

4.3. *Scale cumulative* Le scale cumulative, dette anche di Guttman dal nome del loro autore, sono le misure ordinali che possiamo considerare più corrette¹⁰. Il criterio usato per ordinare gli stati della proprietà è quello cumulativo nel senso che, se gli stati di una proprietà P si pongono in una struttura ordinata, allora un individuo x che manifesta un certo grado r di P , manifesterà anche i gradi precedenti di P a intensità minore di r .

Scalogrammi e riproducibilità. Utilizzando liberamente un noto esempio¹¹, costruiamo una scala della paura, considerando tre indicatori supposti in ordine crescente:

- accelerazione cardiaca (AC);
- nausea (N);
- vomito (V).

Se la struttura è cumulativa, un individuo che manifesta vomito, manifesterà anche nausea e accelerazione cardiaca, che sono sintomi di paura a intensità minore del vomito. Attribuendo il valore 1 o

0 alle risposte date a ciascun indicatore, se il risultato di una misurazione è 3, allora l'individuo interrogato avrà manifestato tutti e tre i sintomi; se il risultato è 2, avrà manifestato accelerazione cardiaca e nausea, ma non vomito; se è 1, solo accelerazione cardiaca; mentre se è 0, nessuno dei tre sintomi.

Considerando per semplicità dieci casi, una scala, che rispetti completamente la cumulabilità, mostra la struttura raffigurata nella figura che segue, dove i casi sono ordinati per grado d'intensità riscosso.

Grado	V	N	AC	Casi
3	1	1	1	1
3	1	1	1	2
3	1	1	1	3
2	0	1	1	4
2	0	1	1	5
2	0	1	1	6
2	0	1	1	7
1	0	0	1	8
1	0	0	1	9
0	0	0	0	10

Si osserva nella tavola, detta scalogramma, che tre casi hanno manifestato i tre sintomi (AC , N , V); quattro solo due (AC , N); due un solo sintomo (AC); uno nessun sintomo. La misura dei primi tre casi sarà dunque 3; 2 quella dei quattro casi successivi; 1 per i tre casi del penultimo gruppo; 0 per gli ultimi tre casi. Dal valore attribuito a ciascun caso potremo prevedere senza errore quali saranno i sintomi dallo stesso esperito: si parla a tale proposito di *riproducibilità* della scala, significando la capacità che ha di informarci dei valori dei singoli indicatori componenti il valore finale di ciascuno dei casi considerati.

Errori di cumulabilità e controllo. Ma una scala, quando applicata, difficilmente mostra una completa riproducibilità, cioè piena coerenza con il principio della cumulabilità: per effetto di errore vi saranno casi che non lo rispettano. Una possibile rappresentazione sarà allora la seguente, dove gli I^1 sono tre indicatori qualsiasi:

Grado	I^3	I^2	I^1	Casi
3	1	1	1	1
3	1	1	1	2
2	1	1	0	3
2	1	1	0	4
2	0	1	1	5
2	0	1	1	6
1	1	0	0	7
1	0	1	0	8
1	0	0	1	9
0	0	0	0	10

In questo scalogramma si osserva che i casi 1, 2, 5, 6, 9, 10 non presentano «errori»; tutti i restanti mostrano qualche forma di incoerenza. Ad esempio, il caso 3, invece di mostrare la presenza di I^1 , mostra quella di I^3 ; pertanto, sulla base del punteggio ottenuto dal caso non possiamo riprodurre i sintomi dallo stesso esperiti.

Per valutare la cumulabilità della scala, la prima operazione da fare riguarda il conteggio degli errori (E)¹². Gli errori possono essere contati in due modi differenti: in funzione del numero minimo di spostamenti che occorre fare per dare completa cumulabilità alla struttura (criterio di Cornell), o in funzione del numero delle previsioni errate che possiamo fare rispetto a un caso, dato il punteggio che ha riportato (tecnica di Goodenough). Questo secondo criterio dà valori più alti di E , perché uno spostamento corrisponde sempre ad almeno due previsioni errate; nell'esempio della figura precedente, il caso 3 con il grado 2 di intensità necessita di un solo spostamento per dare piena cumulabilità alle risposte (il valore 1 di I^3 passa a I^1), mentre dà luogo a due previsioni errate (su I^1 che ha valore 0 invece che 1, su I^3 che ha valore 1 invece che 0).

Un importante coefficiente per controllare la cumulabilità è dato dal cosiddetto *coefficiente di riproducibilità* (C_r), che presenta la seguente formula:

$$C_r = 1 - \frac{E}{I \times N}$$

N sta per campione, con un valore pari al numero dei casi considerati; I per indicatori, di valore pari al numero degli indicatori impiegati; E è il numero degli errori che si sono verificati. La formula

rapporta il numero degli errori (E) al numero delle risposte ($I \times N$). Il coefficiente raggiunge il valore 1 nel caso di massima riproducibilità, quando $E = 0$.

Utilizzando il criterio di Cornell nel caso di cui sopra:

$$C_r = 1 - \frac{4}{3 \times 10} = 0,87$$

Secondo questo criterio, che è il più usato, C_r è considerato accettabile quando il suo valore non scende sotto lo 0,90¹³. Nel nostro caso, dunque, la scala dovrebbe essere rigettata. Ma possedendo numerosi indicatori, possiamo diminuire sensibilmente le incoerenze, eliminando quelle domande che non risultano cumulative fino a giungere a una sistemazione soddisfacente. Il risultato è una scala garantita sul piano dell'ordinabilità e dell'unidimensionalità.

Le scale cumulative sono di tipo fondamentale, non presupponendo precedenti misurazioni, e utilizzano il criterio della manifestazione a indicatori multipli per la loro costruzione. Si deve anche notare la correttezza delle operazioni che vengono effettuate. Nell'esempio, il valore «3» attribuito a chi mostra tutti e tre i sintomi della paura, è chiaramente un numerale, perché significa soltanto che mostra il valore più alto, senza alcun riferimento alla distanza tra gli intervalli.

4.4. Scale per somma

Indici di atteggiamento. Queste scale misurano il valore dell'atteggiamento di un individuo verso una proprietà tramite la somma dei punteggi connessi alle risposte che lo stesso ha dato a un insieme di domande (indicatori) relative alla proprietà in questione. Volendo costruire una scala di liberismo, utilizzo un insieme di domande vertenti sulla libera iniziativa, il profitto, la nazionalizzazione ecc.; assegno a ciascuna risposta un determinato punteggio, e calcolo il grado di liberismo posseduto da un individuo sommando i punteggi relativi alle risposte che ha dato.

Nel complesso, la scala è costruita percorrendo cinque fasi:

1. definizione della proprietà da misurare;
2. costituzione di un insieme di indicatori;
3. determinazione del peso da attribuire a ciascuno di essi;
4. controllo degli indicatori e dei pesi attribuiti;
5. applicazione dello strumento.

I punti più critici attengono agli indicatori utilizzati riguardo alla loro: a) validità (essi possono rifarsi a proprietà teoriche diverse); b) eterogeneità sintattica (quando consistono in scale strutturate diversamente: ad esempio, con un numero di intervalli diversi); c) rilevanza (quando hanno «importanza» diversa per la proprietà da misurare).

– L'*unidimensionalità* degli indicatori può essere controllata per contenuto, tramite un accurato esame semantico, o per coerenza (come si fa per le scale di Likert, che tratterò successivamente). La pertinenza teorica, oltre che per contenuto, mediante criterio, utilizzando un secondo strumento o raffrontando i risultati della misurazione con gli atteggiamenti di gruppi noti. Se in un indice di liberismo gli operai risultano tutti con un alto punteggio e gli imprenditori con uno assai basso, potremo sostenere di essere di fronte a una sensazionale scoperta, ma con una verosimiglianza piuttosto bassa: converrà forse rivedere lo strumento.

– L'*eterogeneità sintattica* può essere fronteggiata omogeneizzando gli intervalli di variazione degli indicatori. Si porta a 100 o a 1000 il valore massimo riscontrato per ogni indicatore, quindi si ricalcolano gli altri valori nel rispetto del rapporto definito, oppure si opera la standardizzazione delle variabili¹⁴.

– Per quanto riguarda la *rilevanza* degli indicatori, quando essi sono considerati differentemente rappresentativi della proprietà da misurare, viene dato un «peso» diverso a ciascuno di essi¹⁵.

Infine, la trasformazione della grandezza vettoriale (l'insieme degli indicatori) in una scalare lascia piuttosto perplessi, perché l'operazione effettua calcoli numerici (di somma) sui valori degli indicatori, che non sono numerici, ma numerali, per ottenere una scala che in definitiva è ordinale. La costituzione di grandezze ordinali (o numeriche) per mezzo di calcoli pseudo-numerici è una procedura che in ogni caso non può dirsi corretta senza l'esplicitazione e l'accettazione di specifici assunti.

Le scale per somma sono di maggiore affidabilità di quelle di valutazione (che però in una qualche misura presuppongono quando i soggetti assegnano valori all'indicatore), perché frutto di una costruzione più pubblica e controllabile. Sono esplicitati i criteri con cui si assegnano i punteggi (tramite gli indicatori) e il metodo con cui viene calcolato il valore finale; la molteplicità degli indicatori permette un giudizio articolato, meno soggetto a errore sistematico.

Scale di Likert. Le scale di Likert migliorano sensibilmente gli indici precedenti nella direzione di una più accurata selezione degli in-

dicatori della proprietà da misurare¹⁶. La loro specificità consiste nell'uso di un insieme di domande a risposta graduata, somministrate a un gruppo di giudici che, tramite le loro risposte, ne permettono la selezione. Ma anche in queste scale resta arbitraria l'assegnazione di valori numerici alle risposte; il criterio di rappresentazione continua a utilizzare operazioni pseudo-numeriche per addivenire a una misurazione ordinale.

Facciamo un esempio. Scelta una proprietà da misurare, supponiamo il clericalismo, viene creato un insieme abbastanza vasto di domande con lo stimolo anticlericale («Alcuni dicono che la Chiesa non dovrebbe possedere alcun bene terreno») e un altro altrettanto vasto di domande con lo stimolo filoclericale («Secondo alcuni, i cittadini dovrebbero seguire i suggerimenti dei sacerdoti riguardo ai loro comportamenti politici»). La formulazione delle domande con stimoli in direzione opposta viene effettuata per impedire il verificarsi di *response set*¹⁷ e, in ogni caso, per attivare l'attenzione del rispondente.

A ciascun giudice viene poi richiesta una valutazione in sette o, più comunemente, in cinque gradi del tipo:

- completamente d'accordo;
- abbastanza d'accordo;
- indifferente;
- abbastanza in disaccordo;
- completamente in disaccordo.

Le graduazioni delle affermazioni vengono trasformate in punteggi, da 1 a 5, in modo che riscuota 5 sia chi si è detto «completamente d'accordo» con un'affermazione filoclericale, sia chi si è detto «completamente in disaccordo» con un'affermazione anticlericale, riscuotendo, nel caso, complessivamente 10 punti per le due domande che esprimono la dimensione in senso opposto¹⁸.

Effettuata la somministrazione e rilevate le risposte, controlliamo l'omogeneità degli indicatori utilizzati tramite la loro *congruenza interna* nel modo seguente. Calcolato il punteggio riscosso da ogni giudice, si formano due gruppi che mostrino rispettivamente il massimo accordo e il massimo disaccordo con la dimensione prescelta (nell'esempio, i più e i meno clericali); si esamina allora il potere discriminante di ciascuna domanda. Se una domanda è apprezzata nello stesso modo dai due gruppi, viene scartata perché supposta porsi in una dimensione diversa da quella che si intende misurare. Questo tipo di controllo ha generato due importanti famiglie di tecniche: quelle che fanno riferimento all'alfa di Cronbach e quelle basate sull'analisi fattoriale¹⁹.

Le domande che superano il vaglio vengono a costituire lo strumento di misurazione che è applicato secondo le usuali modalità degli indici per somma.

5. Matematiche e scienze sociali

La matematica ha dato luogo a lunghi dibattiti, quasi tutti però basati sul presupposto che la matematica sia solo quella numerica; quindi, che la misurazione si riduca alla costruzione e all'applicazione di strumenti per il calcolo numerico. Ma la matematica è un insieme più vasto di tale calcolo che, se è la parte (finora) storicamente più importante della matematica, non è l'unica. Non vi è alcuna ragione per respingere dal regno della matematica i calcoli a due valori e ammettere solo quelli a più valori; rifiutare le operazioni della logica o dell'insiemistica e includervi solo quelle basate sulle strutture numeriche. Anzi, tale esclusione comporterebbe contraddizioni non risolvibili.

Se dunque tutti questi calcoli appartengono alla matematica, non esiste un mondo qualitativo non matematico e uno quantitativo matematico; esistono, semmai, mondi matematici qualitativi e quantitativi: variabili e relazioni formalizzabili con assiomi e teoremi diversi. Il problema allora non è quello dell'applicabilità della matematica, ma quello della *sua buona applicazione*, cioè della costruzione corretta degli strumenti di misurazione, e della pertinenza delle teorie matematiche da utilizzare per la descrizione dei fenomeni sociali. Dobbiamo trovare funzioni con cui la rappresentazione possa essere effettuata in modo corretto, cioè tale da «anticipare» linguisticamente ciò che può avvenire empiricamente nell'universo.

Anche sul significato della misurazione gravano parecchi malintesi. Al di là dell'infelicità del termine, la misurazione non è che un sistema di definizioni dei singoli stati di una proprietà, coerente con i rapporti che corrono tra essi. I valori formali indicano la presenza di tali stati. Alcuni sostengono che la misurazione inizia con le scale numeriche o, al massimo, ordinali; mai con quelle nominali. L'affermazione è condivisibile se per misurazione si intende la *costruzione* dello strumento di misurazione che, almeno per le scale nominali, non richiede alcunché di specifico. Ma non lo è se con essa s'intende che i predicati nominali non sono scale, perché consistono sempre nella rappresentazione di una struttura empirica (a due stati) su una struttura formale (a due valori), logica invece che numerica.

In questa ottica, alcune critiche alla «matematica» sono certamente accettabili: molte misurazioni sono *per fiat*; la matematica in uso è quella fondamentalmente importata dalle scienze naturali, molte volte estranea alla problematica sociale. Gli scienziati sociali, quando lo hanno fatto, si sono occupati più dell'applicazione della matematica *esistente* che della *creazione* di una nuova, più capace a risolvere i loro problemi²⁰. Del resto, difficilmente poteva accadere qualcosa di diverso, considerata la scarsa formazione matematica che ha il ricercatore sociale, che è, al massimo, di tipo statistico.

6. Le relazioni di dipendenza

L'uso della matematica non si esaurisce nella misurazione delle proprietà nei casi, si estende alla scelta delle funzioni che legano il variare degli stati di una proprietà a quelli di un'altra, tramite relazioni (enunciati) di dipendenza. Il livello di misurazione delle variabili restringe solo l'area dei teoremi utilizzabili, perché, a seconda del loro livello, le variabili sono suscettibili di connessioni di natura sintattica diversa. In questa area di fattibilità si opera la scelta della funzione che deve collegare le variabili in modo da descrivere adeguatamente il fenomeno che stiamo studiando.