

BIBLIOTECA DI SCIENZE STATISTICHE

SERVIZIO BIBLIOTECARIO NAZIONALE

BID P0V09/2950 BID

ACQ. 265/03 INV. 83225

COLL. 5-Coll. WP. 4/2003

**Adaptive and full profile
conjoint analysis: alcune
considerazioni e software
statistici a confronto**

R. Arboretti Giancristofaro

2003.4

**Dipartimento di Scienze Statistiche
Università degli Studi
Via C. Battisti 241-243
35121 Padova**

Aprile 2003

ADAPTIVE AND FULL PROFILE CONJOINT ANALYSIS: ALCUNE CONSIDERAZIONI E SOFTWARE STATISTICI A CONFRONTO

Rosa Arboretti Giancristofaro

Dipartimento di Scienze Statistiche, Università di Padova

Riassunto

Le tecniche statistiche multivariate, specie a partire dalla fine degli anni '70, sono divenute uno strumento fondamentale per l'analisi dei dati relativi a studi di mercato. In molti casi, proprio il particolare ambito di applicazione ha suggerito lo sviluppo di tecniche specifiche e di strategie particolari. La Conjoint Analysis (C.A.) rientra in questa categoria e si delinea come un modello di preferenze multiattributivo, che include una serie di tecniche statistiche, che consentono al ricercatore di trasformare le risposte soggettive date dagli intervistati in stime di utilità di livelli di attributi di prodotti o servizi.

Questo lavoro si concentra su un'analisi critico-qualitativa della Conjoint Analysis con riferimento a due tipi di C.A.: quella adattativa e quella full-profile. Il confronto avviene anche mediante l'analisi critica di due software di riferimento: ACA e SPSS-Conjoint, al fine di meglio comprenderne gli algoritmi che li regolano ed evidenziare limiti e vantaggi delle due strategie di C.A..

1. Introduzione

Con il termine Conjoint Analysis (C.A.) si indica un insieme di metodologie a contenuto in prevalenza statistico finalizzate allo studio dei modelli di scelta dei consumatori a partire da giudizi di preferenza espressi da questi ultimi relativamente a diversi profili di un prodotto/servizio che si intende sviluppare (Gustafson et al., 2001). Le tipologie di problemi affrontati mediante l'uso della C.A. sono molteplici e principalmente riguardano il supporto alla definizione di prodotti/servizi e l'identificazione di segmenti di clientela.

A partire dagli anni '70, lo sviluppo della Conjoint Analysis e la sua applicazione a problemi di ricerca sulle preferenze degli individui è stato notevole, sia perché ha coinvolto diverse discipline sia per il fatto che la sua elaborazione ha richiesto lo sforzo di figure professionali di diversa estrazione, a partire dagli accademici, consulenti, uomini d'azienda fino

agli sviluppatori di software. L'applicazione di tale metodologia negli studi aziendali è cresciuta notevolmente nel tempo (Wittink e Cattin, 1989; Green, Krieger e Wind, 2001), uscendo dagli ambiti esclusivi dell'impresa privata. Negli anni '90, infatti, si sono realizzati i primi studi nel campo dei servizi pubblici e da allora anche in tale ambito l'utilizzo della C.A. e i risultati ottenuti con la sua applicazione sono stati notevoli. La Conjoint Analysis è tutt'ora però una metodologia di difficile implementazione nella pratica aziendale e tale difficoltà è riconducibile ad un'ampia gamma di fattori. Il sistema di preferenze del consumatore è infatti un oggetto di misurazione di natura complessa. Infatti è difficile identificare un modello comportamentale univocamente determinato che permetta di modellizzare adeguatamente l'approccio decisionale dell'individuo nei confronti dei nuovi prodotti/servizi. Tale fatto è confermato dalla presenza in letteratura di diversi modelli comportamentali che devono tra l'altro tenere in considerazione l'eventuale esistenza di interazioni rilevanti tra alcuni attributi del prodotto/servizio. Relativamente a questo ultimo aspetto è necessario rilevare che, sebbene la letteratura accademica abbia sottolineato e dimostrato empiricamente l'importanza dei modelli di scelta non compensatori, nei software che attualmente trovano applicazione in campo manageriale questi aspetti non sono affatto considerati. Oltre alle difficoltà connesse alla modellizzazione del problema di C.A., è necessario anche considerare la multidimensionalità intrinseca dei prodotti e dei servizi, che sono oggetto del sistema di preferenze degli individui. Essi infatti si identificano per una serie di attributi difficilmente individuabili con esattezza, oltre che per il fatto di essere caratterizzati da una eventuale parziale percettività.

Nella fase di realizzazione dell'esperimento di CA, che consiste nella somministrazione di un questionario ad un gruppo di rispondenti, i dati vengono usualmente raccolti mediante interviste faccia a faccia, a volte con supporti di tipo grafico per rappresentare in maniera più efficace i prodotti/servizi sottoposti a valutazione da parte dell'intervistato. Un obiettivo fondamentale da conseguire negli studi di C.A. è quello di ottenere un numero sufficiente di dati per le stime dei parametri del modello con un numero quanto più possibile limitato di domande. Questa considerazione conduce alla tematica della definizione del disegno sperimentale (tipica del Design of Experiments) che spesso è molto trascurata soprattutto dai practitioner della Conjoint Analysis che si

affidano più ai software commerciali di C.A. che ad una attenta analisi preliminare per l'impostazione dell'indagine. In merito a questa ultima problematica, una parziale soluzione è offerta dai metodi adattativi di Conjoint Analysis che hanno come obiettivo quello di porre all'intervistato una sequenza di domande nella quale la domanda successiva dipende dalla risposta data dal rispondente nella domanda precedente.

Tra le proposte di metodi di Conjoint Analysis sviluppati in letteratura ne consideriamo in particolare due: la traditional o full-profile C.A. e l'adaptive C.A..

La traditional Conjoint Analysis si basa su giudizi espressi in forma di punteggi o graduatorie relativamente all'insieme completo (full design) di tutte le possibili alternative del prodotto oggetto di analisi oppure solo su una parte delle alternative (reduced design). I giudizi dei rispondenti sono rilevati in genere secondo un ranking o un rating (De Luca, 2002). I metodi di stima generalmente associati alla C.A. sono la regressione ai minimi quadrati (OLS), il modello lineare generalizzato (GLM), nel caso di giudizi metrici, e la regressione monotona, il LINMAP o il MONANOVA nel caso di giudizi ordinali.

Al fine di approfondire la valutazione delle differenze tra i due approcci e di comprendere meglio gli algoritmi sottostanti, nel presente lavoro, vengono confrontati i risultati di una applicazione della Conjoint Analysis effettuata mediante l'uso di due distinti software: il pacchetto ACA "Adaptive Conjoint Analysis" e il pacchetto SPSS-CONJOINT basato sulla Traditional Conjoint Analysis. Si noti che, seguendo lo svolgimento di un semplice esempio, viene completamente descritta la procedura statistica di analisi implementata in ciascuno dei due software considerati, con particolare riferimento al software ACA la cui procedura statistica è spesso descritta in modo piuttosto confuso e frammentato nei manuali e negli articoli di riferimento.

2. L'ADAPTIVE CONJOINT ANALYSIS: APPLICAZIONE DEL SOFTWARE ACA

In questo paragrafo viene brevemente introdotta l'*Adaptive Conjoint Analysis*, così chiamata, perché, a differenza dell'approccio tradizionale, nella presentazione dei profili successivi ai rispondenti, tiene conto delle preferenze precedentemente ottenute, "adattando", di conseguenza, i confronti successivi.

Il rispondente, infatti, opera dei giudizi sulle preferenze tra le diverse coppie proposte nei confronti, partendo da un giudizio sui livelli per ciascun attributo e poi pian piano arrivando a dei giudizi globali sui profili del prodotto. Questo metodo si è sviluppato per far fronte alla necessità di trattare e gestire problematiche con un elevato numero di fattori e conseguentemente di domande da sottoporre a rispondenti.

L'utilità viene indicata UTL_{ij} , dove con $i = 1, \dots, M$, si identificano gli attributi; con $j = 1, \dots, L_i$, si identificano i livelli di ogni attributo. Il numero di profili è dato dalla seguente produttoria: $\prod_{i=1}^M L_i$ e la funzione di utilità è così definita:

$$UTL_{ij} = \sum_{j=1}^{L_i} \sum_{i=1}^M w_{ij} x_{ij} \quad (1)$$

dove w_{ij} è il coefficiente di regressione che esprime il peso attribuito dal rispondente al j-esimo livello dell'i-esimo attributo e x_{ij} è una variabile dummy che rappresenta il j-esimo livello dell'i-esimo attributo.

L'ACA riduce considerevolmente la complessità dell'esperimento di Conjoint Analysis, limitando il numero di profili da sottoporre all'intervistato. Un'altra differenza, rispetto alla metodologia classica è quella di avere un piano sperimentale diverso per ogni intervistato, nel senso che questo si adatta alle risposte fornite dall'intervistato; questo approccio prevede inoltre la necessità di usare il computer per presentare l'intervista, sicuramente un vantaggio in termini di velocità nella raccolta delle informazioni.

Sostanzialmente, l'esperimento non viene più ad essere di vera e propria Conjoint Analysis, ma risulta ibrido, e la vera fase di C.A. risulta solamente l'ultima delle quattro; queste fasi verranno descritte di seguito trattando il successivo esempio dimostrativo.

2.1 Introduzione al software ACA

Il pacchetto software ACA "Adaptive Conjoint Analysis" è utilizzato per esperimenti di C.A. ed è stato sviluppato dalla Sawtooth Software. Il termine "adaptive" si riferisce al fatto che il computer gestisce l'intervista "personalizzandola" per ogni rispondente; ad ogni iterazione, infatti, i giudizi di preferenza precedenti sono usati per determinare le domande successive, in modo da ottenere sempre maggiori informazioni sulle scelte dell'intervistato. L'ACA ha due caratteristiche importanti:

- Consente all'analista di progettare una intervista interattiva, grazie all'uso del PC e di gestire l'intervista con i rispondenti. L'intervista può presentare un numero elevato di attributi e livelli, prestando particolare attenzione a quelli che i rispondenti considerano i più importanti. La fase dell'intervista è strutturata in una maniera "intelligente", in quanto le utilità dei rispondenti sono continuamente ristimate man mano che le domande vengono presentate. Le utilità finali sono disponibili solo una volta completata l'intervista.
- 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 d'acquisto per ciascun prodotto.

Il pacchetto ACA si compone delle seguenti funzioni:

- un "*interviewing module*" che ha il compito di raccogliere i dati di ciascun intervistato;
- un "*utility calculator*" per stimare i valori delle "utilities" per ogni livello di attributo;
- un "*market simulator*" al fine di testare, all'interno di differenti scenari competitivi, le alternative di prodotto.

Il questionario di C.A. si articola sostanzialmente in quattro fasi qui di seguito presentate, e che verranno successivamente riprese e in parte ampliate nella trattazione dell'esempio applicativo del software.

1. In questa prima fase ciascun rispondente ordina (in modalità ranking) secondo le sue preferenze i vari livelli dei diversi attributi presentati. Questa fase permette inoltre al rispondente di definire quei livelli considerati inaccettabili ai quali il software associa per default un valore di -9.999.
2. In questa seconda fase al rispondente sono presentati il migliore e il peggiore livello (ottenuti dalla fase di ranking precedente) di ciascun attributo presentato. Nel fare questo viene utilizzata una scala Likert da 1 a 4, dove 4 identifica l'importanza maggiore.
3. Nella terza fase, al rispondente vengono presentati due concetti di prodotto, intesi come combinazione di due o più attributi; l'intervistato dovrà esprimere il suo grado di preferenza per l'uno o per l'altro profilo, sempre attraverso una scala Likert, ma da 1 a 9.

4. E' la fase detta di "calibrazione", in cui il software presenta dei concetti di prodotto, utilizzando gli attributi e i livelli ritenuti rilevanti. Queste combinazioni sono proposte singolarmente all'intervistato, a cui viene chiesto di indicare la probabilità di acquisto di ognuno di questi profili. Il primo concetto di calibrazione presentato è una soluzione di prodotto a cui corrisponde il minor indice di gradimento, mentre il secondo concetto dovrebbe rispondere appieno alle esigenze del consumatore, mentre il terzo ed il quarto sono soluzioni intermedie. Le informazioni raccolte in questa ultima parte dell'intervista servono, principalmente, al sistema per calibrare le "utilities" calcolate precedentemente e per valutare l'attendibilità delle risposte date.

L'algoritmo utilizzato per la stima provvisoria delle "utilities" di ogni livello di ciascun attributo è il seguente:

1. Se è stato chiesto di ordinare i livelli di ogni attributo, l'ordinamento viene invertito; se è stato chiesto di valutare i livelli mediante una scala, questa viene mantenuta.
2. Viene sottratta la media dei valori associati ad ogni attributo dai giudizi espressi, per centrare il valore sullo zero; ad esempio i valori 3, 2, 1, vengono convertiti in 1, 0, -1, rispettivamente.
3. I valori vengono scalati per rientrare in un range di variazione pari a 1; ad esempio 1, 0, -1 diventano 0.5, 0 e -0.5.
4. Il valore dell'importanza di ogni attributo viene riportato in un range da 1 a 4; in seguito sarà usato come moltiplicatore per i livelli di desiderabilità calcolati nel punto 3; se un attributo ha un'importanza pari a 3 e desiderabilità nell'intervallo 0.5; 0; -0.5, assumerà i valori 1.5, 0, -1.5.

Le stime delle "utilities" calcolate in questo modo godono di due proprietà:

1. per ogni attributo il range di utilità è proporzionale all'importanza; due valori infatti differiscono al massimo di un fattore pari a 4;
2. ogni attributo ha un valore medio uguale a zero; le "utilities" sono proporzionali al livello di desiderabilità espresso nella sezione "rating/ranking".

Nella sezione dei "confronti a coppie", le "utilities" calcolate nella prima sezione dell'intervista vengono aggiornate dopo ogni risposta.

Il software, in questa seconda sezione, per l'aggiornamento delle "utilities" derivanti dalla prima parte dell'intervista, impiega il modello di regressione ai minimi quadrati in modo simile a quanto avviene nella Ridge Regression.

Per il calcolo delle "utilities" è necessario definire quanto segue:

X: matrice di variabili indipendenti, con n righe per n osservazioni;

- b_n : vettore dei coefficienti di regressione che si otterrebbero con il metodo dei minimi quadrati sulle n osservazioni;
- Y : vettore contenente le n risposte od osservazioni;
- z^T : vettore riga contenente i valori assunti dalle variabili indipendenti per una nuova osservazione;
- r : risposta per una nuova osservazione.

2.2 Esempio applicativo su ACA

Presentiamo ora un semplice esempio applicativo del software ACA, che permetterà di chiarire in maniera più concreta i punti sopra descritti; alcuni di questi punti verranno ampiamente ripresi nell'analisi affinché la comprensione degli algoritmi che gestiscono il software risulti più chiara.

L'esempio in questione riguarda un personal computer portatile, in cui vengono individuati tre attributi, quali: peso, durata della batteria, prezzo.

Ciascuno di questi attributi ha un certo numero di livelli; inoltre si assume come ipotesi di avere un solo rispondente e di non avere alcun livello non-accettabile.

Nel dettaglio si ha quindi:

<u>Attributo 1: PESO</u>		<u>Attributo 2: DURATA BATTERIA</u>		<u>Attributo 3: PREZZO</u>	
Level 1	1,5 kg	Level 1	1 ora	Level 1	1000 euro
Level 2	2 kg	Level 2	2 ore	Level 2	2000 euro
		Level 3	4 ore	Level 3	3000 euro
		Level 4	6 ore		

Fig 1 – Attributi e rispettivi livelli.

Riprendiamo quindi le quattro fasi del questionario presentato al rispondente:

1. Questa prima fase del questionario è necessaria per dare un ordine di preferenza ai livelli dei vari attributi. L'intervistato deve infatti ordinare i livelli di ciascun attributo secondo un suo ordine di preferenza. Questa fase può essere evitata per certi attributi che, come ad esempio il prezzo, hanno già un ordine logico dei relativi attributi.
2. In questa fase viene determinata l'importanza relativa di ogni attributo. Vengono proposti due prodotti che si differenziano solo per una caratteristica e si chiede quindi di stimare l'importanza che questa

caratteristica assume nella scelta finale, assegnando un punteggio che va da 1 a 4.

3. Vengono ora presentati due concetti di prodotto, intesi come combinazione di due o più attributi; l'intervistato dovrà quindi esprimere il suo grado di preferenza per l'uno o per l'altro profilo.

Il software fornisce una prima stima delle utilities, ricavate dalle fasi precedenti, e le aggiorna dopo ogni risposta. Ogni domanda sfrutta i dati in possesso in quel momento e viene formulata in modo da fornire il maggior numero di informazioni. I concetti proposti nel confronto hanno un valore di total utility simile; l'intervistato noterà, infatti, che diventerà sempre più difficile operare una scelta sulle combinazioni di profili presentate, ed avrà come l'impressione che il software faccia attenzione alle sue risposte e cerchi di andare in profondità per percepire le differenze di preferenza.

La scelta risulterà sempre più difficile e la valutazione del rispondente si sposterà sempre più verso la parte centrale della scala. Ogni attributo ed ogni livello vengono presentati con uguale frequenza.

Il sistema segue la seguente procedura:

- conta il numero di volte che ogni coppia di attributi è comparsa insieme in un concetto di prodotto presentato, e sceglie a caso (random) tra gli attributi apparsi meno frequentemente;
- per ogni attributo scelto, il sistema segue lo stesso procedimento per selezionare i diversi livelli;
- esamina tutte le possibilità di combinazione dei livelli, e cerca i due concetti più vicini come attrattiva in base alle utilities stimate fino a quel momento;
- determina a caso quale concetto posizionare a sinistra dello schermo e quale a destra.

Il numero di confronti a coppie è dato dalla seguente formula:

$$3 \times (N-n-1) \cdot N \quad (2)$$

con N = somma dei livelli previsti, ed n = numero di attributi considerati.

La giustificazione di questa formula è la seguente: $(N-n-1)$ rappresenta il numero di profili contenuti in un generico piano frazionato di tipo Plackett-Burman, da questo si sottrae N , in quanto si è già ottenuta una osservazione per ogni livello da stimare, attraverso la prima parte dell'intervista (denominata "priors"). La moltiplicazione per il numero 3, si spiega col fatto di voler ottenere tante osservazioni, pari a tre volte il numero di parametri da stimare.

4. In questa ultima fase il software ACA presenta dei concetti, utilizzando gli attributi e i livelli ritenuti più rilevanti. Queste combinazioni sono proposte

I blocchi evidenziano i dati che corrispondono alle quattro fasi del questionario (A, B, C, D) e i risultati delle utilities calcolate (blocco E).

Ora si illustreranno le fasi, in relazione all'esempio considerato.

1° FASE: Ordine di preferenza degli attributi

Viene assegnato un ordine di preferenza ai livelli dei vari attributi, ponendo una domanda del tipo: "Esprima la sua preferenza in merito alla caratteristica seguente (ad esempio il peso)". Nello specifico si ha quindi:

Attributo: PESO

1.5 kg → 1

2 kg → 2

Questo evidenzia come l'intervistato abbia dato come ordine di preferenza prima 1.5 kg e poi 2 kg. Analogamente per gli altri livelli degli altri due attributi.

L'ordine di preferenza viene successivamente invertito e questo lo si nota nelle righe del blocco A del file .ACD. Quello che conta, infatti, viene ad essere la posizione del valore e non il valore stesso. Ad esempio, la seconda riga (1, 2, 3, 4), dice che l'ordine di preferenza dei livelli, in relazione all'attributo durata batteria, è: 6 ore, 4 ore, 2 ore, 1 ora..

2° FASE: Importanza relativa degli attributi

In questa seconda fase, come è stato detto in precedenza, viene determinata l'importanza relativa di ogni attributo. In relazione all'esempio, questa fase fa riferimento al blocco del file evidenziato con la lettera B.

Le tre righe si interpretano in questo modo, rispettivamente:

"Se relativamente agli altri attributi, i prodotti sono entrambi accettabili, quanto è importante, in una scala da 1 a 4, la differenza tra:

I riga) 1.5 kg vs 2 kg ?

Risposta : 2

II riga) 6 ore vs 1 ora ?

Risposta : 4

III riga) 1000€ vs 3000€ ?

Risposta: 3

Come si può vedere, si è data maggiore importanza all'attributo durata della batteria e minore importanza all'attributo peso. terminate queste prime due fasi, il software possiede i dati necessari per determinare il primo "set di utilities", che corrisponde alla riga centrale del blocco E. La procedura per determinare

queste utilities prevede i seguenti passi, già visti nella parte introduttiva del software:

- 1) Se è stato chiesto di ordinare i livelli di ogni attributo, l'ordinamento viene invertito; se è stato chiesto di valutare i livelli mediante un punteggio, questo viene mantenuto.
- 2) Viene sottratta la media dei valori associati ad ogni attributo dai giudizi espressi, per centrare il valore sullo zero; ad esempio i valori 3, 2, 1, vengono convertiti in 1, 0, -1, rispettivamente.
- 3) I valori vengono riscaldati per rientrare in un range di variazione pari a 1; ad esempio 1, 0, -1, diventano 0.5, 0 e -0.5.
- 4) Il valore dell'importanza di ogni attributo viene riportato in un range da 1 a 4; in seguito sarà usato come moltiplicatore per i livelli di desiderabilità calcolati nel punto 3; se un attributo ha un'importanza pari a 3 e desiderabilità nell'intervallo 0.5; 0; -0.5, assumerà i valori 1.5, 0, -1.5.

Nell'esempio in questione, applicando uno alla volta i quattro punti, si ha (Tab. 1):

Livelli	Ord. iniziale	1°	2°	3°	4°
1,5 kg	1	2	0.5	0.5	1
2 kg	2	1	-0.5	0.5	-1
1 ora	4	1	-1.5	-0.5	-2
2 ore	3	2	-0.5	-0.167	-0.668
4 ore	2	3	0.5	0.167	0.668
6 ore	1	4	1.5	0.5	2
1000 euro	1	3	1	0.5	1.5
2000 euro	2	2	0	0	0
3000 euro	3	1	-1	-0.5	-1.5

Tabella. 1 - Passaggi che portano alle stime iniziali di utilità parziale, evidenziate nell'ultima colonna in grassetto.

In questo schema si osserva lo sviluppo dell'algoritmo utilizzato per determinare le stime iniziali delle utilità parziali. Dapprima viene invertito l'ordine (passo 1), successivamente viene sottratta la media dei valori di ogni attributo dai giudizi espressi (passo 2); il passo 3 consiste nello riscaldare i valori in un range di variazione 1 ed infine, il passo 4, prevede di moltiplicare i valori ottenuti, per l'importanza data nella fase 2 del questionario. Nell'esempio in questione, si avrà che le utilities relative all'attributo "peso" saranno moltiplicate per il valore 2, quelle relative all'attributo "durata batteria" per il valore 4, ed infine, quelle relative all'attributo "prezzo" moltiplicate per il valore 3, in altre parole, il giudizio di preferenza dato precedentemente, assume il valore di un peso, utilizzato per la determinazione delle utilities finali.

In questo modo si ottiene l'ultima colonna, con i valori delle stime iniziali.

3° FASE: confronto a coppie

Questa è la cosiddetta fase del confronto a coppie. Nell'esempio in questione si fa riferimento al blocco caratterizzato dalla lettera C.

Al rispondente vengono presentati due concetti di prodotto e lui dovrà esprimere la sua preferenza in una scala da 1 a 9.

Ad esempio, la prima riga del blocco C fa riferimento al confronto:

"Quale profilo preferisce tra i seguenti?"

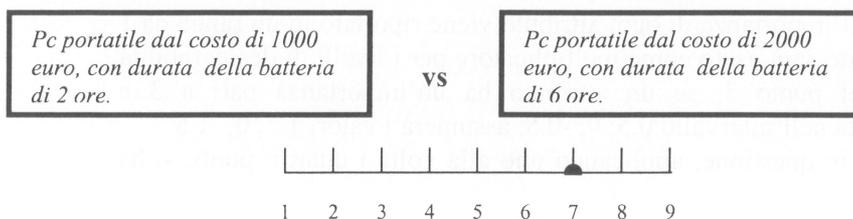


Fig. 3. – Schermata presentata dal software all'intervistato.

La risposta data è stata 7, cioè una discreta preferenza a destra, verso il PC da 2000 euro e con batteria da 8 ore. Il numero massimo di confronti è dato per default dal software ed è pari a tre volte il numero di parametri da stimare, ma può essere variato. Il sistema ACA, in questa fase del questionario, per il calcolo delle utilities, impiega il modello della regressione ai minimi quadrati OLS (Ordinary Least Squares).

Per il calcolo delle "utilities", prima di iniziare ad aggiornare le stime, si precisa quali valori assumono gli elementi precedentemente definiti, in relazione all'esempio trattato:

$X = I$ (matrice identità)

Y : vettore delle osservazioni

b_n : vettore delle stime iniziali

z : vettore formato da +1, -1 o 0; un elemento assume valore +1 se il livello corrispondente appare nel concetto di prodotto a destra dello schermo; -1 se appare a sinistra dello schermo e 0 se non viene presentato nella sezione "confronti a coppie".

r : assume valori da -4 a +4, dove +4 significa "forte preferenza a destra", -4 "forte preferenza a sinistra" e 0 significa "parità di giudizio".

Dopo le prime n osservazioni, si ha:

$$E(Y) = Xb_n \implies Xb_n \approx Y \quad (3)$$

Dove

$$\mathbf{b}_n = (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{Y}) \quad (4)$$

è il vettore dei coefficienti ottenuti con il metodo dei minimi quadrati, basato sulle prime n osservazioni. Se si aggiunge una osservazione si ha:

$$\begin{bmatrix} \mathbf{X} \\ z^T \end{bmatrix} \mathbf{b}_{n+1} \approx \begin{bmatrix} \mathbf{Y} \\ r \end{bmatrix} \quad (5)$$

dove
$$\mathbf{b}_{n+1} \approx (\mathbf{X}^T \mathbf{X} + z^T z)^{-1} (\mathbf{X}^T \mathbf{Y} + zr) \quad (6)$$

è la stima dei minimi quadrati sulle $n+1$ osservazioni.

Si supponga di conoscere anche $\mathbf{b}_n, \mathbf{Y}, z$ e r e di voler calcolare \mathbf{b}_{n+1} , ossia il vettore delle stime iniziali, dopo la prima iterazione.

Introducendo:

$$\mathbf{v} = (\mathbf{X}^T \mathbf{X}) z \quad (7)$$

si ottiene:

$$(\mathbf{X}^T \mathbf{X} + z^T z)^{-1} = (\mathbf{X}^T \mathbf{X})^{-1} - \frac{\mathbf{v}^T \mathbf{v}}{1 + \mathbf{v}^T z} \quad (8)$$

e sostituendo nella prima equazione:

$$\mathbf{b}_{n+1} = \mathbf{b}_n + \mathbf{v} \frac{r - z^T \mathbf{b}_n}{1 + \mathbf{v}^T z} \quad (9)$$

questa formula permette di aggiornare le “utilities” dopo ogni risposta della sezione “Pairs” ossia della 3° fase.

Nell'esempio trattato, prima di iniziare ad aggiornare le “utilities”, si ha quindi:

$$X = \begin{bmatrix} 100000000 \\ 010000000 \\ 001000000 \\ 000100000 \\ 000010000 \\ 000001000 \\ 000000100 \\ 000000010 \\ 000000001 \end{bmatrix} \quad Y = \begin{bmatrix} 1 \\ -1 \\ -2 \\ -0,668 \\ 0,668 \\ 2 \\ 1,5 \\ 0 \\ -1,5 \end{bmatrix} \quad X = \begin{bmatrix} 1 \\ -1 \\ -2 \\ -0,668 \\ 0,668 \\ 2 \\ 1,5 \\ 0 \\ -1,5 \end{bmatrix}$$

I vettori riga z^T , da aggiungere alla matrice identica dopo ogni iterazione, esprimono i confronti a coppie, e sono facilmente ricavabili dal blocco **C** del file **ACD** e assumono, pertanto, i valori:

$$\begin{aligned} z_1^T &= (0 \ 0 \ 0 \ -1 \ 0 \ 1 \ -1 \ 1 \ 0), & z_2^T &= (-1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ -1), \\ z_3^T &= (-1 \ 1 \ -1 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0), & z_4^T &= (-1 \ 1 \ -1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0), \\ z_5^T &= (-1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ -1), & z_6^T &= (0 \ 0 \ 0 \ 1 \ -1 \ 0 \ 0 \ 1 \ -1). \end{aligned}$$

I valori di r , da inserire nella formula, vengono calcolati come deviazione da una scala zero centrale.

Ad esempio, sempre dalla prima riga del blocco **C**, si ha una risposta di valore 7, che assumerà nella formula il valore 2; vediamo perché:

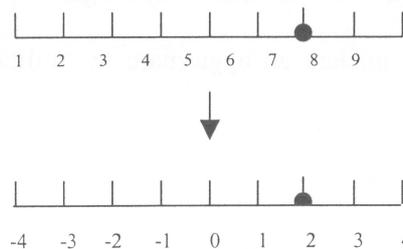


Fig.4. – Viene evidenziato “visivamente” come vengono riscaltati i valori di r .

Si osserva che il valore 7, riportato alla scala zero centrale, ha distanza 2 da 0 e questo è il valore che assume r_1 .

Con il medesimo procedimento, si ricava il valore delle altre variabili, si avrà quindi:

$$r_2 = 3, r_3 = 3, r_4 = 1, r_5 = 3, r_6 = 1$$

Ora, si hanno tutti gli elementi necessari per ricavare i set di utilities mancanti, che sono rispettivamente la prima e la terza riga del blocco **D** del file **.ACD**.

La prima riga, detta “pre calibrated utilities”, si calcola applicando iterativamente la formula (9) con \mathbf{b}_n alla prima iterazione coincidente con le utilities iniziali.

La terza riga, invece, fa riferimento alle utilities derivanti dai confronti a coppie solamente, ed è ottenuta applicando iterativamente sempre la stessa formula, ma con \mathbf{b}_n alla prima iterazione coincidente col vettore nullo.

4° FASE: calibrazione

Questa ultima fase, detta di “calibrazione”, permette di calcolare un’intercetta e due coefficienti di regressione, per determinare i pesi che meglio stimano la probabilità d’acquisto.

I parametri qui ricavati, sono usati poi per stimare le “utilities finali”.

Questo si spiega con il fatto che nelle versioni precedenti di **ACA**, le informazioni ricavate dalla prima parte dell’intervista, ed espresse mediante le utilities iniziali, avevano lo stesso peso di quelle estratte dalla sezione dei confronti a coppie.

Nella versione 4 del software, invece, la fase finale del questionario, in cui si chiede di esprimere la probabilità d’acquisto per i prodotti proposti, viene utilizzata per calibrare le “utilities”, dando pesi diversi alle due stime precedentemente calcolate.

Nell’esempio in questione, in riferimento al blocco **D**, i numeri 4 e 3 stanno rispettivamente ad indicare:

- il numero di concetti di calibrazione presentati (sono le quattro righe);
- il numero di livelli dei concetti di calibrazione presentati.

La prima riga, ad esempio, si legge in questo modo:

“Con quale probabilità acquisterebbe un pc portatile, del peso di 1.5 kg, con la batteria della durata di un’ora e il costo di 3000 €?”

La risposta corrisponde all’ultimo valore della riga e nella fattispecie è del 5%.

I valori restanti della riga fanno riferimento all’utilità del profilo proposto, e rispettivamente il valore:

- 2.5 fa riferimento alla somma delle utilities iniziali, mentre il valore - 1.702 alla somma delle utilities derivanti dai confronti a coppie.

Per determinare i valori dell'intercetta e dei due coefficienti, il software utilizza la seguente formulazione della di probabilità d'acquisto, come funzione dei pesi degli intervistati:

$$\ln \frac{p}{1-p} \approx a + b_1 x_1 + b_2 x_2 \quad (10)$$

dove:

p : probabilità d'acquisto prevista per un certo prodotto;

x₁ : peso stimato per il prodotto, usando le stime iniziali;

x₂ : peso stimato per il prodotto, usando le stime derivanti dai confronti a coppie;

b₁ : coefficiente usato come peso per le stime iniziali;

b₂ : coefficiente usato come peso per le utilities derivanti dai confronti a coppie;

a : parametro di intercetta.

In riferimento all'esempio in questione, i valori di **a**, **b₁**, **b₂**, sono rispettivamente:

$$-0.200 \quad 0.638 \quad 0.565.$$

Una volta ricavati i valori di **a**, **b₁**, **b₂**, il software procede al calcolo delle utilities finali, che sono contenute nel file di estensione **.UTL**.

2.2.2 Il file .UTL

Il software, nel calcolo dei "parth worths" finali, cioè i valori delle "utilities" da utilizzare nell'analisi dei dati, usa il seguente modello di regressione:

$$\text{Utility finale} = \frac{a}{n^{\circ} \text{attributi}} + b_1 x_1 + b_2 x_2 \quad (11)$$

Da qui si ricavano i valori delle "utilities" finali che assumono i seguenti valori:

$$U1 = 0.1204; U2 = -0.2538; U3 = -1.442;$$

$$U4 = -0.774; U5 = 0.300; U6 = 1.65;$$

$$U7 = 0.98; U8 = 0.25; U9 = -1.435$$

Questi sono i dati finali, di cui si serve l'analista, che danno un'indicazione di come sono posizionate le preferenze del rispondente in riferimento ai livelli dei vari attributi rispetto al sistema di preferenze totale e possono, in seguito, essere applicate in un numero considerevole di campi, di cui si è già discusso nei precedenti paragrafi.

3. IL PACCHETTO SPSS-CONJOINT: ALCUNI ESEMPI APPLICATIVI

In questo paragrafo verrà illustrato brevemente tramite il pacchetto software SPSS, l'utilizzo della Traditional Conjoint Analysis, basata su un disegno di indagine di tipo fattoriale completo o frazionato (full profile C.A.).

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

- Le procedure per lo sviluppo e la determinazione di profili usati negli studi di Conjoint.
- Quattro differenti modelli per la Conjoint Analysis.
- Un sommario grafico per i risultati di analisi.

3.1 Il metodo "full profile"

Il metodo full-profile implementato in SPSS, prevede che il rispondente assegni un ordine ai profili presentati, in accordo con le sue preferenze.

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, ovvero caratterizzato da livelli di tutti gli attributi che lo compongono.

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 (part worth) per ciascun livello di ciascun attributo.

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.

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 profilo presentato.

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 ordinando ciascun profilo, dove n indica il numero di profili presentati e il rank più basso indica la preferenza più elevata.

Di seguito viene implementato con il software SPSS l'esempio applicativo già precedentemente trattato con il software ACA.

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

Peso (in kg)	Durata	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

Tab. 2 – Piano ortogonale dell'esempio di applicazione

Compito dell'analista è quello ora di gestire i diversi profili per ciascun rispondente. La procedura implementata in SPSS-Conjoint consente due principali metodi 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.

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.

L'output della Conjoint Analysis è evidenziato in figura 5.

Come si può notare, sono messi in evidenza i valori delle utilità per ogni livello di ogni attributo.

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.

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
6.36	.1871		PESO
	-.1871		1.5 kg
			2 kg
52.57	-1.3755		DURATA
	-.7075		1 ora
	.3665		2 ore
	1.7165		4 ore
			6 ore
41.06	1.0483		PREZZO
	.3183		1000 €
	-1.3667		2000 €
			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

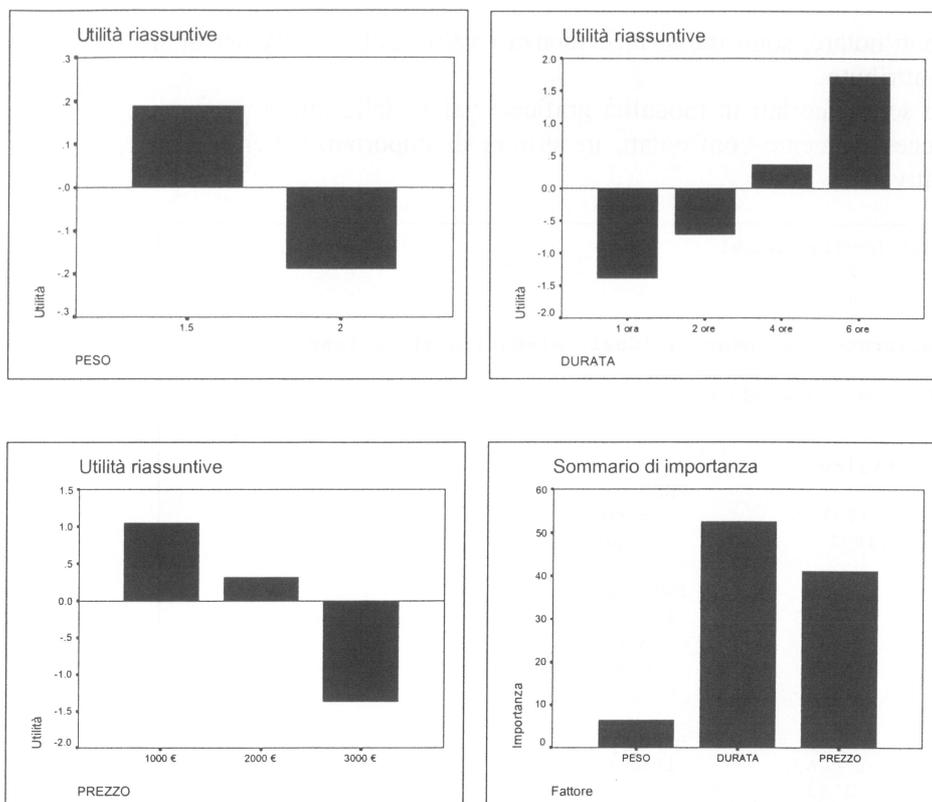


Fig. 6 – Utilità degli attributi

Di seguito, sempre dallo stesso esempio, si riportano i risultati di altri due esperimenti di C.A., che prevedono dieci rispondenti; nel primo caso si ha un'analisi aggregata dei dati (Fig. 7), mentre nel secondo si presenta un'analisi di tipo disaggregato (Fig. 8).

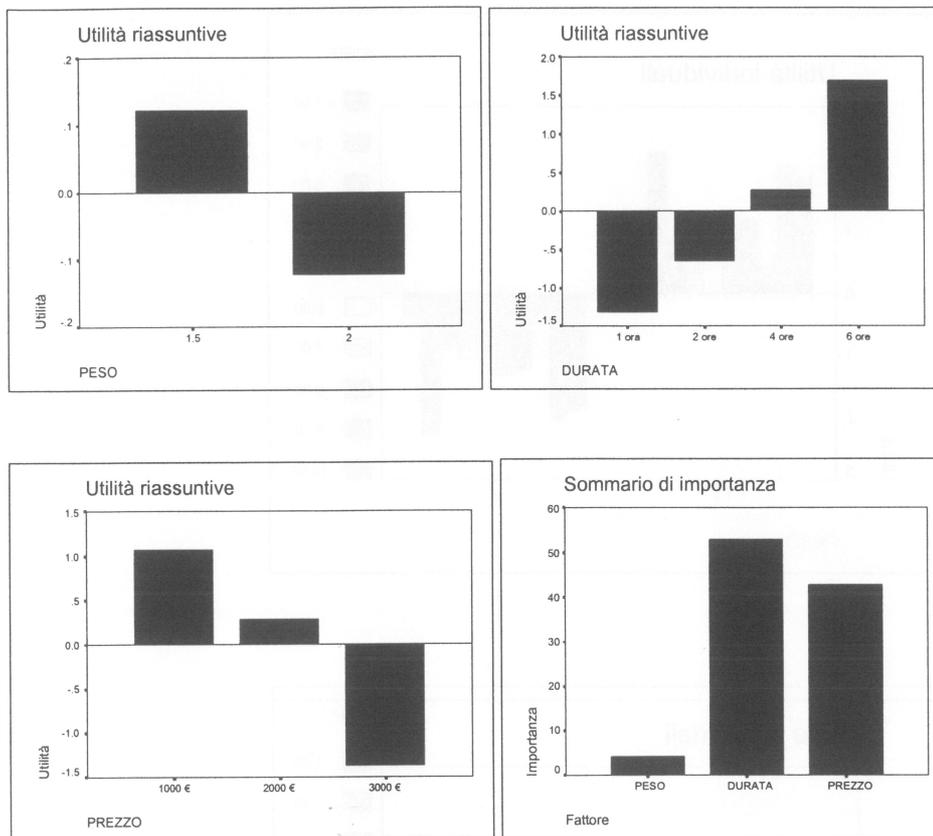


Fig. 7 – Utilità dei tre fattori e confronto complessivo delle utilità dei tre fattori con 10 rispondenti e con confronto di tipo aggregato

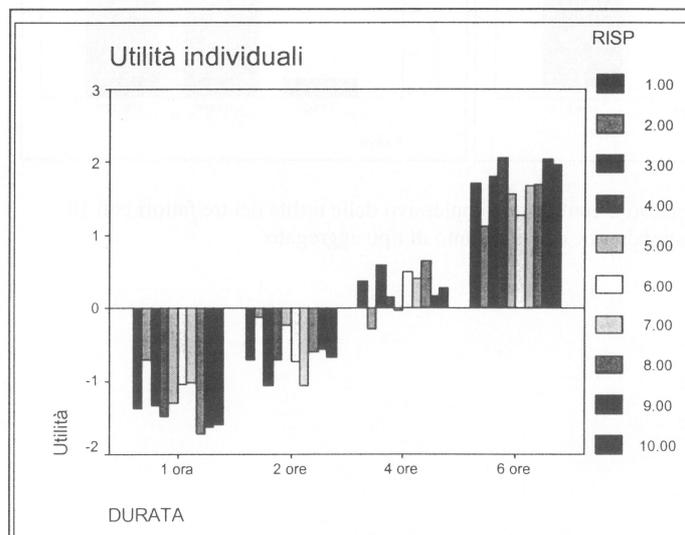
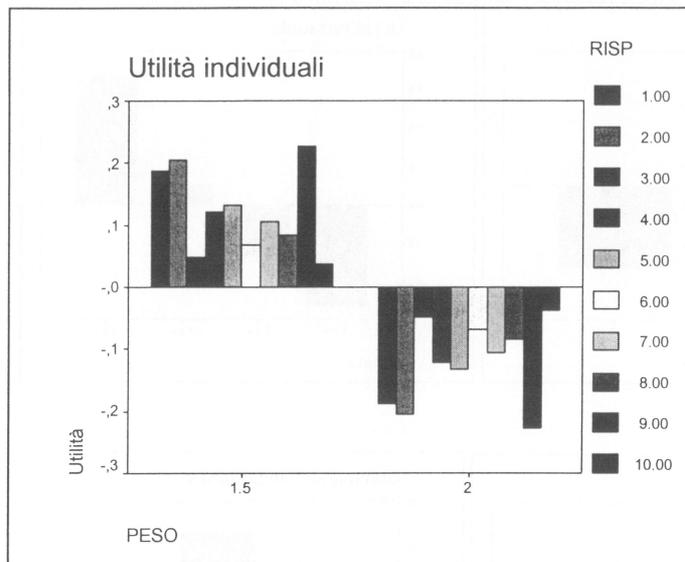


Fig. 8 – Valori di utilità, in modalità disaggregata, per i dieci rispondenti.

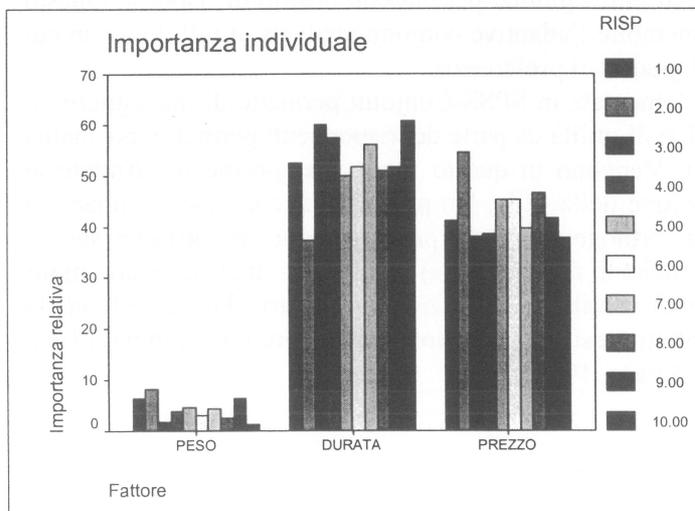
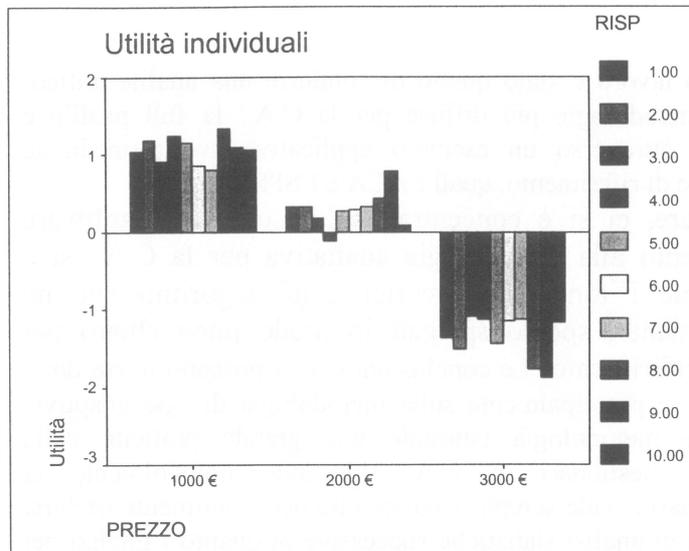


Fig. 8 segue – Valori di utilità, in modalità disaggregata, per i dieci rispondenti.

4. CONCLUSIONI

Lo scopo di questo lavoro è stato quello di condurre una analisi critico-qualitativa delle due metodologie più diffuse per la C.A.: la full profile e l'adaptive C.A. anche attraverso un esempio applicativo svolto mediante l'utilizzo di due software di riferimento, quali l'ACA e l'SPSS.

In modo particolare, ci si è concentrati sullo studio del software ACA, che fa riferimento alla metodologia adattativa per la C.A., si è cercato di estrapolarne i fondamenti teorici e gli algoritmi che ne regolano il funzionamento, spesso spiegati in modo poco chiaro nei manuali e articoli di riferimento. Le conclusioni che si possono trarre dopo questo studio sull'ACA e principalmente sulla metodologia di tipo adaptive, sono le seguenti: tale metodologia consente una grande praticità nella somministrazione dei questionari di C.A. riducendo notevolmente la complessità del questionario. Tale semplificazione può però facilmente tradursi in perdita di affidabilità di analisi statistiche successive in quanto i giudizi dei rispondenti sono espressi su profili molto parziali di tutti gli attributi oggetto di analisi. Qualora infatti l'analista volesse supportare un'analisi statistica successiva (ad esempio di segmentazione di mercato) sulla base dei risultati forniti sulle utilità parziali, dovrebbe stimare molti giudizi di preferenza non potendo infatti lavorare su tutti i giudizi per ciascun profilo di risposta. Questo aspetto rende molto vulnerabile l'adaptive conjoint analysis in tutti i casi in cui vi sia alta variabilità nei giudizi di preferenza;

La full profile C.A. implementata in SPSS-Conjoint permette di raccogliere un maggior numero di giudizi di utilità da parte dei rispondenti permettendo analisi statistiche più affidabili. Vengono in questo modo maggiormente sfruttate le caratteristiche decompositive della C.A.. Un problema ancora aperto rimane la definizione di opportuni orthogonal arrays per il disegno di indagine spesso ancora troppo oneroso (anche in forma ridotta) in presenza di attributi con molti livelli. Sarebbe auspicabile l'utilizzo di frazioni irregolari che permettano la stima delle utilità di attributi presi singolarmente e delle loro interazioni (si veda ad esempio Kounias e Salmaso, 1998).

Bibliografia

1. "ACA 5.0 TECHNICAL PAPER", 2002, Sawtooth Software, sito internet: www.sawtoothsoftware.com.
2. BRASINI S., TASSINARI F., TASSINARI G., 1996, "Marketing e Pubblicità: Metodi di Analisi Statistica", Bologna: Il Mulino.
3. CAMILLO, F., STEFANO SFORZA, 2000, "Indagini sulla soddisfazione del cliente: applicazione della Conjoint Analysis". Luis Management, Agorà Consulting.
4. DE LUCA A., 2002, "Le applicazioni dei metodi statistici alle analisi di mercato", Franco Angeli, pp. 325-407.
5. DOLAN R., 1993, "Managing the New Product Development Process", Reading, Mass: Addison Wesley.
6. DUSSAIX ANNE-MARIE, 1993, «Analyse Conjointe et Identification des critères de choix des consommateur », SPAD MC-Manuel de l'Utilisateur, CISIA, Saint-Mandè.
7. GREEN P.E., SRINIVASAN V., 1978, "Attitudes and Prediction of Behaviour", in: Fishbein M. (a cura di)- Readings in Attitude theory and Measurements, John Wiley and Sons, New York, pp. 477-492.
8. GREEN P.E., SRINIVASAN V., 1990, "Conjoint Analysis in Marketing: new Development with Implications for Research and Practice". Journal of Marketing, 54, October, pp. 3-19.
9. GREEN P.E., KRIEGER A.M., AGARWAL M.K., 1991, "Adaptive Conjoint Analysis: Some Caveats and Suggestions": Journal of Marketing Research, Vol. XXVIII, pp. 215-222.
10. GREEN, P. E., A. M. KRIEGER, AND Y. WIND, 2001, "Thirty years of conjoint analysis: Reflections and prospects," Interfaces, 31(3), S56-S73.
11. GUSTAFSSON, A., ANDREAS HERRMANN, AND F. HUBER, 2001, "Conjoint measurement: methods and applications". Berlin: Springer.
12. JOBSON J.D., 1991, "Regression and experimental design: applied multivariate data analysis", Springer-verlag New York, vol 1, pp. 500-502.
13. JOHNSON RICHARD M., 1996, "Accuracy of Utility Estimation in ACA", Sawtooth Software, sito internet: www.sawtoothsoftware.com.
14. JOHNSON RICHARD M., 1987, "Adaptive Conjoint Analysis, Sawtooth Software, Inc.
15. KOUNIAS, S., SALMASO, L., 1998. Orthogonal plans of resolution IV and V. Journal of the Italian Statistical Society, Vol. 7, pp. 57-75.

16. LAMBIN JEAN-JACQUES, 1991, "Marketing ", McGraw-Hill Libri Italia, Milano.
17. LUCE R.D.M, TUKEY J., 1964, "Simultaneous conjoint measurement: a new type of fundamental measurement", Journal of market psychology, vol 1, pp. 1-27.
18. MOLTENI L., MANOFORTE R., 1998, "La Conjoint Analysis e il problema delle interazioni fra gli attributi: un'evidenza empirica". Liuc Papers n.58, Serie Metodi quantitativi.
19. SPSS CONJOINT 8.0 Manuale d'uso.
20. SAPORTA G., CAMILLO F., 1997, "Classification, Data Analysis and Conjoint Analysis Models: integrated tools for marketing research". Atti del Convegno SIS "La statistica per le imprese", Torino, aprile 1997. Tirrenia Stampatori.
21. SAWTOOTH SOFTWARE, 1994, "ACA System", Margo Metegrano Editor, manuale d'uso.
22. WITTINK D., CATTIN P., 1989, "Commercial Use of Conjoint Analysis: Results and Critical Reflections". International Journal of Research in Marketing, 11, pp. 41-52.