

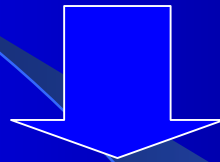
# Conjoint Analysis - 2

Questa metodologia parte quindi dall'ipotesi che un consumatore valuti un prodotto/servizio attraverso una combinazione delle utilità dei singoli attributi che compongono lo stesso (Lancaster considera il prodotto come un insieme di caratteristiche/attributi ed il giudizio di valore sul prodotto come la somma dei giudizi sulle singole parti).

# SEGMENTAZIONE DEL MERCATO - 1

Per definire gli obiettivi e le strategie commerciali un'impresa deve valutare:

- i bisogni ed i comportamenti probabili degli acquirenti potenziali;
- il mercato di riferimento



## SEGMENTAZIONE DEL MERCATO

-in gruppi di consumatori:

- distinti
- omogenei

verso i quali indirizzare specifiche politiche di vendita.

# SEGMENTAZIONE DEL MERCATO - 2

## MARKET SEGMENTATION:

adeguamento dei prodotti e delle strategie di marketing alle differenze individuabili entro l'insieme delle esigenze manifestate dai consumatori.

Attenzione!

In questo contesto la nozione di segmentazione non deve essere confusa con la differenziazione delle attività di un'impresa, e dei relativi prodotti, da quelli della concorrenza.

# SEGMENTAZIONE DEL MERCATO - 3

L'efficace attuazione delle politiche di marketing esige che ogni segmento di mercato a cui l'azienda si rivolge presenti i seguenti requisiti:

- uniformità delle risposte dei potenziali consumatori rispetto agli strumenti classici del marketing-mix (come la qualità del prodotto, la pubblicità, la distribuzione,...);
- la profittabilità, nel senso che la dimensione del segmento devono essere tali da garantire un profitto;
- L'accessibilità, ossia la possibilità per dell'azienda di raggiungere il segmento di mercato d'interesse, con costi addizionali o perdite contenute.

# SEGMENTAZIONE DEL MERCATO - 4

Da un punto di vista operativo la segmentazione del mercato si articola nelle tre seguenti fasi:

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

# Scopi della Conjoint Analysis - 1

Gli usi della Conjoint Analysis possono essere classificati sulla base:  
dello scopo da perseguire con l'analisi stessa.

# Scopi della Conjoint Analysis - 2

In base allo **scopo** dell'analisi:

- Segmentazione del mercato;
- Decisioni relative al prodotto;
- Analisi competitiva;
- Decisioni sul prezzo;
- Decisioni sulla promozione e sulla distribuzione

# Scopi della Conjoint Analysis - 3

## **Segmentazione del mercato:**

Esistono forme diverse di segmentazione del mercato che si differenziano in relazione alla base utilizzata per suddividere il mercato stesso. Generalmente si considerano i benefici che il prodotto offre ai clienti e come questi ultimi li percepiscono (approccio diretto ed orientato al consumatore).

La segmentazione del mercato può realizzarsi in due modi tra loro diversi: a priori attraverso un campionamento stratificato della clientela oppure, a posteriori, creando dei cluster (gruppi) individuati in base alla similarità dei benefici (i benefici percepiti dai clienti sono numericamente rappresentati dai pesi ottenuti con la Conjoint Analysis).

# Scopi della Conjoint Analysis - 4

## Decisioni relative al prodotto:

Questa metodologia risulta utile anche per ottenere informazioni e suggerimenti di supporto a decisioni del tipo:

- modifiche ad un altro prodotto, come la modifica di alcuni attributi;
- progettazione ottimale di un prodotto, in questo caso vengono presi in considerazione tutti gli attributi;
- riprogettazione della linea;
- valutazione dei concetti di un nuovo prodotto.

# Scopi della Conjoint Analysis - 5

## Analisi Competitiva:

Permette di prevedere i cambiamenti, per esempio, della quota di mercato derivanti da variazioni apportate a prodotti o a linee di prodotto. Questo è possibile tramite dei modelli di simulazione basati sui dati ottenuti tramite la Conjoint Analysis.

## Decisioni sul prezzo

Tramite modelli di Conjoint Analysis abbastanza sofisticati è possibile ottenere informazioni e stime dell'elasticità rispetto al prezzo, utili sia dove non vi siano dati storici sia per supportare decisioni di marketing.

# Scopi della Conjoint Analysis - 6

## Decisioni sulla promozione e sulla distribuzione:

Questi sono scopi che possono essere definiti indiretti della Conjoint Analysis. Infatti, tramite questa metodologia si possono ricavare informazioni utili anche rispetto a queste categorie di decisioni, per fare questo si possono utilizzare come attributi variabili relative ad aspetti promozionali, come pubblicità, modalità di pagamento,...rilevando i giudizi che i distributori danno a certe combinazioni.

Si possono, inoltre, analizzare alcuni problemi relativi alla distribuzione come i tempi di trasporto o la disponibilità del prodotto.

# Scopi della Conjoint Analysis - 7

Generalizzando, la Conjoint Analysis permette di rispondere alle seguenti domande:

- Quali nuovi prodotti avranno successo?
- Quali attributi del prodotto/servizio guidano le decisioni di acquisto?
- Quali cambiamenti nel progetto del prodotto faranno aumentare le preferenze dei consumatori?
- Qual è il prezzo ideale?
- Esiste un segmento di mercato per un certo prodotto?
- Quale sarà la risposta ad un certo tipo di pubblicità (o qual è la migliore?)

# Caratteristiche della Conjoint Analysis

-la sua natura “decompositiva”: da un giudizio su un prodotto intero si ricavano le preferenze per ogni attributo particolare e per ogni suo livello, ossia dei “*part-worth*” che combinati tra loro secondo una certa regola (es. additiva), più si avvicinano al giudizio globale del cliente;

-le stime possano essere fatte a livello individuale, così da ricavare un modello predittivo per ogni consumatore esaminato; in effetti, generalmente ci sono molte variazioni nelle preferenze tra persone diverse;

-la flessibilità nelle relazioni tra variabili dipendenti e indipendenti: la relazione esistente non è definita a priori (in genere negli altri metodi si assume una relazione lineare semplice); si possono fare stime separate degli effetti per ogni livello della variabile indipendente senza assumere che esista un legame tra livelli. La flessibilità dipende così dalle scelte fatte dall’analista (come dice Louviere, ’94), non esiste un approccio unico e la conjoint analysis è “ciò che l’analista stesso fa”).

# Terminologia della Conjoint Analysis - 1

Quelli che seguono sono i termini comunemente usati per questa metodologia con i relativi significati:

-**Fattori e attributi:** sono le variabili che il ricercatore controlla per misurare l'effetto su un'altra variabile.

Il fattore è la variabile che rappresenta un certo attributo o caratteristica del prodotto/servizio in esame.

Nella C.A. spesso i fattori spesso sono variabili non metriche e devono essere espresse mediante due o più valori chiamati livelli.

# Terminologia della Conjoint Analysis - 2

- Profilo o combinazione o Stimolo:** è una specifica combinazione di livelli degli attributi.
- Livello:** è ogni possibile valore che può assumere un certo fattore.
- Pesi parziali:** sono i giudizi su un particolare livello di un attributo ricavabili dal peso complessivo attribuito dal singolo rispondente alle alternative di prodotto.

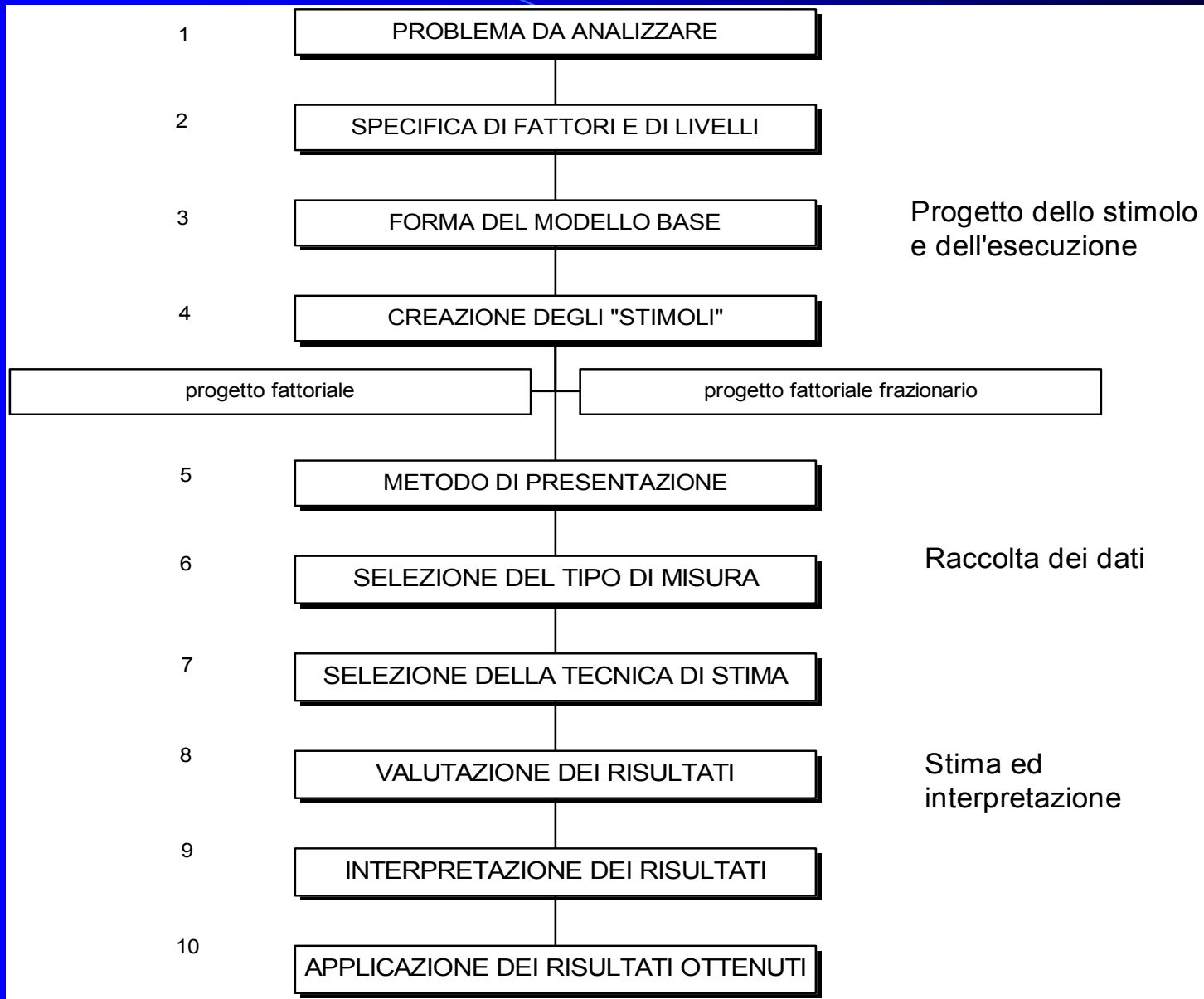
# Terminologia della Conjoint Analysis - 3

- **La misura della preferenza:** questa rappresenta la variabile di osservazione. Questa misura può essere realizzata tramite due diverse modalità: il **Rating** ed il **Ranking**. Il primo prevede la formulazione da parte del consumatore di un punteggio di utilità o preferenza. Il secondo prevede che il rispondente componga un ordinamento sempre in base alla preferenza o all'utilità percepita.

# Sequenza di attività - 1

Consideriamo ora la **sequenza di attività** necessarie per ottenere una conjoint analysis, così come suggerito da Hair et al.(1992). Si tratta di uno schema molto dettagliato, che comprende tutti i passi da compiere per un'analisi completa, suggeriti già da Green e Srinivasan ('78).

# Sequenza di attività - 2



# Sequenza di attività - 3

## Fase 1: Il problema

Prima di pianificare l'analisi vera e propria è necessario inquadrare il tipo di problematica che si andrà a valutare. Si tratta di individuare, magari tramite una ricerca "esplorativa" (*focus group*, interviste telefoniche, interviste "in profondità"), eventuali "trabocchetti" (Louviere, '88) nelle generalizzazioni che l'analista sarebbe portato a fare circa il comportamento del cliente: è importante, per esempio, capire come quest'ultimo identifichi ed esprima a parole certe caratteristiche del prodotto; in generale, si tratta di elaborare un "modello preventivo" di come il cliente si comporti in particolari condizioni, sul quale far poggiare le fasi successive (a cominciare proprio dalla seconda).

Alcuni quesiti tipo possono essere:

- si possono descrivere tutti gli attributi che danno utilità e valore al prodotto in esame?
- qual è il criterio di scelta base per questo particolare tipo di prodotto?
- si deve dare precedenza alla spiegazione del processo di scelta da parte del consumatore (focalizzarsi quindi sull'importanza degli attributi) oppure si cerca di prevedere la quota di mercato per il prodotto?

# Sequenza di attività – 3a

## Fase 2: Selezione di fattori e livelli

In genere si hanno a disposizione i risultati di un'analisi preliminare per effettuare questa attività. Nella scelta degli attributi rilevanti si devono considerare:

- il valore globale del prodotto: vanno inclusi tutti gli attributi che creano potenziale valore nel prodotto e anche quelli che eventualmente ne tolgono, permettendo così una valutazione il più possibile corretta;

# Sequenza di attività – 3b

Vanno poi selezionati i livelli per ogni fattore, tenendo presente che:

- le caratteristiche selezionate devono esprimere un concetto ben preciso, il significato dell'una non deve “sconfinare” in quello di un'altra;
- il concetto deve essere ben espresso per evitare fraintendimenti;
- il numero di livelli va quanto più possibile bilanciato per i diversi fattori;
- il campo di oscillazione dei livelli va stabilito un po' fuori dai valori reali, così da ridurre la “correlazione tra attributi”, senza però arrivare a livelli non realistici: nel caso di attributi “continui”, i livelli scelti devono essere sufficientemente distinti tra loro; anche nel caso di attributi “categorici” i livelli vanno scelti attentamente in modo da essere rappresentativi del fattore stesso (Green e Srinivasan, '78);
- bisogna evitare la “collinearità multipla” tra attributi: ogni variabile indipendente deve essere slegata dall'andamento delle altre (non deve esserci regressione di una sulle altre). Gli attributi devono quindi essere ortogonali. Se esistono legami che creano poi stimoli non validi (es. due attributi compresenti nello stimolo ma mai nella realtà), si può creare un “superattributo” che racchiuda gli aspetti degli attributi correlati, oppure si deve eliminare uno dei fattori.

# Sequenza di attività – 4

## Fase 5: Il metodo di raccolta dati e di presentazione

I metodi più comuni per raccogliere i dati sono due:

-l'approccio “*trade-off*”: vengono comparati solo due attributi alla volta; è poco realistico (un prodotto ha più di due attributi in genere), richiede di esprimere molti giudizi, permette solo risposte non metriche (giudizi non numerici) e non può far uso del metodo frazionario per ridurre il numero di combinazioni. Per questi motivi è usato sempre meno;

-l'approccio “*full profile*”: in questo caso si può usare il metodo fattoriale frazionario e ogni stimolo viene descritto separatamente (in genere su un cartellino); si possono chiedere vari tipi di giudizio, da quello sulla volontà di prova a quello sull'intenzione d'acquisto. Tuttavia, se ci sono molti fattori si rischia che l'intervistato ne consideri solo alcuni, dando così informazioni sovrabbondanti; è importante poi che l'ordine con cui si presentano i fattori sui cartellini cambino di volta in volta per ogni intervistato

# Sequenza di attività – 5

In generale se i fattori sono pochi (6 o meno) si usa (come suggerito da Green e Srinivasan, '78) il secondo metodo; se i fattori sono in numero da 7 a 10 si ricorre al trade-off. Green e Srinivasan ('78) ipotizzano che l'approccio “*full-profile*” dia stime più accurate laddove i fattori siano pochi, e le interazioni tra fattori siano ampie.

Esistono poi approcci alternativi come: i modelli “*self-explicated*” ed i modelli “*ibridi/adaptive*”.

# Sequenza di attività – 6

## Fase 6: Scelta del tipo di misura di preferenza

Il metodo *trade-off* permette solo un giudizio di tipo numerico (es. su una scala da 1 a 10), mentre con il *full-profile* si possono anche ordinare le combinazioni presentate secondo il giudizio di preferenza (dallo stimolo che piace meno a quello preferito). Questo secondo approccio è più facile per l'intervistato e più flessibile (si può fare una stima con tipi diversi di regole di composizione), ma l'analisi poi è più difficile da gestire poiché si deve mettere a punto un'intervista personale alla volta, vista la soggettività dell'ordinamento dei cartellini. Nel caso di giudizi numerici, è buona cosa proporre molte categorie di risposta (es. da 1 a 11) se ci sono molti stimoli da valutare.

# Sequenza di attività – 7

## Fase 7: Selezione della tecnica di stima

In generale, Green e Srinivasan ('78) suddividono i metodi di stima in tre classi:

- - Metodi in cui si assume che la variabile dipendente sia, al più, ridotta in scala di tipo ordinale (*rank-ordering*): si hanno allora la MONANOVA (analisi di varianza monotona), il LINMAP o il PREFMAP;
- -Metodi per variabile dipendente in scala ad intervallo (*rating*): in questa classe c'è la regressione multipla di tipo OLS (*ordinary least square*, metodo dei minimi quadrati);
- Metodi che relazionano dati da confronti a coppie con modelli di probabilità di scelta: qui ci sono il LOGIT e il PROBIT soprattutto.

La MONANOVA viene applicata solo a modelli di tipo “*part-worth function*”, mentre il LINMAP è utile nel caso di modelli “a punto ideale”, poiché usa la programmazione lineare. La MONANOVA, introdotta da Kruskal nel '65 (Louviere, '88) assume un modello di tipo additivo ed usa un algoritmo iterativo per stimare i pesi degli attributi così che il *rank order* per ogni combinazione di livelli di attributi sia il più possibile correlato al *rank order* osservato per quella data combinazione. Il PREFMAP è un altro algoritmo che stima anche le interazioni tra attributi.

# Sequenza di attività – 8

## Fase 8: Valutazione dei risultati

Poiché si misura una grandezza non fisica (la preferenza) utilizzando un questionario, è importante valutare la bontà dello strumento di misura stesso: questo deve essere *affidabile*, cioè deve produrre le stesse misure qualora cambino le condizioni esterne (es. l'intervistatore) e deve anche essere *valido*, cioè deve misurare proprio quanto si propone di considerare.

# Sequenza di attività – 9

## Fase 9: Interpretazione dei risultati

L'analisi può essere fatta a livello aggregato o disaggregato: in genere si usa questo secondo approccio, per cui per ogni persona intervistata viene costruito un modello. Si viene a conoscenza così del comportamento di ogni consumatore. Con l'approccio aggregato, si prepara un modello unico, perdendo informazioni circa i gusti del singolo: è un'analisi quindi poco appropriata, soprattutto quando la popolazione è disomogenea, anche se è utile qualora si voglia stimare per esempio la quota di mercato di un prodotto (obiettivo di ricerca dove conta il comportamento generale).

Oltre a descrivere l'impatto di ogni livello, si stima anche l'importanza relativa di ogni fattore (in percentuale) per ogni intervistato.

# Sequenza di attività – 10

## Fase 10: Applicazione dei risultati ottenuti

I modelli ottenuti (uno per persona) possono servire per:

- decisioni di segmentazione: si raggruppano i consumatori con giudizi simili;
- analisi di redditività relativa di un certo progetto di prodotto: conoscendo il costo di ogni caratteristica, si può combinare il costo di ogni “prodotto” (combinazione) con la quota di mercato prevista per osservarne la fattibilità;
- simulazioni: si può prevedere la quota di mercato che può raggiungere un certo “stimolo” in diversi scenari competitivi; si può stimare cosa succeda se si aggiunge un nuovo prodotto nel settore; si può capire la dinamica di una strategia multiprodotto (effetto della cannibalizzazione).

# Pianificazione dell'indagine di Conjoint Analysis

E' una delle fasi più delicate della Conjoint Analysis e consiste nella selezione degli attributi e nella formulazione dei profili da sottoporre al giudizio degli intervistati. Esistono diversi tipi di approcci tra cui:

**-Traditional Conjoint Analysis (Full profile):** questo metodo prevede l'utilizzo di un piano fattoriale completo costituito da tutte le possibili combinazioni dei livelli di tutti i fattori o di un piano fattoriale frazionato che consente di esprimere un numero più ridotto di profili;

# Le Fasi Operative della Conjoint Analysis - 1

Le fasi in cui si articola la Conjoint Analysis sono le seguenti:

- Individuazione degli attributi del prodotto/servizio e della loro più corretta suddivisione in modalità o livelli;
- Definizione dei profili di prodotto, (combinazione o stimoli), basati su un piano sperimentale da sottoporre al giudizio degli intervistati;
- Scelta di un campione di valutatori;
- Somministrazione dei profili di prodotto ai rispondenti. La presentazione di questi profili può essere fatta nelle seguenti modalità: a) descrizione scritta (elenco delle combinazioni dei livelli dei vari fattori, presentate sui cartellini); b) descrizione verbale; c) materiale illustrativo come modellini, prototipi.; d) mezzi multimediali;
- Stima dei parametri associati a ciascuna modalità degli attributi di prodotto, attraverso la funzione di utilità parziale degli attributi;
- Stima dell'importanza relativa di ciascun attributo/fattore;
- Valutazione dell'utilità totale associata a profili virtuali, non considerati inizialmente nel piano di rilevazione.

# Le Fasi Operative della Conjoint Analysis - 2

Da un punto di vista puramente matematico-statistico lo scopo della Conjoint Analysis è quello di scomporre, in relazione a ciascun rispondente, le valutazioni globali, espresse attraverso le combinazioni di modalità degli attributi o profili di un prodotto, nelle valutazioni delle singole modalità degli attributi stessi. La valutazione delle utilità parziali consente di determinare le funzioni di utilità individuali, dato il modello fissato (attributi e livelli considerati).

# Modelli di preferenza e funzioni di risposta - 1

La scelta del numero di modalità dei fattori, da cui deriva il numero di combinazioni o stimoli sperimentali, viene realizzato in base alle ipotesi formulate sulla **forma della funzione di risposta** considerata (ad esempio, lineare, quadratica, ecc...).

Tale funzione può essere anzitutto di tipo **additivo**, in cui vengono considerati gli effetti principali, o **misto**, cioè, ad effetti principali e ad effetti di interazione.

Green e Srinivasan (1990) suggeriscono che i modelli di preferenza principali, in ordine crescente di generalità:

- il modello **vettore**, lineare
- il modello **punto-ideale**, lineare quadratico
- il modello **part-worth**, discontinuo

# Il modello part-worth

Nel modello **additivo a coefficienti separati** (part worths), in corrispondenza di ogni modalità di un fattore viene stimato un coefficiente di “utilità parziale”. Dati  $P$  fattori, con  $m_1, m_2, \dots, m_p$  modalità o livelli (che possono essere sia qualitative che quantitative), il numero delle possibili combinazioni delle stesse è dato dal prodotto cartesiano  $\prod_{p=1}^P m_p$ . Il numero di coefficienti di utilità da stimare, nel modello ad effetti principali, risulta quindi pari a:

$$\sum_{p=1}^P (m_p - 1) = \sum_{p=1}^P m_p - P$$

Pertanto, ad esempio, dati 5 fattori a 2 livelli ciascuno, cioè  $p = 1, 2, \dots, P = 5$ , con  $m_p = 2$ , si hanno  $2^5 = 32$  possibili combinazioni di livelli ed un numero di coefficienti di utilità, cioè un numero di incognite, pari a  $\sum_{p=1}^5 m_p - P = 10 - 5 = 5$ ; in questo caso si necessita di almeno 5 combinazioni sperimentali per giungere a tante equazioni quante sono le incognite. La funzione di utilità sarà poi del tipo:

$$y_j = \sum_{p=1}^P s_i f_{ij}$$

dove  $s_i$  sia una funzione discontinua definita per l'insieme della modalità degli attributi.

# Tipi di Conjoint Analysis -1

I fattori del prodotto possono essere qualitativi (sconnessi od ordinali) e/o quantitativi. I giudizi di risposta si possono rilevare in forma:

- - di posizioni o ranghi (**rank order**), cioè di graduatoria di preferenza dei profili del prodotto sottoposti a giudizio;
- - di punteggio (**rating scales**), con giudizi assegnati su scala di punteggio.

Nella pratica aziendale le valutazioni su scala di punteggio vengono ipotizzate in genere su scala ad intervalli, applicando, di conseguenza, la cosiddetta **C.A. metrica**.

Quando la variabile risposta è ritenuta e trattata su scala puramente ordinale, si usa, invece, la **C.A. non metrica**. In ambito aziendale, tuttavia, si è imposto l'approccio metrico, sotto l'ipotesi semplificatrice, per la stima dei parametri del modello, che la variabile risposta sia su scala ad intervalli. La **C.A. metrica** sarà anche l'approccio utilizzato nel prosieguo del presente corso.

# Modello Lineare per la Conjoint Analysis Metrica - 1

Disponibili i dati di base, il 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 degli attributi componenti il singolo profilo, dia luogo a valori di preferenza/utilità totali che riproducano il più fedelmente possibile l'ordinamento di preferenza o i punteggi di gradimento espressi.

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 tutto sommato equivalenti.

# Modello Lineare per la Conjoint Analysis Metrica - 2

La **regressione lineare multipla** con variabili di tipo *dummy* costituisce la tecnica maggiormente applicata quando le valutazioni fornite dagli intervistati sono misurate su una scala che può essere ragionevolmente ritenuta quantomeno ad intervallo. Vista la sua ampia diffusione, è il metodo illustrato dettagliatamente nel seguito.

# Modello Lineare per la Conjoint Analysis Metrica - 3

Nella **C.A. metrica** la stima delle funzioni di utilità individuali viene effettuata “direttamente” sui valori dei giudizi di valutazione globale ipotizzati in maniera semplicistica su scala metrica, tramite la tecnica di regressione lineare multipla, applicando preliminarmente alle categorie delle variabili esplicative qualitative (fattori) il seguente schema di codifica binaria disgiuntiva completa:

$$\mathbf{d}_{mkp} = \begin{cases} 1, & \text{se il profilo } m \text{ presenta l'attributo } k, \text{ con livello } p \\ 0 & \text{altrove} \end{cases}$$

Questa è sostanzialmente la definizione delle **variabili dummy**, dove è indicato con:

**m** la generica combinazione  $\mathbf{m} = 1, 2, \dots, M$  dei livelli dei vari attributi;

**k** il generico attributo del prodotto;  $\mathbf{k} = 1, 2, \dots, K$ ;

**p** il generico livello del fattore  $k$ ,  $\mathbf{p} = 1, 2, \dots, P_k$ .

# Modello Lineare per la Conjoint Analysis Metrica - 4

La funzione di utilità parziale (non lineare) del fattore  $k$ , per il profilo  $m$ , in riferimento ad un generico consumatore  $i$ , si esprime nel modo seguente (Lilien Kotter, 1983):

$$u_{ik} = \sum_{p=1}^{P_k} w_{ikp} \cdot d_{mkp}$$

$u_{ik}$  indica l'utilità che il fattore  $k$  procura all' $i$ -esimo rispondente;

$w_{ikp}$  è il coefficiente di regressione che esprime l'importanza attribuita dall' $i$ -esimo rispondente al fattore  $k$ , considerato al livello  $p$ ; si può altresì intendere come un peso relativo all' $i$ -esimo rispondente. Da notare che, poiché  $d_{mkp} = 1$  solo per un livello del fattore  $k$ ,  $u_{ik}$  corrisponde all'utilità del livello medesimo per il fattore  $k$ , con riferimento al profilo  $m$ ;

$D_{mkp}$  è la modalità della variabile dummy, già definita in precedenza.

# Modello Lineare per la Conjoint Analysis Metrica - 5

In riferimento a tutti i  $k$  fattori, la funzione di utilità totale  $R_{im}$  dell' $i$ -esimo rispondente, per il profilo  $m$  del prodotto si esprime secondo il seguente modello:

$$R_{im} = \sum_{k=1}^k \sum_{p=1}^{P_k} w_{ikp} \cdot d_{mkp} + e_{im}$$

dove:

$R_{im}$  è l'utilità globale dell' $i$ -esimo rispondente

$w_{ikp}$  è il peso relativo all' $i$ -esimo rispondente per l'attributo  $k$  e livello  $p$ .

$e_{im}$  è la componente di errore casuale, relativa all' $i$ -esimo rispondente, ed è così definita  $e \sim N(0, \sigma^2)$ . Questa importante assunzione viene fatta sulla base di due considerazioni:

- Sul fatto che il numero delle risposte sia sufficientemente elevato (generalmente maggiore di 10).
- Sul concetto di randomizzazione, e cioè sull'ordine casuale con cui i profili vengono presentati ai rispondenti.

# Modello Lineare per la Conjoint Analysis Metrica - 6

Questa componente è inserita al fine di ottenere che i risultati siano generalizzabili. Si può fare questo dal momento che si assume di conoscere la legge di probabilità sottostante che in qualche modo governa le preferenze dei rispondenti e trasferire quindi la generalizzabilità dal campione alla popolazione.

La risposta  $R_{im}$  ha quindi due componenti:

- una componente che deriva dalla variabile dummy
- una componente che presenta una parte casuale.

A questo punto si è in grado di stimare i valori dei  $w_{ikp}$  ed estenderli alla popolazione di riferimento.

# Le combinazioni dei livelli dei fattori

Nel determinare la scelta del tipo di approccio per la C.A., va tenuto presente come l'uso di un **piano fattoriale completo**, che consente di stimare sia gli effetti principali che quelli di interazione delle varie modalità degli attributi di prodotto, comporti un numero di risposte considerevole per l'intervistato, appena il numero dei fattori o delle corrispondenti modalità supera le poche unità. Combinando in tutti i modi possibili i livelli  $P_1, P_2, \dots, P_k$  di  $K$  fattori ( $k = 1, 2, \dots, K$ ), secondo un "piano fattoriale completo", si ottiene, un numero di combinazioni pari a  $\prod_k P_k$ . Appena il numero  $K$  dei fattori e/o dei livelli aumenta, le combinazioni dei profili del prodotto analizzato crescono numericamente in modo considerevole e, oltre un certo limite non sono più proponibili per il carico di risposte che essi comportano al giudizio dei valutatori. Per tale ragione, anziché fare riferimento all'insieme di tutte le possibili combinazioni, si considera una sua frazione (**piano fattoriale frazionato**, ad esempio, di  $M$  combinazioni). I piani frazionati non consentono, tuttavia, di stimare gli effetti di interazione dei fattori, che risultano, in questo modo, confusi con gli effetti principali. Punto di forza della metodologia è quello di stimare ex post le valutazioni "virtuali" delle combinazioni non sottoposte a giudizio, di numero pari a  $\prod_k P_k - M$ , giungendo in conclusione, comunque, ad una stima di utilità per tutti i possibili profili di prodotto.

# Il Modello da Stimare - 1

Il modello di regressione di utilità totale, da stimare per il singolo rispondente (si stimano tanti modelli quanti sono i rispondenti) è, in notazione algebrica e con simbolismo matriciale:

$$\mathbf{y} = \mathbf{Z} \boldsymbol{\beta} + \mathbf{e}$$

dove :

- $\mathbf{y}$  è il vettore colonna dei punteggi di valutazione osservati.
- $\mathbf{Z}$  è la matrice del piano sperimentale, cioè delle variabili indicatrici delle variabili indipendenti.
- $\boldsymbol{\beta}$  è il vettore colonna dei coefficienti incogniti.
- $\mathbf{e}$  è il vettore colonna dei residui.

# Il Modello da Stimare - 2

Il sistema di equazioni normali associato alla formula precedente risulta indeterminato, essendo la matrice  $\mathbf{Z}$  del piano sperimentale singolare. Per rendere risolvibile il sistema si deve introdurre una condizione complementare (De Luca, 2000), che può realizzarsi, ad esempio, nel sopprimere, con scelta arbitraria, una colonna in ciascun blocco di **variabili dummy** e nel porre pari a zero il corrispondente coefficiente. Nella matrice  $\mathbf{Z}$  viene soppresso, perciò, un numero di colonne pari al numero di fattori. Si giunge così alla matrice del piano sperimentale  $\mathbf{Z}_0$ .

Il modello di regressione si rappresenta, quindi, nel modo seguente:

$$\mathbf{y} = \mathbf{Z}_0 \boldsymbol{\beta}_0 + \mathbf{e}$$

- $\mathbf{y}$  è il vettore colonna dei punteggi di valutazione;
- $\mathbf{Z}_0$  è la matrice del piano sperimentale, in cui è stata soppressa la variabile indicatrice corrispondente all'ultima colonna, in ciascun blocco di **variabili dummy**;
- $\boldsymbol{\beta}_0$  vettore colonna dei coefficienti incogniti; corrisponde al vettore delle utilità parziali cercate;
- $\mathbf{e}$  è il vettore colonna dei residui.

# Il Modello da Stimare - 3

Con questo accorgimento, la matrice  $Z_0$  non risulta più singolare e quindi il vettore delle stime dei coefficienti cercati, indicato con  $b_0$  si ricava da questa formula:

$$b_0 = (Z_0' Z_0)^{-1} Z_0 y$$

dove  $Z_0'$  indica la matrice trasposta di  $Z_0$ .

Gli elementi della matrice  $Z_0$  sono fissi per tutti i rispondenti (gli stimoli sperimentali sono infatti i medesimi per tutti i valutatori), variano invece (normalmente) i valori assunti dalla variabile dipendente  $Y$  che corrispondono ai punteggi di valutazione.

# Il Modello da Stimare - 4

Applicando la formula precedente ai dati di ciascun rispondente si ottengono i coefficienti di utilità parziale (parth worth) delle modalità dei fattori. Tali coefficienti sono stimati con il principio dei minimi quadrati ordinari, minimizzando cioè la somma dei quadrati degli scarti tra i punteggi assegnati  $\mathbf{Y}$  dal rispondente ed i corrispondenti valori stimati  $\hat{\mathbf{Y}}$ . Il principio appena illustrato si esprime con la seguente espressione:

$$\sum_m (\mathbf{Y}_m - \hat{\mathbf{Y}})^2 = \text{minimo}$$

dove:

- $\mathbf{M}$  è il numero di profili del disegno fattoriale;
- $\mathbf{Y}_m$  indica il valore di punteggio osservato relativamente al profilo m-esimo;
- $\hat{\mathbf{Y}}$  indica il valore del punteggio calcolato con la funzione di regressione, sempre relativamente al profilo m-esimo.

# Le Baseline

E' da osservare che considerando congiuntamente la categoria soppressa dei vari attributi, si ottiene la “classe di riferimento” o **baseline**, rispetto alla quale vengono misurati gli effetti delle restanti categorie dei diversi fattori.

Le **baseline** sono, in altre parole, le utilità eliminate, che forniscono il livello di utilità di riferimento. Le utilità dei livelli degli attributi rimasti sono calcolate rispetto a queste in termini di distanza dalla **baseline** corrispondente

# Valutazione della bontà del modello - 1

Per valutare il **grado di bontà del modello**, si calcola il *coefficiente di determinazione*:

dette  $s_y^2$  e  $s_{y'}^2$ , rispettivamente, la varianza dei valori stimati  $y_i$  e la varianza dei valori osservati  $y_i$ ?

$$s_y^2 = \frac{1}{n} * \left[ \sum (y_i - y_{medio})^2 \right]$$

$$s_{y'}^2 = \frac{1}{n} * \left[ \sum (y_i - y'_{medio})^2 \right]$$

il coefficiente di determinazione è:

$$R^2 = \frac{s_y^2}{s_{y'}^2}$$

Qualora  $R^2$  fosse pari a 1 significherebbe che il modello rappresenta perfettamente l'andamento della variabile dipendente nella realtà.

# Valutazione della bontà del modello - 2

La radice quadrata del coefficiente di determinazione, R, è il *coefficiente di correlazione multipla*.

I *coefficienti di correlazione parziale*, ossia quelli che permettono di misurare il legame tra la variabile dipendente e una variabile indipendente a parità di tutte le altre, sono dati da:

$$R_{YX_1, X_2, X_3} = \sqrt{1 - \frac{1 - R_{X_1 X_2 X_3}^2}{1 - R_{X_2 X_3}^2}}$$

dove i primi due indici di R sono le variabili di cui si considera la correlazione, e i secondi indici sono le variabili rimanenti.

Il *coefficiente di correlazione lineare* può essere espresso dalla formula dei momenti misti:

$$R = \frac{(\sum x * y)}{\sqrt{(\sum x^2) * (\sum y^2)}}$$

dove  $x = X - X_{\text{medio}}$  e  $y = Y - Y_{\text{medio}}$ .

# Valutazione della bontà del modello - 3

Poiché si misura una grandezza non fisica (la preferenza) utilizzando un questionario, è importante valutare la bontà dello strumento di misura stesso: questo deve essere *affidabile*, cioè deve produrre le stesse misure qualora cambino le condizioni esterne (es. l'intervistatore) e deve anche essere *valido*, cioè deve misurare proprio quanto si propone di considerare.

Riguardo quest'ultimo aspetto, si tratta di valutare i risultati per esaminare l'accuratezza dei modelli ottenuti nel predire le preferenze dei consumatori (validità interna). Per giudizi metrici si può usare la correlazione. Per dati di tipo “*rank-order*” (non metrici) si usano correlazioni basate sugli ordinamenti reali e previsti, come per esempio il *rho* di Spearman o il *tau* di Kendall (Hair et al., '92).

# Valutazione della bontà del modello - 4

Il rho di Spearman è un *coefficiente di cograduazione* per misurare la relazione esistente tra due variabili graduate per dimensione o per importanza crescenti. Nel nostro caso, si considera da un lato l'ordinamento dei concetti così come rilevato nella realtà (secondo il giudizio di preferenza dei clienti), dall'altro l'ordinamento ricavabile dal modello ottenuto. Detta D la differenza tra ranghi di valori corrispondenti delle due osservazioni (reale e stimata) e N il numero di coppie di dati, si ha:

$$r_s = 1 - \frac{(6 * \sum D^2)}{[N * (N^2 - 1)]}$$

# Valutazione della bontà del modello - 5

Il *tau di Kendall* è un metodo per esprimere la corrispondenza tra le preferenze espresse dagli intervistati e i pesi stimati. Consideriamo per semplicità un esempio: siano A, B, C, D quattro concetti valutati con giudizi “*rank order*” di cui si siano stimati i pesi (“*utilities*”)

<u>concetto</u>	<u>rank</u>	<u>utility</u>
A	3	2.3
B	4	4.5
C	7	1.7
D	9	-6.1

I pesi stimati dovrebbero rispecchiare l'ordine di preferenza (pesi maggiori ai concetti preferiti) e così dovrebbe essere:

A preferito a B    A preferito a C

A preferito a D    B preferito a C

B preferito a D    C preferito a D

In realtà, le differenze tra le “*utilities*” stimate sono:

$$2.3 - 4.5 = -2.2$$

$$2.3 - 1.7 = 0.6$$

$$2.3 + 6.1 = 8.4$$

$$4.5 - 1.7 = 2.8$$

$$4.5 + 6.1 = 10.6$$

$$1.7 + 6.1 = 7.8$$

# Valutazione della bontà del modello - 5

Si vede così che una coppia non soddisfa quanto espresso dal cliente (compare il segno negativo); *tau* è pari a:

$$\tau = \frac{n^{\circ} \text{ confrontigliusti} - n^{\circ} \text{ confrontisbagliati}}{n^{\circ} \text{ confronti}} = \frac{5 - 1}{6} = 0,666$$

In realtà è un indicatore poco utile per procedere poi con un'ottimizzazione del modello.

# L'importanza relativa dei Fattori - 1

Nell'ambito della metodologia di Conjoint Analysis è consuetudine valutare, anche se in maniera piuttosto semplicistica, l'importanza relativa dei fattori. Questo viene fatto con lo scopo di rendere comparabili le utilità parziali e di pervenire a valori di importanza dei fattori compresi tra 0 e 1 (o ovviamente in percentuale).

La procedura seguita è la seguente:

- a) per ciascun fattore si determina la differenza tra l'utilità parziale più elevata e l'utilità più bassa delle sue modalità;
- b) si sommano i campi di variazione, descritti al punto a), di tutti i fattori;
- c) si calcola, per ciascun fattore, il rapporto tra il campo di variazione descritto al punto a) e la somma dei campi di variazione introdotti nel punto b);

# L'importanza relativa dei Fattori - 2

In modo formale, indicando con  $F_i$  un generico fattore, la sua importanza relativa  $\mathbf{Ir}_{F_i}$ , con  $P$  numero di fattori, viene espressa nel modo seguente:

$$\mathbf{Ir}_{F_i} = \frac{(\text{Utilità parziale più grande di } F_i - \text{Utilità parziale più piccola di } F_i)}{\sum_{i=1}^P (\text{Utilità parziale più grande di } F_i - \text{Utilità parziale più piccola di } F_i)}$$