

● LA CASSETTA DEGLI ATTREZZI

STRUMENTI PER LE SCIENZE UMANE

Gianmaria Bottoni

INTRODUZIONE ALL'ANALISI MULTILIVELLO

Teoria e pratica dei modelli
di regressione multilivello

FrancoAngeli



La cassetta degli attrezzi. Strumenti per le scienze umane

Direttore

Giovanni Di Franco, Università di Roma “La Sapienza”

Comitato editoriale

Elena Battaglini, Ires-Cgil

Sara Bentivegna, Università di Roma “La Sapienza”

Claudio Bezzi, valutatore professionista

Alberto Marradi, Università di Firenze

Federica Pintaldi, Istat

Luciana Quattrocioni, Istat

Marta Simoni, Iref-Acli

La collana, rivolta a ricercatori accademici e professionisti, studiosi, studenti, e operatori del variegato mondo della ricerca empirica nelle scienze umane, si colloca sul versante dell’alta divulgazione e intende offrire strumenti di riflessione e di intervento per la ricerca.

Obiettivo è consolidare le discipline umane presentando gli strumenti di ricerca empirica, sia di raccolta sia di analisi dei dati, in modo intellegibile e metodologicamente critico così da consentirne l’applicazione proficua rispetto a definiti obiettivi cognitivi.

I testi sono scritti da professionisti della ricerca che, attingendo alla personale esperienza maturata in anni di attività, offrono ai lettori strumenti concettuali e tecnici immediatamente applicabili nella propria attività di ricerca.

Tutti i volumi pubblicati sono sottoposti a referaggio.

I lettori che desiderano informarsi sui libri e le riviste da noi pubblicati possono consultare il nostro sito Internet: *www.francoangeli.it* e iscriversi nella home page al servizio “informazioni” per ricevere via e-mail le segnalazioni delle novità o scrivere, inviando il loro indirizzo, a: “FrancoAngeli, viale Monza 106, 20127 Milano”.

Gianmaria Bottoni

INTRODUZIONE ALL'ANALISI MULTILIVELLO

**Teoria e pratica dei modelli
di regressione multilivello**

**La cassetta degli attrezzi
Strumenti per le scienze umane/1.7**

FrancoAngeli

Progetto grafico di copertina di Maria Teresa Pizzetti

1a edizione. Copyright © 2022 by FrancoAngeli s.r.l., Milano, Italy

L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore. L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni della licenza d'uso dell'opera previste e comunicate sul sito www.francoangeli.it.

119. La cassetta degli attrezzi. Strumenti per le scienze umane

Volumi pubblicati:

1. Giovanni Di Franco, *L'analisi dei dati con SPSS. Guida alla programmazione e alla sintassi dei comandi.*
2. Silvia Cataldi, *Come si analizzano i focus group.*
3. Federica Pintaldi, *Come si analizzano i dati territoriali.*
4. Giovanni Di Franco, *Il campionamento nelle scienze umane. Teoria e pratica.*
5. Lucia Coppola, *NVivo: un programma per l'analisi qualitativa.*
6. Simone Gabbriellini, *Simulare meccanismi sociali con NetLogo. Una introduzione.*
7. Giovanni Di Franco, *Dalla matrice dei dati all'analisi trivariata. Introduzione all'analisi dei dati.*
8. Giovanni Di Franco, *Tecniche e modelli di analisi multivariata.*
9. Federica Pintaldi, *Come si interpretano gli indici internazionali. Istruzioni per l'uso a favore di ricercatori, giornalisti e politici.*
10. Maria Paola Faggiano, *Gli usi della tipologia nella ricerca empirica.*
11. Danilo Catania, *Dati e rappresentazioni territoriali con ArcGis.*
12. Claudio Bezzi, *Fare ricerca con i gruppi. Guida all'utilizzo di focus group, brainstorming, Delphi e altre tecniche.*
13. Giovanni Di Franco (a cura di), *Il poliedro coesione sociale. Analisi teorica ed empirica di un concetto.*
14. Elena Battaglini, *Sviluppo territoriale. Dal disegno della ricerca alla valutazione dei risultati.*
15. Claudio Bezzi, *Domanda e ti sarà risposto. Costruire e gestire il questionario nella ricerca sociale.*
16. Elena Pavan, *La Network Analysis con Nodexl*
17. Maria Paola Faggiano, *L'analisi del contenuto di oggi e di ieri. Testi e contesti on e offline.*
18. Giovanni Di Franco, *Usi e abusi dei sondaggi politico-elettorali in Italia. Una guida per giornalisti, politici e ricercatori.*
 - 1.1. Sergio Mauceri, *Qualità nella quantità. La survey research nell'era dei Mixed Methods*
 - 1.2. Alberto Marradi, *Tutti redigono questionari. Ma è davvero così facile?*
 - 1.3. Giovanni Di Franco, *Introduzione alla sociologia empirica. Metodi e tecniche della ricerca sociale*
 - 1.4. Vittorio Cotesta, *Millennials. Avere vent'anni a Latina*
 - 1.5. Claudio Bezzi, *Manuale di ricerca valutativa*

1.6 Luca Salmieri (a cura di), *Servizi sociali e misure di contrasto alla povertà*

Volumi pubblicati in Open Access

1. Giovanni Di Franco (a cura di), *Giovani a tempo indeterminato. Valori e atteggiamenti dei giovani romani.*

Indice

1. Introduzione	pag.	9
1.1 Cosa leggere per saperne di più	»	14
2. Introduzione teorica all'analisi multilivello	»	15
2.1 Cosa leggere per saperne di più	»	24
3. Il modello di regressione multilivello	»	27
3.1 I presupposti dell'analisi multilivello	»	27
3.2 Il modello di regressione multilivello	»	34
3.3 Il modello nullo	»	38
3.4 Il modello a intercetta casuale e con coefficiente di regressione casuale	»	39
3.5 La complessità delle equazioni di regressione multilivello	»	45
3.6 Cosa leggere per saperne di più	»	48
4. Un semplice esempio di analisi di regressione multilivello	»	50
4.1 Cosa leggere per saperne di più	»	72
5. Alcuni aspetti pratici dell'analisi multilivello	»	74
5.1 La procedura per stimare un modello multilivello	»	74
5.2 La varianza riprodotta di primo e secondo livello	»	83
5.3 Centrare o standardizzare le variabili esplicative	»	94
5.4 Quale stimatore utilizzare?	»	101
5.5 Il campione nell'analisi multilivello	»	107
5.6 Come comparare modelli alternativi	»	112
5.7 Cosa leggere per saperne di più	»	119

6. Conclusioni: oltre il modello di regressione multilivello	pag.	123
6.1 Cosa leggere per saperne di più	»	127
7. Appendice: i software per l'analisi multilivello	»	129

1. Introduzione

Obiettivo del libro è introdurre in modo chiaro e per quanto possibile semplice l'analisi multilivello in generale e in modo particolare la regressione multilivello. Il lavoro è principalmente indirizzato a tutti gli studenti delle discipline inerenti le scienze sociali ed economiche, nonché a tutti quei ricercatori che fanno dell'analisi dei dati il loro principale strumento di analisi.

In coerenza con l'obiettivo di rendere il libro accessibile alla maggioranza dei lettori – e non solo a quelli dotati di una buona preparazione statistico-matematica – abbiamo ridotto per quanto possibile l'apparato di formule e la notazione matematica. Tuttavia in alcuni punti del manuale la formalizzazione matematica si è resa necessaria ai fini della chiarezza dell'esposizione. Laddove è stato necessario riportare le formule, queste sono state illustrate in termini discorsivi. Tale descrizione è volta a semplificare la comprensione anche a coloro i quali hanno meno familiarità con i concetti statistici descritti nel manuale. Pertanto, ai fini della comprensione del testo non è richiesta ai lettori alcuna particolare conoscenza pregressa. Ovviamente, la conoscenza di alcuni concetti statistici di base ed, in particolare, la familiarità con i modelli di regressione classici aiutano nel processo di comprensione del testo e degli argomenti qui presentati. Pertanto, si consiglia una lettura preliminare di un qualsiasi testo introduttivo alla regressione multipla e più in generale all'analisi multivariata (Di Franco 2017).

In conformità con lo stile della collana che ci ospita, in questo testo adottiamo i termini proposti, a partire dai primi anni '90 del secolo scorso, dalla

sezione di metodologia dell'Ais (associazione italiana di sociologia). Questa scelta è in gran parte determinata dalla volontà di separare il piano teorico-sostantivo e metodologico da quello prettamente tecnico-statistico. Per tale ragione il lettore leggerà 'riprodurre' e non 'spiegare' la varianza; 'rilevare' o 'quantificare' e non 'misurare'; 'tecnica' o 'procedura' o 'procedimento' e non 'metodo' usato per definire un procedimento codificato e standardizzato; 'controllare' e non 'verificare'; 'scarto-tipo' o 'scarto quadratico medio' e non 'deviazione standard'; 'intervallo di fiducia' e non 'intervallo di confidenza'. Per una ampia critica dell'uso ormai troppo inflazionato dei termini che qui abbiamo deciso di non usare, a partire dall'eccessiva estensione semantica del termine 'misurazione', si vedano, tra gli altri, i lavori di Maradi (1981, 2007).

Occorre chiarire innanzitutto cosa si intende per *analisi multilivello*. Ovviamente per *analisi* intendiamo un'analisi statistica dei dati volta ad analizzare le relazioni fra variabili. Fin qui nulla di nuovo. La particolarità risiede nel termine *multilivello*. Infatti, come meglio vedremo, le variabili da analizzare sono rilevate su diversi livelli gerarchici di analisi che si riferiscono ad unità di analisi differenti. L'esempio classico riguarda gli studenti inclusi in diverse classi scolastiche, a loro volta incluse in diverse scuole. In questo caso si analizzeranno le relazioni che sussistono fra le proprietà individuali, quelle riferite alle classi scolastiche e quelle riguardanti le scuole per valutare sia gli effetti dei diversi contesti sugli individui che le differenze sussistenti fra tali contesti.

L'analisi multilivello permette di analizzare le relazioni fra le proprietà dei diversi livelli e rispondere a domande di ricerca che sono alla radice di tutte le scienze sociali. Queste ultime infatti si prefiggono di studiare il comportamento umano in relazione agli altri gruppi sociali, al contesto in cui tali individui agiscono e, in modo ancor più generale, alla società in cui le persone sono inserite.

In maniera ancora più generica, possiamo dire che le scienze sociali studiano l'interazione fra l'individuo e il contesto sociale, l'influenza esercitata dal contesto sul singolo individuo e come a sua volta l'individuo influenzi il contesto sociale in cui è inserito.

Se, da un punto di vista operativo/epistemologico, distinguiamo gli individui dai contesti sociali e li assegniamo a diversi *livelli gerarchici* capiamo il funzionamento di base dell'analisi multilivello.

Tali livelli sono detti *gerarchici* in quanto è possibile immaginare un ordine al loro interno che, generalmente, prevede l'individuo come livello di base. Per esempio, pensiamo allo studio delle dinamiche che sussistono fra un medico (livello individuale), il reparto in cui è inserito (secondo livello) e l'intera struttura ospedaliera (terzo livello) e come tale medico eserciti un'influenza sul suo reparto e come questo a sua volta influisca sul singolo medico e come queste due entità influiscano sull'intero ospedale e viceversa. Se pensiamo a tali entità (medici, reparti, ospedali) come oggetti di analisi posti a diversi livelli gerarchici, capiamo come sia possibile rilevare in modo quantitativo le proprietà di tali oggetti di analisi per ogni livello della gerarchia.

Hox (2010), sulla base della tipologia proposta da Lazarsfeld e Menzel (che identificano a livello individuale variabili cosiddette assolute, relazionali, comparative e contestuali e a livello di gruppo variabili analitiche, strutturali e globali), propone una tipologia semplificata dei diversi tipi di variabili che possiamo avere nell'analisi multilivello a seconda del livello gerarchico a cui si riferiscono. Le variabili *globali* sono quelle variabili rilevate al livello nel quale sono definite. Per esempio il genere, l'altezza di un individuo, il suo livello di scolarizzazione sono tutte variabili globali a livello individuale. Mentre, per esempio, la dimensione di un reparto ospedaliero, la numerosità di una classe sono variabili globali di secondo livello che si riferiscono a unità di tipo aggregato. Poi abbiamo le variabili *strutturali*, ovvero quelle variabili rilevate ad un livello inferiore rispetto a quello a cui si riferiscono. L'esempio più semplice è la variabile intelligenza media in una classe scolastica, dove la variabile è riferita ad una unità di secondo livello (la classe) ma è rilevata a livello individuale ed è il risultato di un processo ulteriore di trasformazione (si rileva l'intelligenza di ogni singolo studente nella classe calcolandone successivamente la media).

Un'altra variabile strutturale è per esempio il reddito medio di una provincia o una regione. È chiaro pertanto come le variabili strutturali siano il frutto di un processo di aggregazione. Esattamente l'opposto abbiamo invece per le variabili cosiddette *contestuali*. Queste ultime sono variabili che si riferiscono ad un livello superiore della gerarchia ma che vengono attribuite alle unità di livello inferio-

re. Per l'esempio precedente abbiamo detto che la variabile intelligenza media di una classe è una variabile strutturale la cui unità di analisi è la classe. Tuttavia, possiamo assegnare ad ogni singolo studente di una specifica classe il valore riferito all'intelligenza media di quella classe (in tal modo, gli studenti di una stessa classe avranno tutti lo stesso valore su quella variabile). Intrinseco al concetto di livelli gerarchici, vi è un'altra nozione che ha a che fare con il concetto di *clustering*. Il clustering è uno dei motivi per i quali è necessario ricorrere all'analisi multilivello. Infatti, individui clusterizzati, ovvero appartenenti ad un medesimo gruppo, potrebbero mostrare comportamenti e atteggiamenti più simili fra di loro rispetto ad individui appartenenti ad altri gruppi. Si pensi ancora all'esempio dei medici, reparti ed ospedali. Gli individui (medici) interagiscono con il contesto sociale e lavorativo in cui sono inseriti, portando la propria specificità in termini di capacità, competenze, comportamenti ed atteggiamenti. Alcuni medici, lavorano a stretto contatto con altri, sono inseriti in reparti, ed ognuno apporta il proprio contributo. Il fatto di appartenere agli stessi reparti fa sì che i medici all'interno di uno stesso reparto condividano alcuni aspetti come pratiche lavorative, atteggiamenti, pratiche di socializzazione e via dicendo. Il contesto (reparto) agisce pertanto sul singolo individuo, ma allo stesso tempo è influenzato dai singoli individui (medici) al suo interno, che ne determinano le dinamiche. In ultima analisi, individui che condividono un medesimo contesto possono essere relativamente più simili fra di loro rispetto ad individui appartenenti ad altri contesti. Pertanto, analizzare gli individui come se questi fossero totalmente disancorati dai loro contesti sociali, non solo è epistemologicamente limitante, ma produce anche distorsioni statistiche come meglio vedremo nel proseguio del libro.

È utile inoltre precisare che l'analisi multilivello è tale solo e quando nell'analisi si ricorre a variabili di livello macro volte a dare conto della variazione dei coefficienti di regressione all'interno delle unità di livello superiore. Se ciò non avviene, siamo in presenza di *modelli di regressione con coefficienti random*. A nostro avviso l'espressione *analisi multilivello* deve essere usata solo nel caso in cui a tale variazione dei coefficienti di regressione si cerca di fornire

una spiegazione inserendo nel modello variabili rilevate su unità di analisi livello superiore.

Inoltre, occorre sottolineare come *analisi multilivello* descriva un insieme di procedimenti di analisi e non, come molti ritengono, solo la regressione multilivello. Infatti, la tecnica riferita alla regressione multilivello è soltanto una fra le tante tecniche disponibili. Notevole importanza stanno assumendo recentemente l'analisi fattoriale multilivello, la *path analysis* multilivello e gli ancor più complessi modelli ad equazioni strutturali multilivello dove i modelli propri dell'analisi fattoriale e i modelli strutturali inerenti la *path analysis* vengono integrati in un unico modello multilivello.

Tali tecniche, saranno descritte in modo molto sintetico nel capitolo 6 poiché richiedono un elevato livello di formalizzazione matematica ed un buon bagaglio di conoscenze pregresse, specialmente per i modelli ad equazioni strutturali multilivello, dove si presume un'ottima conoscenza degli stessi modelli nella loro forma classica (Di Franco 2016).

È necessario anche un ulteriore appunto. Nel testo non si farà riferimento in particolare ad alcun *software* statistico quando si presenteranno gli esempi di analisi, né tale testo presenterà come procedere attraverso il *software* mostrando la sintassi o l'interfaccia del programma e le procedure da effettuare per eseguire un'analisi multilivello. Questo per due motivi molto semplici. Innanzitutto, l'interfaccia dei *software* cambia col tempo e ciò che è valido ora non è detto che lo sia anche dopo un breve lasso di tempo nel caso un aggiornamento venga rilasciato, mentre ciò non risulta essere vero per i concetti di base presentati nel testo. Specialmente per modelli del genere, i *software* sono in costante aggiornamento.

Inoltre, una volta appresi i concetti di base si potrà utilizzare in maniera consapevole – e non meccanica – qualsiasi *software* il mercato metta a disposizione. Pertanto, lasciamo la scelta al lettore su quale *software* utilizzare. Ci limitiamo a dire che fra quelli disponibili i più diffusi che permettono di stimare modelli multilivello sono, STATA, MLwiN, HLM, R, SPSS (rinviamo i lettori interessati all'appendice dedicata ad una rapida rassegna del software idoneo a lavorare con i modelli multilivello). Ovviamente, tali programmi statistici non sono tutti sullo stesso piano, si va da quelli più com-

plici e generalisti (nel senso che permettono di stimare un ampio spettro di modelli, non solo ristretto all'analisi multilivello) come R, a quelli specializzati come MLwiN, fino a quelli meno complessi (ma che allo stesso tempo offrono meno opzioni) come SPSS.

1.1 Cosa leggere per saperne di più

Di Franco, G. (2016). *I modelli di equazioni strutturali. Concetti, strumenti e applicazioni*. Milano, FrancoAngeli.

Tale contributo rappresenta un'ottima introduzione ai modelli di equazioni strutturali.

Di Franco, G. (2017). *Tecniche e modelli di analisi multivariata*. Milano, FrancoAngeli.

Il testo rappresenta un'ottima introduzione ai modelli di analisi multivariata in genereale nonché dei modelli di regressione classici.

Di Franco, G., Marradi, A. (2020). *L'analisi bivariata*. FrancoAngeli.

Il volume presenta le forme principali di analisi della relazione fra due variabili. Le tecniche e i relativi coefficienti sono illustrati attraverso la loro applicazione in esempi di ricerca. Particolarmente utile per acquisire familiarità con i concetti basilari dell'analisi dei dati.

Lazarsfeld, P. F., Menzel, H. (1961). *On the relation between individual and collective properties*. In A. Etzioni (Ed.), *Complex organizations: A sociological reader*. New York: Holt, Rhinehart & Winston.

In tale contributo si definiscono i principi base da un punto di vista metodologico dell'analisi contestuale. Gli autori si interrogano sul rapporto fra individuo e contesto e pongono le basi concettuali per l'analisi multilivello.

Marradi, A., (1981). *Misurazione e scale: qualche riflessione e una proposta*. «Quaderni di Sociologia», XXIX, 4: 595-639.

Si veda questo articolo per una discussione critica del concetto di misurazione.

Marradi Alberto (2007). *Metodologia delle scienze sociali* (a cura di Rita Pavsic e Maria Concetta Pitrone). Bologna: Il Mulino.

Il testo fornisce gli strumenti di base e allo stesso tempo interessanti approfondimenti sulla metodologia delle scienze sociali. Il contributo si distingue per il suo approccio critico e ragionato alle maggiori problematiche inerenti il metodo e le tecniche di analisi dei dati.

2. Introduzione teorica all'analisi multilivello

Fin dalle origini delle scienze sociali, e della sociologia in particolare, gli studiosi si sono divisi sulla base dell'approccio ritenuto privilegiato nello studio del comportamento umano. Secondo la prospettiva olistica la società è intesa come una totalità con specifiche proprietà che le appartengono, non riducibili a quelle dei suoi membri. Al contrario, per l'approccio individualista, tutti i fenomeni sociali possono essere ridotti all'azione individuale. Nel primo caso, i fenomeni sociali hanno una propria specifica esistenza; nel secondo caso, quegli stessi fenomeni sono invece il prodotto dell'azione individuale.

Durkheim e Weber sono i rappresentanti principali di due tradizioni teoriche opposte: le teorie strutturali o sistemiche e le teorie dell'azione (Collins 1988).

Secondo Durkheim (1987) l'intenzionalità dell'individuo non è in nessun modo analizzabile in quanto non è né direttamente osservabile né attendibile al punto che le motivazioni di un'azione spesso sfuggono allo stesso soggetto agente. Per Durkheim le motivazioni degli individui non sono altro che un epifenomeno: le azioni individuali in realtà sono la causa di un fenomeno sociale. In particolare, il fatto sociale si configura come fenomeno che si differenzia da una semplice somma delle parti (individui) ma si pone con le sue caratteristiche riferite all'esteriorità rispetto all'individuo e alla forza coercitiva che esercita su esso.

Weber ha una visione opposta ritenendo come la sociologia debba avere come oggetto di analisi l'individuo e il suo agire (1922). L'autore individua quattro tipi di agire sociale: 1) l'agire razionale rispetto allo scopo; 2) l'agire razionale rispetto al valore; 3) l'agire affettivo; 4) l'agire tradizionale.

I tipi di azione sociale non sono né esaurienti dei modi dell'agire umano né sono direttamente osservabili nel comportamento concreto, piuttosto

rappresentano dei tipi puri – costruzioni sociologiche – di scopo eminentemente euristico (Weber 1922).

L'agire razionale, per le sue caratteristiche di intelligibilità, rappresenta l'oggetto privilegiato per l'analisi sociologica. Gli aspetti sociali di carattere collettivo, invece, sono dovuti ai singoli individui che indirizzano il proprio agire in base alle rappresentazioni che si fanno di tale collettività.

In un certo senso, l'analisi multilivello consente di superare l'opposizione fra la prospettiva microsociologica, il cosiddetto approccio microsociologico, e quella macrosociologica permettendo di connettere (non solo statisticamente ma anche teoricamente) gli individui alla società e ai contesti in cui essi agiscono.

Tale approccio di analisi, infatti, rende esplicito il fatto che fattori di livello superiore (ossia collocati ad un livello epistemologico al di sopra di quello individuale) possano esercitare un'influenza diretta, o indiretta, sul comportamento individuale. Ovviamente, tali livelli di analisi non devono essere pensati in nessun modo come "livelli ontologici", piuttosto come livelli epistemologici, ossia utili alle procedure di indagine. Tanto è vero che, contrariamente al senso comune, il livello di analisi individuale non necessariamente costituisce il livello base, ma possono essere individuati facilmente livelli sub-individuali – per esempio le diverse occorrenze o punteggi su scale di atteggiamento forniti da un individuo in un'indagine longitudinale. In questo caso il livello base è costituito dalle diverse osservazioni – o stati sulla proprietà – effettuate in diversi punti nel tempo (*repeated measures*) e il secondo livello è costituito dall'individuo al quale quelle osservazioni appartengono o sul quale sono state rilevate.

In aggiunta, i livelli di cui sopra possono variare a seconda delle discipline prese in considerazione. Per fare qualche esempio, possibili livelli di analisi nelle diverse discipline possono essere: l'insegnante, la classe e la scuola negli studi sull'efficienza del sistema educativo (tre livelli di analisi), la regione o lo Stato in sociologia, il reparto e l'ospedale negli studi sull'organizzazione aziendale.

Al fine di studiare il comportamento dell'individuo adottando una prospettiva globale e non parziale, non si può in alcun modo ignorare l'influenza che, per esempio, un gruppo sociale (o qualsiasi altra forma di aggregato) esercita sui comportamenti ed atteggiamenti dei

suoi membri e, di rimando, l'influenza che il singolo membro esercita sull'intero gruppo.

Ignorare tale effetto, da un lato, potrebbe condurre ad un'analisi incompleta e statisticamente errata, con la possibilità di cadere in errori di primo tipo e di secondo tipo, ossia individuando relazioni che in realtà non sussistono o viceversa affermando erroneamente l'assenza di una relazione.

Dall'altro lato, ignorare l'influenza reciproca fra individui e aggregati produce effetti negativi anche ad un livello teorico. Infatti, ignorando tale aspetto si perderebbe una grande occasione nel cercare di riprodurre i fenomeni sociali solo attraverso il ricorso a proprietà individuali, escludendo il contributo di proprietà contestuali e strutturali dal modello teorico sviluppato. In altre parole, si rinuncia a priori a studiare l'effetto del contesto sui comportamenti e atteggiamenti individuali nonché l'interazione fra ambiente e individuo. E tale rinuncia non è una scelta teoricamente neutra ma anzi rappresenta una forte presa di posizione teorica, ossia equivale ad affermare aprioristicamente che gli effetti contestuali non esercitano alcuna influenza, ovvero in termini statistici i loro coefficienti non sono statisticamente diversi da zero.

Nella prospettiva multilivello, così come da noi intesa in questo volume (la maggioranza degli autori che si sono occupati di analisi multilivello hanno totalmente tralasciato il piano teorico ed epistemologico, ignorandone i risvolti e le implicazioni, e si sono esclusivamente focalizzati sull'analisi statistica), la società (il contesto, l'aggregato, etc.) non è vista come un'entità che esercita la sua influenza indipendentemente dal comportamento individuale, dalle ragioni e dalle motivazioni che guidano il singolo individuo – così avviene per esempio nella prospettiva olistica durkheimiana. Piuttosto, la struttura sociale viene intesa come sistema di possibilità e costrizioni. In tal modo, gli individui sono il risultato dell'azione congiunta esercitata dalle proprie credenze/opinioni/atteggiamenti e dalla struttura delle possibilità/costrizioni del sistema sociale (Giddens 1984).

L'approccio multilivello, permettendo esplicitamente di considerare comportamenti ed atteggiamenti una funzione non solo delle caratteristiche individuali ma anche delle costrizioni e/o condiziona-

menti dell'ambiente (la maggioranza dei quali esercitano i loro effetti sull'individuo senza che costui ne abbia consapevolezza), consente fra le altre cose anche di esaminare i cosiddetti effetti perversi dell'azione sociale (Boudon 1977).

In tale prospettiva, il modello prevede, inoltre, una pluralità di livelli – teoricamente infinita ma praticamente limitata da numerosi fattori operativi come il disegno campionario, l'eccessiva complessità statistica, la difficoltà di reperire i dati a livelli superiori a quello individuale – che, come detto, possono variare da disciplina a disciplina e in base all'obiettivo cognitivo prefissato.

Non c'è un limite al numero di livelli individuabili da un ricercatore, per esempio gli individui appartenenti a specifici quartieri, contenuti in città, province, regioni, stati e, infine, organi sovranazionali. Ma in uno studio sulla partecipazione elettorale, per esempio, il livello riferito ai quartieri o alle città potrebbe perdere di rilevanza a favore delle circoscrizioni elettorali. La pluralità di livelli e il suo variare in base agli obiettivi prefissati è fondamentale anche perché permette di liberarci dalla prospettiva che pone i soggetti in opposizione alla società e viceversa.

In altre parole, la contrapposizione olistico *vs.* individualismo perde di pregnanza a favore di un approccio volto ad esplorare i diversi livelli in cui un individuo è inserito e ad indagare l'interazione fra tali livelli e le proprietà individuali.

L'analisi multilivello, permettendo di rilevare l'influenza che le caratteristiche di un ambiente e/o dei gruppi sociali esercitano sul comportamento e gli atteggiamenti individuali, ci consente in tal modo di evitare di incorrere sia nella fallacia ecologica sia nel suo inverso – la fallacia atomistica (Robinson 1950).

Nel primo caso, disporre dei dati raccolti a livello individuale (ossia delle distribuzioni congiunte e non solo marginali) scongiura il rischio di fallacia ecologica; nel secondo caso, la fallacia atomistica è eliminata inserendo gli individui nei contesti in cui agiscono.

Inoltre, i contesti individuabili, ovvero i livelli dell'analisi multilivello, non necessariamente devono coincidere con spazi fisici geograficamente delimitati. Tali contesti possono essere anche delimitati socialmente, come le reti relazionali in cui gli individui sono inseriti, gli ambienti lavorativi, gli status socio-economici a cui appartengono.

Il modello in fig. 2.1 illustra come l’approccio multilivello può connettere il livello micro e il livello macro d’analisi. Sono mostrati tre diversi modelli fra loro indipendenti. Il primo modello, quello contrassegnato dalla linea superiore non tratteggiata, rappresenta il modello aggregato, ovvero quando si fa ricorso ad indicatori che sono il frutto di un’aggregazione, come generalmente avviene nelle indagini ecologiche che fanno ricorso agli indicatori sociali.

Un esempio di tale modello potrebbe essere il tasso di natalità di uno Stato fatto regredire sulla percentuale di immigrati di quel Paese. Al contrario, la linea inferiore non tratteggiata rappresenta il modello individuale – ossia il tipo di modello più diffuso fra i ricercatori. Tale modello si riferisce generalmente alla spiegazione di una proprietà individuale attraverso diverse proprietà esplicative individuali (per esempio il livello di felicità di un individuo fatto regredire su una serie di predittori come l’età, il genere, il livello di istruzione, il reddito, lo stato civile, etc.).

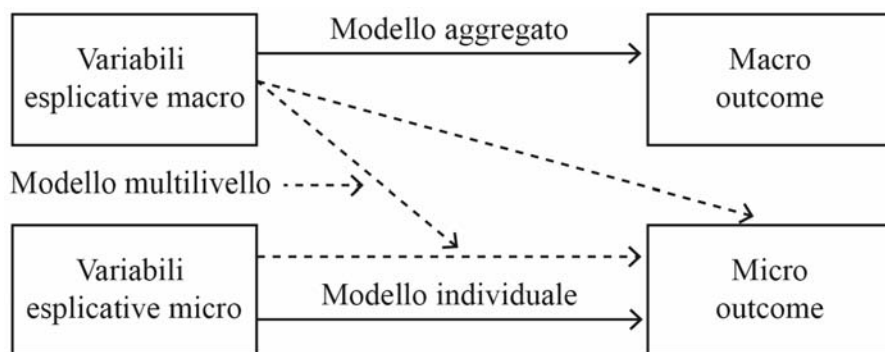


Figura 2.1 – Modello aggregato, individuale e multilivello

Le linee tratteggiate, invece, mostrano concettualmente il modello multilivello base e come tale modello riesca a connettere i livelli micro e macro.

Supponiamo che un ricercatore sia interessato a studiare i fattori che influenzano il risultato ad un test che rileva la conoscenza matematica (*micro outcome*) somministrato ad un campione di studenti delle scuole superiori appartenenti (*nested*) a diverse scuole del territorio italiano. Ad un livello individuale un possibile insieme di varia-