



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

# FLORE

## Repository istituzionale dell'Università degli Studi di Firenze

### Rilevazione e analisi statistica del dato soggettivo

Questa è la Versione finale referata (Post print/Accepted manuscript) della seguente pubblicazione:

*Original Citation:*

Rilevazione e analisi statistica del dato soggettivo / F. MAGGINO. - ELETTRONICO. - (2007), pp. 1-304.

*Availability:*

This version is available at: 2158/328150 since:

*Publisher:*

Firenze University Press, Archivio E-Prints

*Terms of use:*

Open Access

La pubblicazione è resa disponibile sotto le norme e i termini della licenza di deposito, secondo quanto stabilito dalla Policy per l'accesso aperto dell'Università degli Studi di Firenze (<https://www.sba.unifi.it/upload/policy-oa-2016-1.pdf>)

*Publisher copyright claim:*

(Article begins on next page)

# Parte Introduttiva

## Premessa

La struttura di una scienza è frutto di una complessa interazione e sintesi tra due livelli attraverso i quali è possibile esprimere le conoscenze: uno formale (che si esprime a livello logico e/o matematico) e uno empirico.

Le due forme di espressione corrispondono a due livelli di analisi scientifica, tra loro legati anche se analiticamente distinti:

- Livello formale: richiede lo *sviluppo di teorie, di ipotesi e di spiegazioni teoriche dei fenomeni studiati*. La teoria può essere considerata come una rete di *concetti o costrutti* – astratti e definiti e delle loro connessioni e relazioni reciproche. Tale livello coinvolge il ricercatore teorico.
- Livello empirico: richiede la *verifica delle ipotesi astratte attraverso osservazioni e dati empirici*; tale livello interessa principalmente il ricercatore. Tale livello coinvolge il ricercatore applicato.

Entrambi i livelli sono fondamentali e inscindibili. Nelle scienze più avanzate i due tipi di conoscenza interagiscono tra loro integrandosi.

La teoria può rappresentare la realtà a livelli diversi di accuratezza. Se esso fornisce un ritratto fedele di un sistema empirico allora le implicazioni teoriche devono risultare confrontabili con la realtà empirica.

L'interazione tra teoria e osservazioni empirica è rappresentata dalla *misurazione*. Infatti, la misurazione da una parte richiede la costruzione di un modello formale che rappresenta una astrazione e che consiste di aspetti selezionati di un sistema definito (Torgerson, 1958) e dall'altro consente di produrre osservazioni che se risultano coerenti con la teoria consentono di concludere che tale teoria fornisce una descrizione accettabile della realtà di interesse.

In questo senso la misurazione rappresenta una verifica di teorie in quanto comporta l'analisi della bontà d'adattamento di un modello astratto a una proprietà posseduta da una serie di oggetti empirici.

Conseguentemente si può dire che la misurazione non è mai immutabile: è sempre un tentativo di definire un'affermazione sulla natura della realtà. Un sistema di misurazione è potenzialmente aperto a qualsiasi revisione a causa della costante necessità di coerenza tra sistema astratto e mondo reale.

Nel trattare della misurazione è necessario distinguere tra:

- *sistemi*, comprendenti gli oggetti e le cose da misurare,
- *proprietà*, che si riferiscono ai loro aspetti osservabili.

La misurazione si riferisce alle proprietà dei sistemi e può essere vista come un modo per definire una proprietà.

Per rendere la misurazione utile è importante sviluppare una serie generale di assunti che riguardano le connessioni tra osservazioni empiriche, dati e metodi di analisi. Tali assunti devono essere sufficientemente generali da poter essere applicati virtualmente a qualsiasi situazione che un ricercatore può incontrare (DeVellis, 1991).

Il termine "misurazione" viene associato a molte situazioni anche molto comuni. In sede scientifica nel trattare di misurazione ci si riferisce al processo e alla logica coinvolti nella costruzione di uno strumento e alle proprietà che possono essere attribuite a questo.

La misurazione richiede un processo di astrazione che riguarda "cosa si misura". Infatti mentre l'osservazione è globale e può riguardare tutto un insieme, l'attività scientifica è rivolta ad individuare singoli aspetti di un fenomeno. Questo vuol dire che la misurazione riguarda sempre particolari caratteristiche o attributi posseduti dall'oggetto e non direttamente gli oggetti stessi.<sup>1</sup>

In realtà, rispetto alla definizione di misurazione, non esiste uniformità di veduta; sono molte infatti

---

<sup>1</sup> Su tale distinzione si fa spesso molta confusione; si pensi a tale proposito alla difficoltà di comprendere come un delinquente ed una persona onesta possano avere lo stesso livello di intelligenza.

le definizioni di misurazione soprattutto quelle che si riferiscono alla misurazione nell'ambito delle scienze sociali; vediamo alcune:

- *la misurazione di grandezze è, in senso molto generale, qualsiasi metodo attraverso il quale viene stabilita una corrispondenza unica e reciproca tra tutte o alcune grandezze di un tipo e tutti o alcuni numeri (reali, razionali, integrali) (B. Russell negli anni trenta);*
- *la misurazione è l'assegnazione di numeri che rappresentino proprietà di sistemi materiali diversi da numeri, in virtù di leggi che governano tali proprietà (D.T. Campbell nei suoi lavori degli anni trenta);*
- *la misurazione è l'assegnazione di numeri a oggetti o elementi secondo regole esplicite e convenzionali (Stevens, 1951);*
- *la misurazione nelle sue forme più semplici consiste nel sostituire simboli o nomi agli oggetti concreti (Coombs, 1950);*
- *data una variabile, la misurazione è stabilire una corrispondenza tra i valori della variabile e un insieme di numeri (Galtung, 1967);*
- *la misurazione è un processo generale attraverso il quale si assegnano numeri ad oggetti in modo che si comprenda anche quali tipi di operazioni matematiche possano essere legittimamente usati (Blalock, 1982).*

La definizione che trova molti d'accordo è quella che indica la misurazione come un processo che *lega concetti astratti a eventi empirici* e che implica un *piano esplicito e organizzato di classificazione* degli oggetti in riferimento ad un concetto generale; in relazione a ciò è possibile identificare due livelli di interesse:

- interesse teorico, concentrato sul *concetto astratto sottostante non direttamente osservabile*,
- interesse empirico, concentrato sul *valore osservabile*.

## ***IL MODELLO GENERALE DI MISURAZIONE DEL DATO SOGGETTIVO***

Quando il sistema misurato è composto da individui e le proprietà riguardano caratteristiche soggettive si parla di misurazione del soggettivo.

Perché la misurazione del soggettivo soddisfi le condizioni della misurazione scientifica, la misurazione e l'analisi del dato soggettivo richiedono la definizione di un disegno complesso composto da due elementi, uno teorico e l'altro metodologico.

L'aspetto teorico riguarda:

1. la definizione del concetto di “**soggettivo**” cioè la definizione e la individuazione delle caratteristiche che possono essere definite “soggettive”, del modo attraverso il quale possono essere osservate e dei problemi che la loro osservazione può produrre.
2. la identificazione della **teoria di misurazione** di riferimento che definisca le caratteristiche teoriche che rendono la misurazione *scientifica*; in altre parole, la teoria di riferimento, definendo il concetto di *errore di misurazione*, consente di identificare i modelli finalizzati alla verifica di:
  - obiettività ovvero la capacità di una procedura di misurare senza alterazioni dovute a fattori estranei, di essere libera da effetti dovuti al ricercatore; tale concetto si estende dalla procedura e di misurazione alla analisi dei dati e alla interpretazione dei risultati;
  - precisione, misurata attraverso il controllo della coerenza del modello di misurazione;
  - accuratezza ovvero la capacità di una procedura di misurare ciò che si intende misurare (contenuto).

Una procedura di misurazione che soddisfa tali requisiti non solo assume una rilevanza scientifica ma può anche essere *standardizzata*.

Gli aspetti metodologici sono legati alla definizione dei processi logici e ai modelli che riguardano la rilevazione (1) e l'analisi (2) del dato soggettivo. Entrambi i modelli sono basati su un fondamentale procedimento logico, detto *modello gerarchico*: tale modello è definito attraverso “componenti sequenziali” (a partire dalla definizione del modello concettuale a quella degli indicatori elementari); nel modello gerarchico, ciascuna componente del modello è definita e trova il suo significato nell'ambito della precedente (da cui il nome); ciò consente di ottenere dati soggettivi corretti, consistenti, interpretabili e complessi in riferimento alla complessità della struttura del modello gerarchico stesso; la sua definizione consente anche di identificare la popolazione di riferimento (e quindi anche il disegno di campionamento) e gli approcci metodologici per la raccolta dei dati.

1. La **rilevazione del dato soggettivo**, che richiede la definizione dei seguenti modelli:
  - il modello di costruzione del dato soggettivo: tale modello necessita di
    - una teoria dei dati,
    - un metodo per la gestione dei dati e
    - un procedimento per la costruzione del continuum sul quale posizionare i casi osservati relativamente alla caratteristica misurata,
  - il modello di attribuzione dei valori: tale modello consente di attribuire al dato rilevato un valore che lo renda interpretabile e trattabile in termini operativi (sistema di misurazione).
2. L'**analisi del dato soggettivo**, che richiede modelli che consentano di gestire e ridurre la complessità del dato rilevato. La strategia per ridurre la complessità è detta **di condensazione**. Dal punto di vista operativo, Il processo che definisce la condensazione corrisponde alla riduzione la dimensione della matrice dei dati. I dati prodotti attraverso l'applicazione del

modello gerarchico possono essere analizzati e condensati secondo due differenti prospettive e procedimenti:

- dagli indicatori elementari all'indicatore sintetico: si tratta della condensazione delle misure singole (indicatori elementari) in nuovi valori sintetici (indicatore sintetico) al fine di ricostruire l'unità del concetto, attraverso un'analisi che ripercorre il modello gerarchico a ritroso; la procedura di riduzione richiede l'adozione di
  - un modello di condensazione (*scaling model*),
  - procedure e di tecniche per
    - assegnare pesi a ciascun indicatore elementare (*weighting criteria*),
    - aggregare gli indicatori elementari (*aggregating-over-items techniques*),
    - verificare la robustezza dell'indicatore sintetico ottenuto,
    - verificare la capacità di discriminare dell'indicatore sintetico ottenuto;
- from individual-points to grouping-point: si tratta della condensazione delle misure che riguardano unità individuali (individual-points) presenti all'interno di aggregazioni significative (raggruppamenti) individuate secondo diverse prospettive (tipologica, geografica, ecc.); in questo caso, il modello di aggregazione richiede
  - l'identificazione del raggruppamento significativo (*condensing criteria*),
  - la definizione del peso da assegnare a ciascun caso individuale, i cui valori dovranno essere condensati in un *grouping-point* (*weighting criteria*),
  - l'adozione di una tecnica che consenta l'aggregazione dei valori individuali appartenenti a ciascun raggruppamento in uno o più valori rappresentativi (*grouping-point*) da attribuire all'insieme (*aggregating-over-individuals techniques*).

## LA DEFINIZIONE DI “DATO SOGGETTIVO”

Per poter procedere alla rilevazione dei dati soggettivi è importante definire e chiarire:

- quali sono le caratteristiche che devono essere osservate a livello di rappresentazione individuale,
- attraverso quali approcci possono essere rilevate,
- quali sono i problemi che sorgono nella misura delle caratteristiche individuali.

### Caratteristiche soggettive

Tradizionalmente, con “caratteristiche soggettive” si fa riferimento principalmente a tre aree di contenuto (Nunnally, 1978):

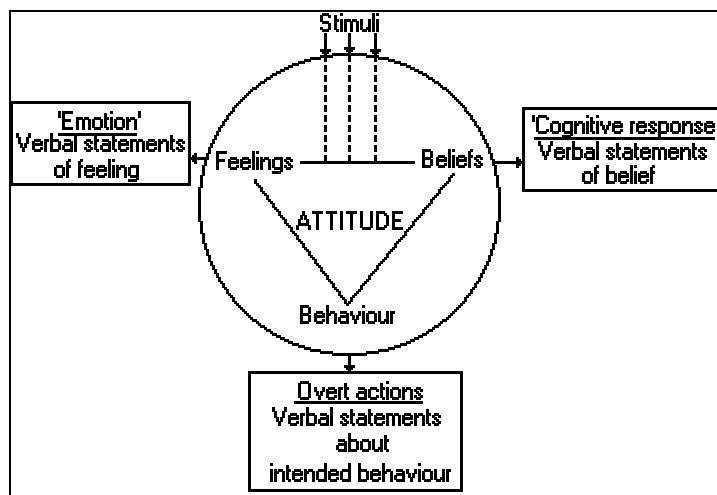
- **le capacità**, intendendo per capacità la abilità nell'eseguire determinati compiti (*performance* valutata rispetto ad uno specifico criterio); le capacità possono essere intellettive (ovvero quelle coinvolte in un'attività di tipo “scolastico” o in un lavoro di tipo “scientifico”) o speciali (ovvero competenze e abilità meccaniche, artistiche, di destrezza fisica); tra le capacità possiamo citare la comprensione verbale, la fluidità verbale, la capacità matematica, il ragionamento (induttivo e deduttivo), la memoria (meccanica, visiva, ecc.), l'orientamento spaziale, la percezione sensoriale;
- **i tratti di personalità**, intendendo per personalità quelle caratteristiche psicologiche che determinano i principi organizzativi e riflettono i modi in cui un soggetto risponde all'ambiente (*locus of control*, ego, introversione, autostima, identificazione, ecc.); è possibile identificare alcune categorie (che possono anche in parte sovrapporsi):

- *tratti sociali*, rappresentati dal comportamento caratteristico degli individui rispetto agli altri; tipici tratti sociali sono l'onestà, la socievolezza, la timidezza, l'autoritarismo, il senso dell'umorismo, la responsabilità, la religiosità, la generosità;
  - *motivazioni e bisogni personali*, intendendo per motivazione uno stato attivato all'interno di un soggetto finalizzato al raggiungimento di un obiettivo e al soddisfacimento di un bisogno non biologico (affiliazione, aggressione, realizzazione e ostilità)<sup>2</sup>;
  - *idee personali*, intendendo con questo i modi attraverso i quali l'individuo giudica e comprende l'ambiente sociale e personale; un soggetto per esempio può (a) vedere gli altri in termini ostili o amichevoli, (b) ritenere che la realizzazione sia dovuta alla fortuna, (c) ritenere importante programmare la propria vita; ecc.;
  - *adattamento*, dimensione rappresentata dalla relativa libertà da bisogni emotivi e/o da comportamenti socialmente distruttivi; questo tratto è legato alle tre precedenti dimensioni (per esempio un tratto sociale ostile rende il soggetto poco adattabile);
  - *dinamiche di personalità*, che consistono dei principi attraverso i quali gli altri tratti sono "messi insieme" (per esempio la capacità di identificarsi in diversi ruoli); tali principi servono per spiegare l'articolazione dei tratti in una unica persona;
- **i sentimenti**, termine generico che si riferisce a:
- *interessi*, riguardanti le preferenze per particolari attività;
  - *valori*, riguardanti le preferenze verso gli ideali di vita e il modo di vedere la (propria) vita; in realtà il termine *valori* si riferisce ad una vasta gamma di contenuti, dai più elementari aspetti intellettuali della vita ai più astratti obiettivi di auto realizzazione;
  - *atteggiamenti*, intesi come sentimenti verso determinati oggetti; tradizionalmente gli atteggiamenti sono visti come composti da tre componenti (descritte nel contesto della misurazione nella successiva figura<sup>3</sup>):
    - *cognitiva*, componente importante seppure non facile da definire, riguardante il modo in cui l'individuo giudica l'ambiente sociale e materiale; si riferisce quindi anche alle **opinioni** che un individuo ha rispetto a particolari oggetti (oggetti fisici, gruppi di persone, politica, istituzioni sociali, ecc.);
    - *affettiva*, che riflette i sentimenti, le valutazioni, le emozioni, le percezioni e le auto-descrizioni di un individuo rispetto a particolari oggetti (per esempio il ruolo professionale); questa componente può comprendere anche le dimensioni di **soddisfazione** e **benessere** per le dimensioni di vita (lavoro, studio, famiglia, relazioni, ecc.) e gli **stati emozionali** (per esempio, felicità);
    - *conativa*, che riflette le tendenze comportamentali di un individuo rispetto ad un certo oggetto; possono essere comprese in questa componente anche le *intenzioni*, intese come le azioni o i comportamenti che l'individuo pianifica e eseguirà nel futuro.

<sup>2</sup> A tale proposito, ricordiamo che Abraham Maslow nel 1954 (vedi traduzione italiana del 1982 Maslow A.H., *Motivazione e personalità*, Armando Armando, Roma) ha definito un modello gerarchico, largamente accettato, dei bisogni e delle motivazioni individuali; se un soggetto non soddisfa i livelli inferiori non è in grado di soddisfare i livelli superiori; la gerarchia definita è la seguente:

1. bisogni fisiologici (cibo, aria, acqua, rifugio),
2. bisogno di sicurezza (protezione, ordine, stabilità),
3. bisogni di appartenenza e di amore (affetto, amicizia),
4. bisogno di autostima (prestigio, auto-rispetto),
5. bisogno di autorealizzazione.

<sup>3</sup> Figura tratta da Spooncer F. (1992 reprint) *Behavioural Studies for Marketing and Business*, Leckhampton, UK: Stanley Thornes (Publishers) Ltd.



Naturalmente tale schema non rappresenta l'unico modo per classificare le caratteristiche che possono essere definite "soggettive". Tale schema non è esaustivo e i diversi elementi individuati per ciascuna area possono sovrapporsi tra loro.<sup>4</sup>

Nel procedere alla misurazione delle caratteristiche soggettive occorre tenere presente che esistono informazioni individuali che non sono propriamente soggettive ma che possono spiegare o essere spiegate dalle precedenti aree (informazioni di contesto); si pensi a tale proposito a:

- le caratteristiche demografiche e socio-economiche (genere, età, stato civile, dimensione e composizione della famiglia, livello di istruzione, occupazione, reddito, luogo di nascita e di provenienza, residenza, domicilio, mobilità geografica e sociale, ecc.);
- lo stile di vita che può essere descritto da aspetti quali
  - l'attività (lavoro, hobby, vacanze, volontariato, sport, acquisti, ecc.)
  - impegno in particolari settori (famiglia, lavoro, volontariato, ecc.)
  - abitudini (orari, uso dei mezzi pubblici, uso degli strumenti di comunicazione, alimentazione, ecc.);
- le conoscenze, intese come le informazioni che l'individuo conserva riguardo a determinati aspetti;
- i comportamenti presenti e/o passati (importante quando si assume essere correlati con comportamenti futuri).

<sup>4</sup> Nell'ambito dell'approccio detto Multiattribute Evaluation esiste un particolare modello (Ajzen, 1980; Fishbein, 1975) che esprime gli atteggiamenti in termini simbolici:

$$A_o = \sum_{i=1}^n b_i e_i$$

dove

- $A_o$  atteggiamento  $A$  verso l'oggetto  $O$
- $b_i$  forza della fiducia che il soggetto ha rispetto all'attributo  $i$
- $e_i$  valutazione dell'attributo  $i$
- $n$  numero di attributi

Analogamente, è stato espresso un modello anche per le intenzioni:

$$B \approx BI = W_1(A_b) + W_2(SN)$$

dove

- $B$  comportamento
- $BI$  intenzione di comportamento
- $A_b$  atteggiamento verso la realizzazione del comportamento  $B$
- $SN$  norma soggettiva/influenza sociale
- $W_1$  e  $W_2$  pesi determinati empiricamente che rappresentano l'influenza relativa delle componenti

Secondo questo modello gli atteggiamenti e le influenze sociali non influenzano direttamente il comportamento individuale ma piuttosto operano a livello di intenzione di comportamento.



Ciò che distingue generiche informazioni individuali da informazioni soggettive è che queste ultime possono essere osservate solamente attraverso/a partire dal soggetto stesso, in altre parole non ammettono figure *proxy*.

### **Approcci per l'osservazione delle caratteristiche soggettive**

Per poter rilevare le caratteristiche soggettive è possibile individuare diverse soluzioni operative; tali soluzioni non sono necessariamente alternative ma possono trovare diverse integrazioni pratiche ed operative. Solo alcune però consentono una analisi statistica.

- Prove di performance: tale approccio è adatto alla misurazione delle capacità; la misura è rappresentata dal risultato ottenuto dalla esecuzione di un compito assegnato; tale risultato viene valutato rispetto ad un criterio definito in termini di *corretta esecuzione*; ciò consente di valutare confronti intra-individuali e inter-individuali;
- Misure dette *inventory*: tale approccio è adatto per misurare tratti di personalità, valori, interessi; la misura è rappresentata dalla risposta individuale alla somministrazione di un certo numero di stimoli; è possibile distinguere due tipi di *inventory*:
  - *self inventory*: in questo caso si chiede al soggetto di descriversi servendosi degli stimoli somministrati;
  - *inventory*: in questo caso non si richiede al soggetto di descriversi ma, per esempio, di indicare ciò che un ipotetico individuo farebbe in una determinata situazione.
- Misure dette *self-report*: tale approccio è adatto per misurare atteggiamenti, opinioni e capacità; la misura è rappresentata dalla risposta data dal soggetto (in termini di accordo, gradimento, ecc.) ad una affermazione definita relative alla caratteristica da misurare.
- Metodi osservativi: in questo caso la misura è rappresentata dal risultato dell'osservazione della caratteristica da misurare fatta da un osservatore (esterno e imparziale) sul soggetto.
- Tecniche proiettive: tale approccio è molto adatto per misurare tratti sociali, motivazioni, adattamento, atteggiamenti; la misura è rappresentata dalla reazione del soggetto a una situazione costruita e incompleta, per esempio un racconto (definito anche attraverso vignette) che si interrompe, e che il soggetto deve proseguire e/o concludere; in questo modo si suscita quella tendenza dei soggetti ad attribuire le personali e, più o meno, indesiderate caratteristiche ad altri soggetti (*proiezione*); tali tecniche si sono sviluppate principalmente in psichiatria e psicologia clinica come *strumenti diagnostici*;<sup>5</sup> esse richiedono solidi approcci interpretativi e per questo sono dette "soggettive", infatti l'interpretazione delle risposte è molto legata all'esperienza del "misuratore". Non essendo standardizzabili, le tecniche proiettive non possono essere considerate misure utilizzabili scientificamente ma possono comunque andare utilmente ad integrare altri approcci.
- Misure fisiologiche: fanno riferimento al legame esistente tra tratti soggettivi e processi fisiologici. Non è però possibile dimostrare con evidenza scientifica tale relazione.

### **Problemi nell'osservazione delle caratteristiche soggettive**

La misura delle dimensioni soggettive presenta delle difficoltà caratteristiche, determinate da alcuni fattori; in particolare, la misurazione di una certa caratteristica può risultare alterata a causa di:

a. *fattori individuali* quali:

- la *desiderabilità sociale*, ossia la tendenza dei soggetti a rispondere alle domande secondo criteri che il soggetto ritiene accettati dalla comunità (socialmente desiderabili); probabilmente tale fattore è il risultato di una combinazione varie componenti (adattamento, conoscenza, ecc.) e può essere in qualche misura controllato garantendo, per esempio,

---

<sup>5</sup> Tecniche ben note in questo ambito sono il *Rorschach test*, che consente di esaminare caratteristiche di personalità attraverso macchie di inchiostro standardizzate, e il *thematic apperception test*, noto anche come la tecnica di interpretazione di figure in quanto utilizza una serie standardizzata di figure ambigue rispetto alle quali il soggetto deve indicare una storia.

- l'anonimato;
  - il *response set*, ossia la tendenza del soggetto a rispondere in modo stereotipato e/o sistematico indipendentemente dalla caratteristica misurata (*acquiescence response set*).
  - l'*atteggiamento del ricercatore* che può non risultare adeguato ed "oggettivo" forse a causa di una carente formazione;
- b. *fattori semantici* che può causare discordanze interpretative tra soggetto e ricercatore;
- c. *fattori situazionali*, quando l'osservazione avviene in situazioni diverse (presenza o meno di altre persone) o in contesti differenti (in casa, a lavoro, per strada, ecc.).

## ***LA TEORIA DELLA MISURAZIONE DEL DATO SOGGETTIVO***

### ***Errore nella misurazione scientifica***

Come si detto, la possibilità di soddisfare i requisiti della misurazione scientifica è strettamente connessa alla possibilità di definire e identificare l'*errore di misurazione*; questo rappresenta una componente ipotizzabile di qualsiasi procedimento di misurazione. L'errore di misurazione è quella quantità per la quale differisce dal suo corrispondente valore atteso (Carmines, 1992):

$$(\text{errore}) = (\text{valore misurato}) - (\text{valore vero})$$

Il modello statistico utilizzato per valutare la presenza dell'errore<sup>6</sup> nel procedimento di misurazione, utilizza – come vedremo – il concetto di variabilità e considera due componenti additive:

- Errore casuale, che può variare da osservazione a osservazione ed prodotto da tutti quei fattori incontrollati che confondono e disturbano la misurazione; l'errore casuale è presente in qualsiasi misura in quantità diverse e il suo effetto, che può essere solo stimato, è completamente asistematico; ciò significa che produce valori che possono sottostimare o sovrastimare quello atteso;
- Errore sistematico (errore metodologico o *bias statistico*), che si verifica sempre con lo stesso valore quando uno strumento viene utilizzato nella stessa maniera; rappresenta quell'errore che produce valori che sono sistematicamente diversi da quello atteso (o troppo alti o troppo bassi); in altre parole, un errore sistematico è qualsiasi effetto che introduce nella procedura osservativi un errore che influenza sempre il risultato della misurazione, impedendo la corretta misurazione della caratteristica di interesse. Questo errore può essere controllato attraverso procedure attentamente standardizzate. Qualsiasi disciplina scientifica definisce e presenta procedure finalizzate alla costruzione di strumenti standardizzati.

### ***Accuratezza e precisione***

In ogni applicazione scientifica, accuratezza e precisione sono strettamente correlati; in particolare

---

<sup>6</sup> Dato che in statistica il concetto di errore è facilmente confuso con quello di residuo, può essere utile differenziarli. Un **errore** rappresenta la quantità per la quale una osservazione differisce dal corrispondente valore atteso; il valore atteso è basato sull'intera popolazione da cui l'unità statistica è stata casualmente estratta. Gli errori, assunti tra loro indipendenti, non direttamente osservabili ma possono essere solo stimati. Un **residuo** rappresenta una stima osservabile dell'errore non osservabile. Si assume che i residui siano non necessariamente indipendenti. Così, la differenza tra

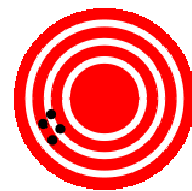
- il valore di ciascun caso in un campione e la media non osservabile della popolazione è un errore,
- il valore di ciascun caso in un campione e la media osservabile del campione è un residuo.

- l'*accuratezza* rappresenta il grado di conformità di una quantità misurata al suo reale e vero valore; valutare l'accuratezza richiede la contemporanea osservazione di un processo noto o la disponibilità di un valore di riferimento (calibrazione);
- la *precisione* rappresenta il grado in cui misurazioni successive mostrano lo stesso risultato o risultati simili; attraverso la determinazione della precisione di uno strumento è possibile verificare, quindi, la capacità di registrare con il minor grado di distorsione possibile; la precisione è correlata ai concetti di *robustezza* e *stabilità* e può essere distinta in
  - *ripetibilità*, ossia la variazione prodotta quando si sono fatti tutti gli sforzi per mantenere costanti le condizioni di misurazione attraverso ripetizioni entro poco tempo;
  - *riproducibilità*, ossia la variazione prodotta utilizzando lo stesso procedimento di misurazione con strumenti e operatori diversi anche con intervalli di tempo lunghi.

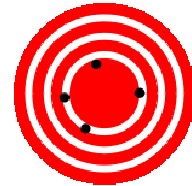
La differenza tra accuratezza e precisione può essere spiegata attraverso l'esempio del bersaglio. In questa analogia, le misure ripetute possono essere paragonate a delle frecce lanciate verso un bersaglio.

Quando tutte le frecce occupano un'area molto ristretta, le misurazioni saranno considerate precise; la dimensione dell'area occupata dalle frecce è interpretata come "grado di precisione".

Quando le frecce sono molto vicine al bersaglio, le misurazioni sono considerate accurate. La distanza del gruppo di frecce dal centro è interpretata come "livello di accuratezza"; in altre parole, più la procedura di misurazione produce valori vicini al valore atteso, più accurata è la procedura.

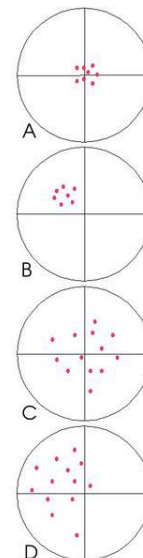


alta precisione ma bassa accuratezza



alta accuratezza ma bassa precisione

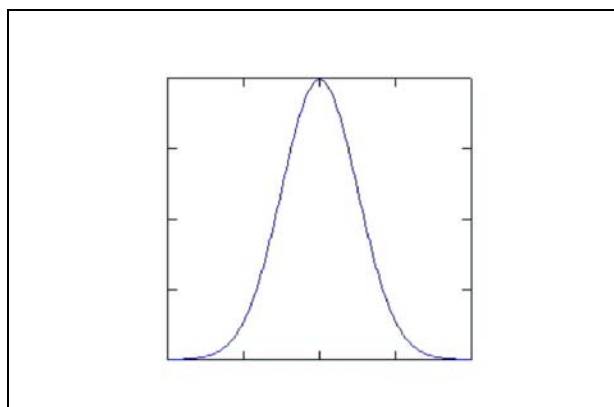
da Wikipedia



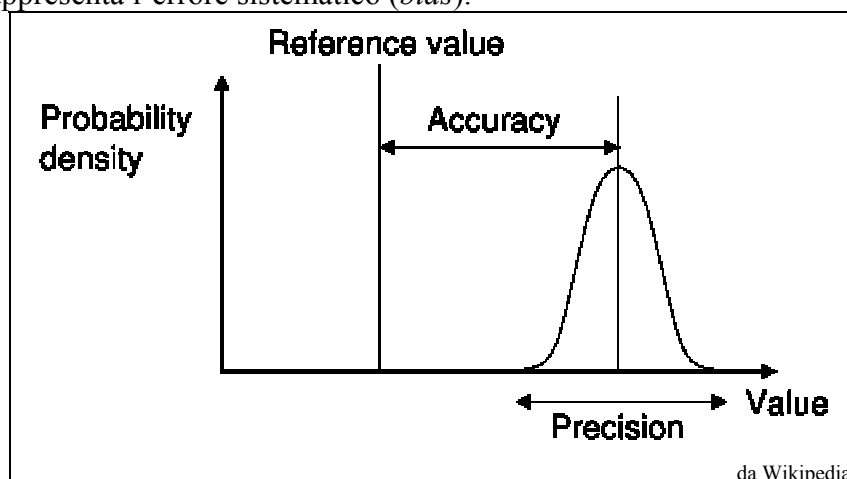
Gli esempi qui accanto mostrano gruppi di misure ripetute che sono abbastanza accurate e precise (A), precise ma non accurate (B), imprecise e non accurate (C and D).

Da ciò che abbiamo visto, possiamo assumere che maggiore è il numero delle ripetizioni, più accurata e precisa è la stima dell'errore. Assumendo misure ripetute, l'errore è detto **incertezza**. Nella distribuzione di frequenza in caso di misure ripetute, si assume che il valore corrispondente alla frequenza più alta è quello con la probabilità più alta di essere vicino al valore vero; è possibile inoltre assumere che gli errori positivi compensano gli errori negativi (anche se ciò non è sempre

vero). Conseguentemente, si assume che la distribuzione dei valori delle misure ripetute sia *normale*.



Maggiore è l'estensione di tale distribuzione, ossia maggiore è l'oscillazione dei punteggi osservati intorno al punteggio medio, maggiore è l'ampiezza dell'errore. Conseguentemente, la deviazione standard delle misure ripetute fornisce una stima dell'incertezza. In particolare, l'incertezza è uguale all'errore standard di tale distribuzione (deviazione standard diviso per la radice quadrata del numero di misurazioni). La differenza tra la media delle misure ripetute e il *reference value* (**calibratura**) rappresenta l'errore sistematico (*bias*).



In statistica, l'effetto dell'incertezza di ciascuna delle misure ripetute sull'incertezza dell'intera misurazione è detta **propagazione dell'incertezza**.

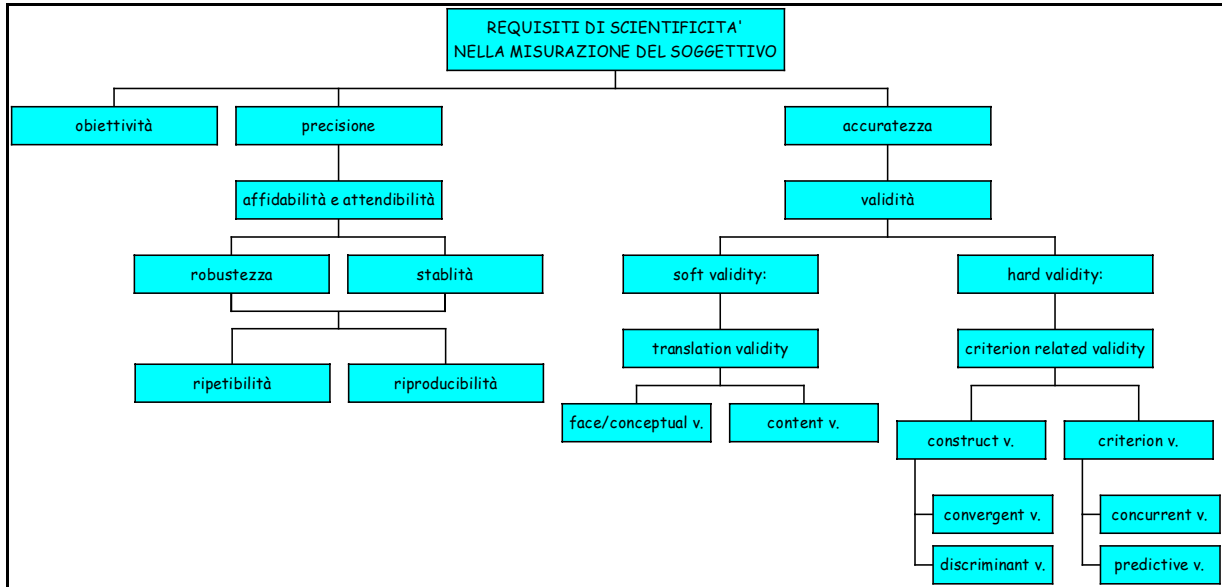
La quantificazione dell'accuratezza e della precisione è importante non solo in termini di correttezza della misurazione; infatti, la valutazione dell'errore consente di fare anche valutazioni rispetto ai costi che la sua presenza comporta all'economia della ricerca (Biemer, 1991; Groves, 1989).

### ***Accuratezza e precisione nella misura del soggettivo***

Come si è detto, la possibilità di soddisfare i requisiti della misurazione scientifica è strettamente connessa alla possibilità di definire e misurare l'errore di misurazione, a sua volta intimamente connesso ai concetti di accuratezza e precisione. Nella misurazione del soggettivo questi due concetti trovano una definizione operativa; in particolare:

- la precisione è misurata dal livello di *affidabilità* che rappresenta la consistenza dello strumento di misurazione (Carmines, 1992; Ghiselli, 1964; Marradi, 1990; Netemeyer, 2003; Thompson, 2003; Traub, 1994);

- l'accuratezza è misurata dal livello di *validità*, ossia la tendenza di un procedimento di misurazione a misurare ciò che si intende misura; in altre parole, la validità si riferisce alla possibilità di ottenere risultati che accuratamente riflettono il concetto che deve essere misurato. L'affidabilità non implica la validità, mentre la validità implica l'affidabilità. Ciò vuol dire che una misura valida deve essere affidabile ma una misura affidabile non necessariamente è valida ossia una misura affidabile misura in modo consistente qualcosa che non necessariamente è ciò che si vuole misurare.



### Affidabilità

Nella misurazione del soggetto, per poter definire la procedura che consenta di stimare e valutare l'affidabilità, è necessario identificare una *teoria della misurazione* (Laveault, 1994; Nunnally, 1978); in particolare, si individuano tre teorie:

- La teoria classica (*Classical Test theory, C-T*): tale teoria assume l'esistenza di un punteggio vero e di un errore e la possibilità di definire *forme parallele* di misurazione per la stima dell'affidabilità; conseguentemente ammette la possibilità che uno strumento abbia diverse stime di affidabilità e ignora l'esistenza di fonti accidentali di errore (DeVellis, 1991; Nunnally, 1978; Spector, 1992). L'affidabilità è matematicamente definita come il rapporto tra la varianza del punteggio vero e la varianza del punteggio osservato o, in modo equivalente, uno meno il rapporto tra la varianza dell'errore e la varianza del punteggio osservato:

$$rho_x = \frac{\sigma_t^2}{\sigma_x^2} = 1 - \frac{\sigma_e^2}{\sigma_x^2}$$

dove

$rho_x$  affidabilità del punteggio osservato,  $X$

$\sigma_x^2, \sigma_t^2, \sigma_e^2$  varianza del, rispettivamente, punteggio osservato, punteggio vero e errore.

Purtroppo, non vi è alcun modo per osservare o calcolare direttamente il punteggio vero, conseguentemente sono stati definiti diversi modelli per la stima dell'affidabilità.

- La teoria della generalizzabilità (*Random Sampling theory, R-S*): secondo questa teoria, la misurazione di un'unica caratteristica richiede un gruppo numeroso di misure; la teoria assume che tale gruppo sia casualmente estratto da un ipotetico universo di misure, che riguardano e definiscono in modo completo la caratteristiche – e possa essere considerato un campione statistico. Le misure estratte consentono di stimare il risultato che si sarebbe ottenuto se fosse stato impiegato l'intero universo di misure; conseguentemente, dato che non è però possibile

fare alcun esplicito assunto rispetto al processo stocastico che genera una risposta (Thompson, 2003), il livello di errore/affidabilità dipende dalla dimensione del campione di misure (DeVellis, 1991). La stima dell'affidabilità deve affrontare principalmente due problemi, la stima del punteggio dell'universo e la generalizzazione di tale stima (fatta attraverso una particolare applicazione dell'analisi della varianza).

- La teoria del tratto latente (*Latent Trait theory, L-T*): secondo questa teoria, la misurazione necessita della definizione di due componenti: la misurazione reale (*indicatore* o *variabile manifesta*) e la corrispondente caratteristica (non misurabile direttamente e per questo detta *variabile latente*). Si assume che 1) le risposte relative agli indicatori sono il risultato della posizione individuale per la variabile latente, 2) gli indicatori non registrano nessuna relazione nel caso in cui si controlla la variabile latente (indipendenza locale) (DeVellis, 1991).

La teoria classica può essere considerata un caso particolare della teoria della generalizzabilità in cui le misure sono fissate anziché essere estratte da un universo; i due approcci però hanno una diversa definizione di affidabilità. Infatti, per la prima l'affidabilità dipende interamente dalla correlazione tra le misure che definiscono la caratteristica da misurare (maggiore è la correlazione media, maggiore è l'affidabilità), per la seconda l'affidabilità dipende dalla dimensione del campione di misure (maggiore è il numero di misure definite, maggiore è l'affidabilità).

Come vedremo (parte I, cap. 1<sup>7</sup>), le tre diverse teorie non sono tra loro completamente incompatibili e trovano una integrazione dal punto di vista applicativo (Bejar, 1983).

## Validità

La validità rappresenta la capacità di una procedura di misurazione di misurare ciò che si intende misurare. È possibile distinguere diversi tipi di validità cui corrispondono metodi diversi di verifica (Campbell, 2001; Carmines, 1992; Cronbach, 1955; DeVellis, 1991; Ghiselli, 1964; Netemeyer, 2003; Nunnally, 1978; Spector, 1992), come mostrato nel seguente schema:

<b>VALIDITÀ</b>	Soft ( <i>translation validity</i> )	La misura è una buona rappresentazione del concetto che si vuole misurare	a. <i>face validity/conceptual validity</i> . apparente, logica, concettuale
			b. <i>content validity</i> . di contenuto, intrinseca, rilevante, rappresentativa
			c. <i>construct validity</i> (della struttura, teorica, fattoriale) <ul style="list-style-type: none"> <li>○ <i>convergent validity</i> (convergente)</li> <li>○ <i>discriminant validity</i> (discriminate)</li> </ul>
	Hard ( <i>statistical validity</i> )	La misura predice altre misure	d. <i>criterion validity</i> , (correlata al criterio): <ul style="list-style-type: none"> <li>○ <i>concurrent validity</i></li> <li>○ <i>predictive validity</i></li> </ul>

- Face Validity: è determinata dalla significatività apparente ed esteriore che una misura presenta, da cui il nome; per valutarla si richiedono giudizi di esperti.
- Content Validity: una misura ha validità di contenuto quando rappresenta in modo accurato l'intero universo di contenuto (*domain*) corrispondente alla caratteristica da misurare; anche questo tipo di validità richiede il giudizio di esperti.
- Construct Validity: è data dal livello in cui la misura da validare è in grado di misurare accuratamente il costrutto teorico considerato; essa è verificata praticamente attraverso metodi statistici che rilevano se un ipotizzato fattore comune esiste tra molte misurazioni. In particolare, la valutazione richiede l'osservazione della correlazione della misura da validare con altre legate ad altri costrutti che si suppone sia correlati con quello di interesse. Le

<sup>7</sup> In questo capitolo sono riportate le indicazioni di parti e capitoli in cui possono essere trovati (nel resto di questo lavoro) gli approfondimenti relativi agli argomenti in questione.

correlazioni che rispecchiano il modello atteso costituiscono l'evidenza della validità di costruito. Poniamo, per esempio, di voler costruire uno strumento che misuri le tendenze razziste di un particolare gruppo di individui; secondo l'ipotesi, l'indice sarà una buona misura di razzismo, ovvero avrà validità di costruito, se risulterà essere inversamente correlato con l'autostima; per poter procedere in questo modo è necessario assumere che la teoria sia corretta, ovvero che esista una relazione tra autostima e razzismo. Casi particolari di validità di costruito sono:

- **Convergent Validity**: è determinata confrontando e correlando i punteggi ottenuti con la misura da validare con quelli ottenuti con la misura di un altro costruito, teoricamente legato al primo. La verifica della validità convergente dipende quindi dalla possibilità di ipotizzare questo tipo di relazioni.
- **Discriminant Validity**: è speculare alla validità convergente; essa è alta quando la misura da validare non correla con le misure di altri costrutti, teoricamente distinti dal primo. Supponiamo, per esempio, di voler predisporre un questionario strutturato per misurare il costruito *patriottismo* in relazione al quale si può ipotizzare che negli individui esista un atteggiamento "reale" e un atteggiamento che risponda a una certa *desiderabilità sociale*; in questo caso è possibile predisporre un gruppo di indicatori che misurano il costruito *risposte socialmente desiderabili*; se le due misure (*patriottismo* e *desiderabilità sociale*) non risultassero correlate, la misura di patriottismo avrebbe più validità discriminante ovvero sarebbe non correlata ad una misura di qualcosa con la quale non dovrebbe essere correlata se fosse valida.

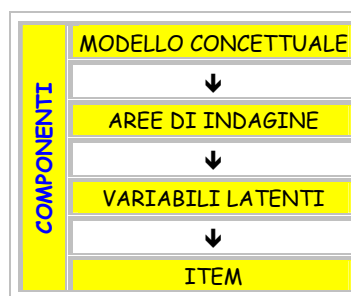
Il concetto di validità discriminante è stato introdotto da Campbell e Fiske (1959) nell'ambito della loro discussione sulla valutazione della validità. Gli autori hanno messo l'accento sull'importanza nella validazione di nuove misure dell'utilizzo tecniche di validazione sia discriminanti che convergenti.

- d. **Criterion Validity**, che rappresenta quanto la misura sia correlata con un definito criterio. E' determinata osservando quanto l'indicatore correla con un'altra misura che il ricercatore ritiene valida nel misurare lo stesso costruito; relazioni alte e significative rappresentano l'evidenza statistica dell'esistenza della validità. Quando il criterio esterno è rilevato:
- contestualmente alla misura da validare, l'obiettivo è quello di stabilire la *concurrent validity*;
  - successivamente alla misura da validare, l'obiettivo è quello di stabilire la *predictive validity* che rappresenta la capacità di uno strumento di fare previsioni accurate del criterio definito.

Vedremo (parte I, cap. 1) attraverso quali strategie di analisi sia possibile verificare la validità.

## IL MODELLO GERARCHICO

Per poter rilevare il dato soggettivo è necessario definire un **modello** di misurazione che, per le sue caratteristiche, è detto **gerarchico** in quanto ciascuna componente è definita e trova significato nell'ambito della precedente. Le componenti che definiscono il **modello gerarchico** possono essere individuate nel modo seguente:



### **Il modello concettuale**

La definizione del modello concettuale rappresenta un vero e proprio processo di astrazione. Rappresenta quindi un momento complesso in quanto richiede la identificazione e la definizione di costrutti che in genere derivano da astrazioni e strutture teoriche che devono trovare riferimenti concreti e di applicabilità. Nelle scienze sociali la descrizione dei concetti varia secondo il punto di vista del ricercatore, gli obiettivi dello studio, l'applicabilità dei concetti, il contesto socio-culturale, geografico, temporale. A tale proposito si pensi a concetti quali salute, educazione, benessere, produzione, commercio, reddito, ecc. Il processo di concettualizzazione consente anche di individuare e definire

- a. il modello per la costruzione del dato,
- b. l'ambito spaziale e l'ambito temporale della rilevazione,
- c. i livelli di aggregazione (degli indicatori e/o delle unità di osservazione),
- d. il modello di condensazione degli indicatori elementari e delle tecniche da utilizzare (criteri di ponderazione, tecniche di aggregazione, ecc.)
- e. i modelli interpretativi e di valutazione.

### **Le aree di indagine**

Le aree di indagine definiscono in termini generali i diversi aspetti che consentono di chiarire e specificare il fenomeno sotto osservazione in modo coerente con gli obiettivi dello studio. Il processo di definizione può essere lungo e impegnativo, soprattutto in presenza di costrutti astratti e complessi, e richiede una analisi e una rassegna della letteratura esistente.

### **Le variabili latenti**

Ciascuna variabile latente rappresenta un aspetto da rilevare e conferisce una rilevanza esplicativa alla corrispondente area definita in precedenza. L'identificazione della variabile latente deve essere basata su presupposti teorici ed empirici (riguardanti, per esempio, la sua omogeneità e/o dimensionalità). Ciascuna delle variabili deve essere sviluppata in modo che rifletta la natura dei fenomeni considerati e il modello concettuale definito.

### **Gli item**

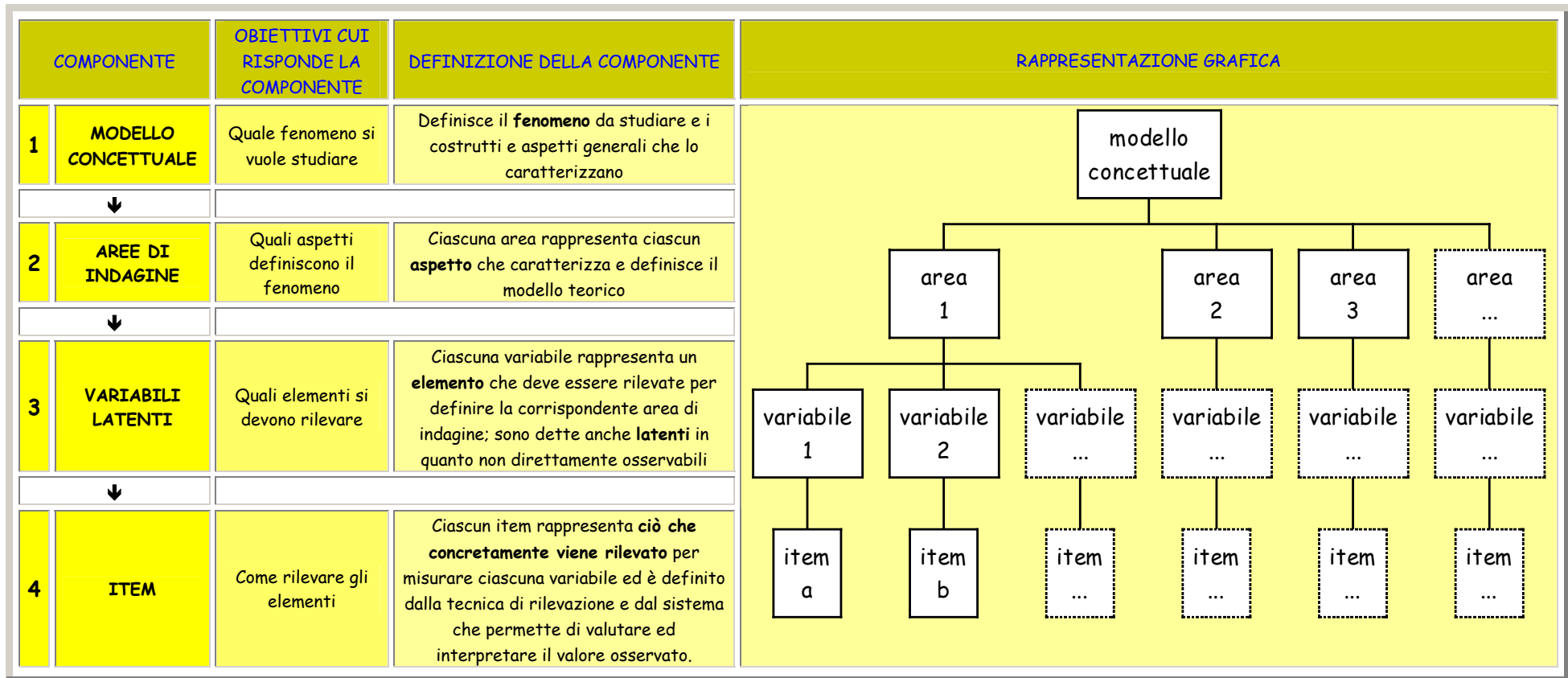
Ciascun item rappresenta ciò che concretamente viene rilevato per misurare la corrispondente variabile latente. Questo vuol dire che ciascun item<sup>8</sup> rappresenta non una misura diretta della variabile ma un **indicatore** della variabile cui si riferisce (DeVellis, 1991). Il processo gerarchico che conduce alla sua individuazione consente di attribuire all'item una precisa collocazione all'interno del modello. In altre parole, ciascun item assume e acquista un significato proprio e può essere conseguentemente interpretato perché inserito all'interno della struttura gerarchica in cui ciascuno di essi misura una distinta componente del fenomeno di interesse. E' importante valutare la possibilità di considerare più forme alternative per ciascun item.

A questo punto, il modello gerarchico e le sue componenti possono essere rappresentati e riassunti nel modo seguente:

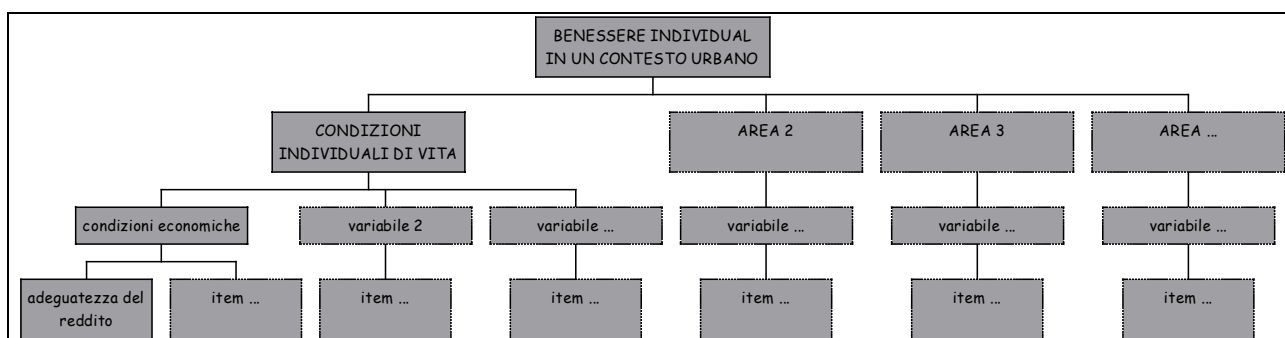
---

<sup>8</sup> Nell'ambito dell'analisi dei dati, ciascun item/indicatore è definito "variabile" di analisi che è quindi concettualmente diversa dalla "variabile latente".





Nella seguente rappresentazione è presentato un esempio:



### Approccio “single item” e approccio “multiple item”

Secondo una semplice ma debole strategia, per ciascuna variabile latente si definisce un unico item (quindi un unico *indicatore* che chiameremo per questo *elementare*: *single indicator approach*).

Tale strategia, pur se utilizzata spesso per la sua praticità e parsimonia, richiede l’adozione di assunti robusti. In ogni caso l’adozione di indicatori singoli è rischiosa in quanto non è sempre definibile la corrispondenza diretta tra variabile non osservabile e indicatore ovvero non sempre l’attributo è direttamente osservabile attraverso un unico indicatore.

Infatti, la definizione e l’adozione di un singolo indicatore per la misurazione di una caratteristica può produrre una grossa componente di errore a causa dell’introduzione di problemi di:

- affidabilità, precisione e accuratezza*, in quanto la misurazione effettuata con un singolo indicatore è molto influenzata dall’errore casuale<sup>9</sup>;
- validità*, in quanto è molto improbabile che un singolo indicatore possa rappresentare e misurare una variabile, specie se complessa e articolata;
- legame e relazione* con altre variabili definite, diverse da quella misurata;
- discriminazione e differenziazione* accurata tra i casi osservati.

Per questa ragione, in molti casi la presenza di variabili latenti complesse richiede la definizione di più indicatori elementari, attraverso l’adozione dell’approccio *multiple indicators* che considera gli item multipli come *misure multiple* (Sullivan, 1981) che contribuiscono a rilevare il maggior numero di aspetti della variabile e consentono di coprirne la variabilità. Questo approccio consente di evitare i problemi prodotti dall’approccio *single item*, o almeno a ridurne il peso.<sup>10</sup>

In termini tecnici, l’intero gruppo di indicatori elementari relativi ad una variabile rappresenta un *set*, mentre l’intero gruppo di indicatori elementari che definiscono un’area sono detti *indicatori tematici*.

### La definizione delle relazioni

La definizione del modello gerarchico può essere completata con la identificazione delle relazioni tra<sup>11</sup>:

- **le variabili latenti**: tali relazioni definiscono il **modello strutturale** e sono ipotizzate e identificate nell’ambito del modello concettuale;

<sup>9</sup> Con l’utilizzo di misure multiple, gli errori causali presenti nelle diverse misure tendono a compensarsi, rendendo la misurazione più accurata. A tale proposito occorre tenere presente che maggiore è la componente d’errore in ogni singola misurazione, maggiore è il numero di misurazioni richiesto per produrre una misura affidabile.

<sup>10</sup> La necessità di disporre di più misurazioni è un aspetto che è alla base, come abbiamo visto, della teoria della generalizzabilità.

<sup>11</sup> Come si potrà notare il modello proposto è analogo a quello che concettualmente è alla base dei modelli a strutture latenti che trovano una soluzione analitica attraverso l’approccio delle equazioni strutturali (Asher, 1983; Bartholomew, 1999; Blalock, 1961, 1974; Bohmstedt, 1994; Lazarsfeld, 1968; Long, 1993a, 1993b; Maggino, 2005; Netemeyer, 2003; Saris, 1990; Sullivan, 1981; Werts, 1974).

- **le variabili latenti e i corrispondenti item:** tali relazioni definiscono il **modello di misurazione** la cui verifica consente di valutare l'affidabilità degli indicatori;<sup>12</sup>
- **gli item:** in questo caso le relazioni definite possono identificare diverse situazioni:
  - gli item correlano con la stessa variabile ovvero contribuiscono alla definizione di tale variabile; in questo caso gli item sono detti **costitutivi** e possono essere sintetizzati;
  - gli item correlano con variabili diverse; in questo caso gli item sono detti **concomitanti**.

Successivamente, la definizione del modello gerarchico consente di individuare la popolazione di riferimento, conseguentemente il disegno di campionamento necessario per procedere alla rilevazione, e la metodologia attraverso la quale si procederà alla rilevazione.

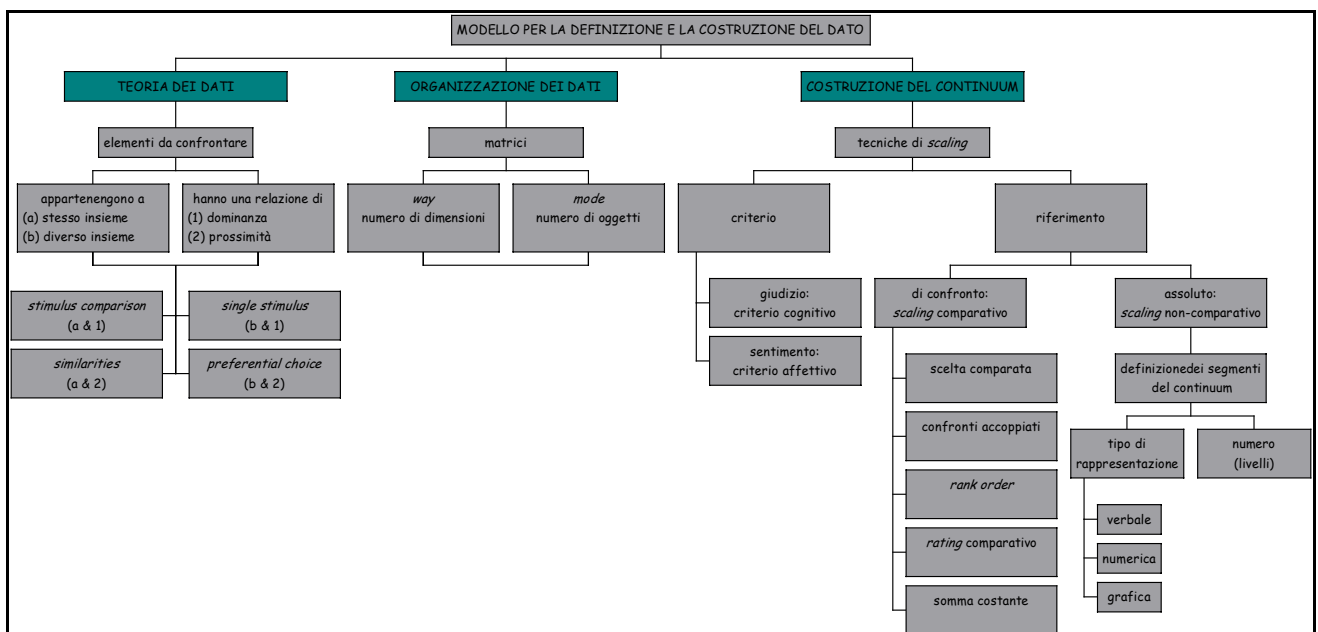
## IL MODELLO DI RILEVAZIONE DEL DATO SOGGETTIVO

### Modello per la definizione e la costruzione del dato

Tale modello consente di trasformare ciascuna *osservazione* empirica, ossia l'informazione che è stata rilevata, in *dato*, ossia una informazione selezionata, analizzabile e interpretabile. I dati consistono quindi di porzioni di informazione estratte dalle osservazioni secondo un modello di riferimento; in questo senso i dati rappresentano una costruzione e una interpretazione del ricercatore.

La definizione del dato soggettivo rappresenta una delle fasi più delicate del procedimento di misurazione e necessita, al fine di evitare l'arbitrarietà e l'eccessiva approssimazione, di un modello (parte I, cap. 2) che definisca:

- la natura dei dati, in riferimento ad una teoria interpretativa (**teoria dei dati**) (Coombs, 1950, 1953, 1964; Flament, 1976; Jacoby, 1991; McIver, 1979),
- l'organizzazione dei dati (**matrici**) (Delli Zotti, 1995; Jacoby, 1991),
- il procedimento per la costruzione del continuum lungo il quale posizionare gli oggetti misurati (**tecniche di scaling**).



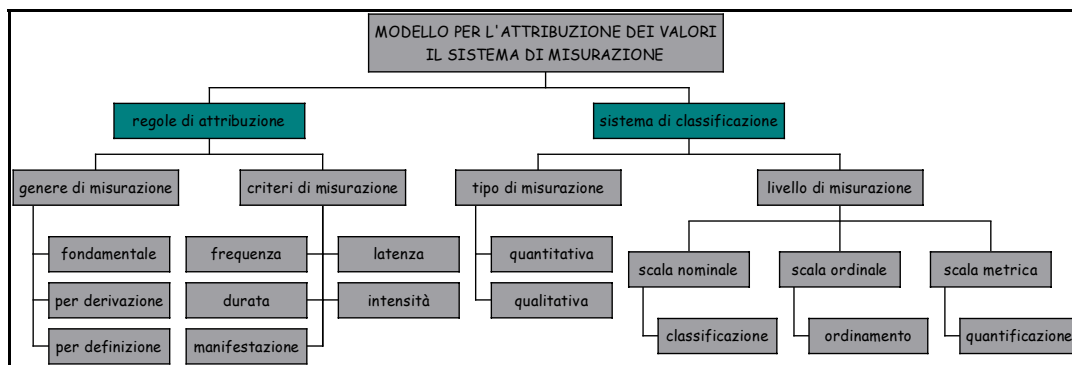
<sup>12</sup> A tale proposito occorre tenere presente che le relazioni tra le variabili latenti possono essere osservate solo quando e se si osservano relazioni significative tra gli indicatori le corrispondenti variabili.

## Modello per l'attribuzione dei valori

Successivamente alla definizione del modello di creazione del dato è necessario procedere alla definizione di un modello per l'assegnazione – a ciascun segmento del continuum identificato – un valore per rendere il dato rilevato analizzabile. In altre parole, tale modello consente di assegnare un valore che renda il dato costruito interpretabile e trattabile in termini operativi.

A tal fine occorre definire regole che chiariscano la procedura di corrispondenza e di assegnazione del simbolo a ciascun livello identificato dal precedente modello. Ciò richiede la definizione di un sistema di misurazione (parte I, cap. 3) che abbia

- regole per l'assegnazione di numeri/simboli che consentano di assegnare i valori in modo standardizzato e secondo una procedura uniforme (**genere e criterio di misurazione**),
- un "sistema di classificazione" che consenta di assegnare a ciascun caso il suo stato relativamente alla proprietà misurata (**tipo e livello di misurazione**).



L'insieme dei valori individuati secondo questo modello definisce una *scala*.<sup>13</sup>

Nella misura del soggettivo, la definizione del sistema di classificazione rappresenta anche uno dei punti più discussi: infatti, eccetto i casi in cui il sistema è così ovvio da non richiedere formulazioni dettagliate (casi che rappresentano delle eccezioni nel generale panorama scientifico), in genere la definizione del sistema di classificazione non è né semplice, né univoca, né ovvia, né intuitiva e pone, soprattutto nelle scienze sociali, il problema dell'arbitrarietà.

<sup>13</sup> Esiste un'altra strategia per l'attribuzione di un valore a ciascun segmento che non richiede la definizione di un sistema a priori. Tale strategia è detta *optimal scaling* e procede in conseguenza di una particolare analisi; alle osservazioni si assegnano punteggi in un modo che soddisfa simultaneamente due condizioni:

- i punteggi assegnati adattano il modello statistico nel modo migliore possibile;
- essi mantengono strettamente le caratteristiche specificate della misurazione.

La strategia dell'*optimal scaling* presenta il vantaggio di fornire ai dati la *migliore* serie di assegnazioni numeriche, verificata attraverso il livello di adattamento tra un modello analitico e una serie di osservazioni empiriche.

Mentre le caratteristiche della misurazione dei dati devono essere specificate prima dell'analisi, il ricercatore può variare i propri assunti sulle caratteristiche e osservare gli effetti dei risultati analitici: se due o più di tali analisi producono precisamente gli stessi risultati allora gli assunti della misurazione più appropriati sono quelli utilizzati per le analisi equivalenti. In questo senso, la strategia di *optimal scaling* incorpora esplicitamente l'idea che tutta la misurazione è una verifica di teorie.

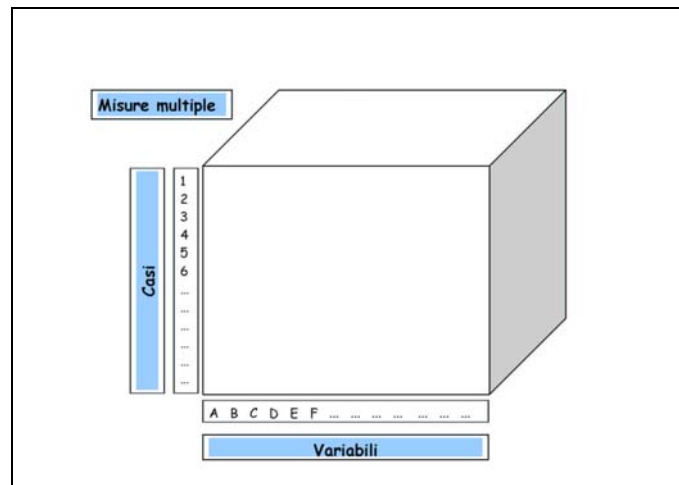
## ***IL MODELLO DI ANALISI DEL DATO SOGGETTIVO***

### ***Riduzione della complessità del dato rilevato: i modelli di condensazione***

Il modello di rilevazione fin qui presentato produce una complessa struttura dei dati in relazione a:

- le *variabili*, secondo il modello gerarchico, ne vengono definite e individuate molte,
- le *misure multiple*, definite per ciascuna variabile,
- più *casì*.

Conseguentemente, anche la struttura logica dei dati richiede una organizzazione complessa, nella quale sono osservabili le tre corrispondenti dimensioni:

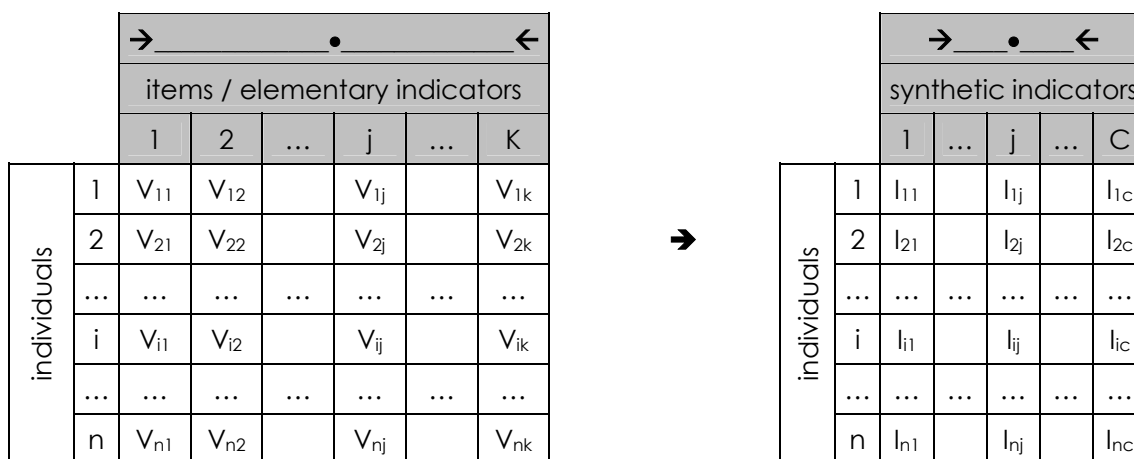


La complessità di tale struttura può essere ridotta attraverso l'applicazione di due ulteriori modelli. Tali modelli sono detti di **condensazione** in quanto hanno l'obiettivo, attraverso la definizione e l'adozione di particolari assunti, di sintetizzare all'interno della matrice tridimensionale la dimensione che si riferisce, rispettivamente alle *misure multiple* e ai casi (punti individuali). Vi possono essere tentativi – come vedremo di seguito – di ridurre anche la terza dimensione (variabili) che però presentano grossi problemi sia metodologici che interpretativi.

### ***Modello di condensazione degli item***

La condensazione degli item, considerati misure multiple, produce nuovi valori sintetici (*indicatori sintetici*) che consentono di ricostruire l'unità del concetto descritto dalla variabile latente; l'analisi fatta secondo tale prospettiva ripercorre a ritroso il modello gerarchico (*from elementary indicators to synthetic indicator*).

Dal punto di vista operativo il processo di condensazione conduce ad una semplificazione della matrice dei dati in termini numero minore di colonne, secondo quanto esemplificato nella seguente figura:



### Complessità della condensazione degli item

Il procedimento di condensazione delle misure multiple che consente di creare nuovi indicatori che riassumono gli indicatori elementari (considerate misure multiple) deve tenere conto di molti fattori (per esempio la presenza di valori missing).

In particolare, l'esplorazione degli indicatori che si intendono coinvolgere nella condensazione (*multivariate and multidimensional analysis*)<sup>14</sup> può aiutare a comprendere la struttura dimensionale sottostante gli item considerati e quindi il livello di complessità dell'aggregazione (Nunnally, 1978). Infatti, la condensazione degli indicatori elementari, considerati misure multiple, può avvenire a diversi livelli di complessità, secondo due diversi criteri:

- *omogeneità*, quando gli indicatori elementari che vengono aggregati misurano la stessa variabile; in questo caso si parla di indicatori sintetici;
- *eterogeneità*, quando gli indicatori elementari che vengono aggregati fanno riferimento a variabili diverse; in questo caso si parla di (Del Vecchio, 1995):
  - indicatore composito (*composite indicator*) quando l'aggregazione è ottenuta attraverso indicatori (elementari e/o sintetici) tra loro non necessariamente legati statisticamente; per le sue caratteristiche, un indicatore composito necessita di una particolare definizione teorica ed una corrispondente e coerente metodologia per la sua costruzione;
  - indicatore comprensivo (*comprehensive/summary indicator*) quando un indicatore composito è costruito con l'intento di essere esauriente rispetto ad un costrutto.



Tale distinzione, valida e applicabile quando si fa riferimento ad indicatori oggettivi, risulta essere valido solo teoricamente nella misurazione del soggettivo. Infatti nella pratica gli indicatori, compositi o comprensivi, che fanno riferimento a variabili complesse pongono seri problemi non

<sup>14</sup> Le tecniche di analisi che possono essere applicate a tale proposito sono *Correlation analysis, Factor Analysis, MultiDimensional Scaling, Cluster Analysis, Item Response Models Analysis*. (Alt, 1990; Anderson, 1958; Bolasco, 1999; Cooley, 1971; Corbetta, 1992, 2003; Cox, 1994; Hair, 1998; Kruskal, 1978; Maggino, 2004a, 2004b, 2005; Sadocchi, 1981).

solo in fase di creazione e di costruzione ma anche, e soprattutto, in sede interpretativa; ciò sconsiglia molto la loro costruzione. In altre parole e come vedremo, il processo di condensazione di dati soggettivi può essere condotto solo fino ad un certo livello.

### **Modelli di scaling**

I modelli che, nella misura del soggettivo, consentono di condensare le misure multiple secondo il criterio di omogeneità sono i modelli di *scaling* (Nunnally, 1978). Ciascun modello di *scaling* è costituito da un insieme di metodi e di tecniche finalizzate a:

- verificare il modello concettuale,
- ricostruire in modo significativo l'unità del concetto di interesse (variabili),
- sintetizzare le misure multiple e attribuire il valore sintetico che ne risulta a ciascun caso,
- identificare il continuum su cui posizionare ciascun caso osservato in modo significativo, interpretabile e gestibile.

In questo senso il modello di *scaling* può essere definito come un "piano internamente consistente di sviluppo di una nuova misura" (Nunnally, 1978).

I modelli di *scaling* si distinguono relativamente a diversi elementi (McIver, 1979):

➤ **Dimensionalità**, legato alla complessità della natura della variabile latente; la definizione di più aspetti per la variabile latente introduce l'idea di «dimensioni». La individuazione di una determinata dimensionalità richiede l'adozione di un particolare modello di *scaling* (McIver, 1979; Netemeyer, 2003). Il concetto di dimensionalità è piuttosto complesso in quanto il suo significato è principalmente ed essenzialmente teorico. E' comunque possibile distinguere tra due diverse "dimensionalità":

- a. **unidimensionalità**: in questo caso la definizione di una certa caratteristica assume una singola, fondamentale dimensione sottostante;
- b. **multidimensionalità**: in questo caso la descrizione di una certa caratteristica assume più aspetti o dimensioni sottostanti.<sup>15</sup>

La corrispondenza tra la dimensionalità individuata e gli indicatori selezionati deve essere dimostrata empiricamente attraverso la verifica del modello di *scaling* applicato.

Fin di primi decenni del Novecento sono stati definiti diversi modelli di *scaling* unidimensionali per ciascuno dei quali sono stati proposti diversi criteri ed applicate varie tecniche che hanno avuto origine da studi di diversi ricercatori (Thurstone, Bogardus, Likert, Guttman, ecc.); l'applicazione di modelli teorico-matematici ha consentito ulteriori sviluppi soprattutto nell'ambito, come vedremo, della misurazione multidimensionale (McIver, 1979).

➤ **Natura dei dati**; a tale proposito ricordiamo che la natura dei dati non è mai predeterminata ma dipende dall'interpretazione che il ricercatore fa delle osservazioni in termini di appropriatezza e corenza. Interpretazioni differenti conducono ad applicazioni di diverse procedure di *scaling*. Vediamo, per ciascuna tipologia di dati, quali sono gli approcci applicabili:

- **Dati del tipo *stimolo-unico***. Essendo i dati più comuni, non meraviglia che per essi sia stata concepita la maggior parte dei modelli di *scaling*, tra i quali ricordiamo il modello additivo e i modelli cumulativi (deterministico e probabilistico) (Flament, 1976; McIver, 1979; Torgerson, 1958).
- **Dati del tipo *confronto di stimoli***. Tra i modelli di riferimento per questo tipo di dati ricordiamo quello definito da Thurstone (Arcuri, 1974; McIver, 1979; Thurstone, 1927, 1959) e la metodologia *Q* (McKeown, 1988).
- **Dati di *somiglianza***. Un modello di riferimento per questo tipo di dati è il *multidimensional scaling* (Cox, 1994; Kruskal, 1978; Torgerson, 1958).
- **Dati del tipo *scelta di preferenza***. Uno dei modelli di riferimento per questo tipo di dati è l'*unfolding* (McIver, 1979).

<sup>15</sup> La nozione di dimensionalità è presente non solo nelle scienze sociali ma anche in altre scienze; si pensi ad attributi unidimensionali quali la lunghezza o il peso e ad attributi multidimensionali quali il colore o lo spazio.

- **Tecnica di *scaling*** utilizzata.
- **Criterio utilizzato per la verifica del modello**, che consente di verificare l'adattamento del modello ai dati; la logica della procedura di verifica è comune a tutti i modelli (parte I, cap. 4) ma i criteri sono diversificati, a seconda del modello prescelto (parte II).
- **Standard di misurazione** che riguarda il trattamento delle misure multiple e l'attribuzione del punteggio sintetico (a chi viene attribuito il punteggio finale, casi o oggetti), secondo il seguente schema:

<i>Standard di misurazione</i>		misure multiple	Rispetto all'attributo, l'obiettivo della misurazione è di classificare	punteggio finale attribuito a
l'insieme delle misure multiple consente di misurare in maniera più accurata	il caso	Stimolo (item)	i casi	Caso
	l'indicatore elementare	Caso	gli indicatori elementari	Stimolo (item)

Il seguente esempio consente di mettere in evidenza come la risposta individuale possa assumere, a seconda dello standard di misurazione, un ruolo, un peso e un significato diverso.

In uno studio sul pregiudizio sociale, una delle variabili che possono essere definite può riguardare "la percezione della distanza sociale da un particolare gruppo sociale"; in questo caso, le misure multiple potrebbero essere rappresentate da diversi item costituiti da affermazioni riguardanti particolari comportamenti ipotetici verso i componenti di tale gruppo sociale (per esempio, "non voglio averci a che fare", "accetterei di sedergli accanto nell'autobus", "lo accetterei come compagno di lavoro", "lo inviterei a casa mia", "lo accetterei come amico", "lo accetterei come parente acquisito"<sup>16</sup>); ciascun soggetto esprimere il proprio accordo ("sì") o disaccordo ("no") rispetto a ciascuna di tali affermazioni. Se l'obiettivo è quello di misurare

- il livello di distanza sociale di ciascun individuo, le misure multiple saranno rappresentate dall'insieme degli item (ossia, l'insieme delle risposte date da un determinato soggetto a tutti gli item può essere sintetizzato e consente di collocare il soggetto sul continuum "distanza sociale percepita");
- il livello di distanza sociale rilevata da ciascun item, le misure multiple saranno rappresentate dall'insieme degli individui (ossia, l'insieme delle risposte registrate per un determinato item da tutti i soggetti può essere sintetizzato e consente di collocare l'item sul continuum "distanza sociale percepita").

- **Contributo alla misurazione di ciascuna misura multipla:** il contributo può essere *uniforme* (ossia, tutte le misure multiple contribuiscono con lo stessa evidenza) o *differenziato* (ossia, le misure multiple contribuiscono con una evidenza diversa); a tale fine, può essere presa in considerazione una particolare caratteristica dell'indicatore elementare, la *trace line*, che definisce la relazione tra il continuum osservato e la frequenza con cui è stato registrato ciascun valore individuato sul continuum. Tale frequenza può essere interpretata in termini di "probabilità di ottenere ciascun valore" (McIver, 1979). In particolare, è possibile associare a ciascun item due distribuzioni corrispondenti a due diverse probabilità:

- *alfa*, probabilità relativa al valore atteso ( "risposta corretta" o "accordo con l'affermazione presentata" o "risposta che va nella direzione della variabile misurata");
- *beta*, probabilità relativa al valore non atteso ( "risposta sbagliata" o "disaccordo con l'affermazione presentata" o "risposta che va nella direzione opposta della variabile misurata").

Di seguito è presentata una tabella che riassume le caratteristiche dei modelli di *scaling*:

<sup>16</sup> Questo esempio fa riferimento alla *Social Distance Scale* di Bogardus, uno strumento psicometrico creato da Emory S. Bogardus al fine di misurare il livello di apertura dei soggetti nei contatti sociali con i componenti di altri gruppi sociali. Tale scala è basata su un modello di *scaling* cumulativo in quanto l'accordo con ciascun item implica l'accordo con quelli precedenti. (v. *Parte II*).



AStRiS 6 – LA RILEVAZIONE E L'ANALISI STATISTICA DEL DATO SOGGETTIVO

Modello di <i>scaling</i>		Dimensionalità	Natura dei dati	Tecnica di <i>scaling</i>	Criterio per la verifica del modello	Standard di misurazione: punteggio finale attribuito a	
Additivo	Unidimensionale	Uni	Stimolo-unico	Non-comparativo	Consistenza interna	Casi	
	Multidimensionale	Multi	Stimolo-unico	Non-comparativo	Dimensionalità degli stimoli	Casi	
Cumulativi	<i>Thurstone scale (differential scale)</i>		Uni	Confronto	Comparativo (confronti accoppiati o <i>rank-order</i> )	Metrica esistente tra gli stimoli	Oggetti
	Metodologia Q		Uni	Confronto	Comparativo ( <i>rank-order</i> o <i>rating</i> comparativo)		Oggetti
	Deterministici	Guttman	Uni	Stimolo-unico	Non-comparativo	Analisi dello scalogramma: riproducibilità, scalabilità e predicibilità	Casi e oggetti
		<i>Multidimensional Scalogram Analysis (MSA)</i>	Bi			Regionalità e contiguità	Casi e oggetti
		<i>Partial Ordered Scalogram Analysis (POSA)</i>	Bi			Corretta rappresentazione	Casi e oggetti
	Probabilistici	Monotoni (con uno o più parametri)		Stimolo-unico	Non-comparativo	<ul style="list-style-type: none"> <li>Stima dei parametri (massima verosimiglianza)</li> <li>Bontà di 'adattamento (analisi del <i>misfit</i> e dei residui)</li> </ul>	Casi e oggetti (senza condensazione)
Perceptual Mapping	<i>Multidimensional scaling</i>		Multi	Somiglianza	Comparativo (confronti accoppiati)	Bontà di adattamento delle distanze alle prossimità (stress, alienazione)	Oggetti
	<i>Unfolding</i>		Uni e Multi	Preferenza	Comparativo	Bontà di adattamento delle distanze alle preferenze ordinali	Casi e oggetti
Modello congiunto		Multi	Preferenza	Comparativo ( <i>rank-order</i> )	Bontà di adattamento del modello ( <i>part-worth</i> ) all'ordinamento degli stimoli osservato	Oggetti	

### **Approcci tecnici alla condensazione**

La condensazione dei valori degli indicatori elementari (successiva alla verifica del modello di *scaling* adottato) in  $d$  valori sintetici (dove  $d$  corrisponde al numero di dimensioni individuate) richiede l'adozione e l'applicazione di particolari approcci tecnici e analitici – che in molti casi accomunano indicatori soggettivi e oggettivi (Nardo, 2005; Sharpe, 2004; Tarantola, 2000) – legati al trattamento dei dati; tali approcci, finalizzati ad ottenere indicatori sintetici non solo significativi ma anche interpretabili, consentono di (parte I, cap. 5):

1. definire l'importanza di ciascun indicatore elementare (item) da condensare (*criteri di ponderazione* detti anche *weighting criteria*),
2. identificare la tecnica di condensazione dei valori degli indicatori elementari che vanno a comporre l'indicatore sintetico (*tecniche di aggregazione* dette anche *aggregating-over-items techniques*),
3. verificare la robustezza dell'indicatore sintetico in termini di capacità di produrre misure corrette e stabili (*uncertainty analysis, sensitivity analysis*),
4. verificare la capacità di discriminare dell'indicatore sintetico (*verifica della selettività e individuazione dei valori-soglia*).

#### **1. Weighting criteria**

L'aver individuato gli indicatori che definiscono il concetto complesso in indicatori elementari non implica che ciascuno di questi contribuisca con la stessa importanza alla valutazione della variabile latente. Per tale motivo, prima di procedere all'aggregazione degli indicatori elementari è necessario definire un sistema di ponderazione che consenta di attribuire un peso a ciascuno degli item considerati (Ghiselli, 1964). Quando non è possibile contare su un sistema implicito di ponderazione, occorre adottare un criterio che consenta di definire un sistema di ponderazione: il sistema di pesi deve riprodurre il più fedelmente possibile il contributo che ciascun item fornisce alla costruzione dell'indicatore sintetico. In questa ottica, la definizione del sistema di ponderazione può costituire un affinamento del modello di misurazione adottato.

Dal punto di vista tecnico, la ponderazione consiste nel definire e attribuire un peso a ciascun indicatore elementare. Nel successivo calcolo del punteggio aggregato, ciascun peso dovrà essere moltiplicato per il valore individuale dell'indicatore cui si riferisce.

La scelta tra i diversi approcci per la definizione dei pesi non è sempre facile da affrontare e spesso è fatta sulla base di valutazioni soggettive. In ogni caso, nel prendere la decisione è necessario considerare (Nardo, 2005):

- la struttura teorica definita e la struttura logica e concettuale su cui è basata la variabile latente e, conseguentemente, l'indicatore sintetico,
- il significato e il contributo di ciascun indicatore elementare all'aggregazione,
- la qualità dei dati e la adeguatezza statistica degli item.

In questo senso, indipendentemente dall'approccio utilizzato, i pesi rappresentano valori di giudizio. E' importante che le ragioni che hanno condotto alla scelta dei pesi, la metodologia adottata e i risultati che si sarebbero ottenuti adottando altre metodologie, ragionevolmente adottabili siano formalmente esplicitate.

Nell'individuare la procedura di identificazione dei pesi si distingue innanzitutto tra:

- attribuzione di pesi uguali (*Equal Weighting - EW*)<sup>17</sup>
- attribuzione di pesi differenziati (*Different Weighting - DW*).

Qualunque decisione venga presa al riguardo, l'indicatore sintetico ne sarà fortemente influenzato. Le relazioni reciproche tra i casi possono cambiare anche profondamente semplicemente cambiando

---

<sup>17</sup> Attribuire "pesi uguali" non significa necessariamente attribuire "pesi unitari".

i pesi attribuiti agli item.

Nella maggior parte delle applicazioni pratiche, si preferisce attribuire lo stesso peso agli indicatori elementari considerati. Ciò succede quando:

- la struttura teorica alla base della costruzione dell'indicatore sintetico attribuisce a ciascuno degli indicatori elementari la stessa adeguatezza nel definire il concetto composito da misurare,
- la struttura teorica non consente di derivare in modo coerente ipotesi di differenziazione dei pesi,
- non esiste una conoscenza statistica o empirica sufficiente sulla quale basare la definizione di pesi,
- non esiste accordo sulle alternative adottabili.

In alcuni casi, attribuire lo stesso peso può essere però rischioso. Ciò accade soprattutto quando:

- nel definire un indicatore sintetico si individuano componenti differenziate definite da un numero diverso di indicatori elementari; in questi casi, attribuire pesi differenti alle diverse componenti (facendo riferimento, per esempio, al numero di indicatori elementari che le compongono) può produrre per l'indicatore sintetico una struttura non bilanciata;
- alcuni tra gli indicatori elementari misurano la stessa componente (alta correlazione tra indicatori elementari): il risultato è equivalente a quello che si otterrebbe se si attribuisse un peso alto ad uno di tali indicatori (*double weighted* o *double counting*).

La decisione di procedere all'attribuzione di pesi differenziati richiede più che la identificazione dei pesi, la selezione del metodo più appropriato per la loro identificazione; tra i metodi che possono essere selezionati ricordiamo i seguenti (Nardo, 2005):

1. **metodi statistici:**

- a. *Correlazione*,
- b. *Analisi delle Componenti Principali (Principal Component Analysis, PCA)*,
- c. *Data Envelopment Analysis (DEA)*,
- d. *Unobserved Components Models (UCM)*.

2. **multi-attribute models:**

- a. *Multi-Attribute Decision Making → Analytic Hierarchy Processes (AHP)*,
- b. *Multi-Attribute Compositional Model → Conjoint Analysis (CA)*.

3. **expertise methods:** si procede attraverso opinioni di esperti in modo che i pesi riflettano meglio le definizioni teoriche. Una tecnica utilizzata in questi casi è quella del *Budget Allocation (BAL)*.

Attribuire pesi differenziati può essere altrettanto rischioso, soprattutto quando la decisione non è sostenuta

- a livello teorico, considerando il significato di ciascun indicatore elementare e il suo impatto sulla sintesi,
- a livello metodologico, adottando le tecniche coerenti con la struttura teorica. Nel prendere qualunque decisione riguardo alla ponderazione è però saggio abbandonare la nozione che possa esistere un insieme di pesi capaci di esprimere in modo perfetto il contributo reale e relativo di ciascun indicatore elementare alla descrizione e misurazione del fenomeno.

## **2. Aggregating-over-items techniques**

La scelta della tecnica di aggregazione deve essere coerente con quanto è stato definito precedentemente. L'applicazione di una tecnica richiede l'adozione di assunti riguardanti:

- il livello di complessità dell'indicatore da creare (*dimensionalità*) espresso in termini di omogeneità tra gli item (relazione tra item),
- il tipo di relazione tra indicatori elementari e variabile latente,
- la *compensabilità* tra item,
- la *omogeneità di scaling* degli item.

La letteratura offre molti esempi di tecniche di aggregazione (parte I, cap. 5). Tra quelle che seguono l'approccio lineare, le tecniche additive sono le più utilizzate. Dall'altra parte, le tecniche

moltiplicative (che seguono l'approccio geometrico) o quelle basate sulla *multi-criteria analysis* (che seguono l'approccio non compensativo) consentono di affrontare i problemi che la compensabilità tra gli item può causare.

		Approcci di aggregazione					
		Aggregazione lineare			Aggregazione geometrica		Aggregazione non-compensativa
		Additiva classica		Cumulativa	Uni	Indipendenza	
Assunti	Dimensionalità (relazione tra item)	Uni	Indipendenza	Uni			Uni
	Relazione "indicatori elementari - variabile latente"	Monotonica		Differenziale	Monotonica		
	Compensabilità tra item	Ammessa		Non ammessa (gradualità, scalabilità)	Ammessa		Non ammessa
	Omogeneità di scaling degli item	Richiesta		Richiesta	Richiesta		Non richiesta

### 3. Verifica della robustezza (*uncertainty and sensitivity analysis*)

Come abbiamo visto, molte sono le scelte da fare nel procedimento di aggregazione delle misure multiple; tali scelte possono influenzare la robustezza ovvero la capacità dell'indicatore sintetico di produrre misure corrette e stabili (Nardo, 2005; Saisana, 2005; Saltelli, 2004; Sharpe, 2004; Tarantola, 2000). Verificare la robustezza consente di valutare il ruolo e le conseguenze della soggettività delle scelte fatte riguardo a:

- il modello di stima dell'errore di misurazione nei dati;
- il procedimento di selezione degli indicatori elementari;
- il procedimento di trattamento dei dati (trattamento dei dati *missing*, la standardizzazione e la normalizzazione dei dati, ecc.);
- il criterio per l'attribuzione dei pesi;
- la tecnica di aggregazione.

Per poter valutare la robustezza dell'indicatore sintetico, è possibile applicare una specifica procedura di analisi che tratti tutte le scelte che possono rappresentare fonti di incertezza. In altre parole, la robustezza viene valutata verificando e confrontando le possibili diverse performance che si sarebbero ottenute prendendo decisioni diverse lungo tutto il procedimento di condensazione. In particolare, il procedimento consente di

- valutare l'applicabilità del modello di misurazione ovvero i fattori che più di altri contribuiscono alla variabilità del punteggio,
- determinare quali sono le scelte che producono valori più accurati,
- comprendere il comportamento del modello adottato,
- determinare la qualità del modello adottato.

Tale procedimento, che si inserisce nel più ampio settore della *what-if analysis*, è condotto attraverso due fasi; ciascuna fase corrisponde ad una diversa metodologia di analisi:

1. ***uncertainty analysis***: l'obiettivo di tale metodo è quello di analizzare quanto l'indicatore sintetico dipende dall'informazione che lo compone. Al fine di valutare come le fonti di incertezza influenzano il punteggio sintetico ottenuto, la procedura prevede l'individuazione per ciascun caso di diversi **scenari**, ciascuno dei quali corrisponde ad una certa combinazione di scelte che produce un determinato valore sintetico;
2. ***sensitivity analysis***: l'obiettivo di tale metodo è quello di valutare il contributo di ciascuna delle potenziali fonti di incertezza scomponendo la varianza totale del punteggio sintetico ottenuto. A tal fine, il metodo prevede la verifica di quanto il punteggio è sensibile alle diverse scelte (piccoli cambiamenti rivelano bassa sensibilità del punteggio).

I due approcci, trattati quasi sempre in contesti separati, sono molto popolari in tutti quei campi in cui si richiede lo sviluppo e la verifica di modelli (applicazioni finanziarie, analisi del rischio, reti neurali, ecc.); inoltre, la *uncertainty analysis* è adottata e utilizzata più spesso della *sensitivity analysis* (Jamison, 2001). L'utilizzo iterativo e sinergico dei due approcci di analisi si è però

rivelato (Saisana, 2005; Saltelli, 2004; Tarantola, 2000) utile e potente nello sviluppo di misure aggregate.<sup>18</sup>

#### 4. Verifica della capacità di discriminare

La verifica della selettività di un indicatore sintetico consiste nell'analisi della sua capacità di

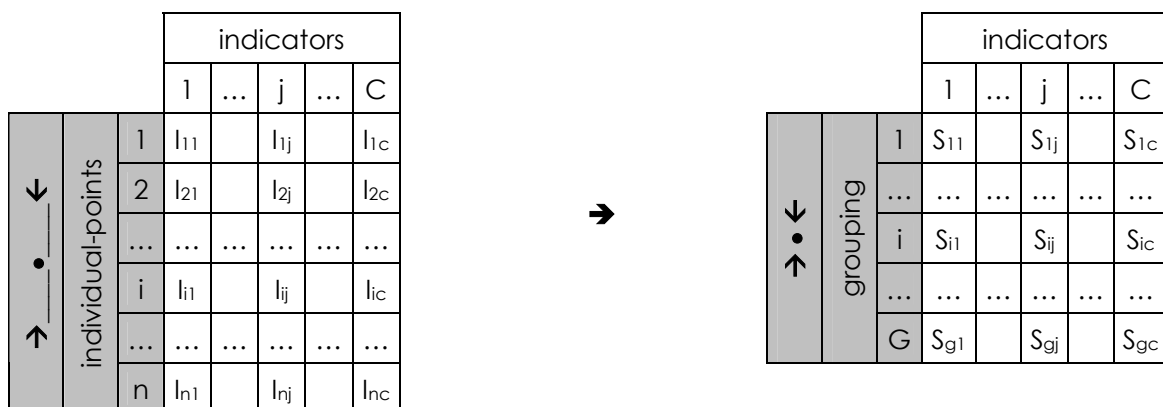
- discriminare tra casi o gruppi (per questo obiettivo si può procedere utilizzando le classiche tecniche statistiche di analisi per la verifica dell'ipotesi),
- distribuire i casi in modo da evitare concentrazioni di punteggi individuali in pochi segmenti del continuum; a tal fine esistono anche alcuni tentativi per definire coefficienti di discriminazione (Guilford, 1954; Maggino, 2003, 2007),
- presentare valori interpretabili in termini di selettività attraverso l'individuazione di particolari valori o punteggi di riferimento che consentano di interpretare i punteggi individuali ed eventualmente di selezionare i casi secondo particolari criteri; l'individuazione di tali valori, che nel caso di dati continui e discreti sono anche detti, rispettivamente, **cut-point** e **cut-off**, è particolarmente utile quando il punteggio sintetico deve essere utilizzato a fini diagnostici e di *screening*.

#### Modello di condensazione dei casi individuali

L'obiettivo in questo caso è quello di condensare i valori individuali, attraverso una procedura di *segmentazione*, e di produrre nuove valori sintetici da attribuire a nuove unità significative (raggruppamenti) identificate secondo prospettive diverse (tipologica, temporale, geografica). L'analisi fatta secondo tale prospettiva mira a sintetizzare i valori individuali presenti all'interno di ciascun raggruppamento ed attribuire il valore sintetico ottenuto al raggruppamento stesso (*from individual-points to grouping-point*).

Questa prospettiva di condensazione è delicata specialmente quando gli indicatori coinvolti riguardano caratteristiche non cumulabili, come quelle soggettive; in presenza di dati soggettivi è per questo necessario definire modelli ad-hoc di condensazione che non prevedano la semplice somma tra i valori individuali.

Operativamente tale processo corrisponde a condensare le informazioni in un numero minore di righe della matrice dei dati, come esemplificato nella seguente figura:



#### 1. Condensing criteria

E' possibile identificare due criteri di condensazione:

<sup>18</sup> Da non trascurare è anche la possibilità di applicare tecniche come la *cluster analysis* che consentono di valutare delle diverse scelte sulle tipologie registrabili tra i casi osservati.

**A. Omogeneità:** i valori vengono condensati se i casi individuali risultano essere omogenei rispetto agli indicatori di interesse; ciò consente il confronto tra i diversi gruppi individuati (**tipologie**) anche attraverso altre variabili di contesto; l'identificazione delle tipologie rappresenta un momento importante che richiede approcci analitici che consentono di individuare tra i casi osservati, gruppi omogenei (Aldenderfer, 1984; Bailey, 1994; Corter, 1996; Lis, 1977; Hair, 1998; Maggino, 2005):

- **segmentation analysis**, che può essere realizzata attraverso approcci diversi (*Hierarchical Cluster Analysis, Q Analysis*);
- **partitioning analysis**, che può essere realizzata attraverso approcci quali *K Means Methods, Iterative Reclassification Methods, "Sift and Shift" Methods, Convergent Methods*; **tandem analysis** (o **Factorial k-means Analysis**) che combina l'Analisi delle Componenti Principali con un *partitioning method (K Means method)*.<sup>19</sup>

Il risultato ottenuto può variare in funzione delle scelte fatte in termini di:

- indicatori selezionati;
- misure utilizzate per valutare le prossimità tra i punti individuali;
- metodo utilizzato per assegnare un caso individuale ad un *grouping*;
- criterio utilizzato per determinare il numero di gruppi;
- criterio utilizzato per verificare l'interpretabilità dei gruppi.

**B. Funzionalità:** i valori vengono condensati se i casi individuali appartengono a raggruppamenti per i quali non è richiesta alcuna procedura analitica per la loro identificazione; essi sono preesistenti e predefiniti in termini di **gruppi** (sociali, generazionali, ecc.), **aree** (geografiche, amministrative, ecc.), **periodi di tempo** (anni, decenni, ecc.).

## 2. Weighting criteria

L'attribuzione di un peso a ciascun caso da condensare nel raggruppamento si pone in particolare modo nel caso di rilevazioni campionarie.

Per tale motivo il problema è affrontato direttamente nell'ambito degli approcci della statistica induttiva con particolare riferimento alle metodologie e alle tecniche di campionamento.

## 3. Aggregating-over-individuals techniques

L'aggregazione dei valori individuali nel caso in cui il raggruppamento sia stato identificato attraverso il criterio di omogeneità presenta relativamente pochi problemi statistici: l'omogeneità, infatti, consente di condensare le informazioni individuali utilizzando semplici strumenti statistici (tecniche di tendenza centrale, univariate, come la media o la mediana, o multivariate, come il centroide).

L'aggregazione dei valori individuali nel caso in cui il raggruppamento sia stato identificato attraverso il criterio di eterogeneità viene realizzata attraverso particolari approcci che presentano comunque non pochi problemi statistici. L'applicazione delle tradizionali tecniche statistiche di sintesi non consente di evidenziare le caratteristiche distributive dei diversi raggruppamenti e conseguentemente non consente confronti tra raggruppamenti. A tale proposito esistono alcuni tentativi di pesare valori medi secondo criteri diversi (Kalmijn, 2005; Veenhoven, 2005).

Il problema dell'aggregazione dei punti individuali è ben noto in molti campi scientifici come in economia<sup>20</sup> e in informatica (dove è affrontato utilizzando particolari approcci analitici come l'analisi di aggregazione probabilistica). Comunque, interpretare le informazioni prodotte dalla

<sup>19</sup> Si procede così ad un adattamento simultaneo di un modello discreto di *clustering* e di un metodo fattoriale continuo per identificare la migliore partizione degli oggetti, descritti attraverso le migliori combinazioni ortogonali lineari delle variabili (fattori) secondo il criterio dei minimi quadrati. Tale approccio presenta molte potenzialità in quanto consente di raggiungere un doppio obiettivo: riduzione dei dati e sintesi, simultaneamente nella direzione dei casi e delle variabili. Attraverso un veloce algoritmo (*alternating least-squares*) è possibile estendere l'applicazione di tale metodo ad ampi dataset (Nardo, 2005).

<sup>20</sup> In campo econometrico, sono state sviluppate particolari metodologie empiriche che consentono di spiegare le differenze sistematiche individuali (*compositional heterogeneity*) che possono avere implicazioni importanti sull'interpretazione dei valori aggregati (Stoker, 1993).

procedura di segmentazione non è semplice. Se per esempio si attribuisce ad una certa area un certo livello di soddisfazione individuale si tende a continuare ad attribuirlo uniformemente a tutti i punti individuali che la compongono che non necessariamente presentano il valore presentato dall'indicatore (**ecological fallacy**).

In altri termini, il pericolo consiste nell'attribuire a tutti i componenti di un gruppo le caratteristiche dell'interno gruppo (stereotipi).