approvazione dei medici veterinari;

X_8: è la variabile di tipo dummy generata dalla modalità sì dell'attributo garanzia del rimborso.

Come è facile constatare, per evitare la perfetta collinearità tra le variabili esplicative, sono state omesse le modalità Formula 9 per quanto riguarda il nome, vasetto di vetro per il tipo di confezione, 1 euro per il prezzo, no per l'approvazione dell'ordine dei medici veterinari e no per la garanzia del rimborso. Poiché ogni stimolo è dato da una specifica combinazione delle 13 modalità di questi 5 caratteri, può essere riespresso, dopo un'opportuna ricodifica, nei termini delle 8 variabili di tipo dummy appena indicate. Le codifiche assunte da tali variabili, relativamente ad ogni profilo di prodotto/servizio, sono le seguenti: 1 se nello stimolo si presenta proprio la modalità alla quale si riferisce la variabile dummy; 0 se nello stimolo è presente non la modalità alla quale si riferisce la variabile dummy, ma quella che corrisponde ad un'altra dummy inserita nel modello; -1 se nello stimolo è presente la modalità che è stata omessa in sede di specificazione del modello e alla quale non corrisponde quindi alcuna dummy. Con riferimento al carattere nome, ad esempio, si avrà allora X1 = 1 e X2 = 0 se nello stimolo si trova la modalità Mamacat; X1 = 1 e X2 = -1 se nello stimolo si trova al modalità Formula 9; X1 = 0 e X2 = 1 se nello stimolo è presente la modalità Plus.

In maniera analoga si procede per tutte le altre variabili di tipo dummy, ottenendo di conseguenza, per le 18 alternative, le codifiche riprodotte in tabella a pagina seguente:

Tab. 12: Sono evidenziate le codifiche di tutte le variabili dummy dell'esempio in questione.

STIMOLI	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
1	1	0	0	1	-1	-1	-1	1
2	-1	-1	0	1	1	0	1	-1
3	0	1	0	1	0	1	-1	-1
4	1	0	-1	-1	-1	-1	1	-1
5	-1	-1	-1	-1	1	0	-1	1
6	0	1	-1	-1	0	1	-1	-1
7	1	0	1	0	1	0	-1	-1
8	-1	-1	1	0	0	1	1	-1
9	0	1	1	0	-1	-1	-1	1
10	1	0	0	1	0	1	-1	-1
11	-1	-1	0	1	-1	-1	-1	-1
12	0	1	0	1	1	0	1	1
13	1	0	-1	-1	1	0	-1	-1
14	-1	-1	-1	-1	0	1	-1	1
15	0	1	-1	-1	-1	-1	1	-1
16	1	0	1	0	0	1	1	1
17	-1	-1	1	0	-1	-1	-1	-1
18	0	1	1	0	1	0	-1	-1

Tali codifiche non sono altro che i valori assunti dalle variabili esplicative nel modello di regressione lineare multipla. E' opportuno sottolineare che, per ciascun intervistato, tali valori sono identici (in quanto gli stimoli in esame sono sempre gli stessi), mentre può variare ovviamente l'insieme di valori assunti dalla variabile dipendente (le valutazioni di preferenza o di gradimento).

12.3 Stime dei parametri

L'applicazione proposta si risolve nella stima dei parametri del modello di regressione con il metodo dei minimi quadrati ordinari che, come noto, rende minima la somma dei quadrati degli scarti tra valutazioni espresse da ciascun intervistato e punteggi corrispondenti previsti tramite il modello.

Se con il simbolo j, come in precedenza, si indica il generico profilo di prodotto/servizio preso in considerazione, si ha in questo caso:

$$\sum_{j=1}^{18}(Y-\hat{Y})^2 \qquad (10)$$

ogni unità del campione fornisce, infatti, 18 giudizi di gradimento su altrettante alternative. Le stime ottenute rappresentano le utilità parziali associate ad ogni modalità di ciascun attributo.

E' importante notare che il modello può essere stimato sia a partire da dati individuali sia da dati aggregati. Nel primo caso si ottengono tante stime dei parametri quanti sono gli intervistati. Se si segue questa opzione, il successivo e più logico passo di analisi è rappresentato dall'applicazione di un algoritmo di C.L.A. alla matrice di dati dove ogni unità del campione viene descritta dai valori dei parametri stimati con il metodo dei minimi quadrati.

In tal modo si possono individuare nella maniera più appropriata gruppi di consumatori che si distinguono per la diversa importanza assegnata a vari attributi e alle singole modalità di questi. Nel secondo caso, invece, si perviene alla stima di parametri medi, a partire da punteggi della variabile dipendente che si corrispondono ai valori medi aritmetici delle valutazioni fornite da tutti gli intervistati.

Sulla base dei dati di gradimento contenuti nell'ultima colonna della tabella prima tabella in esempio, si è proceduto alla stima del modello di regressione, tenendo conto delle codifiche relative alle variabili di tipo dummy riprodotte nella tabella seguente la relazione stimata è:

$$Y = 64,13 - 2,83X_1 + 4,83X_2 + 19,33X_3 + 1,83X_4 + 18,27X_5 + 2,33X_6 + 2,75X_7 + 10,63X_8 \qquad (11)$$

con un valore dell'indice di determinazione lineare R^2 corretto pari a $0,94$.

Anche il valore del coefficiente di correlazione lineare tra punteggi di gradimento osservati e previsti tramite il modello è molto elevato e quasi prossimo all'unità (vale, infatti, 0,98). I parametri di questa equazione indicano le utilità parth worth cercate.

La codifica applicata alle variabili di tipo dummy è stata disegnata in modo da far sì che la somma delle utilità parziali riferite a tutte le modalità di uno stesso attributo sia pari a zero. Pertanto l'utilità parth worth relativa alla modalità di ogni caratteristica non esplicitamente inclusa nel modello di regressione sotto forma di variabili di tipo dummy può essere calcolata come complemento a zero della somma dei parametri stimati in corrispondenza delle modalità incluse. Ad esempio, l'utilità parziale associata alla modalità Formula 9 dell'attributo nome è pari a:

Utilità parziale (Formula 9) = - [ut. parziale (Mamacat)+ut. parziale (Plus)]= -$[-2,83 + 4,83] = -2,00$

Procedendo in maniera analoga, è immediato determinare gli altri valori; le utilità parth worth in questione sono riportate nella seconda colonna della tabella a seguire:

Tab. 13: Sono evidenziate le utilità dei vari livelli, prima e dopo la codifica.

ATTRIBUTI	Ut. parziali originarie	Ut. parziali trasformate
NOME		
Mamacat	**-2,83**	**0,45**
Formula 9	**-2,00**	**0,47**
Plus	**4,83**	**0,64**
CONFEZIONE		
Scatoletta	**1,83**	**0,57**
Vasetto	**-21,16**	**0,00**
Tetrapack	**19,33**	**1,00**
PREZZO		
0,55 euro	**18,17**	**0,97**
0,80 euro	**2,33**	**0,58**
1 euro	**-20,50**	**0,02**
APPROVAZIONE DELL'ORDINE		
Sì	**2,75**	**0,59**
No	**-2,75**	**0,45**
GARANZIA DEL RIMBORSO		
Sì	**10,63**	**0,79**
No	**-10,63**	**0,26**

Come si può osservare, Plus è il nome prediletto, tetrapack è il tipo di confezione preferito, $0,55$ euro più gradito e sia l'approvazione dell'ordine dei medici veterinari sia la garanzia del rimborso sono modalità desiderate dal consumatore.

E' possibile effettuare un controllo di validità esterna dei risultati ottenuti se è stata programmata fin dall'inizio la rilevazione di valutazioni da impiegare come campione di conferma (holdout sample). Ad ogni consumatore intervistato si può chiedere, infatti, di fornire giudizi di preferenza per qualche altra

alternativa di prodotto/servizio selezionata tra quelle non comprese all'interno del disegno fattoriale frazionato (in genere non più di tre o quattro). I punteggi così raccolti non contribuiscono a determinare l'insieme dei dati su cui avviene la stima dei parametri del modello, e dunque non entrano in gioco nel calcolo delle utilità parziali. Proprio a partire dai valori stimati delle utilità part worth si può ricostruire algebricamente il giudizio di preferenza che spetterebbe, se il modello applicato fosse valido, alle alternative che fanno parte del campione di conferma. Il confronto tra punteggi rilevati sul campo e punteggi previsti sulla base dei parametri stimati per tali alternative mette a disposizione una misura del grado di coerenza del modello per ogni intervistato.

12.4 Comparazione delle utilità parziali delle modalità degli attributi

Al fine di facilitare una comparazione diretta delle utilità parziali in molti casi si preferisce apportare una trasformazione di scala ai valori originari, in modo che tutti quanti siano compresi nell'intervallo $0 - 1$. se indichiamo con V_{ij} l'utilità parziale associata alla modalità j-ma dell'attributo i-mo, con G la più grande delle utilità parziali, con P la più piccola di queste e con U_{ij} l'utilità parziale della modalità j-ma per l'attributo i-mo dopo la trasformazione di scala, si può applicare la seguente formula:

$$U_{ij} = \frac{V_{ij}-P}{G-P}$$

Nella terza colonna della tabella precedente sono riprodotti i valori delle utilità parziali trasformate secondo la procedura appena descritta; la loro rappresentazione grafica compare nella figura a pagina seguente.

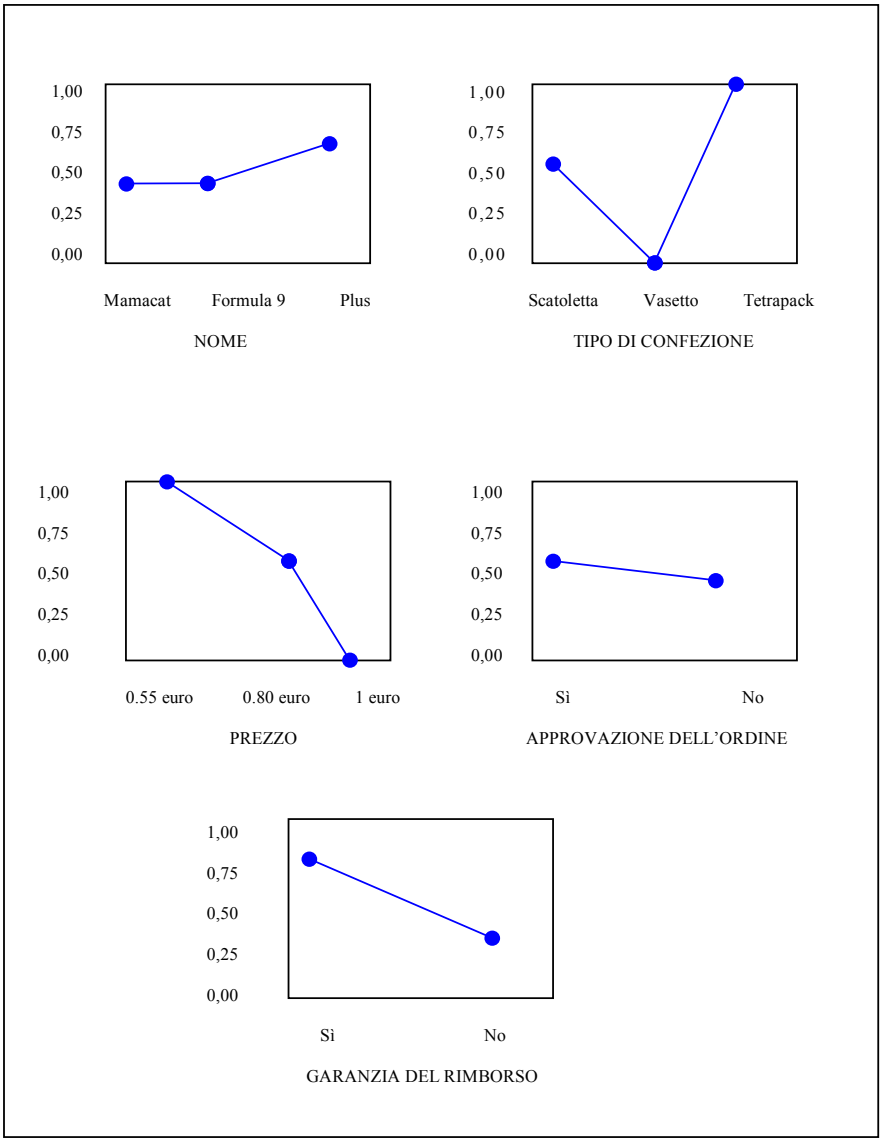


Fig.19: Utilità parziali trasformate

L'importanza relativa do ogni attributo Ai (con il simbolo $IRA_i, i = 1, \ldots, 5$) può essere valutata rapportando la differenza tra i valori di utilità parziale più grande e più piccolo associati alle modalità di quell'attributo e la somma di tali differenze rispetto a tutti gli attributi:

$$IRA_i = \frac{[\text{ut. parziale più grande di } A_i - \text{ut. parziale più piccola di } A_i]}{\sum_{i=1}^{5}[\text{ut. parziale più grande di } A_i - \text{ut. parziale più piccola di } A_i]}$$

Se la scelta del consumatore il esame è fortemente influenzata da una variazione di prezzo da 1 euro a $0,55$ euro (vi è infatti un notevole scarto tra le corrispondenti utilità parziali stimate), tale attributo deve essere considerato più importante di un altro (ad esempio l'approvazione dell'ordine dei medici veterinari) per il quale un cambiamento di modalità produce un minor impatto sul giudizio di gradimento.

I valori di importanza relativa per gli attributi, calcolati nel modo descritto, compaiono nella tabella qui di seguito presentata:

Tab. 14: Importanza relativa degli attributi

ATTRIBUTI	Utilità parziale più grande (1)	Utilità parziale più piccola (2)	Differenza (3) = (1) – (2)	Somma delle differenze	Importanza relativa (5) = (3) / (4)
Nome	4,83	-2,83	7,66	113,58	0,07
Tipo di confezione	19,33	-21,16	40,49	113,58	0,35
Prezzo	18,17	-20,50	38,67	113,58	0,34
Approvazione dell'ordine	2,75	-2,75	5,50	113,58	0,05
Garanzia del rimborso	10,63	-10,63	21,26	113,58	0,19

Nel contesto del caso presentato, il consumatore intervistato assegna l'importanza maggiore al tipo di confezione e al prezzo, valuta di importanza modesta ma non trascurabile la garanzia del rimborso e ritiene praticamente irrilevanti sia il nome sia l'approvazione dell'ordine dei medici veterinari.

Le indicazioni che si traggono dalla lettura di questi dati devono comunque essere interpretate con cautela, soprattutto perché possono essere fortemente influenzate dalle modalità prescelte di rappresentazione di rappresentazione degli attributi.

Se, ad esempio, i livelli di un carattere fossero poco realistici, ossia troppo elevati o bassi, ciò potrebbe contribuire ad alterare artificialmente l'importanza relativa di tale attributo rispetto a tutti gli altri.

La possibilità di assegnare una valutazione di preferenza/utilità totale a tutte le combinazioni degli attributi, anche a quelle non giudicate direttamente dagli intervistati, e di mettere a confronto il desiderio di un consumatore di scambiare elevati livelli di un attributo con più bassi livelli di un altro rende, in definitiva, la C.A. particolarmente efficace.

12.5 Interpretazione dei risultati

Posto che la combinazione preferita tra le 18 prese in considerazione sia, come si è già detto, la dodicesima (Plus, scatoletta, $0,55$€, approvazione dell'ordine dei medici veterinari e garanzia del rimborso), può avere interesse cercare di capire a questo punto se qualcuna delle restanti $(108 - 18) = 90$ combinazioni viene preferita a questa. In virtù del disegno fattoriale frazionato di tipo ortogonale alla base dell'indagine ha senso analizzare anche le combinazioni non valutate sul campo. Osservando con attenzione i valori delle utilità parth worth, ci si rende conto che la preferenza/utilità totale prevista dal modello in corrispondenza del dodicesimo stimolo è pari a $38,21$. Se il tipo di confezione viene cambiato da scatoletta in tetrapack, lasciando inalterate le modalità degli altri quattro attributi, la preferenza/utilità totale sale a $55,71$, che rappresenta la combinazione più gradita in assoluto da quell'intervistato.

Sempre grazie ad un approfondito esame dei risultati dell'applicazione è possibile dare risposta anche ad altri tipi di quesiti. Ci si può domandare, ad esempio, se il consumatore sarebbe disposto a pagare 1 euro anziché $0,80$ euro in cambio dell'inserimento dell'approvazione del prodotto da parte dell'ordine dei medici veterinari, a parità di ogni altro requisito. In questo caso la risposta è negativa, perché passare da un prezzo di $0,80$ euro ad uno di 1 euro, comporta una diminuzione di preferenza/utilità totale pari a $22,83$, mentre aggiungere l'approvazione produce un aumento pari solo a $5,50$, e in definitiva si avrebbe una perdita netta pari a $17,33$.

I passi di analisi descritti devono naturalmente essere replicati per ciascuno dei 150 consumatori appartenenti al campione e per l'insieme dei dati aggregati. Le indicazioni che si possono trarre dall'interpretazione dei risultati dell'applicazione si possono poi tradurre in termini di strategie aziendali e di marketing, con l'obiettivo di calibrare nel modo migliore la messa a punto del profilo della nuova marca da aggiungere alla linea di produzione già esistente e di stimarne la quota di mercato potenzialmente acquisibile.

13 Parametri di valutazione del modello ed esempi di applicazioni di Analisi della Varianza Multifattoriale

Esistono una serie di parametri statistici che permettono di verificare la bontà del modello, tali parametri fanno spesso parte dell'output fornito dai software più comunemente utilizzati per effettuare analisi della varianza.

Vediamo, innanzitutto, come generare una tavola ANOVA con l'utilizzo del software MINITAB.

A differenza del software SPSS, che presenta due tipologie diverse di output: un profilo per ciascun rispondente ed un profilo di sintesi; il software Minitab fornisce come output di un disegno degli esperimenti la Tavola ANOVA con i test relativi all'influenza degli effetti principali e delle interazioni.

In Minitab è possibile creare, con la massima semplicità, una vasta gamma di disegni fattoriali.

Vediamo meglio tramite un esempio di un piano con tre fattori e due livelli, replicato tre volte. Generiamo ora il piano in Minitab:

- una volta entrati nel software dal menù **STAT** scegliere: **DOE Factorial**;

scegliere **Create Factorial Design** Type of Design: 2-level factorial (default generator) con numero di fattori pari a 3;

Design: Full Factorial con 3 repliche;

Dando OK, si ottiene:

StdOrder	RunOrder	CenterPt	Blocks	A	B	C
1	1	1	1	-1	-1	-1
2	2	1	1	1	-1	-1
3	3	1	1	-1	1	-1
4	4	1	1	1	1	-1
5	5	1	1	-1	-1	1
6	6	1	1	1	-1	1
7	7	1	1	-1	1	1
8	8	1	1	1	1	1
9	9	1	1	-1	-1	-1
10	10	1	1	1	-1	-1
11	11	1	1	-1	1	-1
12	12	1	1	1	1	-1
13	13	1	1	-1	-1	1
14	14	1	1	1	-1	1
15	15	1	1	-1	1	1
16	16	1	1	1	1	1
17	17	1	1	-1	-1	-1
18	18	1	1	1	-1	-1
19	19	1	1	-1	1	-1
20	20	1	1	1	1	-1
21	21	1	1	-1	-1	1
22	22	1	1	1	-1	1
23	23	1	1	-1	1	1
24	24	1	1	1	1	1

Si inseriscono ora in una colonna i risultati relativi alle interviste (punteggi da 1 a 10).

Indichiamo la colonna con RISP ed inseriamo i seguenti risultati:2,3,5,4,3,6,7,8, 9,10,2,3,5,6,3,6,7,10,1,1,3,5,2,4.

Per ottenere la tavola ANOVA:

-dal menù **STAT** scegliere: **DOE Factorial**;

scegliere: **Analyze Factorial design**. Scegliamo la variabile risposta e come termini si possono scegliere sia i fattori principali che le interazioni.

-Come risultati si desiderano i coefficienti e la tavola ANOVA.

La prima parte dell'output fornita da Minitab è la seguente:

Term	Effect	Coef	SE Coef	T	P
Constant		4.7917	0.5035	9.52	0.000
A	1.4167	0.7083	0.5035	1.41	0.179
B	-1.9167	-0.9583	0.5035	-1.90	0.075
C	0.0833	0.0417	0.5035	0.08	0.935
A*B	-0.4167	-0.2083	0.5035	-0.41	0.685
A*C	0.5833	0.2917	0.5035	0.58	0.570
B*C	2.2500	1.1250	0.5035	2.23	0.040
A*B*C	0.4167	0.2083	0.5035	0.41	0.685

Il P-value è utilizzato per determinare quale degli effetti compresi nel modello sia statisticamente significativo. Con significativo, da un punto di vista statistico, si intende che la relazione tra il predittore specifico e la variabile dipendente (risposta) non risulta essere di tipo casuale.

Per effettuare una valutazione del valore di p-value, è necessario stabilire un valore di soglia, detto spesso α-value. Come α-value è generalmente usato un valore di 0.05. Se il p-value risulta inferiore o uguale ad α, si conclude che l'effetto è significativo. In caso contrario l'effetto non è significativo.

Una prima valutazione del p-value va fatta, ancor prima che per verificare la significatività dei fattori, per verificare la bontà del modello nel suo complesso. Anche per il modello, infatti, è condotto un F-test che ne verifichi la reale capacità predittiva.

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	34.125	34.1250	11.375	1.87	0.175
2-Way Interactions	3	33.458	33.4583	11.153	1.83	0.182
3-Way Interactions	1	1.042	1.0417	1.042	0.17	0.685
Residual Error	16	97.333	97.3333	6.083		
Pure Error	16	97.333	97.3333	6.083		
Total	23	165.958				

Dalla lettura dell'output otteniamo le seguenti informazioni:

Dal test F vediamo che sia gli effetti principali che le interazioni non risultano significativi, di conseguenza il modello dovrebbe essere scartato;

- Volendo comunque proseguire nella lettura dei risultati (per puro esercizio) vediamo che solo l'interazione di secondo ordine relativa a BC risulta avere effetti significativi;

Nella figura seguente vediamo l'output relativo ad un secondo esempio, tratto dalla guida di MINITAB, sempre in riferimento al modello adottato per un caso di analisi della varianza multifattoriale, questo mette in luce come:

Analyze Factorial Design : Two-Level Factorial Designs : Analysis of Variance Table

▸ P-values

Analysis of Variance for Strength (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Covariates	1	83.767	3.591	3.591	1.58	0.278
Main Effects	4	347.281	304.587	76.147	33.39	0.002
2-Way Interactions	6	20.309	20.309	3.385	1.48	0.366
Residual Error	4	9.121	9.121	2.280		
Total	15	460.478				

Fig.20: Output di Minitab

· gli effetti di interazione inclusi nel modello non siano significativi;

· gli effetti principali sono invece significativi, avendo nel complesso valore di p-value pari a 0.02, a confermare che ci sarà almeno un effetto principale tra quelli considerati che risulterà significativo nel predire il valore della variabile di osservazione;

· i covariates inclusi nel modello non sono significativi, avendo p-value superiore al valore di soglia.

Prima di guardare al valore degli effetti individuali (o di interazione) delle variabili, si deve osservare, come visto in precedenza, nella tabella dei risultati dell'analisi della varianza il valore assunto dal p-value di ciascuno di detti fattori.

Analyze Factorial Design : Two-Level Factorial Designs : Estimated Effects and Coefficient Table

▸ P-values
▸ Effects
▸ Coefficients

Term	Effect	Coef	SE Coef	T	P
Constant		56.043	21.0469	2.66	0.056
MeasTemp		-1.229	0.9793	-1.25	0.278
Material	5.316	2.658	0.6780	3.92	0.017
InjPress	5.645	2.822	0.4008	7.04	0.002
InjTemp	4.355	2.177	0.3783	5.76	0.005
CoolTemp	-3.457	-1.729	0.4198	-4.12	0.015
Material*InjPress	-0.723	-0.361	0.4146	-0.87	0.433
Material*InjTemp	-1.025	-0.512	0.4434	-1.16	0.312
Material*CoolTemp	-0.208	-0.104	0.5099	-0.20	0.848
InjPress*InjTemp	-0.837	-0.419	0.5099	-0.82	0.458
InjPress*CoolTemp	-0.055	-0.027	0.3824	-0.07	0.947
InjTemp*CoolTemp	1.933	0.966	0.3807	2.54	0.064

Fig.21: Tavola ANOVA, output di Minitab

Una volta identificati quelli significativi, si può passare all'esame dei valori assunti dal singolo effetto significativo, partendo da quelli di interazione, dato che influenzeranno anche il modo di interpretare gli effetti principali che vi partecipano.

Nel caso della tavola precedente, per esempio:

i. gli effetti di interazione di secondo ordine contenuti nel modello non risultano significativi;

ii. il modello contiene quattro effetti principali significativi, che possono essere valutati trascurando i loro effetti di interazione; si potrà quindi valutarne il valore di "Effect" per determinarne il relativo peso;

iii. dei due valori di costate sperimentale inseriti nel modello ("Constant" e "MeasTemp", il secondo mostra di non essere significativo agli effetti della determinazione del valore osservato, avendo p-value pari a 0.278, superiore al livello di soglia di 0.05.

Gli Effetti determinano il grado di importanza della singola variabile nell'influenzare il valore di osservazione, ed in particolare:

i. il valore assoluto dell'effetto determina la forza relativa della variabile stessa – più elevato è, e più la variabile è determinante nel costruire il valore finale –;

ii. il segno dell'effetto indica quale livello del fattore comporti un aumento del valore della variabile indipendente; ossia un segno positivi sta ad indicare che ad un livello alto (crescente) della variabile dipendente corrisponde un livello maggiore – a parità di altri fattori – del valore osservato; viceversa nel caso di un segno negativo.

Quindi, nell'esempio:

· i primi tre fattori risultanti significativi mostrano di influenzare positivamente la variabile indipendente, nel senso che, al loro crescere / al passare al livello superiore, aumenta anche il valore dell'osservazione;

· il quarto fattore è invece in relazione di proporzionalità inversa con la variabile osservata;

· si rileva inoltre come l'aumento del quarto fattore abbia un effetto meno negativo sulla variabile osservata di una diminuzione di uno qualsiasi degli altri tre (in particolare del secondo, che risulta essere il più determinante, avendo valore assoluto massimo).

In fine, i valori dei Coefficienti rappresentano appunto i coefficienti da dare alle variabili – opportunamente codificate – nell'equazione del modello, al fine di predire il valore dell'osservazione a partire dai valori delle singole variabili.

Va tenuto ben presente come nell'equazione del modello, che nel caso dell'esempio riportato assumerà per esempio la forma:

$$Strength = 56.043 - 1.229(MeasTemp) + 2.658(Material) + 2.822(InjPress) + 2.177(InjTemp) - 1.729(CoolTemp) - 0.361(Material * InjPress) - 0.512(Material * InjTemp) - 0.104(Material * CoolTemp) - 0.419(InjPress * InjTemp)$$

$$-0.027(InjPress * CoolTemp) + 0.966(InjTemp * CoolTem)$$

dovranno essere inseriti i valori nello stesso modo in cui sono stati codificati per essere sottoposti all'analisi (e quindi al pacchetto statistico con il quale questa è stata condotta), e non inserendo direttamente i valori delle variabili.

14 METODOLOGIE DI CONJOINT ANALYSIS: un confronto

14.1 INTRODUZIONE

La Conjoint Analysis è una delle tecniche statistiche di analisi multivariata che hanno conosciuto maggiore successo applicativo nell'ultimo ventennio, in particolar modo nell'ambito delle ricerche di marketing. Questa metodologia, che trova una propria collocazione naturale tra i modelli di segmentazione flessibile del mercato, viene adottata per comprendere ed investigare come i consumatori valutino gli attributi di indeterminato prodotto/servizio o marche, già esistenti o in fase di progettazione, al fine di identificare segmenti di mercato composti da clienti con profili di risposta simili in termini di preferenza. La tecnica basa i suoi fondamenti concettuali nelle considerazioni sviluppate da Lancaster in relazione alla teoria del consumatore, secondo le quali l'utilità d'uso di un bene deriva dalle singole caratteristiche che lo compongono. Ne consegue che l'utilità di un prodotto/servizio può essere scomposta in diverse utilità aventi come origine le differenti caratteristiche del prodotto/servizio medesimo.

Ed è proprio su questo concetto di natura decompositiva dell'utilità globale di un bene che si basa la metodologia d'interesse.

La tecnica consiste, infatti, nello scomporre i giudizi globali espressi dai consumatori in relazione ad un set di alternative complesse, in valori di utilità distinti per ogni caratteristica del prodotto/servizio, attraverso i quali possono essere ricostruite le valutazioni globali. In termini pratici, questo significa che al campione opportunamente selezionato vengono somministrate un insieme di schede, contenenti ciascuna la descrizione di uno o più profili di prodotto, a seconda della metodologia di Conjoint Analysis utilizzata, dove ogni profilo è basato sulle varie combinazioni delle caratteristiche considerate, dette fattori o attributi, del prodotto in analisi.

Dai giudizi espressi dai consumatori nei confronti delle diverse descrizioni di prodotto, la tecnica permette di determinare l'importanza dei singoli attributi nel processo decisionale del valutatore ed identifica e valuta, per ciascuna caratteristica, quali siano le migliori alternative o livelli.

Esistono diverse metodologie per la raccolta dei dati in uno studio di Conjoint Analysis, vengono di seguito elencate e successivamente analizzate le principali:

1_ Trade Off Matrix;
2_ Full Profile;
3_ Hybrid Conjoint Analysis;
4_ Adaptive Conjoint Analysis (ACA);

5_ Choice-based Conjoint (CBC);

L'ordine in cui sono presentate è di tipo cronologico, dalle prime metodologie sviluppate sino alle più recenti.

14.2 TRADE OFF MATRIX

La metodologia Trade Off Matrix viene anche denominata "Due attributi per volta" ed è la prima metodologia di Conjoint Analysis che è stata sviluppata. Essa consiste nel sottoporre agli intervistati una sequenza di tabelle nelle quali vengono considerate tutte le possibili combinazioni dei livelli di due attributi per volta, corrispondenti appunto alle celle delle matrici. Il rispondente è chiamato ad ordinare le suddette celle secondo un grado di preferenza che accorda a ciascuna coppia di livelli.

Il Trade Off Matrix è la risposta al problema del ranking-order, dove veniva considerato un attributo alla volta, e non due attributi congiunti.

Vediamo meglio tramite un esempio relativo a delle confezioni di birra. Gli attributi considerati, ciascuno su tre livelli, sono la confezione ed il grado alcolico.

Le prime indagini di mercato venivano svolte chiedendo ai rispondenti di ordinare in rank-lists le varie caratteristiche del prodotto. Nel caso in esame verrebbero proposte le due liste seguenti, in modo consequenziale, chiedendo di ordinare per ciascuna di esse le opzioni disponibili.

Tab. 15: Esempio d rank-lists

CONTENITORE		GRADO ALCOOLICO	
Bottiglia di vetro	?	3% alcool	?
Bottigia di plastica	?	4% alcool	?
Lattina	?	5% alcool	?

Supponiamo ora che un rispondente abbia assegnato i seguenti valori in base alle proprie preferenze:

Tab. 16: Rank-lists con valutazione

CONTENITORE		GRADO ALCOOLICO	
Bottiglia di vetro	1	3% alcool	3
Bottigia di plastica	3	4% alcool	2
Lattina	2	5% alcool	1

In questo caso risulta che il rispondente ha espresso la sua preferenza per le birre in bottiglie di vetro ed ad alto contenuto alcolico. Il problema nasce nell'individuazione della seconda preferenza, in quanto sia la birra in bottiglia di vetro con un grado alcolico del 4% che le birre in lattina con un grado alcolico del 5% ottengono lo stesso punteggio corrispondente al secondo posto in graduatoria.

Il Trade Off Matrix permette, tramite l'analisi congiunta di due attributi, di individuare non solo la prima scelta, ma anche tutte le altre. Gli attributi ed

i relativi livelli non vengono in caso presentati in tabelle distinte, ma tramite una matrice dove le righe e le colonne corrispondono ai livelli dei due attributi considerati. Nell'esempio in esame la matrice risulta la seguente:

Tab. 16: Esempio d Trade-off con due matrici

	3% alcool	4% alcool	5% alcool
Bottiglia di vetro	?	?	?
Bottigia di plastica	?	?	?
Lattina	?	?	?

Al rispondente viene quindi chiesto di compilare la matrice inserendo in ciascuna cella la relativa valutazione, dal valore "uno" per l'opzione maggiormente preferita sino a "nove", in questo caso, nella cella corrispondente all'opzione considerata la peggiore. Un ipotetico rispondente potrebbe rispondere nel modo seguente:

Tab. 17: Matrice Trade-off completa di ranking del rispondente

	3% alcool	4% alcool	5% alcool
Bottiglia di vetro	5	3	1
Bottigia di plastica	9	8	7
Lattina	6	4	2

Risulta immediato dalle risposte ottenute che il valutatore attribuisce maggiore importanza al grado alcolico del 5% rispetto al contenitore inoltre, si può notare come il rispondente non apprezzi le bottiglie in plastica. Tramite l'utilizzo della matrice di Trade Off non occorre, quindi, ispezionare nuovamente i dati espressi da ogni rispondente per comprendere i risultati della ricerca.

Per trasformare le preferenze del rispondente in set di utilità è possibile utilizzare un approccio basato sul modello di regressione. Nella tabella a seguire vengono presentati i risultati delle stime delle utilità per ogni singolo livello di ciascun attributo ottenute tramite l'utilizzo della regressione e dello scaling, per il rispondente considerato.

Tab. 18: Utilità per singolo livello di attributo

CONTENITORE		GRADO ALCOOLICO	
Bottiglia di vetro	214	3% alcool	0
Bottigia di plastica	0	4% alcool	71
Lattina	171	5% alcool	143

L'utilità totale, di questo valutatore, per ogni combinazione può essere stimata sommando le utilità dei livelli dell'attributo contenitore e le utilità del secondo attributo grado alcolico.

La tabella seguente mostra come le utilità per i singoli livelli possono essere combinate per calcolare il valore delle utilità parziali.

Tab. 19: Calcolo delle utilità parziali

	3% alcool utilità=0	4% alcool utilità=71	5% alcool utilità=143
Bottiglia di vetro	214 (5°)	285 (3°)	357 (1°)
Bottigia di plastica	0 (9°)	71(8°)	143 (7°)
Lattina	171 (6°)	242 (4°)	314 (2°)

Il Trade Off Matrix presenta le seguenti problematiche:

- La matrice risulta una strumento di richiesta delle preferenze ai rispondenti, completamente artificiale. Inoltre, se i livelli degli attributi considerati si presentano numerosi, la tabella risultante rappresenterà un compito particolarmente oneroso per il rispondente.
- Risultano particolarmente rari i casi in cui si presentano solo due attributi. L'aggiunta di un solo attributo comporterebbe, a titolo di esempio, la presentazione di tre matrici, rendendo molto più lungo e laborioso il ruolo del valutatore.
- Questa metodologia sembra non idonea ad un utilizzo in ambiente internet, ad oggi, infatti, non sono stati sviluppati approcci per l'utilizzo di questa metodologia sul web.

Tab. 20: Vantaggi e svantaggi del Trade -Off Matrix

TRADE OFF MATRIX	
VANTAGGI	**SVANTAGGI**
Notevole semplicità di compilazione da parte dell'intervistato.	Eccessiva mancanza di realismo, il metodo non suscita un reale interesse nel rispondente: si valutano semplici combinazioni di caratteristiche, e non un vero prodotto.
Facile e veloce elaborazione dei dati da parte dei ricercatori.	Il compito diventa lungo e noioso quando le tabelle proposte sono in numero elevato.
La preparazione dell'intervista è elementare, visto che si tratta di costruire semplici tabelle, quindi non richiede addestramenti particolari per il personale che svolge l'intervista.	Rispondente può svolgere il compito in modo meccanico, abbassando di molto la qualità e attendibilità delle risposte, in quanto non gli viene presentato un prodotto reale da valutare.
Non servono molte spiegazioni per chiarire al rispondente come svolgere il suo semplice compito.	Non fornisce giudizi di preferenza globali.
Permette di valutare congiuntamente due attributi alla volta.	Si può utilizzare solo con un numero limitato di attributi e di livelli.
La stima delle utilità si basa su un algoritmo molto facile e immediato.	Non è utilizzabile via web e non è adattato per l'ambiente internet.

14.3 FULL PROFILE

E' il metodo più diffuso di raccolta dei dati e si basa sulla somministrazione ai rispondenti di un insieme di scenari completi descritti, cioè, facendo riferimento a tutti gli attributi presenti contemporaneamente. Questi sono combinati tra loro seguendo un disegno sperimentale precedentemente costruito dal ricercatore. Questi scenari completi o profili, o "prop cards" vengono dapprima ordinati dal rispondente secondo le sue preferenze, attraverso un "rank order", in questo modo si costruisce una graduatoria dei profili del prodotto/servizio sottoposto a giudizio. Successivamente il rispondente stima ogni "prop card" su una scala di probabilità di acquisto, ad esempio da 0-100, secondo un "rating scale", attraverso il quale il soggetto assegna un punteggio ad ogni profilo. Successivamente, con l'ausilio di metodi di stima come LINMAP o MONANOVA, si elaborano e si ricavano i dati cercati. Questo metodo può essere una buona risposta alle limitazioni del Trade-off matrix, precedentemente presentato. Viene di seguito ripreso l'esempio con riferimento all'indagine sulla vendita della birra, in cui gli attributi sono:

Tab. 21: Attributi e livelli dell'indagine svolta

GRADO ALCOOLICO			CONTENITORE			COLORE	
5% alcool	4% alcool	3% alcool	Bottiglia di vetro	Bottiglia di plastica	Lattina	Birra scura	Birra chiara

Due possibili esempi di profili da somministrare al rispondente sono:

Tab. 22: Esempio di profili

PRODOTTO A	PRODOTTO B
Lattina	Bottiglia di plastica
Birra scura	Birra chiara
4% alcool	5% alcool

Con il Full Profile il numero di potenziali profili cresce in modo estremamente veloce con il numero di attributi e di livelli. Con riferimento al caso presentato, figurano due attributi con tre livelli ciascuno ed un attributo con due livelli, in totale si dovranno somministrare 18 profili (= 32 * 21) per rappresentare tutte le possibili opzioni. Con l'ausilio di idonee tecniche statistiche non risulta necessario presentare al rispondente ogni combinazione ottenuta.

In progetti caratterizzati da un elevato numero di attributi, è possibile costruire un piano di profili parziali avente la stessa efficienza di un piano completo,questo consente di ridurre notevolmente il numero di profili che saranno poi aggetto di analisi da parte del valutatore. Questi disegni parziali permettono ad ogni attributo e livello di essere somministrato al rispondente con sufficiente rotazione ed assicura che possa essere stimata l'utilità di ogni livello. Il numero di carte necessario per una corretta analisi è determinato sulla base di :

- numero totale di livelli (di tutti gli attributi);
- la valutazione del sistema;
- se vi è la necessità di calcolare "higher-order effects", che descrivono il modo in cui alcuni attributi interagiscono tra di loro .

Le maggiori interazioni volute dai ricercatori richiedono la valutazione di un maggior numero di profili ed il sistema di valutazione varia in relazione al metodo utilizzato per ordinare le carte: "pair-wise", "sorted" or "rated".

Il Full Profile può facilmente essere applicato in web usando sia il "pair-wise comparisons" o "rating tasks" senza l'uso di specifici software di "data collection".

15 Tab. 23: Vantaggi e svantaggi del Full-Profile

FULL PROFILE	
VANTAGGI	SVANTAGGI
Permette sia il rating che il ranking di profili. L'importanza degli attributi e dei livelli è dedotta indirettamente.	Pericolo di "Information overload", cioè si può sovraccaricare l'intervistato con troppe informazioni, e con un numero eccessivo di profili somministrati, creando confusione che si ripercuote sulle risposte.
Si ha la possibilità di utilizzare più tipi di giudizi di preferenza, come l'intenzione di acquisto e la probabilità d'acquisto.	L'aumento del numero degli attributi porta ad un incremento esponenziale di profili, che va a diminuire la qualità delle risposte e la loro variabilità.
Permette il "trade-off" tra gli attributi più significativi.	Richiede l'uso, anche se minimo, di un software per la regressione multipla.
Ha alti potere esplicativi e predittivi.	Richiede un addestramento specifico per la progettazione, per chi implementa l'analisi e interpretazione dei dati.
Approccio decompositivo.	Richiede tempi di compilazione e di analisi molto lunghi.
Fornisce dei giudizi di preferenza globali.	Il rispondente necessita di molte informazioni e chiarimenti per svolgere il suo compito.
I rispondenti valutano dei concetti di prodotti completi: è un approccio molto realistico, valuta dei prodotti e non singole caratteristiche di prodotto.	La costruzione del questionario richiede molto tempo e molte valutazioni per lo studio dell'ordine di presentazione dei livelli elencati nei profili e della sequenza di somministrazione dei profili.
E' dotato di un'alta convergenza di validità per la stima delle "part-worth" e dell'"importance ratings"	Il compito del rispondente è molto impegnativo all'aumentare del numero degli attributi e dei livelli, serve più tempo per la lettura dei profili e per la risposta, sia essa in termini di ordinamento o di valutazione attraverso un punteggio.
I rispondenti non sono influenzati dalla desiderabilità sociale e dalle inclinazioni.	Tende a far aumentare il numero dei fattori e quindi fornisce informazioni in più; l'intervistato tende così a semplificare il processo di scelta e a concentrarsi solo su alcuni fattori, mentre nella realtà dovrebbe considerarli tutti.
Utilizzazione estesa in accademia e in industria.	L'ordine con il quale i fattori sono elencati potrebbe avere effetti sulla valutazione del consumatore.

15.1 HYBRID CANJOINT ANALYSIS

Un problema pratico nelle applicazioni della CA è l'onerosità del compito richiesto ai rispondenti ed è proprio per far fronte a questo inconveniente che recentemente sono stati introdotti i modelli ibridi di conjoint analysis.

Gli hybrid models sviluppati, quindi, per ridurre lo sforzo e il tempo della raccolta dati nel caso di un ampio numero di attributi, combinano le caratteristiche del self-explication task con le caratteristiche della conjoint analysis (CA) più tradizionale.

Negli anni '70 le principali procedure di raccolta dati erano il tradeoff matrix ed il full profile. Ciascuna di queste metodologie ha comporta un numero relativamente elevato di valutazioni del rispondente per assicurare la stima delle funzioni di utilità individuali. Quindi per il trade-off e il full profile diventa problematico manipolare un grande numero di attributi e di livelli all'interno degli attributi. Inoltre, negli ultimi anni è cresciuto l'interesse commerciale per la misura degli effetti di interazione (interaction effect) per classi di prodotti in cui considerazioni di tipo sensoriale ed estetico sono importanti (food,abbigliamento...). Il calcolo degli effetti di interazione necessitano di stime di molti parametri e di conseguenza subentra il rischio eccessivo aumento della complessità e del tempo dell'intervista.

I modelli ibridi di conjoint analysis sono progettati espressamente per semplificare il compito del rispondente nei casi in cui il numero di stime sia elevato. Originariamente sviluppati per procedure full profile, gli hybrid models possono essere adattati anche ad altri approcci. L'idea centrale di questi modelli consiste nel combinare la semplicità del self-explicated task (tecnica composizionale) con una metodologie di CA pura e sviluppare funzioni di utilità individuali che considerino le differenze fra individui.

Per ogni rispondente devono essere raccolti tre tipi diversi di dati :

- i valori di desiderabilità dei livelli di ciascun attributo;
- l'importanza degli attributi singoli;
- le valutazioni su un limitato numero di profili.

L'idea principale consiste nell'individuare le funzioni di utilità individuali in cui alcuni aspetti (parametri) di esse sono misurati a livello individuale, mentre altri a livello di campione totale. I modelli ibridi inizialmente utilizzavano profili completi nel secondo stadio del questionario e limitavano a 8-9 il numero dei profili stessi da valutare, in contrasto con il tradizionale approccio di CA che prevedeva valutazioni su molti più profili (anche 30).

Ma, per semplificare ulteriormente il compito del rispondente, si iniziò a ridurre anche il numero di attributi che variavano nel set di profili e, in molti casi, si adottò nell'ambito delle tecniche ibride il partial-profile (ACA).

In sintesi, con questa terminologia si intende indicare un'insieme di tecniche che combinano il SELF-EXPLICATED TASK (compositional modeling) e la valutazione di un SUBSET DI CARDS del FULL PROFILE (decompositional modeling), (Green, Goldberg and Montemayar 1981).

Il SELF EXPLICATION TASK, o SELF-EXPLICATED PREFERENCE, è una tecnica compositiva come il casemap procedure (Srinivasan 1988). Con questo metodo non si ottiene una valutazione completa del prodotto o del servizio analizzato. Ogni rispondente deve valutare la desiderabilità di ogni livello di attributo su una scala di range da 0 a 10, che significa 0 (= non accettabile) 10 (=il preferito). Successivamente deve indicare l'importanza relativa degli attributi suddividendo 100 punti attraverso tutti gli attributi presenti.

Riassumendo si ha un "rating" dei singoli livelli (0-10) in base alla desiderabilità e un "rating scale" dei singoli attributi (0-100) a livello di importanza relativa.

Successivamente, sulla base delle scelte di desiderabilità e di importanza di ogni soggetto, il ricercatore definisce un subset di profili da sottoporre alla sua valutazione.

Ambedue i passaggi descritti sono svolti tramite l'utilizzo di personal computer con software specifici.

Esistono diversi approcci per hybrid conjoint modeling in base a:

- Tipo di scala di misurazione delle preferenze (ranking, rating,..);
- Livello di aggregazione (individuale, gruppo, intero campione);
- Caratteristiche del profilo (completo o parziale);
- Tipo di procedura di stima (OLS, LOGIT, PROBIT, LINMAP,...).

MODELLI IBRIDI TRADIZIONALI

Comparsi sulla scena mondiale verso la fine degli anni '70, i primi modelli combinavano self-explicated data e valutazioni su profili completi per stimare le utilità parziali.

Tutt'ora, nel primo stadio dei modelli ibridi più tradizionali i rispondenti devono dare un punteggio a ogni livello di attributo su una scala di desiderabilità (esempio:0-10) e allocare un certo numero di punti (esempio:100 punti) fra gli attributi esprimendo l'importanza di ciascuno di essi.

Nel secondo stadio i rispondenti ricevono un set di full-profile relativamente piccolo (8-10 profili). Nella maggior parte dei casi all'intervistato è chiesto di esaminare i profili, classificarli in categorie e poi ordinarli all'interno di ogni categoria.

Alla fine, il rispondente deve esprimere la probabilità di acquisto (scala 0-100) dei profili.

NUOVI MODELLI IBRIDI

Recentemente Srinivasan e Park (1997) hanno proposto un nuovo approccio chiamato CUSTOMIZED CONJOINT ANALYSIS (CCA) che prevede alti costi e tempi lunghi per la raccolta dei dati.

La CCA può essere suddivisa nelle seguenti fasi:

1. i rispondenti inizialmente identificano i livelli di attributo inaccettabili e successivamente esprimono le loro preferenze per i livelli e l'importanza di ogni attributo;

2. basandosi sui risultati della prima fase vengono identificati dagli analisti i "core attributes", cioè gli attributi più importanti per gli intervistati stessi. I "core attributes" sono inclusi in una full-profile conjoint analysis.

Poiché gli attributi principali differiscono da rispondente a rispondente, il set di stimoli somministrati (8-16 profili) è personalizzato. La scala di misurazioni utilizzata è un rating o un ranking.

3. nella terza fase, detta di calibrazione, al rispondente viene chiesto di valutare (rating o ranking) alcuni profili supplementari.

Quindi, anche la CCA combina un self-explication task con una conjoint analysis più tradizionale.

Solitamente l'intero questionario viene somministrato in 2 sessioni, con uno scarto temporale di 2 settimane fra la prima e la seconda sessione stessa, per dare

all'analista la possibilità di identificare i "core attributes" per ogni soggetto.

Uno svantaggio di questa metodologia è rappresentato dal costo in termini di tempo e denaro poichè richiede due interviste separate e la costruzione di cards con profili personalizzati per ogni rispondente.

Sorprendentemente, per quanto riguarda la validità predittiva della Customized Conjoint Analysis, Srinivasan e Park (1997) indicano che essa è inferiore a quella del self-explicated approach, ma in questa sede non è opportuno approfondire tali risultati.

Una estensione della CCA è il CUSTOMIZED COMPUTERIZED CONJOINT ANALYSIS trattata da Hensel-Börner e Sattler nel loro articolo "Validity of the Customized Computerized Conjoint Analysis".

Questa metodologia è la versione avanzata della CCA e sfrutta l'intervista a computer di tipo adattativi, che viene eseguita in una unica sessione per tutti i rispondenti.

Tab. 24: Vantaggi e svantaggi del Hybrid Conjoint Analysis

HYBRID CONJOINT ANALYSIS	
VANTAGGI	**SVANTAGGI**
Gestisce un numero medio di attributi tra i 15 e i 20 attributi.	Richiede l'uso di software specifici per l'elaborazione.
Elevato potere predittivo ed esplicativo.	Si devono sostenere elevati costi a livello di investimenti iniziali per l'acquisto dei pacchetti software.
Non richiede un elevato sforzo per il rispondente.	Richiede l'uso di personale specializzato e particolarmente addestrato per l'utilizzo del software.
Propone un numero molto ridotto di profili al rispondente.	
La progettazione dello studio richiede meno tempo.	
Il rispondente valuta un subset di profili.	
Approccio realistico.	

15.2 ADAPTIVE CONJOINT ANALYSIS (ACA)

Questa metodologia fa parte delle tecniche "hybrid" sviluppate dalla Sawtooth Software (Johnson 1987). Ad ogni rispondente viene chiesto di esprimere la propria preferenza in modo sequenziale, su opportune combinazioni di attributi

messe a confronto, svolgendo un SELF EXPLICATION TASK si valuta poi un set parziale di profili, proposti due alla volta.

Il rispondente valuta ogni coppia di profili parziali su un scala di confronto o graduatoria chiamata PAIRED COMPARISON SCALE. I due steps vengono descritti in dettaglio di seguito:

1- Il SELF EXPLICATION TASK, o SELF-EXPLICATED PREFERENCE, è una tecnica compositiva come il casemap procedure (Srinivasan 1988). Con questo metodo non si ottiene una valutazione completa del prodotto o del servizio analizzato. Ogni rispondente deve valutare la desiderabilità di ogni livello di attributo su una scala di range da 0 a 10, che significa 0 (= non accettabile) 10 (=il preferito). Successivamente deve indicare l'importanza relativa degli attributi suddividendo 100 punti attraverso tutti gli attributi presenti. Riassumendo si ha un "rating" dei singoli livelli (0-10) in base alla desiderabilità e un "rating scale" dei singoli attributi (0-100) a livello di importanza relativa.

2- Il PAIRED COMPARISON, non è altro che un confronto di due profili di prodotto presentati contemporaneamente. Il rispondente ne sceglie uno e l'altro viene scartato. In base a questa scelta, e alla preferenza espressa, viene definita l'intervista successiva, che viene così personalizzata. Il vantaggio, in questo caso, è che il rispondente non valuta profili descritti da tutti gli attributi e i livelli, ma solo da quelli che nelle sue risposte precedenti sono risultati i più preferiti e importanti.

Tutti e due i punti descritti sono amministrati da un computer attraverso il pacchetto software ACA.

Il termine adaptive si riferisce al fatto che il computer gestisce l'intervista personalizzandola per ogni rispondente. Le due caratteristiche importanti sono:

1- Consente all'analista di progettare una intervista interattiva, grazie all'uso del personal computer e di gestire l'intervista con i rispondenti in modo automatizzato. L'intervista può presentare un numero elevato di attributi e di livelli, prestando particolare attenzione a quelli che i rispondenti considerano i più desiderabili e i più importanti. L'intervista è strutturata in maniera sequenziale, in quanto le utilità dei rispondenti sono continuamente ristimate man mano che le domande vengono presentate. Ogni risposta aggiuntiva va ad incrementare le informazioni a disposizione. Le utilità finali sono calcolate solo una volta che l'intervista è completata.

2- Consente all'analista di simulare le preferenze dei potenziali clienti per un prodotto nuovo o modificato. L'ACA Simulator può essere utile per esplorare gli scenari di tipo what if, come ad esempio, cambiamenti di prezzo, definizione di nuovi prodotti o attività di marketing. L'analista opera specificando ciascun livello di prodotto su ciascun attributo e le utilità dei rispondenti sono a loro volta usate per stimare le preferenze o la verosimiglianza per ciascun prodotto.

Tab. 25: Vantaggi e svantaggi dell'Adaptive Conjoint Analysis

ADAPTIVE CONJOINT ANALYSIS	
VANTAGGI	**SVANTAGGI**
Interviste personalizzate, che migliorano la qualità delle risposte.	Richiede l'uso di software per l'impostazione dello studio e per l'analisi dei dati.
Non richiede uno sforzo elevato da parte del rispondente.	Richiede un addestramento dei ricercatori che applicano questo metodo.
Viene notevolmente ridotto il numero di profili da somministrare al rispondente.	Richiede tempi molto lunghi per l'elaborazione a computer e per le interviste.
I risultati sono di elevata qualità ed è un metodo con elevato potere esplicativo.	Il costo è molto elevato, a livello di investimento iniziale e di costi di personale specializzato.
Elevato controllo sull'andamento dei lavori.	
Ingloba nello studio un numero elevato di attributi e livelli.	
Permette di variare livelli di attributi e di capire automaticamente l'incidenza di questa variazione.	

15.3 CHOICE-BASED CONJOINT (CBC)

La Choice-based Conjoint (d'ora in poi CBC) rappresenta una metodologia relativamente nuova per condurre ricerche di mercato, e rientra in uno degli ultimi sviluppi della CA, considerato che la sua vera e propria diffusione si è verificata in ambito industriale solo negli ultimi dieci anni. La CBC sfrutta le basi teoriche della Conjoint Analysis tradizionale ma, a differenza di questa, richiede ai rispondenti di segnalare la preferenza su un profilo di prodotto/servizio tra un insieme di alternative simili (choice task). Risulta importante notare che tra le varie opzioni di scelta può essere utilizzata anche l'opzione "none", cioè nessuna delle offerte risulta gradita.

Nella maggior parte dei casi questa metodologia è applicata a livello aggregato e la stima delle utilità parziali avviene, a differenza dagli approcci più tradizionali di CA, tramite le tecniche logit, probit o Hierarchical Bayes.

Di seguito si propone l'esempio (inerente al settore dei televisori) di un tipico choice task comprendente l'opzione "None". Ogni televisore è descritto attraverso i seguenti attributi: Marca (JVC, RCA, SONY), Dimensione dello schermo in pollici (25", 26", 27"), Qualità del suono (Mono sound, Stereo sound, Surround sound).

Come si può notare, ogni set di scelta è costituito da un numero variabile

di concetti di prodotto, ciascuno descritto attraverso tutti gli attributi considerati. Il compito del rispondente consiste semplicemente nell'indicare il profilo di prodotto che reputa migliore. Un questionario di CBC prevede siano somministrati un certo numero di questi set di scelta. Raccolte le risposte di tutti gli intervistati queste si analizzano tramite tecniche statistiche come il Logit. Dal punto di vista applicativo, la Choice Based Conjoint Analysis è una buona metodologia per produrre risultati relativamente precisi quando esistono pochi attributi. Infatti, per come è strutturata questa tecnica, un numero troppo elevato di caratteristiche rischierebbe di indurre confusione nell'individuo.

Se si prevede la presenza d'interazioni fra gli attributi, la CBC è senza dubbio la metodologia più conveniente. La presente tecnica assume una certa omogeneità di risposta da parte degli individui ed è quindi adatta principalmente ad analizzare i dati a livello aggregato.

Due importanti decisioni devono essere prese al momento della costruzione dei choice tasks:

1. quanti concetti presentare per ogni choice set;
2. quanti concetti presentare in totale al rispondente.

Troppe domande possono produrre risultati distorti perché il rispondente presenta segni di noia o fatica, poche domande riducono la precisione dei risultati.

Recenti ricerche hanno mostrato che i rispondenti sono in grado di "processare" informazioni su molti concetti ma è opportuno mantenersi attorno ai 3-5 concetti per task e ai 12-20 task complessivi.

Tab. 26: Esempio di Choice Set

QUALE TV PREFERISCI?			
1	2	3	4
JVC	RCA	SONY	NONE
27"	25"	26"	Io non scelgo
Surround Sound	Stereo Suond	Mono Sound	Nessuna di queste
Euro 30	Euro 350	Euro 400	

La CBC si avvale di una raccolta dati molto differente rispetto alla CA tradizionale, ovviamente questo comporta che anche l'analisi degli stessi risulti diversa. Il questionario per un'analisi CBC viene solitamente somministrato tramite PC, ma questo non esclude la possibilità di uno sviluppo dello stesso per via cartacea. La CBC risulta la metodologia di CA che più facilmente può utilizzare per la raccolta dati la "Web based research".

Tab. 26: Vantaggi e svantaggi del Conjoint Based Conjoint Analysis

CONJOINT BASED CONJOINT ANALYSIS	
VANTAGGI	**SVANTAGGI**
Approccio realistico, presenta al rispondente delle descrizioni complete di prodotto.	Molto costoso a livello di investimenti iniziali.
Compito ridotto e più facile per il rispondente.	Tempi lunghi per l'elaborazione e per la somministrazione dei "choice set".
Individua i profili migliori in base alle preferenze dei consumatori.	Personale specializzato in grado di riuscire ad implementare il metodo a mano o con i software.
Elevata capacità nel calcolo degli "higher-order effects".	La tecnica non produce un set di utilità per ogni rispondente.
Viene calcolato anche un livello "none".	Può manipolare un numero abbastanza limitato di attributi.
La scelta rispecchia quello che accade nel mondo reale.	Se non si hanno tutte le utilità è impossibile applicare la segmentazione.
Si richiede al rispondente di fare una scelta decisiva, di "pick" una proposta o di rifiutarle tutte.	Non è possibile utilizzare "hold-out cards" per accertare la correttezza delle risposte date.
Facilmente implementabile via web.	Per avere buoni risultati i rispondenti sono convocati in particolari sedi e retribuiti.