STATISTICA, anno LXII, n. 4, 2002

MODELLI LINEARI AD INTERCETTA CASUALE, STIMATORI E VALUTAZIONE DI SISTEMI FORMATIVI

G. Ghilardi, N. Orsini

1. INTRODUZIONE

I modelli di regressione possono essere impiegati per lo studio di dati dipendenti e nell'ultimo decennio sono stati pubblicati diversi lavori sull'argomento (Bryk e Raudenbush, 1992; Longford, 1993; Goldstein, 1995; Kreft e De Leeuw, 1998; Snijders e Bosker, 1999). Il caso di dati dipendenti si verifica quando le osservazioni individuali sono correlate, come accade con una certa frequenza per le ricerche nei campi dell'istruzione, della medicina e della biologia. Tale dipendenza può derivare dall'esistenza di una struttura gerarchica nei dati (*nested data*) e spesso caratterizza il contesto in cui avviene la rilevazione (dati raggruppati, quali ad esempio gli studenti appartenenti a diverse classi) oppure l'operazione di rilevazione (dati longitudinali, quali ad esempio le misure effettuate da diversi soggetti sullo stesso paziente). In entrambi i casi è opportuno considerare il legame esistente tra le singole osservazioni all'interno di ciascun gruppo di unità statistiche, ricorrendo all'impiego di modelli statistici e stimatori adeguati per l'analisi di questo tipo di dati.

In particolare, nel caso delle ricerche effettuate nel campo dell'istruzione, ci possiamo proporre di rilevare l'esistenza di differenze tra classi (gruppi di unità statistiche) di studenti (unità statistiche) sulla base di una certa misura individuale di un risultato (Y), tenendo conto del fatto che le caratteristiche (x) degli studenti e quelle (z) delle classi possono essere rilevanti nel determinare tale risultato. Nella letteratura (Aitkin e Longford, 1986; Goldstein e Spiegelhalter, 1996) questa operazione di confronto viene definita una valutazione di efficacia relativa, che può essere svolta attraverso le cosiddette tecniche di analisi multilivello (Hox, 1995).

La base logica di tali tecniche deriva dalla considerazione che il risultato individuale Y dipende da fattori riferibili all'unità statistica oggetto di studio (unità di primo livello) e da fattori riferibili al gruppo di appartenenza (unità statistica di secondo livello) delle unità statistiche. Ciò che possiamo osservare (fattori osservabili) è rappresentato da una o più variabili x riferite all'unità di primo livello e da una o più variabili z riferite all'unità di secondo livello. Invece, tutto ciò che non è osservabile (fattori non osservabili o non osservati) viene considerato afferente ad un termine di errore, che può essere indicato con la lettera R, quando si fa riferimento all'unità statistica di primo livello, e con la lettera U nel caso delle unità di secondo livello.

In questo lavoro viene riportata sinteticamente una descrizione del modello statistico lineare a due livelli e dei risultati ottenuti mediante la versione del modello che è stata messa a punto per un'applicazione nel campo della misura del disagio manifestato dagli studenti nei confronti degli insegnanti sulla base di un insieme di dati, che sono stati raccolti nell'ambito di un progetto finalizzato di un consorzio interuniversitario interessato allo studio del disagio scolastico degli studenti delle classi elementari.

2. IL MODELLO STATISTICO LINEARE MULTILIVELLO O GERARCHICO

Il modello statistico applicato per lo studio di dati dipendenti è individuato in letteratura come un modello multilivello o gerarchico (Kreft e De Leeuw, 1998; Snijders e Bosker, 1999). Data l'unità statistica i appartenente al gruppo j (con $i = 1, ..., n_j$ e $j = 1, ..., N$) per N gruppi di numerosità n_j, facendo riferimento al caso di una variabile dipendente Y, di una variabile individuale indipendente x e di una variabile z di gruppo, il modello lineare multilivello può essere rappresentato mediante un'equazione di primo livello,

$$Y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + R_{ij} \tag{1}$$

che rappresenta in base ai coefficienti β_{0j}, β_{1j} la relazione esistente tra le variabili (Y e x) riferite all'unità statistica di primo livello, nella quale il termine d'errore R_{ij} si assume con distribuzione normale, con media pari a zero e varianza costante pari a σ^2, e mediante le equazioni di secondo livello,

$$\beta_{0j} = \gamma_{00} + \gamma_{01}z_j + U_{0j} \tag{2}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}z_j + U_{1j} \tag{3}$$

che rappresentano in base ai coefficienti γ_{00}, γ_{01} e γ_{10}, γ_{11} le relazioni esistenti tra la variabile osservata (z) su ciascun gruppo (unità di secondo livello) e i coefficienti β_{0j}, β_{1j} inseriti nell'equazione di primo livello. Nelle equazioni (2) e (3) i termini di errore U_{0j}, U_{1j} si assumono con media zero, varianze costanti τ_0^2, τ_1^2 e indipendenti tra loro e dal termine di errore R_{ij}, che figura nell'equazione di primo livello (1).

Il caso particolare del modello lineare a due livelli nel quale solo l'intercetta β_{0j} può assumere dei valori diversi in funzione del gruppo di appartenenza delle unità statistiche è conosciuto in letteratura con diversi nomi: modello a componenti di varianza (*Variance Components;* Searle *et al.,* 1992); modello di analisi della covarianza ad effetti casuali (*Covariance Model with Random Effects;* Bryk e Raudenbush, 1992); modello a coefficienti casuali (*Random Coefficient Model*; Longford, 1993) e modello ad intercetta casuale (*Random Intercept Model;* Snijders e Bosker, 1999). Esso può essere utilizzato per valutare le differenze esistenti tra i gruppi delle unità statistiche, considerando le equazioni (1) (2) e assumendo costante il coefficiente $\beta_{1j} = \gamma_{10}$.

Così facendo, per l'impiego del modello si assume che il coefficiente β_{0j} sia una variabile casuale e ciò corrisponde a ritenere che ciascun gruppo costituisca un gruppo scelto a caso da una popolazione ipotetica di gruppi. Inoltre, all'interno di ciascun gruppo si assume che le unità statistiche rappresentino un campione casuale semplice estratto da una popolazione ipotetica di unità di primo livello.

Dalle caratteristiche del modello si ricava la varianza

$$\mathrm{Var}(Y_{ij} \mid x_{ij}) = \mathrm{Var}(R_{ij}) + \mathrm{Var}(U_{0j}) = \sigma^2 + \tau_0^2 \qquad \forall\ i \text{ e} \forall\ j \tag{4}$$

della variabile Y_{ij} dato il valore x_{ij}, che è uguale alla somma delle varianze σ^2 e τ_0^2 dei termini R_{ij}, U_{0j} di errore. In particolare, la varianza

$$\tau_0^2 = \mathrm{Cov}(Y_{ij}, Y_{i'j} \mid x_{ij}, x_{i'j}) = \mathrm{Var}(U_{0j}) = \mathrm{Var}(\beta_{0j}) \qquad i \neq i' \text{ e} \forall\ j \tag{5}$$

misura la covarianza fra i valori Y_{ij}, $Y_{i'j}$ della variabile *Y* all'interno del gruppo *j*.

Per quanto riguarda la componente σ^2 (detta *within*) della varianza della variabile *Y* tra le unità statistiche all'interno dei gruppi e la componente τ_0^2 (detta *between*) della varianza esistente tra i gruppi, il rapporto

$$\rho(Y \mid x) = \frac{\tau_0^2}{\tau_0^2 + \sigma^2} \tag{6}$$

tra la componente di varianza fra gruppi (τ_0^2) e la varianza totale ($\tau_0^2 + \sigma^2$) viene definito coefficiente di correlazione intragruppo, ed è utilizzato come una misura della relazione esistente tra i valori della variabile *Y* all'interno di ogni gruppo o unità di secondo livello.

Passando a considerare il problema della stima dei parametri incogniti γ e delle componenti σ^2, τ_0^2 di varianza, che figurano nel modello lineare ad intercetta casuale, i metodi proposti per ottenere gli stimatori di tali parametri (Bryk e Raudenbush, 1992), sono quello dei minimi quadrati e quello della massima verosimiglianza. Gli stimatori di massima verosimiglianza sono ritenuti più efficienti di quelli ottenuti con il metodo dei minimi quadrati ordinari (Longford, 1993) e il valore positivo del coefficiente di correlazione intraclasse non è ritenuto influente nel pregiudicare la proprietà di correttezza di tali stimatori, ma un valore elevato di questo indice determina generalmente una sottostima della loro variabilità, che può far sorgere qualche problema nella fase della verifica delle ipotesi (Kreft e De Leeuw, 1998). Inoltre, tra gli stimatori di massima verosimiglianza (Bryk e Raudenbush, 1992), sono stati proposti due tipi di stimatori. Il primo tipo viene definito di massima verosimiglianza completa (FML) ed è caratterizzato dal fatto che la stima dei coefficienti di regressione e delle componenti di varianza avviene in due fasi distinte, mentre il secondo tipo è definito di massima verosimiglianza ristretta (REML), perché consente di stimare simultaneamente tutti i parametri.

Dalla letteratura sull'argomento risulta che questi due tipi di stimatori non differiscono tanto per i valori delle stime dei parametri relativi ai coefficienti (γ) di regressione, quanto per quelli della stima delle componenti (σ^2, τ_0^2) di varianza e, in particolare, che con lo stimatore REML tali componenti vengono stimate tenendo conto di un numero inferiore di gradi di libertà, rispetto a quello che caratterizza lo stimatore FML, che sottostima le componenti di varianza di una certa quantità, la quale diminuisce con l'aumento del numero delle unità campione di secondo livello (Snijders e Bosker, 1999) e che è ritenuta trascurable oltre un certo valore del numero N dei gruppi (N>30). Inoltre, notiamo che dai risultati ottenuti mediante diversi studi di simulazione (Kreft e De Leeuw, 1998) non è stato possibile individuare il metodo migliore di stima delle componenti di varianza, poiché è stato osservato che lo stimatore REML pur essendo corretto è meno efficiente dello stimatore FML. Circa gli algoritmi disponibili per la determinazione di tali stime di massima verosimiglianza (Bryk e Raudenbush, 1992; Longford, 1993; Goldstein, 1995), essi sono diversi e per questo lavoro abbiamo utilizzato gli stimatori IGLS (*Iterative Generalized Least Squares*; Goldstein, 1986) e quelli RIGLS (*Restricted Iterative Generalized Least Squares*; Goldstein, 1989), i quali nel caso di normalità della distribuzione dei termini di errore, corrispondono, rispettivamente, agli stimatori di massima verosomiglianza completa e ristretta.

A partire da questa sintetica descrizione del modello multilivello e di alcuni problemi che si incontrano per il suo impiego, nel paragrafo successivo riportiamo la descrizione della versione del modello che è stato usato per lo studio del disagio scolastico manifestato dagli studenti di alcune scuole elementari nei confronti degli insegnanti.

3. L'IMPIEGO DEL MODELLO LINEARE MULTILIVELLO PER LO STUDIO DEL DISAGIO SCOLASTICO DEGLI STUDENTI

L'impiego del modello lineare multilivello appare adeguato per l'analisi dei dati ottenuti nell'ambito di ricerche effettuate nel campo dell'istruzione, quando ci si propone di rilevare l'esistenza di differenze tra classi scolastiche (unità di secondo livello), ciascuna delle quali accoglie un certo numero di studenti (unità di primo livello), sulla base della misura individuale di una variabile Y rilevata su ogni studente, tenendo conto del fatto che le caratteristiche x degli studenti e quelle z delle classi possono essere rilevanti nel determinare il valore osservato della variabile Y. Tale operazione di confronto tra classi, come si è detto, viene definita col nome di valutazione di efficacia relativa (Aitkin e Longford, 1986; Goldstein e Spiegelhalter, 1996) e rappresenta l'oggetto di questo paragrafo, nel quale riportiamo la descrizione della versione del modello che è stato sperimentato per l'analisi dei dati ottenuti nell'ambito di un progetto finalizzato allo studio del disagio scolastico da parte del Consorzio Interuniversitario F.I.T. (Formazione Innovazione e Tecnologie) su richiesta del Comune di Massa. Questo progetto ha coinvolto 592 ($M=\sum n_j$) studenti delle scuole elementari provenienti da 38 (N) classi del distretto scolastico di competenza. La variabile dipendente prescelta è rappresentata dal disagio manifestato (Y_{ij}) dallo studente i della classe j nei confronti degli insegnanti ed è stata utilizzata per effettuare un confronto tra le classi, facendo riferimento al disagio che gli studenti di ogni classe hanno manifestato nei confronti dei propri insegnanti e che è stato rilevato come una variabile di tipo qualitativo con nove modalità ordinabili. Oltre al disagio manifestato dagli studenti nei confronti degli insegnanti, i dati messi a nostra disposizione riguardano diverse caratteristiche individuali (x) degli studenti, ma non è stato possibile disporre delle caratteristiche (z) riferite alle classi di appartenenza, che tuttavia potrebbero essere facilmente prese in considerazione attraverso il modello descritto nel paragrafo precedente.

In base ai dati a nostra disposizione il modello adeguato per lo studio del disagio scolastico può essere rappresentato, considerando: l'equazione di primo livello (1), nella quale il valore osservato Y_{ij} per lo studente i della classe j è dato dalla combinazione lineare della variabile individuale x secondo i coefficienti β_{0j}, β_{1j} e dal termine di errore R_{ij}; le equazioni di secondo livello

$$\beta_{0j} = \gamma_{00} + U_{0j} \tag{7}$$

$$\beta_{1j} = \gamma_{10} \tag{8}$$

mediante le quali si assume che l'intercetta β_{0j} dell'equazione (1) sia una variabile casuale, il cui valore è diverso da classe a classe e misura il disagio manifestato dagli studenti all'interno della classe j, tenendo conto che le caratteristiche individuali x degli studenti sono diverse a seconda della classe di appartenenza. In particolare, notiamo che il coefficiente β_{1j} misura la relazione tra la variabile esplicativa x e la variabile risultato Y ed è costante (pari a γ_{10}) per tutte le classi osservate. Il modello ad intercetta casuale rappresentato dalle equazioni (1), (7) e (8) può essere visto, come un insieme di modelli di regressione (uno per ogni classe scolastica), attraverso il quale si possono evidenziare le differenze tra le classi mediante quelle esistenti tra i valori dell'intercetta β_{0j}. In questo modello con effetti fissi e casuali lo studente e la classe di provenienza sono considerate due fonti di variabilità distinte. La componente (*within*) σ^2 della varianza della variabile R_{ij} rappresenta la variabilità del disagio degli studenti all'interno delle classi, mentre la componente *(between)* τ_0^2 della varianza del termine di errore U_{0j} rappresenta la variabilità tra classi. Il rapporto tra la componente di varianza fra classi (τ_0^2) e la componente di varianza totale ($\tau_0^2 + \sigma^2$) determina il coefficiente ρ di correlazione intraclasse, che può essere utilizzato come una misura della relazione esistente tra i valori Y_{ij} del disagio manifestato dagli studenti all'interno delle classi.

Il modello lineare descritto a due livelli è stato utilizzato per lo studio del disagio scolastico manifestato dagli studenti nei confronti dei loro insegnanti, considerando diverse versioni del modello stesso in funzione di tre variabili esplicative individuali rilevate su ciascuno studente, ovvero il disagio (rilevato con nove modalità ordinabili) che lo studente vive in famiglia (x_{1ij}), il sesso (x_{2ij}) e il disagio (rilevato con dieci modalità ordinabili) manifestato dallo studente verso la scuola (x_{3ij}) nel suo complesso. Queste variabili sono state inserite di volta in volta nell'equazione di primo livello (1), considerando altrettanti coefficienti $\beta_{1j} = \gamma_{10}$, $\beta_{2j} = \gamma_{20}$, $\beta_{3j} = \gamma_{30}$ ed ottenendo diverse versioni del modello prescelto a partire da quella priva di variabili esplicative (tabella 1). Con tale procedimento è possibile evidenziare il ruolo delle tre variabili individuali ai fini della valutazione del disagio manifestato all'interno di ciascuna classe dagli studenti nei confronti dei loro insegnanti e della formazione di una graduatoria delle classi (diversa a seconda delle variabili individuali inserite nel modello) in funzione dei valori dei coefficienti β_{0j}, che misurano il disagio rilevato all'interno di ciascuna classe, tenendo conto delle variabili esplicative prese in esame.

Pertanto, per il confronto tra le classi in termini di disagio manifestato dagli studenti verso gli insegnanti si rende necessaria la stima del parametro β_{0j} dell'equazione (1), che è considerato un indicatore sintetico idoneo al confronto *ceteris paribus* tra classi (Goldstein, 1997).

TABELLA 1

Modelli ad intercetta casuale con diverse variabili esplicative

Modello	Equazione	Coefficienti
MIC-0	$Y_{ij} = \beta_{0j} + R_{ij}$	$\beta_{0j} = \gamma_{00} + U_{0j}$
MIC-1 (x_1)	$Y_{ij} = \beta_{0j} + \beta_{1j}x_{1ij} + R_{ij}$	$\beta_{1j} = \gamma_{10}$
MIC-2 (x_2)	$Y_{ij} = \beta_{0j} + \beta_{1j}x_{1ij} + \beta_{2j}x_{2ij} + R_{ij}$	$\beta_{2j} = \gamma_{20}$
MIC-3 (x_3)	$Y_{ij} = \beta_{0j} + \beta_{1j}x_{1ij} + \beta_{2j}x_{2ij} + \beta_{3j}x_{3ij} + R_{ij}$	$\beta_{3j} = \gamma_{30}$

4. STIMATORI DELL'INTERCETTA DI UN MODELLO LINEARE A DUE LIVELLI

Nella prima fase della stima dei coefficienti del modello lineare i parametri incogniti da stimare sono rappresentati dai parametri γ e dalle componenti di varianza di primo (σ^2) e di secondo livello (τ_0^2). Facendo ricorso al metodo della massima verosimiglianza si sono presi in esame due tipi di stimatori: quello di massima verosimiglianza completa (FML) e quello di massima verosimiglianza ristretta (REML). In pratica, i calcoli necessari possono essere effettuati mediante degli algoritmi opportuni di calcolo automatico, rispettivamente, l'algoritmo IGLS (*Iterative Generalized Least Squares*; Goldstein, 1986) e quello RIGLS (*Restricted Generalized Least Squares*; Goldstein, 1989), che, nel caso di validità dell'assunzione fatta sulla normalità delle distribuzioni, forniscono dei risultati corrispondenti a quelli che si hanno, rispettivamente, per gli stimatori FML e REML (Bryk e Raudenbush, 1992).

In particolare (Neter e Wassermann, 1974), uno stimatore

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} + \hat{U}_{0j} \tag{9}$$

del coefficiente β_{0j} può essere ricavato come funzione dello stimatore $\hat{\gamma}_{00}$ della media generale della variabile Y e del residuo $\hat{U}_{0j}$ riferito alla j-ma classe. Nel caso di un modello ad intercetta casuale con una variabile esplicativa x_1, il residuo

$$\hat{U}_{0j} = \overline{Y}_j - \hat{\gamma}_{00} - \hat{\gamma}_{10}(\overline{x}_{1j} - \overline{x}_1) \tag{10}$$

della j-ma classe è funzione della media $\overline{Y}_j$ di classe della variabile risultato Y, del valore degli stimatori $\hat{\gamma}_{00}$, $\hat{\gamma}_{10}$ e della differenza tra il valor medio $\overline{x}_{1j}$ di classe e la media generale $\overline{x}_1$ della variabile x_1.

Dalle equazioni (9) e (10) si ricava lo stimatore

$$\hat{\beta}_{0j}^{ORD} = \overline{Y}_j - \hat{\gamma}_{10}(\overline{x}_{1j} - \overline{x}_1) \tag{11}$$

del coefficiente β_{0j} che possiamo definire stimatore ordinario (ORD) dell'intercetta (tabella 2).

TABELLA 2

Stimatore ordinario $\hat{\beta}_{oj}^{ORD}$ per un modello ad intercetta casuale secondo il numero di variabili esplicative individuali x

Numero di variabili	Stimatore ordinario $\hat{\beta}_{oj}^{ORD} = \hat{\gamma}_{00} + \hat{U}_{0j}$
0 (MIC-0)	$\overline{Y}_j$
1 (MIC-1)	$\overline{Y}_j - \hat{\gamma}_{10}(\overline{x}_{1j} - \overline{x}_1)$
2 (MIC-2)	$\overline{Y}_j - \hat{\gamma}_{10}(\overline{x}_{1j} - \overline{x}_1) - \hat{\gamma}_{20}(\overline{x}_{2j} - \overline{x}_2)$
3 (MIC-3)	$\overline{Y}_j - \hat{\gamma}_{10}(\overline{x}_{1j} - \overline{x}_1) - \hat{\gamma}_{20}(\overline{x}_{2j} - \overline{x}_2) - \hat{\gamma}_{30}(\overline{x}_{3j} - \overline{x}_3)$

Pertanto, l'inserimento nel modello statistico della prima variabile esplicativa individuale x_1 modifica la valutazione del valore medio $\overline{Y}_j$ di classe in base a due elementi:

1) lo stimatore $\hat{\gamma}_{10}$, che misura la relazione esistente tra la variabile esplicativa inserita x_1 e la variabile risultato *Y*;

2) il valore dello scostamento $(\overline{x}_{1j} - \overline{x}_1)$, relativo alla variabile esplicativa x_1 tra la media $\overline{x}_{1j}$ della classe e la media generale $\overline{x}_1$.

La logica di fondo nella costruzione dell'indicatore di efficacia risiede nel fatto che se la variabile esplicativa inserita nel modello è correlata positivamente con la variabile risultato ($\hat{\gamma}_{10}$ >0), allora quanto maggiore è il valore medio di classe della variabile esplicativa [$(\overline{x}_{1j} - \overline{x}_1)$ >0] tanto minore sarà la valutazione del valore medio $\overline{Y}_j$ di classe della variabile risultato. L'inserimento di più variabili esplicative nel modello statistico estende questa procedura di aggiustamento del valor medio $\overline{Y}_j$ di classe, ma la logica di fondo nella costruzione dell'indicatore di efficacia rimane inalterata (tabella 2).

La direzione e l'entità dell'aggiustamento della media $\overline{Y}_j$ di ciascuna classe, dipende dal segno e dal valore dello stimatore ($\hat{\gamma}$) e dello scostamento [$(\overline{x}_j - \overline{x})$] riferito alla variabile *x*.

La differenza tra una valutazione della classe in base ad un indicatore grezzo (senza variabili esplicative individuali) e netto (con variabili esplicative individuali) di efficacia dipende dunque dagli elementi indicati e ciò implica che se le variabili esplicative introdotte nel modello sono scarsamente correlate con la variabile risultato ($\gamma \to 0$) e/o non si registrano grosse differenze tra classi in ordine a tali variabili esplicative [$(\overline{x}_j - \overline{x}) \to 0$], allora possiamo confrontare le classi sulla base del semplice valor medio $\overline{Y}_j$ di classe della variabile risultato. Lo stimatore ordinario dell'intercetta è corretto, ma non è ritenuto quello più efficiente tra gli stimatori dell'intercetta impiegati attualmente nell'analisi multilivello (Bryk e Raudenbush, 1992; Snijders e Bosker, 1999).

Oltre lo stimatore ordinario $\hat{\beta}_{0j}^{ORD}$, per la stima dell'intercetta β_{0j} è stato proposto anche uno stimatore empirico-bayesiano, che in qualche modo utilizza contemporaneamente due tipi di informazioni (Efron e Morris, 1975): quelle che si hanno dai dati provenienti dalla *j*-ma classe e quelle provenienti dalle *N* classi. In questo caso, lo stimatore

$$\hat{\beta}_{0j}^{EB} = \lambda_j \hat{\beta}_{0j}^{ORD} + (1 - \lambda_j)\hat{\gamma}_{00} \tag{12}$$

empirico-bayesiano (EB) dell'intercetta è funzione dello stimatore $\hat{\beta}_{0j}^{ORD}$, dello stimatore $\hat{\gamma}_{00}$ della media generale della variabile Y e dipende dal peso (*shrinkage factor*),

$$\lambda_j = \frac{\tau_0^2}{\tau_0^2 + \dfrac{\sigma^2}{n_j}} \qquad \text{con} \quad 0 \le \lambda_j \le 1 \tag{13}$$

che varia tra 0 ed 1 in funzione della numerosità n_j di classe; della componente τ_0^2 di varianza tra classi e della componente σ^2 di varianza dentro le classi.

La combinazione lineare (12) viene proposta come la soluzione ottimale per la stima dell'intercetta β_{0j}, è definita intercetta a posteriori o stimatore empirico-bayesiano dell'intercetta (Snijders e Bosker, 1999) e assume un'espressione diversa (tabella 3) a seconda del numero delle variabili esplicative *x* inserite nel modello (1).

TABELLA 3

Stimatore empirico-bayesiano $\hat{\beta}_{0j}^{EB}$ per un modello ad intercetta casuale secondo il numero di variabili esplicative individuali x

Numero di variabili	Stimatore empirico-bayesiano $\hat{\beta}_{0j}^{EB} = \hat{\gamma}_{00} + \hat{U}_{0j}$
0 (MIC-0)	$\lambda_j \overline{Y}_j + (1-\lambda_j)\hat{\gamma}_{00}$
1 (MIC-1)	$\lambda_j[\overline{Y}_j - \hat{\gamma}_{10}(\overline{x}_{1j} - \overline{x}_1)] + (1-\lambda_j)\hat{\gamma}_{00}$
2 (MIC-2)	$\lambda_j[\overline{Y}_j - \hat{\gamma}_{10}(\overline{x}_{1j} - \overline{x}_1) - \hat{\gamma}_{20}(\overline{x}_{2j} - \overline{x}_2)] + (1-\lambda_j)\hat{\gamma}_{00}$
3 (MIC-3)	$\lambda_j[\overline{Y}_j - \hat{\gamma}_{10}(\overline{x}_{1j} - \overline{x}_1) - \hat{\gamma}_{20}(\overline{x}_{2j} - \overline{x}_2) - \hat{\gamma}_{30}(\overline{x}_{3j} - \overline{x}_3)] + (1-\lambda_j)\hat{\gamma}_{00}$

Il peso λ_j può essere interpretato come una misura dell'affidabilità (***reliability***) dello stimatore $\hat{\beta}_{oj}^{ORD}$ del parametro β_{0j} (Bryk e Raudenbush, 1992) ed esso sarà tanto più affidabile ($\lambda_j \to 1$), quanto più grande è la numerosità campionaria n_j delle classi e quanto più grande è la componente τ_0^2 della varianza tra classi rispetto alla componente σ^2 della varianza dentro le classi, ossia quanto più grandi sono le differenze tra i valori medi di classe della variabile risultato *Y*. In altre parole, il peso λ_j fa tendere il valore dello stimatore empirico-bayesiano $\hat{\beta}_{0j}^{EB}$ dell'intercetta verso quello dello stimatore ordinario $\hat{\beta}_{oj}^{ORD}$. Viceversa, quando il peso λ_j è piccolo ($\lambda_j \to 0$) il valore dello stimatore empirico-bayesiano $\hat{\beta}_{0j}^{EB}$ tende verso quello dello stimatore $\hat{\gamma}_{00}$ della media generale di tutte le classi della variabile risultato. Infine, vale la pena di sottolineare che lo stimatore $\hat{\beta}_{0j}^{EB}$ empirico-bayesiano (Lindley e Smith, 1972) è più efficiente di quello ordinario $\hat{\beta}_{oj}^{ORD}$ e che ai fini dell'interpretazione del suo significato esso guadagna forza (*borrowing strenght*) per le classi più povere di informazione ($n_j \to 0$), attingendo alle informazioni contenute nell'intero campione delle classi osservate (Kreft e De Leeuw, 1998).

Per avere un riscontro empirico sulle caratteristiche degli stimatori menzionati, si è ritenuto opportuno applicarli ai dati sul disagio manifestato dagli studenti nei confronti degli insegnanti, per poi procedere nel confronto tra i risultati ottenuti, anche perché in pratica le proprietà di tali stimatori potrebbero essere parzialmente vanificate, dal fatto che le ipotesi formulate per il modello talvolta non sono perfettamente conformi alla natura del problema affrontato.

5. LA VALUTAZIONE DI EFFICACIA RELATIVA DI UN SISTEMA FORMATIVO IN BASE AL DISAGIO SCOLASTICO

Dato che il disagio manifestato dagli studenti di ciascuna classe scolastica può essere misurato mediante il parametro β_{0j}, il valore di uno stimatore adeguato $\hat{\beta}_{0j}$ è utile per costruire una graduatoria delle diverse classi. A tale scopo, partendo dalle stime puntuali dell'intercetta β_{0j} del modello di regressione lineare a due livelli con tre variabili individuali *x*, si possono individuare degli intervalli di confidenza per effettuare in maniera appropriata un confronto tra le classi scolastiche (Goldstein e Healy, 1995).

In particolare, ricorrendo allo stimatore $\hat{\beta}_{0j}^{EB}$ empirico-bayesiano (12) si sono ricavati i valori di tali stime puntuali dell'intercetta, che sono da ritenere significativamente diversi se, e solo se, i rispettivi intervalli di confidenza risultano disgiunti (*non-overlap criterion*). Ordinando in senso crescente le classi sulla base delle stime puntuali dell'intercetta β_{0j} possiamo fare dei confronti tra ciascuna classe e la media generale $\hat{\gamma}_{00}$ del disagio rilevato, individuando almeno tre gruppi di classi corrispondenti a tre livelli di disagio: alto, medio e basso. L'operazione di confronto è agevolata ricorrendo alla rappresentazione grafica dei risultati in un diagramma cartesiano (figura 1), che è ottenuta riportando sull'asse delle ascisse il rango occupato dalle classi in termini del disagio rilevato e sull'asse dell'ordinate il valore della differenza ($\hat{\beta}_{0j}^{EB}$ - $\hat{\gamma}_{00}$) tra gli stimatori $\hat{\beta}_{0j}^{EB}$ e $\hat{\gamma}_{00}$.

Il primo gruppo delle differenze con dei valori minori di zero, per i quali l'intervallo di confidenza non comprende la linea orizzontale ($\hat{\gamma}_{00}$), contribuisce al disagio manifestato dagli studenti nei confronti degli insegnanti in misura inferiore rispetto alla media generale del disagio e consente di individuare il gruppo delle classi migliori. Il secondo gruppo di valori vicini a zero, che corrisponde alla media generale del disagio verso gli insegnanti e per i quali l'intervallo di confidenza comprende la linea orizzontale, individua il gruppo delle classi intermedie. Infine, il terzo gruppo delle differenze con dei valori superiori a zero ed il cui intervallo di confidenza non comprende la linea orizzontale corrisponde alle classi per le quali in disagio degli studenti nei confronti degli insegnanti risulta più elevato e consente di individuare il gruppo delle classi peggiori.

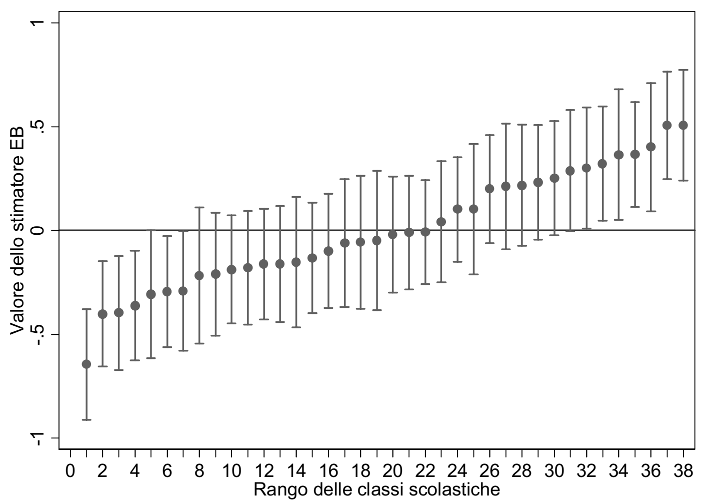

Figura 1 – Intervalli di confidenza determinati mediante il modello MIC-3 per il disagio β_{0j} manifestato dagli studenti delle classi scolastiche secondo il rango occupato da ciascuna classe nella loro graduatoria costruita in base al disagio osservato (i valori dello stimatore $\hat{\beta}_{0j}^{EB}$ sono centrati rispetto a quello dello stimatore $\hat{\gamma}_{00}$).

In altre parole, il criterio d'inserimento nell'uno o nell'altro gruppo dipende dalla collocazione dell'intervallo di confidenza intorno al valore medio assunto dallo stimatore ($\hat{\gamma}_{00}$) del disagio rilevato per tutte le classi. Tale intervallo di confidenza

$$Pr[(\hat{\beta}_{0j}^{EB} - t_{\alpha/2}\, SE_j < \beta_{0j} < \hat{\beta}_{0j}^{EB} + t_{\alpha/2}\, SE_j)] = (1-\alpha)$$

è stato determinato facendo riferimento al livello di confidenza del 95% e la sua ampiezza dipende (Longford, 1993) dai valori dello scarto quadratico medio

$$SE_j = \frac{1}{\sqrt{\tau_0^{-2} + n_j \sigma^{-2}}}$$

dello stimatore $\hat{\beta}_{0j}^{EB}$ ed è funzione delle componenti σ^2 e τ_0^2 della varianza del disagio e della numerosità n_j delle classi. Notiamo (figura 1) che l'ampiezza dell'intervallo di confidenza non è uguale per tutte le classi, in particolare, poiché le componenti di varianza sono uguali per tutte le classi, quanto minore è la numerosità n_j della classe, tanto maggiore è l'ampiezza dell'intervallo di confidenza. Questo fatto rende più probabile l'inserimento della classe povera d'informazione (poco numerosa) nel gruppo delle classi intermedie, ed evita una sua sottovalutazione (inserimento nelle classi peggiori) o una sua sopravvalutazione (inserimento nelle classi migliori) in termini di disagio scolastico.

Passando a considerare il problema del calcolo delle stime dei parametri incogniti dei modelli lineari applicati per lo studio del disagio degli studenti (tabella 1), esso è stato effettuato mediante il programma di elaborazione statistica MLWIN (Rasbash *et al.*, 2000) secondo due diverse procedure: la procedura IGLS (*Iterative Generalized Least Squares*; Goldstein, 1986) e quella RIGLS (*Restricted Generalized Least Squares*; Goldstein, 1989).

TABELLA 4

Valori ottenuti con i procedimenti IGLS *e* RIGLS $^{\Delta}$ *per gli stimatori dei parametri di alcuni modelli ad intercetta casuale*

Modelli	Stimatori	$\hat{\gamma}_{00}$	$\hat{\gamma}_{10}$	$\hat{\gamma}_{20}$	$\hat{\gamma}_{30}$	$\hat{\tau}_0^2$	$\hat{\sigma}^2$
MIC-0	igls	2,162** (0,093)	-	-	-	0,159* (0,074)	2,445** (0,147)
	rigls	2,162** (0,094)	-	-	-	0,168* (0,076)	2,445** (0,147)
MIC-1	igls	2,160** (0,094)	0,120** (0,040)	-	-	0,169* (0,076)	2,402** (0,144)
	rigls	2,160** (0,095)	0,120** (0,040)	-	-	0,177* (0,078)	2,406** (0,144)
MIC-2	igls	2,162** (0,092)	0,119** (0,039)	0,479** (0,130)	-	0,163* (0,073)	2,349** (0,141)
	rigls	2,162** (0,093)	0,120** (0,039)	0,479** (0,130)	-	0,171* (0,076)	2,358** (0,141)
MIC-3	igls	2,161** (0,089)	0,122** (0,039)	0,412** (0,130)	0,257** (0,064)	0,149* (0,069)	2,292** (0,138)
	rigls	2,161** (0,091)	0,122** (0,039)	0,412** (0,130)	0,257** (0,064)	0,157** (0,071)	2,303** (0,138)

$^{\Delta}$ *tra parentesi è riportato l'errore standard asintotico della stima del parametro.*
* *valore significativo al livello di probabilità P = 0,05.*
** *valore significativo al livello di probabilità P = 0,01.*

Ricordiamo che in determinate condizioni (normalità delle distribuzioni dei termini di errore del modello lineare a due livelli), le due procedure forniscono gli stessi risultati che si hanno, rispettivamente, col metodo della massima verosimiglianza completa (FML) e con quello della massima verosimiglianza ristretta (REML). Circa il confronto tra i due tipi di stimatori (FML e REML), in base ai risultati da noi ottenuti per le stime dei parametri incogniti γ, σ^2 e τ_0^2 con le procedure IGLS e RIGLS (tabella 4), si nota che essi non forniscono dei valori sensibilmente diversi per la stima dei parametri γ e σ^2, mentre si registra una sottostima (che si aggira intorno al 5%) dei valori ottenuti con la procedura IGLS per la componente τ_0^2 della varianza del disagio tra le classi scolastiche, a differenza di quanto risulta dalla letteratura sull'argomento e cioè che quando il numero N delle unità di secondo livello è superiore a 30 i due tipi di stimatori forniscono praticamente gli stessi risultati per tutti i parametri menzionati. Inoltre, vale la pena di rilevare che l'inserimento di una variabile esplicativa individuale x nel modello modifica i valori delle stime delle componenti di varianza σ^2 e τ_0^2 e di conseguenza quelli del coefficiente di correlazione intraclasse (tabella 5). In generale, si è osservato che l'inserimento di una variabile esplicativa (x) comporta la diminuzione del valore della stima della componente σ^2 della varianza entro le classi ed una variazione di segno positivo o negativo del valore della stima del coefficiente di correlazione intraclasse e di quello della componente τ_0^2 della varianza fra le classi in funzione delle dissomiglianze/somiglianze esistenti tra le classi rispetto alla variabile individuale x rilevata su ogni studente.

TABELLA 5

Valori delle stime del coefficiente di correlazione ($\rho|x$) secondo i modelli MIC e i procedimenti di stima adottati per lo studio del disagio scolastico

Modelli	MIC-0		MIC-1		MIC-2		MIC-3	
Procedimenti	IGLS	RIGLS	IGLS	RIGLS	IGLS	RIGLS	IGLS	RIGLS
$\hat{\rho}\|x$	0,061	0,064	0,066	0,068	0,065	0,067	0,061	0,064

In particolare, i valori delle stime dei parametri γ, σ^2 e τ_0^2 sono risultati significativamente diversi da zero e si può ritenere che le variabili esplicative relative al disagio che lo studente vive in famiglia (x_{1ij}), il sesso (x_{2ij}) e il disagio manifestato dallo studente verso la scuola nel suo complesso (x_{3ij}) contribuiscono a spiegare la componente (σ^2) della variabilità disagio manifestato dagli studenti nei confronti degli insegnanti all'interno delle classi e quella (τ_0^2) tra i valori del disagio (β_{0j}) manifestato dalle classi.

Per quanto riguarda il confronto tra lo stimatore ordinario $\hat{\beta}_{0j}^{ORD}$ e quello empirico-bayesiano $\hat{\beta}_{0j}^{EB}$, notiamo che i valori di questo stimatore sono ottenuti attraverso l'equazione (12) a partire dallo stimatore ordinario $\hat{\beta}_{0j}^{ORD}$, che in pratica può essere calcolato indifferentemente con la procedura IGLS o con quella RIGLS. In base ai risultati ottenuti per questi due stimatori (tabella 6), si osserva che le differenze esistenti tra le coppie di valori dipendono dal modello adottato e dal numero degli studenti appartenenti alle classi.

TABELLA 6

Valori della stima ordinaria ed emprico-bayesiana del disagio β_{0j} delle classi secondo i modelli MIC e gli stimatori ($\hat{\beta}_{0j}$) usati per lo studio del disagio scolastico

		MIC-0		MIC-1		MIC-2		MIC-3	
N. Classe	n_j	$\hat{\beta}_{0j}^{ORD}$	$\hat{\beta}_{0j}^{EB}$	$\hat{\beta}_{0j}^{ORD}$	$\hat{\beta}_{0j}^{EB}$	$\hat{\beta}_{0j}^{ORD}$	$\hat{\beta}_{0j}^{EB}$	$\hat{\beta}_{0j}^{ORD}$	$\hat{\beta}_{0j}^{EB}$
1	10	3,100	2,544	3,075	2,548	3,011	2,519	3,150	2,562
2	19	2,010	2,076	1,943	2,036	1,907	2,014	1,925	2,028
3	19	2,000	2,070	1,939	2,034	1,941	2,034	1,874	1,999
4	14	1,607	1,890	1,537	1,852	1,596	1,877	1,563	1,869
5	13	2,923	2,521	2,900	2,512	2,852	2,497	2,774	2,449
6	21	1,857	1,982	1,810	1,952	1,793	1,939	1,842	1,973
7	11	2,654	2,374	2,706	2,397	2,766	2,430	2,656	2,373
8	17	2,030	2,091	2,027	2,088	1,997	2,071	1,978	2,063
9	13	2,154	2,158	2,194	2,176	2,257	2,208	2,248	2,202
10	17	2,117	2,138	2,103	2,129	2,071	2,112	2,142	2,151
11	7	1,429	1,924	1,301	1,878	1,375	1,897	1,486	1,943
12	13	2,885	2,503	2,793	2,461	2,879	2,510	2,802	2,462
13	17	2,470	2,328	2,538	2,365	2,551	2,377	2,630	2,413
14	6	2,165	2,163	2,102	2,143	2,195	2,172	1,996	2,113
15	16	2,313	2,241	2,384	2,278	2,328	2,251	2,123	2,141
16	19	3,105	2,696	3,205	2,756	3,154	2,737	3,059	2,668
17	10	2,120	2,145	2,065	2,121	1,984	2,087	2,011	2,100
18	18	1,779	1,950	1,779	1,948	1,798	1,956	1,440	1,764
19	12	1,666	1,938	1,629	1,918	1,693	1,944	1,692	1,950
20	13	2,616	2,376	2,582	2,361	2,543	2,347	2,625	2,379
21	17	1,705	1,916	1,681	1,900	1,744	1,931	1,826	1,981
22	17	2,942	2,582	2,912	2,568	2,946	2,595	2,761	2,483
23	9	2,445	2,270	2,445	2,270	2,365	2,242	2,432	2,264
24	19	1,369	1,713	1,422	1,739	1,536	1,799	1,640	1,867
25	23	2,782	2,542	2,747	2,522	2,746	2,527	2,760	2,527
26	21	3,000	2,657	3,013	2,667	2,970	2,65	3,020	2,667
27	20	1,450	1,750	1,371	1,700	1,420	1,723	1,533	1,799
28	9	1,557	1,931	1,608	1,947	1,661	1,964	1,759	2,008
29	10	1,502	1,893	1,401	1,848	1,339	1,816	1,404	1,854
30	19	0,959	1,481	0,976	1,485	0,959	1,465	1,016	1,515
31	16	1,876	2,012	1,927	2,037	1,972	2,06	1,850	1,999
32	8	2,001	2,105	1,976	2,094	1,930	2,077	1,999	2,104
33	9	3,167	2,546	3,300	2,600	3,195	2,570	3,121	2,526
34	17	2,353	2,265	2,446	2,315	2,575	2,390	2,593	2,393
35	24	2,125	2,139	2,208	2,190	2,296	2,247	2,325	2,263
36	24	2,125	2,139	2,126	2,139	2,116	2,133	2,146	2,152
37	23	1,435	1,717	1,371	1,674	1,417	1,696	1,501	1,758
38	22	2,546	2,393	2,612	2,434	2,373	2,292	2,493	2,360

In generale si può dire che per ogni modello lo scarto quadratico medio

$$\overline{\sigma} = \sqrt{\frac{\sum_{j=1}^{N}(\hat{\beta}_{0j} - \overline{\beta})^2 n_j}{\sum_{j=1}^{N} n_j}}$$

delle stime ottenute con lo stimatore ordinario $\hat{\beta}_{0j}^{ORD}$ ha assunto un valore quasi doppio di quello che caratterizza lo stimatore empirico-bayesiano $\hat{\beta}_{0j}^{EB}$ (tabella 7). Un'altra considerazione interessante che si può fare riguarda la minore sensibilità dei valori delle stime di questo secondo stimatore rispetto all'introduzione nel modello di regressione lineare delle variabili individuali *x* rilevate per gli studenti delle classi scolastiche.

TABELLA 7

Scarto quadratico medio $\overline{\sigma}$ dei valori dello stimatore ordinario ed empirico-bayesiano dell'intercetta β_{0j} dei modelli MIC

Modelli	MIC-0		MIC-1		MIC-2		MIC-3	
Stimatori	ORD	EB	ORD	EB	ORD	EB	ORD	EB
$\overline{\beta}$	2,155	2,155	2,154	2,154	2,156	2,156	2,157	2,157
$\overline{\sigma}$	0,561	0,299	0,581	0,312	0,560	0,305	0,543	0,288

Tuttavia, a prescindere dal modello adottato la numerosità n_j delle classi, che figura implicitamente nell'espressione (12) dello stimatore empirico-bayesiano $\hat{\beta}_{0j}^{EB}$ attraverso il coefficiente λ_j, mostra un effetto abbastanza evidente sui valori delle stime del disagio scolastico β_{0j} delle classi, cosicchè verosimilmente tale stimatore è da preferire rispetto allo stimatore ordinario $\hat{\beta}_{0j}^{ORD}$, perchè esso tiene conto di alcune informazioni rilevanti, quali le numerosità n_j delle singole classi e delle componenti di varianza tra e dentro le classi attraverso il peso λ_j (*shrinkage factor*), il cui impiego comporta anche una diminuzione della variabilità delle stime del parametro β_{0j}, che rappresenta un indicatore di efficacia delle classi in termini di disagio scolastico.

In particolare, è facile intuire che quanto maggiori sono le differenze tra le classi rispetto alla variabile risultato *Y*, tanto minore è la riduzione della variabilità dei valori dello stimatore ordinario dell'intercetta. In altre parole, se il valore stimato della componente τ_0^2 della varianza tra classi fosse molto più grande di quello stimato per la componente σ^2 della varianza dentro le classi (ad esempio $\hat{\tau}_0^2 > 100\,\hat{\sigma}^2$), allora la stima del coefficiente ρ di correlazione intraclasse ed il peso λ_j sarebbero quasi uguale ad uno, facendo coincidere i valori dello stimatore empirico-bayesiano $\hat{\beta}_{0j}^{EB}$ e di quello ordinario $\hat{\beta}_{0j}^{ORD}$ dell'intercetta. In questo caso, anche i valori degli scarti quadratici medi ($\overline{\sigma}$) dei valori degli stimatori ordinario ed empirico-bayesiano coinciderebbero, ma in pratica ciò non si verifica per le ricerche nel campo dell'istruzione, perchè in questo caso (Snijders e Bosker, 1999) generalmente il coefficiente di correlazione intraclasse assume un valore piuttosto piccolo (compreso tra i limiti 0,05 e 0,20).

Per i diversi modelli ad intercetta casuale applicati, lo stimatore del coefficiente ρ di correlazione intraclasse (tabella 5) ha assunto un valore medio pari a 0,065 e i pesi λ_j hanno assunto un valore medio pari a 0,5. Questa circostanza consente di dire che mediamente il valore dello stimatore $\hat{\beta}_{0j}^{EB}$ empirico-bayesiano relativo a ciascuna classe dipende per il 50% dallo stimatore ordinario $\hat{\beta}_{0j}^{ORD}$ dell'intercetta e per il restante 50% dallo stimatore $\hat{\gamma}_{00}$ della media generale della variabile dipendente *Y*.

Volendo mettere in evidenza il ruolo delle variabili individuali dello studente nella misura del disagio delle classi nei confronti degli insegnanti, si possono considerare i valori delle stime dell'intercetta nel caso del modello ad intercetta casuale privo di variabili esplicative individuali e nel caso del modello ad intercetta casuale che include tutte le variabili esplicative individuali prese in esame. A partire da tali valori è utile considerare la variazione percentuale

$$\text{VP} = \frac{\hat{\beta}_{0j}^{MIC-3} - \hat{\beta}_{0j}^{MIC-0}}{\hat{\beta}_{0j}^{MIC-0}} \bullet 100$$

tra le coppie di valori delle stime dell'intercetta che si hanno passando dal modello più semplice (MIC-0) a quello più complesso (MIC-3). A tale proposito si rileva (figura 2) che per tutti e due gli stimatori, ordinario ed empirico-bayesiano, l'uso delle variabili individuali incide in misura più o meno consistente sulla valutazione del disagio manifestato dalle singole classi verso gli insegnanti. Inoltre, è abbastanza evidente che la variazione è relativamente più accentuata per i valori delle stime ottenute ricorrendo allo stimatore ordinario $\hat{\beta}_{0j}^{ORD}$ e che in tutti e due i casi l'uso delle variabili individuali modifica la graduatoria delle classi in funzione del disagio stimato $\hat{\beta}_{0j}$.

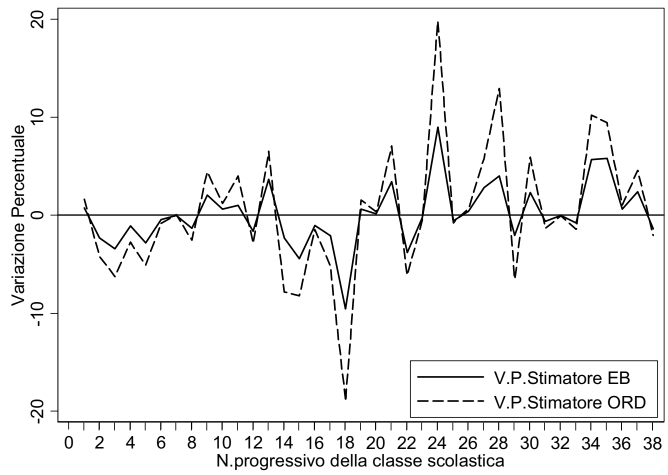

Figura 2 – Variazione percentuale per classe scolastica dei valori delle stime dell'intercetta β_{0j} che si ha passando dal modello MIC-0 (senza variabili individuali) al modello MIC-3 (con tre variabili individuali) per lo stimatore ordinario $\hat{\beta}_{0j}^{ORD}$ e per quello empirico-bayesiano $\hat{\beta}_{0j}^{EB}$.

Pertanto, non è consigliabile confrontare le classi in termini di disagio scolastico senza tenere conto delle disomogeneità esistenti al loro interno con riferimento alle caratteristiche individuali degli studenti. Infatti, considerare o non considerare le variabili esplicative relative allo studente nella valutazione del disagio delle classi verso gli insegnanti, comporta una variazione della stima dell'intercetta (che raggiunge talvolta il 20% per lo stimatore ordinario e il 10% per lo stimatore empirico bayesiano) e può modificare la posizione delle classi stesse nella loro graduatoria secondo il valore stimato dell'intercetta, fino al punto di far assegnare qualche classe ad un gruppo diverso tra quelli individuati come migliore, intermedio o peggiore in termini del disagio manifestato verso gli insegnanti.

6. CONCLUSIONI

Dopo una breve descrizione delle caratteristiche principali del modello multilivello di regressione lineare con riferimento al caso di due livelli con intercetta casuale ed ai problemi di stima dei parametri incogniti, è stata proposta una versione del modello per lo studio del disagio manifestato dagli studenti nei confronti degli insegnanti, con l'intento di effettuare una valutazione di efficacia relativa delle classi in termini di disagio scolastico, considerando che le caratteristiche individuali degli studenti possono essere diverse a seconda della classe di appartenenza e che esse possono concorrere a determinare il disagio manifestato da una determinata classe.

Dai risultati ricavati attraverso l'impiego del modello prescelto ai dati rilevati nell'ambito di una ricerca sul campo è emerso che le variabili individuali prese in esame (il disagio dello studente in famiglia, il sesso e il disagio verso la scuola nel suo complesso) sono significative nel determinare il disagio manifestato dagli studenti verso i propri insegnanti e, pertanto, che un confronto di efficacia relativa tra le classi in termini di disagio scolastico non può essere effettuato senza tenere conto della disomogeneità esistente tra le classi rispetto alle caratteristiche individuali degli studenti.

Inoltre, il procedimento sperimentato per valutare il disagio manifestato dagli studenti verso gli insegnanti è apparso adeguato ai fini dello studio, perchè lo stimatore dell'intercetta può esssere ritenuto un indicatore valido per effettuare un confronto tra le classi in condizioni di parità rispetto alle variabili individuali considerate. Esso ha permesso di costruire una graduatoria delle classi scolastiche, in base alla quale sono stati individuati tre gruppi (migliore, intermedio e peggiore) di classi in base al disagio stimato attraverso uno stimatore empirico-bayesiano.

Per quanto riguarda le caratteristiche di questo e altri stimatori utili ai fini della stima dei parametri incogniti del modello con intercetta casuale, la sua applicazione ha consentito di mettere in evidenza che, come accade spesso per altri tipi di dati, anche per quelli presi in esame gli stimatori di massima verosimiglianza completa (FML) e ristretta (REML) hanno fornito praticamente gli stessi risultati nella stima dei coefficienti di regressione e della componente σ^2 della varianza, che caraterizzano il modello adottato. Invece, a differenza di quanto risulta dalla letteratura sull'argomento nel caso di un numero di classi (unità di secondo livello) maggiore di 30, alcune differenze non trascurabili si sono riscontrate con riferimento al problema della stima della componente τ_0^2 della varianza del disagio tra le classi, perchè col metodo della massima verosimiglianza completa (FML) questa componente è apparsa sottostimata del 5%.

Infine, dal confronto tra lo stimatore ordinario di massima verosimglianza ristretta (REML) e lo stimatore empirico-bayesiano, che è stato effettuato attraverso i valori delle stime del parametro relativo al disagio osservato per le diverse classi scolastiche, si può dire che il secondo stimatore può essere preferito per la sua sensibilità rispetto alla numerosità delle classi scolastiche e per la sua efficienza in termini statistici.

Dipartimento di Statistica e Matematica applicata all'Economia GILBERTO GHILARDI
Università degli Studi di Pisa

Dipartimento di Statistica e Matematica applicata all'Economia NICOLA ORSINI
Università degli Studi di Pisa

RIFERIMENTI BIBLIOGRAFICI

M. AITKIN, N. LONGORD (1986), *Statistical Modelling Issues in School Effectiveness Studies,* "Journal of the Royal Statistical Society", A, vol.149, part. 1, pp. 1-43.

A.S. BRYK, S.W. RAUDENBUSH (1992), *Hierarchical Linear Models:Applications and data analysis methods,* Sage Publications, Newbury Park, CA.

B. EFRON, C.N. MORRIS (1975), *Data analysis using Stein's estimator and its generalizations*, "Journal of the American Statistical Association", 74, pp. 311-319.

H. GOLDSTEIN (1986), *Multilevel mixed linear model analysing using iterative generalised least squares*, "Biometrika", 73, pp. 43-56.

H. GOLDSTEIN (1989), *Restricted unbiased iterative generalised least squares estimation*, "Biometrika", 78, pp. 622-623.

H. GOLDSTEIN (1995), *Multilevel Statistical Models,* 2nd edition, Edward Arnold, London.

H. GOLDSTEIN, M.J.R. HEALY (1995), *The graphical presentation of a collection of means*, "Journal of the Royal Statistical Society", A, vol. 158, pp. 175-177.

H. GOLDSTEIN, D.J. SPIEGELHALTER (1996), *League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance*, "Journal of the Royal Statistical Society", A, vol. 159, pp. 385-443.

H. GOLDSTEIN (1997), *Methods in school effectiveness research*, "School effectiveness and school improvement", 8, pp. 369-95.

J.J. HOX (1995), *Applied Multilevel Analysis*, TT-Publikaties, Amsterdam.

I.G.G. KREFT, J. DE LEEUW (1998), *Introducing Multilevel Modeling*, Sage Publications, London.

D.V. LINDLEY, A.F.M. SMITH (1972), *Bayes estimates for the linear model*, "Journal of the Royal Statistical Society", B, 34, pp. 1-41.

N.T. LONGFORD (1993), *Random Coefficient Models,* Clarendon Press, Oxford.

J. NETER, W. WASSERMANN (1974), *Applied Linear Statistical Models*, Richard D.Irwin, Inc. Illinois.

J. RASBASH, W. BROWNE, H. GOLDSTEIN, M. YANG, I. PLEWIS, D. DRAPER, M.J.R. HEALY, G. WOODHOUSE (2000), *A User's Guide to MLwiN*, version 2.0.008, Institute of Education, London.

S.R. SEARLE, G. CASELLA, C.E. MCCULLOCH (1992), *Variance components*, Wiley Series in Probability and Mathematical Statistics, New York.

T. SNIJDERS, R. BOSKER (1999), *Multilevel Analysis. An introduction to Basic and Advanced Multilevel Modellimg*, Sage Publications, London.

RIASSUNTO

Modelli lineari ad intercetta casuale e stimatori per la valutazione dei sistemi formativi

Questo lavoro presenta due stimatori dell'intercetta, uno ordinario e l'altro empirico-bayesiano, per un modello lineare multilivello di regressione, dove l'intercetta è considerata una variabile casuale. Le stime di tale parametro possono essere impiegate come indicatori di efficacia nella valutazione delle classi scolastiche, le quali differiscono per le caratteristiche individuali degli studenti e per il disagio che esse manifestano nei confronti degli insegnanti. Le differenze tra le dimensioni delle classi e il valore del coefficiente di correlazione intraclasse evidenziano la maggiore efficienza dello stimatore empirico-bayesiano rispetto allo stimatore ordinario dell'intercetta.

SUMMARY

Random intercept linear models and estimators for the evaluation of educational systems

This papers presents two intercept estimators, an ordinary one and an empirical Bayesian one, for a multilevel linear regression model where the intercept is treated as a random variable. These estimators can be used as effectiveness indicators for the evaluation of the classes, which may vary in terms of the students' characteristics as well as the problems in the student/teacher relationship. The differences between class size and the intraclass correlation coefficient value, show that empirical Bayes estimator is more efficient than the ordinary estimator.