



Università degli Studi di Pisa
Dipartimento di Statistica e Matematica
Applicata all'Economia

Report n. 220

**Sull'uso dei modelli statistici lineari
nella valutazione dei sistemi formativi**

Gilberto Ghilardi Nicola Orsini

Pisa, Gennaio 2002

- Stampato in Proprio -

SULL'USO DEI MODELLI STATISTICI LINEARI NELLA VALUTAZIONE DEI SISTEMI FORMATIVI

GILBERTO GHILARDI * NICOLA ORSINI **

Parole chiave: Modelli lineari multilivello, stimatori dei minimi quadrati, indicatori di efficacia relativa.

1. INTRODUZIONE

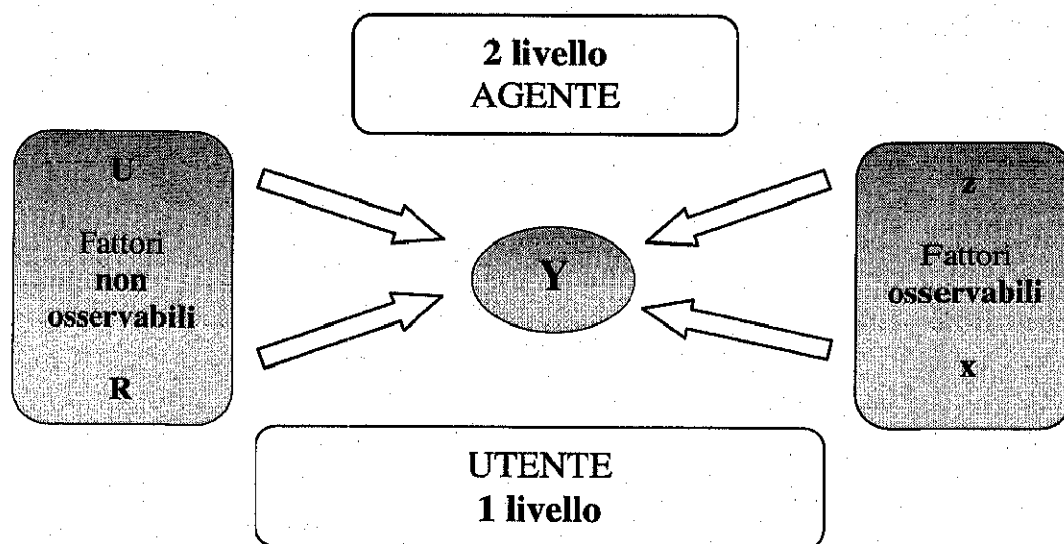
In questi ultimi anni è stato preso in esame il problema del miglioramento della formazione offerta dalle istituzioni educative, perché i processi di cambiamento tecnologico, economico e culturale intervenuti nella società moderna hanno determinato un aumento del fabbisogno formativo, specialmente dei giovani e dei disoccupati. La formazione rappresenta un aspetto importante, oltre che per la crescita personale e professionale di tutte le categorie dei lavoratori anche per la possibilità di far fronte alla domanda di un mercato del lavoro sempre più esigente e flessibile. In questo contesto la *valutazione dell'attività formativa* costituisce uno strumento essenziale per lo sviluppo delle **politiche** educative e per la **gestione** stessa degli Agenti formativi (Stame, 2001; Gori, 2000). Il presente lavoro riporta i risultati dell'applicazione di una metodologia statistica per la determinazione di indicatori sintetici idonei al confronto di Agenti formativi diversi. Tale

* Dipartimento di Statistica e Matematica Applicata all'Economia, Università di Pisa, Via Cosimo Ridolfi, 10, 56124 Pisa.

** Dipartimento di Statistica e Matematica Applicata all'Economia, Università di Pisa, Via Cosimo Ridolfi, 10, 56124 Pisa.

operazione di confronto è svolta mediante una valutazione di **efficacia relativa** (Aitkin & Longford, 1986) e deve essere condotta tenendo conto del fatto che tali Agenti formativi operano con risorse umane (studenti) diverse. Il problema indicato può essere affrontato definendo una variabile risultato Y che viene osservata sull'Utente (studente) e rispetto alla quale si effettua il confronto tra gli Agenti¹ (classi o scuole). Tale risultato Y (Fig.1.1) dipende da fattori riferibili all'Utente, che rappresenta il primo livello di osservazione e da fattori riferibili all'Agente formativo che a sua volta rappresenta invece il secondo livello di osservazione.

Figura 1.1 I fattori determinanti il risultato Y dell'Utente.



Ciò che possiamo osservare (*fattori osservabili*) è rappresentato da una o più variabili x riferite all'Utente e da una o più variabili z riferite all'Agente formativo. Invece, tutto ciò invece che non è osservabile (*fattori inosservabili o non osservati*) è considerato afferente ad un termine d'errore, indicato con la lettera R al livello dell'Utente e con la lettera U al livello dell'Agente formativo.

¹ Nelle valutazioni di efficacia relativa viene di solito scelta come variabile Y un risultato di medio / lungo periodo (*outcome*) che l'Utente raggiunge grazie all'attività svolta dall'Agente formativo (Gori & Vittadini, 1999).

2. IL MODELLO STATISTICO

Il *modello statistico lineare* adeguato per trattare simultaneamente diversi livelli di osservazione e le variabili osservate è definito in letteratura come *modello multilivello* o *gerarchico* e scaturisce dall'impiego dei metodi di analisi della varianza e di regressione (Snijders & Bosker, 1999). Esso è composto (Tab.2.1) da una equazione di *primo livello* che rappresenta la relazione tra le variabili osservate sull'Utente (Y e x) e da una o più equazioni di *secondo livello* che studiano invece le relazioni esistenti tra le variabili osservate sull'Agente (z) e i coefficienti (β_{0j} e β_{1j}) inseriti nell'equazione di primo livello. I fattori, *osservati e quelli non osservati*, riferiti all'Agente formativo contribuiscono a determinare i valori dei coefficienti di regressione dell'equazione di primo livello.

Tabella 2.1 Equazioni ed assunzioni nel modello statistico lineare a due livelli.

Livelli	Equazioni	Assunzioni	Indici
1	$Y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + R_{ij}$	$R_{ij} \sim \text{i.i.d. } N(0, \sigma^2)$	$i = 1.., n_j$
2	$\beta_{1j} = \gamma_{10} + \gamma_{11}z_{1j} + U_{1j}$ $\beta_{0j} = \gamma_{00} + \gamma_{01}z_{0j} + U_{0j}$	$\begin{bmatrix} U_{0j} \\ U_{1j} \end{bmatrix} \sim \text{i.i.d. } N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_0^2 & 0 \\ 0 & \tau_1^2 \end{bmatrix}\right)$	$j = 1.., N$
		$R_{ij} \perp \begin{bmatrix} U_{0j} \\ U_{1j} \end{bmatrix}$	$M = \sum_{j=1}^N n_j$

Fonte: Bryk & Raudenbush, 1992 pag.29

Gli indici (i e j) delle variabili e dei coefficienti indicano le unità di osservazione alle quali si riferiscono i valori delle variabili (Y_{ij} , x_{ij} , z_j) e dei parametri (β_{0j} e β_{1j}). In particolare, l'indice i individua l' i -mo Utente all'interno del j -mo Agente e varia da 1 a n_j , ossia il numero di Utenti osservati con riferimento al j -mo Agente. L'indice j indica l'Agente formativo di provenienza dell'Utente e varia da 1 a N , ossia il numero complessivo degli

Agenti formativi presi in esame. La dimensione M del campione di Utenti è data dunque dalla sommatoria delle numerosità n_j di classe. Le variabili casuali che rappresentano i termini dell'errore di osservazione e che sono associate alle equazioni di primo e secondo livello sono rispettivamente i termini R ed U . Per tali variabili si assume una distribuzione normale con valore atteso nullo ed una struttura di varianza e covarianza costante. In particolare si assume: la distribuzione identica e indipendente di ciascun termine d'errore (i.i.d.); una varianza costante, pari a σ^2 per la variabile casuale R ; una varianza costante, pari a τ_0^2 e τ_1^2 per le variabili casuali U e l'indipendenza stocastica (indicata con il simbolo \perp) tra i termini d'errore di primo (R) e di secondo (U) livello.

Il modello lineare a due livelli nel quale solo l'intercetta β_{0j} è libera di assumere valori diversi tra Agenti formativi (Tab.2.2) è conosciuto in letteratura con diversi nomi: modelli a componenti di varianza (*Variance Components Models*; Searle *et al.*, 1992); modelli di analisi della covarianza ad effetti casuali (*Covariance Models Random Effects*; Bryk & Raudenbush, 1992) e modelli ad intercetta casuale (*Random intercept Model*; Snijders & Bosker, 1999).

Tabella 2.2 Equazioni e assunzioni del modello ad intercetta casuale.

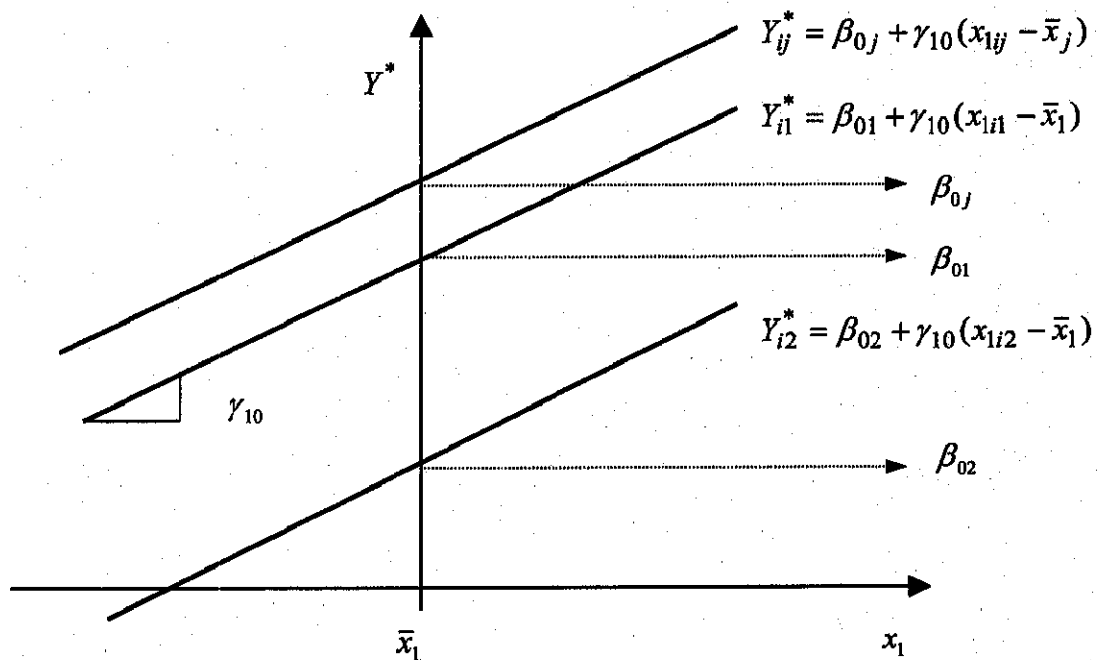
Livelli	Equazioni	Assunzioni	Indici
1	$Y_{ij} = \beta_{0j} + \beta_{1j}x_{ij} + R_{ij}$	$R_{ij} \sim \text{i.i.d. } N(0, \sigma^2)$	$i = 1.., n_j$
2	$\beta_{0j} = \gamma_{00} + \gamma_{01}z_{0j} + U_{0j}$ $\beta_{1j} = \gamma_{10}$	$U_{0j} \sim \text{i.i.d. } N(0, \tau_0^2)$	$j = 1.., N$
		$R_{ij} \perp U_{0j}$	$M = \sum_{j=1}^N n_j$

Fonte: Snijders & Bosker, 1999 pag.38

Considerare dunque l'intercetta β_{0j} come una variabile casuale implica ritenere che gli Agenti formativi (classi) osservati costituiscono un campione

casuale semplice estratto da popolazioni di Agenti (scuole) indipendenti tra di loro². Il coefficiente β_{1j} misura la relazione tra la variabile esplicativa x e la variabile risultato Y ed è tenuto costante (pari a γ_{10}) per tutte le classi osservate. Il modello ad intercetta casuale (Tab.2.2) può essere visto quindi come un insieme di modelli di regressione, uno per ogni classe osservata, dove le differenze tra le classi corrispondono a quelle esistenti tra i valori β_{0j} dell'intercetta (Fig.2.1).

Figura 2.1 Esempio di rette di regressione per un modello ad intercetta casuale.



I parametri incogniti del modello sono dunque i parametri costanti su tutte le classi indicati con la lettera γ ; i parametri relativi a ciascuna classe indicati con la lettera β_{0j} e le componenti di varianza di primo livello (σ^2) e di secondo livello (τ_0^2). Dunque in un modello ad **effetti fissi e casuali** lo studente e la classe di provenienza sono considerate due fonti di variabilità

² D'ora in poi parleremo di studenti (utenti); classi (agenti) e scuole (popolazioni di agenti).

distinte (Tab.2.3). La componente di varianza σ^2 rappresenta la variabilità dei risultati degli studenti **all'interno** delle classi (*within*) mentre la componente di varianza τ_0^2 rappresenta la variabilità dei risultati degli studenti **tra** classi (*between*). 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 intraclassa*, ossia la correlazione tra le risposte date dagli studenti all'interno della stessa classe (Longford, 1993).

Tabella 2.3 Componenti di varianza e correlazione intraclassa.

$Var(Y_{ij}) = Var(R_{ij}) + Var(U_{0j}) = \sigma^2 + \tau_0^2$	$\forall i \text{ e } \forall j$
$Cov(Y_{ij}, Y_{rj} x_{ij}, x_{rj}) = Var(U_{0j}) = \tau_0^2$	$i \neq r \text{ e } \forall j$
$\rho x = \frac{\tau_0^2}{\tau_0^2 + \sigma^2}$	$\forall i \text{ e } \forall j$

Fonte: Snijders & Bosker, 1999 pag.46

Il modello descritto è stato oggetto di un'applicazione con i **dati** ottenuti nell'ambito di un progetto finalizzato allo studio del disagio scolastico da parte Consorzio F.I.T. (Formazione Innovazione e Tecnologie) su richiesta del Comune di Massa. Questo progetto ha coinvolto 592 studenti delle scuole elementari provenienti da 38 classi del distretto scolastico. La variabile Y presa in esame è rappresentata dal disagio manifestato dallo studente nei confronti degli insegnanti ed in base ad essa è stata effettuata una valutazione delle classi. In questa operazione di confronto si è tenuto conto di alcune caratteristiche individuali dello studente, quali il disagio che lo studente vive in famiglia, il sesso e il disagio manifestato dallo studente verso la scuola nel suo complesso. Purtroppo l'assenza di variabili esplicative al livello di classe, ossia delle variabili z, non ci ha permesso di tener conto di alcuni fattori rilevanti, quali ad esempio il metodo d'insegnamento.

Mediante il programma di elaborazione statistica SPSS (*Statistical Package for Social Sciences*) abbiamo stimato con il metodo dei minimi

quadrati ordinari quattro diversi modelli ad intercetta casuale (Tab.2.4), considerando di volta in volta un numero differente di variabili esplicative individuali rilevate sugli studenti.

Tabella 2.4 Modelli ad intercetta casuale con diverse variabili esplicative.

Variabili esplicative individuali inserite	Modelli ad intercetta casuale $\beta_{0j} = \gamma_{0j} + U_{0j}$
0	$Y_{ij} = \beta_{0j} + R_{ij}$
1 Disagio in famiglia	$Y_{ij} = \beta_{0j} + \gamma_{10}\hat{x}_{1ij} + R_{ij}$
2 Sesso	$Y_{ij} = \beta_{0j} + \gamma_{10}\hat{x}_{1ij} + \gamma_{20}\hat{x}_{2ij} + R_{ij}$
3 Disagio a scuola	$Y_{ij} = \beta_{0j} + \gamma_{10}\hat{x}_{1ij} + \gamma_{20}\hat{x}_{2ij} + \gamma_{30}\hat{x}_{3ij} + R_{ij}$

I valori delle variabili esplicative sono stati centrati (*centering*) sul rispettivo valor medio per semplificare l'interpretazione da dare alla stima del parametro β_{0j} relativo all'intercetta (Neter & Wassermann, 1974; Bryk & Raudenbush, 1992). In tal modo è possibile osservare il contributo apportato dall'uso di diverse variabili esplicative individuali ai fini del confronto tra le classi prese in esame.

3. STIMATORI DEI MINIMI QUADRATI

Per descrivere il metodo con il quale il programma di elaborazione statistica (SPSS o SAS) tratta i dati per la stima dei parametri incogniti del modello ad intercetta casuale, occorre formalizzare un modello lineare di regressione multipla in cui sono presenti contemporaneamente variabili quantitative e qualitative.

L'appartenenza o meno di uno studente ad una delle N classi osservate è codificata nel modello attraverso $N-1$ variabili binarie di comodo (*dummy variables*). La considerazione di $N-1$ variabili binarie, anzichè N , tante quante

sono le classi, è la soluzione più utilizzata per rimediare al problema della non unicità degli stimatori dei coefficienti di regressione nella situazione descritta.

La formalizzazione del modello ad intercetta casuale può essere scritta in maniera completa mediante l'equazione (3.1), o in termini matriciali in base all'equazione (3.2), con una simbologia il cui significato è riportato in tabella (3.1), dove al posto dei parametri β_{0j} e β_{1j} si considera il secondo membro delle equazioni (3.3) e (3.4), nelle quali figurano i parametri α , γ e λ .

$Y_{ij} = \alpha_{00}x_0 + \gamma_{10}c_{ij} + \lambda_1x_1 + \lambda_2x_2 + \dots + \lambda_{N-1}x_{N-1} + R_{ij}$	(3.1)
<p>► <i>L'equazione del modello in termini matriciali</i></p> $ \begin{matrix} Y & = & X & \beta & + & R \\ M \times 1 & & M \times p & p \times 1 & & M \times 1 \end{matrix} $ <p style="text-align: right;">(3.2)</p> $ Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_M \end{bmatrix} \quad X = \underbrace{\begin{bmatrix} 1 & c_{11} & x_{11} & x_{12} & \dots & x_{1,N-1} \\ 1 & c_{12} & x_{21} & x_{22} & \dots & x_{2,N-1} \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot \\ 1 & c_{1M} & x_{M1} & x_{M2} & \dots & x_{M,N-1} \end{bmatrix}}_{p=[1+q+(N-1)]} \quad \beta = \begin{bmatrix} \alpha_{00} \\ \gamma_{10} \\ \lambda_1 \\ \cdot \\ \cdot \\ \lambda_{N-1} \end{bmatrix} \quad R = \begin{bmatrix} R_1 \\ R_2 \\ \cdot \\ \cdot \\ \cdot \\ R_M \end{bmatrix} $ <p style="text-align: center;"> $q = \text{numero di var. esplicative}$ </p> <p>► <i>Le assunzioni di indipendenza e omoschedasticità</i></p> $ \begin{matrix} \sigma^2(R_i) & = & \sigma^2 I \\ M \times M & & M \times M \end{matrix} $	

$$\beta_{0j} = \alpha_{00}x_0 + \lambda_1x_1 + \lambda_2x_2 + \dots + \lambda_{N-1}x_{N-1} \quad (3.3)$$

$$\beta_{1j} = \gamma_{10} \quad (3.4)$$

Tabella 3.1 Significato dei simboli del modello ad intercetta casuale con una variabile esplicativa nella formalizzazione completa.

Simboli	Significato
Y_{ij}	variabile risultato o dipendente osservata sull' i -mo studente appartenente alla j -ma classe
α_{00}	intercetta che esprime una costante per tutti gli M studenti
γ_{10}	parametro che esprime la relazione tra Y e x , costante su tutte le classi
λ_j	parametro della j -ma variabile binaria
x_0	costante che assume il valore uguale a 1 e che rappresenta il coefficiente dell'intercetta
$c_{ij} = (x_{1ij} - \bar{x}_1)$	variabile indipendente (covariata) centrata sul valore medio e osservata sull' i -mo studente appartenente alla j -ma classe
x_j	variabile binaria che assume valore 1 se l' i -mo studente appartiene alla j -ma classe e 0 altrimenti
R_{ij}	variabile casuale su cui si fanno determinate assunzioni (Tab.2.2)

I nuovi parametri incogniti α_{00} , γ_{10} e λ_j vengono stimati con il metodo dei minimi quadrati ordinari (*Ordinary Least Squares*). Con riferimento all'equazione del modello in termini matriciali (3.2) indichiamo con b ,

$$b = \begin{bmatrix} \hat{\alpha}_{00} \\ \hat{\gamma}_{10} \\ \hat{\lambda}_1 \\ \cdot \\ \cdot \\ \hat{\lambda}_{N-1} \end{bmatrix} \quad (3.4)$$

il vettore delle stime $\hat{\alpha}_{00}$, $\hat{\gamma}_{10}$, $\hat{\lambda}_1, \dots, \hat{\lambda}_{N-1}$ dei parametri rispettivi.

Dato il sistema di *equazioni normali* in forma matriciale,

$$\begin{matrix} (X'X) & b & = & X' & Y \\ p \times p & p \times 1 & & p \times M & M \times 1 \end{matrix} \quad (3.5)$$

si possono ricavare gli stimatori

$$\begin{matrix} b & = & (X'X)^{-1} & X' & Y \\ p \times 1 & & p \times p & & p \times 1 \end{matrix} \quad (3.6)$$

dei minimi quadrati ordinari dei **coefficienti di regressione**.

In questo modello di regressione quando sono valide le assunzioni fatte (Tab.2.2) gli stimatori dei minimi quadrati coincidono con quelli di massima verosimiglianza (Neter & Wassermann, 1974, pag.226) e godono di alcune proprietà come la correttezza, l'efficienza e la consistenza.

Per quanto riguarda la stima dei minimi quadrati della varianza σ^2 del termine di errore R di primo livello e la varianza τ_0^2 del termine d'errore U di secondo livello dobbiamo far riferimento alla scomposizione nella devianza tra e dentro le classi della somma dei quadrati degli scarti della variabile Y dalla propria media. Il punto di partenza è la scomposizione della somma dei quadrati di un modello di analisi della varianza (o modello ad intercetta casuale privo di variabili esplicative), dove, com'è noto, si considerano solo le caratteristiche della variabile risultato (Y). L'introduzione di una o più variabili esplicative (x) nel modello modifica³ le somme dei quadrati tra e dentro le classi della variabile Y ed è importante conoscere l'effetto che questa operazione ha sulle componenti di varianza.

In particolare, l'inserimento di una variabile esplicativa individuale (x) comporta generalmente la diminuzione della componente di varianza σ^2 di primo livello (dentro le classi) ed una variazione della componente di varianza τ_0^2 di secondo livello (fra classi). Questo secondo effetto positivo/negativo dipende dalle dissomiglianze/somiglianze tra classi portate alla luce dall'inserimento della variabile esplicativa x individuale.

³ Per le formule si veda Neter & Wassermann (1974, pag.706) e Searle (1997, pag.340).

Per comprendere quest'ultimo punto si può osservare che (Neter & Wassermann, 1974, pp.429-447) il valore atteso della varianza dell'errore tra le classi (*MSB*),

$$E(MSB) = E(MSE) + E \left[\frac{\sum_{j=1}^N n_j (\hat{\beta}_{0j} - \hat{\gamma}_{00})^2}{N-1} \right] = E \left[\sigma^2 + \frac{\sum_{j=1}^N n_j (\hat{U}_{0j})^2}{N-1} \right] \quad (3.7)$$

$$= E [\sigma^2 + \tilde{n} \hat{\tau}_0^2] = \sigma^2 + \tilde{n} \tau_0^2$$

è dato dalla somma tra la componente di varianza σ^2 di primo livello ed un termine che è funzione della differenza \hat{U}_{0j} tra la stima $\hat{\beta}_{0j}$ dell'intercetta e la stima $\hat{\gamma}_{00}$ della media generale della variabile risultato. Inoltre è possibile dimostrare (Neter & Wassermann, 1974, pag. 528) che la stima della **componente di varianza tra classi** (*between*) del parametro,

$$\tau_0^2 = \frac{E(MSB) - E(MSE)}{\tilde{n}} \quad (3.8)$$

è data da un rapporto nel quale il denominatore

$$\tilde{n} = \frac{1}{N-1} \left[\sum n_j - \frac{\sum n_j^2}{\sum n_j} \right]$$

dipende dal numero di unità appartenenti alle diverse classi. Possiamo notare dall'equazione (3.7) che il valore della stima $\hat{\tau}_0^2$ della componente di varianza tra le classi dipende dagli scostamenti delle stime $\hat{\beta}_{0j}$ dell'intercetta rispetto alla stima $\hat{\gamma}_{00}$ della media generale della variabile risultato. Come si vedrà più avanti, lo stimatore dei minimi quadrati dell'intercetta comprende il valore dello scostamento tra il valore medio \bar{x}_j , valor medio di classe, e la media generale \bar{x} della variabile esplicativa. Pertanto, quanto maggiore è il valore dello scostamento di ciascuna classe (dissomiglianze tra le classi), tanto maggiore sarà l'aumento della componente di varianza fra classi (effetto positivo sulla varianza tra classi). Una volta stimati i parametri incogniti del

modello possiamo determinare le rette di regressione associate a ciascuna classe. Ad esempio, per la j -ma classe tale retta è data dall'equazione

$$\begin{aligned}\hat{Y}_{ij} &= \hat{\alpha}_{00} + \hat{\gamma}_{10}c_{ij} + \hat{\lambda}_1(0) + \dots + \hat{\lambda}_j(1) + \dots + \hat{\lambda}_{N-1}(0) = \\ &= (\hat{\alpha}_{00} + \hat{\lambda}_j(1)) + \hat{\gamma}_{10}c_{ij}\end{aligned}$$

dalla quale si ricava l'interpretazione da dare alla stima $\hat{\lambda}_j$ del parametro che indica di quanto è maggiore la media delle risposte della j -ma classe rispetto alla N -ma classe, ossia la classe che non è stata considerata mediante le variabili binarie x_j di comodo. Inoltre, possiamo definire il valore $\hat{\alpha}_{00}$

$$\hat{\alpha}_{00} = \hat{\gamma}_{00} + \hat{U}_{0N} \quad (3.9)$$

come somma tra la media generale $\hat{\gamma}_{00}$ e il residuo $(\hat{\beta}_{0j} - \hat{\gamma}_{00})$ riferita alla N -ma classe; il valore $\hat{\lambda}_j$

$$\hat{\lambda}_j = \hat{U}_{0j} - \hat{U}_{0N} \quad (3.10)$$

come lo scarto tra il residuo della j -ma classe e il residuo della N -ma classe; il valore $\hat{\gamma}_{00}$

$$\hat{\gamma}_{00} = \sum_{j=1}^N w_j \bar{Y}_j \quad (3.11)$$

come media pesata delle medie \bar{Y}_j di ciascuna classe, dove i pesi w_j sono rappresentati dalla frequenza relativa

$$w_j = \frac{n_j}{M}$$

del numero degli studenti per ciascuna classe.

Quest'ultima definizione comporta la proprietà di rendere pari a zero la somma ponderata

$$\sum_{j=1}^N w_j \hat{U}_{0j} = 0 \quad (3.12)$$

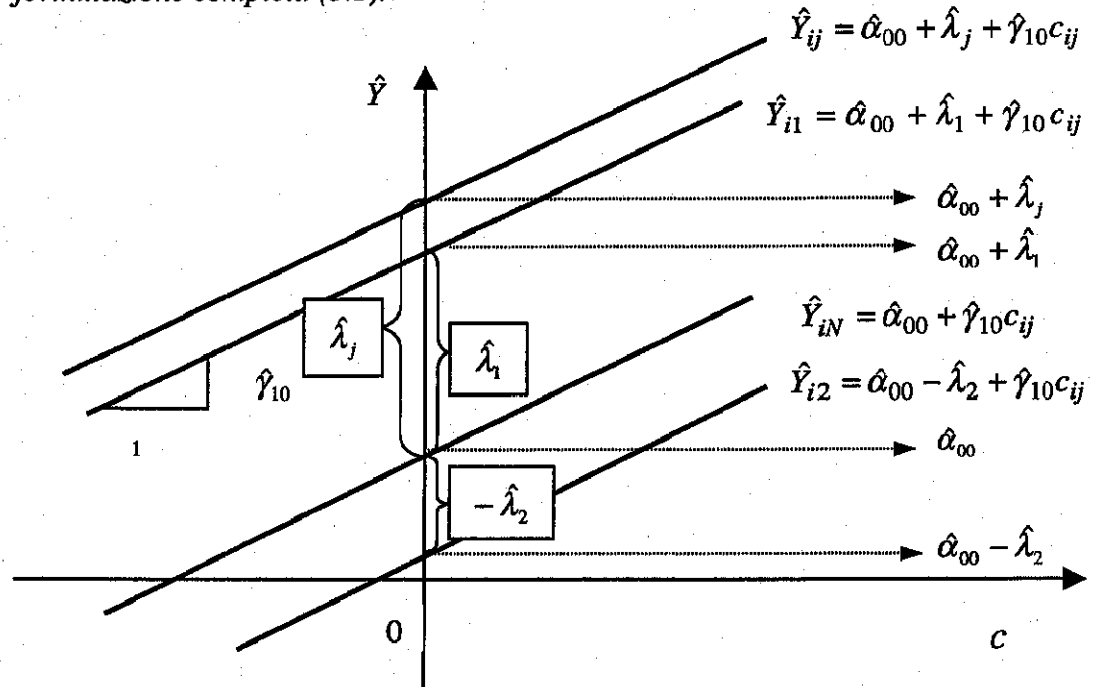
dei residui \hat{U}_{0j} di classe e conserva l'unicità delle soluzioni del sistema di

equazioni⁴ (3.5). La formalizzazione riportata per la descrizione del modello ad intercetta casuale e l'interpretazione dei coefficienti stimati può essere sintetizzata brevemente mediante l'espressione della retta di regressione

$$\begin{aligned}\hat{Y}_{ij} &= \hat{\alpha}_{00} + \hat{\lambda}_j + \hat{\gamma}_{10}c_{ij} = (\hat{\gamma}_{00} + \hat{U}_{0N}) + (\hat{U}_{0j} - \hat{U}_{0N}) + \hat{\gamma}_{10}c_{ij} = \\ &= (\hat{\gamma}_{00} + \hat{U}_{0j}) + \hat{\gamma}_{10}c_{ij} = \hat{\beta}_{0j} + \hat{\gamma}_{10}c_{ij}\end{aligned}\quad (3.13)$$

referita alla *j*-ma classe, dove la differenza tra i valori delle intercette delle rette di regressione è misurabile mediante il valore stimato $\hat{\beta}_{0j}$ dell'intercetta oppure attraverso il valore residuo \hat{U}_{0j} che rappresenta lo scostamento ($\hat{\beta}_{0j} - \hat{\gamma}_{00}$) dell'intercetta associata alla *j*-ma classe dalla media generale di tutte le classi. Le rette di regressione stimate per ogni classe si possono riportare in unico grafico (Fig.3.1).

Figura 3.1 Rette di regressione del modello ad intercetta casuale nella sua formulazione completa (3.1).



⁴ Altrimenti la matrice XX è singolare e quindi non ha inversa e il sistema d'equazioni non ha un'unica soluzione (Freund & Wilson, 2001, pag.513).

Dal momento che la stima dei minimi quadrati dei coefficienti di regressione $\hat{\lambda}_j$ potrà assumere sia valori positivi che negativi le rette di regressione delle classi si troveranno rispettivamente al di sopra o al di sotto della N -ma retta di regressione, la quale costituisce il riferimento per valutare la situazione registrata per ciascuna classe. Infine, notiamo che le equazioni (3.9) e (3.10) permettono di stabilire la corrispondenza esistente (Neter & Wassermann, 1974, pag. 536) tra il grafico della figura (2.1) e quello della figura (3.1), ovvero tra la formulazione *sintetica* (Tab.2.2) e quella *completa* (3.1) del modello ad intercetta casuale.

4. STIMATORE DELL'INTERCETTA

L'elemento principale del metodo di analisi è rappresentato dallo stimatore dell'intercetta, poiché esso è assunto come un indicatore sintetico idoneo ad un confronto *ceteris paribus* tra le classi (Goldstein, 1997). Nel modello (Tab.4.1) ad intercetta casuale privo di variabili esplicative (*empty model*) lo stimatore $\hat{\beta}_{0j}^{OLS}$ dell'intercetta è dato semplicemente dalla media \bar{Y}_j di classe dei valori della variabile risultato ed è considerato un *indicatore grezzo* di efficacia. L'inserimento della prima variabile x_{1j} esplicativa modifica la valutazione del valore medio di classe, \bar{Y}_j in base a due elementi:

- 1) il parametro $\hat{\gamma}_{10}$, che dipende dalla relazione esistente tra la variabile esplicativa inserita e la variabile risultato;
- 2) il valore della differenza $(\bar{x}_{1j} - \bar{x}_1)$, che rappresenta lo scostamento più o meno grande, con riferimento alla variabile esplicativa, tra la media della classe e la media generale di tutte le classi.

La logica di fondo nella costruzione dell'*indicatore netto* 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 $[(\bar{x}_{1j} - \bar{x}_1) > 0]$ tanto minore

sarà la valutazione $\hat{\beta}_{0j}$ ottenuta attraverso il valore medio di classe della variabile risultato. L'inserimento di più variabili esplicative nel modello statistico estende questa procedura di aggiustamento del valor medio di classe, ma la logica di fondo nella costruzione dell'indicatore netto di efficacia *rimane inalterata* (Tab.4.1).

Tabella 4.1 *Stimatore dei minimi quadrati in un modello ad intercetta casuale.*

<i>Variabili esplicative individuali inserite</i>	Stimatore dei minimi quadrati dell'intercetta $\hat{\beta}_{0j}^{OLS} = \hat{\gamma}_{00} + \hat{U}_{0j}$
0	\bar{Y}_j
1	$\bar{Y}_j - \hat{\gamma}_{10}(\bar{x}_{1j} - \bar{x}_1)$
2	$\bar{Y}_j - \hat{\gamma}_{10}(\bar{x}_{1j} - \bar{x}_1) - \hat{\gamma}_{20}(\bar{x}_{2j} - \bar{x}_2)$
3	$\bar{Y}_j - \hat{\gamma}_{10}(\bar{x}_{1j} - \bar{x}_1) - \hat{\gamma}_{20}(\bar{x}_{2j} - \bar{x}_2) - \hat{\gamma}_{30}(\bar{x}_{3j} - \bar{x}_3)$

Nel caso in cui la variabile esplicativa inserita è *correlata* con la variabile Y_{ij} e si registra uno *scostamento* tra i valori medi di classe rispetto alla media generale derivante dalla presenza di tale variabile esplicativa, allora dovremo confrontare le classi tenendo conto contemporaneamente di questi due elementi (**correlazione e scostamento**). La *direzione* e l'*entità* dell'aggiustamento della media \bar{Y}_j di ciascuna classe, dipende dal segno e dal valore del coefficiente stimato ($\hat{\gamma}$) e dal segno e valore dello scostamento $[(\bar{x}_j - \bar{x})]$. 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 appena descritti e ciò implica che se le variabili esplicative introdotte nel modello sono scarsamente correlate con la variabile risultato ($\gamma \rightarrow 0$) e/o non si registrano grosse differenze tra classi in ordine a tali variabili esplicative $[(\bar{x}_j - \bar{x}) \rightarrow 0]$, allora possiamo confrontare

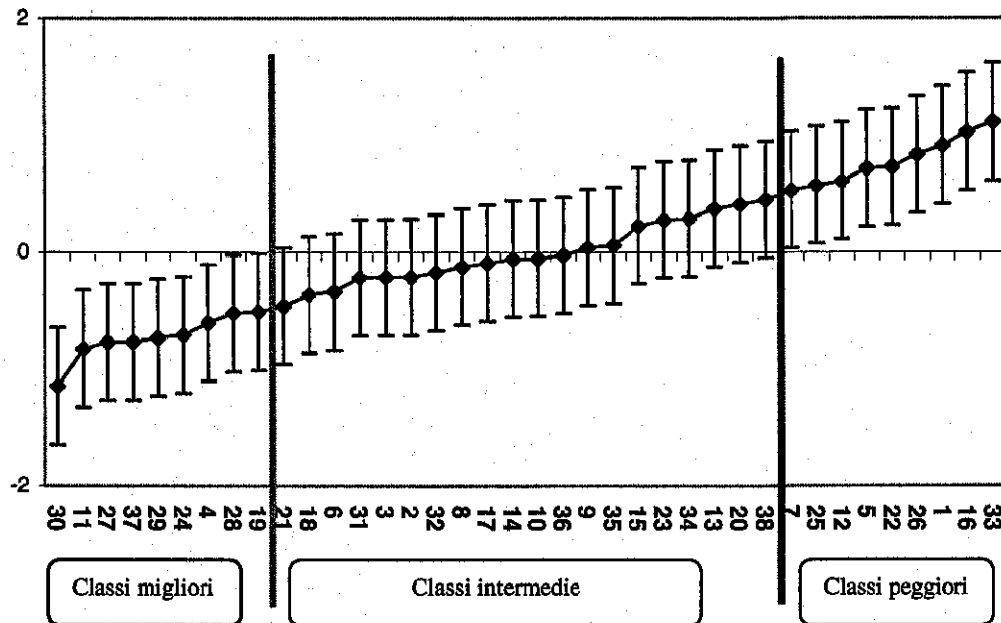
le classi sulla base del semplice valor medio di classe della variabile risultato, altrimenti è necessario far riferimento al cosiddetto indicatore netto di efficacia.

Il metodo dei minimi quadrati non è comunque l'unico metodo di stima dell'intercetta e si può pensare per esempio, di utilizzare nella stima le informazioni provenienti dalle singole classi e quelle derivanti dall'intero campione di classi osservate, ponderandole in base ad un fattore che assicura una maggiore affidabilità dell'indicatore di efficacia relativa (*Empirical Bayes Estimate*, Snijders & Bosker, 1999). Tuttavia, in questo contesto ci siamo limitati a considerare lo stimatore dei minimi quadrati ordinari che può essere utile come punto di partenza, prima di considerare dei metodi di stima alternativi.

5. VALUTAZIONE DI EFFICACIA RELATIVA

A partire dalla stima puntuale dell'intercetta del modello ad intercetta casuale si possono individuare degli intervalli di confidenza per il confronto tra classi. Tali stime puntuali dell'intercetta sono da considerarsi significativamente diverse se e solo se i rispettivi intervalli risultano disgiunti (*non-overlap criterion*) (Goldstein & Healy, 1995). Ordinando in senso crescente le classi in un diagramma cartesiano (Fig.5.1) sulla base delle stime puntuali $\hat{\beta}_{0j}$ dell'intercetta possiamo fare dei confronti tra ciascuna classe e la media generale $\hat{\gamma}_{00}$ delle classi, individuando tre gruppi di classi. La rappresentazione grafica è ottenuta riportando sull'asse delle ascisse il numero identificativo della classe e sull'asse dell'ordinate il valore $(\hat{\beta}_{0j} - \hat{\gamma}_{00})$ dell'intercetta centrata sul valore stimato $\hat{\gamma}_{00}$ della media generale di tutte le classi rispetto al disagio manifestato dagli studenti nei confronti degli insegnanti.

Figura 5.1 Confronto tra le classi in base alla stima per intervalli dell'intercetta nel modello ad intercetta casuale.



Il primo gruppo di valori minori di zero per i quali l'intervallo di confidenza *non comprende* la linea tratteggiata contribuisce al disagio dello studente nei confronti degli insegnanti in misura inferiore rispetto alla media di tutte le classi ed è quindi il gruppo delle **classi migliori**. Il secondo gruppo di valori prossimi a zero che si trova nell'intorno della media generale di disagio con gli insegnanti ed il cui intervallo di confidenza *comprende* la linea tratteggiata, costituisce il gruppo delle **classi intermedie**. Il terzo gruppo di valori superiori allo zero ed il cui intervallo di confidenza *non comprende* la linea tratteggiata contribuisce al disagio dello studente nei confronti degli insegnanti in misura superiore alla media ed è quindi il gruppo delle **classi peggiori**. Il criterio d'inserimento nell'uno o nell'altro gruppo dipende dalla collocazione dell'intervallo di confidenza intorno alla media generale stimata ($\hat{\gamma}_{00}$) per tutte le classi.

6. I RISULTATI

In questo paragrafo presentiamo in maniera più articolata i risultati ottenuti con il programma di elaborazione statistica SPSS (*Statistical Package for Social Science*) nell'applicazione del modello ad intercetta casuale per la misura del disagio degli studenti nei confronti degli insegnanti a parità di determinate condizioni soggettive o ambientali. La variabile risultato Y_{ij} scelta nei nostri quattro modelli statistici lineari (Tab.2.4), come già detto in precedenza (Par.2) è il disagio manifestato nei confronti degli insegnanti da parte di 592 (M) studenti appartenenti a 38 (N) classi di scuole elementari. A partire da questi dati si può cercare di capire se ci sono delle differenze tra classi e quantificare l'effetto classe attraverso gli stimatori dei minimi quadrati dell'intercetta. Le variabili esplicative che consideriamo nella misura del disagio nei confronti degli insegnanti sono il disagio manifestato dallo studente nei confronti della famiglia, il sesso e il disagio manifestato dallo studente nei confronti della scuola nel suo complesso. La correlazione tra queste variabili non è molto stretta ma è piuttosto significativa.

6.1 Il modello privo di variabili esplicative

La stima di un modello ad intercetta casuale privo di variabili esplicative (MIC-0) è il primo passo (Tab.2.4) da compiere per accertare (Snijders & Bosker, 1999, pag. 23) la presenza di differenze tra classi in ordine al disagio manifestato dagli studenti nei confronti degli insegnanti. Tale modello statistico lineare è conosciuto in letteratura (Searle *et al.*, 1992) come modello di analisi della varianza ad effetti casuali (*ANOVA One-way with Random effects*).

$$\begin{aligned} Y_{ij} &= \beta_{0j} + R_{ij} \\ \beta_{0j} &= \gamma_{00} + U_{0j} \end{aligned} \tag{6.1.1}$$

I calcoli necessari per la stima dei parametri incogniti in un modello di analisi della varianza ad effetti fissi e casuali⁵ sono gli stessi (Freund & Wilson, 2001), solo che in quest'ultimo caso dobbiamo stimare due componenti di varianza:

- 1) $\text{Var}(R_{ij}) = \sigma^2$ componente di varianza dentro le classi o *within*.
- 2) $\text{Var}(U_{0j}) = \tau_0^2$ componente di varianza fra classi o *between*.

Tabella 6.1.1 Risultati dell'analisi della varianza del disagio nei confronti degli insegnanti mediante il modello MIC-0.

Disagio con gli insegnanti					
Componenti	Somme dei quadrati	Gradi di libertà	Medie dei quadrati	F	Valore P
Tra classi	186,531	37	5,041	2,058	,000
Dentro le classi	1356,885	554	2,449		
Totale	1543,417	591			

L'interpretazione che viene data (Neter & Wassermann, 1974, pp.526-530) dell'uso del test F , con riferimento al problema trattato, è la seguente:

- in base al modello ANOVA con effetti fissi, se il test F è significativo si rifiuta l'ipotesi nulla che le medie di classe siano tutte uguali, quindi si assume che esista almeno una media che differisce significativamente dalle altre;
- in base ad un modello ANOVA con effetti casuali, se il test F è significativo si rifiuta l'ipotesi nulla $H_0: \tau_0^2 = 0$.

Le procedure inferenziali allora non riguardano le medie delle classi, ma la varianza delle medie delle classi.

⁵ La differenza principale tra un modello di analisi ad effetti fissi ed uno ad effetti casuali (Neter & Wassermann, 1974) è la metodologia di campionamento e di conseguenza l'interesse del ricercatore. Nel primo caso oggetto di studio sono le sole classi osservate (effetti fissi), nel secondo caso sono le popolazioni di classi dalle quali il campione casuale semplice (effetti casuali) è stato estratto (Bohrsted & Knoke, 1998; Snijders & Bosker, 1999).

In questo caso le ipotesi possono essere formulate come segue:

$$H_0: \tau_0^2 = 0;$$

$$H_1: \tau_0^2 \neq 0.$$

La quinta colonna (Tab.6.1.1) riporta il valore di F^* dato dal rapporto tra le medie dei quadrati tra e dentro le classi.

$$F^* = 5,041 / 2,449 = 2,058$$

La regola (Neter & Wassermann, 1974, pag. 448) per decidere quale ipotesi accettare facendo riferimento ad un certo livello di significatività α si basa sulle disuguaglianze,

$$F^* \leq F[1-\alpha; (38-1); (592-38)]$$

$$F^* \geq F[1-\alpha; (38-1); (592-38)]$$

che consentono rispettivamente, di accettare o rifiutare l'ipotesi nulla H_0 . Altrimenti in base al *valore⁶ della probabilità P ($P < \alpha$)* (ossia la *Pr* $\{F > 2,058\}$) possiamo rifiutare l'ipotesi nulla H_0 , concludendo che **esiste una variabilità nelle medie di classe**. A questo punto è interessante stimare le componenti di varianza fra classi, (*between*) $\hat{\tau}_0^2$, e dentro le classi, (*within*) $\hat{\sigma}^2$ per determinare la parte della variabilità di Y_{ij} da attribuire rispettivamente alle differenze tra classi ($\hat{\tau}_0^2$) e alle differenze tra le caratteristiche degli individui ($\hat{\sigma}^2$) all'interno delle classi (Tab.6.1.2). Dai risultati ottenuti

$$\hat{\sigma}^2 = 2,449$$

$$\hat{\tau}_0^2 = \frac{E(MSB) - E(MSE)}{\tilde{n}} = \frac{5,041 - 2,499}{15,541} = 0,167$$

⁶ D'ora in poi faremo riferimento al *valore p* per stabilire la significatività del parametro stimato.

dove il denominatore

$$\tilde{n} = \frac{1}{N-1} \left[\sum n_j - \frac{\sum n_j^2}{\sum n_j} \right] = \frac{1}{38-1} \left[592 - \frac{9696}{592} \right] = 15,541$$

è funzione della dimensione n_j delle classi, è possibile determinare il valore del coefficiente di correlazione intraclassa ($\rho = 6,38\%$).

Tabella 6.1.2. *Stima delle componenti di varianza del modello MIC-0 e relativo coefficiente di correlazione intraclassa.*

Componenti di varianza	Stima dei minimi quadrati
Tra classi (τ_0^2)	0,167
Dentro le classi (σ^2)	2,449
Totale	2,616

Coefficiente di correlazione intraclassa	Valore percentuale
$\rho = \frac{\tau_0^2}{\sigma^2 + \tau_0^2}$	$\frac{0,167}{2,616} = 6,38\%$

Questo significa che la variabilità complessiva del disagio manifestato dagli studenti nei confronti dei loro insegnanti è da attribuire per il 6,38% alle differenze tra le classi di appartenenza ed il 93,62% alle differenze tra le caratteristiche degli studenti all'interno delle classi. Per la stima puntuale dell'intercetta che rappresenta il nostro indicatore di efficacia relativa, dobbiamo stimare il modello completo di analisi della varianza.

$$Y_{ij} = \alpha_{00}x_{i0} + \lambda_1x_{i1} + \lambda_2x_{i2} + \dots + \lambda_jx_{ij} + \dots + \lambda_{37}x_{i37} + R_{ij} \quad (6.1.2)$$

I parametri incogniti α_{00} e λ_j sono stimati con il metodo dei minimi quadrati ordinari (Tab. 6.1.3). L'interpretazione dei parametri stimati ($\hat{\alpha}_{00}$, $\hat{\lambda}_j$) fa riferimento al parametro stimato per la 38-ma classe omessa, ovvero il valore $\hat{\alpha}_{00}$ (Par.3).

Tabella 6.1.3 Stime dei minimi quadrati dei parametri del modello MIC-0.

Nome variabile	Parametro	Stime OLS	Errore Std	Valore P	Nome variabile	Parametro	Stime OLS	Errore Std.	Valore P
Intercetta	$\hat{\alpha}_{00}$	2,545	0,334	0,000	Classe 19	$\hat{\lambda}_{19}$	-0,879	0,562	0,118
Classe 1	$\hat{\lambda}_1$	0,555	0,597	0,353	Classe 20	$\hat{\lambda}_{20}$	0,070	0,547	0,898
Classe 2	$\hat{\lambda}_2$	-0,537	0,490	0,273	Classe 21	$\hat{\lambda}_{21}$	-0,840	0,505	0,097
Classe 3	$\hat{\lambda}_3$	-0,545	0,490	0,266	Classe 22	$\hat{\lambda}_{22}$	0,396	0,505	0,434
Classe 4	$\hat{\lambda}_4$	-0,938	0,535	0,080	Classe 23	$\hat{\lambda}_{23}$	-0,101	0,619	0,870
Classe 5	$\hat{\lambda}_5$	0,378	0,547	0,491	Classe 24	$\hat{\lambda}_{24}$	-1,177	0,490	0,017
Classe 6	$\hat{\lambda}_6$	-0,688	0,477	0,150	Classe 25	$\hat{\lambda}_{25}$	0,237	0,467	0,612
Classe 7	$\hat{\lambda}_7$	0,105	0,578	0,856	Classe 26	$\hat{\lambda}_{26}$	0,455	0,477	0,342
Classe 8	$\hat{\lambda}_8$	-0,516	0,505	0,308	Classe 27	$\hat{\lambda}_{27}$	-1,095	0,484	0,024
Classe 9	$\hat{\lambda}_9$	-0,392	0,547	0,475	Classe 28	$\hat{\lambda}_{28}$	-0,990	0,619	0,110
Classe 10	$\hat{\lambda}_{10}$	-0,428	0,505	0,398	Classe 29	$\hat{\lambda}_{29}$	-1,045	0,597	0,080
Classe 11	$\hat{\lambda}_{11}$	-1,117	0,679	0,101	Classe 30	$\hat{\lambda}_{30}$	-1,590	0,490	0,001
Classe 12	$\hat{\lambda}_{12}$	0,339	0,547	0,536	Classe 31	$\hat{\lambda}_{31}$	-0,670	0,514	0,193
Classe 13	$\hat{\lambda}_{13}$	-0,075	0,505	0,882	Classe 32	$\hat{\lambda}_{32}$	-0,545	0,646	0,399
Classe 14	$\hat{\lambda}_{14}$	-0,379	0,721	0,599	Classe 33	$\hat{\lambda}_{33}$	0,621	0,619	0,316
Classe 15	$\hat{\lambda}_{15}$	-0,233	0,514	0,651	Classe 34	$\hat{\lambda}_{34}$	-0,193	0,505	0,703
Classe 16	$\hat{\lambda}_{16}$	0,560	0,490	0,254	Classe 35	$\hat{\lambda}_{35}$	-0,420	0,462	0,363
Classe 17	$\hat{\lambda}_{17}$	-0,430	0,597	0,472	Classe 36	$\hat{\lambda}_{36}$	-0,420	0,462	0,363
Classe 18	$\hat{\lambda}_{18}$	-0,768	0,497	0,123	Classe 37	$\hat{\lambda}_{37}$	-1,111	0,467	0,018

La stima $\hat{\beta}_{0j}$ dell'intercetta per ciascuna classe del modello MIC-0 pu  essere determinata a partire dalle stime dei parametri

$$\hat{\alpha}_{00} = \hat{\gamma}_{00} + \hat{U}_{0,38} = 2,545$$

$$\hat{U}_{0,38} = 2,545 - 2,155 = 0,390$$

$$\hat{\gamma}_{00} = \sum_{j=1}^{38} \frac{n_j}{592} \bar{Y}_j = 2,155$$

considerando che il valore stimato $\hat{\alpha}_{00}$ è dato dalla somma della media $\hat{\gamma}_{00}$ generale di disagio nei confronti degli insegnanti e il residuo stimato $\hat{U}_{0,38}$ associato alla 38-ma classe. Il valore residuo $\hat{U}_{0,j}$ riferito alla j -ma classe può essere determinato come somma $(\hat{\lambda}_j + \hat{U}_{0,38})$ tra la stima del parametro $\hat{\lambda}_j$ e la stima $\hat{U}_{0,38}$ del 38-mo residuo di classe e rappresenta il disagio medio manifestato dagli studenti della j -ma classe nei confronti degli insegnanti (Tab.6.1.4).

Tabella 6.1.4 Valori delle stime di alcuni parametri del modello MIC-0.

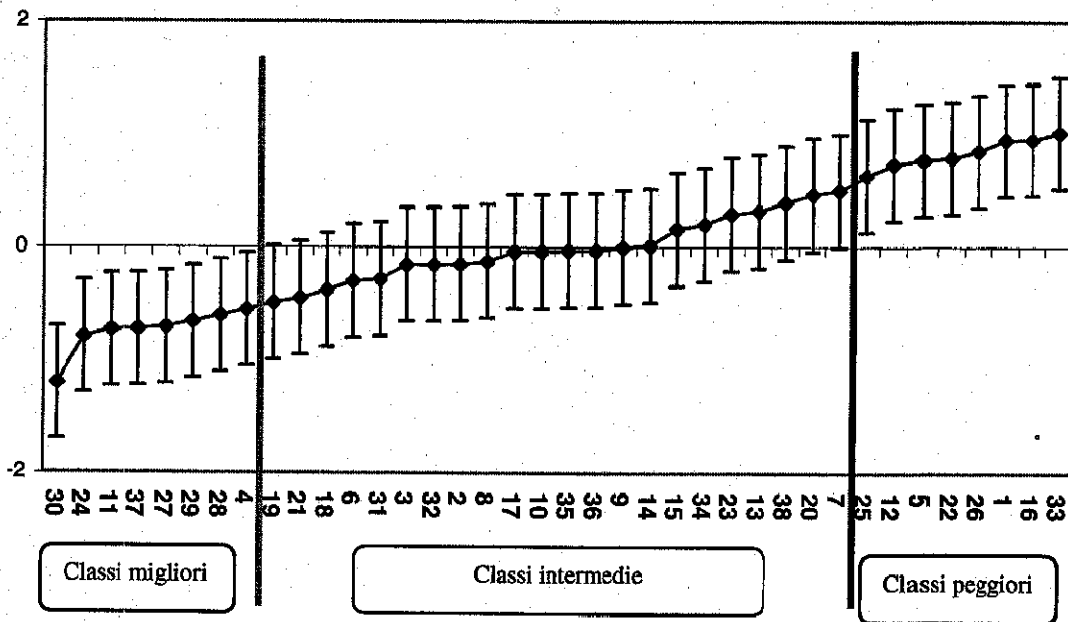
Numero Classe	Parametri stimati $\hat{\lambda}_j$	Residuo di classe $\hat{U}_{0,j}$	Intercetta $\hat{\beta}_{0j}$	Numero Classe	Parametri stimati $\hat{\lambda}_j$	Residuo di classe $\hat{U}_{0,j}$	Intercetta $\hat{\beta}_{0j}$
1	0,555	0,945	3,100	20	0,070	0,460	2,615
2	-0,537	-0,147	2,008	21	-0,840	-0,450	1,705
3	-0,545	-0,155	2,000	22	0,396	0,786	2,941
4	-0,938	-0,548	1,607	23	-0,101	0,289	2,444
5	0,378	0,768	2,923	24	-1,177	-0,787	1,368
6	-0,688	-0,298	1,857	25	0,237	0,627	2,782
7	0,105	0,495	2,650	26	0,455	0,845	3,000
8	-0,516	-0,126	2,029	27	-1,095	-0,705	1,450
9	-0,392	-0,002	2,153	28	-0,990	-0,600	1,555
10	-0,428	-0,038	2,117	29	-1,045	-0,655	1,500
11	-1,117	-0,727	1,428	30	-1,590	-1,200	0,955
12	0,339	0,729	2,884	31	-0,670	-0,280	1,875
13	-0,075	0,315	2,470	32	-0,545	-0,155	2,000
14	-0,379	0,011	2,166	33	0,621	1,011	3,166
15	-0,233	0,157	2,312	34	-0,193	0,197	2,352
16	0,560	0,950	3,105	35	-0,420	-0,030	2,125
17	-0,430	-0,040	2,115	36	-0,420	-0,030	2,125
18	-0,768	-0,378	1,777	37	-1,111	-0,721	1,434
19	-0,879	-0,489	1,666	38	0,000	0,390	2,545

Il valore residuo $\hat{U}_{0,j}$ associato alla j -ma classe ha un valore positivo o negativo a seconda che la j -ma media di classe rispetto al disagio manifestato nei confronti degli insegnanti si trovi al di sopra o al di sotto della media generale delle classi ($\hat{\gamma}_{00} = 2,155$) per la medesima variabile.

Sommando il valore residuo $\hat{U}_{0,j}$ di classe al valore della media generale $\hat{\gamma}_{00}$ (2,155) si ottiene il valore stimato $\hat{\beta}_{0,j}$ dell'intercetta del modello ad intercetta casuale privo di variabili esplicative (RIM-0), la quale coincide con il valor medio \bar{Y}_j di classe rispetto al disagio degli studenti nei confronti degli insegnanti.

L'operazione di confronto delle classi in base alla stima $\hat{\beta}_{0,j}$ dell'intercetta può essere fatta ordinando in modo crescente tali valori e rappresentando graficamente le stime con intervalli dei valori residui $\hat{U}_{0,j}$ (Fig. 6.1.1). L'asse orizzontale contrassegnato con il valore zero corrisponde al valore della media generale delle classi ($\hat{\gamma}_{00}=2,155$) e la losanga corrisponde al valore dei residui $\hat{U}_{0,j}$ di classe.

Figura 6.1.1 *Graduatoria delle classi in base alla stima dell'intercetta nel modello MIC-0.*



In tal modo si possono individuare tre gruppi di classi a seconda della collocazione dell'intervallo di confidenza intorno al valore centrale. Queste valutazioni tuttavia non tengono conto delle caratteristiche degli studenti presenti nelle varie classi e non sarebbero quindi confronti *ceteris paribus*. Possiamo immaginare che in una classe di allievi di 8 anni circa, tutti maschi o che vivono una situazione di disagio in famiglia rispondano in modo diverso da altri studenti, quando gli si chiede ad esempio se gli insegnanti sono ingiusti con loro. Dobbiamo stimare quindi il valore dell'intercetta tenendo conto di alcune variabili individuali, ossia passare da indicatori grezzi a indicatori netti d'efficacia.

6.2 Il disagio dello studente in famiglia

Ai fini della misura del disagio dello studente nei confronti degli insegnanti è utile tenere conto del disagio che lo studente manifesta nei confronti della famiglia perché esso può essere determinante e dar luogo a dei valori diversi dell'intercetta che compare nel modello statistico. Pertanto, è opportuno predisporre un modello di analisi della covarianza ad effetti casuali (MIC-1) con una variabile esplicativa o covariata, x_{ij} che rappresenta il disagio dello studente nei confronti della famiglia. Per facilitare l'interpretazione dell'intercetta si può esprimere la variabile esplicativa in termini di scarti dal valore medio \bar{x}_1 ottenendo la variabile c_{ij} .

$$Y_{ij} = \beta_{0j} + \gamma_{10} c_{ij} + R_{ij} \quad (6.2.1)$$

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

Dalla tabella di analisi della varianza (Tab.6.2.1) possiamo osservare dal valore p la significatività delle stime dei parametri associati alle variabili disagio dello studente in famiglia (valore $p = 0,002$) e classe d'appartenenza (valore $p = 0,000$).

Tabella 6.2.1. Risultati dell'analisi della varianza del disagio nei confronti degli insegnanti mediante il modello MIC-1.

Dependent Variable: Disagio con gli insegnanti

Fonte di variabilità		Somme dei quadrati	Gradi di libertà	Medie dei quadrati	F	Valore P
Intercetta	Hypothesis	2429,373	1	2429,373	502,168	,000
	Error	202,009	41,757	4,838		
Disagio in famiglia	Hypothesis	24,282	1	24,282	10,077	,002
	Error	1332,603	553	2,410		
Classe	Hypothesis	191,022	37	5,163	2,142	,000
	Error	1332,603	553	2,410		

Entrambe le variabili sono statisticamente rilevanti nel determinare il disagio degli studenti nei confronti degli insegnanti e quindi una parte della sua variabilità

complessiva. Inoltre, possiamo notare che si sono modificate le componenti di varianza (Tab.6.2.2) calcolate in funzione della variabile esplicativa inserita e quindi il coefficiente di correlazione intraclassa ($\rho | x = 6,88\%$).

Tabella 6.2.2. Stima delle componenti di varianza nel modello MIC-1 e relativo coefficiente di correlazione intraclassa.

Componente di varianza	Stima minimi quadrati
Tra classi $\hat{\tau}_0^2$	0,178
Dentro le classi $\hat{\sigma}^2$	2,410
Totale	2,588

Coefficiente di correlazione intraclassa	Valore
$\rho x = \frac{\tau_0^2}{\sigma^2 + \tau_0^2}$	$\frac{0,178}{2,588} = 6,88\%$

Rispetto al modello di analisi della varianza ad effetti casuali (MIC-0) la componente *between* della varianza aumenta da 0,168 a 0,178, mentre la componente *within* della varianza si riduce da 2,449 a 2,410.

Il coefficiente di correlazione intraclassa passa quindi dal 6,38% al 6,88% con un aumento di quasi mezzo punto percentuale.

Tutto questo significa che la variabile disagio in famiglia:

- dà luogo ad un aumento del grado di eterogeneità tra classi (*between*) ed evidenzia le **dissomiglianze tra classi** rispetto al disagio degli studenti nei confronti degli insegnanti;
- **contribuisce a spiegare il disagio** manifestato dagli studenti nei confronti degli insegnanti, riducendo quindi la variabilità interna alle classi (*within*).

La figura (6.2.1) mostra le differenze tra classi in termini di disagio medio per classe manifestato dagli studenti in famiglia.

Per quanto riguarda le stime ottenute col metodo dei minimi quadrati ordinari dei parametri incogniti (Tab.6.2.3) del modello ad intercetta casuale nella versione completa

$$Y_{ij} = \alpha_{00}x_{i0} + \gamma_{10}c_{1ij} + \lambda_{01}x_1 + \lambda_{02}x_2 + \dots + \lambda_{37}x_{37} + R_{ij} \quad (6.2.2)$$

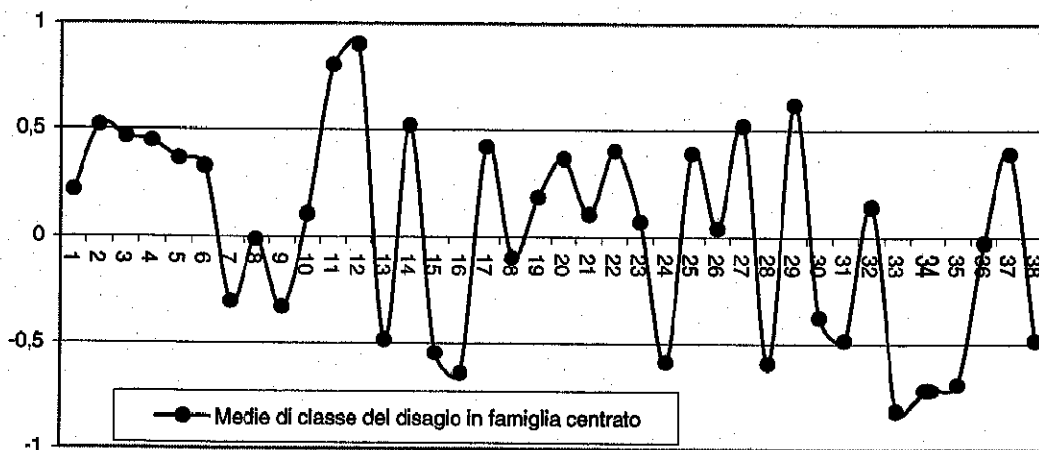
e la loro interpretazione, attraverso la dei parametri

$$\hat{\alpha}_{00} = \hat{\gamma}_{00} + \hat{U}_{0;38} = 2,155 + \hat{U}_{0;38} = 2,608$$

$$\hat{U}_{0;38} = 2,608 - 2,155 = 0,453$$

relativi all'intercetta e all'effetto classe per la classe 38, possiamo determinare ed interpretare i parametri restanti che compaiono nel modello completo (6.2.2).

Figura 6.2.1 Valori medi di classe del disagio manifestato dallo studente in famiglia centrato sulla media generale.



Il valore dell'intercetta ($\hat{\alpha}_{00}=2,608$) rappresenta la somma del valore medio del disagio con gli insegnanti con il valore residuo della classe 38, quindi in funzione del primo valore ($\hat{\gamma}_{00}=2,155$) si trova per differenza il secondo ($\hat{U}_{0;38}=0,453$).

Tabella 6.2.3 Stime dei minimi quadrati dei parametri del modello MIC-1.

Nome variabile	Parametro	Stime OLS	Errore Std	Valore P	Nome variabile	Parametro	Stime OLS	Errore Std	Valore P
Intercetta	$\hat{\alpha}_{00}$	2,608	0,332	0,000	Classe 19	$\hat{\lambda}_{19}$	-0,965	0,558	0,084
Disagio in famiglia	$\hat{\gamma}_{10}$	0,129	0,041	0,002	Classe 20	$\hat{\lambda}_{20}$	-0,039	0,544	0,943
Classe 1	$\hat{\lambda}_1$	0,464	0,593	0,434	Classe 21	$\hat{\lambda}_{21}$	-0,915	0,502	0,069
Classe 2	$\hat{\lambda}_2$	-0,666	0,488	0,173	Classe 22	$\hat{\lambda}_{22}$	0,282	0,503	0,575
Classe 3	$\hat{\lambda}_3$	-0,667	0,488	0,172	Classe 23	$\hat{\lambda}_{23}$	-0,173	0,615	0,779
Classe 4	$\hat{\lambda}_4$	-1,058	0,532	0,047	Classe 24	$\hat{\lambda}_{24}$	-1,163	0,486	0,017
Classe 5	$\hat{\lambda}_5$	0,269	0,544	0,622	Classe 25	$\hat{\lambda}_{25}$	0,125	0,464	0,788
Classe 6	$\hat{\lambda}_6$	-0,793	0,475	0,096	Classe 26	$\hat{\lambda}_{26}$	0,387	0,474	0,415
Classe 7	$\hat{\lambda}_7$	0,082	0,573	0,887	Classe 27	$\hat{\lambda}_{27}$	-1,224	0,481	0,011
Classe 8	$\hat{\lambda}_8$	-0,577	0,502	0,251	Classe 28	$\hat{\lambda}_{28}$	-0,976	0,614	0,113
Classe 9	$\hat{\lambda}_9$	-0,411	0,543	0,449	Classe 29	$\hat{\lambda}_{29}$	-1,187	0,594	0,046
Classe 10	$\hat{\lambda}_{10}$	-0,504	0,502	0,316	Classe 30	$\hat{\lambda}_{30}$	-1,603	0,486	0,001
Classe 11	$\hat{\lambda}_{11}$	-1,282	0,676	0,058	Classe 31	$\hat{\lambda}_{31}$	-0,670	0,510	0,189
Classe 12	$\hat{\lambda}_{12}$	0,161	0,546	0,768	Classe 32	$\hat{\lambda}_{32}$	-0,626	0,641	0,330
Classe 13	$\hat{\lambda}_{13}$	-0,075	0,501	0,881	Classe 33	$\hat{\lambda}_{33}$	0,664	0,614	0,280
Classe 14	$\hat{\lambda}_{14}$	-0,508	0,716	0,479	Classe 34	$\hat{\lambda}_{34}$	-0,162	0,501	0,746
Classe 15	$\hat{\lambda}_{15}$	-0,225	0,510	0,659	Classe 35	$\hat{\lambda}_{35}$	-0,394	0,458	0,391
Classe 16	$\hat{\lambda}_{16}$	0,580	0,486	0,233	Classe 36	$\hat{\lambda}_{36}$	-0,479	0,459	0,296
Classe 17	$\hat{\lambda}_{17}$	-0,546	0,593	0,358	Classe 37	$\hat{\lambda}_{37}$	-1,223	0,464	0,009
Classe 18	$\hat{\lambda}_{18}$	-0,818	0,494	0,098					

Gli altri valori residui $\hat{U}_{0,j}$ (Tab.6.2.4) si trovano come somma tra il valore del j -mo parametro stimato ($\hat{\lambda}_j$) e il residuo $\hat{U}_{0,38}$ relativo alla classe 38.

Il valore stimato del coefficiente di regressione della variabile disagio in famiglia ($\hat{\gamma}_{10} = 0,129$) è significativo (*valore p* = 0,002) e pertanto si può dire che

l'incremento unitario del disagio in famiglia comporta un aumento, in media, di 0,129 punti nel disagio manifestato dallo studente nei confronti degli insegnanti.

Tabella 6.2.4 Valori delle stime di alcuni parametri del modello MIC-1.

Numero Classe	Parametri stimati $\hat{\lambda}_j$	Residuo di classe $\hat{U}_{0,j}$	Intercetta $\hat{\beta}_{0,j}$	Numero Classe	Parametri stimati $\hat{\lambda}_j$	Residuo di classe $\hat{U}_{0,j}$	Intercetta $\hat{\beta}_{0,j}$
1	0,464	0,917	3,072	20	-0,039	0,414	2,569
2	-0,666	-0,213	1,942	21	-0,915	-0,462	1,693
3	-0,667	-0,214	1,941	22	0,282	0,735	2,890
4	-1,058	-0,605	1,550	23	-0,173	0,280	2,435
5	0,269	0,722	2,877	24	-1,163	-0,710	1,445
6	-0,793	-0,340	1,815	25	0,125	0,578	2,733
7	0,082	0,535	2,690	26	0,387	0,840	2,995
8	-0,577	-0,124	2,031	27	-1,224	-0,771	1,384
9	-0,411	0,042	2,197	28	-0,976	-0,523	1,632
10	-0,504	-0,051	2,104	29	-1,187	-0,734	1,421
11	-1,282	-0,829	1,326	30	-1,603	-1,150	1,005
12	0,161	0,614	2,769	31	-0,670	-0,217	1,938
13	-0,075	0,378	2,533	32	-0,626	-0,173	1,982
14	-0,508	-0,055	2,100	33	0,664	1,117	3,272
15	-0,225	0,228	2,383	34	-0,162	0,291	2,446
16	0,580	1,033	3,188	35	-0,394	0,059	2,214
17	-0,546	-0,093	2,062	36	-0,479	-0,026	2,129
18	-0,818	-0,365	1,790	37	-1,223	-0,770	1,385
19	-0,965	-0,512	1,643	38	0,000	0,453	2,608

Il disagio nei confronti degli insegnanti stimato per l' i -mo studente appartenente alla j -ma classe

$$\hat{Y}_{ij} = (\hat{\gamma}_{00} + \hat{U}_{0,j}) + 0,129c_{ij} = \\ = \hat{\beta}_{0,j} + 0,129c_{ij}$$

dipende dal valore costante stimato $\hat{\beta}_{0,j}$ per tutta la classe e dal valore della variabile scarto c_{ij} relativo al disagio manifestato dallo studente in famiglia. Le differenze tra classi si possono misurare in base ai valori residui $\hat{U}_{0,j}$ di classe, o identicamente in base ai valori della stima $\hat{\beta}_{0,j}$ dell'intercetta (Tab.6.2.4). Nel paragrafo precedente abbiamo visto che lo stimatore $\hat{\beta}_{0,j}$ dell'intercetta

rappresenta la media del disagio \bar{Y}_j della classe nei confronti degli insegnanti.

L'introduzione di una variabile esplicativa modifica tale valor medio in funzione di un termine aggiustamento

$$- \hat{\gamma}_{10}(\bar{x}_{1j} - \bar{x}_1)$$

che dipende da due elementi:

- la stima $\hat{\gamma}_{10}$ del coefficiente associato alla variabile disagio dello studente in famiglia ($\hat{\gamma}_{10}=0,129$);
- lo scostamento $(\bar{x}_{1j} - \bar{x}_1)$, con riferimento alla variabile disagio in famiglia, tra la media di classe (\bar{x}_{1j}) e la media generale ($\bar{x}_1 = 3,483$) di tutte le classi.

Questo consente di giustificare il fatto che nell'analisi della covarianza spesso si parla di **effetti classe aggiustati, corretti o netti** (Tab.6.2.5).

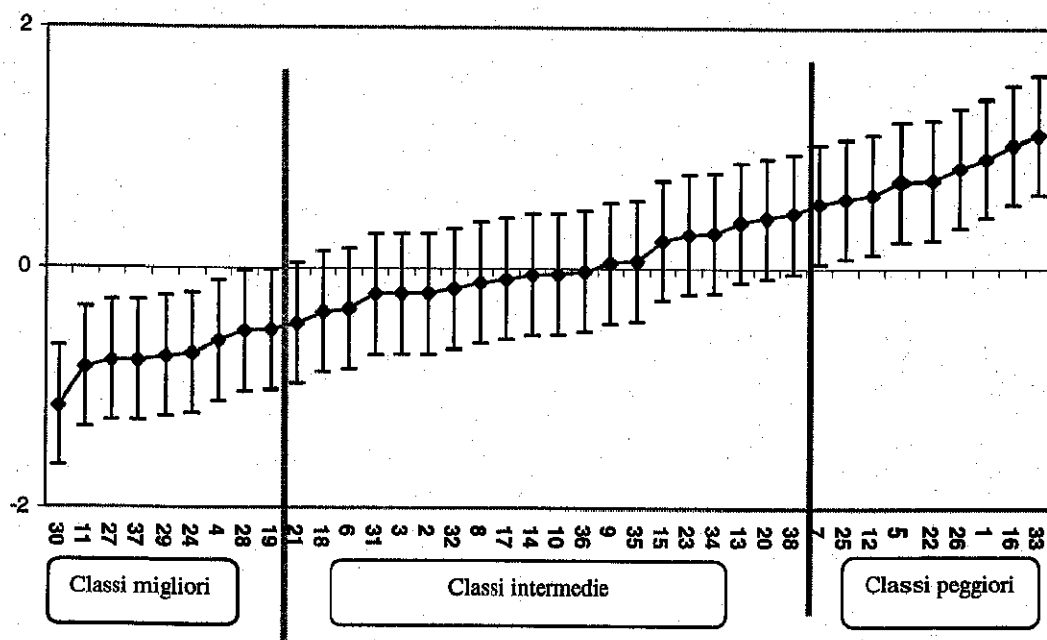
Tabella 6.2.5 *Aggiustamento del valor medio di classe nel modello MIC-1.*

Num. Classe	Disagio medio in famiglia \bar{x}_{1j}	Medie di classe \bar{Y}_j	Medie di classe aggiustate	Num. Classe	Disagio medio in famiglia \bar{x}_{1j}	Medie di classe \bar{Y}_j	Medie di classe aggiustate
1	3,700	3,100	3,072	20	3,846	2,615	2,569
2	4,000	2,008	1,942	21	3,588	1,705	1,693
3	3,947	2,000	1,941	22	3,882	2,941	2,890
4	3,929	1,607	1,550	23	3,556	2,444	2,435
5	3,846	2,923	2,877	24	2,895	1,368	1,445
6	3,810	1,857	1,815	25	3,870	2,782	2,733
7	3,182	2,650	2,690	26	3,524	3,000	2,995
8	3,471	2,029	2,031	27	4,000	1,450	1,384
9	3,154	2,153	2,197	28	2,889	1,555	1,632
10	3,588	2,117	2,104	29	4,100	1,500	1,421
11	4,286	1,428	1,326	30	3,105	0,955	1,005
12	4,385	2,884	2,769	31	3,000	1,875	1,938
13	3,000	2,470	2,533	32	3,625	2,000	1,982
14	4,000	2,166	2,100	33	2,667	3,166	3,272
15	2,938	2,312	2,383	34	2,765	2,352	2,446
16	2,842	3,105	3,188	35	2,792	2,125	2,214
17	3,900	2,115	2,062	36	3,458	2,125	2,129
18	3,389	1,777	1,790	37	3,870	1,434	1,385
19	3,667	1,666	1,643	38	3,000	2,545	2,608

La logica di fondo dell'aggiustamento deriva dalla constatazione, che quanto **maggiore** è il disagio medio di classe manifestato dagli studenti in famiglia tanto

minore sarà la valutazione del disagio medio di classe manifestato dagli studenti nei confronti degli insegnanti. L'operazione di confronto tra le classi in base alla stima $\hat{\beta}_{0j}$ dell'intercetta può essere fatta ordinando in modo crescente tali valori e rappresentando graficamente le stime dell'intercetta con i relativi intervalli di confidenza (Fig. 6.2.2). L'asse orizzontale, tracciato in corrispondenza del valore zero rappresenta la media generale delle classi ($\hat{\gamma}_{00} = 2,155$) e la losanga rappresenta lo scostamento da tale valore della stima $\hat{\beta}_{0j}$ dell'intercetta di classe.

Figura 6.2.2 *Graduatoria delle classi in base alla stima dell'intercetta nel modello MIC-1.*



Confrontando la graduatoria delle classi ottenute mediante il modello MIC-0 (Modello ad Intercetta Casuale con 0 variabili esplicative) e quella ottenuta col modello MIC-1 (Modello ad Intercetta Casuale con 1 variabile esplicativa) è facile rilevare l'effetto dovuto all'introduzione della variabile esplicativa c_{1j} relativa al disagio dello studente in famiglia nella misura del disagio manifestato dalle classi nei confronti degli insegnanti.

6.3 Il sesso dello studente

Un'altra variabile interessante da considerare per questo campo di studio è sicuramente il sesso dello studente. In questo paragrafo ci proponiamo di prenderla in esame per poter rilevare la sua influenza nella misura del disagio manifestato con gli insegnanti delle singole classi. A tal fine introduciamo nel modello statistico (MIC-2) un'altra variabile esplicativa x_{2ij} che identifica il sesso dello studente. Si tratta di una variabile qualitativa dicotomica che può essere inserita senza particolari problemi nel modello di regressione (Freund & Wilson, 2001, pag.513). Anche in questo caso (Bryk & Raudeunbush, 1992, pag.28) è utile centrare la variabile sesso sul suo valore medio considerando il modello

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \gamma_{10} c_{1ij} + \gamma_{20} s_{2ij} + R_{ij} \\ \beta_{0j} &= \gamma_{00} + U_{0j} \end{aligned} \tag{6.3.1}$$

dove:

β_{0j} = intercetta della j -mo piano di regressione;

$c_{1ij} = (x_{1ij} - \bar{x}_1)$, disagio in famiglia centrato sulla media generale;

$s_{2ij} = (x_{2ij} - \bar{x}_2)$, sesso centrato sulla media generale;

γ_{00} = media generale del disagio degli studenti con gli insegnanti;

γ_{10} = variazione di Y_{ij} associata ad una variazione unitaria di x_{1ij} ,
tenendo costante x_{2ij} ;

γ_{20} = variazione di Y_{ij} associata ad una variazione unitaria di x_{2ij} ,
tenendo costante x_{1ij} ;

x_{1ij} = disagio dello studente in famiglia;

x_{2ij} = sesso dello studente (1 se maschio e 0 se femmina).

Anche in questo caso, dalla tabella di analisi della varianza (Tab.6.3.1) possiamo stabilire immediatamente dal *valore p* la significatività dei parametri

associati alle variabili sesso (*valore p* = 0,000), disagio in famiglia (*valore p* = 0,001) e la classe d'appartenenza (*valore p* = 0,000). Da notare che l'inserimento della seconda variabile esplicativa, il sesso dello studente, non riduce la significatività del parametro stimato per la prima variabile esplicativa inserita, il disagio dello studente nei confronti della famiglia.

Tabella 6.3.1. Risultati dell'analisi della varianza del disagio nei confronti degli insegnanti mediante il modello MIC-2.

Dependent Variable: Disagio con gli insegnanti

Fonte di variabilità		Somme dei quadrati	Gradi di libertà	Medie dei quadrati	F	Valore P
Intercetta	Hypothesis	2429,856	1	2429,856	516,657	,000
	Error	196,066	41,689	4,703		
Disagio in famiglia	Hypothesis	24,153	1	24,153	10,230	,001
	Error	1303,304	552	2,361		
Sesso	Hypothesis	29,299	1	29,299	12,409	,000
	Error	1303,304	552	2,361		
Classe	Hypothesis	185,340	37	5,009	2,122	,000
	Error	1303,304	552	2,361		

Per quanto riguarda le variazioni intervenute nelle componenti di varianza (Tab.6.3.2) condizionate dalle variabili esplicative inserite e quindi il relativo coefficiente di correlazione intraclassa, possiamo notare che la componente di varianza *between* diminuisce da 0,178 a 0,171, mentre la componente di varianza *within* si riduce da 2,410 a 2,361. Dunque il coefficiente di correlazione intraclassa passa dal 6,88% al 6,75%.

Tutto questo significa che la variabile sesso:

- **diminuisce il grado di eterogeneità tra classi (*between*), in quanto porta alla luce delle somiglianze tra classi rispetto al disagio degli studenti nei confronti degli insegnanti;**
- **contribuisce a spiegare il disagio manifestato dagli studenti nei confronti degli insegnanti riducendo quindi la variabilità residua (*within*).**

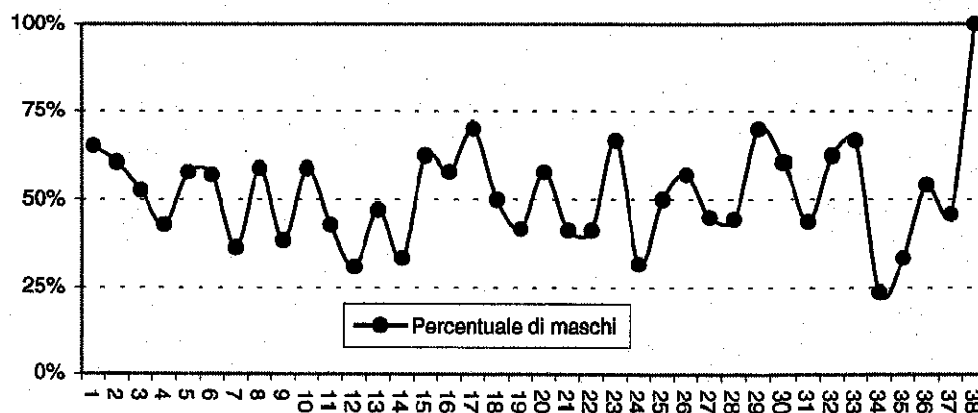
Tabella 6.3.2. Stima delle componenti di varianza nel modello MIC-2 e relativo coefficiente di correlazione intraclassa.

Componente di varianza	Stima minimi quadrati
Tra classi $\hat{\tau}_0^2$	0,171
Dentro le classi $\hat{\sigma}^2$	2,361
Totale	2,532

Coefficiente di correlazione intraclassa	Valore
$\rho x = \frac{\tau_0^2}{\sigma^2 + \tau_0^2}$	$\frac{0,171}{2,532} = 6,75\%$

E' interessante notare la diversa presenza di maschi e di femmine nelle classi (Fig.6.3.1). La linea tratteggiata centrale rappresenta la composizione "ideale" per sesso della classe, ossia il 50% di maschi e quindi di femmine.

Figura 6.3.1 Valori medi di classe della variabile sesso o percentuali di maschi presenti in classe.



In realtà, tranne le classi 3, 18 e 25 tutte le altre classi oscillano intorno a tale valore "ideale" ed in particolare alcune di esse hanno una percentuale di maschi superiore (quindi inferiore di femmine) al 50% mentre altre classi hanno una percentuale di maschi inferiore al 50%. La natura binaria della variabile sesso rende

indifferente parlare di valor medio di classe o percentuale di maschi presenti in classe in quanto abbiamo codificato con il valore 1 i maschi e con il valore 0 le femmine per cui facendo una media di classe rispetto a tale variabile otteniamo il valore percentuale dei maschi presenti in quella stessa classe.

Ad esempio nella classe 29 composta da 10 studenti, di cui 7 maschi (1) e 3 femmine (0), il valor medio è dato da

$$\bar{x}_{2,29} = \frac{7+0}{10} = 0,7$$

e ciò equivale a dire che la classe 29 è composta dal 70% di maschi. Tutte le classi oscillano intorno ad una composizione paritaria tra maschi e femmine, mentre si distingue la classe 34 con il 76% di femmine e la classe 38 con il 100% di maschi.

Per quanto riguarda le stime ottenute col metodo dei minimi quadrati ordinari dei parametri incogniti (Tab.6.3.3) del modello ad intercetta casuale nella versione completa

$$Y_{ij} = \alpha_{00}x_{i0} + \gamma_{10}c_{1ij} + \gamma_{20}s_{2ij} + \lambda_{01}x_1 + \lambda_{02}x_2 + \dots + \lambda_{37}x_{37} + R_{ij} \quad (6.3.2)$$

e la loro interpretazione, attraverso la stima dei parametri

$$\hat{\alpha}_{00} = \hat{\gamma}_{00} + \hat{U}_{0,38} = 2,155 + \hat{U}_{0,38} = 2,380$$

$$\hat{U}_{0,38} = 2,380 - 2,155 = 0,225$$

relativi all'intercetta e all'effetto classe per la classe 38, possiamo determinare ed interpretare i parametri restanti che compaiono nel modello completo (6.3.2).

La stima dell'intercetta ($\hat{\alpha}_{00}=2,380$) rappresenta la somma tra il valore medio $\hat{\gamma}_{00}$ della variabile disagio nei confronti degli insegnanti con il valore residuo $\hat{U}_{0,38}$ della classe 38, conoscendo il primo valore si trova per differenza il secondo ($\hat{U}_{0,38}=0,225$). Gli altri valori residui $\hat{U}_{0,j}$ si trovano come somma tra il valore del j -mo parametro stimato ($\hat{\lambda}_j$) e il residuo $\hat{U}_{0,38}$ relativo alla classe 38 (Tab.6.3.4).

Tabella 6.3.3 Stime dei minimi quadrati dei parametri del modello MIC-2.

Nome variabile	Parametro	Stime OLS	Errore Std.	Valore P	Nome variabile	Parametro	Stime OLS	Errore Std.	Valore P
Intercetta	$\hat{\alpha}_{00}$	2,380	0,334	0,000	Classe 18	$\hat{\lambda}_{18}$	-0,583	0,493	0,238
Disagio famiglia	$\hat{\gamma}_{10}$	0,128	0,040	0,001	Classe 19	$\hat{\lambda}_{19}$	-0,691	0,558	0,216
Sesso	$\hat{\gamma}_{20}$	0,469	0,133	0,000	Classe 20	$\hat{\lambda}_{20}$	0,159	0,542	0,769
Classe 1	$\hat{\lambda}_1$	0,628	0,589	0,286	Classe 21	$\hat{\lambda}_{21}$	-0,639	0,503	0,204
Classe 2	$\hat{\lambda}_2$	-0,481	0,486	0,323	Classe 22	$\hat{\lambda}_{22}$	0,559	0,504	0,268
Classe 3	$\hat{\lambda}_3$	-0,445	0,487	0,361	Classe 23	$\hat{\lambda}_{23}$	-0,016	0,610	0,979
Classe 4	$\hat{\lambda}_4$	-0,789	0,532	0,139	Classe 24	$\hat{\lambda}_{24}$	-0,842	0,490	0,086
Classe 5	$\hat{\lambda}_5$	0,467	0,542	0,389	Classe 25	$\hat{\lambda}_{25}$	0,360	0,464	0,439
Classe 6	$\hat{\lambda}_6$	-0,591	0,473	0,212	Classe 26	$\hat{\lambda}_{26}$	0,588	0,473	0,214
Classe 7	$\hat{\lambda}_7$	0,380	0,574	0,508	Classe 27	$\hat{\lambda}_{27}$	-0,966	0,482	0,046
Classe 8	$\hat{\lambda}_8$	-0,383	0,500	0,443	Classe 28	$\hat{\lambda}_{28}$	-0,715	0,613	0,244
Classe 9	$\hat{\lambda}_9$	-0,122	0,544	0,822	Classe 29	$\hat{\lambda}_{29}$	-1,046	0,589	0,076
Classe 10	$\hat{\lambda}_{10}$	-0,310	0,500	0,535	Classe 30	$\hat{\lambda}_{30}$	-1,419	0,484	0,004
Classe 11	$\hat{\lambda}_{11}$	-1,014	0,673	0,133	Classe 31	$\hat{\lambda}_{31}$	-0,406	0,510	0,426
Classe 12	$\hat{\lambda}_{12}$	0,486	0,548	0,375	Classe 32	$\hat{\lambda}_{32}$	-0,450	0,637	0,480
Classe 13	$\hat{\lambda}_{13}$	0,174	0,501	0,729	Classe 33	$\hat{\lambda}_{33}$	0,821	0,610	0,179
Classe 14	$\hat{\lambda}_{14}$	-0,194	0,714	0,786	Classe 34	$\hat{\lambda}_{34}$	0,197	0,507	0,698
Classe 15	$\hat{\lambda}_{15}$	-0,049	0,507	0,923	Classe 35	$\hat{\lambda}_{35}$	-0,081	0,462	0,861
Classe 16	$\hat{\lambda}_{16}$	0,778	0,485	0,109	Classe 36	$\hat{\lambda}_{36}$	-0,264	0,458	0,564
Classe 17	$\hat{\lambda}_{17}$	-0,405	0,589	0,492	Classe 37	$\hat{\lambda}_{37}$	-0,968	0,465	0,038

Il valore stimato del coefficiente di regressione associato alla variabile disagio dello studente in famiglia ($\hat{\gamma}_{10} = 0,128$) è, confermando quanto già detto in precedenza, significativo (*valore p* = 0,001). Il valore stimato del coefficiente di regressione della variabile sesso ($\hat{\gamma}_{20} = 0,469$) è anch'esso significativo (*valore p* = 0,000).

Tabella 6.3.4 Valori delle stime di alcuni parametri del modello MIC-2.

Numero Classe	Parametri stimati $\hat{\lambda}_j$	Residuo di classe $\hat{U}_{0,j}$	Intercetta $\hat{\beta}_{0j}$	Numero Classe	Parametri stimati $\hat{\lambda}_j$	Residuo di classe $\hat{U}_{0,j}$	Intercetta $\hat{\beta}_{0j}$
1	0,628	0,853	3,008	20	0,159	0,384	2,539
2	-0,481	-0,256	1,899	21	-0,639	-0,414	1,741
3	-0,445	-0,220	1,935	22	0,559	0,784	2,939
4	-0,789	-0,564	1,591	23	-0,016	0,209	2,364
5	0,467	0,692	2,847	24	-0,842	-0,617	1,538
6	-0,591	-0,366	1,789	25	0,360	0,585	2,740
7	0,380	0,605	2,760	26	0,588	0,813	2,968
8	-0,383	-0,158	1,997	27	-0,966	-0,741	1,414
9	-0,122	0,103	2,258	28	-0,715	-0,490	1,665
10	-0,310	-0,085	2,070	29	-1,046	-0,821	1,334
11	-1,014	-0,789	1,366	30	-1,419	-1,194	0,961
12	0,486	0,711	2,866	31	-0,406	-0,181	1,974
13	0,174	0,399	2,554	32	-0,450	-0,225	1,930
14	-0,194	0,031	2,186	33	0,821	1,046	3,201
15	-0,049	0,176	2,331	34	0,197	0,422	2,577
16	0,778	1,003	3,158	35	-0,081	0,144	2,299
17	-0,405	-0,180	1,975	36	-0,264	-0,039	2,116
18	-0,583	-0,358	1,797	37	-0,968	-0,743	1,412
19	-0,691	-0,466	1,689	38	0,000	0,225	2,380

Questo significa che l'incremento unitario della variabile sesso, ossia un maschio in più in classe, tenendo costante il disagio in famiglia, comporta un aumento in media di 0,469 punti il disagio manifestato nei confronti degli insegnanti. Considerando che la variabile disagio con gli insegnanti ha una media generale pari a 2,155, il numero di maschi in classe assume un certo rilievo nel determinare la valutazione delle classi rispetto al disagio manifestato dagli studenti nei confronti degli insegnanti.

Il disagio nei confronti degli insegnanti stimato per l' i -mo studente appartenente alla j -ma classe

$$\begin{aligned} \hat{Y}_{ij} &= (\hat{\gamma}_{00} + \hat{U}_{0j}) + 0,128c_{1ij} + 0,469s_{2ij} = \\ &= \hat{\beta}_{0j} + 0,128c_{1ij} + 0,469s_{2ij} \end{aligned}$$

dipende dal valore stimato $\hat{\beta}_{0j}$ per tutta la classe; dal valore della variabile c_{1ij} relativo al disagio manifestato dallo studente in famiglia e dal valore della variabile

s_{2ij} relativa al sesso dello studente. Le differenze tra classi si possono misurare in base ai valori residui $\hat{U}_{0,j}$ di classe o identicamente in base ai valori della stima $\hat{\beta}_{0,j}$ dell'intercetta (Tab.6.3.4).

In maniera simile a quanto fatto nel paragrafo precedente illustriamo il procedimento d'aggiustamento delle medie di classe in un modello ad intercetta casuale con due variabili esplicative (Tab.4.1). L'entità dell'aggiustamento del j -mo valor medio di classe,

$$-\hat{\gamma}_{10}(\bar{x}_{1j} - \bar{x}_1) - \hat{\gamma}_{20}(\bar{x}_{2j} - \bar{x}_2)$$

dipende da:

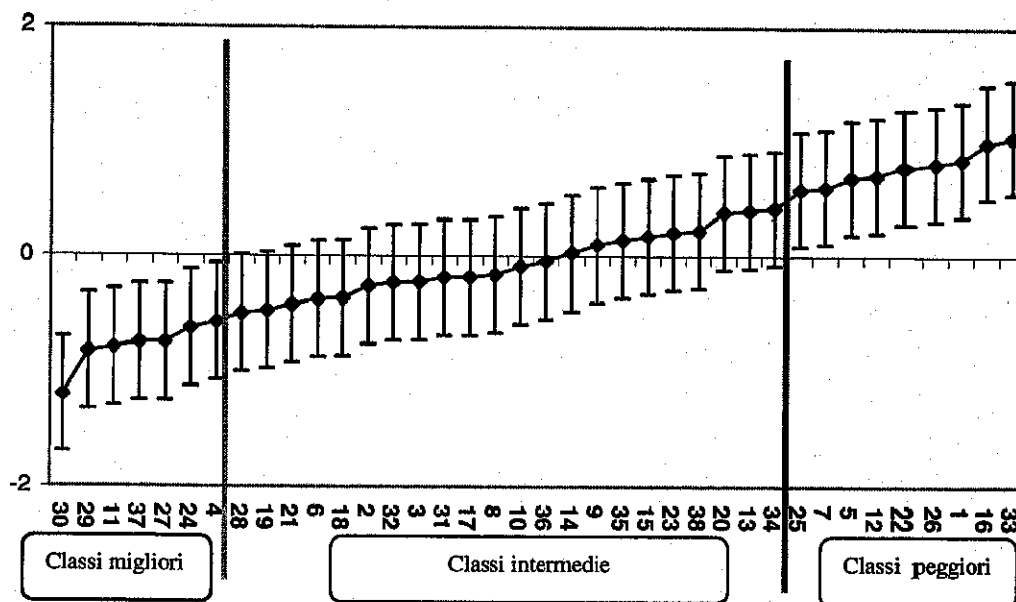
- la stima del parametro $\hat{\gamma}_{10}$ relativo al coefficiente della variabile disagio dello studente in famiglia ($\hat{\gamma}_{10}=0,128$);
- lo scostamento $(\bar{x}_{1j} - \bar{x}_1)$, con riferimento alla variabile disagio in famiglia, tra la media di classe (\bar{x}_{1j}) e la media generale ($\bar{x}_1=3,483$) di tutte le classi;
- la stima del parametro $\hat{\gamma}_{20}$ relativo al coefficiente della variabile sesso dello studente ($\hat{\gamma}_{20}=0,469$);
- lo scostamento $(\bar{x}_{2j} - \bar{x}_2)$, con riferimento alla variabile sesso, tra la media della classe (\bar{x}_{2j}) e la media generale ($\bar{x}_2 = 0,516$) di tutte le classi.

La logica dell'aggiustamento si estende con l'inserimento della nuova variabile esplicativa inserita ed in particolare quanto maggiore è la percentuale dei maschi presenti in classe e tanto minore sarà la valutazione del disagio medio manifestato dagli studenti nei confronti degli insegnanti.

L'operazione di confronto tra le classi in base alla stima $\hat{\beta}_{0,j}$ dell'intercetta può essere fatta ordinando in modo crescente tali valori e rappresentando graficamente le stime dell'intercetta con i relativi intervalli di confidenza (Fig. 6.3.2). L'asse orizzontale, tracciato in corrispondenza del valore zero rappresenta sempre la media

generale delle classi ($\hat{\gamma}_{00} = 2,155$) e la losanga rappresenta lo scostamento da tale valore della stima $\hat{\beta}_{0j}$ dell'intercetta di classe.

Figura 6.3.2 *Graduatoria delle classi in base alla stima dell'intercetta nel modello MIC-2.*



Anche in questo caso dal confronto tra la graduatoria delle classi ottenute mediante il modello MIC-1 (Modello ad Intercetta Casuale con 1 variabili esplicative) e quella ottenuta col modello MIC-2 (Modello ad Intercetta Casuale con 2 variabili esplicative) si può osservare l'effetto dovuto all'introduzione della variabile esplicativa s_{2ij} relativa al sesso dello studente nella misura del disagio manifestato dalle classi nei confronti degli insegnanti.

6.4 Il disagio nei confronti della scuola

In questo paragrafo vogliamo considerare il legame esistente tra il disagio manifestato dallo studente nei confronti della scuola nel suo complesso e il disagio manifestato dallo studente nei confronti degli insegnanti e come tale legame modifica le valutazioni delle singole classi. Perciò, introduciamo nel modello statistico (MIC-3) un'ulteriore variabile esplicativa x_{3ij} che rappresenta il disagio dello studente nei confronti della scuola e, come abbiamo fatto per tutte le altre variabili esplicative, la esprimiamo in termini di scarto dal suo valore medio ($x_{3ij} - \bar{x}_3$) indicandola con q_{3ij} secondo il modello

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \gamma_{10} c_{1ij} + \gamma_{20} s_{2ij} + \gamma_{30} q_{3ij} + R_{ij} \\ \beta_{0j} &= \gamma_{00} + U_{0j} \end{aligned} \tag{6.4.1}$$

dove:

β_{0j} = intercetta della j -mo iperpiano di regressione;

$c_{1ij} = (x_{1ij} - \bar{x}_1)$, disagio in famiglia centrato sulla media generale;

$s_{2ij} = (x_{2ij} - \bar{x}_2)$, sesso centrato sulla media generale;

$q_{3ij} = (x_{3ij} - \bar{x}_3)$, disagio con la scuola centrato sulla media generale;

Y_{ij} = valore osservato di disagio nei confronti degli insegnanti dello i -mo studente appartenente alla j -ma classe;

γ_{00} = media generale del disagio degli studenti con gli insegnanti;

γ_{10} = variazione di Y_{ij} associata ad una variazione unitaria di x_{1ij} ,
tenendo costante x_{2ij} e x_{3ij} ;

γ_{20} = variazione di Y_{ij} associata ad una variazione unitaria di x_{2ij} ,
tenendo costante x_{1ij} e x_{3ij} ;

γ_{30} = variazione di Y_{ij} associata ad una variazione unitaria di x_{3ij} ,
tenendo costante x_{1ij} e x_{2ij} .

Anche in questo caso dalla tabella di analisi della varianza (Tab.6.4.1) possiamo stabilire dal *valore p* la significatività dei parametri associati alle variabili disagio nei confronti della scuola (*valore p* = 0,000), sesso (*valore p* = 0,003), disagio in famiglia (*valore p* = 0,001) e la classe d'appartenenza (*valore p* = 0,000). Da notare che l'inserimento della terza variabile esplicativa relativa al disagio dello studente nei confronti della scuola non riduce la significatività dei parametri associati alle variabili esplicative precedentemente inserite.

Tabella 6.4.1. Risultati dell'analisi della varianza del disagio nei confronti degli insegnanti mediante il modello MIC-3.

Dependent Variable: Disagio con gli insegnanti

Fonte di variabilità		Somme dei quadrati	Gradi di libertà	Medie dei quadrati	F	Valore P
Intercetta	Hypothesis	2424,392	1	2424,392	545,395	,000
	Error	185,245	41,673	4,445		
Disagio in famiglia	Hypothesis	24,902	1	24,902	10,791	,001
	Error	1271,517	551	2,308		
Sesso	Hypothesis	20,709	1	20,709	8,974	,003
	Error	1271,517	551	2,308		
Disagio a scuola	Hypothesis	31,787	1	31,787	13,775	,000
	Error	1271,517	551	2,308		
Classe	Hypothesis	174,392	37	4,713	2,042	,000
	Error	1271,517	551	2,308		

Le componenti di varianza condizionate dalle variabili esplicative inserite e quindi il coefficiente di correlazione intraclassa (Tab.6.4.2) assumono dei valori diversi da quelli ottenuti in precedenza e la componente di varianza *between* diminuisce da 0,171 a 0,156 mentre la componente di varianza *within* si riduce da 2,361 a 2,308. Tutto questo significa che la variabile disagio dello studente nei confronti della scuola:

- **diminuisce il grado di eterogeneità tra classi (*between*), in quanto porta alla luce delle somiglianze tra classi rispetto al disagio manifestato dagli studenti nei confronti degli insegnanti;**

- contribuisce a spiegare il disagio manifestato dagli studenti nei confronti degli insegnanti riducendo quindi la variabilità residua (*within*).

Tabella 6.4.2. Stima delle componenti di varianza nel modello MIC-3 e relativo coefficiente di correlazione intraclasse.

Componente di varianza	Stima minimi quadrati
Tra classi $\hat{\tau}_0^2$	0,156
Dentro le classi $\hat{\sigma}^2$	2,308
Totale	2,464

Coefficiente di correlazione intraclasse	Valore
$\rho \mid x = \frac{\tau_0^2}{\sigma^2 + \tau_0^2}$	$\frac{0,156}{2,464} = 6,33\%$

Per quanto riguarda le stime ottenute col metodo dei minimi quadrati ordinari dei parametri incogniti (Tab.6.4.3) del modello ad intercetta casuale nella versione completa

$$Y_{ij} = \alpha_{00}x_{i0} + \gamma_{10}c_{1ij} + \gamma_{20}s_{2ij} + \gamma_{30}q_{3ij} + \lambda_{01}x_1 + \lambda_{02}x_2 + \dots + \lambda_{37}x_{37} + R_{ij} \quad (6.4.2)$$

e la loro interpretazione, la stima dei parametri

$$\hat{\alpha}_{00} = \hat{\gamma}_{00} + \hat{U}_{0,38} = 2,155 + \hat{U}_{0,38} = 2,498$$

$$\hat{U}_{0,38} = 2,498 - 2,155 = 0,343$$

relativi all'intercetta e all'effetto classe per la classe 38, consentono di determinare ed interpretare i parametri restanti che compaiono nel modello completo (6.4.2).

La stima dell'intercetta ($\hat{\alpha}_{00} = 2,498$) rappresenta la somma tra il valore medio $\hat{\gamma}_{00}$ della variabile disagio nei confronti degli insegnanti con il valore residuo $\hat{U}_{0,38}$ della classe 38, conoscendo il primo valore si trova per differenza il secondo ($\hat{U}_{0,38} = 0,343$).

Tabella 6.4.3 Stime dei minimi quadrati dei parametri del modello MIC-3.

Nome variabile	Parametro	Stime OLS	Errore Std.	Valore P	Nome variabile	Parametro	Stime OLS	Errore Std.	Valore P
Intercetta	$\hat{\alpha}_{00}$	2,498	0,332	0,000	Classe 18	$\hat{\lambda}_{18}$	-1,048	0,503	0,038
Disagio famiglia	\hat{c}_{1ij}	0,130	0,040	0,001	Classe 19	$\hat{\lambda}_{19}$	-0,810	0,552	0,143
Sesso	\hat{s}_{2ij}	0,399	0,133	0,003	Classe 20	$\hat{\lambda}_{20}$	0,119	0,535	0,824
Disagio scuola	\hat{q}_{3ij}	0,249	0,067	0,000	Classe 21	$\hat{\lambda}_{21}$	-0,680	0,497	0,172
Classe 1	$\hat{\lambda}_1$	0,645	0,582	0,268	Classe 22	$\hat{\lambda}_{22}$	0,261	0,504	0,604
Classe 2	$\hat{\lambda}_2$	-0,580	0,481	0,228	Classe 23	$\hat{\lambda}_{23}$	-0,069	0,603	0,908
Classe 3	$\hat{\lambda}_3$	-0,628	0,484	0,195	Classe 24	$\hat{\lambda}_{24}$	-0,862	0,484	0,076
Classe 4	$\hat{\lambda}_4$	-0,941	0,528	0,075	Classe 25	$\hat{\lambda}_{25}$	0,256	0,460	0,579
Classe 5	$\hat{\lambda}_5$	0,274	0,538	0,611	Classe 26	$\hat{\lambda}_{26}$	0,518	0,468	0,268
Classe 6	$\hat{\lambda}_6$	-0,662	0,468	0,158	Classe 27	$\hat{\lambda}_{27}$	-0,976	0,477	0,041
Classe 7	$\hat{\lambda}_7$	0,154	0,571	0,787	Classe 28	$\hat{\lambda}_{28}$	-0,741	0,606	0,221
Classe 8	$\hat{\lambda}_8$	-0,521	0,495	0,293	Classe 29	$\hat{\lambda}_{29}$	-1,101	0,583	0,059
Classe 9	$\hat{\lambda}_9$	-0,251	0,539	0,641	Classe 30	$\hat{\lambda}_{30}$	-1,483	0,479	0,002
Classe 10	$\hat{\lambda}_{10}$	-0,360	0,494	0,466	Classe 31	$\hat{\lambda}_{31}$	-0,642	0,509	0,207
Classe 11	$\hat{\lambda}_{11}$	-1,024	0,665	0,124	Classe 32	$\hat{\lambda}_{32}$	-0,503	0,630	0,425
Classe 12	$\hat{\lambda}_{12}$	0,292	0,545	0,592	Classe 33	$\hat{\lambda}_{33}$	0,631	0,605	0,298
Classe 13	$\hat{\lambda}_{13}$	0,131	0,496	0,792	Classe 34	$\hat{\lambda}_{34}$	0,094	0,502	0,852
Classe 14	$\hat{\lambda}_{14}$	-0,507	0,711	0,476	Classe 35	$\hat{\lambda}_{35}$	-0,174	0,458	0,705
Classe 15	$\hat{\lambda}_{15}$	-0,365	0,509	0,474	Classe 36	$\hat{\lambda}_{36}$	-0,354	0,453	0,435
Classe 16	$\hat{\lambda}_{16}$	0,568	0,482	0,240	Classe 37	$\hat{\lambda}_{37}$	-1,005	0,460	0,029
Classe 17	$\hat{\lambda}_{17}$	-0,497	0,582	0,394					

Gli altri valori residui $\hat{U}_{0,j}$ si trovano sempre come somma tra il valore del j -mo parametro stimato ($\hat{\lambda}_j$) e il residuo $\hat{U}_{0,38}$ relativo alla classe 38 (Tab.6.4.4).

Tabella 6.4.4 Valori delle stime di alcuni parametri del modello MIC-3.

Numero Classe	Parametri stimati $\hat{\lambda}_j$	Residuo di classe $\hat{U}_{0,j}$	Intercetta $\hat{\beta}_{0j}$	Numero Classe	Parametri stimati $\hat{\lambda}_j$	Residuo di classe $\hat{U}_{0,j}$	Intercetta $\hat{\beta}_{0j}$
1	0,645	0,968	3,143	20	0,119	0,462	2,617
2	-0,580	-0,237	1,918	21	-0,680	-0,337	1,818
3	-0,628	-0,285	1,87	22	0,261	0,604	2,759
4	-0,941	-0,598	1,557	23	-0,069	0,274	2,429
5	0,274	0,617	2,772	24	-0,862	-0,519	1,636
6	-0,662	-0,319	1,836	25	0,256	0,599	2,754
7	0,154	0,497	2,652	26	0,518	0,861	3,016
8	-0,521	-0,178	1,977	27	-0,976	-0,633	1,522
9	-0,251	0,092	2,247	28	-0,741	-0,398	1,757
10	-0,360	-0,017	2,138	29	-1,101	-0,758	1,397
11	-1,024	-0,681	1,474	30	-1,483	-1,14	1,015
12	0,292	0,635	2,79	31	-0,642	-0,299	1,856
13	0,131	0,474	2,629	32	-0,503	-0,16	1,995
14	-0,507	-0,164	1,991	33	0,631	0,974	3,129
15	-0,365	-0,022	2,133	34	0,094	0,437	2,592
16	0,568	0,911	3,066	35	-0,174	0,169	2,324
17	-0,497	-0,154	2,001	36	-0,354	-0,011	2,144
18	-1,048	-0,705	1,45	37	-1,005	-0,662	1,493
19	-0,810	-0,467	1,688	38	0,000	0,343	2,498

Le stime dei coefficienti di regressione associati alle variabili esplicative inserite sono tutte significative e l'incremento unitario della variabile disagio nei confronti della scuola, tenendo costanti i valori delle altre due variabili, comporta un aumento, in media, di 0,249 punti il disagio manifestato dagli studenti nei confronti degli insegnanti.

Il disagio nei confronti degli insegnanti stimato per l' i -mo studente appartenente alla j -ma classe

$$\begin{aligned} \hat{Y}_{ij} &= (\hat{\gamma}_{00} + \hat{U}_{0j}) + 0,130c_{1ij} + 0,399s_{2ij} + 0,249q_{3ij} \\ &= \hat{\beta}_{0j} + 0,130c_{1ij} + 0,399s_{2ij} + 0,249q_{3ij} \end{aligned}$$

dipende dal valore stimato $\hat{\beta}_{0j}$ per tutta la classe; dal valore della variabile c_{1ij} relativo al disagio manifestato dallo studente in famiglia; dal valore della variabile s_{2ij} relativo al sesso dello studente e dal valore della variabile q_{3ij} relativo al disagio dello studente nei confronti della scuola nel suo complesso.

Le differenze tra le classi si possono misurare in base ai valori residui $\hat{U}_{0,j}$ di classe, o identicamente in base ai valori della stima $\hat{\beta}_{0,j}$ dell'intercetta (Tab.6.4.5).

In maniera analoga a quanto è stato fatto in precedenza, si può illustrare il procedimento d'aggiustamento delle medie di classe nel modello di analisi della covarianza con tre variabili esplicative, considerando che la logica di fondo rimane inalterata (Tab.4.1). L'entità dell'aggiustamento del j -mo valore medio di classe

$$-\hat{\gamma}_{10}(\bar{x}_{1j} - \bar{x}_1) - \hat{\gamma}_{20}(\bar{x}_{2j} - \bar{x}_2) - \hat{\gamma}_{30}(\bar{x}_{3j} - \bar{x}_3)$$

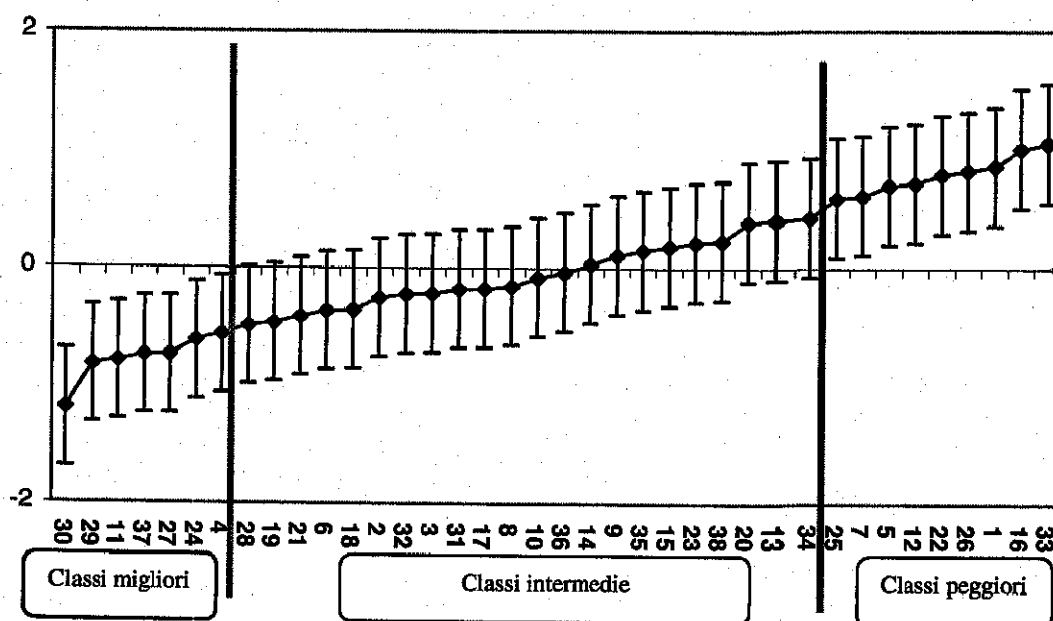
dipende da:

- la stima del parametro $\hat{\gamma}_{10}$ relativo al coefficiente della variabile disagio dello studente in famiglia ($\hat{\gamma}_{10}=0,130$);
- lo scostamento $(\bar{x}_{1j} - \bar{x}_1)$, con riferimento al disagio in famiglia, tra la media della classe (\bar{x}_{1j}) e la media generale ($\bar{x}_1=3,483$) di tutte le classi;
- la stima del parametro $\hat{\gamma}_{20}$ relativo al coefficiente della variabile sesso dello studente ($\hat{\gamma}_{20}=0,399$);
- lo scostamento $(\bar{x}_{2j} - \bar{x}_2)$, con riferimento alla variabile sesso, tra la media della classe (\bar{x}_{2j}) e la media generale ($\bar{x}_2 = 0,516$) di tutte le classi;
- la stima del parametro $\hat{\gamma}_{30}$ relativo al coefficiente della variabile disagio dello studente nei confronti della scuola ($\hat{\gamma}_{30}=0,249$);
- lo scostamento $(\bar{x}_{3j} - \bar{x}_3)$, con riferimento alla variabile disagio nei confronti della scuola, tra la media della classe (\bar{x}_{3j}) e la media generale ($\bar{x}_3 = 0,606$) di tutte le classi.

La **logica dell'aggiustamento** richiede la considerazione della nuova variabile esplicativa inserita, il disagio nei confronti della scuola, in particolare quanto maggiore è il disagio medio di classe manifestato verso la scuola, tanto minore sarà la valutazione del disagio medio di classe manifestato nei confronti degli insegnanti.

L'operazione di confronto tra le classi in base alla stima $\hat{\beta}_{0j}$ dell'intercetta è stata fatta ordinando nuovamente in modo crescente tali valori e rappresentando graficamente i relativi intervalli di confidenza (Fig. 6.4.1). L'asse orizzontale, tracciato in corrispondenza del valore zero rappresenta sempre la media generale delle classi ($\hat{\gamma}_{00} = 2,155$) e la losanga lo scostamento da tale valore della stima $\hat{\beta}_{0j}$ dell'intercetta di classe.

Figura 6.4.1 *Graduatoria delle classi in base alla stima dell'intercetta nel modello MIC-3.*



Il confronto tra la graduatoria delle classi ottenute mediante il modello MIC-2 (Modello ad Intercetta Casuale con 2 variabili esplicative) e quella ottenuta col modello MIC-3 (Modello ad Intercetta Casuale con 3 variabili esplicative) consente di osservare l'effetto dovuto all'introduzione della variabile esplicativa q_{3ij} relativa al disagio dello studente nei confronti della scuola nella misura del disagio manifestato dalle classi nei confronti degli insegnanti.

7. Conclusioni

In questo lavoro ci si siamo proposti di illustrare uno strumento statistico per lo studio di un tema che è oggetto di interesse attualmente, cioè la valutazione di un sistema formativo. Facendo riferimento alle esperienze internazionali di valutazione dell'efficacia scolastica⁷ (*Multilevel Project: Institute of Education, University of London*) abbiamo considerato alcune variabili, modelli e programmi statistici rilevanti per gli obiettivi di ricerca. Avvalendoci della collaborazione dei responsabili⁸ di un progetto innovativo di rilevazione sul disagio scolastico abbiamo applicato alcuni modelli multilivello per la misura del disagio manifestato dallo studente nei confronti degli insegnanti. A partire dai dati raccolti nell'ambito di questo progetto si è proposto un metodo per valutare il disagio manifestato dalle classi nei confronti degli insegnanti, tenendo conto di alcune variabili esplicative individuali dello studente: il disagio in famiglia, il sesso e il disagio nei confronti della scuola.

I risultati ottenuti dall'analisi riguardano:

- l'esistenza di differenze tra classi in termini di disagio manifestato dagli studenti;
- l'esistenza di una correlazione tra le variabili esplicative individuali degli studenti e il disagio manifestato dagli studenti nei confronti degli insegnanti.

In particolare, è emerso che il sesso, il disagio in famiglia e quello manifestato verso la scuola rappresentano delle variabili rilevanti al fine di spiegare il disagio manifestato dalle classi nei confronti degli insegnanti.

Inoltre, l'uso dei parametri che permettono di misurare il disagio manifestato dalle classi nei confronti degli insegnanti consente di individuare

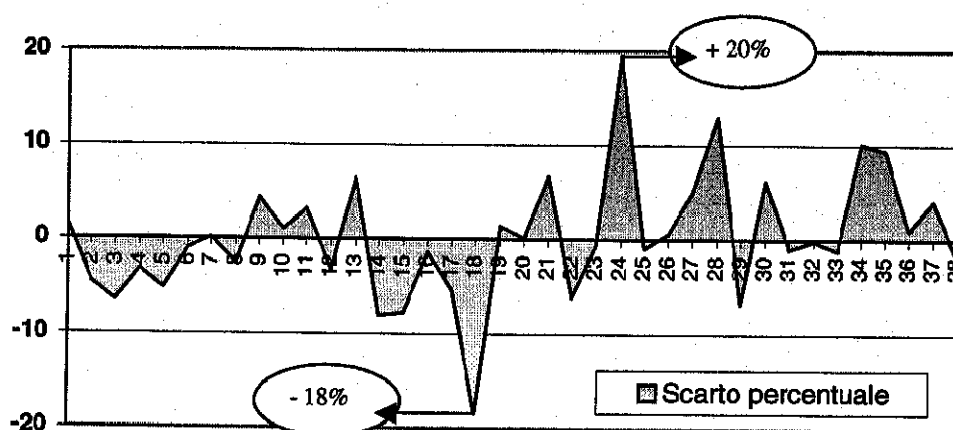
⁷ Multilevel Project: <http://multilevel.ioe.ac.uk>

⁸ Prof.ssa Nuti (responsabile e coordinatrice del progetto di rilevazione, Consorzio interuniversitario Formazione Innovazione e Tecnologie) e prof. Bianchi (responsabile didattico per tutte le scuole del Comune di Massa).

tre gruppi di classi (*migliori, intermedie e peggiori*) in funzione di tale disagio. Infine, sulla base dei diversi modelli formalizzati (MIC-0; MIC-1; MIC-2; MIC-3) si è potuta osservare una differenza tra le valutazioni grezze e nette nella misura del disagio delle classi tenuto conto di volta in volta di un numero diverso di variabili esplicative individuali dello studente.

Al fine di 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 dei parametri, nel caso del modello ad intercetta casuale *privo* di variabili esplicative individuali o nel caso del modello ad intercetta casuale che include *tutte* le variabili esplicative individuali prese in esame. Dalla variazione percentuale del parametro relativo all'intercetta che misura per ogni classe il disagio nei confronti degli insegnanti, passando dal modello più semplice (MIC-0) a quello più complesso (MIC-3), si nota (Fig.7.1) che le variabili individuali incidono sensibilmente nelle valutazioni del disagio manifestato dalle classi verso gli insegnanti.

Figura 7.1 *Variazione percentuale dei valori delle stime dell'effetto classe nel modello MIC-0 e MIC-3.*



Perciò, non si possono fare simili confronti tra classi senza tenere conto delle disomogeneità esistenti al loro interno con riferimento alle caratteristiche individuali degli studenti. Infatti, nella valutazione del disagio delle classi

verso gli insegnanti la variazione della stima del parametro relativo all'effetto classe raggiunge il 20% e modifica la posizione delle classi stesse nella loro graduatoria formata in base alla misura di questo tipo di disagio.

Pertanto, l'uso dei modelli lineari ad intercetta casuale per la valutazione dell'efficacia è certamente utile ai fini conoscitivi, ma per la sua applicazione è necessario disporre di informazioni adeguate sulle caratteristiche degli Utenti e degli Agenti formativi, informazioni che possono essere ricavate programmando opportunamente delle indagini finalizzate alla raccolta dei dati statistici interessanti.

BIBLIOGRAFIA

- AITKIN M., LONGFORD N., (1986), Statistical Modelling Issues in School Effectiveness Studies, *Journal of the Royal Statistical Society, A*, vol.149, part.1, pp.1-43.
- BOHRNSTEDT G.W. & KNOKE D., (1998) *Statistica per le scienze sociali*, Ed.Mulino, Bologna.
- BRYK, A.S. & RAUDENBUSH, S.W. (1992), *Hierarchical Linear Models*, Newbury Park, CA: Sage.
- FREUND R.J., WILSON W.J. (2001), *Metodi statistici*, Piccin Nuova Libreria, Padova.
- GOLDSTEIN H. (1997), Methods in school effectiveness research. *School effectiveness and school improvement*. 8: 369-95.
- GOLDSTEIN H., HEALY M.J.R. (1995), The graphical presentation of a collection of means, *Journal of the Royal Statistical Society, A*, vol.158, pp.175-177.
- GORI E. (2000), La valutazione nei servizi alla persona: prospettive, problemi e impostazione. *Atti del Convegno su "Valutazione dei servizi alla persona"*, 25-26 Febbraio, Dipartimento di Statistica, Firenze.
- GORI E., VITTADINI G., a cura di, (1999), *Qualità e valutazione nei servizi di pubblica utilità*, serie Gestione d'impresa – Direzione, ETAS, Milano.
- LONGFORD, N. (1993), *Random Coefficient Models*. Clarendon Press, Oxford.
- NETER J., WASSERMANN W., (1974), *Applied Linear Statistical Models*, Richard D.Irwin, Inc. Illinois.

SEARLE S.R., (1997), *Linear Models*, Wiley and Sons, New York.

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

SNIJDERS, T.A.B. & BOSKER, R. (1999), *Multilevel Analysis. An introduction to Basic and Advanced Multilevel Modelling*. Sage, London.

STAME N., (2001) *Approcci e problemi nella valutazioni di politiche e di programmi nella realtà del nuovo millennio. Atti del Convegno Intermedio SIS, 4-6 giugno, Roma.*