

Un modello ad effetti misti con dati longitudinali per il *Progress Test* dei corsi di laurea italiani in odontoiatria

Laura Antonucci^{1*}, Giulio Biscardi², Najada Firza³,
Leonardo Grilli²

¹Università di Foggia, ²Università di Firenze, ³Università di Bari.

Riassunto: Il lavoro svolto ha lo scopo di illustrare l'analisi statistica dei Corsi di Laurea in Odontoiatria mediante i risultati dei Progress Test. La trattazione ha inizio ponendo l'attenzione sul concetto di apprendimento dello studente, come misurarlo e come poterlo migliorare; uno dei principali problemi che affrontano le Università, in particolare la Facoltà di Medicina, le quali hanno come obiettivo finale formare i medici del futuro in modo che siano competenti e preparati. Per questo è stato istituito il Progress Test, uno strumento idoneo a valutare l'andamento delle conoscenze degli studenti nei vari anni di corso e comparare i risultati tra i vari atenei.

Keywords: Progress Test; indici sintetici; Modello ad effetti misti per dati longitudinali; Raw Residuals.

1. Introduzione

Da alcuni anni i corsi di laurea in odontoiatria e protesi dentaria italiani, così come quelli in Medicina e Chirurgia hanno adottato un sistema di monitoraggio delle competenze teoriche acquisite dai propri studenti utilizzando il metodo Progress Test, o PT (Albanese, 2016).

Il Progress Test utilizza 300 domande a scelta multipla, 150 delle quali riguardano le discipline di base, mentre le altre riguardano le scienze cliniche.

Esso serve per valutare l'acquisizione e la conservazione nel tempo delle conoscenze in relazione agli obiettivi del core curriculum e permette di misurare

* Autore corrispondente: laura.antonucci@unifg.it

Il lavoro qui descritto è frutto di un progetto comune, ma L. Antonucci ha provveduto alla redazione dei parr. 1-3, G. Biscardi ha redatto i parr. 8-9, N. Firza ha redatto i parr. 6 e 7, L. Grilli i parr. 4 e 5.

l'incremento o il decremento annuale delle competenze teoriche acquisite dagli studenti iscritti nei 6 anni di corso.

Le domande sono a risposta multipla precompilata, con una sola risposta esatta su 5 opzioni, ma le 4 risposte errate non comportano penalizzazioni.

Le domande relative alla Scienze di Base riguardano 15 materie (si veda tabella 1) per cui vi sono 10 domande per ogni disciplina. Lo stesso dicasi per le discipline cliniche. Il test è somministrato con cadenza annuale a tutti gli studenti, dal primo al sesto anno di corso.

Le risposte fornite dagli studenti del 1° anno rappresentano il paradigma di riferimento rispetto al quale misurare la crescita di competenze negli anni successivi. Si presuppone, infatti, che gli studenti del primo anno selezionino le risposte in base alle conoscenze che hanno acquisito durante le scuole superiori, oppure all'elemento "fortuna" (cioè al caso).

La somministrazione del questionario agli studenti degli anni successivi invece riflette le informazioni acquisite gli anni precedenti, ad esempio le risposte degli studenti del 2° anno rispecchiano le conoscenze acquisite il 1° anno e così via.

Le informazioni raccolte includono:

- il grado di partecipazione al test;
- i risultati globali sia per Scienze di Base che per Scienze Cliniche;
- l'andamento dei risultati delle discipline individuali dal 1°anno fino all'ultimo anno.

Al termine si ottengono informazioni non solo sull'acquisizione di nuove conoscenze ma anche sul permanere di quelle apprese in precedenza. Osservando l'andamento di ogni studente rispetto al proprio gruppo di appartenenza, è possibile delineare l'evoluzione temporale dell'apprendimento di ciò che lo studente ricorda a distanza di tempo, soprattutto, come riesce ad utilizzare le conoscenze apprese.

I risultati ottenuti non hanno valore certificativo ma formativo per lo studente che lo aiuta a comprendere i suoi punti di forza e debolezza nel percorso di apprendimento e ad aumentare la consapevolezza di quali saranno le sfide a cui saranno esposti una volta finito il percorso di Studi.

2. Descrizione dei Progress Test per l'analisi di performance del corso di studi in Odontoiatria e Protesi Dentaria delle Università Italiane

Le analisi descrivono i risultati ottenuti nei PT somministrati per i corsi di laurea in Odontoiatria e Protesi Dentaria per tutti gli Atenei Italiani per i due gruppi di insegnamenti: Scienze di Base e Scienze Cliniche.

Tabella 1. *Insegnamenti di Scienze di Base e di Scienze Cliniche considerati per il Progress Test dei corsi di Odontoiatria e Protesi Dentaria degli Atenei Italiani.*

SCIENZE DI BASE	SCIENZE CLINICHE
Scienze Comportamentali	Principi di Odontoiatria
Chimica e Biochimica	Materiali Dentari
Fisica	Tecnologie di Laboratorio
Biologia e Genetica	Patologia Orale
Istologia e Anatomia	Chirurgia Orale
Fisiologia	Parodontologia
Patologia Generale	Protesi
Farmacologia	Gnatologia
Medicina Interna	Ortodonzia
Anestesiologia e Chirurgia Generale	Conservativa
Anatomia Patologica	Endodonzia
Medicina Legale	Maxillo-Facciale, Chirurgia Plastica e Otorino
Igiene	Implantologia
Neurologia e Psichiatria	Odontoiatria Pediatrica e Pediatria
Radiologia	Clinica Odontostomatologia

3. Il modello ad effetti misti proposto

I dati analizzati provengono da un database di nostra costruzione e sono stati rilevati dai questionari dei PT somministrati agli studenti iscritti dal primo al sesto anno del corso di laurea in Odontoiatria e Protesi Dentaria per tutti gli Atenei Italiani, negli anni 2017 e 2018.

Per l'analisi delle performance delle Università Italiane, si è proceduto dapprima al calcolo dei indici sintetici di performance per ciascun anno di corso, per ogni università considerata:

Il data base contenente i risultati del progress test per gli anni 2017 e 2018 riporta il numero delle risposte corrette fornite da ciascuno studente partecipante al test per cui sono stati costruiti 3 diversi indici sintetici:

Indice A: va ad indicare la performance di ogni università per ciascun anno di corso nelle scienze di base, calcolato con la seguente formula:

$$\frac{\sum \text{Punteggio Totale Scienze di Base}}{15 \times n^{\circ} \text{ Rispondenti Scienze di Base}}$$

Indice B: va ad indicare la performance di ogni università per ciascun anno di corso nelle scienze cliniche, calcolato con la seguente formula:

$$\frac{\sum \text{Punteggio Totale Scienze Cliniche}}{15 \times \text{n}^\circ \text{ Rispondenti Scienze Cliniche}}$$

Indice C: indicherà la performance di ogni università per ciascun anno di corso nelle scienze di base ed in quelle cliniche, con la seguente formula:

$$\frac{15 \times \text{n}^\circ \text{ rispondenti} \times (\text{Indice A} + \text{Indice B})}{15 \times (\text{n}^\circ \text{ Rispondenti A} + \text{n}^\circ \text{ Rispondenti B})}$$

Avendo a nostra disposizione solamente i punteggi relativi a due anni si può verificare che uno stesso gruppo di studenti, appartenente ad un certo ateneo ed iscritto ad un certo anno del corso, potrà svolgere il test al più due volte. Dunque, non abbiamo una misurazione dei punteggi del test realizzata progressivamente nel tempo. In ogni caso, pur non essendo esattamente all'interno di un contesto longitudinale, abbiamo ritenuto di utilizzare una tipologia di modelli adatta ai dati panel. Tutti i modelli che abbiamo pensato di utilizzare all'interno dell'analisi sono racchiusi nel seguente:

$$y_{jdt} = (\alpha_{0d} + u_{0j}) + (\alpha_{1d} + u_{1j}) * t_1 + (\alpha_{2d} + u_{2j}) * t_2 + \varepsilon_{jdt}$$

dove:

$$\begin{aligned} j &= 1 \dots 31, && \text{sono gli atenei} \\ t &= 0 \dots 5, && \text{anni di iscrizione} \\ d &= 1 \dots 30, && \text{sono le discipline} \\ \alpha_{0d}, \alpha_{1d}, \alpha_{2d} &&& \text{sono gli effetti fissi di disciplina} \\ u_{0j}, u_{1j}, u_{2j} &&& \text{sono gli effetti casuali di ateneo} \end{aligned}$$

Questa è la base di partenza per il modello in cui viene inserito l'effetto casuale per le Università. In questo modello generale si ha, per ogni componente del polinomio di secondo grado, un effetto fisso di disciplina e un effetto casuale di ateneo. Per limitare la complessità del modello si è deciso di adottare una specificazione parsimoniosa della componente di secondo grado che dipende solo dalla variabile binaria c_d che indica se la disciplina d è clinica piuttosto che di base. Questo è il modello che si userà per studiare le traiettorie del Progress Test sia nell'anno di svolgimento del test 2017 che 2018:

$$y_{jdt} = (\alpha_{0d} + u_{0j}) + (\alpha_{1d} + u_{1j}) * t_1 + (\beta + \gamma c) * t_2 + \varepsilon_{jdt} \quad (1)$$

I termini di errore sono assunti normalmente distribuiti:

$$\begin{aligned} [u_{0d}, u_{1d}] &\sim N(0, \Sigma) \\ \varepsilon_{jdt} &\sim N(0, \sigma_\varepsilon^2), \end{aligned}$$

Le varianze dovranno essere poi stimate. Si può osservare come sia presente una parte dell'intercetta, un effetto lineare fisso e differente per ogni disciplina.

L'interpretazione dei coefficienti è legata al *random effect* che viene utilizzato per prendere in considerazione la dipendenza nei risultati dei test svolti all'interno di una stessa università.

Per la scelta del modello abbiamo utilizzato i criteri AIC, BIC confrontando un modello meno complesso in cui non era stato inserito l'effetto fisso lineare per le discipline con uno più complesso per cui avevamo stimato gli effetti fissi per il termine quadratico delle discipline. Alla fine si è scelto il modello meno complesso poiché i valori AIC e BIC non erano molto diversi rispetto a quelli relativi al modello più complesso.

I metodi di stima utilizzati sono basati sulla Massima verosimiglianza e sono stati implementati col software Stata.

4. Residui grezzi e residui empirici bayesiani

Nel nostro caso di studi realizzeremo delle previsioni per gli effetti casuali sul livello iniziale e sulla crescita dei punteggi per ciascun Ateneo. La previsione per il valore degli effetti casuali solitamente viene calcolata in due modi: utilizzando i *Raw Residuals* oppure con gli *Empirical Bayes Residuals*.

Supponiamo di essere in un contesto semplificato e di avere un modello con solamente l'intercetta random:

$$y_{it} = \alpha + \beta t + u_{0i} + \varepsilon$$

Se volessimo prevedere semplicemente il valore previsto degli effetti random utilizzando i *raw residuals* potremmo calcolarlo semplicemente prendendo la media campionaria per l'*i*-esimo gruppo e sottraendogli la previsione che si è realizzata per quel cluster. Il valore ottenuto in un certo senso sarebbe il valore che avremmo ottenuto facendo una regressione classica sul gruppo selezionato. Nel secondo caso invece dobbiamo prendere la distribuzione della media condizionata degli effetti random:

$$\widehat{u_{0i}^{EB}} = E \left[u_{0i} | y_{1i}, \dots, y_{t_i, i}, ; \hat{\alpha}, \hat{\beta}, \hat{\sigma}_\varepsilon^2, \hat{\sigma}_u^2 \right]$$

Questi ultimi sono conosciuti in letteratura come *residui ritirati* dato che esiste un fattore di restringimento dipendente dalle varianze di entrambi i livelli e dalla dimensione del cluster con cui possiamo calcolare gli Empirical bayes residual partendo dai raw residuals. Quindi in questo contesto semplificato ci basterà calcolare prima il valore dei raw come in precedenza e poi ricavare il fattore di *Shrinkage*. In questo caso ciò che si è scritto può essere sintetizzato con una semplice formula:

$$\hat{u}_j^{EB} = \left(\frac{\sigma_u^2}{\sigma_u^2 + \frac{\sigma_\varepsilon^2}{n_j}} \right) \hat{u}_j^{RAW}$$

dove:

$$0 < \frac{\sigma_u^2}{\sigma_u^2 + \frac{\sigma_\varepsilon^2}{n_j}} < 1 \quad \text{è il fattore di restringimento.}$$

$$\hat{u}_j^{RAW} \quad \text{sono i residui raw.}$$

Si può vedere come il fattore di restringimento sia compreso fra 0 e 1, dato che al numeratore è presente la varianza a livello di gruppo, al denominatore invece è sommata a quest'ultima componente la varianza dei residui di primo livello pesata per il numero di osservazioni presenti nel gruppo j .

Quindi abbiamo due quantità positive, e quella al denominatore non potrà essere più piccola rispetto a quella sopra. Dunque, il fattore di restringimento sarà tanto più grande quanto più grande sarà n_j . Quando questo valore è vicino ad 1 non c'è differenza fra queste due stime. Questo ci fa comprendere perché tale fattore viene definito di restringimento. Infatti, si potrebbe vedere che:

$$\hat{\beta}_{0j}^{EB} = \frac{\sigma_u^2}{\sigma_u^2 + \frac{\sigma_\varepsilon^2}{n_j}} \hat{\beta}_{0j}^{OLS} + \left(1 - \frac{\sigma_u^2}{\sigma_u^2 + \frac{\sigma_\varepsilon^2}{n_j}} \right) \alpha$$

Ovvero, gli EB sono semplicemente una media ponderata dal fattore di restringimento fra il valore dell'intercetta stimato con OLS per uno specifico gruppo e una stima globale dell'intercetta. Dunque, quanti più il j -esimo gruppo sarà piccolo tanto più grande sarà il fattore di restringimento avvicinandosi così alla stima generale. Utilizzando gli EB siamo in grado di prendere in considerazione il diverso numero di percentuali di risposte corrette all'interno di ciascuno dei 31 Atenei. Ci avvarremo di BLUP (Best Linear Unbiased Prediction) che minimizzano il mean square error e infatti in generale sono preferiti rispetto agli effetti fissi. Nel software Stata, implementeremo questo metodo.

5. Specificazione alternativa: effetti casuali discipline, effetti fissi Atenei

Grazie alla disponibilità di variabili a nostra disposizione, si sarebbe potuta scegliere una specificazione alternativa come la seguente:

$$y_{jdt} = (\alpha_{0d} + u_{0j}) + (\alpha_{1d} + u_{1j}) * t + (\beta + \gamma c) * t_2 + \varepsilon_{jdt} \quad (2)$$

In questo caso le osservazioni vengono raggruppate sulla base delle discipline e si inseriscono degli effetti fissi per le università. I risultati di questo modello alternativo sono stati inseriti all'interno dell'elaborato perché possono comunque essere utilizzati per verificare la plausibilità dell'assunzione di Normalità degli effetti random per gli atenei nel modello che si utilizzerà per l'analisi.

6. Modelli per proporzioni

Quest'ultimo modello (2), insieme a quello visto in precedenza (1) presenta delle criticità legate al fatto che si tratta la variabile di risposta, che è una proporzione che può assumere valori fra 0 e 100, utilizzando un modello lineare nei parametri. In letteratura sono presenti delle trasformazioni della variabile dipendente che, dopo aver variato il range di valori che quest'ultima può assumere (nel nostro caso fra 0 e 1), portandola nei reali, ci permettono di ricavare una relazione diretta fra il valore atteso della nostra variabile di risposta e il predittore lineare. In Fiksel (2021) si ritrovano due metodi con cui possiamo trattare variabili dipendenti e indipendenti che rappresentano proporzioni. Partendo da una situazione nella quale abbiamo un gruppo di variabili definite in un insieme:

$$S^D = \left\{ (x_1, x_2, \dots, x_d)' \mid x_j \geq 0, i = j, \dots, D; \sum_{i=j}^D x_j = 1 \right\}$$

a prescindere dal loro essere semplicemente variabili di risposta o variabili dipendenti, possono essere portate da S^D a R^{D-1} utilizzando due tipologie di trasformazioni, la trasformazione ILR:

$$ilr(x)_j = \sqrt{\frac{D-j}{D-j+1}} \ln \left(\frac{x_j}{(\prod_{k=j+1}^D x_k)^{\frac{1}{D-j!}}} \right)$$

e la trasformazione ALR:

$$alr(x)_j = \ln\left(\frac{x_j}{x_d}\right), j = 1, \dots, D - 1$$

Riadattando queste due trasformazioni al nostro caso di studio, nel quale abbiamo per ogni Ateneo e anno di iscrizione la proporzione di risposte corrette in tutte le discipline per un dato gruppo di studenti, si potrebbe inserire e trattare la nostra variabile direttamente in un modello di regressione lineare senza imporre nessun vincolo. Il problema nell'utilizzo di queste strategie è legato nel primo caso alle difficoltà interpretative che si potrebbero riscontrare: infatti si sarebbe costretti ad avere un'interpretazione dei coefficienti del modello sempre in relazione a tutte le altre variabili. Inoltre, per ottenere tutti i coefficienti, dovremmo realizzare più di un modello e dunque non potremmo interpretare i coefficienti congiuntamente. La seconda trasformazione ha il medesimo problema, ovvero si riesce a stimare un unico modello ma in certi casi possono comunque esserci diverse difficoltà di tipo interpretativo. L'ultima strada da poter percorrere è data dall'applicazione diretta di un modello lineare in seguito all'imposizione di determinate condizioni sui valori che possono essere assunti dalla variabile dipendente. Questo è ciò che viene solitamente fatto quando si utilizza la *Beta regression*. Quest'ultima viene utilizzata perché è estremamente utile nel caso in cui si vogliono studiare degli outcome definiti fra 0 e 1 e dunque le previsioni che verrebbero realizzate a partire da questo modello sarebbero comunque definite fra 0 e 1. Naturalmente questo si verifica perché si assume che la variabile di risposta si distribuisca come una Beta. Inoltre, sappiamo anche che questa metodologia viene spesso utilizzata nell'ambito delle scienze economiche e sociali. Anche in questo caso si sta scrivendo di un modello per la media condizionata al valore delle covariate e quindi si deve trovare una funzione link per la media condizionata in modo tale che questa non possa assumere valori maggiori di 1 o minori di 0. Dunque, assumendo che:

- $g(\cdot)$ sia la funzione link e $g(\cdot)^{-1}$ la sua inversa,
- μ_x il valore atteso condizionato,

si assume che $g(x\beta) = \mu_x$, dove il link potrebbe essere di qualsiasi tipo, ad esempio se fosse un link logit si otterrebbe:

$$\mu_x = \frac{\exp(x\beta)}{\{1 + \exp(x\beta)\}}.$$

Invece la varianza condizionata di questa distribuzione è data da:

$$Var(y|x) = \left\{ \frac{\mu_x(1-\mu_x)}{1+\psi} \right\},$$

dove il valore di $\psi > 0$ è un fattore di scala per la varianza condizionata. Dunque, si assume che la variabile di risposta abbia una densità come quella riportata:

$$f(y; \mu_x, \psi_x) = \frac{\Lambda[\psi_x]}{\Lambda[\mu_x, \psi_x]\Lambda[(1-\mu_x), \psi_x]} y^{\mu_x \psi_x - 1} (1 - y)^{(1-\mu_x)\psi_x - 1}.$$

Questa regressione può essere implementata molto facilmente su Stata e la stima dei parametri viene realizzata utilizzando sempre dei metodi basati sulla verosimiglianza. In ogni modo, si è scelto di non utilizzarla e proseguire con il modello presentato in precedenza.

7. Risultati relativi alle Performance delle Università

Grazie al calcolo degli indici A e B abbiamo ottenuto l'indice C da cui, per entrambi gli anni considerati, per ciascun anno di corso, abbiamo stilato una classifica dei corsi di studio attivi presso le diverse università,, in base al numero di risposte corrette date dagli studenti che hanno partecipato al PT.

- I dati relativi al **primo anno** di corso, riportati nella Tabella 2, evidenziano che nel 2017 l'Università più performante è stata quella di Napoli "Federico II", seguita dalla Sapienza e dall'Università di Parma. Le peggiori risultano essere l'Università degli Studi di Roma "Tor Vergata" e quella di Torino, che presentano un indice pari a 0.

Nel 2018 l'Università più performante è stata l'Aquila, seguita dall'Università Vita-Salute San Raffaele. Entrambe hanno notevolmente migliorato il loro indice rispetto all'anno precedente. Le peggiori risultano essere le università di Cagliari, Chieti, Messina, Palermo, Pisa, e Roma "Tor Vergata", tutte con un indice pari a 0.

- Per quanto riguarda le performance del **secondo anno** di corso si evidenzia che nel 2017 l'Università più performante è stata l'Università Vita-Salute San Raffaele seguita dalle Università di Bari e di Parma. La peggiore risulta essere l'Università degli Studi di Messina, che presenta un indice pari a 0.

Nel 2018 l'indice cala per tutte le Università, ma la più performante risulta ancora l'Università della Campania, seguita dall'Università di Napoli "Federico II". La peggiore risulta essere l'Università di Roma "Tor Vergata".

- Nel 2017 le performance relative al **terzo anno** di corso indicano che l'Università più performante è stata l'Università di Chieti seguita dalle Università di Padova e di Bari. La peggiore risulta essere l'Università degli Studi di Torino che presenta un indice pari a 0,91.

Tabella 2 – *Indice C, performance per tutti gli anni di corso, per tutte le Università italiane osservate, periodo 2017, 2018, in base al numero di risposte corrette fornite dagli studenti che hanno partecipato al PT.*

Università	1° anno		2° anno		3° anno		4° anno		5° anno		6° anno	
	2017	2018	2017	2018	2017	2018	2017	2018	2017	2018	2017	2018
Università degli Studi dell'Aquila	2,0	5,0	2,5	2,7	4,1	4,6	5,0	6,3	5,6	6,3	5,5	6,6
Università degli Studi di Bari	3,5	4,2	6,1	3,1	6,1	7,5	6,5	8,0	7,4	7,7	7,3	7,7
Università degli Studi di Bologna	3,5	2,0	3,3	3,5	4,9	4,6	5,5	6,1	7,0	6,8	6,1	6,9
Università degli Studi di Brescia	3,7	1,5	4,1	2,1	5,4	4,7	6,2	5,4	7,5	6,3	7,4	8,4
Università degli Studi di Cagliari	2,8	0,0	3,0	2,1	5,3	2,6	6,1	5,7	6,0	6,3	5,3	6,2
Università degli Studi della Campania	3,0	1,5	4,1	3,1	5,3	2,7	6,6	7,3	7,5	8,1	7,4	7,8
Università degli Studi di Chieti	3,2	0,0	3,7	4,9	7,2	4,2	6,4	5,7	6,5	6,1	6,6	7,3
Università degli Studi di Ferrara	1,9	2,6	3,0	3,0	5,2	3,7	5,5	4,6	5,4	6,4	6,4	6,2
Università degli Studi di Firenze	2,5	2,1	3,7	2,4	4,5	4,2	5,6	4,7	6,6	5,9	6,7	5,4
Università degli Studi di Foggia	3,1	1,5	5,3	4,2	4,9	5,1	6,2	4,3	6,3	7,1	7,4	6,7
Università degli Studi di Genova	1,4	2,7	2,9	2,4	3,8	4,2	5,8	5,2	5,7	6,6	5,0	5,5
Università degli Studi dell'Insubria	3,6	3,3	3,7	3,7	4,2	4,5	5,4	5,3	6,3	5,7	5,9	6,5
Università Politecnica delle Marche	3,6	3,8	2,6	3,7	4,5	5,1	5,5	4,9	6,0	6,1	5,6	6,4
Università degli Studi di Messina	3,1	0,0	0,0	2,7	4,0	3,5	6,6	6,2	7,6	6,8	7,5	7,2
Università degli Studi di Milano	3,0	2,2	2,6	2,9	5,6	3,4	6,0	4,8	6,9	5,6	7,3	6,2
Univ. degli Studi di Milano "Bicocca"	2,4	1,7	4,7	3,2	5,6	5,6	6,6	6,1	7,1	6,6	7,2	6,8
Università degli Studi di Modena	3,6	3,1	2,7	3,7	4,6	4,5	5,7	5,4	6,1	6,1	6,4	6,0
Università degli Studi di Napoli "Federico II"	4,3	4,3	5,0	4,9	5,8	6,0	6,7	7,5	7,0	7,5	7,3	7,2
Università degli Studi di Padova	3,5	2,9	4,9	3,7	6,3	5,4	6,2	6,0	7,1	6,5	7,2	5,4
Università degli Studi di Palermo	1,4	0,0	1,7	2,4	3,2	2,7	5,7	4,5	5,5	5,4	5,4	5,6
Università degli Studi di Parma	3,8	2,2	5,4	4,8	4,8	6,0	6,2	6,3	7,5	7,3	6,5	7,6
Università degli Studi di Pavia	2,4	1,6	4,5	2,7	4,8	3,6	5,5	5,2	6,2	6,1	6,1	6,2
Università degli Studi di Pisa	1,9	0,0	2,6	1,9	3,7	2,7	5,0	4,3	5,7	5,3	5,8	5,9
Sapienza Università di Roma	3,9	3,8	4,2	4,0	5,6	5,3	6,5	5,9	6,7	7,0	7,1	6,9
Univ. Studi di Roma "TorVergata"	0,0	0,0	2,1	1,8	4,1	3,9	6,4	4,8	5,5	6,8	5,8	5,6
Università Vita-Salute San Raffaele	3,1	4,5	6,1	4,2	5,8	7,3	7,0	6,9	7,0	7,4	6,9	7,4
Università degli Studi di Sassari	2,1	2,2	2,2	2,2	3,7	3,1	4,7	4,7	6,1	5,8	6,6	6,2
Università degli Studi di Torino	0,0	1,3	1,5	2,5	0,9	2,1	1,9	2,9	1,4	3,5	0,0	0,0
Università degli Studi di Trieste	2,7	2,5	2,8	2,7	5,5	4,1	5,2	5,3	5,9	5,8	6,1	5,8
Università degli Studi di Verona	3,5	3,0	4,4	3,7	5,0	4,4	5,7	5,1	6,2	5,4	5,7	5,9

Nel 2018 l'Università più performante risulta essere quella di Bari, che dal terzo posto dell'anno precedente arriva al primo, seguita dall'Università Vita-Salute San Raffaele. La peggiore risulta essere sempre l'Università di Torino, anche se ha incrementato la propria performance rispetto all'anno precedente.

- Le performance per il **quarto anno** di corso nel 2017 indicano che l'Università più performante risulta essere l'Università Vita-Salute San Raffaele seguita dall'Università di Napoli "Federico II". La peggiore risulta essere l'Università degli Studi di Torino che presenta un indice pari a 1,93. Nel 2018 l'Università più performante risulta essere quella di Bari, che dal settimo posto dell'anno precedente arriva al primo.

Il peggiore indice di performance compete, ancora una volta, all'Università di Torino.

- Le performance delle Università per il **quinto anno** di corso, nel 2017 vedono in pole position l'Università di Messina, mentre l'ultimo posto della graduatoria spetta all'Università degli Studi di Torino che presenta un indice pari a 1,37. Nel 2018 l'Università più performante risulta essere quella della Campania, che dal quarto posto dell'anno precedente arriva al primo. La peggiore risulta essere sempre l'Università di Torino, nonostante un incremento di performance rispetto al 2017.
- In ultimo, osserviamo le performance per il **sesto anno** di corso. Nel 2017 l'Università più performante risulta essere l'Università di Messina, seguita dalle Università di Brescia e della Campania. Ancora una volta all'ultimo posto si posiziona l'Università degli Studi di Torino, con un indice pari a 0.

Nel 2018 l'Università di Brescia brilla per i suoi risultati. Il peggior risultato spetta ancora all'Università di Torino, che mantiene il medesimo valore pari a 0.

8. Risultati relativi al modello di crescita

L'analisi è stata condotta utilizzando un modello ad effetti misti nel quale abbiamo inserito degli effetti fissi per ogni disciplina presente nel test insieme ad una variabile dicotomica con cui distinguere fra discipline cliniche e di base, e degli effetti random per i diversi atenei in cui il test è stato somministrato. Sono state analizzate le traiettorie di crescita per le diverse discipline per tutti gli atenei. Abbiamo così studiato lo sviluppo medio dei punteggi nelle specifiche discipline durante i sei anni di iscrizione sempre distinguendo fra discipline cliniche e di base. Si sono ottenute le previsioni per i componenti dei *random effects* tramite il metodo EB implementato in *Stata*. Avendo le previsioni per le componenti aleatorie, siamo stati in grado di osservare quanto le traiettorie degli specifici atenei deviassero dall'andamento medio per quanto riguarda la crescita dei punteggi nei diversi anni di iscrizione rispetto al livello di partenza.

Abbiamo ottenuto buone traiettorie di crescita per le discipline a livello nazionale, in linea con le analisi descrittive e con precedenti studi sui *Progress Test* delle facoltà di Odontoiatria e protesi dentaria del 2017 (Crocetta et al. 2018). Infatti, analizzando l'anno 2017 e rappresentando le traiettorie delle singole discipline, abbiamo osservato la presenza di importanti differenze nella percentuale di risposte corrette fra queste. Per quanto riguarda le discipline di base: istologia ed anatomia è stata la

disciplina di base per cui si è ottenuto il punteggio medio più elevato. Infatti, la percentuale di risposte corrette raggiunge un suo massimo solamente nel quinto anno e nell'anno successivo questo valore non diminuisce enormemente. Le stesse considerazioni si possono fare per Patologia generale. Molto probabilmente le conoscenze trasmesse da questi insegnamenti vengono riprese durante tutto il corso di laurea. Scienze comportamentali è invece la disciplina di base per cui si ottiene il punteggio medio più basso. La sua traiettoria ha infatti una curvatura molto poco accentuata nei 6 anni di iscrizione e quindi il valore nei punteggi non varia molto. Radiologia, Anestesiologia e Chirurgia generale sono le discipline di base per cui è presente una crescita più sostenuta.

Per la maggior parte delle discipline di base come chimica, biochimica e fisica il punto di massimo si ha nell'anno in cui il corso viene erogato agli studenti. La distribuzione delle previsioni per le discipline di base può essere modellata come un polinomio di secondo grado. In molti casi il valore dei punteggi aumenta nei primi anni fino al raggiungimento di un plateau. In altri assume proprio la forma di una parabola con la concavità rivolta verso il basso (fisica e chimica). Dunque, lo sviluppo delle traiettorie previste negli anni di iscrizione è esattamente in linea con quanto si può leggere nella teoria in quest'ambito di studio (McNeish et al. 2022).

Anche per le discipline cliniche si sono osservati risultati coerenti con il precedente studio realizzato in Italia. Infatti, si sono stimati livelli di partenza per ogni disciplina clinica molto bassi rispetto ai valori di partenza delle discipline di base ed anche in questa analisi abbiamo riscontrato che: Gnatologia è la clinica per cui la percentuale di risposte media è più bassa. La diversa natura delle discipline caratterizza le traiettorie delle materie cliniche rendendole molto simili ad una linea retta. In questo caso non si raggiunge un punto di "massima conoscenza". Queste materie vengono insegnate più tardi e sono indispensabili per il lavoro futuro che gli studenti dovranno svolgere. Per queste materie all'ultimo anno non si registra un calo significativo nella percentuale di risposte esatte.

Risultati simili a questi sono stati confermati dall'analisi delle traiettorie per le discipline sui dati del 2018. Dall'analisi sulle previsioni degli effetti casuali si è constatato come i valori iniziali dei punteggi in ciascun ateneo italiano e allo stesso modo i tassi di crescita non sono uguali poiché esistono differenze iniziali significative fra i vari atenei che producono disuguaglianze, spesso non colmate. Questa è una caratteristica che abbiamo riscontrato negli anni 2017 e 2018 e presumibilmente è dovuta al sistema del punteggio minimo con cui si accede alle facoltà e al prestigio che in certi casi viene esercitato da alcuni atenei. Gli studenti meno preparati tendono ad entrare in quelle facoltà in cui il punteggio minimo al test

d'ingresso è più basso e quindi è semplice accedervi. Per questo motivo, questi atenei avranno punteggi peggiori dei primi anni. Esistono infatti tre atenei per cui si ottengono, nel 2017 e nel 2018, dei valori di partenza molto elevati: Bari, San Raffaele Milano e Napoli.

In ogni modo abbiamo sempre stimato una relazione negativa fra le due componenti casuali. Questo è uno dei risultati più interessanti fra quelli ottenuti perché significa che non è scontato che tutti gli atenei con un valore poco elevato inizialmente abbiano un livello basso nella percentuale di risposte corrette al test debbano mantenere punteggi bassi anche negli anni successivi. Spesso si prevede un effetto aleatorio negativo sulla crescita dei punteggi per ciascuno di questi atenei. Al contrario atenei come Messina, Brescia e Milano Bicocca per cui non abbiamo previsto valori iniziali molto elevati hanno una crescita piuttosto sostenuta che li porta a raggiungere al sesto anno la percentuale di risposte corrette di Bari e Salerno. Nel 2018 l'ateneo che al sesto anno ottiene il numero di risposte corrette più elevato è quello barese. Naturalmente non tutti gli atenei presenti nello studio hanno effetti sulla crescita elevati e quindi non raggiungono uno stesso livello di punteggio finale. Al contrario si vede molto chiaramente come esistano più livelli verso i quali convergono le traiettorie dei vari atenei. In generale nonostante delle differenze legate alla magnitudo delle previsioni i risultati osservati nei due anni sono piuttosto simili.

Dato che, con questa specificazione, la stima dell'effetto è libera e non si impone la normalità, si è utilizzato questo modello per controllare se la normalità degli effetti random dipende dalle assunzioni fatte in partenza oppure no. Il fatto che i risultati vengano simili significa che l'assunzione di normalità è plausibile.

Criticità.

La specificazione utilizzata presenta molteplici criticità. Ciascuno studente ha partecipato al Progress Test al massimo due volte e per questo sono presenti all'interno del data-set originale studenti appartenenti a molteplici coorti di laurea. Il risultato è che non vengono seguiti i punteggi realizzati dallo stesso gruppo di studenti per l'intera durata del corso di Odontoiatria.

Un altro problema è legato al trattamento della nostra variabile dipendente (la percentuale di risposte corrette, quindi una proporzione) con un modello lineare. Questo porta a previsioni prive di un reale significato perché minori di 0 o maggiori di 100. Nelle due analisi per l'anno 2017 e 2018, nel 2017 solamente lo 0.722% delle previsioni doveva essere corretto mentre nel 2018 l'1.25%. Quindi in ogni caso il problema non si è presentato molto spesso. L'ideale sarebbe stato trattare il nostro

caso di studio utilizzando un modello con una variabile di risposta binomiale ma purtroppo si riscontrano altri problemi nella stima del modello. Un'altra soluzione è il trattamento della variabile di risposta come una proporzione e dunque l'utilizzo di adeguate trasformazioni come la ILR o ALR attraverso le quali variamo il supporto della variabile di risposta dall'intervallo 0,1 ai numeri reali. Altra metodologia, già ricordata nel par. 6, è la Beta Regression, nella quale imponiamo il supporto fra 0 e 1 alla variabile dipendente. In ogni modo abbiamo continuato ad utilizzare il modello lineare visto in precedenza, per ragioni principalmente legate alla facilità con cui esso può essere interpretato, e come si è già sottolineato, la bassa proporzione di valori previsti maggiori di 100 o minori di 0.

9. Conclusioni

Questo studio ha ampliato sull'analisi dei Progress Test, strumento con cui si valuta il livello delle conoscenze acquisite dagli studenti nel proprio percorso di studi, di 31 Facoltà italiane di Odontoiatria e Protesi Dentaria negli anni 2017 e 2018. Il Progress Test è un test composto da 300 domande la metà delle quali riguarda discipline cliniche, l'altra metà riguarda discipline di base. Viene erogato nello stesso formato e con le stesse domande per tutti gli studenti di una certa Facoltà indipendentemente dall'anno di corso al quale sono iscritti. L'utilità del Progress Test e della sua analisi è legata alla possibilità di capire quali siano le principali lacune per gli studenti e trovarne un rimedio, formando così operatori sanitari migliori.

Dall'analisi effettuata il PT risulta essere un buon indicatore per la valutazione dei progressi e/o regressi delle conoscenze acquisite nel corso degli anni di formazione. In particolare, dall'analisi statistica sui Corsi di Laurea in Odontoiatria e Protesi Dentaria si pone l'attenzione sull'andamento delle discipline individuali degli studenti dal 1° all'ultimo anno di corso e la valutazione riguardante la crescita e/o decrescita del livello di preparazione degli studenti e i risultati totali per le Scienze di Base e le Scienze Cliniche per ogni Ateneo ma, col focus sulla classificazione CENSIS di questi in Piccoli e Grandi Atenei; da cui si evince che gli studenti a prescindere se facciano parte di un Piccolo o di un Grande Ateneo disperdono le loro conoscenze mano a mano che avanzano nel percorso di studi, questi risultano essere ancora molto legati al modello del "teacher-centred" ovvero la loro formazione è basata sul docente e sul superamento dell'esame in sé, segno che gli studenti italiani hanno bisogno di cambiare il loro metodo di apprendimento

e il modo di erogazione della didattica da parte delle Università in modo che i futuri medici mantengano le conoscenze apprese anche al termine del loro percorso di studi, ovvero quando si immetteranno nel mondo del lavoro, in quanto rappresentano conoscenze imprescindibili da possedere.

Dal confronto delle performance si evidenzia quali sono le discipline di Scienze di Base e Cliniche che hanno un minor rendimento così da sapere su quali insegnamenti puntare per migliorare la didattica in modo che gli studenti alla fine del sesto anno abbiano una preparazione adeguata che gli permetta di affrontare qualsiasi problema che gli si presenti, sia tramite la conoscenza della materia che attraverso lo sviluppo di un pensiero critico che si può raggiungere solo attraverso uno studio indotto dall'efficacia dei docenti nell'insegnamento.

Anche dal calcolo degli indici sintetici, A, B e C, si evince come la performance delle università basate sulle risposte corrette date dagli studenti partecipanti al Progress Test non sia correlata alla dimensione degli Atenei come dimostrano le Università dell'Aquila, di Brescia, di Foggia e l'unica università privata dell'analisi l'Università Vita-Salute San Raffaele, che presentano nei sei anni di corso delle performance alquanto elevate. Da questa analisi si evince che le Università meno performanti sono due Grandi Atenei quali l'Università di Roma "Tor Vergata" e l'Università di Torino. I risultati spingono a continuare con questa metodologia al fine di continuare a migliorare sotto tutti i punti di vista, la didattica e l'apprendimento degli studenti delle facoltà di Odontoiatria e Protesi Dentaria, sfruttando quanto più possibile uno degli obiettivi del Progress Test, fornire agli studenti la capacità di consapevolezza e riflessione delle proprie conoscenze e competenze e di riuscire a tenerle senza dispendere negli anni.

Riferimenti bibliografici

- Albanese M., Case S.M. (2016). *Progress testing: critical analysis and suggested practices*. In: *Advances in Health Sciences Education* 21.1, pp. 221-234.
- Cerasti S., Broggi F., Ricci C., Monaco E., Nacinovich R., Tagliabue L. (2019). *Il Progress Test come strumento di verifica del processo formativo nei CDL in terapia della neuro e psicomotricità dell'età evolutiva: analisi dei risultati a livello nazionale e integrazione nel progetto teco di ANVUR*. In: *TNPE* 1, pp. 63-92
- Coombes L., Ricketts C., Freeman A., Stratford J. (2010). *Beyond assessment: feedback for individuals and institutions based on the progress test*. In: *Medical teacher* 32.6, pp. 486-490

- Crocetta C., Brindisi M., Lo Muzio L. (2018). *Analisi dei risultati del progress test 2017 dei Corsi di Laurea in Odontoiatria e Protesi Dentaria*. In: Journal of Italian Medical Education (2018) 2.78, pp. 3487-3493
- Fahrmeir L., Kneib T., Lang S., Marx D. (2013). *Regression models*. In: *Regression*. Springer, pp. 348–412
- Fiksel J., Zeger S., Abhirup D. (2021). *A transformation-free linear regression for compositional outcomes and predictors*. In: Biometrics
- Hedeker D. (2004). *An introduction to growth modeling*. In: *The Sage handbook of quantitative methodology for the social sciences*, pp. 215-234
- Grimm K. J., Ram N. (2009). *Nonlinear growth models in M plus and SAS*. In: Structural Equation Modeling 16.4, pp. 676–701
- McNeish D., Dumas D., Torre D., Rice N. (2022). *Modelling time to maximum competency in medical student progress tests*. In: Journal of the Royal Statistical Society. Series A: Statistics in Society
- Rabe-Hesketh S., Skrondal A. (2008), *Multilevel and longitudinal modeling using Stata*. In: STATA press, pp. 264-267
- Tenore A., Basili S., Lenzi A., Marangon M., Proietti M. (2012). *Il Progress Test 2011*. In: Journal of Italian Medical Education 56, pp. 2487-2509
- Tenore A., Basili S., Proietti M. (2013). *Il Progress Test 2012*. In: Journal of Italian Medical Education 60, pp. 2699-2704.