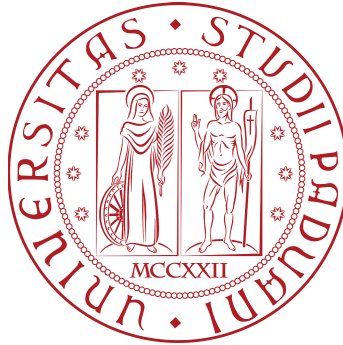


Università degli Studi di Padova
Dipartimento di Scienze Statistiche



**Corso di Laurea Triennale in Statistica
per le Tecnologie e le Scienze**

Relazione finale

**Sorveglianza di un processo multilivello: un caso di
studio per prevedere il successo o l'insuccesso degli
studenti**

Relatore: Prof. Guido Masarotto
Dipartimento di Scienze Statistiche

Laureanda: Bacelle Lisa
Matricola N. 1224555

Anno Accademico 2022/23

Indice

Introduzione	5
1 Descrizione dei dati	7
1.1 Il sistema delle scuole superiori olandesi	7
1.2 Variabili e dataset	8
1.3 Determinanti del rendimento degli studenti	11
2 Modello gerarchico	14
2.1 Il modello	14
2.2 Stima dei parametri	16
2.3 Risultati e misure di accuratezza	17
3 Monitoraggio delle prestazioni degli studenti	21
3.1 Monitoraggio statistico dei processi	21
3.1.1 Diagramma di controllo a tre livelli	22
3.1.2 Esempio	24
3.2 Monitoraggio predittivo	25
3.2.1 Monitoraggio predittivo multilivello	26
3.2.2 Esempio	27

Conclusioni	33
Bibliografia	35

Introduzione

Negli ultimi anni, a causa di rapidi cambiamenti della società attuale e del sistema scolastico, è emerso il bisogno di monitorare gli studenti durante tutta la loro carriera scolastica. Attraverso i rapporti sugli indicatori di allerta precoce, vengono avvertiti gli insegnanti e il personale, degli studenti con elevati rischi di fallimento o abbandono scolastico.

La modellazione multilivello è un buon metodo in contesti educativi e può essere utilizzato per il monitoraggio predittivo. Con l'aumentare della quantità e della complessità dei dati disponibili è sorta l'esigenza di sostituire l'utilizzo di metodi univariati per il monitoraggio di un singolo indicatore di qualità con metodi per il monitoraggio di complessi processi multivariati, come ad esempio il monitoraggio del profilo.

Questa tesi prende in considerazione l'approccio introdotto da Leo C.E. Huberts, Marit Schoonhoven e Ronald J.M.M. Does nell'articolo: *Multilevel process monitoring: A case study to predict student process or failure* (2020).

Questa relazione ha lo scopo di rispondere a 3 domande di ricerca relative ai progressi degli studenti delle scuole superiori durante un anno accademico: (1) cosa determina il rendimento degli studenti? (2) come può essere utilizzato il monitoraggio del processo statistico per sorvegliare i progressi degli studenti? (3) quale metodo può essere utilizzato per il monitoraggio predittivo dei risultati degli studenti?

L'elaborato si compone di 3 capitoli.

Nel primo verrà introdotto il sistema della scuola superiore olandese, il dataset con le relative variabili studiate e i fattori determinanti del rendimento degli studenti.

Il secondo capitolo definirà il modello utilizzato con particolare attenzione alla stima dei parametri, ai risultati e alle misure di accuratezza.

Il terzo e ultimo capitolo verrà suddiviso in due sezioni: il monitoraggio statistico dei processi, il quale include la definizione del diagramma di controllo a 3 livelli, ed il monitoraggio predittivo con la spiegazione del procedimento; per entrambi verrà fornito un esempio.

Capitolo 1

Descrizione dei dati

In questa prima sezione viene illustrato il sistema scolastico delle scuole superiori olandesi, per poi introdurre il dataset e le relative variabili. Si analizza, inoltre, attraverso dei grafici, quali di esse possono essere considerate possibili determinanti del rendimento degli studenti.

1.1 Il sistema delle scuole superiori olandesi

Il sistema scolastico olandese consiste in 8 anni di scuola primaria, o *Basisschool*, che è il percorso di studi che accoglie bambini dai 4 ai 12 anni, seguito da 4, 5 o 6 anni di scuola superiore. Negli ultimi anni sono stati definiti due criteri fondamentali per determinare il livello di scuola superiore che un ragazzo può frequentare: il consiglio dell'insegnante, il quale ha un peso rilevante dal momento che indirizza lo studente verso la scuola più adatta in base alle sue competenze; il test del National Institute for Test Development: si tratta di un test, anche detto *CITO*, avviato nei Paesi Bassi nel 1966 da A.D de Groot, scientificamente progettato per analizzare le capacità di un bambino. Dal 2014 ogni scuola è tenuta a condurlo al termine dell'ottavo anno di scuola primaria. Se gli studenti dovessero trovarsi in disaccordo con la strada consigliata hanno comunque la possibilità di frequentare un anno di orientamento.

La scuola secondaria olandese si divide in tre percorsi di studio: il VMBO, o scuola secondaria professionale, che copre la fascia d'età 12-16 anni; l'HAVO, o scuola secondaria generale, che copre la fascia d'età 12-17 anni; il VWO, o scuola secondaria pre-universitaria, che copre la fascia d'età 12-18 anni. Il sistema scolastico olandese permette di scegliere alcune materie per personalizzare il proprio piano di studi. In media gli alunni seguono una decina di corsi.

La scala dei voti va da 1 a 10 e per superare l'anno scolastico devono essere soddisfatte alcune condizioni stabilite dalla scuola. Un corso è considerato *insufficiente* se il voto arrotondato risulta minore di 6. Una scuola potrebbe far ripetere l'anno se si ottengono più di due punti di errore ("*failpoints*"), ossia quando alla fine dell'anno si ha una materia con voto pari a 3, oppure tre materie con voto pari a 5, o ancora, una materia con il 4 ed una con il 5.

Dal quarto anno delle superiori in poi, gli studenti si dividono in diversi profili di scuola superiore, nei quali hanno maggiore rilevanza i corsi specifici di quel profilo.

1.2 Variabili e dataset

Per procedere con l'analisi, vengono riportate di seguito le variabili utilizzate e i loro effetti. Vengono usati i nomi in inglese per comodità in quanto le figure presenti di seguito fanno riferimento ad essi:

- *SES*: variabile esplicativa relativa allo Stato socioeconomico degli studenti;
- *Disability*: variabile esplicativa relativa alla presenza di disabilità;
- *Language*: variabile esplicativa relativa alla lingua;
- *Non-native*: variabile esplicativa relativa alla presenza di non madrelingua nella scuola;
- *Student effort*: variabile che consente l'eterogeneità non osservata e che indica lo sforzo degli studenti;

- *Peer associations*: variabile che consente l'eterogeneità non osservata relativa alle associazioni tra studenti;
- *Parent involvement*: variabile che consente l'eterogeneità non osservata e che indica il coinvolgimento dei genitori nelle questioni scolastiche;
- *School climate*: variabile che consente l'eterogeneità non osservata relativa al clima scolastico;
- *Intelligence*: variabile esplicativa inosservata relativa all'intelligenza degli studenti;
- *Grades*: variabile esplicativa dipendente dal tempo che fornisce i voti degli studenti;
- *Absences*: variabile esplicativa dipendente dal tempo che indica le assenze degli studenti durante l'anno.

Lo Status socioeconomico (SES) influisce in modo significativo sul rendimento degli studenti. Geiser e Santelices nel 2007 hanno sostenuto che l'omissione di questo fattore può portare ad una sovrastima del potere predittivo delle variabili accademiche. L'ipotesi è basata su uno studio di Rothstein, il quale ha affermato che l'esclusione di esso dai modelli di previsione gonfia l'apparente validità dei test di ammissione all'Università del 150%.

Per quanto concerne le disabilità, anch'esse hanno un effetto significativo, in quanto influiscono negativamente sulle performance degli studenti. I bambini dislessici non riescono a raggiungere voti scolastici commisurati alla loro intelligenza. Anche asma, epilessia e autismo sono fattori rilevanti. Da non sottovalutare i problemi medici, come ad esempio problemi di vista, di udito i quali possono causare delle notevoli difficoltà. Un altro importante aspetto da considerare riguarda la lingua. Nel 1995 Collier ha dichiarato che gli immigrati e gli studenti delle minoranze linguistiche hanno bisogno di 4-12 anni di sviluppo della seconda lingua. Inoltre, la presenza di non madrelingua nelle scuole danneggia le prestazioni degli studenti madrelingua.

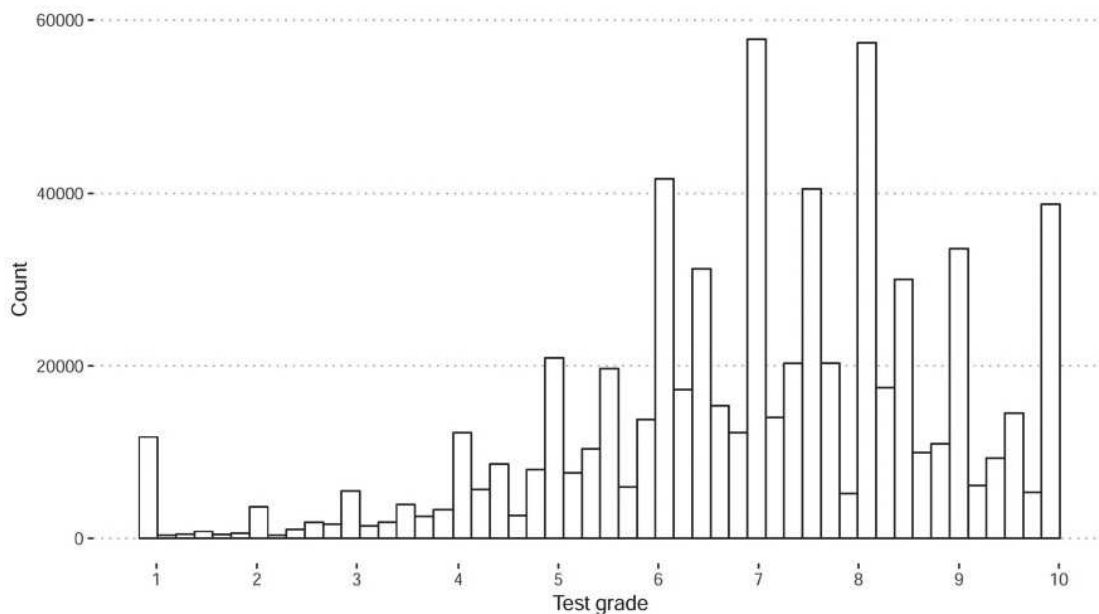


Figura 1.1: Istogramma dei singoli voti del test nei dati.

Alcune variabili, invece, rimangono inosservate ma possono essere incorporate nei modelli consentendo l'eterogeneità non osservata.

La prima di queste è lo sforzo degli studenti, caratterizzato dal coinvolgimento e dall'impegno scolastico. Anche le associazioni tra studenti influiscono molto sul rendimento di quest'ultimi. Inoltre, il coinvolgimento dei genitori risulta un importante elemento nell'analisi. Sui-Chu e Willms hanno affermato che la discussione a casa è un aspetto fondamentale, in quanto vengono fornite informazioni concrete ai genitori sui metodi di insegnamento e sui programmi previsti per l'anno scolastico. Due variabili considerate difficili da misurare risultano essere il clima scolastico e l'intelligenza, approssimata utilizzando i punteggi dei test della scuola primaria.

Per finire, si analizzano due variabili che cambiano nel tempo: i voti e le assenze. La prima è considerata l'indicatore più importante del rendimento degli studenti. In qualsiasi momento durante l'anno, questi test determinano un voto medio ponderato corrente per ogni studente e corso. Dalla Figura 1.1 si può notare come alla maggioranza degli studenti nei test individuali venga attribuito un voto pari a circa $7/8$.

Per quanto riguarda le assenze, sono anch'esse significative poiché frequentare le lezioni aiuta i ragazzi a comprendere il materiale e motiva la loro partecipazione. La modellazione terrà conto della variazione delle caratteristiche degli studenti nel tempo: essi cambieranno l'allocazione del loro impegno e del loro tempo in base al loro voto medio attuale, al voto medio per altri corsi, agli effetti stagionali, ai cambiamenti all'interno della scuola e ai fattori esterni.

Si hanno a disposizione 8 anni di dati i quali comprendono: 36 differenti materie, 1700 studenti delle superiori (di cui il 51% ragazze) e 711.653 test individuali. Gli studenti sono nati in 38 paesi diversi, parlano 18 lingue differenti e hanno avuto 110 insegnanti. Dei 1700 studenti, 326 hanno disabilità, 162 non risultano essere di nazionalità olandese e 51 possiedono una grave barriera linguistica. Il numero di studenti con genitori che hanno frequentato l'università o accademici di livello superiore è di 261.

Per incorporare lo stato socioeconomico (SES) vengono utilizzati i dati sullo stato sociale a livello nazionale forniti dal governo olandese. Sono stati aggiunti al dataset il punteggio SES di ogni studente e i disturbi dell'apprendimento, tra cui la Dislessia e il Disturbo da Deficit dell'Attenzione/Iperattività.

1.3 Determinanti del rendimento degli studenti

Dal Boxplot in Figura 1.2 si possono osservare le diverse distribuzioni per le variabili predittive categoriali più importanti.

La prima coppia di boxplot è relativa al sesso degli studenti; si nota che le ragazze sembrano superare i ragazzi in termini di voti finali. Procedendo con l'analisi, si può osservare come gli studenti con disabilità ottengano voti di fine anno più bassi. Per quanto riguarda, inoltre, i bambini di genitori altamente istruiti sembrano avere prestazioni leggermente migliori con voto di fine anno pari a circa 7.4. Concludendo, un'affermazione differente si può fare, invece, per gli studenti con una lingua madre diversa, i quali ottengono voti di poco inferiori; al contrario, i bambini nati al di fuori dei Paesi Bassi sembrano non avere prestazioni

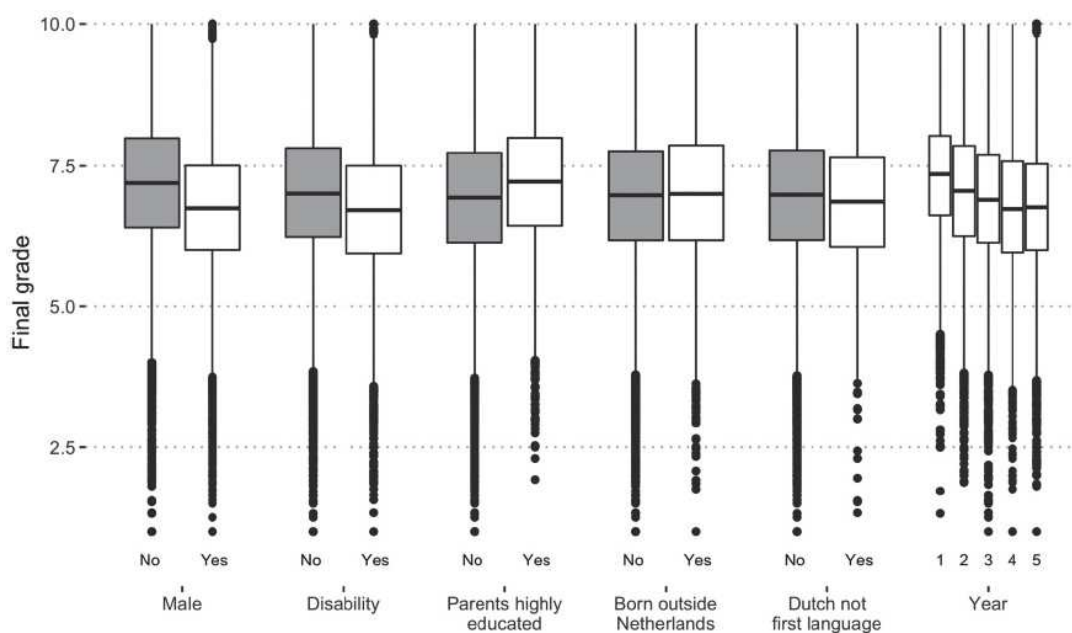
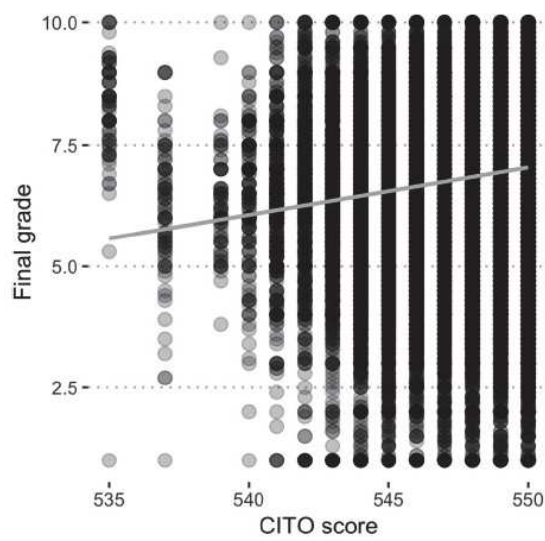


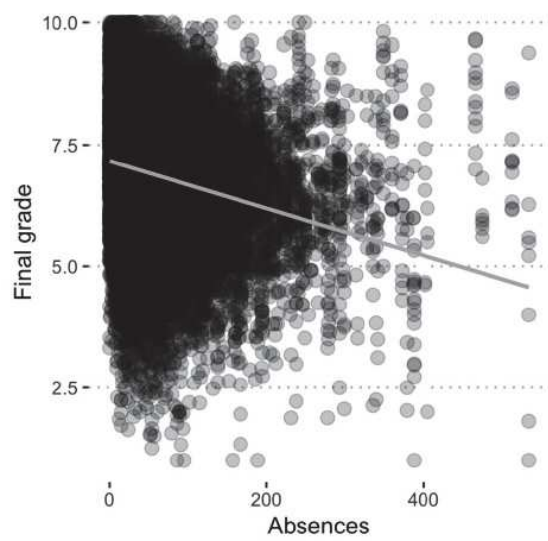
Figura 1.2: Boxplot voti finali per le più importanti variabili predittive categoriche.

diverse dai bambini olandesi. Infine, osservando i boxplot riportati all'estrema destra, un'importante considerazione può essere fatta a proposito dell'andamento scolastico e dei relativi voti; si nota, infatti, come le valutazioni siano più basse verso la fine del liceo.

Analizzando i grafici in Figura 1.3 si può osservare come il punteggio CITO abbia un'evidente correlazione positiva con i voti scolastici, si parla infatti di *tendenza lineare positiva*; mentre al contrario, si nota una *tendenza lineare negativa* a proposito delle assenze, in quanto maggiore sarà il numero di ore non frequentate dallo studente, più bassi saranno i voti scolastici finali.



(a) Final grades versus CITO score



(b) Final grades versus total absences

Figura 1.3: Grafici a dispersione dei voti finali e delle variabili numeriche più importanti con una linea di tendenza lineare.

Capitolo 2

Modello gerarchico

In questa sezione, viene introdotto un modello gerarchico a tre livelli per i voti degli studenti e vengono confrontate le sue prestazioni con modelli più semplici. L'obiettivo dello studio è costruire un modello che permette di monitorare i progressi degli studenti durante l'anno scolastico, dove l'interesse principale consiste nel segnalare gli studenti "eccezionali", ovvero coloro che risultano avere prestazioni sia inferiori che superiori alla maggior parte dei loro compagni.

2.1 Il modello

Durante tutto l'anno scolastico, gli studenti effettuano dei test per ciascun corso $i=1, \dots, n_0$.

I voti sono definiti come $g_{ki} \in [1, 10]$ con $k=1, \dots, K_i$, dove K_i è il numero di test effettuati nel corso i .

Poiché questi voti sono ottenuti per i test individuali, si hanno una serie di voti medi ponderati cumulativi $y_{i,j[i],h[j[i]]}$ per il corso i , lo studente j e la classe h .

Per leggibilità vengono eliminati gli indici j e h .

I risultati dei test individuali g_{ki} e i pesi dei test w_{ki} determinano il voto medio

$$y_i = \frac{\sum_{k=1}^{K_i} w_{ki} g_{ki}}{\sum_{k=1}^{K_i} w_{ki}} \quad (2.1)$$

con $y_i \in [1, 10]$.

I pesi sono assegnati dalle scuole sulla base delle specificità dei percorsi didattici, e, grossomodo, dipendono dalle ore settimanali assegnate alle singole materie e dalla collocazione nell'anno scolastico (i test verso fine dell'anno pesano di più).

Si considera un *modello gerarchico a tre livelli* e si utilizzano:

- l'indice i ($i=1,2,\dots,n_0$) per denotare il livello del corso individuale;
- l'indice j ($j=1,2,\dots,n_1$) per denotare il livello del singolo studente;
- l'indice h ($h=1,2,\dots,n_2$) per denotare il livello della classe.

Si hanno p_0 predittori per il livello del corso individuale, p_1 predittori per il livello dello studente, p_2 predittori per il livello della classe.

Si definiscono, inoltre, i vettori di riga $X_i^{(L_0)}$, $X_j^{(L_1)}$ e $X_h^{(L_2)}$ che consistono rispettivamente nei valori di *intercetta* e *predittore* per il corso, lo studente e il livello di classe.

Si modella il voto medio ponderato cumulativo y_i per il corso i come

$$y_i \sim N(X_i^{(L_0)} \beta_{j[i]}^{(L_0)}, \sigma^2) \quad i = 1, \dots, n_0 \quad (2.2)$$

dove il livello dello studente è modellato come

$$\beta_j^{(L_0)} \sim N(\beta_{h[j]}^{(L_1)} X_j^{(L_1)'}, \Sigma^{(L_1)}) \quad j = 1, \dots, n_1 \quad (2.3)$$

e il livello della classe è specificato come

$$\text{vec}(\beta_h^{(L_1)}) \sim N(\beta^{(L_2)} X_h^{(L_2)'}, \Sigma^{(L_2)}) \quad h = 1, \dots, n_2 \quad (2.4)$$

dove:

- $X_i^{(L_0)}$ è il vettore riga $1 \times (p_0+1)$ di variabili specifiche della materia (come il contenuto e il livello del corso);
- $\beta_{j[i]}^{(L_0)}$ è il vettore $(p_0+1) \times 1$ di parametri per lo studente j che segue il corso i ;
- σ^2 è la varianza per il livello del corso;
- $\beta_{h[j]}^{(L_1)}$ è la matrice di parametri $(p_0+1) \times (p_1+1)$ determinata dalla classe h in cui si trova lo studente j ;
- $X_j^{(L_1)}$ è il vettore riga $1 \times (p_1+1)$ di variabili specifiche dello studente (come l'età, le assenze e il QI);
- $\Sigma^{(L_1)}$ è la matrice di covarianza per i parametri $\beta_j^{(L_0)}$;
- $\text{vec}(\beta_h^{(L_1)})$ è la versione vettorializzata di $\beta_h^{(L_1)}$ con dimensioni $(p_0+1) \times (p_1+1) \times 1$;
- $\beta^{(L_2)}$ è la matrice di parametri $(p_0+1) \times (p_1+1) \times (p_2+1)$ a livello di classe;
- $X_h^{(L_2)}$ è il vettore riga $1 \times (p_2+1)$ di variabili specifiche della classe (come la dimensione di essa);
- $\Sigma^{(L_2)}$ è la matrice di covarianza per i parametri $\beta_h^{(L_1)}$.

2.2 Stima dei parametri

I parametri di un modello multilivello possono essere stimati utilizzando la massima verosimiglianza, i minimi quadrati generalizzati e la teoria bayesiana. Secondo quanto affermato da Browne e Draper (2006) e condiviso dagli autori Huberts, Schoonhoven e Does, quest'ultima fornisce un miglioramento in termini sia di stime puntuali e di intervallo, sia di distribuzioni a posteriori per i parametri. Verrà, quindi, utilizzata per la stima dei parametri del modello.

L'insieme completo dei parametri $\{\beta^{(L_0)}, \sigma^2, \beta^{(L_1)}, \Sigma^{(L_1)}, \beta^{(L_2)}, \Sigma^{(L_2)}\}$, dove $\beta^{(L_0)}$ e $\beta^{(L_1)}$ sono costruiti impilando le matrici dei parametri $\beta_j^{(L_0)}$ e $\beta_h^{(L_1)}$ rispettivamente per tutti i gruppi j e h , può essere stimato sulla base di dati considerati

rappresentativi, cioè in controllo.

Viene utilizzato un metodo bayesiano applicando i metodi *Markov Chain Monte Carlo* (MCMC), i quali fanno uso della procedura di *Gibbs*.

Poiché il numero di parametri aumenta rapidamente con l'aggiunta di livelli di gruppo, anche il tempo di stima aumenta notevolmente. Pertanto, quando si definisce un modello multilivello, esiste un compromesso tra la precisione aggiunta e il tempo di stima aggiuntivo per un livello di gruppo.

- In un modello a 2 livelli, il numero di parametri da stimare è:
1 per σ^2 , $(p_0+1)(p_1+1)$ per $\beta^{(L_1)}$ e $\frac{1}{2}(p_0+1)(p_0+2)$ per $\Sigma^{(L_1)}$;
- In un modello a 3 livelli, il numero di parametri da stimare aumenta:
1 per σ^2 , $\frac{1}{2}(p_0+1)(p_0+2)$ per $\Sigma^{(L_1)}$, $(p_0+1)(p_1+1)(p_2+1)$ per $\beta^{(L_2)}$ e
 $\frac{1}{2}(p_0+1)(p_1+1)((p_0+1)(p_1+1)+1)$ per $\Sigma^{(L_2)}$.

Ad esempio, se sono presenti 3 parametri per livello, il numero di parametri è 27 per un modello a 2 livelli e 211 per un modello a 3 livelli.

Si ottengono, quindi, le stime che si indicano con $\{\hat{\beta}^{(L_0)}, \hat{\sigma}^2, \hat{\beta}^{(L_1)}, \hat{\Sigma}^{(L_1)}, \hat{\beta}^{(L_2)}, \hat{\Sigma}^{(L_2)}\}$.

2.3 Risultati e misure di accuratezza

In questa sezione si considera l'accuratezza delle stime dei voti medi di fine anno per $N=3$, 839 corsi e 268 studenti durante l'anno scolastico 2014/2015. Questo sottoinsieme è costituito dagli studenti del primo, secondo e terzo anno. Al quarto anno gli studenti scelgono un profilo, che cambia la composizione delle classi. Per la stima dei parametri vengono utilizzati i cinque anni scolastici dal 2009 al 2014.

Come benchmark, si considerano l'utilizzo del voto medio ponderato (y_i) e un semplice modello di regressione lineare a un livello (\hat{y}_{sr}) per prevedere.

La regressione lineare a un livello si adatta a $y_i = X_i\beta + \varepsilon_i$ utilizzando gli stessi predittori della modellazione multilivello.

Come misure di accuratezza sono riportati il *Root Mean Squared Errors* (RMSE) ed il *Nearest Neighbors proportions* (NN).

L'errore quadratico medio (RMSE) è calcolato come

$$RMSE = \sqrt{\frac{1}{N} \sum_{i \in N} (y_i - \hat{y}_i)^2} \quad (2.5)$$

con i che identifica tutti i voti previsti e N il numero totale di voti. Questo punteggio punisce fortemente i grandi errori.

La seconda misura (NN) dell'accuratezza, invece, è calcolata come

$$NN = \frac{1}{N} \sum_{i \in N} I(\hat{y}_i - 1 \leq y_i \leq \hat{y}_i + 1), \quad (2.6)$$

che fornisce la proporzione delle previsioni "quasi giuste", ovvero le proporzioni dei valori osservati che danno un errore minore di 1.

Un'alternativa all'RMSE potrebbe essere la Deviazione Media Assoluta (MAD), la quale produce risultati comparabili ad esso.

In Tabella 2.3 vengono riportati RMSE e NN per le previsioni del modello gerarchico (\hat{y}_H), adattando un modello di regressione lineare a un livello (\hat{y}_{sr}) e per la media ponderata (y_i) a 5 punti nel tempo $t=0, 0.1, 0.3, 0.5, 0.7$.

Le due misure di performance indicate mostrano la superiorità del metodo gerarchico \hat{y}_H nella previsione dei voti di fine anno all'inizio dell'anno ($t=0$).

Con l'avanzare del tempo, il vantaggio del modello diminuisce man mano che si accumulano più voti ed il voto finale risulta meno incerto. Si nota che al tempo $t=0$ non è presente un valore di RMSE in quanto \hat{y}_i risulta essere una media mobile, dunque non possiede valori prima del tempo 0.

Per concludere, quindi, il modello gerarchico sembra essere quello che descrive in modo più accurato le previsioni dei voti di fine anno 2014/2015 dei 268 studenti.

Time	RMSE			NN		
	\hat{y}_i	\hat{y}_{sr}	\hat{y}_H	\hat{y}_i	\hat{y}_{sr}	\hat{y}_H
0	-	1.152	0.860	-	0.802	0.902
0.1	1.526	1.069	0.835	0.699	0.830	0.908
0.3	1.037	0.831	0.741	0.856	0.917	0.940
0.5	0.773	0.668	0.648	0.931	0.956	0.957
0.7	0.511	0.478	0.474	0.980	0.983	0.984

Tabella 2.3: Risultati RMSE e NN per le previsioni dei voti di fine anno 2014/2015 di 268 studenti utilizzando il voto medio (y_i), la regressione semplice (\hat{y}_{sr}) e la specificazione gerarchica (\hat{y}_H).

	Actual grades								
	3	4	5	6	7	8	9	10	
Predicted	6	0	1	0	0	0	0	2	0
	7	9	53	208	722	962	747	283	33
	8	0	6	20	134	255	252	140	12

Tabella 2.4: Matrice di confusione delle previsioni per le classi di fine anno 2014/2015 di 268 studenti basata sul modello di regressione lineare semplice a $t=0$.

Il confronto delle Tabelle 2.4 e 2.5 riportate chiarisce il vantaggio del modello di regressione gerarchica a tre livelli rispetto al modello a un livello. Entrambe mostrano i voti di fine anno previsti e realizzati prima dell'inizio dell'anno.

La differenza in RMSE di 0.292 potrebbe non sembrare degna del disturbo, ma mettendo a confronto i valori delle due tabelle, la 2.5 mostra una granularità molto maggiore. Ciò significa che il modello gerarchico identifica molta più struttura nei dati; risulta essere utile, quindi, per prevedere voti molto al di sopra e al di sotto della media. In Tabella 2.4 si può notare, inoltre, come il modello lineare non preveda voti negativi causando una distorsione verso voti più alti; la proporzione NN conta quanti valori osservati si trovano sulla diagonale, oppure una riga sotto o una sopra ad essa.

		RMSE = 0.860							
		Actual grades							
		3	4	5	6	7	8	9	10
Predicted	3	0	1	1	0	0	0	0	0
	4	0	1	3	2	0	0	0	0
	5	3	10	19	27	11	2	0	0
	6	4	36	114	358	182	55	10	0
	7	2	10	83	425	749	434	79	3
	8	0	2	8	43	267	464	213	14
	9	0	0	0	1	8	44	118	22
	10	0	0	0	0	0	0	5	6

Tabella 2.5: Matrice di confusione delle previsioni per le votazioni di fine anno 2014/2015 di 268 studenti basata sul modello a tre livelli a $t=0$.

Osservando i valori riportati in Tabella 2.5, ad esempio, si nota che il valore più alto (749) si trova in corrispondenza dei voti previsti correttamente pari a 7. In generale, si osserva che si ha una concentrazione dei valori più alti di previsioni sulla diagonale della matrice di previsione; ciò porta a preferire il modello gerarchico a tre livelli al modello di regressione lineare semplice.

Capitolo 3

Monitoraggio delle prestazioni degli studenti

Questa sezione riguarda il monitoraggio delle prestazioni degli studenti utilizzando i voti accumulati nei test. Vengono prese in considerazione tecniche di monitoraggio statistico dei processi e monitoraggio predittivo.

3.1 Monitoraggio statistico dei processi

Per utilizzare una tecnica classica della carta di controllo (ad esempio carte Shewhart, CUSUM o EWMA) è necessario disporre di un set di dati di Fase I che funge da set di "addestramento" ed un set di dati di Fase II che sarà un set di test.

La Fase I viene utilizzata per analizzare il modello scelto e per stimare i parametri coinvolti; si presume, inoltre, che i dati siano in controllo e che il monitoraggio inizi nella Fase II.

In questo caso, non ci sono dati disponibili per svolgere la Fase I, dunque si sceglie di utilizzare i dati degli studenti degli anni precedenti. Tuttavia, quest'ultimi non sono disponibili per le matricole, per i nuovi corsi e in caso di dati limitati.

Risulta difficile quindi identificare una chiara configurazione delle due fasi. Questi problemi sono anche amplificati dal fatto che y_i non è i.i.d. (indipendente e identicamente distribuita), violando i presupposti dell'uso di base delle carte di controllo.

Si modella y_i all'ora t utilizzando tutti i voti dei test prima dell'ora t , con $t \in \{t_i, T\}$, dove t_i indica l'inizio dell'anno scolastico e T la fine.

Si calcola, quindi, un valore atteso \hat{y}_i : la differenza tra il valore atteso e il valore effettivo osservato y_i al momento t può essere monitorata in un set di dati di Fase II utilizzando un'impostazione dello schema di sorveglianza residuo.

3.1.1 Diagramma di controllo a tre livelli

Si valuta se le relazioni date dal modello a tre livelli sono ancora valide. A tal fine, si monitorano i residui a tre livelli.

Per ogni gruppo, si hanno le stime dell'insieme completo dei parametri $\{\hat{\beta}^{(L_0)}, \hat{\sigma}^2, \hat{\beta}^{(L_1)}, \hat{\Sigma}^{(L_1)}, \hat{\beta}^{(L_2)}, \hat{\Sigma}^{(L_2)}\}$.

Si calcolano i *residui* per i tre livelli di $\{y_i, X_i^{(L_0)}, X_j^{(L_1)}, X_h^{(L_2)}\}$:

$$r_i^{(L_0)} = y_i - X_i^{(L_0)} \hat{\beta}_{j[i]}^{(L_0)} \quad (3.1)$$

$$r_j^{(L_1)} = \hat{\beta}_j^{(L_0)} - \hat{\beta}_{h[j]}^{(L_1)} X_j^{(L_1)} \quad (3.2)$$

$$r_h^{(L_2)} = \text{vec}(\hat{\beta}_h^{(L_1)}) - \hat{\beta}^{(L_2)} X_h^{(L_2)} \quad (3.3)$$

dove $r_i^{(L_0)}$, $r_j^{(L_1)}$, $r_h^{(L_2)}$ sono i vettori residui ai tre livelli di dimensione 1, (p_0+1) e $(p_0+1)(p_1+1)$ rispettivamente.

Si vuole determinare se una nuova osservazione deriva dalla distribuzione in controllo di Fase I, ottenuta utilizzando i dati di stima $\{X_I^{(L_0)}, X_I^{(L_1)}, X_I^{(L_2)}, y_I\}$ di dimensione n_0 , dove $X_I^{(L_0)}$ è la matrice $n_0 \times (p_0+1)$ con l' i -esima riga contenente i valori di intercetta e predittore per il corso i .

Si utilizza un *diagramma di controllo di Shewhart* prendendo le stime della media e della varianza della Fase I per $r_i^{(L_0)}$, con $\widehat{UCL}_y = 3\hat{\sigma}^2$ e $\widehat{LCL}_y = -3\hat{\sigma}^2$.

Il grafico segnala un allarme quando il residuo supera uno dei limiti di controllo.

Per $r_j^{(L_1)}$ e $r_h^{(L_2)}$ sono necessarie carte di controllo multivariate perchè sono *residui multidimensionali*. La carta di controllo *Hotelling* T^2 multivariata offre una soluzione con le statistiche dei test:

$$T^2_{(L_1)} = n_0 r_j^{(L_1)' \hat{\Sigma}^{-1(L_1)} r_j^{(L_1)} \quad (3.4)$$

$$T^2_{(L_2)} = n_0 r_h^{(L_2)' \hat{\Sigma}^{-1(L_2)} r_h^{(L_2)} \quad (3.5)$$

dove n_0 è il numero di osservazioni usate per stimare la matrice di covarianza.

Si nota un errore nell'articolo originale citato inizialmente, nel quale non compare l'inversa sulla matrice di covarianza.

Il limite di controllo inferiore è $LCL=0$.

I limiti di controllo superiore che garantiscono alla carta di controllo una percentuale di falso allarme α sono:

$$LCL_{(L_1)} = \frac{p_1(n_0 - 1)}{n_0 - p_1} F_{\alpha, p_1, n_0 - p_1} \quad (3.6)$$

$$UCL_{(L_2)} = \frac{p_2(n_0 - 1)}{n_0 - p_2} F_{\alpha, p_2, n_0 - p_2} \quad (3.7)$$

- Se il grafico di $T^2_{(L_2)}$ fornisce un segnale, l'analisi della causa principale può concentrarsi sul livello di classe;
- Se il grafico di $T^2_{(L_1)}$ fornisce un segnale, l'analisi della causa principale può concentrarsi sul livello dello studente;
- Se la carta di Shewhart fornisce un segnale, l'analisi può concentrarsi sul livello del corso.

Allo stesso modo, si potrebbero monitorare anche le stime dei parametri.

3.1.2 Esempio

Per illustrare l'approccio di monitoraggio a tre livelli, si monitora la media ponderata cumulativa y_i 15 volte durante l'anno scolastico 2014/2015.

La Fase I comprende i cinque anni scolastici dal 2009 al 2014; la Fase II è l'anno scolastico 2014/2015 per i 3839 corsi seguiti da 268 studenti delle classi prime, seconde e terze.

Lo scopo della scuola è quello di rilevare corsi e studenti "eccezionali".

Considera i corsi eccezionali come voti finali inferiori a 6 o superiori a 8. Ogni punto inferiore a 6 viene conteggiato come "failpoint". Un singolo corso con un voto di fine anno 5 equivale a 1 failpoint; un singolo corso con un voto di fine anno 3 equivale a 3 failpoint e un corso con voto 4 e uno di 3 equivale a 5 failpoint.

D'altra parte, ogni punto superiore a 8 viene conteggiato come un "excelpoint". Pertanto il voto massimo di 10 per un corso equivale a 2 excelpoint. Uno studente eccezionale è uno studente con almeno quattro failpoint e/o almeno quattro excelpoint.

Si applica il modello di regressione gerarchica e si monitorano i residui usando un *diagramma di controllo di Shewhart*.

Le stime del modello a tre livelli hanno un RMSE complessivo di 1.172. Come mostrato in Figura 3.1, il grafico di Shewhart che monitora i residui del primo livello $r_i^{(L_0)}$ segnala 4 volte verso la fine dell'anno. In totale, i grafici dei residui segnalano 190 volte (di cui 88 sono corsi eccezionali) per 112 studenti (di cui 36 sono studenti eccezionali). Si nota che i limiti di controllo non risultano essere costanti ma variano nel tempo al modificarsi della varianza.

Come specificato dall'Equazione [3.4], si possono anche monitorare i residui a livello di studenti utilizzando un grafico Hotelling T^2 . Utilizzando gli stessi dati del precedente, il grafico T^2 segnala almeno una volta per 105 studenti (di cui 38 studenti eccezionali). I grafici segnalano casi eccezionali durante tutto l'anno.

Tuttavia non si può determinare retrospettivamente se al momento di un segnale ci fosse qualche incognita che ha influenzato le prestazioni dello studente j per il corso i . Non si è, quindi, in grado di distinguere i segnali falsi da quelli veri. L'approccio di monitoraggio statistico identifica le anomalie nelle medie pondera-

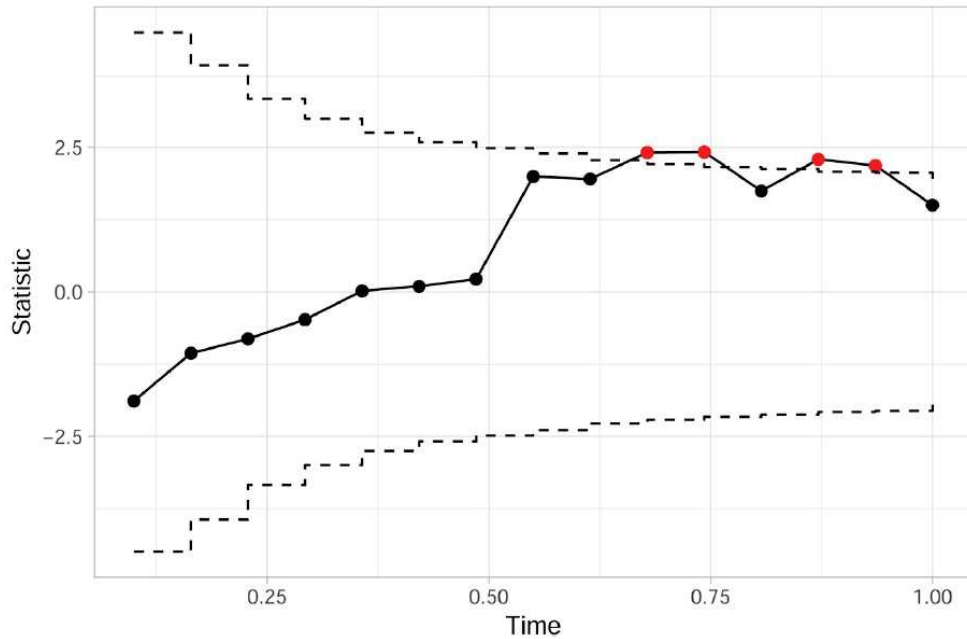


Figura 3.1: Grafico di controllo residuo di Shewhart che monitora $r_i^{(L_0)}$ sulla base di una regressione a tre livelli (segnali in rosso).

te. L'obiettivo principale della scuola è identificare gli studenti che necessitano di supporto o corsi più impegnativi.

Questo approccio di monitoraggio, però, è insufficiente per tale obiettivo. Pertanto, nella sezione successiva, verrà utilizzato il modello gerarchico per monitorare i risultati di fine anno attesi dagli studenti per identificare coloro che ottengono prestazioni inferiori o superiori.

3.2 Monitoraggio predittivo

In questo caso di studio, la scuola superiore mira a prevedere i voti di fine anno dei suoi studenti. Ciò consente alla scuola di ricevere avvisi tempestivi su studenti eccezionali. Nella sezione seguente viene considerato, quindi, il monitoraggio predittivo delle prestazioni degli studenti.

3.2.1 Monitoraggio predittivo multilivello

Si vuole monitorare $P(E)_t$, definita come la probabilità di un evento E al tempo t . Riassume il risultato del modello in un'unica probabilità predittiva al tempo t con $t \in \{t_i, T\}$.

Il grafico segnala quando $P(E)_t$ supera la soglia C , ossia la massima probabilità consentita che si verifichi l'evento ($0 < C < 1$).

L'evento E riguarda i valori di y_i , che dipendono dal contesto e possono assumere molte forme: ($y_i = e$, $y_i \geq e$, $y_i \leq e$, $e_1 \leq y_i \leq e_2$, $\sum_{i=a}^b y_i \geq e$, etc. dove e , e_1 , e_2 sono costanti arbitrarie mentre a e b sono interi compresi tra 1 e n_0).

Seguendo la stima MCMC delle densità a posteriori dei parametri $\theta = \{\beta^{(L_0)}, \sigma^2, \beta^{(L_1)}, \Sigma^{(L_1)}, \beta^{(L_2)}, \Sigma^{(L_2)}\}$, possono essere utilizzate per calcolare $P(E)_t$.

I passaggi per il monitoraggio predittivo sono:

1. Definire l'evento E e la soglia C ;
2. Specificare il modello multilivello per y_i ;
3. Stimare i parametri per ottenere $\hat{\theta}_I$ utilizzando i dati di Fase I al tempo t_i usando MCMC;
4. Calcolare $P(E)_t$ utilizzando le nuove osservazioni disponibili all'istante $t > t_I$;
5. Segnalare se $P(E)_t > C$;
6. Ristimare i parametri per ottenere $\hat{\theta}_t$ utilizzando tutti i dati disponibili all'istante t , di seguito tornare al punto 4 per un nuovo punto temporale $t_{II} > t$.

Si suppone di avere un grande insieme di dati di Fase I in controllo $\{X_I^{(L_0)}, X_I^{(L_1)}, X_I^{(L_2)}, y_I\}$ al tempo $t = t_I$.

Al tempo $t < t_I$ si ottengono le stime dei parametri $\{\hat{\beta}^{(L_0)}, \hat{\sigma}^2, \hat{\beta}^{(L_1)}, \hat{\Sigma}^{(L_1)}, \hat{\beta}^{(L_2)}, \hat{\Sigma}^{(L_2)}\}$ sulla base delle osservazioni nella Fase I.

Utilizzando le stime, in ogni momento $t > t_I$ si ha una distribuzione prevista per la variabile di risultato $\hat{y}_{i,t}$

$$\hat{y}_{i,t} \sim N(\Psi_{i,t}, \Omega_{i,t}) \quad (3.8)$$

dove

$$\begin{aligned} \Psi_{i,t} &= (X_{i,t}^{(L_0)} \otimes X_{j[i,t]}^{(L_1)'}) \hat{\beta}^{(L_2)} X_{h[j[i,t]]}^{(L_2)'} \\ \Omega_{i,t} &= (X_{i,t}^{(L_0)} \otimes X_{j[i,t]}^{(L_1)'}) \hat{\Sigma}^{(L_2)} (X_{i,t}^{(L_0)} \otimes X_{j[i,t]}^{(L_1)'}) + X_{i,t}^{(L_0)} \hat{\Sigma}^{(L_1)} X_{i,t}^{(L_0)'} + \hat{\sigma}^2 \end{aligned}$$

con \otimes che indica il prodotto di Kronecker.

Si può usare questo risultato per stimare la probabilità $P(E)_t$.

Si suppone che $y_i \leq e$ (cioè di studiare che il grado y_i sia minore dell'evento e). Lo schema di monitoraggio proposto utilizza la distribuzione a posteriori di $\hat{y}_{i,t}$ per calcolare la probabilità $P(E)_t$. Il grafico segnala quando $P(E)_t > C$.

Il monitoraggio di $P(E)_t$ richiede una ristima periodica dei parametri per incorporare le nuove informazioni disponibili al momento t .

Intorno al verificarsi dell'evento E la probabilità $P(E)_t$ converge a 1 se $t \rightarrow T$.

Il principale vantaggio del monitorare $P(E)_t$ invece di $\hat{y}_{i,t}$ consiste nel fatto che, a seconda della capacità predittiva del modello multilivello, lo schema fornisce allerta precoce e l'opportunità di intervenire prima che si verifichi l'evento E .

Se si verifica un intervento, è importante includerlo nei preditorri $\{X^{(L_0)}, X^{(L_1)}, X^{(L_2)}\}$ includendo una variabile aggiuntiva per estrarre l'effetto dell'intervento sull'esito E .

Inoltre, non sono necessarie n_0 carte di controllo. Serve solo una singola carta di controllo che traccia i valori di $P(E)_t$ e segnala le osservazioni o i gruppi per i quali $P(E)_t$ supera C .

3.2.2 Esempio

Seguendo i passaggi delineati in precedenza, si definiscono due eventi: E^f definito come lo studente che fallisce l'anno, E^e definito come lo studente che eccelle l'anno

scolastico. E^f si verifica se uno studente ha quattro o più failpoint, mentre E^e se uno studente ha quattro o più excelpoint.

Il voto arrotondato di fine anno dello studente j per il corso i è definito come y_{ij} . Al tempo t la probabilità che uno studente bocci l'anno è $P(E_j^f)_t = P(\sum_{i=1}^{n_j} \max(0, (6 - y_{ij})) \geq 4)_t$, dove n_j è il numero di corsi per lo studente j . Mentre, la probabilità che uno studente eccelga l'anno è $P(E_j^e)_t = P(\sum_{i=1}^{n_j} \max(0, (y_{ij} - 8)) \geq 4)_t$.

La Figura 3.2 sotto riportata mostra una carta di controllo di $1 - P(E_j^f)_t$ per $J=268$ studenti in 15 punti nel tempo. La soglia $C=0.05$ è indicata con una linea nera tratteggiata. Si nota che $1 - P(E_j^f)_t$ rappresenta la probabilità di passare l'anno scolastico. I $J_p=238$ studenti che l'hanno superato sono segnati in blu, mentre le probabilità dei $J_f=30$ studenti che hanno fallito in rosso.

Nel complesso il modello stima in modo coerente le probabilità; ciò si può osservare anche dai dati riportati in Tabella 3.2.

La tabella riporta i valori di $\frac{1}{J_p} \sum_{j \in J_p} P(E_j^f)_t$ nella prima riga (ovvero la probabilità media stimata di insuccesso per gli studenti che superano l'anno) e $\frac{1}{J_f} \sum_{j \in J_f} P(E_j^f)_t$ nella seconda riga (ossia la probabilità media stimata di insuccesso per gli studenti che non superano l'anno).

Proseguendo con l'analisi, la Figura 3.4 traccia $P(E_j^e)_t$ per gli stessi $J=268$ studenti. Gli studenti $J_n=222$ che non hanno eccelso sono rappresentati in rosso e le probabilità degli studenti $J_e=46$ che hanno eccelso sono rappresentate in blu. La soglia $C=0.95$ è rappresentata da una linea tratteggiata.

Il modello risulta avere delle prestazioni importanti, mostrate dalle differenze nelle probabilità medie nel tempo tra gli studenti che eccellono $\frac{1}{J_e} \sum_{j \in J_e} P(E_j^e)_t$ e quelli che no $\frac{1}{J_n} \sum_{j \in J_n} P(E_j^e)_t$, come si può notare dalla Tabella 3.3.

A seconda della soglia C che determina se lo schema di monitoraggio segnala, il modello identifica correttamente diversi studenti che falliranno/eccelleranno così come alcuni *falsi positivi*. Le Tabelle 3.4 e 3.5 riportano i valori di precisione e richiamo che monitorano rispettivamente E^f e E^e .

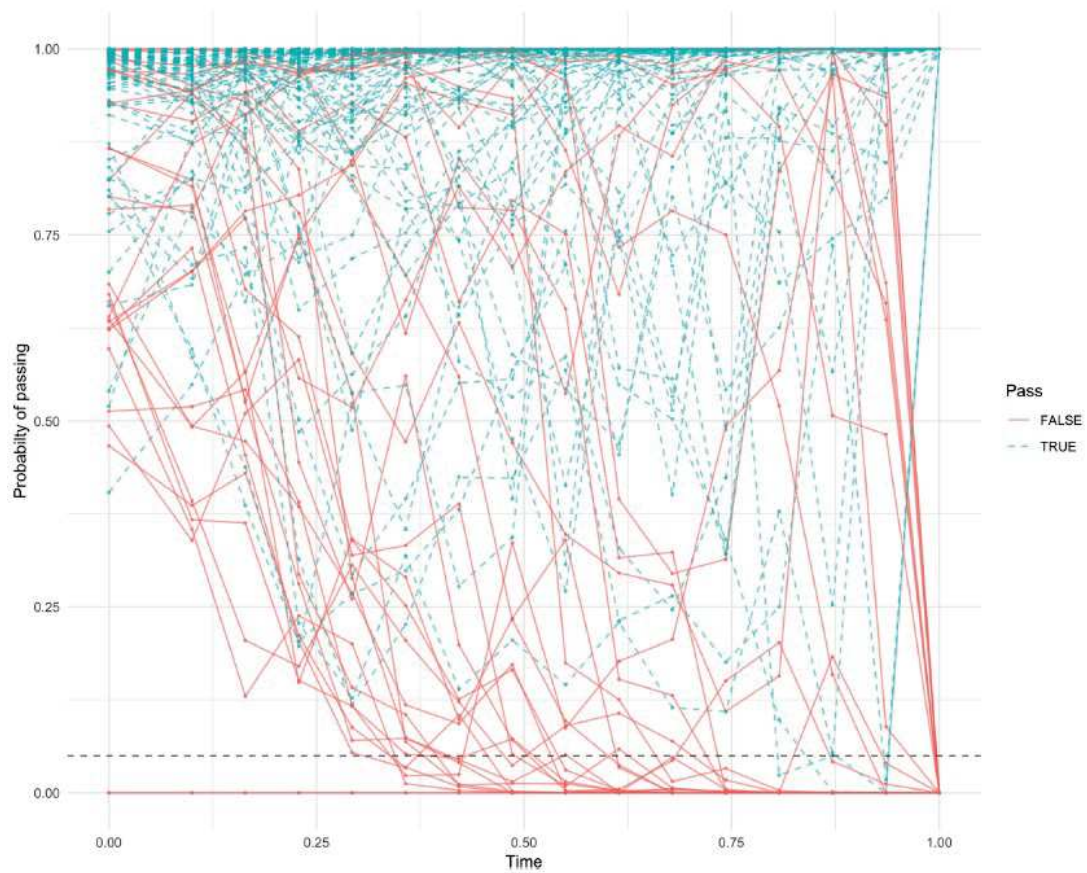


Figura 3.2: Una carta di controllo che monitora le probabilità stimate di superare $1 - P(E^f)_t$ per 268 studenti nel 2014/2015, con soglia tratteggiata $C=0.05$ in nero. Le linee blu tratteggiate rappresentano gli studenti che hanno superato, le linee continue rosse gli studenti che hanno fallito.

La *precisione* è definita come:

$$\text{Precision}_t(C) = \frac{tp_t(C)}{tp_t(C) + fp_t(C)}$$

con $tp_t(C)$ uguale al numero di veri positivi al tempo t per la soglia C e $fp_t(C)$ uguale al numero di falsi positivi al tempo t per la soglia C .

Il *richiamo* è definito come:

$$\text{Recall}_t(C) = \frac{tp_t(C)}{tp_t(C) + fn_t(C)}$$

Failed	Time						
	0	0.1	0.3	0.5	0.7	0.9	1
No	0.02	0.02	0.04	0.03	0.03	0.01	0.00
Yes	0.27	0.28	0.52	0.61	0.75	0.79	1.00

Tabella 3.2: Probabilità medie stimate di insuccesso $P(E^f)_t$ per 268 studenti nel 2014/2015, suddivise per risultato osservato.

con $fn_t(C)$ uguale al numero di falsi negativi al tempo t per la soglia C .

Osservando la Tabella 3.4 si può notare come la procedura identifichi correttamente gli studenti che falliranno all'inizio dell'anno. A seconda del livello scelto di C vengono generati più avvisi precoci che aiutano nel sistema di supporto degli studenti. Per esempio, ponendo C a 0.75, la procedura individua quasi la metà (14 su 30) degli studenti che falliranno prima dell'inizio dell'anno con solo il 26% (5) di falsi positivi.

La Tabella 3.5, invece, mostra i valori di precisione e richiamo quando si prevedono studenti eccellenti. A seconda delle preferenze della scuola è possibile ottenere un'elevata precisione o richiamo all'inizio dell'anno. Ad esempio, ponendo C a 0.50, la procedura individua la metà (23 su 46) degli studenti che eccelleranno prima dell'inizio dell'anno con solo il 15% (4) di falsi positivi.

La procedura di monitoraggio multilivello ha mostrato il suo valore in un contesto di scuola superiore, in quanto fornisce adeguatamente i voti previsti di fine anno per tutti gli studenti e le materie. Questo può aiutare a classificare gli studenti a rischio che necessitano di supporto, nonché le aree in cui hanno bisogno di aiuto. Dall'altro lato, il modello identifica con successo gli studenti eccellenti che possono beneficiare di compiti scolastici più impegnativi. Inoltre, il modello fornisce risultati facilmente interpretabili, nonché una buona spiegabilità per i parametri.

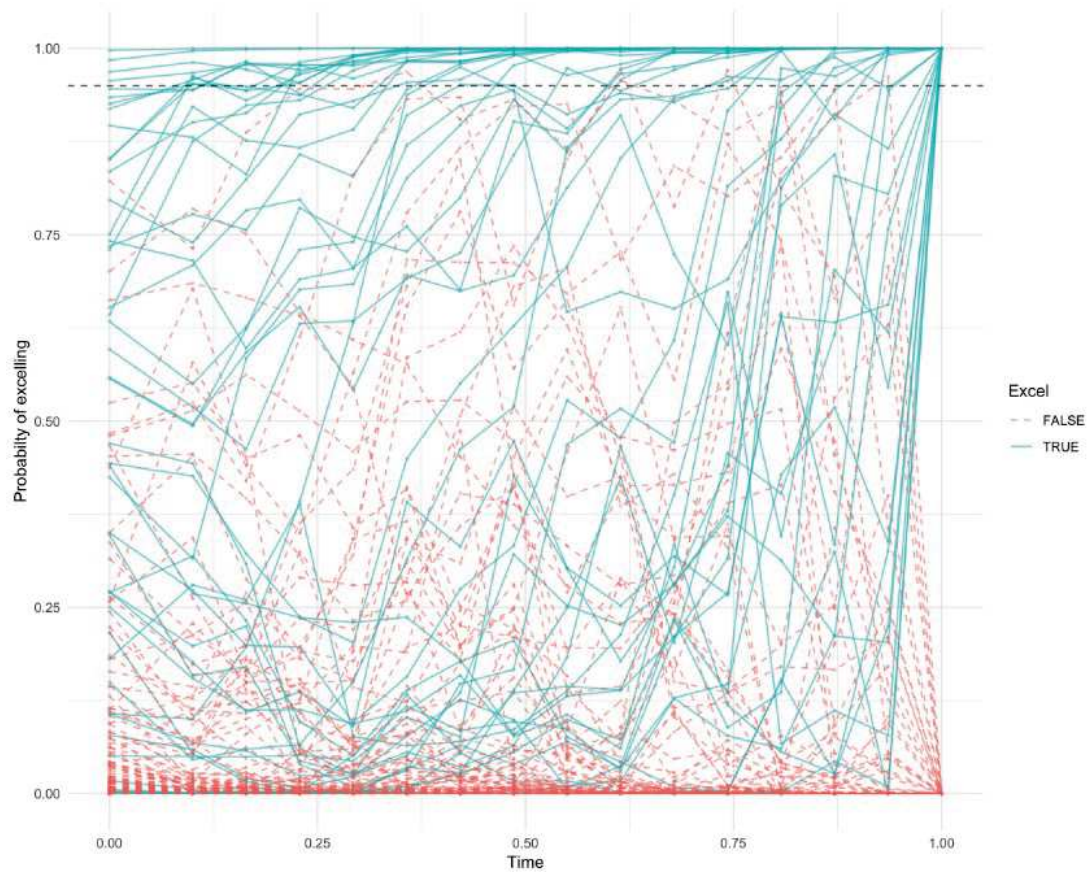


Figura 3.4: Una carta di controllo che monitora le probabilità stimate di eccellere $P(E^e)_t$ per 268 studenti nel 2014/2015, con soglia tratteggiata $C=0.95$ in nero. Le linee blu continue rappresentano gli studenti che hanno eccelso, le linee tratteggiate rosse gli studenti che non hanno eccelso.

Excelled	Time						
	0	0.1	0.3	0.5	0.7	0.9	1
No	0.05	0.04	0.04	0.06	0.04	0.03	0.00
Yes	0.50	0.49	0.50	0.61	0.67	0.81	1.00

Tabella 3.3: Probabilità medie stimate di eccellere $P(E^e)_t$ per 268 studenti nel 2014/2015, suddivise per risultato osservato.

C							
		0.05	0.1	0.25	0.5	0.75	0.999
Time	0	1 (0.07)	1 (0.07)	1 (0.07)	0.67 (0.13)	0.74 (0.47)	0.25 (0.93)
	0.1	1 (0.07)	1 (0.07)	1 (0.07)	1 (0.27)	0.71 (0.40)	0.25 (0.93)
	0.3	1 (0.10)	1 (0.20)	0.85 (0.37)	0.76 (0.53)	0.67 (0.67)	0.27 (1)
	0.5	1 (0.33)	1 (0.43)	0.94 (0.53)	0.79 (0.63)	0.67 (0.67)	0.34 (0.97)
	0.7	1 (0.57)	1 (0.63)	0.88 (0.73)	0.77 (0.70)	0.70 (0.77)	0.40 (0.97)
	0.9	0.90 (0.63)	0.86 (0.63)	0.88 (0.70)	0.81 (0.70)	0.81 (0.73)	0.59 (0.90)
	1	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)

Tabella 3.4: Risultati di $\text{Precision}_t(C)$ ($\text{Recall}_t(C)$) durante il monitoraggio di $P(E^f)_t$ con vari valori di C e t utilizzando le previsioni del modello a tre livelli dei voti di fine anno per 268 studenti nel 2014/2015.

C							
		0.99	0.95	0.75	0.5	0.25	0.01
Time	0	1 (0.02)	1 (0.09)	0.93 (0.3)	0.85 (0.5)	0.69 (0.72)	0.38 (0.89)
	0.1	1 (0.04)	1 (0.2)	0.94 (0.35)	0.72 (0.46)	0.71 (0.65)	0.45 (0.87)
	0.3	1 (0.09)	1 (0.28)	0.89 (0.37)	0.83 (0.54)	0.68 (0.54)	0.45 (0.85)
	0.5	1 (0.37)	1 (0.41)	0.92 (0.52)	0.77 (0.59)	0.65 (0.67)	0.47 (0.96)
	0.7	1 (0.43)	1 (0.48)	0.89 (0.54)	0.88 (0.61)	0.72 (0.78)	0.57 (0.93)
	0.9	1 (0.57)	1 (0.63)	0.94 (0.74)	0.88 (0.83)	0.8 (0.87)	0.64 (1)
	1	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)

Tabella 3.5: Risultati di $\text{Precision}_t(C)$ ($\text{Recall}_t(C)$) durante il monitoraggio di $P(E^e)_t$ con vari valori di C e t utilizzando le previsioni del modello a tre livelli dei voti di fine anno per 268 studenti nel 2014/2015.

Conclusioni

Questo caso di studio ha considerato tre domande di ricerca riguardanti le prestazioni degli studenti di una scuola superiore olandese. Di seguito, vengono discussi i risultati relativi alle tre domande che hanno portato all'uso di un modello a tre livelli in uno schema di monitoraggio predittivo per dati gerarchici.

- Cosa determina il rendimento degli studenti?

Il set di dati dettagliato utilizzato ha mostrato interessanti determinanti del rendimento degli studenti. Questi sono generalmente in linea con la letteratura educativa e sono utili per monitorare i progressi degli studenti.

Le studentesse hanno ottenuto voti finali più alti; successivamente, per quanto riguarda gli studenti con disabilità, quest'ultimi ottengono risultati di poco peggiori. I bambini con genitori altamente istruiti superano i loro coetanei con genitori meno istruiti. Per quanto concerne la lingua, una seria barriera linguistica sembra produrre voti leggermente inferiori. Non si evidenziano, invece, risultati differenti tra gli studenti nati all'estero ed i coetanei nati in loco.

Gli studenti mostrano una diminuzione delle prestazioni durante la loro carriera liceale, con circa mezzo punto di differenza nei voti tra il primo e il quarto anno di scuola superiore. Inoltre, le assenze sembrano avere una forte correlazione negativa con i voti. Per finire, la relazione tra i punteggi dei test della scuola primaria (CITO) ed i voti degli studenti dovrebbe essere considerata nell'attuale discussione sulle determinanti del livello di scuola superiore.

- Come può essere utilizzato il monitoraggio del processo statistico per monitorare i progressi degli studenti?

Le classiche tecniche di monitoraggio statistico dei processi sono spesso insufficienti se applicate a processi complessi. Quando è presente una struttura gerarchica, la modellazione multilivello migliora l'affidabilità; in particolare migliora l'accuratezza e la spiegabilità della stima rispetto ai normali modelli di regressione lineare. Il metodo è essenziale per la modellazione predittiva di nuovi studenti/classi (inclusi campioni di piccole dimensioni).

Le tecniche di monitoraggio del processo statistico univariato si sono rivelate insufficienti e i modelli di regressione lineare a un livello non hanno fornito risultati soddisfacenti. Dunque, si è deciso di utilizzare grafici di controllo residuo a tre livelli come metodo di monitoraggio statistico multilivello, il quale ha fornito risultati promettenti.

- Quale metodo può essere utilizzato per il monitoraggio predittivo dei risultati degli studenti?

Si è sviluppato un metodo di monitoraggio predittivo per consentire un sistema di monitoraggio di allerta precoce. Ciò ha permesso di monitorare la probabilità di un evento. Il modello a tre livelli è stato utilizzato per prevedere continuamente i voti individuali di fine anno.

Tramite un modello gerarchico bayesiano, si ottengono le distribuzioni di probabilità per i risultati degli studenti. Queste possono essere utilizzate per osservare i risultati indesiderati, ovvero prestazioni insufficienti e superiori. E' stata fornita un'analisi completa del modello a tre livelli e della sua stima sfruttando il campionamento di Gibbs per stimare i parametri.

Per concludere, i sistemi di indicatori di allerta precoce hanno il potenziale per migliorare il sistema educativo a basso costo. Possono aggiungere un livello di sofisticazione alla valutazione delle prestazioni scolastiche e degli insegnanti e lavorare per soddisfare le esigenze dei singoli studenti.

Bibliografia

Leo C. E. Huberts, Marit Schoonhoven Ronald J. M. M. Does. “Multilevel process monitoring: A case study to predict student success or failure”. In: *Journal of Quality Technology* 54.2 (2022), pp. 127–143. DOI: 10.1080/00224065.2020.1828008.

Montgomery, Douglas C. *Introduction to Statistical Quality Control*. Sixth Edition. John Wiley & Sons, 2020.