

# 1. Indagine statistica

## 1.1 Il disegno

### 1.1.1 Caratteristiche del disegno

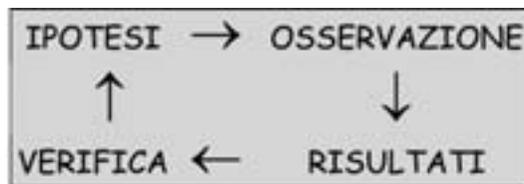
Il disegno di indagine rappresenta la struttura che definisce il progetto di indagine e le procedure necessarie per la sua realizzazione. Il disegno richiede la definizione:

1. del modello di indagine,
2. delle aree e delle variabili di indagine,
3. delle procedure di misurazione,
4. dello strumento per la raccolta dei dati,
5. del campione (procedimento di estrazione e numerosità),
6. del piano per l'analisi dei dati.

#### 1.1.1.1 Disegno e ipotesi

Ciò che caratterizza il disegno è principalmente il tipo di ipotesi su cui è basata l'indagine stessa.

Il concetto di ipotesi ha trovato molte definizioni lungo la storia del pensiero filosofico e scientifico: secondo una di queste, l'ipotesi è un'affermazione relativa ad un evento il cui risultato è sconosciuto al momento in cui l'affermazione viene fatta. Essa può essere vista come «un'opera creativa e avventurosa della ragione». L'ipotesi si condensa in quegli enunciati sperimentabili (se riducibili a elementi di osservazione) e verificabili. In quest'ottica, l'ipotesi può essere definita come l'insieme di supposizioni e congetture che mirano a spiegare un certo fenomeno di cui si ha una conoscenza incompleta. Nella pratica essa si limita spesso a un'enumerazione dei fatti osservabili direttamente o indirettamente. L'ipotesi di ricerca può essere già formulata e quindi solo da verificare oppure può essere dedotta da altri risultati che conducono alla formulazione di nuove congetture da verificare e inserire in una nuova formulazione di ipotesi e verifica. Tale processo circolare può essere rappresentato graficamente (figura I. 1.1).



*Fig. I. 1.1 Il processo circolare della ricerca*

L'ipotesi deve essere formulata in modo che possa risultare non vera e quindi respinta (o viceversa) e quindi che possa essere ugualmente *confutata* o *confermata* attraverso – la *sperimentazione*, ovvero la possibilità di osservare, controllare, riprodurre e condurre (attraverso processi matematici) l'analisi dei dati;

– il *modello statistico-matematico*: la risposta all'ipotesi non è mai assoluta per ragioni legate alle possibilità e al caso; ne consegue la necessità di soluzioni *statistiche* per la stima dei parametri e/o delle costanti. Ciò vuol dire ricorrere, per l'analisi dei dati, a sistemi di generalizzazione e misure d'interconnessione di tipo probabilistico.

In particolare nella formulazione di un'ipotesi occorre definire e occorre inserire

- il problema e gli aspetti da rilevare,
- la validazione teorica,
- gli elementi di base per l'attuazione dell'indagine. Infatti per giungere ad una decisione riguardo all'accettabilità di una ipotesi definita<sup>1</sup> si procede alla rilevazione dei dati empirici.

Nella fase successiva è necessario effettuare una rassegna della letteratura esistente per verificare l'esistenza di altri studi riguardanti il problema e di strumenti di rilevazione già validati. Segue la definizione delle ipotesi di interconnessioni.

In sintesi quindi l'ipotesi di ricerca può svolgersi in un piano strutturato riguardante tutte le fasi dello studio: dall'idea del perché fare una certa ricerca all'oggetto dello studio, dalla delimitazione e limitazione dei dati alla loro rilevazione, dall'analisi dei dati all'interpretazione dei risultati.

### 1.1.1.2 Classificazione dei disegni

I disegni possono essere classificati principalmente in due tipi (figura I. 1.2):

- disegno esplorativo
- disegno conclusivo, a sua volta distinguibile in:
  - *disegno descrittivo*<sup>2</sup>: il cui scopo è, in termini forse riduttivi ma essenziali, quello di *contare*; si effettuano su popolazioni o su campioni rappresentativi (che consentono di inferire e di estendere i risultati a tutta la popolazione); spesso vengono effettuate per confrontare i risultati con altri studi fatti in altre aree o in altri momenti (per esempio per definire dei *trend*);
  - *disegno causale (sperimentale)*, il cui obiettivo è essenzialmente quello di evidenziare le relazioni causali; in particolare con questo disegno si cerca di capire

<sup>1</sup> La verifica delle ipotesi costituisce uno dei momenti più importanti all'interno del metodo sperimentale, nato e teorizzato come metodo generale di indagine nel campo delle scienze naturali. Come si è visto, il metodo sperimentale si basa su un ragionamento ipotetico-deduttivo (confronto tra esperienza e induzione) che consente di giungere all'enunciazione di una legge generale.

<sup>2</sup> Nell'ambito dell'approccio descrittivo è possibile fare un'ulteriore distinzione tra criterio – analitico-relazionale; in questo caso l'indagine è pianificata allo scopo di esplorare le relazioni esistenti tra determinate variabili; l'indagine è meno orientata verso lo studio della rappresentatività e della descrizione e più verso l'analisi delle associazioni tra fenomeni e delle loro possibili predizioni; – operativo; in questo caso l'indagine è effettuata su popolazioni o gruppi definiti e non su campioni; produce risultati che consentono di effettuare precise scelte operative nell'ambito indagato e nei riguardi della popolazione oggetto di indagine.

quali variabili sono la causa (variabili indipendenti) e quali l'effetto (variabili dipendenti) e di determinare la natura della relazione tra i due tipi di variabili. Nella figura I. 1.3 sono sintetizzate le principali caratteristiche dei disegni.



Fig. I. 1.2 Tipologie dei disegni di ricerca

		Disegno			
		Esplorativo	Conclusivo		
			Descrittivo	Causale/sperimentale	
Obiettivi		<ul style="list-style-type: none"> <li>• fornire intuizioni e comprensioni</li> <li>• consentire la formulazione e la definizione più precisa di un problema</li> <li>• identificare percorsi alternativi di una determinata azione</li> <li>• sviluppare ipotesi</li> <li>• isolare variabili e relazioni per ulteriori studi</li> <li>• stabilire priorità per ulteriori studi</li> </ul>	<ul style="list-style-type: none"> <li>• formulare e verificare ipotesi ed esaminare relazioni</li> <li>• aiutare chi prende decisioni a valutare, determinare e selezionare il migliore percorso da fare in una determinata situazione.</li> </ul>		
Caratteristiche	procedimento	flexibile e versatile	pre-pianificato e strutturato e formalmente impostato.		
	aree di indagine	definizione approssimativa	definizione chiara		
	campione	dimensione ridotta e scarsa rappresentatività	dimensione ampia e rappresentatività statistica.		
	analisi dei dati	essenzialmente qualitativa	essenzialmente quantitativa		
Tipo di risultati		spesso il disegno esplorativo rappresenta la parte iniziale di un modello complesso che prosegue con un nuovo disegno esplorativo o conclusivo	risultati possono essere utilizzati come informazioni nel processo di decisione.		
				<ul style="list-style-type: none"> <li>• manipolazione di una o più variabili indipendenti</li> <li>• controllo di altre variabili</li> </ul>	

Fig. I. 1.3 Principali caratteristiche dei diversi disegni di ricerca

Anche se definiti separatamente, i disegni non sono nettamente distinti tra loro; per rispondere alle esigenze conoscitive, il progetto di uno studio può prevedere l'ado-

zione di più disegni. Infatti, non sempre la scelta di un singolo approccio rappresenta il miglior modo di procedere; tutto dipende da ciò che si intende indagare. Molti studi utilizzano due o più disegni in fasi diverse. All'interno di un ampio studio, il disegno esplorativo può rappresentare la fase iniziale che consente di giungere alla definizione dell'ipotesi, da verificare in seguito attraverso un disegno descrittivo o causale.

Se la fase di formulazione e di concettualizzazione del problema è ad uno stadio iniziale, l'approccio esplorativo può essere considerato il migliore; esso consente di elaborare e sviluppare concetti e ipotesi e di definire le variabili fondamentali, da indagare in seguito in uno studio sperimentale, la cui pianificazione richiede una formulazione molto precisa di ciò che si vuole studiare.

### 1.1.1.3 Affinità tra disegni

Pur differenziandosi nettamente, i disegni presentano alcuni aspetti comuni.

- a. Essi sono alla ricerca di *regolarità*, di modelli che aiutino a comprendere le interconnessioni tra gli elementi osservati e tra le variabili osservate. Differente è l'evidenza indagata dai diversi disegni.
- b. Essi esaminano unità *rappresentative* di un numero più vasto di unità. Si vedrà come la rappresentatività assume diverse definizioni e connotazioni.
- c. Essi sviluppano *concetti*, o adottano concetti esistenti, per esprimere le regolarità notate tra le unità osservate, considerate rappresentative.
- d. Essi sono finalizzati allo studio di fenomeni che variano o soggetti a variazione, ovvero studiano *variabili* e prevedono, definiscono e individuano *costanti* ossia quelle caratteristiche comuni che tendono a non variare.
- e. Essi consentono di costruire e verificare una *teoria*, una spiegazione di un insieme di relazioni.
- f. Ciascun disegno prevede e richiede una *pianificazione* che comprende anche strategie e ipotesi sui risultati attesi.

### 1.1.2 Il disegno descrittivo

Come indica il nome, il principale obiettivo di questo disegno è quello di descrivere una particolare realtà o situazione. Al contrario di quello esplorativo, il disegno descrittivo richiede una chiara definizione del problema e, successivamente, dell'ipotesi e una dettagliata descrizione delle procedure<sup>3</sup>.

<sup>3</sup> A tale proposito la letteratura anglo-sassone individua nelle cosiddette sei *W* gli elementi che devono essere precisati nell'ambito del progetto:

- *who*: chi è il committente,
- *why*: perché condurre lo studio,
- *what*: quali sono le informazioni che si vogliono ottenere,
- *when*: quando svolgere l'indagine,
- *where*: dove realizzare l'indagine,
- *way*: in che modo ottenere le informazioni.

### 1.1.2.1 Disegni trasversali e longitudinali

I disegni possono essere classificati rispetto ai tempi di rilevazione; in tale ottica possiamo distinguere i seguenti tipi:

- **Disegno trasversale** (o *cross-sectional*): l'obiettivo di questo approccio è quello di rilevare le stesse informazioni in un dato momento per poter valutare le caratteristiche di un particolare gruppo in un particolare momento. Per studiare e comprendere il fenomeno il gruppo viene successivamente analizzato suddividendolo sulla base di caratteri (variabili) indipendenti e confrontando i sottogruppi. In questo tipo di indagine è necessario pianificare in anticipo i confronti e renderli parte integrante del progetto. Il disegno trasversale può essere a sua volta distinto in
  - singolo, quando coinvolge un unico gruppo (campione),
  - multiplo, quando coinvolge più campioni.

Un particolare tipo di disegno trasversale multiplo è quello detto *analisi di coorte* (*cohort analysis*) basato su una serie di rilevazioni condotte a intervalli definiti di tempo; la coorte, utilizzata come unità base di analisi, rappresenta un gruppo di soggetti che vive lo stesso evento nello stesso intervallo di tempo. Per esempio, una coorte di età può essere composta da soggetti nati durante lo stesso intervallo di tempo.

Gli studi trasversali forniscono solo una evidenza indiretta degli effetti del tempo e devono essere utilizzati con grande cautela quando si vogliono fare deduzioni riguardanti cambiamenti.

- **Disegni ripetuti**: tali rilevazioni sono dette anche periodiche o ricorrenti, in quanto ripetute in modo programmato nel tempo. Il gruppo osservato può essere:
  - parzialmente sovrapposto a quello osservato nella precedente rilevazione,
  - senza alcuna sovrapposizione con quello osservato nella precedente rilevazione.
 L'obiettivo è quello di osservare cambiamenti a livello di gruppo e non a livello individuale.
- **Disegno longitudinale**: al contrario delle indagini trasversali, attraverso le quali non è possibile determinare chiaramente le relazioni di causa-effetto, ma solo informazioni sulle correlazioni, i disegni longitudinali sono finalizzati a seguire un particolare gruppo nel tempo. In generale le indagini *longitudinali* rilevano una serie di osservazioni per le stesse variabili sugli stessi soggetti in momenti diversi. Tali disegni possono essere utilizzati anche nell'ambito sperimentale dove spesso assumono la particolare connotazione di verifica *prima-dopo*: un gruppo di casi viene sottoposto alle influenze della variabile sperimentale; il confronto tra i valori registrati prima e dopo nelle variabili indipendenti consente di valutare gli effetti dell'esperimento.

I disegni longitudinali possono essere:

- senza rotazione, che hanno l'obiettivo di seguire un particolare gruppo nelle modifiche che possono intervenire a livello individuale; i risultati possono essere riferiti solo alla popolazione di partenza; questo vuol dire che se nel frattempo la popolazione di riferimento si è modificata, tali studi non consentono di estendere i risultati;

- con rotazione, che, a differenza degli studi del precedente tipo, prevedono l'introduzione periodica di nuovi casi in modo da mantenere rappresentativo il gruppo rispetto alla popolazione di riferimento; ciò consente di fare valutazioni sia longitudinali che trasversali.

Un particolare tipo di studio longitudinale è quello detto *panel*, non limitato a due sole rilevazioni ma rappresentato da più osservazioni per le stesse variabili sugli stessi casi in diversi periodi di tempo. Tali disegni consentono di studiare e valutare l'impatto di un certo fenomeno lungo il tempo e in particolare di analizzare le influenze e i processi di cambiamento e di adattamento o di studiare il cosiddetto *turnover* ovvero gli spostamenti tra misurazioni per ciascun soggetto. Il modello *panel*, utilizzato soprattutto in medicina ed epidemiologia in molti studi detti di *follow-up*, può essere distinto in:

- studio prospettico, quando si segue per un certo periodo una selezionata popolazione "a rischio" per scoprire le possibili cause di una certa malattia confrontando i soggetti nei quali si sviluppa con quelli nei quali non si sviluppa;
- studio retrospettivo, quando si osservano i soggetti che hanno già sofferto di una malattia o una patologia per cercare i fattori antecedenti collegati con essa.

Gli studi *panel* presentano particolari problemi legati

- alla loro durata, che può produrre, nei soggetti studiati, cambiamenti non direttamente collegati al fenomeno studiato,
  - agli elevati costi, sia in termini economici che di tempo,
  - alle difficoltà di individuazione delle unità da seguire (perdite per mortalità, spostamenti territoriali, ecc.).
- **Disegno sociometrico:** si tratta di un particolare disegno applicabile nel campo della ricerca sociale che consente di ottenere una rappresentazione dei legami tra i componenti di un gruppo. L'ideazione di tale tecnica è dovuta a J.L. Moreno che la sviluppò negli anni '50<sup>4</sup>. Tale approccio può essere utilizzato solo su gruppi chiusi, anche se alla fine si giunge ad una descrizione individuale dei soggetti appartenenti al gruppo studiato. Per poter applicare correttamente tale modello è necessario che il gruppo risulti chiuso, ovvero che i legami studiati avvengano al suo interno:
- ciascun componente fa o riceve una scelta da un altro componente,
  - ciascun componente sceglie almeno uno dei componenti del gruppo.
- Non sono permesse scelte al di fuori del gruppo.

### 1.1.3 Il disegno sperimentale

Se da una parte il disegno descrittivo consente di determinare il grado di associazione tra variabili, dall'altra esso non è appropriato per esaminare le relazioni causali. Per soddisfare tale esigenza si richiede il disegno causale/sperimentale. Il

<sup>4</sup>La prima applicazione di tale tecnica riguardava lo studio del livello di comportamento cooperativo in una comunità di adolescenti.

disegno sperimentale è detto anche causale in quanto consente di evidenziare le relazioni di tipo causa-effetto rispetto ad un certo fenomeno e di determinare la natura della relazione tra variabili causali e l'effetto da predire. Il concetto di causalità, fondamentale nel disegno sperimentale, non è semplice da definire scientificamente ed è difficile da tradurre in termini operativi, anche se intrinseco allo stesso pensiero scientifico, nell'ambito del quale è stato messo spesso in discussione da filosofi ed epistemologi.

Qualsiasi disegno sperimentale è in grado non tanto di provare e di dimostrare in modo conclusivo la causalità quanto di inferire in termini probabilistici la relazione di causa-effetto.

Per poter assumere la causalità occorre verificare se è possibile soddisfare una o più tra le seguenti condizioni necessarie, che però non risultano sufficienti a dimostrare la causalità:

- a. *Variazione concomitante*: causa ( $X$ ) ed effetto ( $Y$ ) si verificano insieme o variano insieme così come previsto dall'ipotesi considerata. L'assenza di un'evidenza iniziale di variazione concomitante non implica la mancanza di causazione. In certi casi è possibile che l'intervento di una terza variabile possa mettere in evidenza un'associazione, precedentemente non visibile.
- b. *Successione temporale delle variabili*: l'evento 'causa' deve verificarsi o prima o in concomitanza dell'evento 'effetto' mai dopo. Per definizione, un effetto non può essere prodotto da un evento che si verifica successivamente ad esso. È comunque possibile che ciascun evento all'interno di una stessa relazione causale sia allo stesso tempo causa ed effetto.
- c. *Depurazione da altri possibili fattori causali*: è necessario che le variabili investigate siano le sole possibili spiegazioni causali. Nell'analizzare un evento dopo che si è verificato, non si è mai sicuri di essere riusciti ad escludere altri eventuali fattori. In molti casi è possibile bilanciare gli effetti di fattori esterni al modello.

L'evidenza di tali condizioni (considerate singolarmente o in combinazione) non consente di dimostrare in modo conclusivo che una relazione causale esiste; se l'evidenza totale è forte e consistente, è possibile ragionevolmente concludere che la relazione causale esiste. L'evidenza accumulata da più sperimentazioni, confortata dalla teoria, consente di incrementare la loro confidenza e significatività. L'adozione del metodo sperimentale in condizioni controllate consente di raggiungere una forte evidenza rispetto alle tre condizioni.

## 1.2 Complessità dell'indagine in ambito sociale

Nel disegnare un'indagine statistica in ambito sociale è necessario tenere presente che, a differenza della ricerca applicata in altre scienze, la ricerca sociale presenta una complessità che deriva principalmente dall'esigenza e dalla difficoltà di:

- isolare l'ambito sperimentale senza alterare le osservazioni,
- realizzare situazioni più vicine possibile all'osservazione pura,
- controllare gli elementi coinvolti,
- estendere i risultati,
- disporre di un certo tipo di risorse economiche, di persone, di tempo.

Ciò non deve condurre alla rinuncia all'indagine, o alla superficialità nella formulazione delle ipotesi, seguendo l'affermazione secondo la quale non è possibile fare ricerca scientifica come nel campo della fisica, ma è possibile condurre solamente ricerche approssimate.

Si può provare ad identificare gli elementi che rendono la ricerca sociale più complessa e difficile rispetto ad altre ricerche (Tesi, 1983):

- Complessità della struttura della ricerca, dovuta all'alto numero di variabili da considerare e alla loro appartenenza a discipline diverse quali discipline biologiche, socio-antropologiche, psicologiche. Per questa complessità la ricerca deve essere necessariamente interdisciplinare per approccio e tecnologia.
- Diversità dei livelli sociali: l'individuo oggetto della ricerca viene ad essere mobilitato in vari aspetti:
  - individuo con i suoi diversi livelli di personalità;
  - membro di un gruppo ristretto: famiglia, gruppo amicale;
  - membro di collettività: partiti, chiese;
  - presente in sistemi olistici: comportamento di massa, storico, ecc.
- Livello di personalità: le domande poste durante l'intervista possono agire sulla personalità a diversi livelli, proiettivo, di temperamento, di interessi/valori, di attitudini sociali, di disposizioni motivazionali, espressivo/stilistico/intellettuale, patologico<sup>5</sup>.
- Memoria: l'individuo vive e ricorda ogni sua esperienza, per cui non si può porre due volte la stessa domanda presupponendo di trovare il soggetto nella stessa identica situazione.
- Esperienza individuale: gli intervistati si possono trovare, nei riguardi delle situazioni prospettate nelle domande, in situazioni diverse per esperienza e maturità di giudizio. Riguardo all'esperienza, l'intervistato può:
  - aver vissuto la situazione per esperienza diretta,
  - aver avuto una esperienza indiretta e da lui vissuta con più o meno partecipazione,
  - averne una conoscenza generale,
  - non averne alcuna conoscenza.

La posizione individuale deve essere chiarita mediante domande aggiuntive. Riguardo alla maturità di giudizio l'intervistato può non aver maturato alcun giudizio o averlo maturato in precedenza. Nei confronti di questi punti non è possibile definire la posizione individuale.

<sup>5</sup> Questo livello non può essere conosciuto se non vengono condotte più indagini parallele con tecniche diverse, che non possono essere validate per l'interazione fra gli stimoli contenuti nei diversi strumenti.

- Osservazione e sperimentazione: l'individuo interagisce con lo sperimentatore per cui non abbiamo mai, negli studi su organismi viventi, una situazione di pura osservazione, ma una condizione intermedia indeterminata tra osservazione (rilevazione non interagente) e sperimentazione (risposta provocata da stimolazione). Questa interazione può agire in una determinata direzione, per esempio le persone si possono sentire lusingate di essere state prescelte e per questo possono mutare il loro comportamento (*effetto Hawthorne*<sup>6</sup>). L'individuo osservato può presentare e manifestare altre tendenze:
  - la *desiderabilità sociale*, ovvero la tendenza di alcuni soggetti a rispondere in un modo socialmente desiderabile o in una direzione accettabile, anziché riferire la vera risposta o il vero atteggiamento;
  - il *response set*, ovvero la tendenza di alcuni soggetti a rispondere in modo sistematico per esempio ad essere d'accordo con tutti gli indicatori indipendentemente dal contenuto (*acquiescence response set*).
- Indeterminazione: come conseguenza dell'impossibilità di condurre ricerche in situazioni di osservazione che non interagiscano con l'osservato, si presenta la situazione sperimentale di *indeterminazione* dell'effetto dell'osservazione-stimolo sulla risposta, ossia l'osservato è modificato in misura non conosciuta per effetto dello stimolo stesso (domanda) necessario per la rilevazione. Per analogia con il principio di Heisenberg *ogni misurazione comporta uno scambio d'energia fra oggetto misurato e strumento di rilevazione*, trasferendosi nel campo dell'indagine sociale si potrebbe affermare che *ogni rilevazione comporta un passaggio d'informazione tra intervistatore e intervistato*. Ciò significa che un'informazione contenuta nella domanda modifica in maniera non valutabile la risposta spontanea. Questa situazione è ancora più grave dal punto di vista sperimentale quando alcune domande inducono determinate risposte, ossia la previsione di una situazione prospettata in una certa direzione, influenza la risposta nella stessa direzione: *effetto Edipo*.
- Contenuto pedagogico: l'inchiesta può assumere il significato di strumento pedagogico utile per diffondere alcuni concetti, prospettando situazioni e soluzioni possibili. Le domande, e l'elencazione delle possibili alternative previste nelle risposte, possono essere presentate in modo da indurre l'intervistato a considerare una situazione in uno schema fissato dal ricercatore. In questi casi non si tratterà di una ricerca conoscitiva, ma di una vera e propria azione pedagogica rispetto alla quale la ricerca può essere estranea.
- Olismo: alcune componenti individuali e/o sociali riunite possono assumere una dimensione nuova che supera la somma delle sue parti, fino a raggiungere un significato nuovo e indipendente dagli elementi che può persistere anche quando

<sup>6</sup> Presso la *Hawthorne Electrical Company* fu condotto uno studio sulla produttività; il gruppo sotto studio migliorò il rendimento non solo quando furono migliorate le condizioni di lavoro, ma anche quando furono ristabilite le condizioni iniziali, così il miglioramento fu attribuito alla partecipazione gratificante all'esperienza.

alcuni di questi, o tutti, non sono più presenti, per cui la rilevazione dei singoli elementi può non essere sufficiente per analizzare o comprendere situazioni e fenomeni complessi.

- **Moralità e induzione:** la posizione gerarchica che il ricercatore può assumere rispetto all'intervistato, deve essere tenuta presente specialmente per le domande che prospettano situazioni che possono turbare l'intervistato stesso.

Tutti questi elementi vanno tenuti presenti sia nel disegno generale della ricerca sia nell'analisi dei risultati che ne scaturiscono. Essi sono sempre presenti in misura diversa da soggetto a soggetto, ed è necessario essere coscienti di questi limiti che riducono la validità della rilevazione nelle scienze sociali, perché le risposte, una volta raccolte, verranno riunite insieme e considerate come omogenee nelle analisi, anche se questa omogeneità non potrà essere provata.

Rispetto a tali problemi il dibattito metodologico è ancora aperto.

### 1.3 Le fasi

Cerchiamo a questo punto di riassumere le fasi che definiscono il processo di indagine; esse potrebbero essere viste come momenti che vanno a comporre una vera e propria *catena di produzione* (figura I. 1.4), all'interno della quale le diverse operazioni sono in continua comunicazione tra loro (sia con quelle precedenti che quelle successive).

#### 1ª FASE

Rappresenta una delle fasi più delicate di tutto il lavoro d'indagine. I momenti più importanti di questa fase sono:

- **Ipotesi di ricerca:** definizione degli obiettivi e dei percorsi da seguire; tale momento comporta
  - l'esame del fenomeno e la definizione degli scopi dell'indagine (perché si fa ricerca e su cosa si indaga),
  - la raccolta e l'esame delle informazioni disponibili che riguardano il fenomeno e la rassegna delle letterature,
  - la definizione degli aspetti e dei caratteri da rilevare,
  - la definizione delle variabili che rappresentano e definiscono il problema (cosa rilevare),
  - la definizione della conoscenza di base della natura delle variabili,
  - l'identificazione delle variabili indipendenti (in quali termini osservare),
  - la validazione teorica,
  - la definizione dei criteri di verifica (come valutare),
  - l'identificazione dei metodi di analisi.

Tutto ciò consente, come abbiamo visto in precedenza, di identificare il disegno di ricerca che meglio risponde all'ipotesi.

- **Definizione, scelta e costruzione dello strumento di rilevazione**: ovvero di un sistema di designazione, rappresentazione e classificazione degli aspetti (definizione di scale e metodi di misurazione, strumenti di rilevazione, ecc.); esso deve essere:
  - a. **affidabile** ovvero *stabile/ripetibile* (lo stesso osservatore che ripete la medesima osservazione/misurazione nelle medesime condizioni deve produrre gli stessi risultati) e *riproducibile* (utilizzabile da più ricercatori in situazioni analoghe con risultati confrontabili);
  - b. **valido** (ovvero rappresentativo, esclusivo ed esaustivo).
 La costruzione dello strumento prevede anche la definizione
  - del sistema di classificazione,
  - dell'apparato di rilevazione,
  - dell'apparato di codifica e di registrazione coerenti con il sistema di classificazione.
 La scelta dello strumento di rilevazione è condizionata oltre che dagli scopi (ipotesi) di ricerca, anche da altri fattori come:
  - la modalità di rilevazione (individuale, collettiva, per telefono, per posta, con intervistatore, senza intervistatore, ecc.),
  - il livello culturale degli intervistati e la complessità degli argomenti da rilevare,
  - il numero dei soggetti da raggiungere,
  - la variabilità delle risposte e la possibilità di prevederle.
 Ovviamente in una ricerca occorre adottare la stessa tecnica di somministrazione per tutti i soggetti coinvolti nell'indagine.
  
- **Campionamento** che riguarda
  - la definizione dell'oggetto da osservare (chi indagare: popolazione di riferimento, unità di rilevazione),
  - l'identificazione dell'ambito spazio-temporale cui i risultati vanno riferiti.

## 2ª FASE

Questa fase presenta i seguenti momenti:

- **Pianificazione della ricerca sul campo - osservazione del fenomeno e raccolta dei dati** attraverso la misurazione delle grandezze definite in precedenza. La raccolta dei dati, in ambiente *naturale* o in laboratorio, deve rispondere a criteri d'accuratezza e uniformità. Nel caso di ricerche sociali la rilevazione può essere preceduta da incontri preparatori con i rilevatori (se utilizzati) per chiarire le modalità di svolgimento della rilevazione con la presentazione dei problemi più ricorrenti.
  
- **Verifica**: dopo la raccolta, i dati devono essere sottoposti ad una serie di verifiche finalizzate al controllo della completezza delle informazioni e conseguente integrazione delle informazioni recuperabili: se tali verifiche vengono effettuate con particolari accorgimenti, consentono di limitare in parte l'influenza degli errori non campionari (errori di misurazione: parziale o totale mancanza di risposte, risposte non pertinenti o non comprensibili).

– **Acquisizione e trattamento dei dati:**

- codifica dei dati rilevati, ovvero trasformazione delle informazioni in simboli;
- l'eventuale analisi e post-codifica delle domande aperte;
- la registrazione, ovvero il trasferimento dei dati codificati su supporti magnetici (registrazione). Tale operazione deve consentire l'integrazione a livello informatico
  - con altri sistemi, ovvero la possibilità di trasferire i dati per esempio tra personal computer e *mainframe*<sup>7</sup>;
  - con altri programmi (notare che uno dei passaggi più dispendiosi dell'analisi statistica automatica dal punto di vista del tempo è rappresentato dal trasferimento dei dati da un programma all'altro in quanto spesso si basa su procedure macchinose).

La registrazione dei dati può essere controllata attraverso, per esempio:

- a. verifica della registrazione;
- b. controllo sull'intervallo delle risposte ammissibili;
- c. impostazione di condizioni logiche sui dati in registrazione.

### 3ª FASE

La fase successiva consiste nella elaborazione dei dati. L'elaborazione automatica tramite computer presenta sicuramente notevoli vantaggi ed ha reso tale fase affidabile, in quanto è stata aumentata la precisione di analisi<sup>8</sup>, ed economica e veloce, in quanto sono stati ridotti i costi in termini di tempo (elaborazione veloce) e di risorse umane (sono impiegate poche persone anche se altamente specializzate).

L'analisi dei dati non per questo è divenuta meno complessa. Le fasi dell'elaborazione possono essere così riassunte:

- lettura dei dati,
- revisione e ricodifica,
- gestione e manipolazione.

La maggior parte dei programmi statistici (*package* statistici) consente di effettuare praticamente tutte le funzioni di gestione e di preparazione dei dati per l'analisi (ricodifica, calcolo di variabili, gestione di valori *missing* e *outlier*, funzioni di *merge* e di ricostruzione di file, ecc.) in modo più semplice per l'utente. Per l'elaborazione dei dati è comunque utile che il ricercatore (o chi per lui) predisponga un programma in un linguaggio comprensibile alla macchina, che contenga tutte le istruzioni per poter eseguire correttamente le procedure di elaborazione.

<sup>7</sup> Si tratta, sinteticamente, di una configurazione che consente, a più utenti, di utilizzare lo stesso elaboratore e di poter eseguire contemporaneamente più programmi (*multi task-multi user*). Si tratta di una configurazione che necessita di calcolatori molto potenti e di *software* (sistemi operativi) dalla logica complessa.

<sup>8</sup> A tale proposito occorre comunque notare che la presenza di un errore comporta la sua propagazione, proprio a causa dell'utilizzazione di sistemi automatici.

#### 4<sup>a</sup> FASE

– **Applicazione di procedure di verifica di affidabilità e validità degli strumenti utilizzati**

– **Analisi statistica**

La ricerca che si avvale del supporto statistico è rivolta principalmente alla generalizzazione, classificazione e spiegazione di una molteplicità di osservazioni, condotte in un ambito definito e delimitato, sulla base di un'ipotesi formulata in precedenza. Le conclusioni di una ricerca possono essere estese a un ambito spazio-temporale più ampio se vi persistono le stesse condizioni di base (cioè sulla base del principio secondo il quale le *condizioni uguali producono risultati simili*). A seconda del tipo di indagine, i principali obiettivi dell'analisi statistica possono essere la descrizione delle caratteristiche dei fenomeni osservati, l'individuazione di modelli per spiegare tali variazioni o la previsione di eventuali variazioni future. Per fare ciò è possibile utilizzare metodi di analisi che consentano di perseguire (come si vedrà) obiettivi descrittivi, esplorativi e induttivi.

L'analisi procede secondo livelli progressivi di complessità. In un primo momento l'analisi può riguardare l'esame dettagliato di ciascuna variabile separatamente (analisi univariata) col proposito di individuare i dati anomali, le distorsioni, le asimmetrie. Un secondo livello può prendere in considerazione le relazioni tra variabili o il confronto tra casi. Un terzo livello è quello che coinvolge gruppi di variabili (analisi multivariata) tentando di far emergere le componenti strutturali comuni ricorrendo successivamente a tecniche di aggregazione e disaggregazione, in modo che le componenti residuali non interferiscono con l'interpretazione delle regolarità riscontrate. A tutti gli approcci è possibile affiancare i metodi che consentono di effettuare l'analisi grafica.

I metodi statistici sono particolarmente rigidi e richiedono un'applicazione rigorosa. La mancata applicazione rigorosa può condurre a molti problemi. In teoria un ricercatore può preparare ed eseguire tutta una ricerca senza conoscere assolutamente nulla di statistica e poi rivolgersi e affidare l'analisi dei dati ad uno statistico; nella pratica è però importante che tutti i componenti del gruppo di ricerca abbiano una sufficiente informazione sui metodi statistici di base; d'altra parte è anche importante che l'esperto statistico venga coinvolto, data la rilevanza della statistica nella ricerca, in tutti i momenti della ricerca, per evitare applicazioni improprie che possano produrre dati non utilizzabili a livello di analisi statistica o risultati non pertinenti.

#### 5<sup>a</sup> FASE

La **comunicazione, descrizione, presentazione e diffusione dei risultati** rappresenta un momento molto importante che dovrebbe ricevere la stessa attenzione delle fasi precedenti. In genere si tratta di stendere un rapporto di ricerca che dovrebbe presentare i seguenti punti:

- definizione del problema e della popolazione oggetto di studio;
- descrizione delle procedure seguite per svolgere la ricerca e che riguardano la rilevazione (costruzione del campione, lavoro sul campo, metodi di controllo), gli strumenti di rilevazione, il trattamento dei dati (codifica, registrazione, eliminazione di dati, ecc.);
- descrizione dei metodi di analisi utilizzati e presentazione delle elaborazioni statistiche<sup>9</sup>;
- risultati ottenuti e confronto con le ipotesi di ricerca;
- valutazione dei risultati e, eventualmente, suggerimenti per ulteriori indagini e verifiche.

### **Le risorse per la ricerca: bilancio e programmazione**

Una volta stabilito e specificato il disegno di ricerca, occorre prendere decisioni riguardo al bilancio e alla programmazione; in questo modo si assicura che il progetto di ricerca sia completato all'interno delle risorse disponibili (finanziarie, di tempo, di personale, ecc.).

Specificando i parametri di tempo all'interno dei quali ciascuna operazione deve essere completata e i costi relativi, il progetto di ricerca può essere gestito ed eseguito nelle condizioni migliori.

Un utile approccio alla gestione di un progetto di ricerca è il cosiddetto *critical path model* (CPM), sistema che richiede la suddivisione del progetto di ricerca in attività componenti, determinando la sequenza di tali attività e stimando il tempo richiesto per ciascuna attività. Le attività e i relativi tempi stimati vengono organizzati all'interno di un diagramma nella forma di rete di flusso. Tale rappresentazione consente di evidenziare i percorsi critici, rappresentati dalle attività che ritardano la loro esecuzione.

Una versione aggiornata del CPM è il cosiddetto *PERT* (*Programm Evaluation and Review Technique*); si tratta di un approccio di programmazione espresso in termini probabilistici, che riconosce e misura l'incertezza dei tempi di completamento del progetto.

Un'altra tecnica è il *GERT* (*Graphical Evaluation and Review Technique*) che consente di gestire e rappresentare contemporaneamente in un grafico sia le probabilità di completamento sia i costi delle attività. Tale argomento è trattato in modo più diffuso nell'appendice A.

### **La documentazione**

Tutte le fasi dovrebbero essere continuamente documentate, in particolare nei seguenti momenti:

- **MODELLO DI CAMPIONAMENTO:** tale documentazione riguarda la popolazione dalla quale è stato estratto il campione, il metodo di campionamento, le procedure e gli eventuali problemi di campionamento e la struttura del campione utilizzato.

<sup>9</sup> Tali informazioni non dovrebbero essere solo fornite, ma anche richieste quando si utilizzano dati predisposti da altre persone ed altri gruppi.

- STRUTTURA DEL QUESTIONARIO: è molto importante documentare la struttura del questionario, l'ipotesi e i modelli d'analisi concepiti all'inizio del lavoro, e conservare copie documentate del questionario utilizzato.
- PREPARAZIONE E PROCEDIMENTO DI RILEVAZIONE: tale documentazione riguarda il corso di preparazione degli intervistatori, le istruzioni e i materiali utilizzati, le informazioni riguardanti il lavoro sul terreno (tempi, problemi incontrati, ecc.) e le procedure per il controllo di qualità della rilevazione.

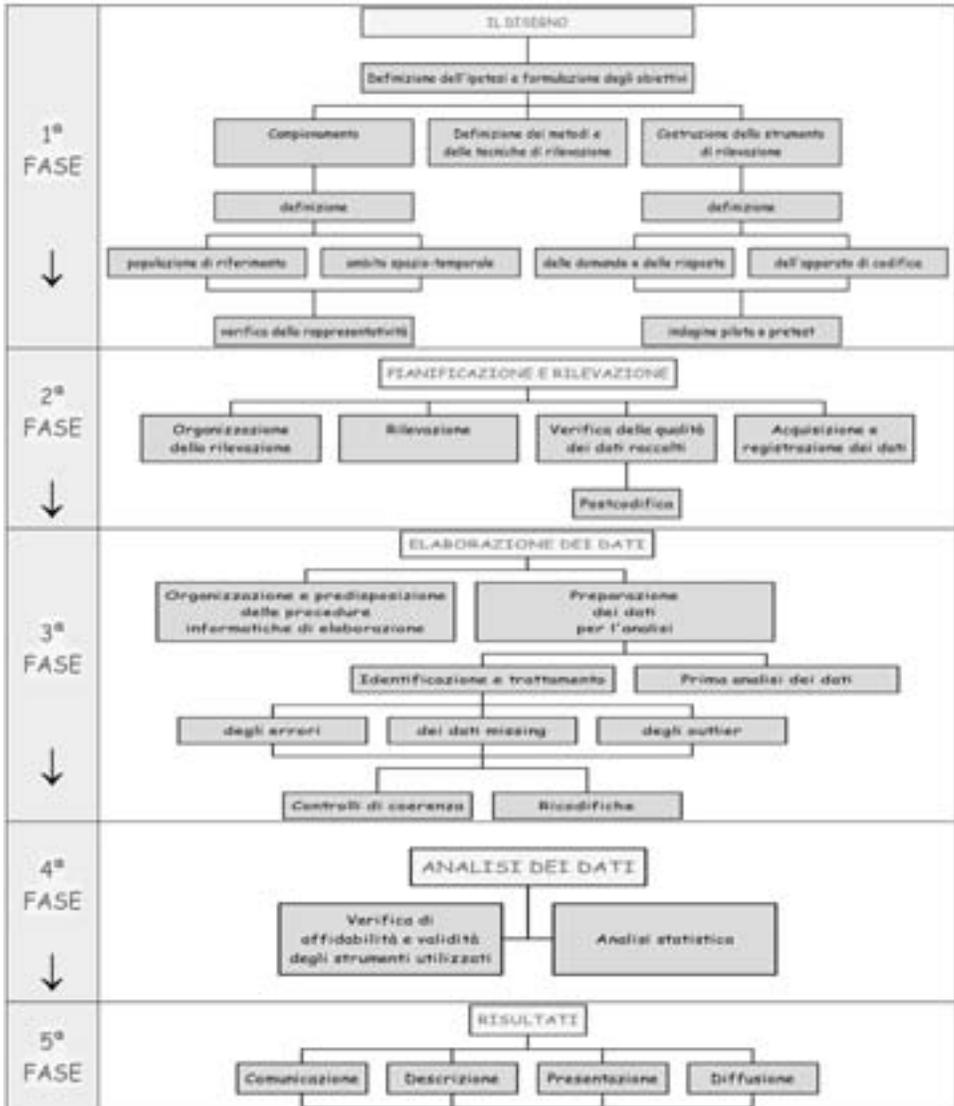


Fig. I. 1.4 La successione delle fasi di un'indagine

- CODIFICA e REGISTRAZIONE DEI DATI: è importante documentare in modo accurato ed esplicito tutti i codici assegnati ai dati raccolti per mezzo del questionario. Relativamente alle informazioni precodificate presenti sul questionario, basterà rendere disponibile una copia dello strumento di rilevazione, mentre è necessario elencare in maniera precisa i codici definiti in seguito.
- STRUTTURA DELL'ARCHIVIO DEI DATI.
- ELABORAZIONE DEI DATI CHE PRECEDE L'ANALISI: questa documentazione riguarda la descrizione delle tecniche utilizzate per gestire i valori *missing* e gli *outlier*, delle trasformazioni effettuate sui dati e delle analisi condotte su aree, indici, scale, ecc. Tutte queste operazioni sono finalizzate all'analisi statistica e possono influenzare i risultati e le conclusioni della ricerca. La documentazione non può essere solo tecnica ma dovrebbe riguardare anche i problemi incontrati e gli assunti sottostanti le operazioni.
- DISTRIBUZIONE DI FREQUENZA DI TUTTE LE VARIABILI: si tratta dei primi risultati e rappresentano un'importante fonte di documentazione per tutte le analisi successive e per questo dovrebbero essere sempre disponibili.