



UNIVERSITÀ DEGLI STUDI DI PALERMO

SCUOLA POLITECNICA.

Statistica per l'Analisi dei Dati

ANALISI LONGITUDINALE DELLE CARRIERE UNIVERSITARIE
NELLE LAUREE SCIENTIFICHE DI BASE IN ITALIA ANNO 2011-2012

RAPPORTO DI STAGE DI

Alessandro Albano

Tutor universitario:

Massimo Attanasio

Tutor aziendale:

Giovanni Boscaino

TRIENNALE



Indice

| | |
|--|----|
| Introduzione..... | 3 |
| Capitolo 1. 1.1 L'istruzione terziaria in italia | 5 |
| 1.2 Il progetto lauree scientifiche | 7 |
| Capitolo 2 2.1 I dati e gli obiettivi | 10 |
| 2.2 Costruzione del dataset | 12 |
| 2.3 Le variabili in studio | 13 |
| Capitolo 3 3.1 Analisi preliminare dei dati..... | 14 |
| 3.2 Analisi del successo..... | 19 |
| Capitolo 4 4.1 Modello logit | 22 |
| 4.2 Modello logit - Analisi dei risultati..... | 24 |
| Conclusioni | 28 |
| Bibliografia..... | 30 |

Introduzione

Lo stage è stato svolto presso il Dipartimento SEAS dell'Università degli Studi di Palermo, in cooperazione con il collega Andrea Priulla, sotto la supervisione del dottor Giovanni Boscaino come tutor Aziendale. Questo lavoro nasce nell'ambito del Protocollo di collaborazione per lo Studio della Mobilità Studentesca Universitaria che è stato stipulato tra l'Ufficio di Statistica e Studi del Ministero dell'Istruzione, dell'Università e della Ricerca e le l'Università degli Studi di : Cagliari, Palermo, Siena e Torino. Con questo protocollo di collaborazione, le Parti intendono analizzare le carriere universitarie e il fenomeno della mobilità studentesca nel nostro Paese.

Il nostro studio, invece, si focalizza sull'analisi delle carriere degli immatricolati della coorte del 2011-2012 nelle lauree scientifiche di base.

Per studiare il fenomeno della mobilità è utile un richiamo al contesto della formazione universitaria e delle sue recenti evoluzioni, in particolare si farà riferimento al Processo di Bologna e alla Conferenza di Lisbona perché rappresentano delle pietre miliari dell'evoluzione delle università degli ultimi 20 anni e perché investono in termini sostanziali le lauree scientifiche.

Il tema della qualità dell'istruzione universitaria ha assunto rilevante importanza del Processo di Bologna del 1999. In quell'occasione i ministri dell'istruzione di 29 paesi europei hanno sottoscritto un accordo, la dichiarazione di Bologna, con il fine di creare una comune cornice europea dell'istruzione terziaria, l'European Higher Education Area(EHEA). Per realizzare questo spazio comune occorre, tra le altre cose, standardizzare i cicli formativi (bachelor – master – Ph.D) e il sistema dei crediti; promuovere la mobilità di docenti e studenti fra università e paesi; accrescere la capacità di attrazione delle università europee nei confronti dei cittadini extra-europei. L'iniziativa, che si fonda sull'adesione volontaria dei paesi, si è successivamente estesa oltre i confini dell'Unione Europea, e coinvolge oggi 47 nazioni. Per favorire la qualità dell'istruzione a livello europeo, nel 2000 è stato istituito l'European Network for Quality Assurance in Higher Education (ENQA) e nel 2008 l'European Quality Assurance Register (EQAR), network transnazionali che hanno lo scopo di assicurare standard qualitativamente alti degli atenei e delle agenzie di valutazione. Proprio l'ENQA è stata incaricata dai ministri degli paesi firmatari del Processo di Bologna di elaborare “una base condivisa di standard, procedure e linee guida sui processi di assicurazione della qualità” (Comunicato di Berlino del 19 settembre 2003). Il risultato è il documento “Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG)”, approvato a Bergen nel 2005, il cui scopo

è fornire una guida per lo svolgimento dei processi di Assicurazione di Qualità, tanto alle istituzioni di istruzione superiore quanto alle agenzie di valutazione esterna. Le ESG sono strutturate su tre livelli: i) l'assicurazione interna della qualità delle istituzioni di istruzione superiore; ii) l'assicurazione esterna della qualità dell'istruzione superiore e iii) l'assicurazione della qualità delle agenzie di assicurazione esterna della qualità. Come in un sistema a matryoska, il livello superiore delle ESG presuppone l'esistenza del livello precedente, per cui la valutazione esterna si basa sulla valutazione interna. Considerata la pluralità di sistemi normativi, politici e accademici presenti all'interno dello Spazio Europeo dell'Istruzione Superiore, ENQA ha scelto di elaborare, quali standard di qualità, dei principi generici piuttosto che norme specifiche o di carattere restrittivo o prescrittivo. Inoltre le ESG sono state formulate in maniera tale da essere applicabili a tutte le università e le agenzie europee di Assicurazione della Qualità, prescindendo dalle loro specificità organizzative e dal contesto nazionale in cui sono inserite. I paesi aderenti al Processo di Bologna hanno creato quindi dei sistemi di Assicurazione di Qualità aderenti alle ESG ma, allo stesso tempo, differenti nelle modalità adottate. In generale si possono riscontrare nelle procedure di accreditamento e valutazione delle tendenze evolutive comuni quali: il passaggio dell'attenzione dai corsi di studio alle corrispondenti istituzioni; la parallela riduzione del carico di lavoro degli atenei connesso alle visite; la centralità degli studenti nei processi di Assicurazione di Qualità e delle informazioni ad essi destinate. Il fondamentale tema dell'istruzione viene approfondito anche nella "strategia di Lisbona" prima e la "strategia Europa 2020". Entrambe pongono la conoscenza e l'innalzamento dei livelli di istruzione al centro della strategia di sviluppo europea. Nel programma Istruzione e Formazione 2020 (ET 2020) è stato delineato un "quadro strategico aggiornato per la cooperazione europea nel settore dell'istruzione e della formazione, con l'obiettivo di affrontare le sfide sostanziali che l'Europa deve superare per diventare un'economia basata sulla conoscenza e rendere l'apprendimento permanente una realtà per tutti" (Consiglio dell'Unione Europea, 2009). Con riferimento all'istruzione terziaria, per l'Unione è stato definito l'obiettivo di innalzare la quota dei laureati nella popolazione di età compresa tra i 30 e i 34 anni al 40% entro il 2020. I singoli paesi sono stati invitati a contribuire al conseguimento degli obiettivi europei definendo obiettivi nazionali che tenessero conto delle condizioni specifiche di ciascun paese oltre che, evidentemente, della capacità e volontà politica di perseguire un generale innalzamento della quota di laureati nella popolazione. L'Italia ha definito per il 2020 un obiettivo del 26%, in linea con i trend di crescita degli anni 2000, ma molto lontano dall'obiettivo medio europeo. Tra i grandi paesi, il Regno Unito, la Spagna e la Francia già superano il target europeo. Il Regno Unito non ha fissato specifici obiettivi; la Spagna si sta avvicinando all'obiettivo nazionale del 44%; la Francia appare in ritardo rispetto all'ambizioso obiettivo del 50%, come pure la Germania, che,

muovendo da livelli inferiori alla media europea, non sembra nelle condizioni di innalzare il tasso di laurea al 42% prefissato.

(Anvur 2016)

Capitolo 1

L'istruzione terziaria in Italia

L'Italia risulta essere rispetto al resto d'Europa ancora in sostanziale ritardo, infatti prosegue il calo degli studenti che dopo il diploma scelgono di proseguire gli studi all'università (fig 1), l'andamento negli anni del Tasso di passaggio Scuola-Università (rapporto fra gli immatricolati in un certo anno accademico e i diplomati dell'anno scolastico precedente) si rileva per l'Italia un trend fortemente decrescente

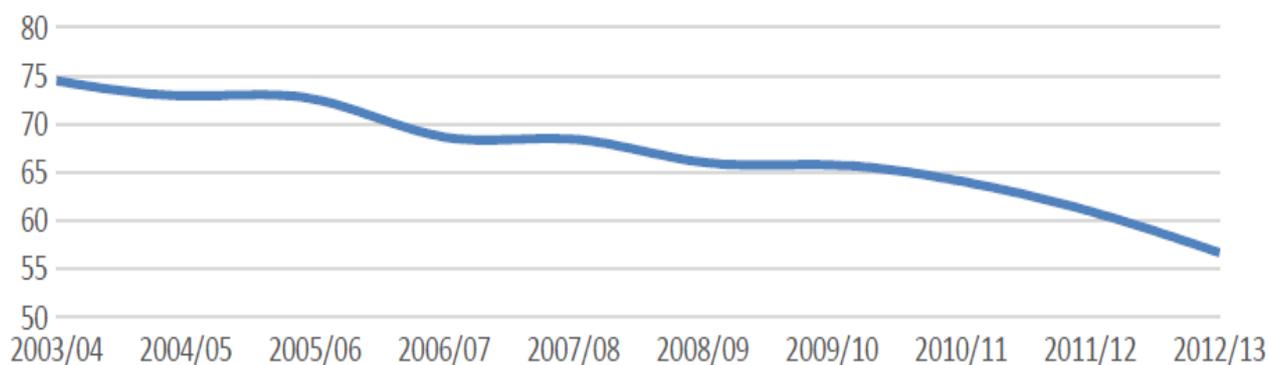
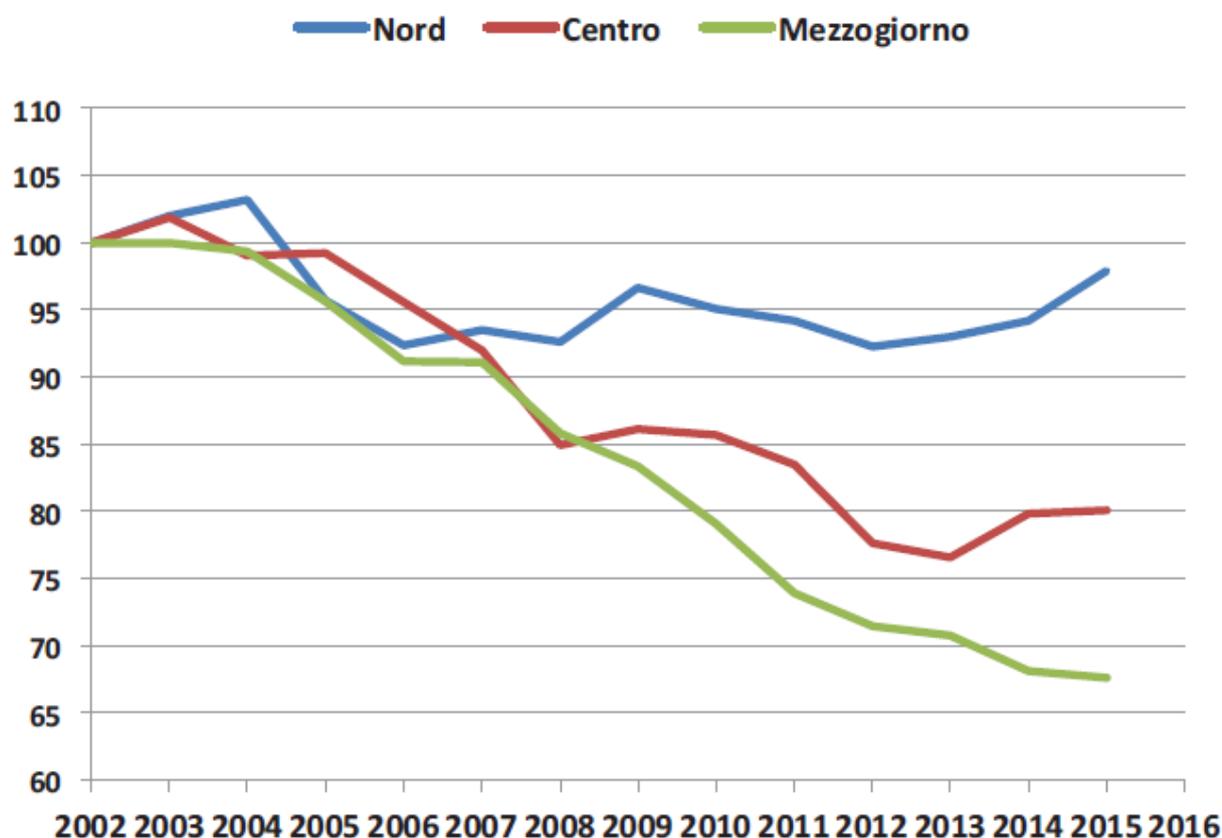


Figura 1 – Italia: Tasso di passaggio Scuola-Università

Fonte: MIUR

Stratifichiamo adesso per ripartizione geografica dividendo in tre macro-categorie le regioni italiane al fine di vedere se questa di differenziano per tassi di passaggio Scuola-Università.

Fig.2– Andamento degli immatricolati per area territoriale di residenza (numeri indice 2002/03=100)



Il calo degli immatricolati è stato nettamente più marcato per gli studenti residenti nel Centro e nel Mezzogiorno. Per i residenti al Nord la riduzione è stata pari al 6,8%. Anche il recupero dell'ultimo anno si concentra tra i residenti al Nord. Considerando i giovani immatricolati, a ridosso della conclusione delle scuole superiori, ovvero gli immatricolati con età pari o inferiore a 20 anni, gli immatricolati sono diminuiti di quasi l'11% tra i residenti nel Mezzogiorno, del 4,5% tra i residenti al Centro e sono aumentati dell'8,5% tra i giovani del Nord. Anche in questo caso nel 2015/16 il recupero è stato maggiore al Nord.

1.2 Il progetto lauree scientifiche

Il Piano Lauree Scientifiche anche se non esplicitamente fa riferimento a temi di qualità e di monitoraggio dei processi formativi (la riduzione dell'abbandono entra in gioco a pieno titolo). Esso è direttamente investito dagli obiettivi di Bologna, perché il numero dei laureati nelle Lauree Scientifiche è estremamente basso in termini assoluti e il numero degli abbandoni copre il 30% circa al primo anno.

Il PLS, istituito a partire dal 2004 su iniziativa del MIUR, della Conferenza dei Presidi di Scienze e Tecnologie e di Confindustria, rappresenta una iniziativa volta a favorire l'acquisizione di competenze scientifiche meglio rispondenti alle sfide della società contemporanea ed alle attese del mondo del lavoro da parte degli studenti e a rafforzare l'impatto della formazione sulla società.

Il Ministero, nell'ambito del Decreto Ministeriale 976/2014 (noto come "Fondo Giovani"), ha confermato e consolidato il proprio investimento nel PLS con riferimento al periodo 2014-2016 secondo le seguenti linee di azione:

- L'estensione del numero di classi di laurea coinvolte dal Piano;
- Il rafforzamento del ruolo delle Università nella realizzazione dei progetti, anche al fine di permettere una maggiore interdisciplinarietà tra le diverse aree coinvolte
- La creazione di progetti più strutturati, in termini di risorse disponibili e di durata
- L'inserimento di un'azione di intervento, mirata a ridurre il tasso di abbandono tra il primo e il secondo anno e, più in generale, al miglioramento della didattica universitaria del I ciclo nei Corsi di Laurea delle discipline scientifiche;
- La creazione di "Progetti Nazionali" per un miglior coordinamento tra le sedi e per una più efficace attività di monitoraggio e valutazione.

Tra le azioni, sono incluse quelle finalizzate a:

A. Mettere a sistema la pratica del "laboratorio" per l'insegnamento delle scienze di base, in particolare al fine dell'orientamento formativo degli studenti dell'ultimo triennio della Scuola secondaria di II grado;

B. Aprire una nuova sperimentazione nelle Scuole e negli Atenei di attività didattiche di autovalutazione e recupero, finalizzate al miglioramento della preparazione degli studenti relativamente alle conoscenze richieste all'ingresso dei corsi di laurea scientifici, in collegamento

con le verifiche e gli obblighi formativi aggiuntivi previsti dai corsi di laurea ai sensi dell'articolo 6, comma 1, del DM 270/04;

- C. Consolidare, sviluppare e mettere a sistema le opportunità di crescita professionale dei docenti di materie scientifiche in servizio nella Scuola secondaria di secondo grado, in relazione ai temi sopra indicati, che già si sono cominciate a realizzare attraverso le azioni del Piano Lauree Scientifiche;
- D. Ridurre il tasso d'abbandono tra il primo e il secondo anno nel corso degli studi universitari attraverso l'innovazione di strumenti e metodologie didattiche.

Le azioni A e C erano già presenti nei precedenti PLS e riguardano le attività di laboratorio con gli studenti e la formazione dei docenti. Il presente lavoro si inserisce tra le attività relative all'Azione D, in quanto mirato a sostenere le attività che possono favorire il completamento degli studi e la riduzione del tasso di abbandono nei Corsi di Laurea delle discipline. Le attività di questa azione saranno anche mirate all'introduzione di strumenti e metodologie didattiche innovative coerenti con l'approccio dello studente al centro delle attività di apprendimento per un miglioramento generale della didattica del I ciclo e la riduzione del tempo necessario per concludere gli studi.

Per affrontare in maniera più efficace possibile il problema dell'abbandono all'interno dei corsi PLS vengono adottate diverse misure, di tipo orientativo, tra cui:

- Offrire agli studenti degli ultimi anni delle scuole superiori opportunità di conoscere temi, problemi e procedimenti caratteristici dei saperi (scientifici), anche in relazione ai settori del lavoro e delle professioni, al fine di individuare interessi e disposizioni specifiche e fare scelte consapevoli in relazione a un proprio progetto personale;
- Mettere in grado gli studenti degli ultimi anni delle scuole superiori di autovalutarsi, verificare e consolidare le proprie conoscenze in relazione alla preparazione richiesta per i diversi corsi di laurea (scientifici).

Le classi di laurea che fanno parte di questo progetto sono:

TAB I Corsi di laurea facenti parte del progetto PLS

| Denominazione | Codice | |
|---------------|--------|--------|
| | DM270 | DM 509 |
| Statistica | L 41 | 37 |
| Geologia | L 34 | |
| Fisica | L 30 | 25 |
| Chimica | L 27 | 21 |
| Matematica | L 35 | 32 |
| Biologia | L 13 | 12 |
| Biotecnologie | L 2 | 1 |

È doveroso precisare che, anche la classe di laurea Scienza dei Materiali è inclusa nel progetto, ma questa corrisponde ad un classe di laurea di Chimica o di Fisica nei diversi atenei italiani.

Capitolo 2

I dati e gli obiettivi

Vediamo come si distribuiscono negli atenei italiani le sedi dei corsi di laurea PLS

TAB II : Distribuzione dei PLS negli Atenei italiani per sede e per disciplina. AA.AA 2016-2018

Legenda: X presenza di un corso di laurea PLS nell'ateneo della città; - assenza di un corso di laurea PLS nell'ateneo della città

| Distribuzione dei PLS negli Atenei italiani per sede e per disciplina. AA.AA 2016-2018 | | | | | | | |
|--|---------|--------|------------|------------|---------------|----------|--------------------------|
| Sede | Chimica | Fisica | Matematica | Statistica | Sc. Materiali | Geologia | Biologia & Biotecnologie |
| Ancona | - | - | - | - | - | - | X |
| Bari | X | X | X | X | X | X | X |
| Basilicata | X | - | X | - | - | X | X |
| Bergamo | - | - | X | - | - | - | - |
| Bologna | X | X | X | X | - | X | X |
| Brescia | - | - | - | - | - | - | X |
| Brescia Cattolica | - | X | X | - | - | - | - |
| Cagliari | X | X | X | - | - | X | X |
| Calabria (Rende-Cosenza) | X | X | X | X | X | X | X |
| Camerino | X | X | X | - | - | X | X |
| Catania | X | X | X | - | - | X | X |
| Chieti-Pescara | - | - | - | - | - | X | - |
| Ferrara | X | X | X | - | - | X | X |
| Firenze | X | X | X | X | - | X | X |
| Genova | X | X | X | - | X | X | X |
| Insubria (Como-Varese) | X | X | X | - | - | - | X |
| L'Aquila | X | X | X | - | - | - | X |
| Lecce | - | X | X | - | - | - | X |
| Messina | X | X | X | - | - | X | X |
| Milano Statale | X | X | X | - | - | X | X |
| Milano Bicocca | X | X | X | X | X | X | X |
| Milano Politecnico | - | - | X | - | - | - | - |
| Modena e Reggio E. | X | X | X | - | - | X | X |
| Molise | - | - | - | - | - | - | X |
| Napoli Federico II | X | X | X | - | X | X | X |

| | | | | | | | |
|--------------------|-----------|-----------|-----------|-----------|----------|-----------|-----------|
| Napoli 2 | - | X | X | - | - | - | X |
| Napoli Parthenope | - | - | - | X | - | - | X |
| Padova | X | X | X | X | X | X | X |
| Palermo | X | X | X | X | - | X | X |
| Parma | X | X | X | - | - | X | X |
| Pavia | X | X | X | - | - | X | X |
| Perugia | X | X | X | - | - | X | X |
| Piemonte Orientale | X | - | - | - | X | - | X |
| Pisa | X | X | X | - | - | X | X |
| Roma La Sapienza | X | X | X | X | - | X | X |
| Roma Tor Vergata | X | X | X | - | X | - | X |
| Roma 3 | - | X | X | - | - | X | X |
| Salerno | X | X | X | - | - | - | X |
| Sannio (Benevento) | - | - | - | X | - | X | X |
| Sassari | X | - | - | - | - | - | X |
| Siena | X | X | X | - | - | X | X |
| Torino | X | X | X | - | X | X | X |
| Torino Politecnico | - | - | X | - | - | - | - |
| Trento | - | X | X | - | - | - | X |
| Trieste | X | X | X | X | - | X | X |
| Udine | - | X | X | - | - | - | X |
| Urbino | - | - | - | - | - | X | X |
| Venezia Ca Foscari | X | - | - | - | - | - | - |
| Verona | - | X | X | - | - | - | X |
| Viterbo | - | - | - | - | - | - | X |
| Numero sedi | 32 | 35 | 39 | 11 | 9 | 29 | 44 |

Si evidenzia che Biologia e Biotecnologie sono i più numerosi, seguiti da Matematica (39), mentre Statistica e Scienza dei materiali sono il fanalino di coda.

In questo elaborato l'analisi è condotta sulle carriere universitarie dalla coorte degli studenti immatricolati negli atenei italiani nel 2011 in un corso di laurea triennale, includendo per semplicità gli studenti che sono alla loro prima carriera universitaria.

Particolare enfasi viene data alle carriere e al successo universitario degli immatricolati della coorte confrontando le lauree triennali e in particolare le LS. E' anche oggetto di interesse la ripartizione geografica. Nel nostro caso come già detto la variabile risposta sarà l'esito universitario a 5 anni dall'immatricolazione, l'obiettivo è valutare se l'esito universitario (Laurea SI/NO) è associato significativamente alle variabili di interesse di tipo individuale, per fare ciò sarà condotta un'analisi statistica servendosi di modelli di regressione logistica.

2.2 Costruzione del dataset

Per la nostra coorte ogni record rappresenta uno studente per il quale sono disponibili delle variabili relative alla sua carriera demografica, e alle caratteristiche socio-demografiche :

Distinguiamo quindi variabili rilevate all'ingresso:

- ANNO NASCITA
- CFU VALIDI ALL'IMMATRICOLAZIONE
- DIPLOMA
- ANNODIPLOMA
- CITTADINANZA
- GENERE
- VOTO DIPLOMA
- MACRO REGIONE RESIDENZA
- REGIONE RESIDENZA
- VOTO DI DIPLOMA

E variabili rilevate annualmente:

- ALTRA CLASSE
- ATENEO
- CICLO UNICO
- CARRIERA RECENTE
- CFU RICONOSCIUTI
- CFU RICONOSCIUTI CUMULATI
- CFU SOSTENUTI
- CFU SOSTENUTI CUMULATI
- DATA LAUREA
- EVENTO
- UNIVERSITA' TELEMATICA

- ANNO CORSO1
- TIPO LAUREA
- MACRO REGIONE
- MEDIA VOTO ANNO
- MOTIVO CHIUSURA CARRIERA
- NUM_LAUREE
- REGIONE
- STATALE
- VOTO LAUREA

Il dataset originale è stato ridotto in quanto conteneva informazioni relative anche a studenti che non fanno parte della popolazione di riferimento, avendo una precedente carriera universitaria o essendo iscritti a corsi di laurea magistrali a ciclo unico.

Occorre quindi escludere:

- Gli immatricolati in corsi di laurea a ciclo unico
- Gli studenti con una carriera pregressa con un numero di cfu maggiore di 10

L'esclusione dei primi è ovvia perché questo gruppo di studenti non permette una comparazione con le lauree triennali visto che il periodo di osservazione è 5 anni, mentre il secondo gruppo è stato escluso perché non è comparabile con gli studenti appena immatricolati. Abbiamo quindi assunto che studenti con un numero minore di 10 cfu (conseguiti in precedenti carriere) siano assimilabili a quelli di una matricola con 0 cfu. Queste operazioni comportano la perdita di 55585 osservazioni dal dataset che in origine ne aveva 279163.

2.3 Le variabili in studio

Il dataset contiene 223583 osservazioni, con 160 variabili. Occorre scegliere un sottoinsieme di variabili di interesse. Visti gli obiettivi, ovvero vengono scelte le variabili che, basandoci su studi precedenti, sono risultate associate al successo universitario:

- Genere
- Diploma
- Voto di diploma

Genere: maschio, femmina.

Diploma : Scientifico, Classico, Tecnico, Professionale, Altro liceo, Estero.

Liceo scientifico e classico sono considerati separati dagli altri licei perché quelli “storici” nati per primi mentre gli altri sono stati istituiti successivamente.

Voto di diploma: 60-70,70-80,80-90,90-100,100+. Quest’ultima variabile è stata divisa in 5 classi considerata la grande dimensione dei dati.

Considerati i nostri obiettivi altre variabili di interesse sono:

- **Regione di residenza e Regione di Immatricolazione**
- **Corso di studi**

Macroregione di residenza e Macroregione di Immatricolazione: per comodità le regioni di residenza e di immatricolazione sono state aggregate in macroregioni seguendo la classificazione ISTAT

Isole → (Sicilia Sardegna),

Sud → (Abruzzo, Molise, Campania, Puglia, Calabria, Basilicata)

Centro → (Toscana, Marche, Lazio, Umbria)

Nord est → (Veneto, Trentino Alto Adige, Friuli Venezia Giulia, Emilia Romagna)

Nord ovest → (Lombardia, Piemonte, Liguria, Valle d’Aosta)

Estero. Per facilità la Macroregione di Residenza e di Immatricolazione si indicheranno con M.Res e M. Imm **Corso di studi:** PLS, NO PLS, Biologia+ Biotecnologie. Le classi di laurea Biologia e Biotecnologie pur facendo parte del progetto PLS vengono analizzate separatamente dato il loro elevato numero di immatricolati.

Capitolo 3

3 Analisi preliminare dei dati

La variabile di maggior interesse per il nostro studio è sicuramente l’appartenenza al corso di laurea, vediamo quindi come si distribuiscono gli studenti nei 3 corsi di laurea di interesse.

TAB III distribuzione degli immatricolati per corso di laurea

| | Tot immatricolati | Frequenza relativa |
|-----------------|--------------------------|--------------------|
| PLS | 11080 | 0,049 |
| Bio+Biot | 11470 | 0,051 |
| NOPLS | 201028 | 0,900 |
| Totale | 223583 | 1 |

I corsi di laurea NO PLS ospitano la maggior parte degli studenti. Prima di studiare l'effetto che le variabili prese in esame hanno sull'esito universitario è opportuno osservarne le caratteristiche distribuzionali. Per le variabili: Genere, Diploma e Voto diploma calcoliamo le percentuali per colonna, mentre distinti i corsi di laurea di PLS (escludendo biologia) studiamo la composizione interna. In un secondo momento aggreghiamo i corsi PLS e li confrontiamo con NO PLS e BIO+BIOT

TAB IV – Distribuzione degli immatricolati 2011-2012 in Italia nei corsi PLS (valori assoluti e frequenze relative per colonna)

| VARIABILE | STAT | | MAT | | CHI | | GEO | | FIS | |
|---------------------|------------|-------------|-------------|----------|-------------|----------|-------------|----------|-------------|----------|
| | Assolute | Relative | Assolute | Relative | Assolute | Relative | Assolute | Relative | Assolute | Relative |
| <i>Genere:</i> | | | | | | | | | | |
| M | 502 | 0,54 | 1088 | 0,44 | 1669 | 0,48 | 821 | 0,63 | 1857 | 0,63 |
| F | 425 | 0,46 | 1361 | 0,56 | 1778 | 0,52 | 479 | 0,37 | 1100 | 0,37 |
| TOT | 927 | 1 | 2449 | 1 | 3447 | 1 | 1300 | 1 | 2957 | 1 |
| <i>Diploma:</i> | | | | | | | | | | |
| Altro liceo | 61 | 0,07 | 211 | 0,09 | 200 | 0,06 | 87 | 0,07 | 186 | 0,06 |
| Classico | 55 | 0,06 | 214 | 0,09 | 507 | 0,15 | 122 | 0,09 | 378 | 0,13 |
| Scientifico | 451 | 0,49 | 1626 | 0,66 | 1853 | 0,54 | 655 | 0,50 | 1881 | 0,64 |
| Professionale | 37 | 0,04 | 47 | 0,02 | 183 | 0,05 | 62 | 0,05 | 82 | 0,03 |
| Tecnico | 284 | 0,31 | 304 | 0,12 | 579 | 0,17 | 344 | 0,26 | 378 | 0,13 |
| Estero | 39 | 0,04 | 47 | 0,02 | 125 | 0,04 | 30 | 0,02 | 52 | 0,02 |
| TOT | 927 | 1 | 2449 | 1 | 3447 | 1 | 1300 | 1 | 2957 | 1 |
| <i>Voto diploma</i> | | | | | | | | | | |
| 60-69 | 241 | 0,26 | 278 | 0,11 | 758 | 0,22 | 505 | 0,39 | 488 | 0,17 |
| 70-79 | 256 | 0,28 | 474 | 0,19 | 1019 | 0,30 | 417 | 0,32 | 621 | 0,21 |
| 80-89 | 240 | 0,26 | 599 | 0,24 | 848 | 0,25 | 247 | 0,19 | 681 | 0,23 |
| 90-99 | 100 | 0,11 | 456 | 0,19 | 425 | 0,12 | 77 | 0,06 | 453 | 0,15 |
| >=100 | 43 | 0,05 | 574 | 0,23 | 296 | 0,09 | 40 | 0,03 | 652 | 0,22 |
| NA | 47 | 0,05 | 68 | 0,03 | 101 | 0,03 | 14 | 0,01 | 62 | 0,02 |
| TOT | 927 | 1,00 | 2449 | 1 | 3447 | 1 | 1300 | 1 | 2957 | 1 |

TAB V Distribuzione degli immatricolati 2011-2012 in Italia (valori assoluti e frequenze relative per colonna)

| VARIABILE | lauree triennali | | | | | |
|-----------------|------------------|----------|---------------|-------------|--------------|-------------|
| | PLS | | NO PLS | | BIOL+BIOT | |
| <i>Genere:</i> | Assolute | Relative | Assolute | Relative | Assolute | Relative |
| M | 5937 | 0,54 | 90177 | 0,45 | 3556 | 0,31 |
| F | 5143 | 0,46 | 110851 | 0,55 | 7914 | 0,69 |
| TOT | 11080 | 1 | 201028 | 1 | 11470 | 1 |
| <i>Diploma:</i> | | | | | | |
| Altro liceo | 745 | 0,07 | 32651 | 0,16 | 1045 | 0,09 |
| Classico | 1276 | 0,12 | 22864 | 0,11 | 2435 | 0,21 |
| Scientifico | 6466 | 0,58 | 72292 | 0,36 | 6485 | 0,57 |
| Professionale | 411 | 0,04 | 12222 | 0,06 | 432 | 0,04 |
| Tecnico | 1889 | 0,17 | 51891 | 0,26 | 880 | 0,08 |
| Estero | 293 | 0,03 | 9108 | 0,05 | 193 | 0,02 |
| TOT | 11080 | 1 | 201028 | 1 | 11470 | 1 |
| Voto diploma | | | | | | |
| 60-69 | 2270 | 0,20 | 53502 | 0,27 | 1982 | 0,17 |
| 70-79 | 2787 | 0,25 | 60738 | 0,30 | 3111 | 0,27 |
| 80-89 | 2615 | 0,24 | 45712 | 0,23 | 3107 | 0,27 |
| 90-99 | 1511 | 0,14 | 21025 | 0,10 | 1732 | 0,15 |
| >=100 | 1605 | 0,14 | 14002 | 0,07 | 1379 | 0,12 |
| NA | 292 | 0,03 | 6049 | 0,03 | 159 | 0,01 |
| TOT | 11080 | 1 | 201028 | 1,00 | 11470 | 1,00 |

Notiamo subito (Tab IV e V) che la distribuzione per genere differisce tra i corsi di laurea; infatti nei corsi PLS sono i maschi raggiungono il 54% del totale, nei corsi NO PLS si riducono al 43% e, infine, nei corsi di Biol+Biot sono solo il 31%. I dati evidenziano come il bacino di provenienza degli studenti sia per lo più quello del Liceo Scientifico, soprattutto nei Corsi di Laurea PLS (58%) e Biol+Biot (57%)

Questo dato è coerente con un percorso formativo già intrapreso con gli studi di scuola secondaria superiore. Nei corsi NO PLS una percentuale rilevante è dovuta agli studenti provenienti dall'istituto tecnico (23%), che probabilmente è da attribuirsi al fatto che i corsi NO PLS includono quelli economici di agraria.

In ultima analisi prendiamo in esame l'informazione proveniente dai voti di diploma, questa variabile potrebbe essere poco attendibile poiché la sua determinazione proviene da una valutazione soggettiva da parte dell'organo di commissione. Data la sua natura quindi un confronto tra individui con voti di maturità diversi non è pienamente informativo.

Nonostante i suoi limiti, questa variabile ha una forte associazione con l'esito universitario; banalmente, studenti che al liceo conseguono buone valutazioni mantengono generalmente la stessa linea all'università.

Analizzando la tabella notiamo che, il corso di laurea che mostra la maggiore percentuale di studenti che hanno preso un voto di diploma maggiore o uguale a 100 è quello PLS (14%) seguito da Biologia+ Biotecnologie (12%) e NO PLS(7%). Per quanto riguarda invece studenti con un voto compreso tra 90 e 99 si ha Biologia (15%), immediatamente dopo PLS (14%) e ancora una volta ultima NO PLS (10%).

Osserviamo quindi che il 28% degli studenti che si sono immatricolati in corsi PLS ha conseguito il diploma con un voto maggiore a 90.

Per le variabili M.Res e M.Imm è più opportuno riportare e commentare la distribuzione percentuale per riga, ovvero, stratificando per macro regione osserviamo come si distribuiscono gli studenti tra PLS, NO PLS e biologia in modo da analizzarne le differenze.

TAB VI- Distribuzione degli immatricolati 2011/2012 rispetto alle regioni di residenza e immatricolazione (valori assoluti e frequenze relative di riga)

| VARIABLE | lauree triennali | | | | | | TOT |
|----------------|------------------|----------|----------|----------|-----------|----------|-------|
| | PLS | | NO PLS | | BIOL+BIOT | | |
| | Assolute | Relative | Assolute | Relative | Assolute | Relative | |
| <i>M. Res:</i> | | | | | | | |
| Isole | 937 | 0,041 | 20688 | 0,898 | 1416 | 0,061 | 23041 |
| Sud | 2490 | 0,040 | 55499 | 0,894 | 4094 | 0,066 | 62083 |
| Centro | 2424 | 0,053 | 40832 | 0,888 | 2743 | 0,060 | 45999 |
| Nord Ovest | 3017 | 0,059 | 45805 | 0,902 | 1960 | 0,039 | 50782 |
| Nord Est | 2142 | 0,055 | 35373 | 0,914 | 1184 | 0,031 | 38699 |
| <i>M. Imm:</i> | | | | | | | |
| Isole | 822 | 0,044 | 16852 | 0,901 | 1025 | 0,055 | 18699 |
| Sud | 1935 | 0,038 | 45001 | 0,895 | 3331 | 0,066 | 50267 |
| Centro | 2601 | 0,048 | 48335 | 0,885 | 3660 | 0,067 | 54596 |
| Nord Ovest | 3067 | 0,055 | 50595 | 0,908 | 2057 | 0,037 | 55719 |
| Nord Est | 2655 | 0,060 | 40164 | 0,908 | 1397 | 0,032 | 44216 |

Notiamo che le preferenze degli studenti variano al variare della regione di residenza, infatti, la composizione è molto diversa per le nostre tre sottocategorie. In prima istanza si nota che la frequenza relativa di studenti che sceglie di immatricolarsi in corsi di laurea NO PLS rimane costante e compresa tra il 88% e il 91%. In tutta l'Italia la propensione a immatricolarsi in biologia, invece,

diminuisce progressivamente salendo lungo lo “stivale”; in particolare passa dal 6% delle Isole, Sud e Centro al crollo del 3,1% e 3,9 % che osserviamo sia al Nord est che al Nord ovest.

Al contrario di biologia, la propensione a immatricolarsi in corsi PLS aumenta se confrontiamo le regione del Centro e Sud (dove si registra tra il 3,8% e il 4,8%) con quelle del Nord Ovest e Nord est (5,5% e 6%).

Cosa ci suggeriscono questi dati? Appare evidente che, gli atenei universitari del Centro/Sud ha una bassa attrattività per quanto concerne l’ambito scientifico.

Consideriamo adesso la distribuzione degli immatricolati rispetto a M. Res e M. Imm. E’ evidente come il numero di studenti residenti nelle Isole e nel Sud sono 85124 (23041+ 62083) e gli immatricolati invece 68966 (18699+50267) denotando una “perdita” di circa 16000 unità che si distribuisce più o meno equamente al Centro, al Nord Ovest, Nord Est .

Calcoliamo adesso i singoli rapporti MRes / MImm per ogni cella della tab VII si ha

Tab.VII : Rapporto MRes / MImm per corso di laurea coorte 2011

| | PLS | NO PLS | BIOL+BIOT |
|------------------|------|--------|-----------|
| Isole | 1,14 | 1,23 | 1,38 |
| Sud | 1,29 | 1,23 | 1,23 |
| Centro | 0,93 | 0,84 | 0,75 |
| NordOvest | 0,98 | 0,91 | 0,95 |
| Nord Est | 0,81 | 0,88 | 0,85 |

Come ci aspettavamo si hanno rapporti maggiori all’unità per gli studenti delle Isole e del Sud, denotando un numero maggiore di residenti rispetto agli immatricolati e invece inferiori all’unità per studenti del Centro, del Nord Est, e Nord Ovest.

Un’ulteriore informazione sulla mobilità degli studenti che si muovono all’interno dell’Italia è data dalla tabella origine-destinazione, in cui l’origine è la macroregione di residenza e la destinazione la macroregione di immatricolazione, offrendo diversi spunti di riflessioni.

Tab VIII: Probabilità di transizione degli immatricolati per MRes (origine) e MImm (destinazione)

| Origine | | Isole | Sud | Centro | Nord Ovest | Nord Est | Tot |
|------------|--|-------|------|--------|------------|----------|-----|
| Isole | | 0,77 | 0,01 | 0,09 | 0,08 | 0,05 | 1 |
| Sud | | 0,01 | 0,79 | 0,12 | 0,04 | 0,04 | 1 |
| Centro | | 0 | 0,03 | 0,92 | 0,02 | 0,03 | 1 |
| Nord Ovest | | 0 | 0 | 0,02 | 0,93 | 0,05 | 1 |
| Nord Est | | 0 | 0 | 0,03 | 0,04 | 0,93 | 1 |

Analizzando la tabella, notiamo subito che le frequenze relative di studenti residenti nelle Isole e nel Sud che decidono di rimanere nella propria terra per proseguire gli studi universitari è molto bassa, ovvero del 76% e 77% per la nostra coorte mentre il Nord est e il Nord ovest hanno un percentuale di studenti che non si trasferiscono maggiori al 90%.

Tale mobilità coinvolge, in piccola parte, anche territori del Sud e Isole, sedi di grandi università che riescono ad attirare, studenti provenienti dalle provincie che sono contigue ad esse e che il più delle volte risultano sprovviste di sedi universitarie.

3.2 Analisi del successo

Analizziamo adesso i tassi di laurea a 5 anni dei corsi di laurea che compongono il PLS

TAB IX . Tasso di Successo/Insuccesso a 5 anni dall'immatricolazione coorte 2011/12

| Corso di laurea | Tasso di successo (laurea) | Tassi di Insuccesso (Censura e Abbandoni) |
|-------------------|----------------------------|---|
| STATISTICA | 54,2 | 45,8 |
| GEOLOGIA | 39,2 | 60,8 |
| FISICA | 48 | 52 |
| MATEMATICA | 47,6 | 52,4 |
| CHIMICA | 40,3 | 59,7 |
| BIOLOGIA | 40,2 | 59,8 |

Si evidenzia un'alta variabilità tra i tassi di laurea dei corsi di laurea PLS, notiamo infatti questi sono compresi in un intervallo piuttosto ampio che varia dal 39% del corso di Geologia al tasso più elevato è Statistica 54,2%.

Aggreghiamo ora i corsi PLS e confrontiamo i tassi di laurea e abbandono a 5 anni dall'immatricolazione degli studenti italiani della coorte del 2011-2012 per corso di laurea triennali espressi in percentuale

TAB X: tassi di Successo/Insuccesso a 5 anni dall'immatricolazione

| % | Successo (laurea) | | | | Insuccesso | Tot immatricolati |
|------------------|-------------------|--------|--------|------|------------|-------------------|
| | Terzo | Quarto | Quinto | Tot | | |
| PLS | 23,9 | 18 | 5,3 | 47,2 | 52,8 | 11080 |
| Bio+ Biot | 15,5 | 21,1 | 3,6 | 40,2 | 59,8 | 11470 |
| NOPLS | 27,8 | 16,1 | 3,8 | 47,7 | 52,3 | 205033 |
| Totale | 27 | 16,4 | 3,9 | 47,3 | 52,7 | 227583 |

La tabella offre diversi spunti di analisi, in primo luogo solo il 47,3% degli studenti italiani immatricolati nell'anno accademico 2011/2012 riesce a conseguire il titolo di laurea triennale entro 5 anni, questo comporta un insuccesso generale abbastanza preoccupante.

Notiamo, inoltre, che gli studenti immatricolati in corsi PLS e NO PLS presentano tassi di laurea rispettivamente uguali a 47,2% e 47,7% maggiori rispetto a biologia 40,2%, questa differenza si evidenzia particolarmente al terzo anno, infatti il tasso di studenti che si laureano in tempo nei corsi di biologia è molto più bassa rispetto agli altri due gruppi.

Quali sono le cause di quanto appena scritto? Sicuramente il basso tasso di laureati in biologia, in particolar modo a 3 anni dall'immatricolazione, è da attribuire in larga parte alla migrazione verso il corso di laurea di medicina e professioni sanitarie. Sappiamo infatti che le richieste di immatricolazione nel corso di laurea di medicina e professioni sanitarie sono in aumento (il corso di

laurea in medicina attrae i giovani laureati sia per lo status sociale sia per il reddito futuro). Per questo motivo molti studenti che non riescono ad entrare nei corsi appena citati, si iscrivono in corsi vicini che permettono loro di studiare materie utili non solo per il test di accesso dell'anno successivo ma anche per una eventuale convalida a medicina. Nella tabella XI si riportano in termini percentuali i passaggi degli studenti delle lauree scientifiche di base e medicina, professioni sanitarie, farmacia.

| aa 2012/13 | | | | | | | | | | |
|----------------|----------|----------|----------|--------|----------|---------|--------|----------------|-------|--------|
| aa 2011/12 | Medicina | Biologia | Mat+Stad | Fisica | Farmacia | Chimica | Altro | Prof.Sanitarie | NA | TOT |
| | 2° | 2° | 2° | 2° | 2° | 2° | 2° | 2° | 2° | 1° |
| Medicina | 98,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,8 | 0,2 | 0,6 | 5688 |
| Biologia | 13,6 | 51,9 | 0,1 | 0,1 | 1,2 | 0,2 | 11,9 | 11,1 | 9,7 | 11470 |
| Mat+Stad | 0,4 | 0,6 | 70,5 | 0,3 | 0,4 | 0,2 | 13,4 | 1,3 | 12,9 | 3376 |
| Fis+Geo | 0,6 | 1,0 | 0,3 | 69,5 | 0,3 | 0,4 | 11,5 | 2,8 | 13,5 | 4257 |
| Farmacia | 8,9 | 0,8 | 0,0 | 0,1 | 70,1 | 0,2 | 6,7 | 6,7 | 6,6 | 9592 |
| Chimica | 3,3 | 2,8 | 0,2 | 0,3 | 3,1 | 59,0 | 12,6 | 6,4 | 12,1 | 3447 |
| Altro | 0,3 | 0,3 | 0,1 | 0,1 | 0,2 | 0,1 | 82,0 | 1,1 | 15,9 | 215913 |
| Prof.sanitarie | 2,3 | 0,3 | 0,0 | 0,0 | 0,2 | 0,0 | 2,9 | 88,4 | 3,0 | 15892 |
| TOTALE | 9144 | 6900 | 2571 | 3131 | 7474 | 2239 | 180995 | 18669 | 37904 | 269635 |

TAB VIII - Distribuzione percentuale degli studenti per corso di studio dal primo anno di carriera(origine) rispetto al secondo (destinazione). (% per riga).

Solamente il 51% degli studenti che si erano immatricolati a biologia prosegue il suo corso di studi nello stesso corso di laurea, assistiamo ad una migrazione verso medicina ma che interessa e professioni sanitarie, gli approfondimenti su questo importante tema sono svolti nei rapporti di stage di Andrea Priulla "Analisi dei passaggi/trasferimenti nelle lauree scientifiche di base nelle coorti 2011 e 2014 degli studenti universitari italiani".

Capitolo 4

Modello di regressione logistica

Il modello di regressione logistica si applica quando si vuole studiare la relazione tra una variabile dipendente Y e un insieme di variabili esplicative. In un modello di regressione si esprime $E(Y|x)$ ovvero la speranza matematica della variabile dipendente Y condizionato ad un dato x , in funzione delle variabili esplicative x_i . Nel caso del modello di regressione logistica, questo valor medio condizionato corrisponde a $P(Y=1|x)$, cioè alla probabilità di possedere l'attributo in esame dato il vettore delle variabili indipendenti x_i .

La funzione che lega tale probabilità, alla funzione delle variabili indipendenti sarà indicata con $\pi(x)$. Il modello di regressione per Y è dunque: $Y = \pi(x) + \varepsilon$.

Un modello di regressione lineare sarebbe del tutto inappropriato a questo scopo, infatti una funzione lineare di x , essendo non limitata (né inferiormente, né superiormente), potrebbe dare luogo a valori stimati di $\pi(x)$ esterni all'intervallo $[0, 1]$, e quindi privi di senso.

Per descrivere la relazione di dipendenza della probabilità $\pi(x) = P(Y = 1 | x)$ dai valori di $x = (x_1, x_2, \dots, x_p)$ Si può usare la distribuzione logistica:

$$\pi(\mathbf{x}) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}} = \frac{e^{\beta_0 + \sum_{j=1}^p \beta_j x_j}}{1 + e^{\beta_0 + \sum_{j=1}^p \beta_j x_j}} \quad [1]$$

Si consideri, ora, la seguente funzione di $\pi(x)$, detta logit,

$$\text{logit}(\pi(\mathbf{x})) = \ln \left[\frac{\pi(\mathbf{x})}{1 - \pi(\mathbf{x})} \right] \quad [2]$$

Che è il logaritmo naturale del rapporto della probabilità condizionata di possedere l'attributo rispetto alla probabilità condizionata di non possederlo. Il rapporto fra probabilità associate ad una dicotomia, cioè fra probabilità complementari, è detto odd .

E' possibile dimostrare che

$$\text{logit}(\pi(\mathbf{x})) = \beta_0 + \sum_{j=1}^p \beta_j x_j$$

Quindi, mentre $\pi(\mathbf{x})$ è funzione non lineare delle variabili x_1, x_2, \dots, x_p , il logaritmo dell'odds, detto logit, è funzione lineare delle stesse.

Dobbiamo stimare la probabilità che un soggetto (o un'altra unità statistica) appartenga ad un gruppo ($Y=0$) oppure ad un altro ($Y=1$), possiamo cercare la combinazione lineare di variabili esplicative che crea la maggiore discriminazione fra le unità del primo e del secondo gruppo, a seconda delle variabili esplicative (o fattori di rischio) inclusi.

I problemi della selezione delle variabili esplicative e del numero di modalità delle stesse, così come quello della scelta del modello, sono condizionati da due fattori tra loro contrapposti: da un lato si vuole un modello che adatti i dati nella maniera migliore possibile e da un altro lato si vuole un modello parsimonioso, ovvero un modello in cui il numero dei parametri sia sufficientemente contenuto per una lettura più agevole del fenomeno.

In generale per i modelli log lineari si vuole un modello che sia più articolato del modello di indipendenza, ma che, nello stesso, non sia saturo, ovvero non contempli tutte le interazioni possibili ma solo quelle utili alla spiegazione della variabile risposta. Questo genere di problema, visto il numero elevato di variabili è presente in questo elaborato.

Se il logit è negativo (positivo) significa che la categoria i -esima è proporzionalmente meno (più) rappresentata rispetto alla categorie I -esima (ovvero l'ultima). Gli antilogaritmi dei logit indicano l'effetto netto corrispondente alla modalità i -esima nei confronti della modalità I -esima.

La denominazione di effetto netto sta ad indicare l'effetto dato da quella componente quando tutte le altre componenti sono tenute costanti. In un modello logit l'effetto netto principale è associato solo alla variabile risposta e misura il livello medio (o baseline) di preferenza della i -esima categoria rispetto alla I -esima, mentre gli effetti netti del secondo ordine, associati all'interazione della coppia i_j (con $i=1, \dots, I-1$; $j=1, \dots, J-1$), misurano la preferenza della i -esima categoria rispetto alla I -esima, tenendo fisso j , e della j -esima categoria rispetto alla J -esima, tenendo fisso i . Quando si hanno interazioni superiori al secondo l'interpretazione è analoga.

Di seguito si riporta la tabella che mostra il modello applicato con i relativi :

- Coefficienti di regressione
- Standard Error
- Valore del test T
- P-valore che indica il livello di significatività

4.2 Modello logit - Analisi dei risultati

Il modello ha come variabile di risposta Laurea a 5 anni dall'immatricolazione si/no

Le variabili esplicative sono: Diploma, Genere, Corso di laurea, M. IMM, Voto di diploma. E' sembrato opportuno costruire una variabile MOB che indichi la mobilità : se lo studente entro il terzo anno cambia corso di laurea: Mover se lo studente prosegue il suo percorso di studi all'interno dello stesso corso : Stayer),

La baseline è rappresentata da

- Diploma; Professionale + Estero
- Genere; femmina
- Corso di laurea ;NO PLS
- MOB; Mover
- R.Imm; Nord- Ovest
- Voto diploma ; ≥ 100

Le stime dei parametri vengono riportate insieme ai loro errori standard anche se questi ultimi non hanno un significato inferenziale in quanto abbiamo dati di popolazione, tuttavia ci sembra opportuno riportarli perchè sono in qualche modo una misura dell'intensità del parametro ad essi corrispondente. Tab XII: Stima dei parametri modello logit. Var risposta laurea entro 5 anni, lauree triennali Italia coorte 2011/2012

| Coefficients: | Stima | Std.Error | Z value | Pr(> z) |
|-----------------------------------|-------|-----------|---------|----------|
| <i>Intercetta</i> | 1,06 | 0,04 | 24,05 | < 2e-16 |
| Scuola: | | | | |
| <i>ALTRO LICEO</i> | 0,85 | 0,02 | 40,52 | < 2e-16 |
| <i>CLASSICO</i> | 1,35 | 0,02 | 59,98 | < 2e-16 |
| <i>SCIENTIFICO</i> | 1,27 | 0,02 | 66,12 | < 2e-16 |
| <i>TECNICO</i> | 0,36 | 0,02 | 18,22 | < 2e-16 |
| Genere : | | | | |
| <i>Maschio</i> | -0,43 | 0,01 | -41,98 | < 2e-16 |
| Voto diploma: | | | | |
| <i>60-69</i> | -1,83 | 0,02 | -91,22 | < 2e-16 |
| <i>70-79</i> | -1,25 | 0,02 | -64,09 | < 2e-16 |
| <i>80-89</i> | -0,75 | 0,02 | -37,99 | < 2e-16 |
| <i>90-99</i> | -0,40 | 0,02 | -18,12 | < 2e-16 |
| Corso di laurea: | | | | |
| <i>Bio+Biot</i> | -0,11 | 0,09 | -1,15 | 0,25 |
| <i>PLS</i> | -0,81 | 0,10 | -8,18 | 2.79e-16 |
| Mob: | | | | |
| <i>Stayer</i> | -0,47 | 0,04 | -13,33 | < 2e-16 |
| R.IMM : | | | | |
| <i>NORD EST</i> | 0,19 | 0,01 | 13,62 | < 2e-16 |
| <i>CENTRO</i> | -0,34 | 0,01 | -26,47 | < 2e-16 |
| <i>SUD</i> | -0,75 | 0,01 | -55,26 | < 2e-16 |
| <i>ISOLE</i> | -0,87 | 0,02 | -46,55 | < 2e-16 |
| Genere * Corso di laurea: | | | | |
| <i>Maschio*Bio+Biot</i> | -0,02 | 0,05 | -0,49 | 0,63 |
| <i>Maschio*PLS</i> | 0,37 | 0,04 | 8,39 | < 2e-16 |
| Diploma * Corso di laurea: | | | | |
| <i>ALTRO LICEO*Bio+Biot</i> | -0,71 | 0,12 | -6,09 | 1.16e-09 |
| <i>CLASSICO*Bio+Biot</i> | -1,22 | 0,10 | -11,68 | < 2e-16 |
| <i>SCIENTIFICO*Bio+Biot</i> | -0,70 | 0,10 | -7,25 | 4.15e-13 |
| <i>TECNICO:*Bio+Biot</i> | -0,25 | 0,12 | -2,04 | 0,04 |
| <i>ALTRO LICEO*PLS</i> | -0,25 | 0,13 | -2,01 | 0,04 |
| <i>CLASSICO*PLS</i> | -0,19 | 0,11 | -1,70 | 0,09 |
| <i>SCIENTIFICO*PLS</i> | 0,16 | 0,10 | 1,64 | 0,10 |
| <i>TECNICO*PLS</i> | 0,14 | 0,11 | 1,27 | 0,20 |

Dall'analisi della tabella sopra riportata emerge:

- Prendendo in esame la variabile Diploma notiamo che gli esiti migliori li ottengono studenti che provengono dal liceo in particolare dal liceo Classico $\beta_2=1,35$ e dal liceo Scientifico $\beta_3=1,27$, infatti i coefficienti che riguardano gli effetti netti risultano essere i maggiori in valore assoluto e positivi, i risultati peggiori si registrano per studenti provenienti dall'Istituto Professionale /Estero utilizzati come baseline.
- La variabile macro-regione di immatricolazione è significativa nella spiegazione della risposta.
Nello specifico, ponendo come baseline Nord Ovest tutti i coefficienti di regressione associate alle macro regioni sono negativi eccetto per Nord Est che presenta un coefficiente con valore $\beta_{13}=0,19$, essa quindi è la modalità che incrementa maggiormente la probabilità di successo.
Gli studenti che conseguono i peggiori risultati universitari sono quelli che si immatricolano alle Isole e al Sud, seguiti da quelli del Centro.
- Il voto di diploma come accennato precedentemente, è associato positivamente con l'esito universitario. Studenti che conseguono la maturità con un voto maggiore o uguale a 100 hanno una probabilità maggiore di laurearsi entro i 5 anni; posta come baseline la modalità di voto "> 100" i coefficienti di regressione associati alle modalità rappresentati le altre fasce di voto risultano tutti negativi e significativi.
- Le studentesse conseguono risultati migliori dei colleghi maschi, queste si laureano con una probabilità maggiore entro i 5 anni dall'immatricolazione. Il coefficiente per i maschi è $\beta_5=-0,43$

- Dai coefficienti degli effetti del primo emerge che studenti che mostrano una propensione alla laurea maggiore sono quelli immatricolati in corsi NO PLS. Per quanto concerne i coefficienti che riguardano l'effetto netto vediamo che (ricordando che la baseline è NO PLS) per PLS $\beta_{11}=-0,81$ mentre per Biologia $\beta_{12}=-0,11$, entrambi significativi e negativi.
- L'analisi delle interazioni tra la scuola di diploma e il corso di laurea scelto mostra qualcosa di interessante perché fa cambiare la probabilità di laurearsi. Analizzando l'interazione Classico*Biologia vediamo che il coefficiente per questa interazione risulta essere $\beta_{20}=-1,22$ denotando una riduzione del tasso di laurea, anche per gli studenti diploma al liceo scientifico si verifica una riduzione del tasso di laurea ma inferiore a quanto visto per il classico, $\beta_{21}=-0,70$.

Analizzando i coefficienti per l'interazione Scientifico*PLS notiamo, invece, che il coefficiente di questa interazione è $\beta_{25}=0,16$ mostrando un incremento della probabilità di laurea, si registra ancora una volta una riduzione della probabilità di successo per studenti del liceo classico che si immatricolano in corsi PLS $\beta_{24}=-0,19$.

Questo è un risultato che ci aspettavamo, è naturale infatti che al liceo scientifico si riceva una preparazione migliore in discipline come matematica, fisica etc.. che sono alla base dei corsi PLS. Anche della variabile genere sono state studiate le interazioni con il corso di laurea. E' interessante notare come i Maschi iscritti ai corsi PLS (senza biologia) hanno il medesimo successo rispetto alle Femmine, infatti l'effetto principale $-0,43$ viene quasi totalmente compensato dall'interazione Maschio*PLS che è pari a $+0,37$ il profilo degli studenti maschi immatricolati in biologia, invece, mostra addirittura una riduzione del tasso di laurea rispetto alle femmine, il coefficiente mostra assume un valore di $\beta_{17}=-0,02$.

Conclusioni

Questo lavoro è stato molto interessante per due punti di vista: il primo è dovuto al fatto che ci siamo imbattuti per la prima volta in un big data set ed il secondo perché abbiamo lavorato con dati di tipo longitudinale. Studi con dati longitudinali e individuali sono molto utilizzati in diversi ambiti; dall'economico al sociale, poiché, nonostante comportino dei costi maggiori rispetto a dati cross-section permettono forti vantaggi; non ultimo quello di valutare gli effetti delle variabili esplicative sulla risposta al netto delle altre variabili.

Lo scopo di questo rapporto di stage è stato quello di descrivere, attraverso analisi di tabelle di contingenza prima e, successivamente, analizzare, attraverso un modello di regressione logistica, le carriere degli studenti universitari della coorte 2011-12 immatricolati in tutti gli Atenei d' Italia rispetto alle caratteristiche individuali degli studenti, con particolare attenzione alle lauree scientifiche di base. L'analisi è stata circoscritta a corsi di laurea triennali in modo da rendere confrontabili i successi/insuccessi degli studenti, e il periodo di osservazione è stato di 5 anni accademici, ovvero la censura è fissata a 5 anni dall'immatricolazione.

La probabilità di laurea risulta più elevata per gli studenti:

- Provenienti da un liceo classico o scientifico
- Femmine
- Immatricolati in corsi NO PLS (Tassi di laurea pari al 47,7%)
- Mover
- Immatricolati in un ateneo del Nord-Est
- Studenti diplomati con un voto di maturità maggiore o uguale 100

Ancora una volta si evidenziano le differenze tra Centro-Nord e Sud e Isole, e il divario delle ultime si manifesta anche nei tassi di laurea. Il Sud e le Isole confermano tragicamente di essere terra di emigrazione e questo genera un circolo vizioso rispetto al recupero del proprio deficit di sviluppo economico, i diplomati meridionali che frequentano l'Università nel Centro-Nord costituiscono un sostanziale impoverimento del capitale sociale ed economico dei loro territori di origine.

Le LS migrano in percentuale maggiore rispetto ai corsi NO PLS.

La migrazione è un fenomeno altamente selettivo e tende ad interessare gli individui più intraprendenti o, come in questo caso, quelli dotati di risorse maggiori (vale a dire provengono da famiglie che possono mantenere gli studi dei propri figli). Ormai da anni si assiste ad un progressivo impoverimento culturale e economico delle regioni del Sud e questa emigrazione è un

ostacolo allo sviluppo del Sud e alla stessa sopravvivenza delle università del Sud. Infatti la probabilità di laurearsi/migrare verso università del Centro-Nord è maggiore per chi possiede un diploma di maturità liceale o scientifica.

Urgono interventi mirati per innalzare la qualità dell'istruzione negli atenei del Mezzogiorno poiché come sappiamo l'istruzione terziaria ha ricadute importantissime sullo sviluppo. Inoltre favorendo lo sviluppo si contrasta il fenomeno della migrazione.

Ringraziamenti

Si ringrazia il dott. Marco Enea dell'Università di Palermo per l'aiuto sull'analisi dei dati.

Bibliografia

Attanasio M. Analisi della selezione universitaria: tre generazioni di immatricolati nell'Ateneo di Palermo. Tesi Dottorato in Scienze Statistiche, Padova, 1991.

S. Fasola Laurea o abbandono: il buon giorno si vede dal mattino? Tesi di laurea 2009/2010

A. Priulla "Analisi dei passaggi/trasferimenti dalle lauree scientifiche di base degli immatricolati delle università italiane anni 2011/12 2014/15. 2016/2017

Anvur; Rapporto biennale sullo stato del sistema universitario e della ricerca 2016

Ministero dell'Istruzione, dell'Università e della Ricerca; LINEE GUIDA Piano Nazionale Lauree Scientifiche 2015

Ministero dell'Istruzione, dell'Università e della Ricerca; Protocollo di collaborazione per lo Studio della Mobilità Studentesca Universitaria 2015

Attanasio M, Boscaino G, Capursi V, Plaia A (2011). Indicators and measures for the assessment of university students' careers. In: Proceedings CLADAG 2011. ISBN: 978-88-96764-22-0

Capursi V., Ghellini G. (Eds), (2008), DOTTOR DIVAGO Discernere, Valutare e Governare la Nuova Università, F. Angeli, Milano.

Enea M., Capursi V., Plaia A., (2014) Modelling students' mobility in Italy: an analysis of the determinants by combining individual and aggregate data, submitted

Giambona F., Porcu M., Sulis I. Measuring university attractiveness. An analysis of students' mobility between competing provinces by pair comparisons, submitted

Sitografia

[1] <http://www.progettolaureescientifiche.eu/>