

Misurazione e scale attitudinali

1. Il problema della misurazione nelle ricerche di mercato

La ricerca di mercato richiede una o più misurazioni funzionali all'oggetto di studio, come ad esempio la probabilità di acquisto di un prodotto di nuova concezione, se un individuo preferisca una marca piuttosto che un'altra, la modalità preferita di un determinato attributo, come il colore di un autoveicolo, etc.

Per misurazione si intende il processo mediante il quale si assegnano valori numerici (punteggi) alle caratteristiche, proprietà o attributi di oggetti, eventi, stati o individui, secondo regole predefinite. E' utile distinguere diversi livelli o scale di misurazione a seconda delle proprietà che soddisfano e dell'insieme delle trasformazioni ammissibili: per queste ultime si intende qualsiasi trasformazione dei valori numerici che lascia inalterato il significato della scala originaria.

Scala nominale Le unità sono classificate in funzione dell'appartenenza ad una particolare classe o modalità (principio di identità). Si pensi alla classificazione di un campione di consumatori in ragione del motivo per il quale è stata accordata la preferenza ad un determinato esercizio commerciale (vicinanza, promozioni, spaziosità, etc.).

Scala ordinale Oltre al principio di identità vale il principio di ordinamento, per cui, date due misurazioni, x e y , si può affermare $x < y$ ovvero $x = y$ ovvero $x > y$. Si possono elencare gli oggetti secondo un certo ordine, ma non è possibile valutare la differenza di intensità con cui il carattere o attributo è presente nell'unità. Tutte le trasformazioni monotoniche sono ammissibili, in quanto preservano l'ordinamento. La scala ordinale risulta importante per le ricerche di mercato, in cui buona parte delle misurazioni cadono in questa categoria.

Scala ad intervallo Oltre alle caratteristiche delle scale precedenti, le differenze sono comparabili, ovvero a distanze numeriche uguali sulla scala corrispondono distanze eguali con riferimento al carattere considerato: tra 1 e 3 passa la stessa differenza che esiste tra 15 e 17. Pertanto, è lecito confrontare l'intensità del fenomeno in unità diverse. Tuttavia, non esiste una origine naturale e l'unità di misurazione è arbitraria ed è ammissibile la classe delle trasformazioni lineari del tipo $y = a + bx$ (un esempio classico riguarda la misurazione delle temperature, che può essere effettuata mediante i gradi Celsius e Fahrenheit; per la conversione si utilizza la relazione $C = (5F - 160)/9$). Un ulteriore esempio è rappresentato dal tempo t di un evento, la cui origine viene fissata a partire da un tempo t_0 arbitrario.

Date due misurazioni effettuate sugli individui i e j mediante la medesima scala ad intervallo, contrassegnate rispettivamente da y_i e y_j ha senso operare i confronti $y_i - y_j$ e $y_j - y_i$, al fine di misurare la differenza di intensità.

Scala di rapporti A differenza della scala precedente, esiste un'origine naturale (zero assoluto) che denota l'assenza del carattere. Esempi di misurazioni su questa scala sono le vendite, i costi di una impresa, la quota di mercato. La scala prende il nome dal fatto che, date due misurazioni effettuate sugli individui i e j , contrassegnate rispettivamente da y_i e y_j , oltre alle differenze $y_i - y_j$, ha senso calcolare i rapporti y_j/y_i o viceversa. I valori numerici sulla scala riflettono non solamente il possesso della caratteristica, l'ordine e le differenze nell'ordine di grandezza, ma anche la dimensione assoluta. La classe delle trasformazioni ammissibili è del tipo $y = cx$ (si pensi alla conversione di una grandezza economica in altra valuta).

Si osservi che l'ordine di presentazione delle scale riflette il fatto che il contenuto informativo delle scale è crescente, nel senso che nel passaggio alla successiva si aggiunge una proprietà non posseduta dalla scala precedente.

I metodi di sintesi ed analisi validi per una particolare scala non sono sempre estendibili alle scale con minore contenuto informativo. Ad esempio, con riferimento al calcolo di un indicatore di sintesi della distribuzione dei valori di una misurazione, è immediato riconoscere che la mediana sia ammissibile per la scala ordinale e successive, la media per la scala ad intervallo e di rapporti, un numero indice solo per quella di rapporto; mentre la moda è definita per tutte le scale.

Questo argomento non deve indurre a tentare di utilizzare sempre la scala di rapporti: solitamente esiste un livello ottimale di misurazione a seconda del fenomeno sottostante, tale per cui la scala naturale per un carattere come il sesso è quella nominale. E' consentito, tuttavia, declassare il livello di misurazione, ad esempio passando da una scala di rapporti ad una ordinale suddividendo in classi un fenomeno come le vendite.

Una classificazione più rozza distingue le misurazioni *qualitative* (nominali e ordinali) da quelle *quantitative* (ad intervallo e di rapporti). Le prime sono anche dette categoriche, dal momento che le modalità di manifestazione sono raggruppate in categorie mutualmente esclusive ed esaustive, sconnesse (sesso) o ordinate (titolo di studio).

Un'ulteriore classificazione è operata dalla distinzione tra misurazioni *continue*, il cui dominio è rappresentato dall'insieme dei numeri reali e quelle discrete. Ad esempio, il numero di esami sostenuti rappresenta una misurazione su scala di rapporti quantitativa e discreta.

2.1 La misurazione delle attitudini

La misurazione può riguardare elementi fattuali (es. consumo di latte nel mese di gennaio) e elementi non fattuali, come le attitudini (es. fedeltà alla marca) e le opinioni. Nella prima fattispecie esiste un valore vero sottostante l'informazione richiesta che può, almeno in linea di principio, essere determinato da fonti esterne alternative alla risposta dell'intervistato. Come importante conseguenza si ha che misurazioni possono essere validate. Nel secondo caso questo non avviene. Un ulteriore elemento di differenziazione è ravvisabile nei fattori che condizionano la risposta e l'errore di misura (si veda [6]): nel primo caso la definizione, comprensione, memoria, desiderabilità sociale; nel secondo la formulazione del quesito, la presenza di alternative bilanciate, l'ordine di presentazione dei quesiti e delle alternative, etc. Una categoria intermedia riguarda il giudizio personale su fatti concreti (es. come giudica lo stato dell'economia?).

Il rilievo della misura delle attitudini sta nel fatto che si ritiene che queste condizionino il comportamento degli utenti/consumatori, per cui ha rilievo per la funzione di marketing conoscere i fattori che presiedono alla loro formazione.

2.1.1 Costruzione delle scale attitudinali

Una scala attitudinale costituisce un caso particolare di misurazione e può essere costruita mediante tecniche diverse. L'idea fondamentale che accomuna queste tecniche è che le attitudini possano essere rappresentate in uno spazio unidimensionale (ad es. lungo una retta): i rispondenti possono essere ordinati in funzione dell'intensità della loro attitudine lungo un continuo.

E' questione assai dibattuta se le misurazioni cui danno origine siano ad intervallo o semplicemente ordinali. Una rilevante distinzione è tra le scale attitudinali semplici, o a quesito singolo, e le scale attitudinali complesse, o a quesiti multipli. Spesso, la comprensione dei fattori che presiedono alla formazione delle attitudini richiede l'impiego di quesiti multipli (es. percezione e valutazione dei servizi offerti da un istituto di credito). E' stato mostrato che i quesiti diretti sono non affidabili, dimostrando scarsa stabilità temporale; i quesiti multipli aumentano l'affidabilità della scala, poiché possono essere viste come misure ripetute del dominio di misurazione; sono inoltre imprecisi poiché restringono il livello di misurazione a quello nominale e tendono a categorizzare i rispondenti in due gruppi senza cogliere le sfumature all'interno; infine, l'attitudine o

il fenomeno da misurare presenta sfumature che non possono essere colte con un quesito singolo.

3 Scale attitudinali semplici o di valutazione (rating scales)

Le scale semplici realizzano una misurazione diretta dell'attitudine o delle sue componenti attraverso un quesito singolo: il rispondente è chiamato ad indicare la sua posizione lungo una scala di misurazione delle attitudini prestabilita.

I possibili campi di applicazione vanno dalla misurazione dell'attitudine o preferenza generale verso un oggetto, prodotto, marca etc, all'intensità con cui un oggetto possiede un determinato attributo, all'atteggiamento individuale verso un attributo (es. gusto), fino all'importanza assegnata ad un particolare attributo (es. contenuto di caffeina).

La caratteristica fondamentale va ravvisata nella circostanza che il soggetto che effettua la valutazione colloca l'oggetto valutato in un punto lungo un continuo o in una delle categorie ordinate in successione; un valore numerico è assegnato al punto o alla categoria.

Solitamente si opera una distinzione tra scale non comparative e scale comparative, a seconda della natura assoluta o relativa della valutazione.

Noncomparative rating scales (scale di valutazione non comparativa) Al rispondente viene richiesto di valutare un oggetto (prodotto) senza far riferimento ad uno standard prefissato o ad oggetti concorrenti; conosciute anche come scale monadiche poiché solo un prodotto o marca è valutato. Il loro principale impiego è nella fase di test della concezione di un prodotto nuovo. Ad esempio, si può condurre un'intervista personale di un numero prefissato di individui presso un centro commerciale; ciascun intervistato prova il prodotto e dichiara una delle seguenti alternative: 1. Sicuramente acquisterò 2. Probabilmente acquisterò 3. Forse acquisterò 4. Forse non acquisterò. 5. Probabilmente non acquisterò 6. Sicuramente non acquisterò. La percentuale di risposte che rientrano nelle valutazioni più favorevoli (prime due) fornisce il *top box score*.

A seconda delle modalità di presentazione distinguiamo:

- Scale grafiche di valutazione non comparativa: (scale di valutazione continue) al rispondente viene richiesto di fornire la sua valutazione apponendo un segno su una linea che si estende da un estremo all'altro dell'attitudine:

E' soddisfatto dei servizi offerti dalla banca x?

Molto _____ Per niente

- Scale categoriche di valutazione non comparativa: il rispondente può scegliere tra un limitato numero di opzioni ordinate a seconda della loro posizione sulla

scala. Costituisce lo strumento di misura delle attitudini più largamente utilizzato. Non esiste una formula che sia univocamente migliore delle altre. Le opzioni - categorie - possono essere descritte verbalmente o possono essere numeriche (es. scala percentuale). Il numero delle categorie dipende dalla natura dell'attitudine misurata.

E' soddisfatto dei servizi offerti dalla banca x?

Molto	Abbastanza	Poco	Per niente
soddisfatto	soddisfatto	soddisfatto	soddisfatto
☐	☐	☐	☐

Una decisione importante riguarda l'uso di opzioni bilanciate (scale simmetriche) o non bilanciate (scale asimmetriche); le prime implicano un identico numero di opzioni favorevoli e sfavorevoli; in questo caso ha rilievo se un numero pari o dispari di opzioni viene scelto. Nel secondo caso la modalità centrale è neutrale. L'esperienza dimostra che l'inclusione della modalità neutrale incide sui risultati. Se ci si attende che una parte del campione non abbia alcuna valutazione da esprimere si dovrebbe porre una opzione esterna alla scala del tipo: nessuna opinione, non so.

E' soddisfatto dei servizi offerti dalla banca x?

Molto	Abbastanza	Poco	Per niente	Non
soddisfatto	soddisfatto	soddisfatto	soddisfatto	so
☐	☐	☐	☐	☐

Comparative rating scales (scale di valutazione comparative) Nel caso precedente la valutazione individuale non fa riferimento ad uno standard comune. Una scala comparativa viene utilizzata nei casi in cui si desidera un comune riferimento per tutti i rispondenti (ad esempio una o più marche concorrenti).

- Misure binarie di preferenza e discriminazione. Sono comunemente impiegate nel campo della misurazione della capacità del consumatore di saper discriminare tra diverse versioni dello stesso prodotto o tra marche e della preferenza tra versioni alternative del prodotto da introdurre sul mercato.

Nella *Paired comparison* al rispondente vengono presentati due oggetti (marche, prodotti, attributi) alla volta tra cui deve selezionarne uno in base ad un criterio prestabilito (similarità o preferenza). Tale tecnica è generalmente limitata ad un attributo e poche marche, poiché per n marche, sono richiesti $n(n-1)/2$ confronti.

La misurazione può essere effettuata su una scala nominale (la marca A è simile alla/migliore della marca B ovvero mediante una scala ordinale o ad intervallo:

Simile	_____	Diversa
Migliore	_____	Peggiora

Nel primo caso si ottiene una graduatoria tra le marche contando il numero delle volte in cui viene preferita (ad esempio, se nei 6 confronti binari concernenti quattro marche - A, B, C, D - si ha che esse vengono preferite rispettivamente 2,3,0,1 volte, allora si ottiene la scala attitudinale di tipo ordinale $B > A > D > C$).

Una variante, *Double-Paired comparison*, mira a misurare sia la capacità di discriminare che la preferenza. Il confronto riguarda 4 oggetti che consistono in due campioni identici dell'oggetto A e dell'oggetto B. L'individuo ripete la comparazione a coppia due volte e si considera la consistenza dei risultati.

Nella *Triangle Discrimination and Triangle Preference* al rispondente vengono forniti due esemplari di un prodotto e un solo esemplare dell'altro; nella prima deve identificare l'esemplare unico per poi rispondere ad una domanda sulla preferenza. Nella seconda deve ordinare i tre esemplari: coloro che hanno capacità discriminatoria collocheranno l'esemplare unico al primo o al terzo posto.

- Misure ordinali di valutazione (Rank Order Rating Scales). Al rispondente viene richiesto di ordinare un insieme di oggetti in base ad un criterio prestabilito o con riferimento ad un oggetto che funge da standard. E' essenziale che nell'insieme delle alternative siano contemplate tutte le marche o tutte le versioni del prodotto rilevanti. Il metodo è ampiamente utilizzato per confronti tra marche e attributi del prodotto, risultando meno *time-consuming* del confronto binario e consentendo un confronto più realistico; di semplice gestione, ha il difetto di produrre soltanto dati ordinali e consente solo un limitato numero di analisi.

Il confronto può essere semplificato richiedendo di effettuare un ordinamento a più stadi, tale che ad esempio, il primo ordinamento consiste nell'allocazione degli oggetti ad un numero limitato di gruppi mutualmente esclusivi caratterizzati da diversi livelli di similarità o preferenza. Successivamente si ordinano gli oggetti all'interno dei gruppi.

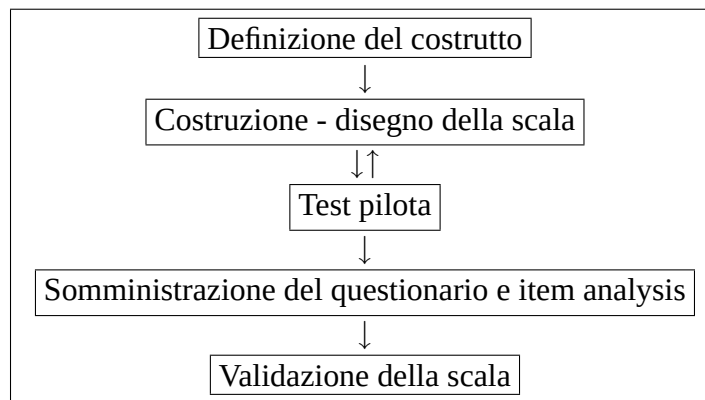
- La scala a somma costante: il rispondente deve distribuire una somma costante (100) tra diverse marche o diversi attributi del prodotto in modo tale da riflettere la preferenza relativa per ciascun oggetto, l'importanza di un attributo o il grado con cui un attributo è presente in un determinato prodotto. I valori individuali sono poi sommati per produrre una scala di rapporti.

Qui di seguito sono elencate 5 caratteristiche di una calzatura da passeggio. Le viene ora richiesto di distribuire 100 punti tra le caratteristiche considerate in relazione alla loro importanza. Quando avrà terminato si assicuri che la somma dei punti attribuiti sia esattamente pari a 100 (se ritiene che una caratteristica sia del tutto irrilevante assegni punteggio 0):

- _____ Comoda da indossare
- _____ Conveniente
- _____ Alla moda
- _____ Di marca prestigiosa
- _____ Conveniente
- 100

4 Scale attitudinali complesse

Lo sviluppo di una scala attitudinale segue un processo che comporta più stadi. I principali possono essere di seguito sintetizzati:



Il primo stadio concerne la definizione coerente dell'oggetto della misurazione (fedeltà alla marca, soddisfazione; nel secondo si rende operativa la misurazione mediante l'individuazione di una batteria di quesiti (*item*) legati al dominio di misurazione. Questi vengono testati su un campione pilota al fine di verificarne in via preliminare la consistenza interna e la validità; ciò consentirà di evitare ridondanze e di eliminare i quesiti ambigui. Segue la fase della somministrazione del questionario al campione e dello spoglio dei risultati, i quali vengono poi sintetizzati in una scala attitudinale. Nella fase finale si valida la scala, nel senso che se ne valuta la capacità di misurare il costrutto per la quale è stata costruita.

Qui di seguito illustriamo alcune delle tecniche disponibili per il disegno della scala. Queste differiscono prevalentemente per la formulazione dei quesiti nel secondo stadio. Approcci alternativi (scalogramma di Guttman, scale fattoriali) sono descritti in [7], cap. 6.

4.1 Scala di Thurstone o equintervallata

La costruzione della scala attitudinale passa attraverso le seguenti fasi: a) Si colleziona un numero m di statement o quesiti sufficientemente elevato, nei confronti dei quali l'intervistato dovrà manifestare il proprio assenso o dissenso. b) viene reclutato un numero

elevato di giudici, ai quali si chiede di ordinare in 11 gruppi i quesiti in ragione della loro connotazione (positiva o negativa) nei confronti del dominio di indagine. c) Per ciascun quesito si calcola il valore mediano ed il campo di variazione. Ad esempio, con riferimento all'*i*-esimo statement, se i giudici sono 100:

Score	1	2	3	4	5	6	7	8	9	10	11	Tot
Frequenze	0	0	0	0	4	11	25	35	20	5	0	100

il punteggio mediano risulta pari a 8 ed il campo di variazione è 5. d) Sono eliminati i quesiti che presentano dispersione più elevata, e che pertanto sono più ambigui e si seleziona un numero ridotto (20-30) di quesiti in modo da coprire l'intero spazio di misurazione delle attitudini. e) I quesiti scelti sono sottoposti al campione di individui f) Il valore della scala attitudinale per il rispondente viene ottenuto considerando il punteggio mediano realizzato sulle affermazioni con le quali egli è d'accordo.

La scala di Thurstone presenta i seguenti punti critici: (i) individui che presentano *pattern* di risposta diversi possono avere lo stesso punteggio attitudinale; questo punto è peraltro comune a tutte le scale attitudinali complesse e consegue dalla monodimensionalità della misurazione attitudinale. (ii) Il metodo non consente di cogliere l'intensità delle preferenze, essendo basato sul semplice accordo/disaccordo con determinate affermazioni. Ciò motiva il ricorso alla formulazione degli *item* su una scala ad intervallo, effettuato nella scala di Likert e del differenziale semantico. (iii) La tecnica è costosa e *time-consuming*. (iv) I risultati sono condizionati dall'operato dei giudici, che non necessariamente riflettono la popolazione di riferimento.

4.2 Scala di Likert (summated ratings)

La scala di Likert differisce dalla precedente per la modalità di presentazione dei quesiti. Il rispondente indica il suo grado di accordo o disaccordo rispetto ad una serie di affermazioni:

Concordo pienamente	Concordo	Non concordo né dissento	Dissento	Dissento fortemente
☐	☐	☐	☐	☐

Le categorie di risposta sono multiple e ciò consente di aumentare la precisione con cui cogliamo le differenze attitudinali. A ciascuna categoria di risposta viene assegnato un punteggio es. $\{1, 2, 3, 4, 5\}$ o $\{2, 1, 0, -1, -2\}$. Si noti che la scala non possiede un'origine naturale e qualunque trasformazione lineare $y = a + bx$ della medesima risulta ammissibile, vale a dire non altera la natura della misurazione. Il punteggio totalizzato da un individuo fornisce il valore numerico dell'attitudine sulla scala prescelta. Per gli *statement* che hanno una connotazione negativa si richiede il rovesciamento della scala dei punteggi attribuiti a quelli con connotazione positiva.

4.3 Scala del differenziale semantico

Si caratterizza per l'impiego di aggettivi bipolari:

Gradevole	---	---	---	---	---	---	---	Sgradevole
Caro	---	---	---	---	---	---	---	Economico
Veloce	---	---	---	---	---	---	---	Lento

Gradevole	1	2	3	4	5	6	7	Sgradevole
Inutile	1	2	3	4	5	6	7	Utile

Il numero di modalità presentate è solitamente dispari e la modalità centrale è neutrale. Il problema fondamentale di questa scala sta nel fatto che non è sempre possibile trovare una coppia di aggettivi che siano in completa opposizione dal punto di vista semantico. Ai fini della misurazione attitudinale in senso stretto, si assegna a ciascun quesito un punteggio tra 1 e 7 o da -3 a +3; come nel caso precedente l'origine dell'intervallo è arbitraria.

Una versione semplificata è la scala di Stapel. Non è bipolare, ma le frasi descrittive e gli aggettivi sono presentati separatamente; inoltre i punti sulla scala sono identificati da numeri e il numero di posizioni lungo la scala è pari (manca la modalità neutrale). Risulta di più facile gestione in alcuni tipi di intervista, ed in particolare in quella telefonica.

Es.: L'orario di apertura è

Soddisfacente	---	---	---	---	---	---
	-3	-2	-1	1	2	3

4.4 Analisi aggregata e del profilo

Il set informativo completo di una misurazione attitudinale può essere organizzato come evidenziato nella tabella 2.1.

Le analisi consentite sono

- *Analisi aggregata*: i punteggi individuali sono sommati per ottenere un punteggio individuale complessivo, $y_{h.}$, che costituisce la misura dell'attitudine per l'individuo h ; i dati individuali o per gruppi di individui sono poi confrontati.
- *Analisi del profilo*: viene calcolato il punteggio totale, medio o mediano assegnato dai rispondenti o da un particolare gruppo a ciascuno degli *item*: nel primo caso si ottiene $y_{.i}$, ed il profilo è identificato nell'ultima riga della tabella 2.1. L'analisi del profilo viene utilizzata per identificare i punti deboli e forti dell'oggetto o costrutto che viene misurato.