

Statistica sociale - 7

Prof. Antonio Mussino

a. a. 2016-2017



SAPIENZA
UNIVERSITÀ DI ROMA

Una specifica modalità di codifica: le scale

Le scale - 1

- Una **scala** è una particolare sequenza di domande con risposte chiuse, chiamate **item**,
- che mirano a indagare il posizionamento dell'intervistato su di un'alternativa graduata di atteggiamenti da una situazione (*sfavorevole, disaccordo, insoddisfazione* e così via) a una opposta (*favorevole, accordo, soddisfazione* e così via) su un tema qualsiasi;

Le scale - 2

- In sintesi è una batteria multipla di variabili (in genere qualitative ordinabili) da elaborare congiuntamente per misurare l'atteggiamento dell'intervistato sul tema considerato.
- Quindi si tratta di uno strumento per ***misurare gli atteggiamenti***.
- Gli item sono un ***campione*** di tutti i possibili punti di vista sul tema.

Le scale - 3

- Ovviamente l'item così codificato può anche essere uno solo, ma in questo caso la percezione dell'atteggiamento può essere fallace:
 - può trattarsi di un **indicatore** non in grado di rappresentare l'atteggiamento che vogliamo misurare,
 - ovvero un singolo pezzo di un **puzzle**, che ci interessa studiare completo!
- Esempio: i Professori del Dipartimento di Scienze statistiche.

Le scale - 4

- I passaggi (*step*) per la costruzione di una scala sono tre:
 - decidere quali sono gli item appropriati per misurare l'atteggiamento (campione, esaustività, validità);
 - definire tecnicamente come proporre gli item;
 - testare le proprietà della scala (attendibilità, validità).

La scala Likert - 1

- La scala più comune è quella di **Likert**, caratterizzata dall'aver un numero **dispari** (5,7, raramente 9) di alternative: ad esempio
 - "molto contrario",
 - "abbastanza contrario",
 - "né d'accordo né contrario",
 - "abbastanza d'accordo",
 - "molto d'accordo".

La scala Likert - 2

- Va segnalato che un numero dispari di alternative prevede sempre una posizione ***intermedia*** (di mancato pronunciamento) in cui, oltre a chi è effettivamente incerto, possono rifugiarsi coloro che non vogliono esprimere o far conoscere il loro reale atteggiamento:
- per ovviare a questa situazione si utilizzano spesso scale con un numero **pari** di alternative (4, 6, raramente 8), che ci permettono sempre di contare i "pro" e i "contro".

La scala Likert - 3

- Come esempio riportiamo la seguente batteria di item, somministrata a un campione di studenti di corsi di laurea in Scienze motorie (Falese, 2013).
- La codifica degli item è la seguente:
 - 1 - molto
 - 2 - abbastanza
 - 3 - poco
 - 4 - per niente

RISPONDERE A TUTTE, PER OGNI RIGA – Quanto sei d'accordo?

1. Il calcio e lo sport professionistico in genere, con il loro giro d'affari, sono una delle cause principali della crisi di valori dello sport italiano	1	2	3	4
2. Un atleta che non sia anche un esempio di vita per i giovani non può essere considerato un grande campione sportivo	1	2	3	4
3. Finché non sarà possibile imporre a tutti gli atleti di alto livello il rispetto delle norme antidoping è inutile e ingiusto controllare soltanto i nostri campioni	1	2	3	4
4. Senza lo stimolo del risultato tecnico e della competizione lo sport perderebbe qualunque motivo di interesse per i praticanti	1	2	3	4
5. Il Coni e le federazioni sportive dovrebbero occuparsi soltanto della preparazione olimpica e delle prove internazionali dei nostri atleti	1	2	3	4
6. Una quota consistente dei guadagni derivanti dalla cessione dei diritti televisivi andrebbe destinata allo sport dilettantistico	1	2	3	4
7. La televisione sta progressivamente trasformando lo sport in una specie di spettacolo da circo	1	2	3	4
8. Le attività fisico-motorie che riguardano il tempo libero, la salute, il turismo ecc. non hanno nulla a che fare con lo sport	1	2	3	4

La scala Likert - 4

- Come si può vedere, ciascun item ha un suo ruolo e un suo significato, ma è il complesso dei profili di risposte che permette di definire le opinioni e l'atteggiamento verso lo sport degli intervistati.
- Come lo stesso Likert suggeriva sarebbe utile proporre frasi con polarità invertita fra di loro, per tenere desta l'attenzione dell'intervistato e, qualora questo non avvenisse, individuare e trascurare le risposte date meccanicamente, senza attenzione al contenuto.

La scala Likert - 5

- Come già detto queste scale **non** sono trattabili come misure sul **continuum**: anche se l'atteggiamento vi può essere riprodotto, la distanza fra i punti rappresentativi delle categorie non è detto che sia uguale per gli intervistati.

La scala Likert - 6

- Nel costruire una scala di questo tipo va tenuto presente il cosiddetto effetto di ***curvilinearità***.
- Infatti, si possono dichiarare "molto d'accordo" su un singolo item intervistati che hanno una visione opposta del fenomeno studiato:
 - ad esempio l'item 8 può essere fortemente condiviso da coloro che sono fautori di una visione dello sport tradizionale (competitivo, organizzato, con impegno costante), così come da appassionati delle attività destrutturate *open air*.

La scala Likert - 7

- Poiché il nostro obiettivo era, invece, quello di costruire una contrapposizione fra questi due atteggiamenti, l'item 8 è, di fatto, fuorviante.
- La soluzione non è, come spesso si suggerisce, di proporre solo item **estremi**, che caratterizzino – anche banalmente – le posizioni da contrapporre, ma di elaborare gli item congiuntamente, per avere le diverse sfaccettature degli atteggiamenti.
- Nell'ultima parte del corso vedremo come si possono elaborare queste scale con metodi statistici di **analisi multivariata**.

Altre scale

- Le scale utilizzabili nei questionari sono molteplici: in particolare gli psicologi utilizzano questo strumento per **quantificare** tratti di personalità.
- Qui riportiamo due esempi: una scala di tipo **numerico** e una di tipo **grafico**.

La scala di Cantril

- Una scala numerica può essere assimilata alla valutazione scolastica tradizionale (scala di **Cantril**):
 - si può, infatti, chiedere di dare un voto da 1 a 10 a un servizio, a una struttura e così via.
- Di seguito una batteria di item che miravano a chiedere la valutazione da parte di iscritti al CUS Roma sugli impianti sportivi utilizzati e sulla organizzazione delle attività in essi svolte.
- Anche in questo caso ciascun item ha una sua autonoma valenza, ma è la sintesi (per esempio con un voto medio) di tutti gli item ad avere maggiore significato per la valutazione.

Consideriamo la sede di Tor di Quinto: dia un voto da 1 a 10 (come a scuola) alle seguenti voci, mettendo ND ove non sia in grado di rispondere o non ne abbia usufruito:

- 1. Pulizia spogliatoi** _____
 - 2. Confort degli spogliatoi** _____
 - 3. Armadietti** _____
 - 4. Docce** _____
 - 5. Attrezzi per le attività** _____
 - 6. Istruttore che la ha seguita** _____
 - 7. Pulizia piscina** _____
 - 8. Spazio acqua disponibile** _____
 - 9. Organizzazione dei corsi** _____
 - 10. Temperatura dell'acqua** _____
-

La scala grafica - 1

- Una scala grafica tende a rappresentare su di un **continuum** la percezione dell'intervistato rispetto ad una sensazione da una soglia minima ad una massima.
- Esempio: in una ricerca sui corsi di preparazione al parto si chiese alle intervistate di collocare la sensazione del dolore che provavano in varie esperienze su un continuum rappresentato da un segmento:

min I-----I max



La scala grafica - 2

- Per misurare la risposta è necessario dotarsi di un ***righello*** e registrare la distanza fra il punto di minimo e quello dove si è posizionato l'intervistato.
- Si può ***normalizzare*** la misurazione assegnando un valore pari a 100 alla distanza fra massimo e minimo e riproporzionare i valori delle altre distanze.
- Potrebbe essere utile anche una ***standardizzazione per unità***, perché alcuni individui tendono a *estremizzare* le loro risposte, altri a *mediarle*.

Il differenziale semantico - 1

- Con questa tecnica l'intervistato è posto di fronte a una batteria di item **bipolari**, che mettono a confronto aggettivi antitetici (Osgood, 1957);

➤ ad esempio:

buono I-----I cattivo

triste I-----I felice

freddo I-----I caldo

largo I-----I stretto

e così via.

Il differenziale semantico - 2

- L'ipotesi è che tra i due estremi ci sia un continuum e che questo possa essere trattato come nella scala grafica, oppure operando con l'utilizzo di intervalli di ampiezza uguale, da trattare come nella scala Cantril, o in quella Likert (con un numero di alternative dispari, o pari).
- In questo caso è ancor più evidente la necessità di individuare ranking o classificazioni di atteggiamenti elaborando **congiuntamente** tutti gli item (analisi multivariata)!

Attendibilità - 1

- L'attendibilità (***reliability***) nelle scienze sperimentali è riferita alla affidabilità di un processo di misurazione in più prove svolte in condizioni controllate.
- Nella ricerca sociale la ripetitività delle prove non è garantibile in condizioni stabili,
- tuttavia il concetto di attendibilità è utilizzabile per valutare se, anche a fronte di indagini in tempi e condizioni diverse, la procedura di misurazione o classificazione può essere considerata affidabile.

Attendibilità - 2

- È il caso del “livello di partecipazione sportiva” proposto da COMPASS, che a fronte di condizioni ambientali, strutturali e culturali diverse, consente di classificare i cittadini su sette diversi livelli di partecipazione sulla base di informazioni*, anche raccolte con questionari parzialmente diversi.

* relative alla quantità, qualità e organizzazione della partecipazione stessa.

Attendibilità - 3

- Alcune scale precedentemente citate sono da considerarsi attendibili, quali il punteggio da 1 a 10 per "dare un voto" a un servizio o a una struttura, oppure una scala di Likert con un numero ridotto di alternative:
- è prevedibile che gli intervistati abbiano piena consapevolezza della votazione scolastica o delle alternative proposte nella scala Likert.

Attendibilità - 4

- Semmai potrebbero esservi differenti valutazioni a seconda della tendenza a estremizzare o meno i voti: anche tra i professori c'è chi utilizza tutta la gamma di possibilità tra 1 e 10 e chi non usa i valori più estremi non considerando *dignitosi* i più bassi e *raggiungibili* i più alti!
- In questi casi una scala standardizzata per unità sarebbe preferibile.

Attendibilità - 5

- In altre situazioni, prevalentemente in Psicometria, l'attendibilità viene valutata con metodi specifici che mettano in correlazione le risposte agli item di una scala fra di loro (*split half method*) o con quelli di un'altra scala costruita con item equivalenti (*alternate form method*).

Validità - 1

- Più facile è definire cosa si intende per validità (***validity***):

“stiamo misurando quello che volevamo misurare?”

- ovvero le differenze fra i punteggi o la classificazione delle unità rappresenta la loro effettiva collocazione nella realtà del fenomeno che volevamo ***misurare***?

Validità - 2

- Ancora una volta ci viene in aiuto l'esempio della classificazione COMPASS:
- se avessimo considerato solo la **quantità**, la **qualità** e l'**organizzazione** delle attività strettamente **sportive**, e non anche di quelle **fisico motorie**, non avremmo classificato correttamente tutti gli individui in vari tempi e paesi,
- in quanto la percezione di sport e di attività fisico motoria è differente nelle diverse realtà, quindi una tale classificazione non sarebbe stata **valida** ai nostri fini, anche se essa sarebbe stata pienamente **attendibile!**

Validità - 3

- Una scala ***non attendibile*** è sicuramente ***non valida***.
- Non è vero il contrario: infatti una scala ***attendibile*** potrebbe misurare qualcos'altro da quello che vogliamo misurare!

Validità - 4

- Esistono vari metodi per analizzare la validità: la **content validity** presuppone che i contenuti del fenomeno sotto studio siano tenuti in considerazione, in maniera completa e bilanciata, nell'indicatore (scala).
- Questo metodo si basa sul parere di esperti che definiscano tali contenuti e verifichino la loro presenza: esempi sono il livello di partecipazione COMPASS e l'**ISU (Indice di Sviluppo Umano)**, indice sintesi che contiene pesate misure della **salute**, dell'**istruzione** e del **reddito**.

Validità - 5

- Anche la **construct validity** si basa sul parere di esperti che individuano il tipo e il grado di associazione dell'indicatore con altre variabili da loro scelte che, sulla base di considerazioni teoriche, sono concettualmente legate con l'indicatore.

Validità - 6

- Altri metodi per stabilire la validità sono basati su ***criteri esterni*** osservabili o misurabili, ***contemporaneamente*** o ***successivamente***:
- ad esempio la ***predictive validity*** potrebbe misurare la validità di una metodologia di allenamento di atleti sulla base dei risultati ottenuti da loro ottenuti in una competizione;
- la ***concurrent validity*** potrebbe misurare quella di un test sulla depressione, mettendolo a confronto con le diagnosi di psichiatri.

Coerenza - 1

- Un'altra caratteristica che può essere valutata in una scala è la sua **coerenza interna**, ovvero il fatto che il campione di item scelto sia legato allo stesso atteggiamento che vogliamo misurare.
- In questo caso le correlazioni fra gli item dovrebbero essere alte!

Coerenza - 2

- Un indice per misurare tale **coerenza interna** è l'**alfa di Cronbach**, che si basa su di un'elaborazione della matrice delle correlazioni fra tutti gli item.

$$\alpha = \frac{nM_r}{1 + M_r(n-1)}$$

dove n è il numero di item e M_r la loro correlazione media.

Coerenza - 3

- **Alfa** non è un coefficiente di correlazione; assume un valore positivo fra 0 e 1, purché il verso degli item sia orientato sempre verso la stessa polarità.
- Più è alto, più forte è la coerenza interna: in genere si richiede un valore di 0,70 (Corbetta, 2003).
- Dipende da "n", ma questo non deve spingere a "gonfiare" la scala, perché l'intervistato, per stanchezza e confusione, potrebbe poi rispondere a caso.

Coerenza - 4

- Per calcolare **alfa** si devono calcolare i coefficienti di **correlazione**.
- Abbiamo, però, visto che molte scale utilizzano item che non sono variabili quantitative, bensì qualitative ordinabili,
- pertanto il coefficiente di correlazione da calcolare non dovrebbe essere quello di **Pearson**, bensì quello di **Spearman**:

$$\rho = 1 - \frac{6\sum_i (D_i)^2}{n(n^2-1)}$$

dove "Di" è la differenza dei ranghi della i-esima osservazione e "n" il numero degli item.

Analisi fattoriale

- In realtà non è detto che la batteria degli item sia unidimensionale, ma potrebbe misurare due o più atteggiamenti diversi.
- Per comprendere e studiare questa situazione si utilizzano strategie di Analisi Multivariata di **tipo fattoriale**, diverse a seconda della natura degli item utilizzati:
 - **Analisi in Componenti principali**
 - **Analisi delle Corrispondenze multiple**