



Università degli Studi di Palermo

UNIVERSITÀ DEGLI STUDI DI PALERMO

Facoltà di Economia

Corso di Laurea Specialistica in
Scienze Statistiche Sociali e Sanitari

**Le misure di importanza relativa e la
valutazione della 'qualità della didattica'
nell'opinione degli studenti**

Tesi di Laurea di:

Clara Romano

Relatore:

Vincenza Capursi

Anno Accademico 2007-2008

INDICE

Introduzione	VII
1. Qualità dei servizi e soddisfazione dell'utente	1
1.1 Definizione di qualità	1
1.2 Le dimensioni della qualità	4
1.3 La Customer Satisfaction	6
1.4 Le dimensioni della Customer Satisfaction	7
2. La valutazione della didattica nell'opinione degli studenti	9
2.1 La valutazione come attività scientifica	9
2.2 La valutazione nelle istituzioni	12
2.3 La valutazione istituzionale del sistema universitario	15
2.4 La valutazione della didattica nell'opinione degli studenti	16
2.5 L'indagine presso l'Ateneo di Palermo	18
2.6 Il questionario	19
3. Analisi preliminari	22
3.1 Gli insegnamenti valutati	22
3.2 Le risposte "positive" e i dati mancanti	23
3.3 Gli item eliminati	25
4. Analisi esplorative	27
4.1 Premessa	27
4.2 Lo scaling multidimensionale	28
4.2.1 Lo scaling multidimensionale non metrico	30
4.2.1.1 La misura di prossimità: il τ_b di Kendall	30

4.2.1.2	L'algoritmo di Kruskal	32
4.2.2	Interpretazioni delle dimensioni	35
4.2.3	Applicazione dello scaling multidimensionale non metrico agli item del questionario	36
4.2.4	I risultati	37
4.2.5	L'interpretazione della configurazione dei punti	42
4.3	L'analisi fattoriale	48
4.3.1	L'analisi fattoriale applicata ai nostri dati	49
4.3.2	Le matrici fattoriali	50
5	Le misure di importanza relativa	64
5.1	Introduzione.	64
5.2	La scomposizione di R^2 nelle metriche di importanza relativa.	65
5.3	La metrica <i>PMVD</i>	67
5.4	I pesi ottenuti con il metodo <i>PMVD</i>	69
	Alcune considerazioni conclusive	77
	Ringraziamenti	80
A	I paradigmi generatori dei dati	81
A.1	I paradigmi generatori dei dati	81
A.2	I tipi di indagine	82
A.2.1	Gli studi osservazionali	84
A.2.2	Le indagini campionarie probabilistiche	87
A.2.3	Gli studi sperimentali	88
B	Le metriche di importanza relativa	90
B.1	I diversi approcci nella storia	90
B.2	Le metriche fornite da <i>relaimpo</i>	91

C	Il questionario per la valutazione della didattica da parte degli studenti dell'Università degli Studi di Palermo	96
D	Scaling multidimensionale non metrico: configurazione degli item nelle diverse facoltà	99
	Bibliografia	100

ELENCO TABELLE

3.1	<i>Tassi di copertura degli insegnamenti rilevati</i>	22
3.2	<i>Percentuale risposte più sì che no e decisamente sì</i>	24
3.3	<i>Percentuale dati mancanti per l'item F4</i>	25
4.1	<i>Scaling multidimensionale non metrico: la bontà di adattamento</i> . . .	37
4.2	<i>I ranghi degli item nelle 11 facoltà</i>	44
4.3	<i>Factor pattern ruotato: facoltà 2</i>	52
4.4	<i>Factor pattern ruotato: facoltà 3</i>	53
4.5	<i>Factor pattern ruotato: facoltà 4</i>	54
4.6	<i>Factor pattern ruotato: facoltà 5</i>	55
4.7	<i>Factor pattern ruotato: facoltà 6</i>	56
4.8	<i>Factor pattern ruotato: facoltà 7</i>	57
4.9	<i>Factor pattern ruotato: facoltà 8</i>	58
4.10	<i>Factor pattern ruotato: facoltà 9</i>	59
4.11	<i>Factor pattern ruotato: facoltà 10</i>	60
4.12	<i>Factor pattern ruotato: facoltà 11</i>	61
4.13	<i>Factor pattern ruotato: facoltà 12</i>	62
5.1	<i>I modelli regressivi</i>	70
5.2	<i>Stime con il metodo dei minimi quadrati dei parametri dei modelli regressivi e dei pesi PMVD</i>	70
5.3	<i>I pesi 'significativi'</i>	75

ELENCO FIGURE

<i>1.1 Il ciclo di Deming</i>	6
<i>3.1 Percentuale dati mancanti per gli item del questionario</i>	24
<i>4.1 Diagramma di Shepard per la facoltà 5</i>	39
<i>4.3 Scaling multidimensionale non metrico: Stress al variare del numero di dimensioni</i>	40
<i>4.4 Scaling multidimensionale non metrico: configurazione degli item nella facoltà 2 con e senza l'item F4</i>	47
<i>5.1 Due modelli di regressione a confronto</i>	67
<i>5.2 I pesi PMVD per le 11 facoltà</i>	76
<i>A.1 I tipi di indagine</i>	82
<i>A.2 I disegni statistici: punti di forza e di debolezza</i>	89

INTRODUZIONE

La valutazione della didattica assume un ruolo importante all'interno delle università italiane dal 1993 quando la legge 537/93, collegata alla finanziaria dell'anno medesimo, istituisce in Italia un Sistema di valutazione universitario nazionale e locale che, tra gli altri, ha il compito di verificare la produttività della ricerca e della didattica. In applicazione di tale legge le università hanno costituito il proprio Nucleo di Valutazione e hanno cominciato ad operare nell'ambito del vasto e complesso insieme di attività di valutazione. Tutto questo ha contribuito alla raccolta di materiali concernenti diversi aspetti: la valutazione della ricerca universitaria, la valutazione della didattica da parte degli studenti, l'analisi della condizione studentesca da parte degli Atenei.

La presenza di sistemi di valutazione ha incoraggiato, se non imposto, la formalizzazione di "procedure per la qualità" che si basano su un'attività interna ed una esterna. La prima si riferisce al processo continuo di mantenimento e miglioramento della qualità, la seconda alla periodica valutazione della qualità da parte di una organizzazione indipendente. Una valutazione significativa, interna o esterna, richiede l'acquisizione dei punti di vista di chi partecipa in modo importante al processo formativo. Quindi non solo le opinioni attribuibili ai professionisti della formazione (docenti a vari livelli e ruoli) ma anche di chi fa dello studio il proprio lavoro (studenti). In tale contesto si colloca la costruzione del questionario comune per la raccolta delle opinioni dei discenti. Attraverso tale questionario, somministrato in forma anonima durante le lezioni, si intendono ottenere informazioni circa gli elementi di forza e/o di debolezza che possono caratterizzare il servizio 'erogazione della didattica'.

In questo scenario si inserisce il presente lavoro che ha lo scopo di misurare il concetto di qualità della didattica nell'opinione degli studenti all'Università di Palermo. La misurazione di un concetto così complesso è stata oggetto di lunghi dibattiti che hanno messo in evidenza aspetti importanti da

tenere in considerazione in qualsiasi analisi riguardante la valutazione della qualità della didattica. Inoltre, la qualità di un corso universitario è considerata una variabile latente che non può essere direttamente osservabile (Capursi, Librizzi 2008). Così, allo scopo di misurare tale concetto, è necessario fare riferimento ad un certo numero di dimensioni, che sono misurate mediante un insieme di variabili indirette o una batteria di item. L'obiettivo è valutare la qualità della didattica ottenendo un'approssimazione della vera misura mediante misurazioni indirette fornite dai giudizi degli studenti. Ovviamente, è necessario sottolineare che il risultato dipende fortemente da fattori soggettivi. Ogni studente è influenzato dai propri bisogni e attese che a loro volta dipendono dai differenti ambienti culturali, dalle differenti condizioni economico-sociali. Indipendentemente dai risultati di carattere metodologico, non bisogna mai dimenticare che il punto di vista dal quale ha origine la valutazione è l'opinione degli studenti. Quindi, non si può pretendere di giungere ad una misura della qualità della didattica basata su un sistema di valori condivisi a causa della presenza dell'eterogeneità dei giudizi degli studenti: la risposta dello studente dipende sia dall'effettiva qualità del servizio sia dalla qualità che egli percepisce.

I dati del presente lavoro sono stati forniti dal Centro Informativo dell'Ateneo di Palermo (C. IN. A. P.) che ha il compito di coordinare l'indagine sulla rilevazione dell'opinione degli studenti sulla didattica; inoltre i dati a disposizione si riferiscono all'anno accademico 2006/2007.

La tesi è complementare a quella di Salvatore Marcantonio che ha come obiettivo un'analisi di indicatori semplici di item del questionario per ogni facoltà; sulla base di tali indicatori si attribuiranno dei pesi di importanza relativa allo scopo di misurare l'importanza che gli item hanno nello spiegare la soddisfazione degli studenti.

Il lavoro è articolato nel seguente modo:

- nel capitolo 1 sono descritti il concetto di qualità e di customer satisfaction;
- il capitolo 2 introduce il concetto di valutazione dei servizi, con particolare riferimento alla valutazione della didattica intesa come attività di customer

- satisfaction, con il fine di verificare le attività di erogazione del servizio didattico che soddisfano le esigenze dei clienti (studenti);
- il terzo capitolo, oltre a presentare alcune analisi preliminari condotte sui dati originari (tasso di copertura degli insegnamenti rilevati, numero di valori mancanti, percentuale di risposte positive), esamina e discute i motivi che hanno condotto all'eliminazione di alcuni item;
 - nel capitolo 4, si effettuano due analisi di tipo esplorativo: lo scaling multidimensionale non metrico che, sulla base dei giudizi espressi dagli studenti, permette di individuare quegli item che sono responsabili della soddisfazione dello studente e conseguentemente della valutazione della didattica; l'analisi fattoriale che, condotta su indicatori semplici relativi agli item, consente di evidenziare la correlazione tra gli item e le dimensioni del concetto che si vuole misurare;
 - nel capitolo 5, a partire dai risultati dell'analisi fattoriale si costruiscono modelli di regressione che presentano come variabili esplicative gli indicatori semplici degli item considerati, e come variabile risposta l'indicatore relativo all'item della soddisfazione C2. A tali modelli si applica il metodo *PMVD* ottenendo i pesi di importanza relativa degli item (Capursi, D'Agata, Librizzi, 2008);
 - si riportano alcune considerazioni finali sui principali risultati ottenuti;
 - infine, si riporta l'appendice divisa in 4 parti: la prima, più descrittiva, fa una rassegna dei principali paradigmi generatori dei dati; la seconda, più metodologica, descrive le principali metriche di importanza relativa esistenti in letteratura, con particolare riferimento a quelle che il pacchetto *relaimpo* del software *R* restituisce; nella terza si riportano le configurazioni degli item per le diverse facoltà risultanti dall'applicazione dello scaling multidimensionale; nell'ultima si riporta il questionario per la valutazione della didattica all'Università degli Studi di Palermo per l'anno accademico 2006/2007.

CAPITOLO 1

Qualità dei servizi e soddisfazione dell'utente

In questo primo capitolo cerchiamo di 'circoscrivere' il concetto che si intende misurare senza la pretesa di definire in maniera esaustiva e incontrovertibile cosa si possa intendere per qualità della didattica. Questa fase è molto importante in quanto da essa dipende una corretta elaborazione delle analisi che seguiranno e soprattutto un'adeguata interpretazione dei risultati.

Il concetto che si intende misurare è quello che viene definito "*qualità della didattica nell'opinione degli studenti*". I termini contenuti all'interno di questa espressione che di seguito sono ampiamente discussi sono *qualità* e *opinione degli studenti*. In particolare quando si parla dell'opinione degli studenti si fa riferimento alla soddisfazione che questi esprimono attraverso le opinioni e quindi in generale si fa riferimento al concetto di customer satisfaction.

1.1 Definizione di qualità

Sono ormai molteplici e sempre più pressanti gli stimoli, nei confronti della Pubblica Amministrazione, all'introduzione di meccanismi che siano in grado di gestire la qualità di beni e servizi prodotti sia dal settore pubblico sia dal settore privato. In particolare esistono due categorie di beni e servizi. Fanno parte della prima categoria i cosiddetti *search goods*: sono beni e servizi per i quali il fruitore è in grado di individuare le caratteristiche qualitative prima di consumare o fruire di essi, poiché la qualità è riferita all'output ed è identica per tutti gli utenti. Gli *experience goods* sono la seconda tipologia di beni e servizi per i quali non è possibile stabilire a priori la qualità. Fra questi rivestono una particolare importanza i Servizi alla Persona di Pubblica Utilità (SPPU). Quest'ultimi sono considerati *relazionali*, ovvero necessitano, per essere prodotti, della

collaborazione tra chi eroga il servizio e chi usufruisce dello stesso. In questo caso il risultato finale del processo di erogazione del servizio non può essere valutato in termini di output, bensì in termini di outcome, cioè in termini di valore aggiunto che la fruizione del servizio produce sul 'benessere' dell'utente. Inoltre i SPPU presentano un'alta componente di *intangibilità*, poiché si tratta di prestazioni non materiali difficilmente misurabili, e sono anche caratterizzati dalla *negoziabilità*, intesa come margine di flessibilità per l'erogatore e i fruitori del servizio nel realizzare la transazione. I SPPU devono essere valutati in più tempi, in quanto la loro qualità deve essere misurata considerando gli esiti di breve, medio e lungo periodo.

La didattica e i servizi erogati dal Sistema Università rientrano nei SPPU e presentano quindi le caratteristiche sopra menzionate e sono fortemente caratterizzati dalla relazione che in aula si determina fra docente e discente.

La Pubblica Amministrazione, così come qualsiasi altro settore, deve orientare i propri meccanismi al controllo della qualità di questi particolari beni e servizi ma anche inserirsi in un processo di miglioramento continuo della qualità delle prestazioni fornite.

Cercheremo adesso di fornire alcune definizioni sul concetto di qualità ed in particolare di qualità percepita dagli studenti e orientata ai loro bisogni.

Di definizioni del concetto di qualità ne esistono moltissime che si prestano anche ad usi ed interpretazioni diverse. Reeves e Bednar (1994) affermano che la ricerca di una definizione universale di qualità non ha condotto a risultati apprezzabili. Secondo Quartapelle la qualità è un valore assoluto, assimilabile all'assenza di difetti e per misurarla è sufficiente confrontare il prodotto con il campione di riferimento. Per Crosby (1994) la qualità è essere conforme alle richieste, mentre per altri autori un servizio (prodotto) è di qualità se non ha difetti (Tagucci 1989). Per Juran la qualità è intesa come "idoneità all'uso" per il quale il servizio è stato costruito.

Nonostante non vi sia una definizione universalmente condivisa, il concetto di qualità sembra sempre più pregnante sia nell'ambito industriale sia in quello delle aziende di servizio.

Per comprenderne il significato può essere utile descrivere come è cambiato nel tempo il concetto di qualità.

In passato il concetto di qualità era associato ad un contenuto di lusso, di non strettamente necessario, di estetica più che di sostanza. In un secondo tempo l'enfasi si è spostata sull'idea di "rispetto delle specifiche", cioè di rispondenza di un particolare prodotto o servizio alle condizioni e ai requisiti previsti.

Secondo quest'approccio la qualità di un prodotto è più elevata nella misura in cui esso risponde alle caratteristiche progettate e alle specifiche dell'utilizzo per il quale è stato costruito.

Successivamente il concetto di qualità si estende e si rivolge alla piena soddisfazione del cliente, si profila in questo modo un passaggio da una concezione oggettiva ad una soggettiva e il passaggio da una valutazione della qualità interna ad un'esterna. De Bono (1999) definisce la qualità come un contenitore che raccoglie nel suo spazio, spesso caricato emotivamente, tutte le possibili interpretazioni. Questa definizione riflette l'aspetto estremamente soggettivo del concetto di qualità, così da rendere il tentativo di una misurazione oggettiva un percorso impossibile, o comunque particolarmente difficoltoso. Per Murto (1999) la qualità è la comparabilità dei servizi sociali offerti con i bisogni dei clienti e che può essere misurata chiedendo agli stessi clienti la loro opinione.

Quindi, pur non trovando un'unica definizione della qualità in letteratura, si deduce un graduale spostamento da una logica centrata sui processi produttivi interni ad una logica di mercato che si focalizza sui bisogni del cliente. Partendo da questi ultimi si cerca di incontrare e di superare le aspettative degli utenti.

La definizione generale di qualità attualmente accettata in tutto il mondo è esplicitata nella norma ISO 8402: "La qualità è l'insieme delle caratteristiche di un'entità che ne determinano la capacità di soddisfare esigenze espresse e implicite".

1.2 Le dimensioni della qualità

Lo sviluppo di un Sistema di Qualità all'interno della Pubblica Amministrazione presenta notevoli difficoltà, soprattutto di carattere organizzativo, economico, professionale, sociale. È quindi opportuno descrivere quali sono le tradizionali dimensioni della qualità per comprendere in che modo lavorare e offrire servizi.

La prima dimensione riguarda la *qualità organizzativa* che si ispira al modello teorico del “*total quality management*”, ovvero della gestione totale della qualità. Mettere in atto questo modello comporta progettare un'organizzazione non solo centrata sulle singole prestazioni, sulle singole funzioni, ma anche sul prodotto finale dell'azienda, ovvero, nel nostro caso sul miglioramento della performance dell'erogazione della didattica. Tutto ciò impone che l'intera organizzazione debba modellarsi attorno alla soddisfazione e ai bisogni dello studente. Questa dimensione della qualità può essere migliorata se i processi organizzativi aziendali si ispirano alla seguente parola d'ordine: EFFICIENZA.

La seconda dimensione è chiamata *qualità tecnica*. Le aziende devono porre molta attenzione su come, tecnicamente, devono prestare servizi, e devono essere in grado di capire quali siano i servizi da realizzare. Da queste poche righe si è già compreso che la parola d'ordine di questa dimensione della qualità è EFFICACIA. Ogni giorno dovrebbero essere svolte prove di efficacia che stanno alla base delle prestazioni didattiche.

Si parla anche di *qualità percepita*: le due dimensioni della qualità appena descritte, quella organizzativa e quella tecnica, si integrano con quest'ultima dimensione. L'introduzione sistematica del questionario della valutazione della didattica rappresenta il tentativo di migliorare quegli aspetti della qualità del servizio che sono direttamente percepibili da parte degli utenti. Concludendo la parola d'ordine che descrive questa dimensione della qualità è: SODDISFAZIONE DEL CLIENTE.

Il primo e più importante principio per implementare la qualità è il **focalizzarsi sul cliente**. L'obiettivo della soddisfazione del cliente è fondamentale

per la qualità totale e si esprime nel tentativo dell'organizzazione di progettare e gestire prodotti e servizi che rispondano ai bisogni del cliente. La soddisfazione del cliente è il requisito fondamentale a lungo termine per il successo dell'impresa. È necessario quindi che l'organizzazione si allinei ai bisogni del cliente.

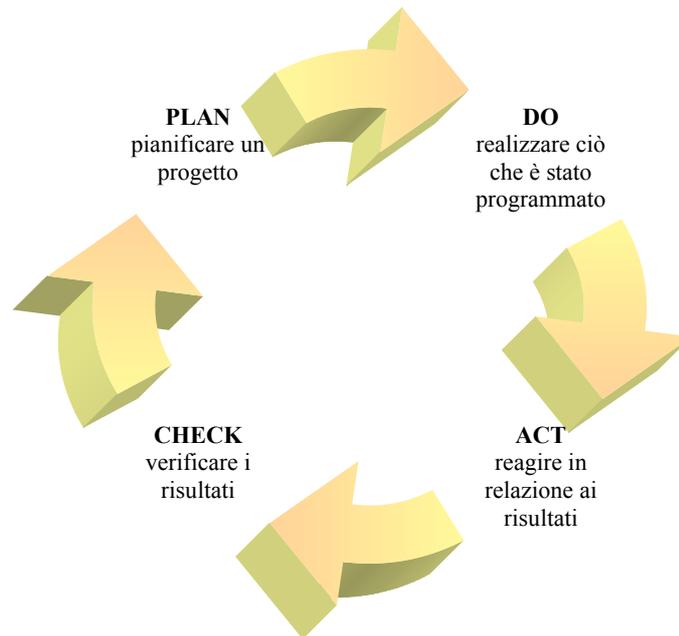
Il secondo principio della qualità totale è l'implementazione del **miglioramento continuo**. Consiste in un costante esame delle tecniche e dei processi amministrativi nella ricerca dei metodi più adeguati. Questi principi implicano il concetto di un'organizzazione intesa come un sistema di parti collegate e la convinzione che, migliorando questi processi, si vada incontro alle aspettative dei consumatori.

Un metodo efficace per operare secondo la logica del miglioramento continuo è quello di applicare a un'iniziativa specifica o all'intera attività dell'organizzazione il ciclo PDCA (Deming, anni '50):

- **PLAN**: pianificare un progetto, un programma o un intervento
- **DO**: realizzare ciò che è stato programmato
- **CHECK**: verificare i risultati di quanto realizzato
- **ACT**: reagire in relazione ai risultati

Ogni Ateneo, così come una qualsiasi organizzazione che può definirsi come un'entità sociale, ha solitamente un momento in cui pianifica la propria attività (PLAN). A questa fase segue l'esecuzione di quanto pianificato (DO). Perché ogni ciclo segni un miglioramento rispetto al precedente è necessaria una verifica (CHECK) della coerenza tra quanto fatto (DO) e quanto si era previsto di fare (PLAN). Se i risultati ottenuti rispondono alle aspettative il processo è 'stabile'; in caso contrario dovranno essere messi in atto i necessari correttivi ricominciando dalla fase della pianificazione (Figura 1.1).

Figura 1.1 – Il ciclo di Deming



1.3 La customer satisfaction

La parola “Customer Satisfaction” entra ogni giorno di più nel nostro vocabolario e soprattutto in quello delle priorità aziendali. Ogni azienda che produce prodotti o servizi si pone come obiettivo la soddisfazione del cliente.

La Customer Satisfaction è cultura del servizio, che è logica del dare per ottenere a favore di un vantaggio reciproco tra personale e cliente.

Essa, quindi, può favorire, a certe condizioni, la strategia del miglioramento continuo.

La necessità di un miglioramento continuo riguarda non solamente le organizzazioni private ma anche le amministrazioni pubbliche, in quanto la soddisfazione del cliente-utente è relativa (dipende dal contesto cui appartiene) e dinamica (varia nel tempo).

Per mantenere, quindi, un certo livello di soddisfazione del cliente-utente o, meglio ancora, aumentarlo, occorre attivare un processo di miglioramento continuo attraverso uno sforzo organizzato e sistematico che coinvolge tutte le funzioni aziendali, a tutti i livelli. Questo sostanzialmente per due motivi: il primo

fa riferimento a tutti gli standard raggiunti, che, con il passare del tempo, tendono a deteriorarsi; pertanto è necessario sviluppare almeno un'attività di mantenimento per conservare il valore dello standard uguale a quello che era stato precedentemente raggiunto; il secondo sottolinea che il cliente, con il passare del tempo, si abitua al valore dello standard e perciò vuole di più e sempre di più, scattano in lui nuovi stati di necessità, pena una sua congenita insoddisfazione.

Si può quindi affermare che l'orientamento sul quale si pone più attenzione è quello strategico, che mette la figura del cliente in primo piano, essendo la vera molla dello sviluppo e della sopravvivenza aziendale. Questa visione condiziona le strategie complessive e competitive, ed è vista sia come fonte del vantaggio competitivo sia come elemento capace di accrescere il valore per tutte le diverse parti in gioco.

È un obbligo strategico la ricerca della corrispondenza tra le esigenze del cliente e le proprie prestazioni. In questo senso, soddisfazione del cliente significa per l'azienda fornire beni/servizi coerenti con le attese del cliente stesso ma, soprattutto, porre quest'ultimo sempre più esigente, selettivo e informato, quale arbitro della qualità (Fabbris, 1999), al punto tale da poter affermare che: "Un servizio è soddisfacente quando è giudicato così da chi lo usa e non da chi lo produce" (Gasco, Casati, 1990).

1.4 Le dimensioni della Customer Satisfaction

La qualità di un servizio, come è noto, può essere definita come la globalità degli aspetti e delle caratteristiche di un servizio, da cui dipendono le sue capacità di soddisfare completamente un dato bisogno.

Esiste dunque uno stretto legame tra soddisfazione del cliente e qualità del servizio, in particolare il cliente è soddisfatto quando le prestazioni ottenute coincidono con le sue aspettative o vanno al di là delle sue attese.

La rilevazione della Customer Satisfaction assume dunque uno specifico significato di strumento utile al miglioramento continuo della qualità del servizio.

Per la soddisfazione del cliente è necessario che vengano appagate, in

maniera ben equilibrata, sia le aspettative funzionali, praticamente i bisogni espliciti, sia le aspettative emotive, ovvero i bisogni impliciti o latenti.

La Customer Satisfaction esprime, quindi, il livello di sovrapposizione tra qualità percepita e qualità prevista.

Quando pensiamo ai mezzi che un'azienda ha a disposizione per soddisfare un cliente, scopriamo che essi si possono dividere in tre categorie:

1. il servizio attraverso i sistemi: i tempi di attesa per lo svolgimento di un esame oppure i tempi di attesa per la correzione di un compito scritto;
2. il servizio attraverso le cose: le attrezzature, le aule;
3. il servizio attraverso le persone: l'offerta di informazioni, la competenza professionale, l'assistenza.

Per progettare bene un servizio è necessario impiegare correttamente, combinandoli con abilità, tutti questi mezzi.

In definitiva, per le aziende, la ricerca della soddisfazione del cliente e della qualità dei propri servizi diventano due facce della stessa medaglia: perché la qualità che conta per le aziende non è solo quella tecnica dei propri prodotti, ma è anche e soprattutto quella percepita dal cliente.

CAPITOLO 2

La valutazione della didattica da parte degli studenti

2.1 La valutazione come attività scientifica

La valutazione o ricerca valutativa, sia per le pressioni che vengono dall'Unione Europea, sia per una maggiore attenzione a vecchi e nuovi problemi sociali, è un settore di studi che si è ormai ampiamente diffuso anche in Italia, paese in cui c'è un gran bisogno, oltre che una grande domanda, di Valutazione, e di valutazione sociale in particolare (Capursi, 2002)

Sicuramente altrove tale disciplina è già consolidata da alcuni decenni, ma in Italia il cammino verso una “maturazione” vera e propria è iniziato solamente negli ultimi anni. E le ragioni di questo ritardo sono diverse, ma non tali da pregiudicare la possibilità di colmarlo, visto che le condizioni per farlo ci sono e sono molteplici.

Ma che cosa è la valutazione? In quanto conoscenza scientifica o nuovo sapere la Valutazione è una disciplina complessa. Può essere filosofia o tecnica, politica o sociologia, economia o statistica ecc. Ancora, “[I]a valutazione è [quanto mai una scienza] [...] interdisciplinare, perché ogni ricerca valutativa oltre ad attingere a uno o più [dei suoi] saperi fondanti [sociologia, economia, statistica, psicologia], dovrà entrare in contatto con le discipline che si occupano delle attività o delle questioni da valutare: la medicina nella valutazione in campo sanitario; la geologia, le altre scienze naturali per la valutazione di impatto ambientale; architettura e ingegneria per la valutazione dei progetti urbanistici, ecc.” (Stame, 1998). Ma per quanto in Italia la valutazione sia fortemente auspicata e voluta, essa non costituisce ancora una *disciplina scientifica*

nell'accezione tradizionale del termine. In generale, una disciplina diventa autonoma quando giunge a definire in modo autonomo rispetto ad altre discipline, affini e non, il proprio campo di studio e di ricerca, un proprio e coerente corpo di conoscenze teoriche ed empiriche, un proprio metodo e una propria strumentazione di ricerca.

La parola “valutare”, da cui deriva valutazione, può assumere differenti significati:

- a) “esprimere un giudizio di valore, di merito su...” che rinvia alla dimensione valoriale della valutazione;
- b) “tener conto di ...” che rinvia al contesto della valutazione;
- c) “misurare, calcolare ...” che rinvia al carattere scientifico, ai metodi e alle tecniche valutative.

Se si vuole pervenire ad una definizione più formale della valutazione come attività di ricerca scientifica, è utile ed indispensabile passare in rapidissima rassegna, alcune delle principali definizioni riscontrabili nella letteratura nazionale.

Il solo termine italiano “valutazione”, affermano D. Fanciullacci, M. Micarelli e G. Pennisi, introducendo il tema della valutazione di progetti, “si riferisce sia ad un esame critico preventivo (valutazione *ex ante*) che ad un esame retrospettivo (fatta durante o dopo l'esecuzione) dell'iniziativa di cooperazione e deve quindi essere associato ad un attributo che ne qualifichi il momento” (Fanciullacci, Guelfi, Pennisi, 1991). “Il termine stesso di “valutazione” implica un certo grado di arbitrarietà, di soggettività” (Bezzi, 1995).

In generale, la valutazione può essere intesa come quel “processo che accompagna lo sviluppo delle decisioni, consentendo al decisore di esprimere un giudizio possibilmente libero da stereotipi e pre-giudizi” (Bertin, 1995), una valutazione, quindi, intesa come funzione di una decisionalità libera da pregiudizi. A. Vergani (2002), individua nella valutazione la funzione atta ad emettere un giudizio in vista di una decisione, intendendo su questa linea il termine valutazione come “*un'attività deliberata e socialmente organizzata che porta alla produzione di un*

giudizio di valore – di norma articolato – in vista di una, più o meno strettamente conseguente, presa di decisione ad esso riconducibile”.

È bene fin d’ora sottolineare l’importanza che riveste il giudizio nella valutazione. In tale ambito ci riferiremo sempre al giudizio “razionale”, cioè al giudizio fondato sulla conoscenza e tale è l’attività valutativa. Valutare significa giudicare ed il giudizio può essere realizzato attraverso il pre-giudizio, vale a dire utilizzando le credenze e le conoscenze che la storia individuale e organizzativa ha consentito di sedimentare nel tempo. Oppure si può sostituire questa impostazione con una raccolta sistematica di informazioni che consentono di aumentare la capacità di giudizio. Per far questo è necessario riconoscere che ogni giudizio è prodotto dall’interazione di alcune assunzioni di valore, messe in relazione con informazioni e conoscenze relative al fenomeno che si intende valutare. In questa logica valutare significa aumentare la capacità di giudizio supportandola con la definizione degli elementi valoriali che lo compongono e con l’attivazione di tecniche di ricerca capaci di aumentare la conoscenza sul fenomeno che deve essere giudicato.

Per completezza riportiamo di seguito la classificazione degli approcci valutativi effettuata da Stame (2001):

- la valutazione basata sugli obiettivi: mira a verificare se il risultato previsto da un programma sia stato raggiunto (efficacia), o di quanto ci si discosti, con le risorse disponibili (efficienza). Ruolo del valutatore: verificare solo se il risultato è stato raggiunto senza entrare nel merito del cambiamento che il programma prevedeva;
- la valutazione basata sui criteri di qualità: vengono stabiliti dei criteri e degli standard, il che implica sia stato definito cosa è un buon programma o un buon servizio. Per esempio se si pensa alla produzione di servizi alla persona di pubblica utilità (SPPU), come è quello universitario, bisogna che siano stati individuati i veri destinatari, le modalità di erogazione del servizio, gli effetti che si vogliono ottenere, i processi che ne consentono la realizzazione. Ruolo del valutatore: non solo recepire i criteri di merito ma individuare le ‘non conformità’;

- la valutazione basata sul riconoscimento del successo: si interviene durante l'attuazione di un programma per individuare eventuali problemi e trovare una soluzione per essi. Ciò implica l'impossibilità a confrontare programmi svolti in luoghi diversi. Ruolo del valutatore: cercare di cogliere cosa abbia prodotto buoni risultati tenendo conto del parere degli attori e, quindi, formulare un giudizio che rappresenti una negoziazione fra le parti.

La valutazione include in sé, in quanto approccio scientifico e metodologico, una forte valenza decisionale e partecipativa oltre che pratica; una valutazione in stretto rapporto con la sfera politica, con i *decision maker*, con le *policy*, con un forte collegamento con la programmazione, proposizione di un modello di ricerca valutativa volta alla trasformazione. Ma anche una valutazione partecipata che abbia come approccio metodologico il coinvolgimento degli *stakeholder*, la percezione degli atteggiamenti e una vocazione all'interesse collettivo.

2.2 La valutazione nelle istituzioni

Lo strumento della valutazione nelle istituzioni può essere considerato come opportunità da cogliere per usare meglio le risorse finanziarie e umane. In ogni caso, esso rappresenta un'importante componente della riforma in atto per la regolazione del sistema istituzionale, e in quanto tale è destinato a consolidarsi anche su sollecitazione della trasformazione organizzativa che si sta verificando a livello politico-istituzionale nel nostro paese e che determina nuove dimensioni politiche e istituzionali come l'autonomia.

È in questo scenario dominato dal cambiamento che si inserisce la trasformazione del sistema universitario. La struttura del sistema di istruzione superiore italiano e il suo modo di funzionare non è sempre stato quello che oggi conosciamo e viviamo. Esso è un prodotto storico-istituzionale del processo di *modernizzazione*, prima, poi di quello *neoistituzionalista* ancora *in fieri*.

In ambito universitario il processo di trasformazione è stato lento, ma grazie anche ad una serie di interventi legislativi che hanno dato vita ad un sistema di valutazione centrale e locale si è generato un "mondo universitario autonomo e

responsabile”. L’autonomia universitaria, causa prima della trasformazione, implica un alto grado di responsabilità e un sistema di valutazione – centrale e locale – come sua condizione necessaria e come suo elemento di equilibrio.

In sintesi, il sistema universitario italiano può essere assunto come modello del processo di istituzionalizzazione della valutazione in virtù di una concezione in base alla quale la stessa valutazione è incorporata (*embedded*) nel sistema universitario e ne costituisce attività essenziale.

A questo punto, è opportuno chiarire l’*iter* che ha seguito la valutazione nel sistema universitario per definirne le caratteristiche.

Nell’università tradizionale, il sistema di istruzione era sistema di élite. Esistevano poche sedi universitarie, dotate di un’élite di accademici che insegnava ad un’élite di studenti. Le università tradizionali gestivano il “piccolo quantitativo”: l’entità della spesa pubblica destinata alle università era modesta, la differenza tra le sedi universitarie era quasi inesistente, sia in termini di *curricula* dei docenti sia in termini di *curricula* dei laureati. Per definizione, il sistema universitario tradizionale si (auto)-gestiva, tanto che gli accademici, professionisti di sicuro successo, avevano una scarsa propensione ad essere gestiti, erano dotati di un forte senso di responsabilità e impegnati nel conseguimento di una eccellenza che, seppure autoreferenziale, era al tempo stesso una garanzia.

Nell’università degli anni ’70, si verifica una crescita del processo di istruzione che “l’università di massa” è chiamata a gestire: la globalizzazione e le conseguenti esigenze di liberalizzazione degli accessi all’istruzione superiore determinano un aumento nella differenza qualitativa tra le diverse sedi universitarie con una caduta nei servizi e un ampliamento delle “aree di parcheggio” della popolazione studentesca. Aumenta la quota dei fondi erogati dallo Stato per il funzionamento degli atenei. E in un sistema, dove le risorse sono per la gran parte di natura pubblica, è necessario che attecchisca e si consolidi un sistema di valutazione che ne regoli i comportamenti e ne controlli i risultati per garantire agli utenti la qualità del servizio erogato.

Nell’università dei nostri giorni si vive una fase di profondo rinnovamento legata alle continue trasformazioni socio-politiche e culturali del nostro paese e

alla presenza dell'Europa. Se la realtà del Paese nella sua globalità è molto complessa, lo è pure quella universitaria con le sue diverse dimensioni – didattica, ricerca, amministrazione e riforme.

L'ultimo decennio del secolo scorso ha segnato un notevole cambiamento nel sistema universitario italiano, a partire dal 1990, fino ad arrivare alla riforma universitaria avvenuta con il DPR 509 del 1999, seguita dalla trasformazione degli anni 2000. Con il DPR 341/1990 e l'istituzione dei Diplomi Universitari, l'Italia recepisce le direttive europee indirizzate a formare figure "giovani" provenienti da percorsi triennali altamente professionalizzanti. In breve il sistema universitario storicamente statico, da tempo contenitore di studenti con iscrizioni decennali, diventa dinamico, ed i programmi didattici dei nuovi corsi universitari, indirizzano lo studente verso specifici sbocchi professionali.

La riforma universitaria del '99 con l'istituzione delle lauree di primo e secondo livello, disattiva i diplomi universitari, facendo nascere la figura del "laureato triennale", modificando in relazione al titolo accademico, i programmi didattici.

Il processo di riforma del sistema universitario, che sembrava essersi concluso con il Decreto 3 novembre 1999, n. 509, è stato riaperto dal Ministro Moratti con il Decreto 22 Ottobre 2004, n. 270 che modifica nella sostanza l'impianto del sistema universitario separando nettamente triennale e specialistica, spezzando la continuità didattica dei 300 crediti; infatti, tale decreto stabilisce che la laurea Magistrale si consegue con 120 crediti.

Il "sistema Paese" evidenzia, quindi, la necessità di innalzare il livello culturale e per farlo ha bisogno di un sistema di istruzione superiore che rappresenti il luogo idoneo dove produrre sia la ricerca finalizzata all'innovazione tecnologica e organizzativa, sia una formazione intellettuale e scientifica capace poi di sviluppare e mantenere livelli occupazionali socialmente accettabili.

In questo scenario, lo strumento/la risposta per comprendere le/alle dinamiche del processo di adattamento istituzionale avviato per far fronte ai cambiamenti ambientali e struttural-istituzionali del nostro paese resta la valutazione istituzionale del sistema universitario.

2.3 La valutazione istituzionale del sistema universitario

Il percorso di attuazione dell'autonomia universitaria ha la sua svolta significativa da un punto di vista normativo nella legge 168/1989 (articolo 6) di istituzione del Ministero dell'Università e della Ricerca Scientifica e Tecnologica (MURST, oggi MIUR). La stessa norma che sancisce l'autonomia prevede per la prima volta il tema della valutazione interna nelle università, tema che sarà ripreso con vigore dalla legge 537/1993 ("Interventi correttivi di finanza pubblica") istitutiva a livello centrale dell'Osservatorio per la valutazione del sistema universitario presso il MURST e a livello locale dei Nuclei di Valutazione Interna (NVI) presso ciascuna università.

In questo ambiente complesso, il processo di autonomia si conferma con la trasformazione dell'organismo nazionale di valutazione del sistema universitario (Osservatorio) in Comitato Nazionale per la Valutazione del Sistema Universitario (CNVSU, legge 370/99). Il Comitato opera collaborando in stretto rapporto con gli uffici del Ministero e con la rete dei Nuclei di Valutazione di Ateneo. Questi ultimi svolgono la duplice funzione di organismi di autovalutazione degli atenei e di diretti referenti del Comitato. Infatti, la legge 370/99 consolida il ruolo dei Nuclei di Valutazione di Ateneo stabilendo che le università debbono assicurare loro accesso alle informazioni e autonomia funzionale ed introduce la relazione tra valutazione e allocazione delle risorse ministeriali. Inoltre, la legge stabilisce che le università, oltre a trasmettere al Comitato i dati, sono tenute alla trasmissione dell'opinione degli studenti frequentanti sull'attività didattica. Con l'istituzione del CNVSU e con gli altri organismi di valutazione, il processo di valutazione del sistema universitario definisce sia la sua missione ultima, che si esemplifica nella tensione del sistema universitario verso la qualità e la corretta concorrenzialità, sia il suo assetto istituzionale che raggiunge la sua "apoteosi" con un sistema nazionale di valutazione centralizzato che si prefigge di restringere la forbice tra norma e prassi valutativa, cercando di porre la valutazione come elemento costitutivo delle decisioni politiche di governo, della prassi amministrativa ministeriale e della gestione degli atenei. Nella sua globalità, il sistema è orientato a individuare gli

elementi di base per un buon governo, al fine di offrirli ai tre diversi livelli decisionali del sistema universitario – Ministero, Ateneo, unità operative all'interno di ciascun ateneo (Dipartimento, Facoltà, Corso di studi, ...) – e di definire le correlazioni tra le attività di valutazione interna ed esterna.

L'attività di valutazione è prevista anche dalla *Dichiarazione di Bologna*. I ministri dei diversi paesi europei, che hanno aderito all'iniziativa ambiziosa di realizzare entro il 2010 lo Spazio europeo dell'istruzione superiore, concordano sulla necessità che i sistemi nazionali per l'assicurazione della qualità debbano includere la valutazioni di corsi di studio, che preveda una valutazione interna, una revisione esterna, la partecipazione degli studenti e la pubblicazione dei risultati.

L'attività di valutazione interna, o *autovalutazione*, è un'attività per così dire endogena, svolta dagli Atenei attraverso organismi interni alle unità da valutare al fine di responsabilizzare quanti vi operano e ottenere migliori prestazioni.

Diversamente, la valutazione esterna è un'attività esogena, è un processo svolto da strutture esterne non coinvolte e specialistiche, così da assicurare rigore e imparzialità al processo valutativo e con l'obiettivo di innescare comportamenti virtuosi da emulare. Inoltre, la realizzazione di un sistema nazionale di valutazione del sistema universitario si pone, tra gli altri, gli obiettivi:

- a. di offrire agli *stakeholder* informazioni trasparenti così da porli nelle migliori condizioni di scelta;
- b. di garantire il valore degli investimenti pubblici e privati nelle università;
- c. di verificare l'efficienza e l'efficacia dell'università italiana in ambito internazionale, con particolare riguardo all'Europa.

2.4 La valutazione della didattica nell'opinione degli studenti

Tra i numerosi elementi di efficienza ed efficacia che possono essere presi in esame per formulare un giudizio sul sistema universitario, un aspetto di particolare interesse è rappresentato dalla soddisfazione espressa dai più immediati e diretti fruitori/creatori del sistema educativo, gli studenti frequentanti.

Conoscere l'opinione degli studenti sugli insegnamenti e sul corso di laurea che stanno frequentando è sicuramente un'informazione di rilievo che può essere utile agli organi decisionali (singolo docente, consiglio di corso di laurea, consiglio di facoltà, ecc.) per individuare elementi di inefficienza e inefficacia del sistema e, quindi, per porre in essere opportune azioni correttive.

Scopo della valutazione da parte degli studenti è evidentemente quello di ottenere una misura della qualità percepita dal cliente, in modo da verificare quanto il servizio che a esso viene erogato ne soddisfi le esigenze. Indubbiamente ciascun docente dispone già di una propria autovalutazione del proprio corso e del modo di tenerlo, tuttavia una valutazione di seconda parte, effettuata quindi direttamente dagli studenti, presenta grandi vantaggi ai fini delle possibilità di miglioramento dell'intero processo di formazione. Questi vantaggi consistono soprattutto nel favorire la comunicazione reciproca fra studenti e docente, nel consentire a quest'ultimo di conoscere il grado di apprezzamento del suo corso e del suo metodo di insegnamento e nel fornire agli studenti un mezzo per fare presenti esigenze che, se soddisfatte, in molti casi potrebbero innalzare il livello di efficienza dell'apprendimento.

In questo contesto, è però necessario porre molta attenzione alla validità e all'affidabilità delle valutazioni degli studenti. La valutazione della qualità della didattica non riflette facilmente l'effettivo apprendimento, poiché gli studenti non si trovano in una posizione adeguata per valutare ciò che stanno studiando. Sono ritenuti buoni giudici per quanto riguarda, ad esempio, la capacità del docente di comunicare con chiarezza, mentre si considera inopportuna la loro valutazione sulla validità dei contenuti.

Gli indicatori derivanti dall'opinione degli studenti vanno considerati congiuntamente ad altre informazioni e analizzati attraverso il filtro di un giudizio competente e coerente con una corretta politica di assicurazione della qualità.

2.5 L'indagine presso l'Ateneo di Palermo

Il piano di indagine per la rilevazione dell'opinione degli studenti nell'Ateneo di Palermo si può sintetizzare come segue:

- a) popolazione di riferimento: studenti frequentanti;
- b) oggetto della rilevazione: il singolo insegnamento;
- c) strumento di misura: questionario delineato dal Senato Accademico (Appendice C);
- d) tempi della rilevazione: la rilevazione si svolge durante ciascun semestre, nel momento in cui i corsi da monitorare volgono al termine, ossia quando hanno raggiunto almeno i tre quarti della loro durata.

In particolare:

- a) la popolazione di riferimento è costituita da quegli studenti che si trovano in aula durante la somministrazione del questionario. Non si tratta quindi di una indagine campionaria in senso probabilistico, ma di una indagine parziale, in quanto rivolta agli studenti presenti in aula in quella determinata ora di lezione;
- b) per quanto attiene l'oggetto di rilevazione, ovvero l'insegnamento da valutare, è opportuno precisare che in teoria dovrebbero essere rilevati tutti gli insegnamenti attivati nel semestre, come da comunicazione delle singole Facoltà al C. IN. A. P. In realtà la copertura degli insegnamenti valutati, come si può vedere dalla tabella 3.1 a pag 22, non è totale. Ciò è addebitabile essenzialmente alle risorse finanziarie destinate dall'Ateneo alla rilevazione, che non risultano del tutto adeguate perché la rilevazione sia censuaria;
- c) la descrizione dello strumento di rilevazione è riportata nel paragrafo successivo;
- d) la rilevazione viene eseguita soltanto dopo che sono stati svolti almeno i tre quarti delle ore complessive previste da ciascun insegnamento, al fine di dare la possibilità allo studente di effettuare la valutazione sulla base di una sufficiente esperienza. Inoltre, nel caso in cui in aula risultavano presenti meno di 10 studenti (dall'anno accademico 2007/2008 il numero è sceso a 5) non si è proceduto ad effettuare la rilevazione.

Quanto appena detto fa emergere il carattere osservazionale di tipo trasversale del nostro studio.

Infatti, né i soggetti in studio, né le variabili di interesse (cioè gli item del questionario) sono manipolati dal ricercatore; non si conoscono a priori le caratteristiche dei soggetti; il criterio sottostante è il realismo (Kish, 1987). Inoltre, poiché la rilevazione si effettua in un preciso istante temporale si parla di studio trasversale nel quale i soggetti hanno in comune il fatto di frequentare quel particolare insegnamento a prescindere dalle loro caratteristiche socio-demografiche (per ulteriori caratteristiche sul tipo di studio si veda l'Appendice A).

Il questionario è somministrato da 200 rilevatori, appositamente formati, reclutati tra gli studenti di tutte le facoltà presenti in Ateneo.

Dopo la rilevazione, i questionari, suddivisi per insegnamento, vengono consegnati al C. IN. A. P. che, dopo averli numerati per un immediato recupero in fase di controllo e correzione degli errori, trasferisce tutte le risposte su supporto informatico tramite lettore ottico e con un adeguato piano di codifica. Infine, i dati (aggregati per facoltà e per Ateneo) vengono inviati al Nucleo di Valutazione d'Ateneo, che provvede alla loro validazione "formale", all'analisi dei risultati e alla stesura della relazione finale. Ai Presidi vengono fornite (semestralmente) tutte le schede dei docenti e (annualmente) gli aggregati per Corso di Studio e Ateneo.

2.5.1 Il questionario

I questionari implementati dagli Atenei contengono 7 item comuni imposti dal CNVSU e altri item liberamente aggiunti da ogni Ateneo.

In particolare, il questionario utilizzato per rilevare l'opinione degli studenti sulla qualità della didattica degli insegnamenti dell'Università degli Studi di Palermo per l'anno accademico 2006/2007 (Appendice C) recepisce le indicazioni predisposte dal CNVSU per tutti gli Atenei italiani al fine avere una base comparativa uniforme su tutto il territorio, almeno per le domande imposte dal Comitato.

Il questionario è suddiviso in sette sezioni che forniscono informazioni sulle caratteristiche degli studenti, sui giudizi che gli studenti esprimono su molti aspetti dei corsi universitari, dagli obiettivi di un corso alla programmazione didattica e le attività di insegnamento. In particolare la prima sezione preliminare, contiene informazioni generali circa il corso (periodo didattico, tipo di laurea, denominazione dell'insegnamento...), seguita dalla sezione A riguardante le caratteristiche personali degli studenti (data di nascita, residenza, età, tipo di diploma,...). Le sezioni (B, C, D, E, F) contengono item riguardanti i diversi aspetti di un corso: le caratteristiche dell'insegnamento (B), la soddisfazione e l'interesse verso il corso (C), gli aspetti organizzativi (D), le infrastrutture (E), le caratteristiche del docente (F). Infine, la sezione G si riferisce a quegli insegnamenti articolati in moduli.

Gli item delle sezioni B, C, D, E, F, G sono tutti di carattere valutativo, ad eccezione degli item B1 e B2 che si riferiscono alle percentuali di ore di lezione e di esercitazione frequentate dallo studente e dell'item F1 che fa riferimento alla percentuale delle ore di lezione svolte dal docente. Nella sezione C troviamo un item, l'item C2, che misura la soddisfazione dello studente nei confronti dell'insegnamento. Ricordiamo che la soddisfazione del cliente, in questo caso dello studente, rappresenta, nella migliore delle ipotesi, la distanza tra qualità percepita e qualità attesa; purtroppo, il questionario non dà informazioni relative alle aspettative dello studente, ne consegue che nel considerare la soddisfazione come espressione della qualità percepita si assume che la qualità attesa sia una costante.

Quasi tutti gli item sono misurati su scala Likert: *decisamente no, più no che sì, più sì che no, decisamente sì*, ad eccezione di B1, B2 e F1. Alcuni item prevedono come modalità di risposta NA (non applicabile) che è stata opportunamente codificata. I dati, opportunamente codificati, sono trattati rispettando la loro natura ordinale.

Spendiamo qualche parola sulla sezione A. Essa contiene domande di carattere informativo di tipo socio-demografico relative allo studente. L'utilità di queste domande sarebbe notevolmente superiore qualora vi fosse l'opportunità di

correlarle con quelle di carattere valutativo. In caso contrario sarebbe preferibile alleggerire il questionario, facendo spazio ad altre domande di carattere valutativo, e ottenere le stesse informazioni tramite dati di segreteria.

Il problema deriva dal fatto che uno studente, naturalmente, può aver compilato più questionari, tanti quanti corsi ha frequentato nel periodo di riferimento. Un altro studente può essersi astenuto dalla compilazione del questionario, poiché non è obbligatoria.

Il risultato finale è che, non essendo in grado, con l'approccio adottato, di identificare i questionari compilati dal medesimo studente, una correlazione tra le variabili di carattere valutativo con quelle di informativo non può essere effettuata in maniera valida, poiché il legame osservato sulle categorie di quegli studenti che (per caso o per necessità!) hanno compilato un numero maggiore di questionari risulterebbe più forte di quanto effettivamente lo sia.

CAPITOLO 3

Analisi preliminari

In questo capitolo si riportano alcune analisi descrittive e preliminari. Si è scelto a priori di eliminare tutti i questionari in cui la percentuale di lezioni frequentate risulta inferiore al 50% (B1). Ciò allo scopo di migliorare l'attendibilità delle valutazioni. Inoltre, spesso, chi frequenta poco non ha le conoscenze sufficienti per una valutazione consapevole. I questionari eliminati sono stati complessivamente 6865, riducendo il database da 68582 osservazioni a 61717.

3.1 Gli insegnamenti valutati

Di seguito si riporta la percentuale di copertura degli insegnamenti per facoltà.

Tabella 3.1 - Tassi di copertura degli insegnamenti rilevati

Facoltà	Tasso di copertura degli insegnamenti rilevati
1	10,9
2	56,3
3	37,3
4	56,8
5	57,6
6	51,9
7	66,2
8	42,4
9	65,1
10	49,3
11	60,7
12	40,7

Dalla tabella 3.1 si osserva che la facoltà che valuta la maggior parte degli insegnamenti attivati è la facoltà 7 (66,2%) seguita dalle facoltà 9 (65,1%) e 11 (60,7%). Un dato negativo per i nostri intenti è sicuramente quello della facoltà 1 nella quale vengono rilevati solamente il 10,9% degli insegnamenti. Dato lo scarso rilievo quantitativo appena sottolineato, la facoltà 1 non sarà contemplata nelle analisi successive.

3.2 Le risposte “positive” e i dati mancanti

Per evidenziare eventuali addensamenti e/o rarefazioni delle percentuali di risposta, si sono analizzate, nel dettaglio le risposte che gli studenti hanno dato alle domande del questionario. In particolare, le risposte date ad ogni item tendono a concentrarsi sulle modalità *più sì che no* e *decisamente sì*. Per tali modalità sono state calcolate le percentuali di risposta per tutti gli item delle sezioni B, C, D, E, F e per tutte le facoltà, escludendo la sezione G che interessa gli insegnamenti articolati in moduli (Tabella 3.2).

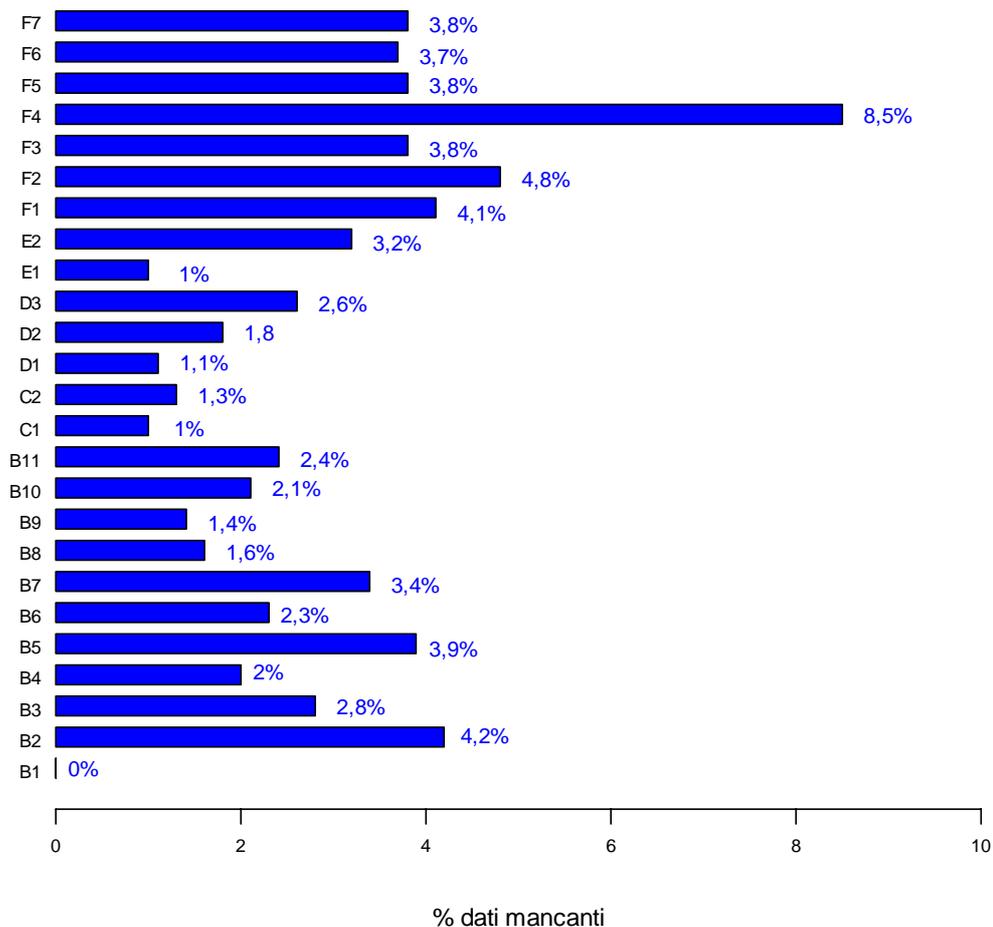
Gli studenti sono molto soddisfatti della disponibilità che i docenti mostrano alle loro richieste di chiarimenti durante le lezioni (F5), infatti gli studenti hanno dato, a questi item, risposte *più sì che no e decisamente sì* nel 92% dei casi. Inoltre si riscontra un'elevata percentuale di risposte positive per gli item F3 e F4, relativi alla puntualità del docente nel rispettare l'orario di svolgimento delle attività didattiche e l'orario per il ricevimento. Complessivamente gli studenti sono molto interessati ai contenuti degli insegnamenti seguiti (C1) con una percentuale di risposte positive pari all'88%. Possiamo qui evidenziare percentuali di risposte positive che vanno dall'85% all'92%. Si notano anche alte percentuali (dal 82% all' 89%) di risposte *più sì che no e decisamente sì* per l'item B6, relativo all'utilità ai fini dell'apprendimento delle attività didattiche. Le percentuali di risposte positive più basse si ottengono per l'item E2, che riguarda l'adeguatezza dei locali e attrezzature per le attività integrative e per l'item D2 concernente la sostenibilità del carico di studio complessivo degli insegnamenti.

Tabella 3.2 - Percentuale risposte più sì che no e decisamente sì

Facoltà	B3	B4	B6	B7	B8	B9	B10	B11	C1	C2	D1	D2	D3	E1	E2	F2	F3	F4	F5	F6	F7
2	78	66	87	76	72	64	68	67	87	74	57	42	70	35	32	80	87	86	90	77	79
3	81	70	89	78	77	72	63	69	87	78	60	43	67	66	55	86	91	90	92	81	82
4	77	68	82	73	78	67	67	66	87	75	62	49	59	56	51	76	84	86	90	76	77
5	79	74	87	79	72	69	66	66	86	72	61	42	68	67	63	85	91	89	92	75	76
6	83	79	88	81	79	74	69	80	91	77	65	43	65	52	51	89	89	93	92	78	80
7	86	76	89	80	78	73	70	75	91	83	65	55	69	63	49	84	90	89	92	84	86
8	83	66	85	71	85	75	72	58	92	82	57	51	45	61	50	87	93	94	95	84	87
9	86	84	87	81	84	79	76	72	90	82	65	52	70	70	59	84	91	93	94	85	86
10	84	64	83	74	85	77	72	75	89	85	62	47	56	61	60	78	90	91	93	84	85
11	73	62	82	68	79	73	66	66	85	78	53	50	49	33	32	74	89	86	91	80	82
12	83	72	84	73	77	72	64	70	88	80	59	43	58	35	46	78	87	90	94	84	87
Ateneo	81	72	86	77	78	72	68	67	88	78	61	46	63	61	54	83	90	90	92	80	81

Appare opportuno avanzare alcune considerazioni sui dati mancanti, anche se in questa tesi il problema non è trattato.

Figura 3.1 - Percentuale dati mancanti per gli item del questionario



La figura 3.1 riporta le percentuali di dati mancanti per item. Essa mette in evidenza come tale percentuale si mantiene sotto la soglia del 5%. L'unica eccezione è rappresentata dall'item F4 ("Il docente rispetta l'orario di ricevimento?") che presenta una percentuale di dati mancati pari all'8,5% con un totale di oltre 5000 osservazioni perse. Se osserviamo la percentuale dei dati mancanti di tale item (Tabella 3.3) si constata che per tutte le facoltà (escluse la 5 e la 6) la percentuale di missing supera abbondantemente la soglia del 5%.

Tab. 3.3 - Percentuale dati mancanti per l'item F4

Facoltà	% dati mancanti
2	17,9
3	9,8
4	7,1
5	4,8
6	4,9
7	8,1
8	7,3
9	9,8
10	8,8
11	8,1
12	5,1

3.4 Gli item eliminati

Gli item che sono stati eliminati ai fini delle nostre analisi sono: B1, B2, B5, B6, B7, B9, C1, E2, F1 e gli item dell'intera sezione G.

Gli item B2, B6, B7, E2, non sono stati considerati nelle nostre analisi, poiché riguardano la valutazione delle attività di laboratorio o altri aspetti che non sono applicabili a tutte le facoltà. Per la stessa ragione la sezione G non è stata considerata, infatti, gli insegnamenti modulari non sono distribuiti tra le facoltà e all'interno delle stesse facoltà in maniera uniforme. Alcune facoltà o corsi di studio possiedono molti insegnamenti suddivisi in moduli, altri non ne hanno o ne hanno pochi. Ciò si traduce in una numerosità molto ridotta rispetto al totale dei questionari. Inoltre in un contesto di analisi orientata alla facoltà appare un dettaglio eccessivo prendere in considerazione domande riguardanti parti di

insegnamento (quali sono i moduli) che appaiono più importanti in un'analisi a livello di corsi di studio o insegnamento.

Per quanto riguarda i rimanenti item è necessario descrivere i motivi dell'esclusione in maggior dettaglio.

L'item C1 ("Sei interessato ai contenuti di questo insegnamento?") non è stato considerato perché appare intrinsecamente "portatore" di qualità, nel senso che è nostra opinione che l'interesse verso una disciplina possa condizionare il "giudizio" verosimilmente verso l'alto, indipendentemente dalla qualità intrinseca del servizio erogato.

L'esclusione dalle analisi dell'item B9 ("Le conoscenze preliminari possedute sono sufficienti per la comprensione degli argomenti trattati?") è motivata in parte dalle stesse considerazioni avanzate per l'item C1 e in parte per le difficoltà che uno studente può incontrare nel definire "sufficienti" le sue conoscenze preliminari.

L'esclusione dell'item B1 è motivata, dalla scelta, già citata, di eliminare tutti quei questionari che presentano una percentuale di lezioni frequentate inferiore al 50%.

L'eliminazione dell'item F1, relativo alla percentuale di lezioni tenute dal docente titolare, trova la sua ragione nella difficoltà di interpretazione dell'item stesso. Si da per scontato che l'elevata percentuale di lezioni condotte dal docente possa essere considerata un arricchimento di nozioni didattiche dell'insegnamento di cui lo studente può usufruire (in questo caso l'item sarebbe orientato positivamente con la qualità della didattica)? L'ambiguità interpretativa è stata 'sciolta' con l'eliminazione di F1.

Si è eliminato dall'analisi l'item B5 ("L'insegnamento ha contenuti sovrapposti a quelli degli altri insegnamenti?") in quanto da precedenti da analisi (Sulis, 2007), si è constatato che gli studenti interpretano (non si capisce il perché) in modo positivo la sovrapposizione ad altri insegnamenti.

CAPITOLO 4

Analisi esplorative

4.1 Premessa

La complessità del concetto che vogliamo misurare richiede un'attenzione in fase esplorativa dei dati ancora più attenta di quanto già normalmente non si faccia. Ciò perché la molteplicità degli item presenti nel questionario, fortemente correlati tra loro, rende difficile la scelta degli item 'responsabili' della qualità della didattica.

Al fine di rendere meno arbitraria tale scelta si fa ricorso a due tecniche esplorative: lo *scaling multidimensionale non metrico* e l'*analisi fattoriale*.

Lo *scaling multidimensionale non metrico* è stato condotto sui dati originali di natura ordinale. Le configurazioni degli item che scaturiscono da esso forniscono indicazioni di massima sulle possibili dimensioni che unitamente prese danno informazioni sul concetto di qualità della didattica.

L'*analisi fattoriale* è stata condotta con lo stesso obiettivo: infatti da un lato si vogliono corroborare i risultati dello *scaling non metrico* e dall'altro si vogliono ottenere dei pesi (i così detti pesi fattoriali) che, mostrando la correlazione tra gli item e le diverse dimensioni ottenute ci permettono di identificare successivamente in un contesto regressivo quali item sono, all'interno di ciascuna dimensione, quelli più 'importanti', ovvero quelli che spiegano di più la soddisfazione dello studente.

L'*analisi fattoriale* è stata applicata a dati metrici ottenuti attraverso la costruzione di un indicatore semplice che ha permesso di sintetizzare i giudizi relativi ad ogni item. Non sarà nostro compito mostrare né tutte le proprietà di tale indicatore, né il modo attraverso il quale è stato ottenuto; si vuole, comunque, riportare per completezza di esposizione la formula dell'indicatore ottenuto:

$$IS_{0,5} = 1 - \left(\frac{1}{3} \sum_{m=1}^3 F_m^{1/2} \right)^2 \quad [1]$$

dove F_m indica la frequenza relativa cumulata dei giudizi in corrispondenza della m -esima modalità della variabile ordinale con $m = 1,2,3$ poiché tutte le variabili considerate hanno quattro modalità (Capusi, Librizzi, 2008)

Una proprietà di tale indicatore è l'invarianza rispetto a tutte le trasformazioni ammissibili per scale ordinali, ovvero le trasformazioni strettamente monotone, pertanto $IS_{0,5}$ dipende soltanto dalla struttura ordinale propria dei dati.

Applicando la formula [1], si ottiene un dato aggregato relativo a ciascun insegnamento per tutte le 11 facoltà; in sostanza le nostre unità d'analisi sono gli insegnamenti delle facoltà considerate, e quindi avremo 11 matrici nelle cui righe sono posizionati gli insegnamenti e nelle cui colonne gli item considerati.

4.2 Lo scaling multidimensionale

Lo *scaling multidimensionale* (*Multidimensional Scaling*, siglato *MDS*) è una tecnica di analisi multivariata sorta inizialmente nell'ambito della Psicometria e successivamente estesa ad altri settori di ricerca quali la sociologia, l'economia e così via.

La tecnica consiste in un insieme di procedure che, partendo da una matrice di prossimità tra n entità o stimoli tra le quali si assumono relazioni monotone, permette di rappresentare geometricamente le medesime in uno spazio S^k di dimensioni minime (quasi sempre k è uguale a 2) in modo tale che le distanze tra i punti nello spazio approssimino il più possibile le dissomiglianze iniziali ottenute. L'insieme dei punti, ognuno individuato dalle coordinate si chiama *configurazione*. Gli scopi che l'analisi si prefigge sono:

- verificare alcune ipotesi teoriche;
- evidenziare i legami esistenti tra le entità, tramite eventuali raggruppamenti delle stesse;
- denominare le dimensioni ottenute.

Inizialmente è necessario procedere alla costruzione di una matrice di prossimità $P_{n,n}$, simmetrica di ordine n (quante sono le entità), a partire dall'originaria matrice entità-variabili $X_{n,t}$ (n entità, t variabili). Il generico elemento δ_{ij} di $P_{n,n}$ sintetizza, in altri termini, il grado di similarità/dissimilarità tra le entità i -esima e j -esima. Si noti che in $P_{n,n}$ la diagonale principale è costituita da valori nulli (essendo nulla la dissimilarità tra un'entità e se stessa); al di sotto della diagonale (e, simmetricamente, al di sopra della stessa) sono riportate le dissimilarità tra le $n(n-1)/2$ coppie di possibili entità oggetto d'analisi.

Solitamente si utilizza la distanza euclidea per modellare la dissimilarità tra le entità i -esima e j -esima:

$$d_{ij}^2 = \sqrt{\sum_{h=1}^t (x_{ih} - x_{jh})^2} \quad [2]$$

dove: $t = \#$ di variabili;

$x_{ih}, x_{jh} =$ valori assunti rispettivamente dalle entità i -esima e j -esima in riferimento alla generica dimensione h ;

$h = 1, \dots, t.$

Successivamente è necessario tramutare le misure di dissimilarità in distanze geometriche.

Come già detto lo scopo dello scaling multidimensionale è prevalentemente esplorativo anche se non sono poche quelle tecniche che assumono una relazione funzionale tra le coordinate della soluzione e le misure di prossimità iniziali. Negli ultimi anni si sono sviluppate anche tecniche di tipo confermativo.

4.2.1 Lo scaling multidimensionale non metrico

I metodi di scaling multidimensionale apparsi per primi sono stati quelli metrici (la natura della trasformazione delle dissomiglianze è metrica), come risposta metodologica all'insufficienza teorica ed empirica dei metodi di scaling unidimensionale (Delvecchio, 1992, Fabbris 1997).

L'introduzione dell'analisi di scaling multidimensionale non metrico è dovuta a Shepard che, generalizzando il modello metrico, considerò misure di prossimità ordinabili (per cui si possono usare come misura delle prossimità anche indici non metrici), supponendo l'esistenza di una relazione monotona fra i ranghi che competono agli indici di distanze e agli indici di prossimità (*condizione di monotonicità*). Quindi, precisando, tale tecnica ha lo scopo di trovare una configurazione di punti in uno spazio di bassa dimensionalità che minimizza una funzione di perdita basata sulle differenze tra le distanze ed una trasformazione non metrica delle dissomiglianze che ne mantiene la monotonia, ovvero ne conserva l'ordine del rango.

Nel nostro caso vogliamo fornire una rappresentazione degli item del questionario le cui risposte presentano una natura ordinale, ciò giustifica l'applicazione di tale tecnica ai nostri dati.

Nei paragrafi successivi è presentata la *misura di prossimità* appropriata per i nostri dati e l'*algoritmo di Kruskal* come procedimento di ricerca della soluzione.

4.2.1.1 La misura di prossimità: il τ_b di Kendall

La prossimità tra due oggetti indica vicinanza nello spazio ma anche somiglianza. Se consideriamo una matrice di prossimità, questa contiene i giudizi di similarità o di dissimilarità che gli intervistati, nel nostro caso gli studenti, esprimono circa le $n(n-1)/2$ coppie di oggetti, gli item del questionario, sottoposti a valutazione. Per variabili di tipo metrico, generalmente, viene usato il coefficiente di correlazione come misura di somiglianza (Delvecchio, 1992). Nel nostro caso, trattandosi di variabili ordinali, la misura di somiglianza deve basarsi sulla corrispondenza fra le modalità degli item, analizzando se alle modalità di ordine più elevato di un item corrispondono più frequentemente le modalità di ordine più basso o più alto dell'altro item considerato. Per tale motivo si vuole un indice che assuma valori positivi nel caso di concordanza tra gli item (l'unità che presenta il rango più alto/basso per la prima variabile ha il rango più alto/basso per la seconda) e valori negativi nel caso di discordanza (l'unità che presenta il rango più alto/basso per la prima variabile ha il rango più basso/alto per la

seconda). In letteratura esistono diversi indici che misurano il grado di associazione tra variabili ordinali e che si basano sul numero di concordanze e discordanze. Un indice di associazione che ha queste caratteristiche è l'*indice γ di Goodmann e Kruskal* che però non tiene conto del numero di coppie di unità che rispetto ad una delle due variabili presentano uguale modalità. Un indice che considera anche il numero di ties è il τ_b di *Kendall* che quindi risulta più appropriato per il nostro lavoro data l'alta presenza di ties (Di Ciaccio, Borra, 1996).

Inoltre è importante ricordare che il coefficiente τ_b è invariante rispetto a tutte le trasformazioni ammissibili per le scale ordinali e quindi rappresenta una misura desiderabile per le misure di prossimità.

L'indice è definito come:

$$\tau_b = \frac{N_s - N_d}{\sqrt{(N_s + N_d + T_x)(N_s + N_d + T_y)}} \quad [3]$$

dove: N_s = # coppie di unità ordinate allo stesso modo su entrambi gli item;

N_d = # coppie di unità con ordinamento discordante;

T_x = # coppie che presentano uguale modalità rispetto al carattere X;

T_y = # coppie che presentano uguale modalità rispetto al carattere Y;

$-1 \leq \tau_b \leq +1$: $\tau_b = +1 \Rightarrow$ ordinamento concorde

$\tau_b = -1 \Rightarrow$ ordinamento discorde

$\tau_b = 0 \Rightarrow$ indipendenza tra coppie di variabili.

L'utilizzo del coefficiente τ_b come misura di somiglianza conduce però ad un problema: una misura di somiglianza, oltre a dover soddisfare la proprietà della simmetria e quella della disuguaglianza triangolare (Fabbris, 1997), deve essere non negativa. Poiché l'indice τ_b varia tra -1 e 1 è necessario applicare una trasformazione monotona che porti τ_b nel range desiderato (si visualizzano solamente 2 valori di prossimità negativa). Inoltre l'analisi dello scaling multidimensionale non metrico è stata svolta con l'ausilio del software R utilizzando una funzione contenuta nel package MASS, isoMDS, che vuole come

argomento una matrice di dissomiglianza; è stato quindi necessario trasformare la misura di somiglianza in misura di dissomiglianza in modo tale da avere lungo la diagonale principale della matrice solo i valori nulli. In conclusione la trasformazione monotona considerata è $1-\tau_b$ che vale 0 per $\tau_b = 1$ (massima concordanza), 1 per $\tau_b = 0$ (indipendenza), 2 per $\tau_b = -1$ (massima discordanza).

4.2.1.2 L' algoritmo di Kruskal

Con i metodi non metrici le soluzioni si ottengono prevalentemente attraverso approssimazioni successive. In tal senso, l' algoritmo più utilizzato fu introdotto da Kruskal (1964) e prevede i seguenti step:

1. si dispongono in ordine non decrescente gli indici di dissimilarità δ_{ij} di $P_{n,n}$;
2. si fissa il numero k di dimensioni e la tipologia di distanza (solitamente euclidea) che intercorrerà tra tutte le entità proiettate nello spazio S^k ;
3. si crea una configurazione iniziale, ad esempio applicando un modello metrico alla matrice $P_{n,n}$ o anche in maniera casuale;
4. si calcolano le distanze d_{ij} fra tutte le possibili coppie di elementi della precedente configurazione;
5. si confronta l' ordinamento delle distanze con quello degli indici di dissimilarità per verificare se è soddisfatta l' ipotesi che l' andamento delle distanze d_{ij} sia funzione monotona dei ranghi $R(\delta_{ij})$ degli indici di dissimilarità δ_{ij} . Ciò può essere messo in luce dal diagramma cartesiano $(d_{ij}; R(\delta_{ij}))$, dal quale si deduce anche il tipo di funzione monotona che lega d_{ij} a $R(\delta_{ij})$;
6. se la relazione fra i suddetti ranghi non è monotona, si calcolano le cosiddette *disparità* $\hat{d}_{ij} = f(\delta_{ij})$ (o pseudo-distanze), con la regressione monotonica di Kruskal (1964b), che consiste nel trasformare le distanze $d_{i,j}$ in modo che:
 - sia rispettato l' ordinamento delle dissimilarità;
 - sia $\hat{d}_{ij} \leq \hat{d}_{hj}$ quando $d_{ij} \leq d_{hj}$;
 - sia minima la funzione Stress (di cui si dirà al successivo punto 7).

In particolare, ordinate le distanze d_{ij} secondo il posto occupato dai corrispondenti valori degli indici δ_{ij} di dissimilarità, se nella successione delle distanze d_{ij} , alcuni valori consecutivi si discostano dalla monotonicità, ad ognuno di essi si sostituisce la loro media: la nuova successione – se è monotona – è quella delle disparità (le quali, quindi, non sono distanze); se, al contrario, detta successione non è monotona, si ripete il procedimento finché non sono soddisfatte le condizioni previste. Riportando in un piano cartesiano i punti di coordinate $(\delta_{ij}; d_{ij})$ si ottiene il cosiddetto diagramma di Shepard che consente di individuare la forma della funzione che lega le distanze alle dissomiglianze.

7. Si valuta la bontà dell'adattamento calcolando il valore dell'indice di Kruskal, denominato Stress:

$$Stress = \sqrt{\frac{\sum_{i=2}^n \sum_{j=1}^i (d_{ij} - \hat{d}_{ij})^2}{\sum_{i=2}^n \sum_{j=1}^i (d_{ij}^2)}} \quad [4]$$

La [4] (che gode della proprietà di essere invariante per rotazioni, traslazioni ed anche variazioni di scala degli assi ortogonali) esprime una misura di quanto i valori di distanza della configurazione iniziale si discostano da quelli della soluzione analitica. L'indice varia tra 0 (quando l'aderenza tra le prossimità iniziali e la configurazione ottenuta è perfetta) ed 1 (quando lo scostamento è massimo). La distribuzione campionaria dell'indice Stress non è nota; tuttavia, sulla base di esperienze empiriche, Kruskal definisce l'adattamento:

<p><i>perfetto</i> se $Stress = 0$</p> <p><i>eccellente</i> se $0,001 \leq Stress \leq 0,025$</p> <p><i>buono</i> se $0,025 < Stress \leq 0,05$</p> <p><i>sufficiente</i> se $0,05 < Stress \leq 0,10$</p> <p><i>scarso</i> se $0,10 < Stress \leq 0,20$</p> <p><i>non significativo</i> se $Stress > 0,20$</p>	[5]
---	-----

Occorre tuttavia precisare che esistono diverse misure del grado di adattamento del modello prodotto con il MDS rispetto ai dati. Tra queste, una delle più impiegate è l'S-stress (Davison, 1983) che, rispetto all'indice proposto da Kruskal, è calcolata sui quadrati delle distanze (e può anch'essa assumere valori compresi tra 0 e 1):

$$S - Stress = \sqrt{\frac{\sum_{i=2}^n \sum_{j=1}^i (d^2_{ij} - \hat{d}^2_{ij})^2}{\sum_{i=2}^n \sum_{j=1}^i (d^2_{ij})}} \quad [6]$$

Il miglior adattamento possibile si avrebbe per $\hat{d}_{ij} = d_{ij}$, con $Stress = S - Stress = 0$; ne consegue che se l'adattamento è buono, i punti di coordinate $(\hat{d}_{ij}; d_{ij})$ dovrebbero disporsi all'incirca sulla bisettrice del primo quadrante.

Un'altra misura di adattamento proposta da Guttman è il *coefficiente di alienazione* (Cox, Cox, 1994). Prima egli definì un coefficiente di monotonicità:

$$\mu = \sqrt{\frac{\left(\sum_{i,j} \hat{d}_{ij} d_{ij} \right)^2}{\sum_{i,j} \hat{d}_{ij}^2 \sum_{i,j} d_{ij}^2}} \quad [7]$$

che misura il grado di associazione ordinale tra i dati originali e le distanze stimate e che ha valore 1 in caso di perfetto adattamento, poi, dalla [7] definì il coefficiente di alienazione:

$$k = \sqrt{1 - \mu^2} \quad [8]$$

Più basso è il valore di k migliore sarà l'adattamento del modello non parametrico ai dati.

8. Se il valore di Stress denota che l'adattamento non è soddisfacente, si determina una nuova configurazione e si reitera la procedura a partire dallo step 4. Il processo iterativo ha termine quando il valore della funzione Stress è non maggiore di un valore predeterminato, oppure quando si registra un numero di iterazioni prefissato, oppure quando la differenza con il valore della funzione Stress del passo precedente è minore di un valore soglia prefissato. Se il minimo valore di Stress (per quel valore di k) non è significativo, vuol dire che il fenomeno non è rappresentabile in S^k ; occorre allora riprendere il procedimento assumendo un numero di dimensioni pari a $k+1$. In tal senso, un metodo per individuare la dimensione ottimale di S^k è quello di rappresentare i punti di coordinate $(k; Stress)$; il valore ottimale di k è quello in corrispondenza di un "gomito" del diagramma, oltre il quale la spezzata si appiattisce.

4.2.2 Interpretazione delle dimensioni

L'interpretabilità "semantica" delle dimensioni trovate gioca un ruolo nella scelta della dimensionalità. Se, per esempio, le dimensioni di una soluzione bi-dimensionale sono interpretabili, mentre non si trova il significato di una soluzione tri-dimensionale, allora la soluzione conveniente è quella a due dimensioni. Inoltre, sempre per migliorare l'interpretazione, si può procedere alla rotazione degli assi, di solito con il criterio *Varimax*, allo scopo di far coincidere la direzione degli assi ortogonali con la direzione del continuum rappresentato dall'insieme di punti della configurazione. La rotazione ortogonale si dimostra

ausilio efficace per l'interpretazione quando il *continuum* ha una forma quasi rettilinea. Caso che, nella realtà, è poco comune. Rotazioni oblique sono, tuttavia, applicate solo nei rari casi in cui le nuvole di punti sono palesemente disposte ad angolo non ortogonale. Non sempre le dimensioni latenti nella configurazione si trovano allineate con gli assi ortogonali anche dopo una o più rotazioni a causa della dominanza esercitata sulla soluzione dalle entità collocate.

Allo scopo di verificare la stabilità di una configurazione si possono per esempio dividere i dati in due metà casualmente determinate, analizzarle separatamente con gli stessi parametri e confrontare i due esiti ottenuti: se questi sono gli stessi, la configurazione è stabile, altrimenti si possono determinare vari gradi di instabilità. Per esempio, possono essere stabili solo le prime due dimensioni e non le successive.

È possibile, inoltre, verificare l'effetto che sul risultato ha l'esclusione di una parte delle informazioni iniziali. Se la configurazione rimane la stessa nonostante l'esclusione di determinati insiemi di entità, significa che le dimensioni trovate sono applicabili all'intera popolazione in esame. Si dimostra (Graef e Spence, 1976), che, in un'analisi dalla quale sia escluso un terzo delle dissomiglianze più piccole, la configurazione multidimensionale mantiene le sue caratteristiche essenziali, ma che l'esame senza una parte delle dissomiglianze più grandi provoca un serio degrado della configurazione. In sostanza, l'applicazione fornisce una buona rappresentazione per la struttura globale, non altrettanto fa per la struttura locale.

4.2.3 Applicazione dello scaling multidimensionale non metrico agli item del questionario

In relazione alle considerazioni fatte nel capitolo precedente, gli item sui quali è stato applicato lo scaling sono: B3, B4, B8, B10, B11, C2, D1, D2, D3, E1, F2, F3, F4, F5, F6, F7. La matrice di partenza presenta nelle righe i 61717 questionari e nelle colonne le 16 variabili, ovvero gli item appena elencati. A partire da questa matrice si sono costruite 120 (il numero delle combinazioni) tabelle di contingenza 4x4 (perché 4 sono le modalità di ogni item).

Per ogni tabella si è calcolato il τ_b di Kendall che ha dato origine alla matrice di somiglianza $P_{16,16}$ sulla quale, dopo aver applicato la già citata trasformazione, si è applicato lo scaling multidimensionale non metrico.

L'analisi è stata svolta distinguendo per facoltà, quindi la metodologia dello scaling è stata applicata 11 volte tante quante sono le facoltà considerate. Ciò ha permesso il confronto tra facoltà.

4.2.4 I risultati

Le configurazioni finali sono state ottenute in uno spazio bidimensionale. Considerando che, la tecnica dello scaling ha natura essenzialmente esplorativa ed è quindi basata su un approccio grafico, la scelta bidimensionale è quella che da un lato porta ad una percezione immediata e sintetica del fenomeno in studio, e dall'altro lato ci consente comunque di apprezzare il miglioramento, in termini di bontà di adattamento, che si ha passando da uno spazio unidimensionale ad uno a 2 dimensioni.

Per ogni configurazione si è calcolato l'indice Stress, il coefficiente di monotonicità e il coefficiente di alienazione riportati nella tabella seguente.

Tabella 4.1 – Scaling multidimensionale non metrico: la bontà di adattamento

Facoltà	Stress	Coeff. monotonicità	Coeff. Alienazione
2	0,08	0,996	0,09
3	0,07	0,997	0,08
4	0,07	0,997	0,07
5	0,08	0,997	0,08
6	0,09	0,996	0,09
7	0,07	0,997	0,07
8	0,09	0,997	0,07
9	0,07	0,998	0,07
10	0,09	0,995	0,10
11	0,12	0,992	0,13
12	0,11	0,993	0,12

Come si può osservare, nonostante non esistano criteri rigorosi ma solo criteri fondati su evidenze empiriche, i valori dello stress sono significativi; il valore più alto si riferisce alla facoltà 11, pari a 0,12 che è comunque accettabile. I

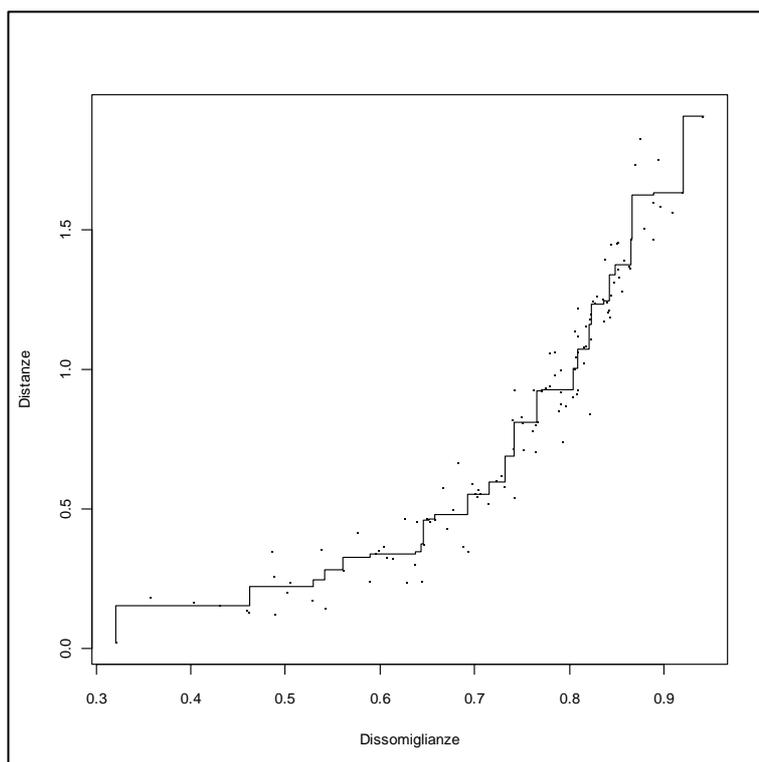
coefficienti di monotonicità presentano valori prossimi all'unità, segnale questo di un buon grado di associazione tra i dati originali e le distanze stimate; inoltre, i coefficienti di alienazione mostrano valori abbastanza bassi che testimoniano un buon adattamento del modello ai dati.

Ad ogni modo, poiché, come detto, non esistono criteri rigorosi per stabilire quanto grandi o piccoli siano questi valori, è opportuno valutare la bontà di adattamento della configurazione finale da un punto di vista grafico.

Attraverso il diagramma di Shepard è possibile osservare la disposizione dei punti (le distanze nella configurazione finale) rispetto alla trasformazione monotona delle dissomiglianze. I diagrammi per le 11 facoltà mostrano una disposizione concentrata rispetto alla trasformazione monotona delle dissomiglianze, quindi l'adattamento sembra buono. Questo risultato era atteso, dati gli alti valori del coefficiente di monotonicità.

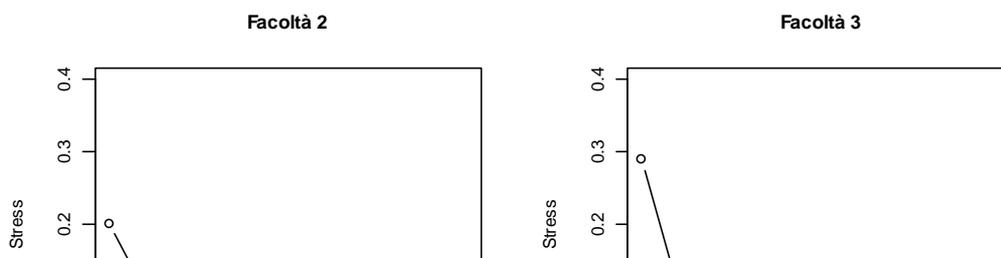
Riportiamo il diagramma di Shepard per una facoltà a mo' di esempio, per osservare il buon adattamento sopra descritto (l'andamento per le altre facoltà è pressoché sovrapponibile).

Figura 4.1 - Diagramma di Shepard per la facoltà 5

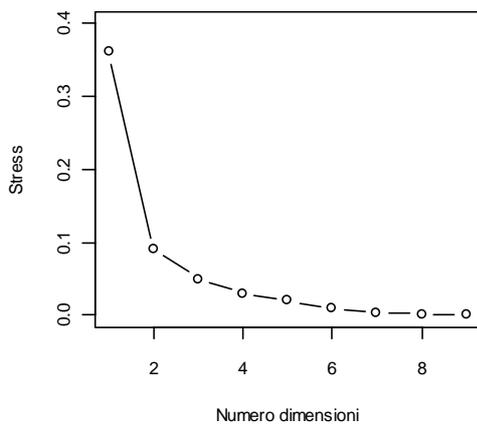


Un'altra rappresentazione grafica utile per valutare la soluzione bidimensionale consiste nel riportare il valore minimo dell'indice Stress, in corrispondenza di un numero variabile di dimensioni, nel nostro caso tra 1 e 9. I grafici per ogni facoltà (Figura 4.2) mostrano un "gomito" in corrispondenza della dimensione pari a 2, oltre la quale la 'pendenza' della spezzata diminuisce notevolmente. Si riportano di seguito i grafici per tutte le facoltà per mettere in evidenza il miglioramento di adattamento che si ha al crescere delle dimensioni.

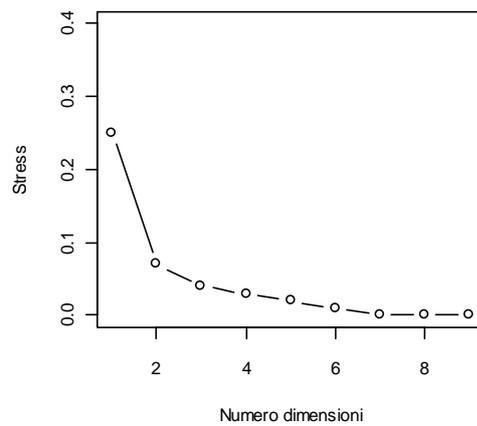
Figura 4.2 - Scaling multidimensionale non metrico: Stress al variare del numero di dimensioni



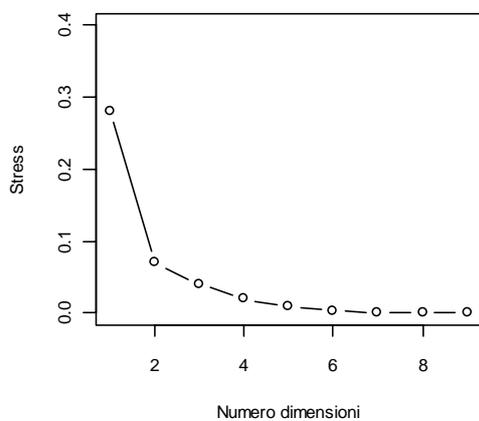
Facoltà 6



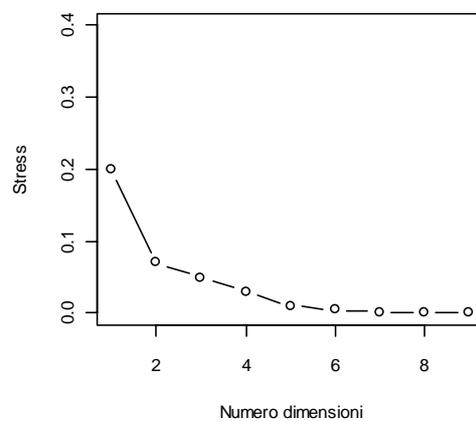
Facoltà 7



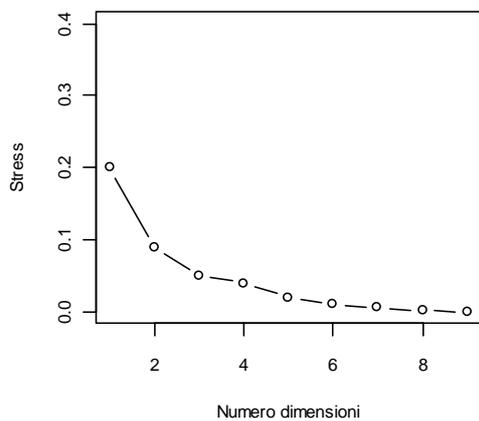
Facoltà 8



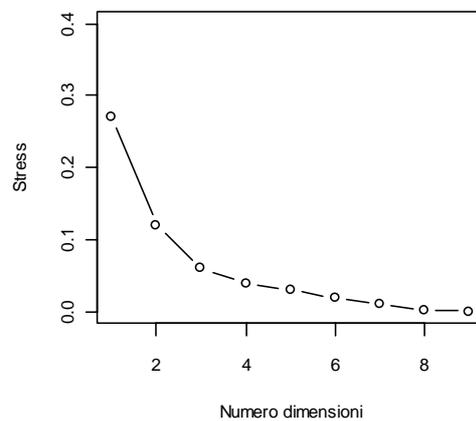
Facoltà 9



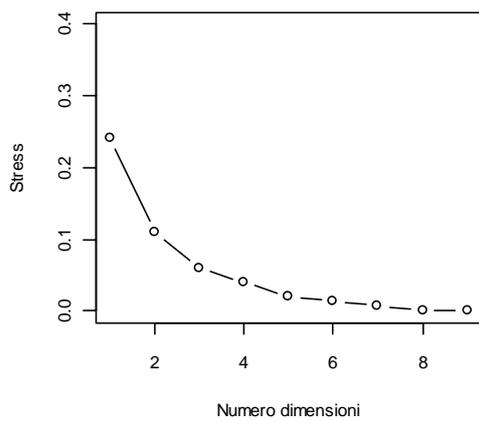
Facoltà 10



Facoltà 11



Facoltà 12



Come si può notare passando dalla situazione unidimensionale a quella bidimensionale si ha una netta diminuzione del valore di stress, mentre dalla terza dimensione in poi esso diminuisce gradualmente. Le facoltà 4 e 6, oltre a presentare un “gomito” in corrispondenza della dimensione 2, hanno un netto miglioramento (rispettivamente 0,30 e 0,27) rispetto alla soluzione unidimensionale.

4.2.5 L’interpretazione della configurazione di punti

La configurazione dei punti (Appendice D), i quali indicano ognuno un item del questionario, permette di fare delle considerazioni sulla qualità della didattica percepita dagli studenti. Naturalmente confrontare la posizione dei diversi item in tutte le 11 facoltà è abbastanza difficile, ma cercheremo di capire se ci sono item per i quali è possibile fare medesime considerazioni nelle diverse facoltà. In particolare, si è scelto di considerare come punto di confronto l’item C2 relativo alla soddisfazione dell’insegnamento, sia perchè si pensa che esso possa esprimere la percezione generale della qualità della didattica da parte di uno studente, sia perché l’item occupa una posizione centrale nelle configurazioni finali risultanti dallo scaling.

Per capire quali item hanno una minore distanza dall’item C2, abbiamo ordinato le distanze e assegnato ranghi crescenti al crescere della distanza da C2. In questo modo per ogni facoltà abbiamo ottenuto l’ordinamento degli item rispetto alla distanza da C2. Quindi, si è calcolato il coefficiente di cograduazione tra due o più ordinamenti *W di Kendall* (Kendall, Gibbons, 1990), che in questo contesto consente di misurare il grado di concordanza fra gli 11 ordinamenti, uno per ogni facoltà, dei 15 item ordinati in relazione alla loro distanza da C2.

La formula del coefficiente di concordanza di Kendall è la seguente:

$$W = \frac{S}{\max S} \quad [9]$$

dove: S = somma dei quadrati degli scarti tra la somma dei ranghi di ciascun item e la media dei ranghi (devianza dei ranghi);

$maxS$ = massimo di S .

Dopo opportuni passaggi algebrici, si dimostra che:

$$max S = \frac{m^2 (n^3 - n)}{12} \quad [10]$$

dove: m = # ordinamenti;

n = # item.

Sostituendo si ottiene:

$$W = \frac{12S}{m^2 (n^3 - n)} \quad [11]$$

Nel nostro caso il valore di W è pari a $0,92$ che testimonia un'elevata concordanza tra gli 11 ordinamenti dei 15 item per le diverse facoltà.

In particolare, dalla tabella 4.2 si osserva che:

- l'item B3 ha rango 1 in 5 facoltà ed è presente nelle prime 3 posizioni in tutte le facoltà, ad eccezione della facoltà 5 per la quale occupa il 4° posto;
- l'item B8 occupa le prime 4 posizioni in tutte le facoltà;
- B4 ha rango al massimo 5 fatta eccezione per le facoltà 10 e 11, dove ha rango rispettivamente 6 e 8;
- F7 si trova nelle prime 4 posizioni in tutte le facoltà, esclusa la facoltà 6 dove ha rango 5;
- l'item F6 è presente nelle prime 5 posizioni in tutte le facoltà;
- l'item F5 è 6° nelle prime 5 facoltà e nella facoltà 9, ha rango 7 nelle facoltà 10, 11, 12 e rango 8 nella facoltà 5;

- B3, B4, B8, F6, F7 hanno rango al massimo pari a 5 nelle facoltà 2, 3, 4, 5, 6, 7, 8, 9, 12 e quindi, globalmente, sono gli item più vicini a C2;
- l'item E1 ha rango massimo in tutte le facoltà, eccetto nelle facoltà 5 e 8, nelle quali è al terzultimo posto (rango 13);
- D3 occupa sempre le ultime 3 posizioni tranne nella facoltà 3, dove è quartultimo (rango 12);
- D2 è sempre quartultimo tranne per le facoltà 4 e 7 dove ha rango 11;
- B11 occupa le ultime 5 posizioni, nelle facoltà 7 e 8 è penultimo;
- D1 occupa le ultime 5 posizioni in tutte le facoltà, tranne per 6 nella quale ha rango 9;
- E1 e D3 occupano le ultime 3 posizioni in tutte le facoltà ad eccezione della facoltà 3 dove D3 ha rango 12.

Tabella 4.2 – I ranghi degli item nelle 11 facoltà

Item	Facoltà										
	2	3	4	5	6	7	8	9	10	11	12
B3	2	1	2	2	4	1	1	1	1	3	2
B4	3	5	5	5	2	2	5	5	6	8	5
B8	1	2	4	1	3	3	3	2	4	4	1
B10	7	9	9	9	11	9	9	7	5	12	6
B11	11	13	13	12	10	14	14	10	10	10	12
D1	10	10	10	10	9	10	11	11	11	11	10
D2	14	14	11	14	12	11	12	12	13	13	14
D3	13	12	14	15	13	13	15	14	14	14	13
E1	15	15	15	13	15	15	13	15	15	15	15
F2	12	11	12	11	14	12	10	13	12	9	11
F3	9	8	8	8	8	7	8	9	9	5	9
F4	8	7	7	7	7	6	6	8	8	6	8
F5	6	6	6	6	6	8	7	6	7	7	7
F6	5	4	3	3	1	5	4	4	3	2	3
F7	4	3	1	4	5	4	2	3	2	1	4

I punti sopra descritti ci permettono di avanzare alcune considerazioni sostantive in merito alla soddisfazione degli studenti e quindi alla percezione della qualità della didattica. Innanzitutto, la soddisfazione nei confronti di un insegnamento è legata alle prime informazioni che in genere devono essere fornite ad uno studente, ovvero gli obiettivi formativi che l'insegnamento si pone (B3) e

le modalità per lo svolgimento dell'esame (B4). In particolare maggiore importanza si dà all'illustrazione degli obiettivi formativi che occupa, come già detto, la prima posizione in 5 facoltà su 11. Gli studenti, per ritenere buona la qualità dell'insegnamento, danno molta importanza all'adeguatezza del materiale didattico per lo studio della materia (B8). Inoltre, gli studenti sono soddisfatti dell'insegnamento sia quando sono stimolati e motivati dal docente verso lo studio della materia (F6), sia quando il docente espone gli argomenti in modo chiaro (F7). La disponibilità del docente a chiarimenti durante le lezioni (F5) è legata, seppur marginalmente, alla soddisfazione dell'insegnamento.

Emerge, ancora, che l'adeguatezza delle aule dove si svolgono le lezioni (E1) occupa un ruolo marginale nella soddisfazione dello studente. A differenza di altre componenti, questa non è direttamente influenzabile dai docenti, dipende dalle dotazioni delle facoltà e non è quindi modificabile prontamente. Un altro ruolo marginale, viene attribuito all'organizzazione (D1, D2, D3), che interessa aspetti influenti sulla qualità dell'insegnamento (l'organizzazione degli insegnamenti, il carico di studio complessivo degli insegnamenti, l'adeguatezza degli orari delle lezioni nei confronti di eventuali spostamenti da un'aula o da una sede all'altra), ma che dipendono prevalentemente dall'organizzazione del corso di studi.

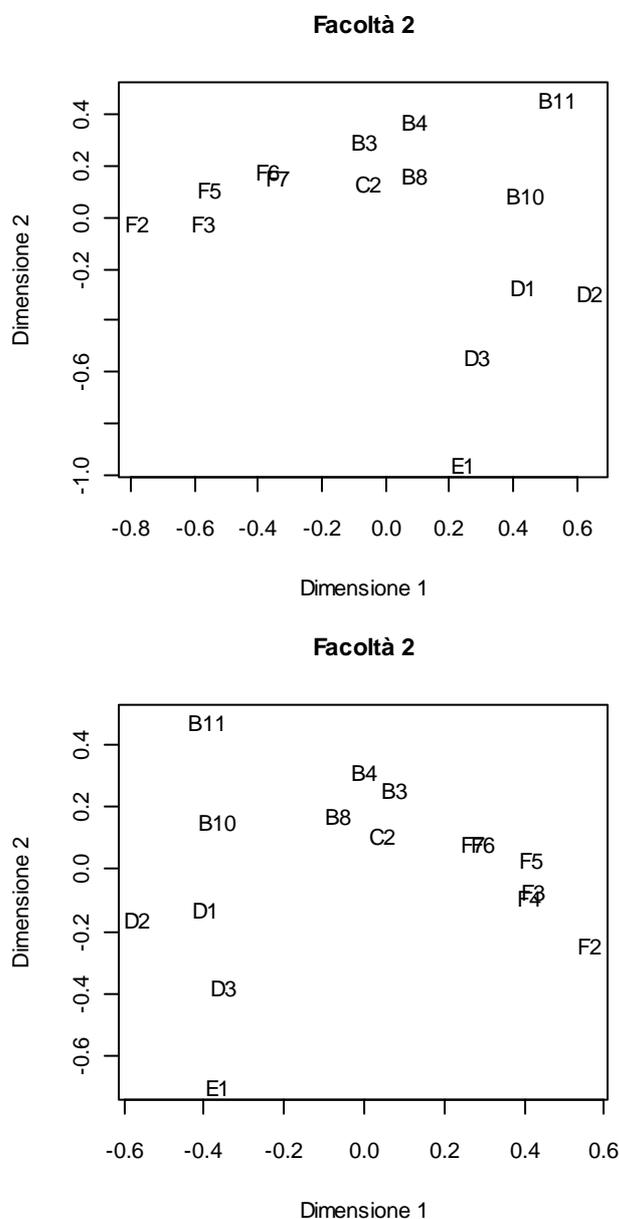
Inoltre, la soddisfazione relativa al corso seguito non è del tutto influenzata dal coordinamento che l'insegnamento in questione ha con altri insegnamenti (B11); probabilmente per uno studente è più importante che il docente spieghi bene, motivi l'interesse, sia disponibile, illustri gli obiettivi formativi e le modalità dell'esame, piuttosto che fornisca informazioni che in qualche modo possano essere collegate con altre informazioni derivanti da altri corsi.

Sostanzialmente, la soddisfazione di un insegnamento è in stretto legame con alcuni item relativi all'insegnamento, in particolare con le informazioni preliminari inerenti la programmazione di un corso stesso: obiettivi formativi, reperimento di materiale didattico e anche modalità dell'esame. Non di meno, la chiarezza del docente e le motivazioni degli studenti verso la disciplina assumono un ruolo di primo piano nella determinazione della soddisfazione. Marginali

rispetto alla soddisfazione sono soprattutto l'adeguatezza delle aule e tutti gli item relativi all'organizzazione dell'insegnamento.

Nell'analisi dei valori mancanti si è osservato che l'item F4 ("Il docente rispetta l'orario di ricevimento?") presenta le percentuali più alte di valori mancanti sia in generale sia per ogni facoltà. Probabilmente, molti studenti si sono astenuti dal rispondere in quanto non hanno mai usufruito del ricevimento studenti, soprattutto se riferito agli studenti del 1° anno. Per tale motivo si è svolta un'analisi di scaling per quelle facoltà dove la percentuale di valori mancanti per l'item F4 è superiore al 5%. Le uniche facoltà con una percentuale inferiore al 5% sono la quinta, la sesta e la dodicesima. Effettuando un'analisi di scaling, dopo aver eliminato l'item F4, non si evidenziano forti cambiamenti nella configurazione degli item; la posizione dei punti cambia direzione solamente nella facoltà 2 (Figura 4.3) che presenta la percentuale più alta di valori mancanti, ma se si controllano le distanze da C2 queste rimangono immutate.

Figura 4.3 - Scaling multidimensionale non metrico: configurazione degli item nella facoltà 2 con e senza l'item F4



Allo scopo di verificare la stabilità della configurazione si sono esclusi gli item E1 e D3 che occupano posizioni periferiche. Le configurazioni sembrano stravolgersi, tuttavia le distanze dall'item C2 rimangono inalterate (alcuni item cambiano rango ma lo scarto è al più di una posizione). Ciò ci conduce ad affermare che la disposizione degli item rimane pressoché costante così come la posizione relativa all'item C2.

4.3 L'analisi fattoriale

Con il termine generale *analisi fattoriale* si denota una classe di tecniche statistiche principalmente usate per ridurre e sintetizzare i dati relativi a più variabili in un insieme più piccolo di *fattori*, che sono generalmente diversi dalle variabili originali benché indotti da queste. I *fattori* sono variabili (o *costrutti latenti*) che non sono direttamente osservabili ma devono essere inferiti dalle variabili-input.

Le relazioni tra un set di molte variabili interrelate sono esaminate e rappresentate in termini di pochi fattori sottostanti.

Nell'analisi della varianza, nella regressione multipla, nell'analisi discriminante, una variabile è considerata dipendente e tutte le altre sono indipendenti o predittori.

Nell'analisi fattoriale, invece, non esiste questa distinzione; piuttosto, in quest'ultima è esaminato un intero set di relazioni interdipendenti.

Il processo d'analisi si basa su una matrice di correlazione tra le variabili originarie. Infatti, affinché l'analisi fattoriale sia appropriata, le variabili devono essere correlate fortemente. Ci dovremmo aspettare che se le variabili fossero molto correlate fra loro, allora dovrebbero essere fortemente correlate anche con lo stesso fattore o fattori.

Per identificare i fattori, esistono in letteratura diverse procedure alternative a seconda delle finalità della ricerca. Nel nostro caso ricorremo **all'approccio delle componenti principali** che è preferibile quando non si ipotizza l'esistenza di un processo comune sottostante il set di misure (cioè quando si vogliono analizzare variabili che sono empiricamente correlate, ma senza specifiche ipotesi di appartenenza delle variabili allo stesso dominio concettuale). Tale metodo risulta, quindi, meno sensibile, rispetto agli altri metodi, alle assunzioni sulla distribuzione dei dati.

Al fine di riassumere l'informazione contenuta nelle variabili originarie, un numero più contenuto di fattori dovrebbe essere estratto.

I fattori ottenuti dall'analisi fattoriale sono indipendenti fra di loro e riproducono la maggior parte della varianza dei dati.

L'autovalore relativo al primo fattore rappresenta la quota maggiore di varianza, l'autovalore relativo al secondo fattore spiega la seconda quota di varianza e così via.

Il problema è, quante variabili latenti devono essere selezionate? Diverse procedure e criteri suggeriscono il numero di fattori da considerare:

- col criterio degli autovalori (eigenvalues), mantenendo solo i fattori con autovalori maggiori di 1 (dove l'autovalore, Ω_i , rappresenta l'ammontare della varianza dell' i -esimo fattore e la somma degli autovalori è uguale alla somma delle varianze, se l'analisi è condotta su una matrice di varianze-covarianze e a p , numero di variabili, se è condotta su una matrice di correlazione);
- col criterio dello scree-plot sulla base della rappresentazione grafica degli autovalori rispetto al numero dei fattori in ordine d'estrazione.
- col criterio della percentuale della varianza, estraendo fattori fino a che la percentuale cumulativa di varianza estratta dai fattori raggiunga un livello soddisfacente, in genere il 75%; tuttavia nell'analisi di dati derivanti da indagini sociali, è accettabile anche una frazione pari al 60%, o anche minore.

4.3.1 L'analisi fattoriale applicata ai nostri dati

Come detto nella premessa al fine di individuare le dimensioni componenti la 'qualità della didattica', con particolare riferimento all'insegnamento si conduce un'analisi fattoriale sulle variabili individuate dagli indicatori semplici.

I dati a nostra disposizione ci hanno permesso di costruire 11 matrici per le 11 diverse facoltà. Ogni matrice presenta nelle righe gli insegnamenti per la facoltà considerata, e in colonna gli item del questionario. Ovviamente le diverse matrici avranno stesso numero di colonne ma diverso numero di righe. L'elemento generico di ogni matrice è rappresentato dal valore che l'indicatore semplice assume per un particolare item in corrispondenza di un determinato insegnamento. Gli indicatori sono relativi agli item B3, B4, B8, B10, B11, C2, D1, D2, D3, E1, F2, F3, F4, F5, F6, F7.

L'estrazione dei fattori, come già detto, è stata effettuata con il metodo delle componenti principali. Sono stati scelti per ogni analisi fattoriale i fattori con autovalore superiore all'unità. Inoltre, poiché nella maggior parte dei casi, la matrice fattoriale (*factor pattern matrix*) che contiene i *factor loadings*, non consente facilmente di interpretare i risultati, poiché ciascun fattore risulta correlato con molte variabili (invece che con poche e diverse, fortemente), si è optato per una trasformazione della matrice attraverso la rotazione dei fattori (*factor rotation*), che consiste in un cambiamento di posizione delle dimensioni estratte nella prima fase dell'analisi allo scopo di facilitare la comprensione del significato dei fattori stessi.

La soluzione ideale, ai fini dell'interpretazione dei fattori, è quella in cui tutti i pesi fattoriali sono prossimi a 0 o a 1. La rotazione comporta pertanto una ridistribuzione delle comunanze delle variabili e della varianza spiegata dai fattori. Il metodo di rotazione che è stato considerato è quello *Varimax*, che effettua una rotazione ortogonale e che modifica le colonne della matrice di pesi fattoriali, agevolando il compito di associare le variabili ai singoli fattori. La percentuale di varianza complessiva dei fattori ruotati rimane inalterata, mentre si modifica la percentuale di varianza spiegata da ciascun fattore. La rotazione *Varimax*, è dunque appropriata per interpretare i fattori.

4.3.2 Le matrici fattoriali

Interpretare un fattore significa dargli un nome, il che però non deve essere dettato dal caso o frutto della fantasia. L'interpretazione è facilitata identificando, nella matrice di rotazioni dei fattori, gli item che sono correlati con uno stesso fattore. Il coefficiente di correlazione tra un fattore e una variabile si chiama anche *saturatione*. Quanto più il valore del coefficiente è elevato in valore assoluto, tanto più l'item si considera determinante per quel fattore. Non esistono delle regole universalmente riconosciute per fissare il limite dopo il quale una saturazione è rilevante. Si può, ad esempio, accettare il suggerimento di Overall e Klett (1972) di prendere come valore minimo in assoluto 0,35.

Di seguito, per ogni facoltà, riportiamo la matrice fattoriale, cercando di fornire un'interpretazione dei fattori ottenuti.

Facoltà 2

La tabella 4.3 mostra i risultati dell'analisi fattoriale per 4 fattori, gli unici che hanno autovalore maggiore di 1. In particolare tali fattori spiegano complessivamente il 73% della varianza. Il primo fattore individua gli aspetti della qualità della didattica legati al singolo 'insegnamento', con riferimento alla 'capacità del docente' (F5, F6, F7) e alla sua 'puntualità' (F2, F3, F4).

Il secondo fattore è legato sempre al singolo 'insegnamento', ma dal punto di vista delle 'informazioni' preliminari che uno studente riceve all'inizio di ogni corso (B3, B4, B8): obiettivi formativi dell'insegnamento, modalità di esame, materiale didattico indicato o fornito. Inoltre il fattore 2 ha coefficiente abbastanza elevato (0,660) con la variabile B11 ("L'insegnamento ha contenuti coordinati con altri insegnamenti?"). Il fattore 2 è anche correlato (0,797) con C2, l'item relativo alla soddisfazione generale dell'insegnamento; ciò porta alla considerazione che la qualità della didattica è fortemente legata, in termini di soddisfazione degli studenti, alle informazioni preliminari che ogni docente dovrebbe fornire nelle prime lezioni di un insegnamento.

Gli item relativi all'organizzazione (D1, D2, D3) determinano il terzo fattore. A tali item è possibile aggiungere anche B10 ("Il carico di studio richiesto da questo insegnamento è proporzionato ai crediti indicati nel piano di studi?"), anche se ha un peso pari a 0,505.

L'item che fa riferimento alle infrastrutture (E1) è l'unico che è correlato con il quarto fattore.

Tabella 4.3 - Factor pattern ruotato: facoltà 2

Item	Componente			
	1	2	3	4
B3	0,303	0,883	0,107	-0,028
B4	0,187	0,733	0,243	-0,077
B8	0,262	0,794	0,221	0,132
B10	0,027	0,348	0,505	0,451
B11	-0,160	0,660	-0,035	0,371
C2	0,378	0,797	0,158	-0,119
D1	0,136	0,185	0,894	-0,060
D2	0,034	0,049	0,903	0,002
D3	0,080	0,163	0,708	0,255
E1	0,121	-0,031	0,110	0,819
F2	0,809	0,123	0,093	0,136
F3	0,838	0,106	0,119	-0,004
F4	0,892	0,021	0,055	0,056
F5	0,858	0,128	0,041	-0,084
F6	0,796	0,334	0,084	0,093
F7	0,778	0,347	0,000	0,023
% of Varianza	28,317	21,737	16,019	7,233
% Cumulata	28,317	50,054	66,073	73,306

Facoltà 3

Osservando la tabella sottostante (Tabella 4.4) si evidenzia che il primo fattore sintetizza gli aspetti della qualità della didattica, connessi al singolo insegnamento, che riguardano sia gli item riguardanti la dimensione ‘informazioni’ (B3, B4, B8), sia gli item riguardanti la ‘capacità del docente’ (F6, F7). L’item C2 con un coefficiente pari a 0,873 è fortemente correlato, come era ovvio attendersi, al primo fattore.

Il secondo fattore è legato agli item sulla ‘puntualità del docente’ (F2, F3, F4). L’item F5 (“Il docente è disponibile alle richieste di chiarimenti durante le lezioni?”) presenta uno scarto modesto tra i pesi fattoriali del primo e del secondo fattore, e dipende da ambedue le dimensioni.

Il terzo fattore individua la dimensione ‘organizzazione’ (D1, D2, D3) del corso di studi al quale si aggiunge un aspetto organizzativo del singolo insegnamento (B10).

Il quarto fattore è specifico per l’item E1 che sintetizza le ‘infrastrutture’.

L'item B11 (“L’insegnamento ha contenuti coordinati con gli altri insegnamenti?”), presenta quasi lo stesso peso fattoriale per il primo e il quarto fattore. Concettualmente ha senso legare questo item con la dimensione relativa al singolo ‘insegnamento’ piuttosto che con la dimensione ‘infrastrutture’ identificata dal singolo item E1.

La percentuale di varianza spiegata da questi 4 fattori, che hanno autovalore maggiore di 1, è il 71%.

Tabella 4.4 - Factor pattern ruotato: facoltà 3

Item	Componente			
	1	2	3	4
B3	0,867	0,201	0,169	0,106
B4	0,663	0,203	0,257	0,231
B8	0,514	0,219	0,370	0,248
B10	0,308	0,076	0,663	-0,003
B11	0,407	-0,145	-0,147	0,452
C2	0,863	0,260	0,226	-0,012
D1	0,031	0,148	0,860	0,056
D2	0,129	-0,021	0,832	-0,039
D3	0,079	0,216	0,685	0,364
E1	-0,055	0,174	0,221	0,836
F2	0,272	0,778	0,109	0,069
F3	0,190	0,874	0,117	-0,003
F4	0,366	0,781	0,109	0,143
F5	0,578	0,598	0,116	0,022
F6	0,864	0,319	0,065	-0,066
F7	0,850	0,340	0,056	-0,051
% of Varianza	28,343	17,981	17,197	7,485
% Cumulata	28,343	46,324	63,521	71,005

Facoltà 4

L'analisi fattoriale condotta sugli indicatori della facoltà 4 (Tabella 4.5) mostra risultati simili a quelli ottenuti nella facoltà 3. In particolare la prima dimensione è sempre quella del singolo ‘insegnamento’ individuata, principalmente dagli item relativi alle ‘informazioni’ (B3, B4, B8) e da quelli attinenti la ‘capacità del docente’ (F5, F6, F7). L'item C2 risulta sempre fortemente correlato (0,873) con il primo fattore.

Gli item F2, F3, F4 determinano il secondo fattore legato alla ‘puntualità del docente’. La terza dimensione è quella relativa all’‘organizzazione’ (sezione D del questionario) con l’inclusione anche dell’item B10 sul carico di studio dell’insegnamento. Il quarto e il quinto fattore sono specificati rispettivamente dall’item B11 (“l’insegnamento ha contenuti coordinati con gli altri insegnamenti?”) e dall’item E1 (“Le aule in cui si svolgono le lezioni sono adeguate?”) che individuano una dimensione che si può etichettare con ‘programma’ e la dimensione ‘infrastruttura’.

Tabella 4.5 - Factor pattern ruotato: facoltà 4

Item	Componente				
	1	2	3	4	5
B3	0,855	0,167	0,155	0,282	0,019
B4	0,655	0,139	0,132	0,487	0,045
B8	0,812	-0,056	0,177	0,068	-0,016
B10	0,407	-0,234	0,512	0,076	-0,368
B11	0,112	0,072	-0,046	0,913	-0,073
C2	0,873	0,241	0,187	0,060	-0,020
D1	0,191	0,086	0,813	0,029	0,116
D2	0,232	0,089	0,812	-0,212	-0,108
D3	-0,023	0,306	0,642	0,149	0,182
E1	0,070	-0,041	0,096	-0,051	0,929
F2	0,199	0,821	0,077	0,191	0,004
F3	0,150	0,890	0,147	0,038	-0,035
F4	0,514	0,661	0,141	-0,139	0,014
F5	0,717	0,421	0,067	-0,112	0,138
F6	0,865	0,210	0,136	0,067	-0,002
F7	0,893	0,172	0,070	-0,047	-0,035
% of Varianza	32,754	15,213	13,692	8,187	6,790
% Cumulata	32,754	47,967	61,659	69,847	76,637

Facoltà 5

I 4 fattori risultanti dall’applicazione dell’analisi fattoriale alla facoltà 5 spiegano più del 73% della varianza totale (Tabella 4.6). In particolare gli item B3, B4, B8, C2, F5, F6, F7 saturano il primo fattore che viene denominato ‘insegnamento’. Mentre il secondo fattore ha coefficienti elevati con le variabili F2, F3, F4 e può essere definito ‘puntualità del docente’. Il terzo fattore è denominato ‘organizzazione’, infatti è determinato dagli item relativi all’organizzazione

complessiva degli insegnamenti (D1), al carico di studio complessivo degli insegnamenti previsti nel periodo didattico di riferimento (D2) e al carico di studio dell'insegnamento valutato (B10).

Il quarto fattore ha coefficienti elevati con gli item D3 e E1. Il primo evidenzia in che misura l'orario di svolgimento delle lezioni tiene conto dei tempi di spostamento da un'aula/sede all'altra. Il secondo fa riferimento all'adeguatezza delle aule. La loro elevata correlazione con il quarto fattore ci consente, forse con una certa forzatura, di denominare il fattore come 'infrastrutture', anche se concettualmente D3 sembra legato alla dimensione 'organizzazione'.

L'item B11 ha un peso pressoché simile per il primo e l'ultimo fattore e per tale motivo non è stato preso in considerazione.

Tabella 4.6 - Factor pattern ruotato: facoltà 5

Item	Componente			
	1	2	3	4
B3	0,871	0,222	0,150	0,152
B4	0,679	0,265	0,027	0,290
B8	0,694	0,193	0,170	0,051
B10	0,306	0,118	0,717	-0,063
B11	0,517	-0,172	-0,085	0,519
C2	0,861	0,269	0,271	0,072
D1	0,138	0,098	0,742	0,450
D2	0,117	-0,013	0,851	0,222
D3	0,083	0,102	0,248	0,744
E1	0,072	0,272	0,149	0,725
F2	0,252	0,828	0,006	0,169
F3	0,309	0,837	0,092	0,071
F4	0,464	0,711	0,109	0,132
F5	0,692	0,402	0,188	0,092
F6	0,858	0,288	0,187	0,108
F7	0,887	0,192	0,139	0,007
% of Varianza	32,658	15,873	13,221	11,188
% Cumulata	32,658	48,531	61,752	72,940

Facoltà 6

La tabella 4.7 evidenzia 4 fattori derivanti dall'analisi fattoriale che spiegano complessivamente il 69% della varianza. L'interpretazione dei primi due fattori è identica a quella della facoltà 5; il terzo fattore può essere definito come

‘organizzazione’ ed è specificato dagli item D1, D3, e E1. Si è escluso l’item D2 in quanto non si può definire se presenta maggiore correlazione con il fattore 3 o con il fattore 4, poiché la differenza dei pesi per questi due fattori è minima. L’ultimo fattore è completamente specificato dall’item B10 (0,887) è ed quindi interpretabile come ‘carico di studio dell’insegnamento’.

Tabella 4.7 - Factor pattern ruotato: facoltà 6

Item	Componente			
	1	2	3	4
B3	0,866	0,164	0,019	-0,025
B4	0,593	0,490	-0,007	0,034
B8	0,714	0,284	0,218	0,272
B10	0,124	-0,083	-0,121	0,887
B11	0,374	-0,138	-0,257	0,186
C2	0,931	0,157	0,126	0,089
D1	0,092	0,105	0,830	-0,056
D2	0,123	-0,022	0,587	0,574
D3	0,010	-0,121	0,864	-0,110
E1	-0,103	0,334	0,547	0,361
F2	-0,053	0,741	0,026	-0,054
F3	0,339	0,802	0,112	-0,088
F4	0,439	0,719	-0,003	0,066
F5	0,653	0,246	-0,176	0,088
F6	0,909	0,034	0,079	0,084
F7	0,928	0,045	0,020	-0,086
% of Varianza	31,834	14,424	14,215	8,856
% Cumulata	31,834	46,257	60,472	69,329

Facoltà 7

Osservando il pattern factor riportato di seguito (Tabella 4.8), il secondo fattore individua chiaramente la dimensione ‘organizzazione’ dove gli item con peso maggiore sono (B10, D1, D2, D3). L’interpretazione del primo e del terzo fattore è più ambigua; infatti gli item B3 (“Gli obiettivi dell’insegnamento sono stati illustrati in aula in modo chiaro?”) e C2 (“Sei soddisfatto di come è stato svolto questo insegnamento?”) hanno coefficienti di correlazione simili con i fattori in questione. L’item B8, che concettualmente appartiene alla dimensione ‘insegnamento’, ha peso elevato per il terzo fattore, insieme all’item B11 e, anche se con un peso inferiore, all’item E1. Ne consegue che il terzo fattore non è

facilmente interpretabile (insegnamento o infrastruttura?) e che il primo fattore è sempre definito ‘insegnamento ma privo dell’item B8 che lo ha determinato nelle restanti facoltà.

Tabella 4.8 - Factor pattern ruotato: facoltà 7

Item	Componente		
	1	2	3
B3	0,620	0,307	0,604
B4	0,568	0,340	0,457
B8	0,376	0,389	0,630
B10	0,251	0,681	0,127
B11	0,162	0,029	0,757
C2	0,562	0,309	0,576
D1	0,094	0,847	0,122
D2	0,158	0,873	0,048
D3	0,026	0,790	0,310
E1	-0,241	0,448	0,532
F2	0,742	0,030	0,322
F3	0,827	0,008	0,193
F4	0,851	0,133	0,124
F5	0,827	0,154	0,031
F6	0,848	0,101	0,106
F7	0,858	0,120	0,040
% of Varianza	34,023	20,575	15,363
% Cumulata	34,023	54,597	69,960

Facoltà 8

I fattori con autovalore maggiore di 1 sono quattro e spiegano quasi il 68% della varianza complessiva (Tabella 4.9). Il primo fattore, che da solo spiega il 29% della varianza, è saturato dalle variabili B3, B4, B8, C2, F5, F6, F7 e può essere definito ‘insegnamento’, inglobando in sé sia le informazioni preliminari che un insegnamento deve fornire, sia la capacità e la motivazione proprie del docente. La seconda dimensione fa riferimento sempre al singolo insegnamento considerando in particolare la ‘puntualità del docente’ (F2, F3, F4). Il terzo e il quarto fattore individuano aspetti diversi dell’organizzazione. Il primo fa riferimento al ‘carico di studio’ (B10, D2), il secondo all’organizzazione della didattica in termini di ‘orario di svolgimento e aule’ (D1, D3, E1). L’item B11 ha coefficienti bassi con il secondo e il terzo fattore.

Tabella 4.9 - Factor pattern ruotato: facoltà 8

Item	Componente			
	1	2	3	4
B3	0,844	0,243	0,130	0,177
B4	0,629	0,238	-0,025	0,105
B8	0,553	0,191	0,423	0,135
B10	0,253	0,012	0,775	-0,090
B11	0,125	0,223	0,270	0,095
C2	0,861	0,178	0,224	0,094
D1	0,064	0,118	0,438	0,729
D2	0,017	-0,037	0,775	0,336
D3	0,058	0,100	0,322	0,712
E1	0,193	0,037	-0,212	0,679
F2	0,195	0,837	0,041	0,149
F3	0,265	0,883	0,020	0,036
F4	0,352	0,811	0,070	0,021
F5	0,738	0,331	0,093	0,037
F6	0,878	0,146	0,147	0,019
F7	0,902	0,116	0,081	0,074
% of Varianza	29,006	15,890	11,905	10,816
% Cumulata	29,006	44,897	56,801	67,618

Facoltà 9

I risultati dell'analisi fattoriale per questa facoltà si discostano un po' da quelli complessivi delle altre facoltà (Tabella 4.10). Infatti la dimensione 'insegnamento' è scomponibile in tre fattori ognuno dei quali individua un aspetto diverso. Il primo fattore fa riferimento alle 'informazioni' iniziali che ogni corso deve fornire con coefficienti elevati per gli item B3, B4 e B8, ai quali si aggiunge l'item B10 relativo al carico di studio che finora ha caratterizzato l'organizzazione del singolo insegnamento. Da sottolineare è il peso elevato per questo fattore dell'item C2 che riassume la soddisfazione globale dell'insegnamento valutato.

Il secondo fattore cattura l'aspetto del singolo insegnamento legato alla 'capacità del docente' nell'espone in modo chiaro gli argomenti (F7), nel motivare l'interesse verso la disciplina (F6) e nell'essere disponibili per chiarimenti (F5). Il quarto fattore è correlato con gli aspetti riguardanti la 'puntualità del docente' (F2, F3, F4). Il terzo fattore ha coefficienti elevati con gli item della sezione D e, anche se misura inferiore con l'item E1, e può essere definito come 'organizzazione'.

Tabella 4.10 Factor pattern ruotato: facoltà 9

Item	Componente			
	1	2	3	4
B3	0,777	0,422	0,055	0,159
B4	0,689	0,217	0,080	0,148
B8	0,810	0,298	0,078	0,160
B10	0,669	-0,022	0,242	0,069
B11	0,226	0,364	0,171	-0,217
C2	0,754	0,458	0,035	0,193
D1	0,303	0,022	0,848	0,106
D2	0,303	-0,052	0,761	0,149
D3	0,016	0,064	0,858	-0,039
E1	-0,089	0,183	0,581	-0,128
F2	0,270	0,049	-0,007	0,820
F3	0,129	0,219	0,001	0,872
F4	0,130	0,469	0,051	0,741
F5	0,135	0,801	0,112	0,352
F6	0,366	0,811	0,050	0,264
F7	0,357	0,797	0,018	0,221
% of Varianza	21,128	18,150	15,596	15,240
% Cumulata	21,128	39,277	54,873	70,113

Facoltà 10

La tabella sottostante (Tabella 4.11) individua 4 fattori. Il primo individua gli aspetti della qualità della didattica legati al singolo ‘insegnamento’: la ‘capacità del docente’ (F5, F6, F7), la sua ‘puntualità’ (F2, F3, F4) e la sua ‘chiarezza’ nell’illustrare gli obiettivi formativi (B3). Il secondo fattore non è facilmente interpretabile in quanto ha coefficienti elevati con B8 e B10 che si riferiscono rispettivamente al materiale didattico e al carico di studio dell’insegnamento valutato.

Il terzo fattore è definito ‘organizzazione’ dei contenuti del corso di studi (D1, D2, D3); il quarto è determinato sia dall’item E1 relativo all’adeguatezza delle aule dove si svolgono le lezioni, sia dall’item B11 che fa riferimento al coordinamento dell’insegnamento con gli altri insegnamenti: tali item non sembrano avere un collegamento dal punto di vista concettuale.

Tabella 4.11 - Factor pattern ruotato: facoltà 10

Item	Componente			
	1	2	3	4
B3	0,657	0,490	-0,015	0,339
B4	0,435	0,421	0,098	0,186
B8	0,157	0,684	0,047	0,369
B10	0,118	0,793	0,180	-0,036
B11	0,312	0,310	0,127	0,647
C2	0,677	0,296	-0,005	0,469
D1	0,125	0,135	0,847	0,258
D2	0,066	0,331	0,797	-0,136
D3	0,123	-0,100	0,751	0,301
E1	-0,166	-0,004	0,365	0,738
F2	0,809	-0,046	0,020	0,020
F3	0,818	-0,183	0,122	-0,074
F4	0,816	0,269	0,144	0,062
F5	0,641	0,398	0,163	-0,274
F6	0,803	0,348	0,035	0,082
F7	0,785	0,271	0,102	0,149
% of Varianza	30,872	14,342	13,659	11,003
% Cumulata	30,872	45,215	58,873	69,876

Facoltà 11

La tabella 4.12 mostra ben 5 fattori che spiegano complessivamente l'81% della varianza. Il primo fattore fa riferimento sempre al singolo 'insegnamento' con pesi elevati per gli item riguardanti le 'informazioni' preliminari (B3, B4, B8), ai quali si aggiunge un solo item relativo al docente ed in particolare alla sua capacità di esporre gli argomenti in modo chiaro (F7). Il secondo fattore ha coefficienti elevati con F3, F4, F5, F6 che riguardano sempre il singolo insegnamento e che inglobano aspetti relativi al docente (puntualità, disponibilità capacità). Gli item D1 e D3 determinano l'aspetto 'organizzativo del corso di studi'. Il fattore 4 è specificato dall'item B10 e si riferisce al carico di studio richiesto dall'insegnamento. L'ultimo fattore è fortemente correlato con gli item D2 ed E1 relativi al 'carico di studio' e all' 'adeguatezza delle aule'. Gli item B11 e F2 non sono stati presi in considerazione in quanto il primo presenta una differenza modesta in termini di peso tra il primo e il terzo fattore e il secondo tra il primo e il secondo fattore.

Tabella 4.12 Factor pattern ruotato: facoltà 11

Item	Componente				
	1	2	3	4	5
B3	0,933	0,101	-0,025	-0,011	0,040
B4	0,755	0,014	0,170	-0,424	-0,009
B8	0,884	0,115	0,082	0,242	-0,173
B10	0,080	0,101	-0,007	0,888	-0,075
B11	0,516	0,083	-0,594	0,330	-0,110
C2	0,765	0,480	-0,191	0,292	-0,118
D1	0,218	0,215	0,753	0,276	0,103
D2	0,016	0,133	0,393	-0,079	0,679
D3	-0,052	-0,021	0,834	-0,110	0,064
E1	-0,085	-0,210	-0,028	-0,023	0,875
F2	0,601	0,568	-0,295	-0,205	0,177
F3	0,384	0,638	0,124	0,295	-0,354
F4	0,236	0,880	0,117	0,066	-0,165
F5	-0,056	0,893	0,075	-0,021	0,046
F6	0,580	0,672	-0,072	0,237	0,062
F7	0,703	0,520	-0,010	0,334	-0,095
% of Varianza	28,329	21,264	12,314	10,115	9,401
% Cumulata	28,329	49,593	61,907	72,023	81,424

Facoltà 12

Il pater factor in tabella 4.13 evidenzia 3 fattori che complessivamente spiegano il 71% della varianza. Così come si è ottenuto per quasi tutte le facoltà il primo fattore può essere definito 'insegnamento'; in particolare è determinato dall'item B3, obiettivi formativi dell'insegnamento, B8, materiale didattico, B10, carico di studio e gli item relativi alla capacità del docente (F5, F6, F7). Il secondo fattore coglie altri aspetti del singolo insegnamento, come le modalità dell'esame (B4) e il suo collegamento in termini di contenuti con altri insegnamenti (B11), oltre che aspetti riguardanti la 'puntualità del docente' (F2, F3, F4). Il terzo fattore ha coefficienti di correlazione elevati con gli item relativi all'organizzazione (D1, D2, D3, E1).

Tabella 4.13 Factor pattern ruotato: facoltà 12

Item	Componente		
	1	2	3
B3	0,820	0,438	0,085
B4	0,414	0,657	0,240
B8	0,707	0,288	0,299
B10	0,612	0,061	0,150
B11	0,287	0,661	0,302
C2	0,814	0,337	0,239
D1	0,323	0,091	0,826
D2	0,244	0,250	0,797
D3	0,178	0,325	0,617
E1	-0,311	-0,001	0,759
F2	0,125	0,840	0,217
F3	0,041	0,761	0,206
F4	0,269	0,832	-0,164
F5	0,832	-0,016	-0,031
F6	0,848	0,331	0,128
F7	0,939	0,187	-0,083
% of Varianza	32,095	22,252	17,132
% Cumulata	32,095	54,347	71,479

Riassumendo i risultati ottenuti dalle diverse analisi fattoriali, si può sottolineare che la qualità della didattica legato al singolo insegnamento è scomponibile in tre aspetti: le informazioni preliminari (B3, B4, B8), la puntualità del docente (F2, F3, F4) e la capacità del docente (F5, F6, F7).

In particolare il primo fattore delle facoltà 3, 4, 5, 6, 8, 12 coglie i primi due aspetti; il primo fattore risultante dall'analisi fattoriale nelle facoltà 2 e 10 è determinato da tutti gli item relativi al docente; per la facoltà 9 il primo fattore è correlato maggiormente con gli item relativi alle informazioni preliminari. Inoltre vi sono alcuni item che non sono correlati sempre con lo stesso fattore. Per esempio, l'item B10 (carico di studio) è un aspetto della didattica legato all'insegnamento che si sta valutando, infatti nelle facoltà 9 e 12 determina il fattore legato al singolo insegnamento; in altre facoltà (2, 3, 4, 5, 7) determina l'aspetto dell'organizzazione: quindi pur essendo un item che si riferisce al singolo insegnamento è in correlazione con l'aspetto organizzativo della didattica

dell'intero corso di studi. Ancora, l'item E1 sintetizza l'aspetto delle infrastrutture, ma nelle facoltà 6, 9 e 12 determina il fattore 'organizzazione'.

Questi risultati spostano la nostra attenzione, nelle analisi che seguiranno, soprattutto sugli aspetti della qualità della didattica legati al singolo insegnamento.

CAPITOLO 5

Le misure di importanza relativa

5.1 Introduzione

Nel capitolo precedente, si sono condotte due analisi di tipo esplorativo. Come detto la prima, lo scaling multidimensionale non metrico, ci ha fornito un'idea delle possibili configurazioni di item in relazione all'item C2; la seconda, l'analisi fattoriale degli indicatori semplici, ci ha permesso di identificare le dimensioni del concetto in studio, evidenziando le correlazioni degli item con i fattori estratti. In particolare, l'analisi fattoriale, ha permesso di identificare quali item sono, all'interno di ciascuna dimensione, più importanti, cioè quelli che spiegano di più la soddisfazione dello studente. La nostra attenzione si è rivolta alla dimensione 'insegnamento' che può essere individuata da tre aspetti diversi: le informazioni preliminari, la puntualità del docente, la capacità del docente.

Il nostro obiettivo, ora è riuscire a individuare il peso in termini di importanza che ciascun item ha nella 'quantificazione' della qualità della didattica in senso stretto. In particolare ciò che si vuole individuare è una 'misura' di capacità predittiva dei singoli item della variabile "soddisfazione complessiva dell'insegnamento" (item C2). Per far ciò facciamo ricorso ad un modello di regressione lineare, la cui variabile predetta è appunto la 'soddisfazione complessiva dell'insegnamento' ed i cui predittori saranno i singoli item così come sono emersi dall'analisi fattoriale. Le variabili sono tutte metriche, visto che tutti gli item sono stati sintetizzati con la [1]. Il problema che rimane tuttavia nella conduzione di una analisi regressiva è quello legato al nostro disegno di indagine che, come già detto, è di tipo osservazionale e i regressori sono correlati in maniera rilevante. Infatti, se tutti i regressori fossero in correlati, il contributo di ciascun regressore si otterrebbe dall'usuale scomposizione di R^2 .

Nel nostro caso diventa difficile scomporre il valore di R^2 nei singoli contributi e l'usuale output del modello di regressione lineare non è adeguato per fornire le misure di importanza relativa di ciascun regressore. È necessario, quindi, fare riferimento a metodi che consentano una corretta scomposizione della varianza spiegata nel rispetto di determinati criteri di scomposizione. Tali metodi sono genericamente riportati in letteratura come 'metriche di importanza relativa' (una breve descrizione di tali metodi è riportata nell'Appendice B). Nella nostra analisi utilizziamo il metodo *PMVD* (Proportional Marginal Variance decomposition).

5.2 La scomposizione di R^2 nelle metriche di importanza relativa

Indichiamo con μ (nel nostro caso è il coefficiente R^2) la misura di performance di un modello statistico Θ , con β il vettore dei parametri stimati e con β^* il vettore dei parametri veri e ignoti. Consideriamo una misura di importanza relativa $\phi_i(\Theta, \mu, \beta)$.

Una opportuna scomposizione di R^2 deve soddisfare i seguenti criteri (Feldmann 2005):

1. Non-negatività.

Tutte le componenti di scomposizione devono essere più grandi o uguali a zero: $\phi_i(\Theta, \mu, \beta) \geq 0$ per la variabile i nel modello.

2. Propria esclusione.

I singoli contributi dovrebbero essere uguali a zero se il vero beta è pari a zero. Se $\beta_i^* = 0$ allora $\phi_i(\Theta, \mu, \beta^*) = 0$.

3. Propria inclusione.

Tutte le variabili che fanno parte del modello dovrebbero ricevere una porzione di scomposizione. Se $\beta_i^* \neq 0$ allora $\phi_i(\Theta, \mu, \beta^*) \neq 0$.

4. Contributo totale.

La varianza del modello deve essere scomposta in tanti contributi, la cui somma deve restituire il valore della misura di performance del modello.

Se indichiamo con S un insieme di variabili tale che, per una variabile i in S e un'altra variabile j che non appartiene a S , $\sigma_{ij} = 0$, allora

$\sum_{i \in S} \phi_i(\theta, \mu, \beta^*) = w(S)$, dove $w(S)$ è il contributo marginale congiunto alla performance del modello (contributo in termini di varianza spiegata dalle variabili del modello).

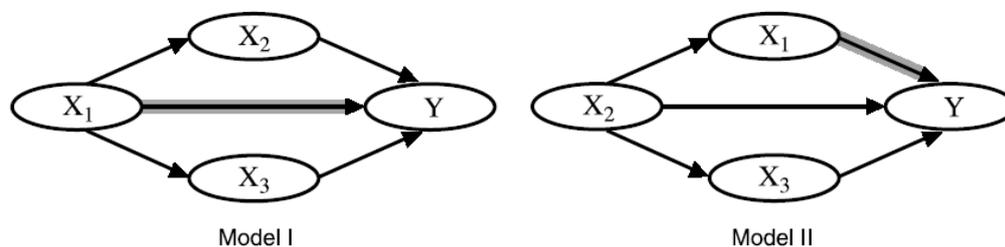
La *non-negatività* è il criterio di base. L'importanza relativa può essere pensata come una misura dell'informazione relativa attribuita al modello da ciascuna variabile. L'informazione è intrinsecamente non-negativa.

Per ciò che concerne il criterio della *propria esclusione* (e conseguentemente della *propria inclusione*), Feldmann (2005) ritiene molto importante tale proprietà poiché egli considera un regressore con coefficiente nullo come spurio. Quando si pensa alla rilevanza predittiva, un regressore con coefficiente uguale a zero nell'equazione lineare non fornisce nessun tipo di contributo, dato che tutti i regressori con coefficiente diverso da zero sono disponibili, cosicché l'esclusione è ragionevole. Se parliamo di importanza relativa e di interpretazione causale e i regressori sono correlati, l'esclusione è un criterio meno convincente.

Facciamo un esempio. La figura sottostante (Figura 5.1) mostra i grafici di due modelli causali (le frecce indicano la relazione causale diretta) derivanti dallo stesso modello di regressione con 3 regressori correlati. Nel primo modello il regressore X_1 influenza direttamente gli altri due regressori e anche la variabile risposta Y . Se provassimo ad eliminare la freccia ombreggiata, il coefficiente β_1 diventerebbe uguale a zero. Nonostante ciò, X_1 esercita un'influenza su Y attraverso gli altri due regressori, e non c'è nessun motivo che porta ad attribuire un contributo nullo. Nel secondo modello, X_2 e X_1 hanno ruoli invertiti. Se la freccia ombreggiata viene eliminata, il coefficiente β_1 diventa zero. In questo caso X_1 non esercita nessun influenza attraverso gli altri due regressori ed è quindi più ragionevole che X_1 attribuisca un contributo nullo.

Il criterio del *contributo totale*, in inglese *full contribution*, è l'essenza dell'importanza relativa. È l'unico criterio che non può essere violato dalle misure statistiche significative. Il *full contribution* richiede anche l'efficienza dello stimatore di importanza relativa. Il potere esplicativo del modello completo basato su tutte le variabili, deve essere diviso tra le variabile stesse.

Figura 5.1 - Due modelli di regressione a confronto



In letteratura è ampio il dibattito su quali siano i criteri cogenti: perché la scomposizione di R^2 sia correttamente condotta, nel nostro caso, tuttavia, non riteniamo opportuno approfondire ulteriormente tale aspetto poiché l'obiettivo che si intende perseguire con l'analisi regressiva è di tipo predittivo. Ciò ha diretto la nostra scelta verso la metrica *PMVD* che appare soddisfare i criteri sopra enunciati.

5.3 La metrica *PMVD*

In questo paragrafo è descritta la metrica *PMVD*. Tale metrica è stata preferita rispetto alle altre presenti in letteratura (Appendice B), in quanto oltre a soddisfare tutti i criteri esposti precedentemente, è anche quella più consona agli obiettivi del presente lavoro.

L'approccio considerato per *PMVD* è basato sui valori degli incrementi sequenziali di R^2 . Per eliminare la dipendenza dell'incremento da un particolare ordinamento dei regressori si effettua, per ciascun regressore, una media ponderata degli incrementi di tutti i possibili ordinamenti.

Per descrivere tale metrica, riportiamo la seguente notazione.

Il coefficiente di determinazione R^2 con i regressori appartenenti all'insieme S, è dato dal rapporto tra devianza dovuta alla regressione e devianza totale.

Il valore di R^2 che si ottiene aggiungendo i regressori, appartenenti ad un'insieme M, al modello con i regressori nell'insieme S è dato da:

$$seqR^2(M | S) = R^2(M \cup S) - R^2(S) \quad [12]$$

L'ordine dei regressori in un modello è una permutazione dei regressori considerati, e si indica con $r = (r_1, \dots, r_p)$.

Indichiamo con $S_k(r)$ l'insieme dei regressori che entrano nel modello prima del regressore X_k secondo l'ordinamento r .

La porzione di R^2 allocata dal regressore X_k in base a r può essere scritta come segue:

$$seqR^2(\{x_k\} | S_k(r)) = R^2(\{x_k\} \cup S_k(r)) - R^2(S_k(r)) \quad [13]$$

Per costruzione, *PMVD* garantisce che ad un regressore con coefficiente stimato nullo è assegnata un'importanza relativa pari a zero, e che il contributo tende a zero asintoticamente se il vero coefficiente è nullo. *PMVD* può essere visto come una media su tutte le permutazioni dei regressori, ma con pesi *data-dependent* per ciascun ordine.

Si riporta di seguito l'espressione di tale metrica:

$$PMVD(x_k) = \frac{1}{p!} \sum_{r=1}^{p!} p(r) seqR^2(\{x_k\} | r) \quad [14]$$

dove $p(r)$ indica i pesi che dipendono dai dati.

In questo caso se tutti i regressori hanno coefficienti diversi da zero, la permutazione r riceve un peso che è proporzionale a

$$L(r) = \prod_{i=1}^{p-1} seqR^2(\{x_{r_{i+1}}, \dots, x_{r_p}\} | \{x_{r_1}, \dots, x_{r_i}\})^{-1} \quad [15]$$

da ciò

$$p(r) = \frac{L(r)}{\sum_r L(r)} \quad [16]$$

dove $p(r)$ è la probabilità associata all'ordine r ed $L(r)$ è la verosimiglianza che l'ordine considerato sia quello effettivo di importanza dei regressori.

Osservando la [15] si comprende come i fattori crescono al crescere di i . $L(r)$ sarà particolarmente grande, se il primo regressore cattura già un grande ammontare di varianza, o se l'ultimo regressore cattura solo una piccola porzione di varianza. In altre parole, la metrica *PMVD* attribuisce un peso maggiore a quelle permutazioni in cui pochi regressori spiegano una porzione elevata di varianza, rendendo il metodo *PMVD* più parsimonioso rispetto per esempio a quello *LMG* (vedi Appendice B).

Il metodo *PMVD*, rispetta i 4 criteri descritti precedentemente, ed in particolare rispetta il criterio dell'esclusione propria, in virtù del quale se un regressore ha coefficiente nullo avrà anche peso nullo, sebbene ciò può comportare una variabilità maggiore della misura assegnata al regressore.

Poiché la metrica *PMVD* si considera più esaustiva dal punto di vista della soddisfazione dei criteri, nell'analisi dei nostri dati la useremo applicata ad un modello di regressione che ha come variabili esplicative alcuni indicatori semplici.

5.4 I pesi *PMVD* degli indicatori di item

In questa fase dello studio si sono implementati 11 modelli di regressione (Tabella 5.1) con variabile spiegata I_{C2} , ovvero l'indicatore semplice relativo all'item C2 e con variabili esplicative gli indicatori semplici di quegli item maggiormente correlati con la dimensione 'insegnamento' considerata nei suoi diversi aspetti. I modelli regressivi cambiano da facoltà a facoltà in relazione ai diversi item che più determinano (si vedano le matrici di pesi fattoriali) il fattore 'insegnamento'.

Per ogni regressione si mostrano le stime dei coefficienti, la significatività dei regressori e i valori dei pesi *PMVD*.

Inoltre sono state, preliminarmente, verificate le ipotesi sulla regressione (multicollinearità, linearità, normalità, omoschedasticità) che si possono accettare per tutti i modelli considerati.

Tabella 5.1 - I modelli regressivi

Facoltà	Modelli regressivi
2	$I_{C2} = \beta_0 + \beta_1 I_{B3} + \beta_2 I_{B4} + \beta_3 I_{B8} + \beta_4 I_{B11} + \varepsilon$
3	$I_{C2} = \beta_0 + \beta_1 I_{B3} + \beta_2 I_{B4} + \beta_3 I_{B8} + \beta_4 I_{F6} + \beta_5 I_{F7} + \varepsilon$
4	$I_{C2} = \beta_0 + \beta_1 I_{B3} + \beta_2 I_{B4} + \beta_3 I_{B8} + \beta_4 I_{F5} + \beta_5 I_{F6} + \beta_6 I_{F7} + \varepsilon$
5	$I_{C2} = \beta_0 + \beta_1 I_{B3} + \beta_2 I_{B4} + \beta_3 I_{B8} + \beta_4 I_{F5} + \beta_5 I_{F6} + \beta_6 I_{F7} + \varepsilon$
6	$I_{C2} = \beta_0 + \beta_1 I_{B3} + \beta_2 I_{B4} + \beta_3 I_{B8} + \beta_4 I_{F5} + \beta_5 I_{F6} + \beta_6 I_{F7} + \varepsilon$
7	$I_{C2} = \beta_0 + \beta_1 I_{B3} + \beta_2 I_{B4} + \beta_3 I_{F2} + \beta_4 I_{F3} + \beta_5 I_{F4} + \beta_6 I_{F5} + \beta_7 I_{F6} + \beta_8 I_{F7} + \varepsilon$
8	$I_{C2} = \beta_0 + \beta_1 I_{B3} + \beta_2 I_{B4} + \beta_3 I_{B8} + \beta_4 I_{F5} + \beta_5 I_{F6} + \beta_6 I_{F7} + \varepsilon$
9	$I_{C2} = \beta_0 + \beta_1 I_{B3} + \beta_2 I_{B4} + \beta_3 I_{B8} + \beta_4 I_{B10} + \varepsilon$
10	$I_{C2} = \beta_0 + \beta_1 I_{B3} + \beta_2 I_{F2} + \beta_3 I_{F3} + \beta_4 I_{F4} + \beta_5 I_{F5} + \beta_6 I_{F6} + \beta_7 I_{F7} + \varepsilon$
11	$I_{C2} = \beta_0 + \beta_1 I_{B3} + \beta_2 I_{B4} + \beta_3 I_{B8} + \beta_4 I_{F7} + \varepsilon$
12	$I_{C2} = \beta_0 + \beta_1 I_{B3} + \beta_2 I_{B8} + \beta_3 I_{F5} + \beta_4 I_{F6} + \beta_5 I_{F7} + \varepsilon$

I risultati dei modelli di regressione sono riportati nella tabella 5.2. La figura 5.2 mostra gli item che più spiegano, in termini di importanza relativa, la soddisfazione dello studente e la tabella 5.3 mostra i pesi ‘significativi’ per le diverse facoltà. Alcuni di questi pesi sono bassi ma non si può negare la loro importanza nel valutare la soddisfazione dell’insegnamento.

Tabella 5.2 - Stime con il metodo dei minimi quadrati dei parametri dei modelli regressivi e dei pesi PMVD

Facoltà 2

Item	Stima	Errore Std.	t	p-valore	PMVD
B3	0,839	0,076	10,974	0,000	0,966
B4	0,030	0,055	0,541	0,589	0,001
B8	0,169	0,074	2,299	0,022	0,029
B11	-0,054	0,039	-1,399	0,163	0,003
R² =72,59%					

Facoltà 3

Item	Stima	Errore Std.	t	p-valore	PMVD
B3	0,404	0,051	7,900	0,000	0,371
B4	-0,037	0,032	-1,130	0,259	0,002
B8	0,217	0,031	6,909	0,000	0,060
F6	0,162	0,052	3,103	0,002	0,081
F7	0,339	0,048	7,012	0,000	0,487
R² =82,73%					

Facoltà 4

Item	Stima	Errore Std.	t	p-valore	PMVD
B3	0,449	0,090	4,991	0,000	0,441
B4	-0,050	0,053	-0,952	0,343	0,003
B8	0,122	0,061	1,998	0,048	0,025
F5	-0,037	0,089	-0,415	0,679	0,001
F6	0,343	0,072	4,728	0,000	0,354
F7	0,189	0,067	2,833	0,005	0,177

R² =82,23%

Facoltà 5

Item	Stima	Errore Std.	t	p-valore	PMVD
B3	0,398	0,053	7,548	0,000	0,371
B4	-0,028	0,035	-0,780	0,436	0,001
B8	0,158	0,031	5,127	0,000	0,036
F5	0,025	0,060	0,418	0,676	0,000
F6	0,313	0,048	6,595	0,000	0,380
F7	0,227	0,040	5,710	0,000	0,212

R² =85,51%

Facoltà 6

Item	Stima	Errore Std.	t	p-valore	PMVD
B3	0,275	0,124	2,223	0,031	0,150
B4	0,047	0,088	0,529	0,599	0,004
B8	0,204	0,089	2,289	0,027	0,097
F5	0,150	0,105	1,437	0,158	0,015
F6	0,205	0,103	1,989	0,053	0,179
F7	0,346	0,105	3,285	0,002	0,557

R² =89,50%

Facoltà 7

Item	Stima	Errore Std.	t	p-valore	PMVD
B3	0,893	0,087	10,315	<2e-16	0,961
B4	0,119	0,070	1,710	0,090	0,018
F2	0,017	0,071	0,244	0,808	0,001
F3	0,011	0,090	0,124	0,901	0,000
F4	0,020	0,124	0,160	0,874	0,000
F5	-0,200	0,115	-1,731	0,086	0,008
F6	0,197	0,093	2,117	0,036	0,008
F7	-0,103	0,090	-1,145	0,255	0,004

R² =80,49%

Facoltà 8

Item	Stima	Errore Std.	t	p-valore	PMVD
B3	0,369	0,068	5,467	0,000	0,411
B4	0,062	0,034	1,802	0,073	0,006
B8	0,184	0,048	3,820	0,000	0,031
F5	-0,154	0,090	-1,713	0,088	0,005
F6	0,225	0,062	3,652	0,000	0,129
F7	0,346	0,064	5,400	0,000	0,419

R² =78,39%

Facoltà 9

Item	Stima	Errore Std.	t	p-valore	PMVD
B3	0,763	0,061	12,458	0,000	0,854
B4	0,056	0,041	1,374	0,171	0,005
B8	0,342	0,065	5,284	0,000	0,141
B10	-0,009	0,034	-0,258	0,796	0,000

R² =75,22%

Facoltà 10

Item	Stima	Errore Std.	t	p-valore	PMVD
B3	0,602	0,121	4,967	0,000	0,680
F2	-0,131	0,072	-1,826	0,072	0,016
F3	0,163	0,106	1,547	0,127	0,037
F4	0,330	0,176	1,871	0,066	0,059
F5	-0,459	0,166	-2,759	0,008	0,035
F6	0,285	0,134	2,128	0,037	0,158
F7	0,084	0,142	0,596	0,554	0,014

R² =72,29%

Facoltà 11

Item	Stima	Errore Std.	t	p-valore	PMVD
B3	0,200	0,233	0,858	0,403	0,064
B4	0,016	0,173	0,092	0,928	0,000
B8	0,122	0,268	0,456	0,655	0,019
F7	0,693	0,158	4,377	0,000	0,916

R² =85,37%

Facoltà 12

Item	Stima	Errore Std.	t	p-valore	PMVD
B3	-0,105	0,193	-0,546	0,590	0,009
B8	0,412	0,130	3,177	0,004	0,312
F5	0,162	0,224	0,721	0,479	0,026
F6	0,510	0,201	2,538	0,019	0,639
F7	0,090	0,259	0,346	0,733	0,014
R² =83,65%					

Come si può osservare, per tutti gli 11 modelli, la bontà di adattamento è buona, infatti il valore di R^2 varia tra 0,72, per la facoltà 10, e 0,89, per la facoltà 6.

I pesi ottenuti con il metodo *PMVD* mettono in evidenza l'importanza relativa degli item.

Per quanto riguarda la facoltà 2, il regressore, che più spiega, in termini di importanza relativa, la soddisfazione dello studente è quello relativo all'illustrazione degli obiettivi formativi dello studente (B3). Gli altri pesi sono bassi. Inoltre B3 è l'unico regressore significativo al 5%.

Per la facoltà 3, il primo regressore in termini di importanza relativa è F7 ("Il docente espone gli argomenti in modo chiaro?") con un peso pari a 0,487; il secondo regressore è B3. È importante fare un'osservazione: nonostante i regressori B8 e F6 siano significativi il peso *PMVD* è basso, ciò ci permette di sottolineare come l'attribuzione di pesi di importanza relativa va oltre l'importanza che le stime ottenute dalla regressione forniscono, infatti non sempre le stime *PMVD* mimano le statistiche *t*.

I regressori che, in ordine di importanza, spiegano di più la soddisfazione dello studente nella facoltà 4, sono B3, F6 ("Il docente stimola/motiva l'interesse verso la disciplina?") e F7. Si può osservare come il metodo *PMVD* attribuisca peso

quasi nullo a quei regressori che hanno coefficienti vicini allo zero, ciò nel rispetto del criterio di propria esclusione.

La facoltà 5 mette in evidenza gli stessi regressori della facoltà 4 con il seguente ordine di importanza relativa: F6, B3, F7. Inoltre, nonostante l'item F5 ("Il docente è disponibile alle richieste di chiarimenti durante le lezioni?") sia correlato con l'item C2 (0,73), il metodo *PMVD* attribuisce peso nullo.

L'item che in assoluto risulta più importante nella facoltà 6, e che quindi meglio predice la soddisfazione di uno studente è l'item F7 (0,557), seguito dall'item F6 (0,179) e da B3 (0,150). L'item B8 è correlato con C2 (0,73), ma presenta un peso di importanza relativa pressoché nullo.

Degli otto item che sono stati inseriti nel modello di regressione per la facoltà 7, solamente l'item B3 ha un peso elevato (0,961), che, in questo caso, è l'unico che ha un p-valore pari a zero.

Anche per la facoltà 8 gli item più importante sono in ordine F7, B3 e F6.

Gli item che hanno una maggiore importanza relativa per la facoltà 9 sono in ordine B3 e B8 ("il materiale didattico è adeguato per lo studio della materia?"); il peso di B3 (0,854) è di gran lunga superiore rispetto a quello di B8 (0,141).

L'importanza relativa degli item B3 e F6 è la più elevata rispetto a quelli degli altri item inseriti nel modello di regressione implementato per la facoltà 10.

L'unico item, per la facoltà 11, che ha grande capacità esplicativa è l'item F7 con un peso elevatissimo (0,916).

Infine, per la facoltà 12, gli item che più spiegano la soddisfazione dello studente e che quindi hanno pesi più elevati sono in ordine di importanza F6 e B8. Tale risultato esclude gli item F7 e B3 che presentano alti coefficienti di correlazione con C2 (rispettivamente 0,80 e 0,77).

Tabella 5.3 – I pesi 'significativi'

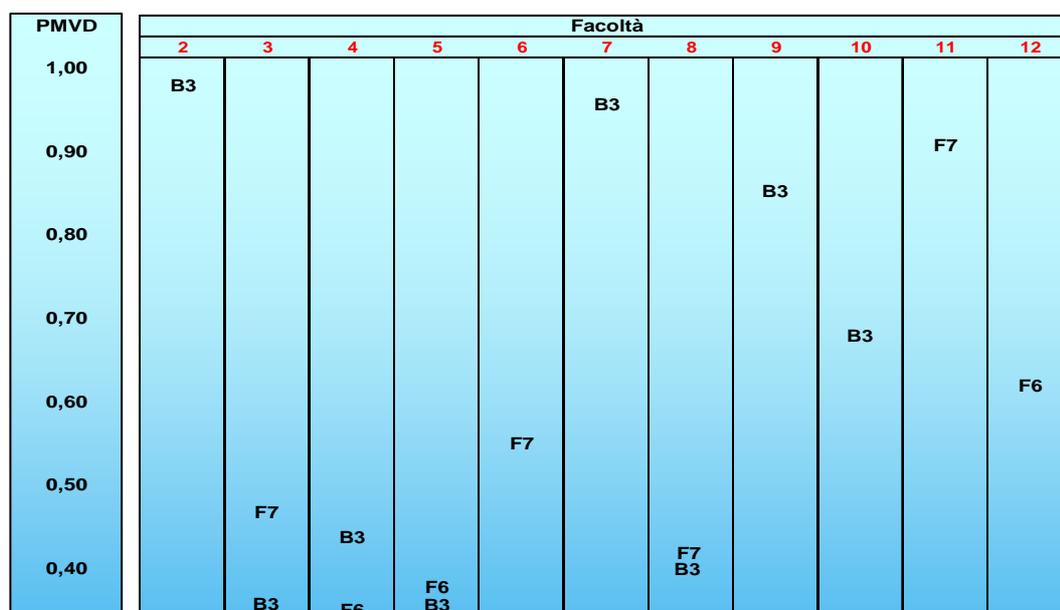
Item	Facoltà										
	2	3	4	5	6	7	8	9	10	11	12
B3	0,996	0,371	0,441	0,371	0,150	0,961	0,411	0,854	0,680		
B4											
B8	0,029	0,060	0,025	0,036	0,097		0,031	0,141			0,312
B10											
B11											
F2											
F3											
F4											
F5							0,005		0,035		74
F6		0,081	0,354	0,380	0,179		0,129		0,158		0,639
F7		0,487	0,177	0,212	0,557		0,419			0,916	

In generale, si può sottolineare che i risultati ottenuti con il metodo *PMVD* sono in linea con quanto emerso nello scaling non metrico applicato ai dati originari. Infatti gli item che presentano un peso *PMVD* elevato sono anche quelli che appaiono più vicini all'item relativo alla soddisfazione complessiva (F02) nella configurazione finale dei punti.

Alcuni risultati meritano una maggiore attenzione. Per esempio l'item B3 relativo alla chiara illustrazione degli obiettivi formativi del corso, è quello che per la maggior parte delle facoltà (escluse la facoltà 11 e 12) ha peso elevato.

Più facile sembra giustificare la grande importanza che hanno gli item F6 e F7: uno studente, probabilmente si sente soddisfatto, quando emergono le capacità del docente; l'esposizione degli argomenti in modo chiaro e anche l'interesse che il docente stimolo verso la disciplina riassumo, da soli, la soddisfazione complessiva. Gli studenti frequentanti i corsi di studio nelle facoltà 9 e 12 ritengono importante anche l'adeguatezza del materiale didattico per lo studio della disciplina (B8). Dai risultati dello scaling forse ci si sarebbe aspettato questo risultato in più facoltà.

Figura 5.2 - I pesi *PMVD* per le 11 facoltà



ALCUNE CONSIDERAZIONI CONCLUSIVE

Al di là degli strumenti statistici utilizzati, quello che appare più rilevante nei risultati ottenuti è la semplificazione a cui si è pervenuti nel definire la qualità

della dimensione strettamente legata alla didattica. È il caso di sottolineare che si è volutamente deciso di sottacere la dimensione organizzativa della ‘qualità della didattica’. Infatti, sono troppe le variabili che intervengono negli aspetti organizzativi sia a livello di corso di studi e, ancor di più, a livello di Facoltà. Inoltre le differenze relative all’organizzazione diventano ancor più marcate in funzione degli aspetti tipicamente formativi che caratterizzano i diversi percorsi di studio: ben diverso è organizzare un laboratorio di discipline tecnico-scientifiche che una classica lezione frontale di discipline umanistiche. È apparso allora più sensato occuparsi di aspetti legati alla didattica avente come perno il docente, dal quale dipende, infatti, la chiarezza espositiva, l’illustrazione degli obiettivi formativi, il suscitare interesse verso la disciplina e così via.

L’omogeneità concettuale della dimensione ‘qualità dell’insegnamento’ ha reso appena accettabile l’aggregazione per Facoltà. Mentre era atteso il forte peso predittivo che gli item F6 (“*Il docente stimola/motiva l’interesse verso la disciplina?*”) e F7 (“*Il docente espone gli argomenti in modo chiaro?*”) hanno nei confronti della ‘soddisfazione’, appare straordinario che l’item B3 relativo agli obiettivi formativi, in talune Facoltà (2,7,9,11), abbia un peso predittivo molto vicino a 1, lasciando intendere che il livello di soddisfazione da parte dello studente sarebbe quasi del tutto spiegato dall’aver fornito gli obiettivi formativi del corso (B3)!

La scelta di un’analisi a livello di facoltà è stata dettata da esigenze del Nucleo di Valutazione. Non si può nascondere che un’analisi più dettagliata potrebbe mettere in luce altri aspetti diversi e soprattutto potrebbe condurre a risoluzioni più mirate dei problemi inerenti la didattica.

Si è tentato di individuare driver della qualità generali per l’intero Ateneo, mediante un’aggregazione per Facoltà che, però, ha probabilmente, condotto a risultati confusi. È sensato pensare che un’aggregazione a livello di Corso di Studi avrebbe messo in risalto maggiori specificità; è sensato pensare che gli studenti stessi possano essere condizionati nell’esprimere opinioni sull’insegnamento da obiettivi formativi comuni, da curricula propri di uno stesso Corso di Studi; è sensato pensare che i Corsi di Studio differiscano tra loro per molteplici aspetti e

analizzarli globalmente a livello di Facoltà può portare a risultati troppo omogenei che nascondono le reali diversità. Diversità non solo riferite all'aspetto organizzativo, ma alla didattica in senso stretto: è il Corso di Studi che gestisce i programmi, che istituisce la commissione didattica, che effettua il coordinamento didattico dei contenuti, che 'sorveglia' l'erogazione della didattica, che può intervenire più celermente per migliorare gli aspetti della didattica più carenti.

È quindi importante evidenziare che ha più senso analizzare un concetto multidimensionale troppo complesso come quello della 'qualità della didattica', considerando come unità di analisi d'aggregazione il Corso di Studi, perché solo entrando dentro il 'mondo' Corso di Studi è possibile controllare, monitorare e migliorare con opportuni mezzi l'erogazione della didattica.

Si suggerisce quindi in futuro di condurre analisi di valutazione della didattica a livello disaggregato allo scopo di migliorare l'erogazione dei servizi agli studenti che sono i veri attori protagonisti del sistema universitario.

Nonostante tutto si possono comunque mettere in risalto alcune somiglianze fra le facoltà nell'opinione degli studenti: l'item F7 è un 'portatore' di qualità in sei facoltà su undici, mentre l'item F6 in sette su undici. Entrambi gli item sono strettamente legati alla capacità del docente, l'uno infatti attiene alla chiarezza e l'altro alla capacità di suscitare l'interesse. Aspetti facilmente riconoscibili come dirimenti nella accezione più stretta di qualità dell'insegnamento.

Anche se qualche risultato ci lascia perplessi, come ad esempio la quantità di varianza spiegata dall'item B3, rimane tuttavia di conforto verificare che gli item individuati nello scalogramma non metrico come quelli più 'vicini' all'item che denota la soddisfazione (C2), siano sovrapponibili a quelli individuati dal metodo *PMVD*, che in più ci consente di quantificarne l'importanza e quindi di trovare un sistema di pesi per l'eventuale costruzione di un indicatore composto di qualità della didattica in senso stretto.

RINGRAZIAMENTI

Ringrazio il dottorando Leonardo Librizzi per il supporto e i numerosi consigli forniti, il collega Salvatore Marcantonio per il sostegno che mi ha offerto.

Grande riconoscenza va alla professoressa Vincenza Capursi per aver contribuito alla mia formazione e crescita offrendomi la possibilità di abbracciare ambiti statistici a me sconosciuti prima della stesura di questa tesi.

Ringrazio con affetto i miei genitori, mio fratello e Tony per la pazienza che hanno avuto nei miei confronti nel periodo di preparazione della tesi.

APPENDICE A

I paradigmi generatori dei dati

A.1 I paradigmi generatori dei dati

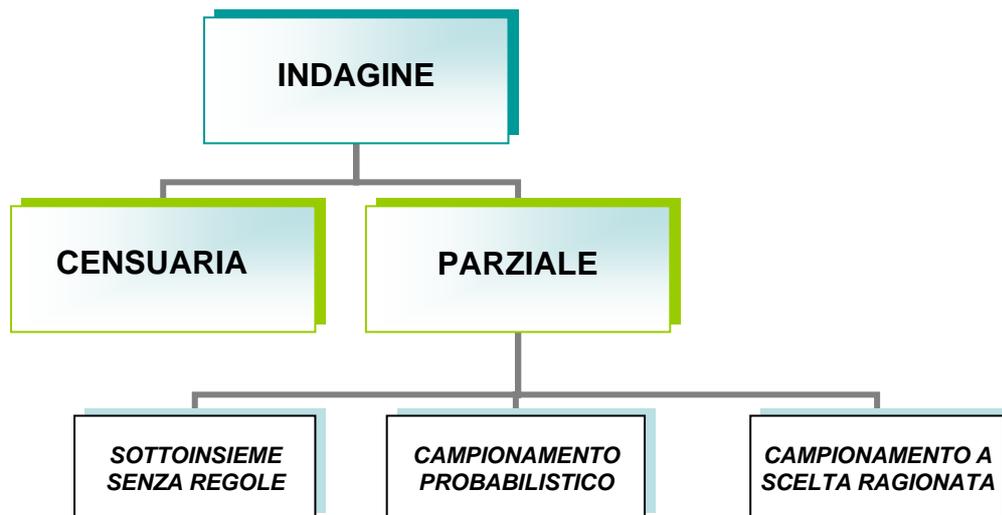
Le elaborazioni condotte in questo lavoro sono state precedute da un'ampia classificazione dei diversi disegni statistici. Ciò allo scopo di individuare le caratteristiche dello studio effettuato, in relazione a ciò che si è osservato e a come è stato osservato. Infatti, prima dell'inizio di ciascuno studio statistico, è necessario non solo fissare e dichiarare in modo implicito gli obiettivi dell'analisi, ma anche, avere chiaro qual è il tipo di studio che si andrà a condurre e conoscere come i dati sono stati generati. O, più specificamente, è necessario conoscere il meccanismo o il *paradigma* sottostante, che ha dato origine ai dati in nostro possesso.

Per ottenere le opportune informazioni sull'evento studiato, attraverso i dati, è necessario prima di tutto stabilire *cosa osservare e come osservare* e poi *come organizzare sistematicamente* le osservazioni effettuate per pervenire ai dati.

Spesso i dati sono il risultato di semplici operazioni di misura di una o più qualità su un insieme di *oggetti* (unità statistiche) sotto osservazione. In questo caso le analisi successive saranno dettate sia dall'obiettivo che ha spinto verso la misurazione, sia dalle proprietà dei dati, associate alle scale di misura che esprimono i dati stessi.

Molto spesso, invece, e soprattutto nelle scienze "*osservazionali*" (quali le Scienze Sociali o Mediche), i dati sono il risultato di una tecnica di raccolta messa a punto per misurare gli oggetti che vengono *scelti* rispetto ad una o più qualità. Tale processo è noto con il nome di *indagine*. È evidente che il modo in cui i dati sono raccolti condiziona l'analisi successiva e la strumentazione statistica che si può legittimamente utilizzare. Riportiamo di seguito uno schema che può sintetizzare, a grandi linee, i possibili tipi d'indagine dalle quali possono scaturire i dati.

Figura A.1 I tipi di indagine



A.2 I tipi di indagine

In questo paragrafo si vogliono descrivere i diversi tipi di indagine, in particolare l'indagine può essere:

- a. *censuaria*, ossia estesa a tutte le unità della popolazione di riferimento (es.: individui di un condominio, o studenti di una facoltà, o popolazione di una nazione);
- b. *parziale*, ossia estesa solo ad una parte, ad un sottoinsieme, della popolazione di riferimento (*parte del tutto*).

Le indagini parziali possono scomporsi secondo i *disegni statistici*, che costituiscono sempre un compromesso tra ciò che è *possibile* e ciò che è *desiderabile* (Kish, 1987).

Il disegno è il modo in cui si ritiene di procedere per acquisire tutte le informazioni in grado di confermare o meno la definizione del problema. Il disegno deve tradursi in una sequenza di azioni strutturate, organiche e coerenti, cioè in una procedura che comprende attività (azioni che coinvolgono l'uso di mezzi e che si svolgono nel tempo) e decisioni (azioni che stabiliscono di volta in volta la direzione da seguire lungo il percorso intrapreso). Al termine di tutte le attività svolte all'interno della procedura adottata si perviene al controllo delle conclusioni.

I compromessi coinvolti nella scelta tra i tipi di disegni si riferiscono ai problemi di base filosofica che tutte le scienze empiriche hanno: come fare inferenza da grandi popolazioni, da universi infiniti, e da sistemi causali provenienti da limitati campioni di osservazioni. Questi problemi e compromessi esistono implicitamente, e qualche volta forniscono giustificazioni vaghe in relazione all'utilizzo di un particolare disegno statistico.

I disegni statistici obbediscono, fondamentalmente, a 3 criteri:

- *realismo*
- *randomizzazione*
- *rappresentatività*

Facciamo un esempio per descrivere questi criteri (Kish, 1987).

Si considerino disegni per confrontare e valutare l'efficacia relativa a due o più tecniche di istruzione nelle scuole.

Inizialmente, compromessi sono necessari per il *realismo* nella scelta delle variabili. I differenti tipi di istruzione per essere confrontati deve essere definiti e 'operazionalizzati'; naturalmente ciò comporta problemi di scelta tra tecniche "ideali" e tecniche fattibili.

In secondo luogo, per l'affidabilità scientifica, dovrebbe essere soddisfatta la *randomizzazione* dei diversi tipi di istruzione per i soggetti, e spesso ciò conduce a grandi difficoltà. Per esempio, si deve considerare la randomizzazione di studenti singoli? O, l'intera classe di studenti? O dobbiamo considerare tutte le scuole come unità di selezione?

Infine, la *rappresentatività* delle unità campionate, di solito, coinvolge compromessi sia per l'individuazione della popolazione obiettivo, sia nella selezione delle unità campionate da essa. Quali età e tipi di studenti dovrebbero essere considerati? Si deve considerare un campione nazionale, o ci si può concentrare solo su un'area più piccola?

Questo esempio, illustra, brevemente, le difficoltà che si incontrano nel soddisfare i tre criteri dei disegni statistici: realismo, randomizzazione e rappresentatività.

E, queste difficoltà sono ancora più grandi nelle ricerche sociali, come per esempio nella valutazione degli effetti della povertà nel lungo periodo, o delle politiche di assistenza sociale.

Molti compromessi devono essere considerati per l'applicazione di questi tre criteri.

Realismo, randomizzazione e rappresentatività mostrano risultati, rispettivamente, negli studi osservazionali, negli studi sperimentali, nelle indagini campionarie.

A.2.1 Gli studi osservazionali

Gli *studi osservazionali* sono indagini scientifiche, in cui né i soggetti in studio né le variabili di interesse sono manipolate dal ricercatore. Il criterio sottostante è il *realismo*, attraverso il quale si tenta di riprodurre oggettivamente le reali condizioni in cui si è manifestato l'evento studiato. Uno studio osservazionale può essere definito semplicemente come un'indagine che non è sperimentale: si effettua una raccolta di dati (accurata e con un opportuno grado di controllo) senza randomizzazione delle unità statistiche. Non si conoscono a priori le caratteristiche dei soggetti inseriti nello studio. Questo tipo di studio non permette, a rigore, di analizzare le relazioni causa-effetto tra le variabili esaminate, ma è molto diffuso nelle indagini sulle relazioni tra variabili (ad esempio, l'effetto di trattamenti clinici), soprattutto quando non si può disporre di un campione probabilistico e non è stato implementato un disegno sperimentale per lo studio in atto.

Le indagini osservazionali si possono distinguere in:

- indagini longitudinali prospettive o di coorte
- indagini longitudinali retrospettive
- indagini cross-sectional o trasversali

Nella tipica *indagine osservazionale cross-sectional*, a un particolare istante di tempo si rilevano più caratteri su un campione di unità. Quindi in dato istante t le unità statistiche selezionate presentano tutte una determinata caratteristica, o sperimentano un certo evento noto. Ad esempio, si può essere interessati a tutti i laureati in un certo anno, o a tutti coloro che condividono una specifica posizione

occupazionale, o a tutti gli studenti che stanno frequentando uno stesso corso durante un semestre specificato, ecc., a prescindere da altre caratteristiche socio-demografiche e dalla distanza temporale che li separa dal momento in cui sono entrati nella particolare classe o stato di interesse (laureato, occupato/disoccupato, sposato, ecc.).

Le statistiche ufficiali più diffuse derivano da indagini trasversali.

Al fine di condurre correttamente un'indagine trasversale è necessario formulare e verificare che siano soddisfatte alcune *assunzioni*, ovvero che:

1. la *composizione della popolazione sia stabile*: ciò è necessario perché la popolazione è intervistata nell'istante t su cose passate, anche se si tratta di un'assunzione *arbitraria*, non effettivamente verificabile;
2. *esistono dati affidabili registrati in tempi precedenti in relazione alle variabili d'interesse*: quest'ipotesi non sempre è vera, perché si possono presentare errori di memoria, oltre che l'*effetto telescoping*; tale fenomeno consiste in una distorsione dovuta ad un'errata collocazione degli eventi nel tempo: di solito si osserva una tendenza a collocare gli eventi accaduti nel passato, nel presente, e non viceversa, soprattutto quando si tratta di avvenimenti negativi.

Negli *studi longitudinali prospettivi* si rileva un campione di unità ed esse vengono seguite nel tempo, il trattamento ricevuto viene registrato, ma l'assegnazione del trattamento non è controllata dal ricercatore. Tutti i soggetti sono seguiti nel tempo (*prospettivamente*) per rilevare informazioni sulle loro caratteristiche, abitudini di vita, o su tutto ciò che possa costituire un'esposizione al fenomeno studiato. Tale processo si protrae nel tempo, sino a quando la quantità delle informazioni raccolte non è ritenuta sufficiente per indagare le relazioni di interesse e per trarre delle conclusioni. Il risultato è la registrazione del numero di soggetti che, in ciascun campione, in un certo punto del tempo, ha sperimentato l'evento e di coloro che non l'hanno sperimentato.

Negli *studi longitudinali retrospettivi*, si rileva un campione di unità e quindi viene esaminata la storia passata di quelle unità per spiegare la risposta osservata

attualmente. Quindi, la rilevazione delle informazioni comincia a partire da un certo anno t e poi si torna indietro di k anni (*lag*) sino a $t-k$.

I soggetti dei campioni sono selezionati tra le categorie della variabile risposta (ad esempio: avere avuto l'infarto); pertanto, al tempo t , il manifestarsi dell'evento studiato è già noto e così il risultato del processo di esposizione ai fattori di rischio, che si assumono legati all'evento e sui quali si indaga. Si conosce, ovviamente, l'ammontare di coloro che hanno sperimentato l'evento e di coloro che non l'hanno sperimentato. Si torna indietro nel tempo, per scoprire se le ipotesi formulate, sia sui fattori di rischio, che sulla loro relazione con l'evento studiato, sono corrette. Tale procedura, è necessaria per sapere quali soggetti sono stati esposti al fattore di rischio e quali no. Un esempio classico è lo *studio caso-controllo* nel quale, al tempo t , si domanda ai soggetti se negli anni precedenti (per un dato numero k di anni) sono stati esposti al fumo, sia passivo che attivo, se hanno assunto farmaci per il controllo della pressione arteriosa oppure l'aspirina con una certa regolarità e così via. Si cerca di ricostruire un percorso temporale che possa chiarire se e come i fattori possono avere giocato un ruolo nell'insorgere dell'infarto, ovvero, si cerca di stabilire se c'è un certo grado di *associazione* tra la variabile risposta ed il fattore di rischio in esame. Questo è più uno studio di natura confermativa; è più rapido da condurre, oltre che più economico, rispetto ad uno studio prospettivo, ma fornisce una minore quantità e qualità di informazioni.

Inoltre, la mancanza di controllo sull'assegnazione dei trattamenti fa sì che negli studi di questo tipo i gruppi di unità non siano mai totalmente comparabili e che quindi le stime dei parametri che si vogliono studiare siano distorte.

Appare quindi evidente, nelle indagini longitudinali, a differenza di quanto avviene nelle indagini trasversali o *cross-section*, l'importanza del tempo, sia perché nel tempo possono cambiare le caratteristiche degli elementi della popolazione oggetto di studio, sia perché può variare la composizione stessa della popolazione. Per queste ragioni, nelle indagini longitudinali il tempo viene considerato una variabile a tutti gli effetti ed influenza la scelta del modello opportuno a seconda della scala su cui viene misurato (discreta o continua).

A.2.2 Le indagini campionarie probabilistiche

Si tratta di indagini in cui è nota *a priori* la probabilità che ha ogni unità della popolazione di essere inserita nel campione. Il criterio sottostante, che ogni campione probabilistico deve soddisfare, è la *rappresentatività* della popolazione studiata: solo in questo caso è possibile e corretto stimare il parametro di interesse, o indagare le relazioni tra le variabili rilevate, e riportare queste informazioni dal campione alla popolazione, attraverso l'inferenza.

E' importante che sia chiaro che il concetto di indagine parziale non coincide con quello di indagine campionaria: il termine "parziale" è inteso come una *parte del tutto*, un sottoinsieme qualsiasi della popolazione; il termine campionario evoca, invece, regole di statistica inferenziale (legate alla rappresentatività), che ci permettono di considerare il collettivo estratto come *una parte per il tutto*. L'indagine campionaria può considerarsi una sottodimensione specifica dell'indagine parziale.

Un'ulteriore distinzione deve essere operata all'interno della classe delle indagini campionarie. Indagini campionarie in senso stretto sono da intendersi esclusivamente quelle probabilistiche, che rispettano il requisito della rappresentatività. Ma altre tipologie di indagini campionarie sono ampiamente diffuse nelle scienze sociali e, pur se non rispondono ai requisiti appena citati, è utile menzionarle. Le indagini campionarie sono essenzialmente dette:

- *campionarie in senso stretto*: quando si estrae un campione rappresentativo della popolazione di riferimento, secondo determinate regole, che consentono in seguito, attraverso l'inferenza, di estendere i risultati campionari all'intera popolazione;
- *senza alcuna regola*: quando si considerano campioni presi a casaccio. In questi casi non si è in grado di fare inferenza.

Anche i campioni possono essere di due tipi: *probabilistici* e *non*. Tra questi ultimi, il tipo più frequente nelle indagini parziali, è il *campione per scelta ragionata*. Questo è molto diffuso nello studio di particolari fenomeni attinenti Scienze quali l'Archeologia, la Medicina, ecc., o quando si è in presenza di una

lista incompleta. In determinate circostanze i *campioni per scelta ragionata* sono l'unica scelta operabile per avere l'opportunità di indagare relazioni che altrimenti non potrebbero essere esaminate, perché non si dispone di un campione probabilistico. Tra le tecniche di campionamento la più usata, per questi fini, è il *campionamento per quote* (si costruisce il campione in modo da rispettare le quote con determinate caratteristiche che si evincono nella popolazione).

Tra i campioni probabilistici più utilizzati si annoverano:

1. il *campione casuale semplice*;
2. il *campione casuale stratificato*;
3. il *campione a grappoli*;
4. il *campione sistematico*.

Questa distinzione è importante per non cadere nell'equivoco di considerare il *campionamento casuale semplice* come unica tecnica di campionamento, anche se questa di solito viene utilizzata come riferimento.

A.2.3 Gli studi sperimentali

Gli *studi sperimentali* (o *esperimenti controllati*): permettono di testare ipotesi, ad esempio, sui risultati e/o le conseguenze di un trattamento clinico o farmacologico. In uno studio sperimentale, le unità statistiche vengono assegnate attraverso un processo casuale ai diversi trattamenti. L'*esperimento ideale* è quello in cui tutte le variabili sono controllate prima che l'esperimento abbia avuto inizio e durante lo svolgersi.

Il criterio sottostante è la *randomizzazione*: senza questa non è né lecito né corretto fare inferenza sulle conclusioni tratte dallo studio sperimentale.

Il termine *controllato* fa riferimento alla suddivisione dei soggetti, coinvolti nello studio, in due gruppi: il gruppo o braccio sperimentale che riceve il trattamento, e il gruppo o braccio di controllo che ne riceve uno diverso o nessun trattamento. I due gruppi devono risultare il più possibile omogenei, almeno per tutte le variabili considerate, e quindi comparabili.

Il termine *sperimentale* indica che le modalità di assegnazione dei soggetti alla popolazione da studiare vengono stabilite dallo sperimentatore (criteri di

inclusione/esclusione). Una volta reclutata la popolazione, si verifica l'effetto di un trattamento (ad esempio, la somministrazione di un farmaco) confrontandolo con l'effetto di un altro diverso trattamento (ad esempio, un altro farmaco, nessun farmaco o un placebo).

Mostriamo, a scopo riassuntivo, lo schema dei disegni statistici presentati con i loro punti di forza e di debolezza:

Figura A.2 I disegni statistici: punti di forza e di debolezza

	PUNTI FORTI	PUNTI DEBOLI
<ul style="list-style-type: none"> • Esperimento 	<ul style="list-style-type: none"> • Controllo delle variabili esplicative (randomizzazione) • Identificazione causa-effetto 	<ul style="list-style-type: none"> • Alto costo • Difficile reclutamento unità • Non rappresentativo di definiti target di popolazione
<ul style="list-style-type: none"> • Studio osservazionale 	<ul style="list-style-type: none"> • Basso costo 	<ul style="list-style-type: none"> • Assenza di controllo • Non rappresentativo di definiti target di popolazione • Distorsione per auto-selezione
<ul style="list-style-type: none"> • Indagine campionaria 	<ul style="list-style-type: none"> • Rappresentatività 	<ul style="list-style-type: none"> • Alto costo • Controllo delle variabili

APPENDICE B

Le metriche di importanza relativa

B.1 I diversi approcci nella storia

Nel 1968 Darlington fornisce un quadro generale degli approcci sull'importanza relativa disponibili in quel periodo, che includevano, per esempio, il così detto “*usefulness*” di un regressore (che è la crescita di R^2 che si ottiene quando un regressore è aggiunto al modello che già includeva tutti gli altri regressori; è equivalente alla statistica t). Altri approcci prendevano in considerazione le correlazioni semplici marginali al quadrato, i coefficienti standardizzati al quadrato e i prodotti di coefficienti standardizzati con correlazioni marginali (Hoffmann, 1960).

Tutti questi approcci conducono agli stessi risultati nel caso di regressori incorrelati, e tutti questi approcci sono stati ampiamente criticati. La necessità era quella di ottenere una scomposizione di R^2 in termini di contributi dei differenti regressori. La proposta di Hoffmann si adatta a questa necessità, ma non è considerata appropriata dalla maggior parte degli autori, principalmente perché qualche contributo può essere negativo. Pratt (1987) fornisce un insieme di criteri attraverso i quali la proposta di Hoffmann sembra possa avere una giustificazione, che ha convinto qualche autore a riconsiderare il metodo, mentre altri (Darlington, 1968; Bring, 1996; Johnson e Lebreton, 2004) considerano questo metodo inappropriato.

Lindeman, Merenda e Gold (1980), proposero di usare somme sequenziali di quadrati provenienti dal modello lineare (*LMG*), che dipendono dall'ordine con cui i regressori sono inseriti nel modello, ottenendo una stima globale effettuata su tutti i possibili ordinamenti dei regressori.

Questa proposta non trova molti consensi soprattutto per due motivi: innanzitutto essa è abbastanza complicata dal punto di vista computazionale; inoltre le sue proprietà non sono ancora del tutto chiare.

Un'altra soluzione è stata avanzata da Feldmann (2005) che introdusse il concetto di “proportional marginal variance decomposition” (*PMVD*), che è simile a *LMG* ma con pesi dipendenti dai dati (Capitolo 5)

Johnson e Lebreton (2004) definiscono l'importanza relativa come “il contributo proporzionale che ciascun predittore dà a R^2 , considerando sia l'effetto diretto, sia il suo effetto combinato con le altre variabili nel modello di regressione”.

Il software R, mediante il pacchetto *relaimpo*, fornisce alcune metriche di importanza relativa oltre alla metrica *PMVD* usata nella nostra analisi.

In particolare *relaimpo* si focalizza sulle metriche di “*dispersion importance*”.

B.2 Le metriche fornite da *relaimpo*

La metrica *first*

Un modo di considerare l'importanza relativa è paragonare, ciò che ciascun regressore da solo è capace di spiegare, cioè confrontare i valori di R^2 che derivano da p modelli di regressione con un unico regressore.

Questi valori univariati di R^2 sono identici alle correlazioni al quadrato dei regressori con la risposta. Questi valori sono disponibili in *relaimpo* sotto il nome di *first*. Se i regressori sono correlati, la somma di questi contributi individuali è spesso più elevata del valore di R^2 derivante dal modello completo con tutti i regressori, ovvero il modello totale spiega meno della somma di tutti i singoli modelli.

$$r^2_{y,x_i} = \left[\frac{\text{cov}(y, x_i)}{\sigma_y \sigma_{x_i}} \right]^2 \quad [\text{b1}]$$

$$\sum_{i=1}^p r^2_{y,x_i} > R^2_{y.12\dots p} \quad i = 1, \dots, p$$

First è la sola metrica che ignora completamente la presenza di altri regressori nel modello senza effettuare alcun aggiustamento. Quindi, non è appropriata per la

scomposizione di R^2 in contributi dei singoli regressori, poiché i contributi non si aggiungono naturalmente a R^2 .

La metrica *last*

Un altro modo di intendere l'importanza relativa è confrontare, quello che ciascun regressore ha la capacità di spiegare rispetto a tutti gli altri regressori che sono disponibili. In questo caso si attribuisce ad ogni regressore l'aumento di R^2 quando il regressore è inserito, nel modello, come ultimo dei p regressori. Questo approccio è implementato in *relaimpo* sotto il nome di *last*. Se i regressori sono correlati, questi contributi non si aggiungono al coefficiente di determinazione totale, ma tipicamente aggiungono meno rispetto a R^2 completo.

I valori di importanza relativa per la metrica *last* sono ottenuti come rapporto tra la somma dei quadrati di tipo III e la somma dei quadrati totale $\sum (y_i - \bar{y})^2$.

In particolare la somma dei quadrati di tipo III si ottiene scartando la somma dei residui al quadrato ottenuta dal modello privo del regressore considerato dalla somma dei residui al quadrato attenuata dal modello completo.

La relazione tra i regressori in termini di *last* è identica alla relazione tra i regressori in termini della statistica t-test. Tale metrica è stata chiamata da Darlington (1986) "usefulness". Anche qui i contributi non scompongono R^2 in modo naturale.

La metrica *betasq*

Un altro approccio per valutare i singoli contributi consiste nel considerare i coefficienti standardizzati. Supponiamo che i regressori con scala da 0 a 100 siano riscalati da 0 a 1. In questo caso i rispettivi coefficienti devono essere moltiplicati per 100. Similmente se la risposta fosse riscalata dalla scala 0–100 alla scala 0-1, tutti i β dovrebbero essere divisi per 100. I coefficienti standardizzati sono introdotti come versioni di scala invarianti dei coefficienti, aggiustati con le deviazioni standard stimate:

$$\hat{\beta}_{i,std} = \hat{\beta}_i \frac{\sqrt{s_{ii}}}{\sqrt{s_{yy}}} \quad [b2]$$

dove s_{kk} = varianza empiriche del regressore x_i

s_{yy} = varianza empiriche della risposta y .

I coefficienti standardizzati al quadrato sono stati proposti come metrica di importanza relativa che è disponibile in *relaimpo* come *betasq*.

Bring (1996) cita un dibattito ancora in corso su considerazioni riguardanti l'uso di tale metrica. Johnson e Lebreton (2004) criticano i coefficienti standardizzati, perché tengono troppo conto dell'effetto diretto. In ogni caso, se l'interesse è nella scomposizione di R^2 , *betasq* non è una metrica appropriata, poiché non fornisce una scomposizione naturale di R^2 .

La metrica *pratt*

Hoffman (1960) propose di moltiplicare i coefficienti standardizzati per la correlazione marginale. Poiché la somma di questi prodotti su tutti i regressori restituisce R^2 , egli considerò questa metrica come una naturale scomposizione del coefficiente di determinazione. Questa proposta è stata criticata, ma ha trovato in Pratt (1987) un suo difensore; per tale motivo è stata inclusa nel pacchetto *relaimpo* sotto il nome di *pratt*.

Spesso è opportuno calcolare la metrica relativizzata d_i come segue:

$$d_i = \beta_{i,std} r_i / R^2 \quad [b3]$$

che indica la proporzione di varianza R^2 , spiegata dalla variabile X_j .

La metrica ha un'importanza relativa significativa e quindi ha senso usarla se è maggiore di $1/2p$, dove p indica il numero di regressori nel modello di regressione,

Sebbene Pratt trovò qualche interessante proprietà, molti ricercatori hanno criticato tale metrica, soprattutto nel contesto di scomposizione di R^2 ; infatti tale metrica può allocare contributi negativi. Pratt concorda con tale critica, ma non nasconde che essa può essere applicata in altre situazioni.

La metrica *lmg*

La metrica LMG si basa su valori di R^2 sequenziali e si concentra sulla dipendenza di tutti i possibili ordinamenti dei regressori effettuando una media sugli ordinamenti attraverso l'utilizzo di semplici medie non ponderate. Per descrivere tali metriche, riportiamo la seguente notazione.

La metrica può essere scritta come segue:

$$LMG(x_k) = \frac{1}{p!} \sum_{r=1}^{p!} seqR^2(\{x_k\} | r) \quad [b4]$$

Gli ordinamenti con lo stesso $S_k(r) = S$ possono essere riassunti in un'unica espressione:

$$LMG(x_k) = \frac{1}{p!} \sum_{S \subseteq \{x_1, \dots, x_p\} \setminus \{x_k\}} n(S)! (p - n(S) - 1)! SeqR^2(\{x_k\} | S) \quad [b5]$$

Christens mostrò che la metrica può anche essere scritta in un modo differente che alcuni ritengono più intuitivo:

$$LMG(x_k) = \frac{1}{p} \sum_{j=0}^{p-1} \left(\sum_{\substack{S \subseteq \{x_1, \dots, x_p\} \setminus \{x_k\} \\ n(S)=j}} \frac{SeqR^2(\{x_k\} | S)}{\binom{p-1}{j}} \right) \quad [b6]$$

Questa formula permette di comprendere come la metrica lmg si basa su una media aritmetica semplice delle somme sequenziali su tutte le permutazioni dei regressori e per questo viene considerata come una metrica i cui pesi sono *data-independent*. Ovvero attribuisce a ciascun regressore un peso pari alla media degli incrementi di R^2 dovuti all'inserimento del regressore considerato, tenendo conto di tutti i possibili ordini di inserimento dati dalle permutazioni dei regressori. Poiché l'incremento di R^2 dipende dai regressori già presenti nel modello, la

misura di importanza relativa viene resa indipendente dal particolare ordine con cui le variabili sono introdotte nel modello.

Il metodo *LMG* produce, rispetto a quello *PMVD* una distribuzione più equa dei pesi così da riflettere l'incertezza della struttura causale, ma non dà una descrizione della stessa.

Inoltre, rispetta tutti i criteri enunciati da Feldman, ma non rispetta quello della propria esclusione, proprio a causa della mancanza di $p(r)$.

APPENDICE C

**Il questionario per la valutazione della
didattica da parte degli studenti
dell'Università degli Studi di Palermo**

Università degli Studi di Palermo

QUESTIONARIO DI VALUTAZIONE DELLA DIDATTICA

PRIMA DELLA COMPILAZIONE LEGGERE ATTENTAMENTE LE ISTRUZIONI RIPORTATE SUL RETRO

INFORMAZIONI GENERALI

Data di compilazione ____/____/____ Anno accademico: ____/____

Primo periodo didattico Secondo periodo didattico

Laurea 1° livello N.O. Laurea specialistica Laurea V.O. Laurea a ciclo unico

Denominazione Corso di Studi: _____

Denominazione Insegnamento⁽¹⁾: _____

N) LO STUDENTE

A1	Età:	<input type="checkbox"/> ≤18 <input type="checkbox"/> 19 <input type="checkbox"/> 20 <input type="checkbox"/> 21 <input type="checkbox"/> 22 <input type="checkbox"/> ≥23
A2	Sesso M = maschio F = femmina	<input type="checkbox"/> M <input type="checkbox"/> F
A3	Scuola secondaria di provenienza	A = Liceo classico D = Liceo scientifico G = Altri licei B = Liceo socio-pedagogico E = Ist. tecnico commerc. H = Ist. tecnico industriale C = Ist. tecnico per geometri F = Ist. professionale I = Altro
A4	Residenza A = In sede B = Fuori sede pendolare C = Fuori sede stanziale	<input type="checkbox"/> A <input type="checkbox"/> B <input type="checkbox"/> C
A5	Anno di corso al quale lo studente è iscritto	In corso <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 Ripetente <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> 5 <input type="checkbox"/> 6 Fuori corso <input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> ≥5
A6	Numero totale di crediti acquisiti alla data della rilevazione	0-30 31-60 61-90 91-120 121-150 151-180 <input type="checkbox"/> 181-210 <input type="checkbox"/> 211-240 <input type="checkbox"/> 241-270 <input type="checkbox"/> 271-300
A7	Insegnamenti frequentati in questo periodo didattico	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 <input type="checkbox"/> ≥5
A8	Attività lavorativa in questo anno accademico A = nessuna B = saltuaria o part-time C = a tempo pieno	<input type="checkbox"/> A <input type="checkbox"/> B <input type="checkbox"/> C

Legenda: 1 = decisamente no 2 = più no che sì 3 = più sì che no 4 = decisamente sì NA = non applicabile

I) L'INSEGNAMENTO

B1	Quante ore di lezione hai frequentato (in percentuale)?	<25% 25-50% 50-75% >75%
B2	Quante ore di esercitazioni hai frequentato (in percentuale)? (se l'insegnamento ⁽¹⁾ non prevede esercitazioni, rispondete non applicabile)	<25% 25-50% 50-75% >75% NA
B3	Gli obiettivi formativi dell'insegnamento ⁽¹⁾ sono stati illustrati in aula in modo chiaro?	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4
B4*	Le modalità dell'esame sono state illustrate in aula in modo chiaro?	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4
B5	L'insegnamento ⁽¹⁾ ha contenuti che si sovrappongono a quelli degli altri insegnamenti ⁽¹⁾ ?	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4
B6*	Le attività didattiche integrative (esercitazioni, laboratori, seminari, ecc...) sono utili ai fini dell'apprendimento? (se non previste attività didattiche integrative, rispondete non applicabile)	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 NA
B7	Le attività didattiche integrative (esercitazioni, laboratori, seminari, ecc...) previste all'interno dell'insegnamento ⁽¹⁾ sono adeguatamente coordinate tra loro? (se non previste attività didattiche integrative, rispondete non applicabile)	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 NA
B8*	Il materiale didattico (indicato o fornito) è adeguato per lo studio della materia?	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4
B9*	Le conoscenze preliminari possedute sono sufficienti per la comprensione degli argomenti trattati?	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4
B10*	Il carico di studio richiesto da questo insegnamento ⁽¹⁾ è proporzionato ai crediti indicati nel piano di studi?	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4
B11	L'insegnamento ⁽¹⁾ ha contenuti coordinati con altri insegnamenti ⁽¹⁾ ?	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4

I) INTERESSE E SODDISFAZIONE

C1*	Sei interessato ai contenuti di questo insegnamento ⁽¹⁾ ? (indipendentemente da come è stato svolto)	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4
C2*	Sei soddisfatto di come è stato svolto questo insegnamento ⁽¹⁾ ?	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4

I) ORGANIZZAZIONE

D1*	L'organizzazione complessiva (sedi, orario, esami, ecc...) degli insegnamenti ⁽¹⁾ ufficialmente previsti in questo periodo didattico è accettabile?	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4
D2*	Il carico di studio complessivo degli insegnamenti ⁽¹⁾ ufficialmente previsti in questo periodo didattico è sostenibile?	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4
D3	L'orario di svolgimento dell'attività didattica tiene conto dei tempi di spostamento fra le sedi/aule didattiche?	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4

I) INFRASTRUTTURE

E1*	Le aule in cui si svolgono le lezioni sono adeguate? (si vede, si sente, si trova posto)	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4
E2*	I locali e le attrezzature per le attività didattiche integrative (esercitazioni, laboratori, seminari, ecc...) sono adeguati? (se non previste attività didattiche integrative, rispondete non applicabile)	<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4 NA

Per la Facoltà di Medicina e Chirurgia si intende Corso Integrato

Responsabile dell'insegnamento/modulo COGNOME: NOME:

Denominazione modulo (solo se previsto):

Scrivere in stampatello e per esteso all'interno degli spazi predisposti

Legenda: 1 = decisamente no 2 = più no che si 3 = più si che no 4 = decisamente si NA = non applicabile

F) RESPONSABILE DELL'INSEGNAMENTO/MODULO

		<50%	50%-90%	>90%	
F1	Quale percentuale delle ore di lezione tra quelle previste per il docente è stata svolta dal docente stesso?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
F2	Nell'impossibilità di svolgere la lezione, il docente avverte con congruo anticipo (o comunque in tempo utile)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> NA
F3	Il docente rispetta l'orario di svolgimento dell'attività didattica previsto dal calendario o concordato con gli studenti?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
F4	Il docente rispetta l'orario previsto per il ricevimento?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
F5*	Il docente è disponibile alle richieste di chiarimenti durante le lezioni?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
F6*	Il docente stimola/motiva l'interesse verso la disciplina?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
F7*	Il docente espone gli argomenti in modo chiaro?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

G) DOMANDE RELATIVE AL MODULO (da compilare solo se l'insegnamento è articolato in moduli)

G1	Il modulo ha contenuti che si sovrappongono a quelli degli altri moduli?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
G2	Le conoscenze preliminari possedute sono sufficienti per la comprensione degli argomenti trattati?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
G3	Sei interessato ai contenuti di questo modulo? (indipendentemente da come è stato svolto)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
G4	Sei soddisfatto di come è stato svolto questo modulo?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

* Le domande contrassegnate con un asterisco compongono un questionario minimo che il Comitato Nazionale per la Valutazione del Sistema Universitario e il Consiglio Nazionale degli Studenti Universitari suggeriscono di adottare, al fine di garantire un'omogenea rilevazione su scala nazionale e assicurare la compatibilità dei dati.

(1) Per la Facoltà di Medicina e Chirurgia si intende Corso Integrato

MARCATO LE CASELLE COSÌ: E NON COSÌ:

ISTRUZIONI PER LA COMPILAZIONE

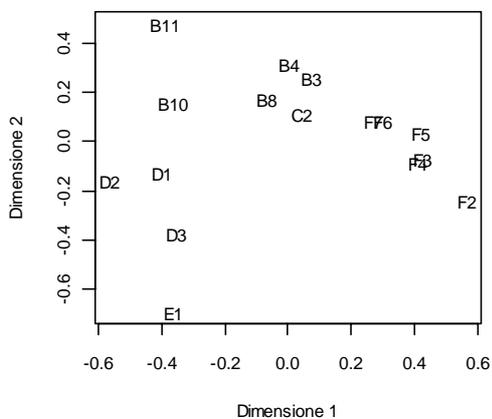
- A) Scrivere esclusivamente con una penna nera o blu
- B) Scrivere in stampatello e per esteso all'interno degli spazi predisposti
- C) Annerire esclusivamente la casella corrispondente alla risposta esatta. Non sono ammesse correzioni di alcun tipo. Saranno considerate errate le risposte per le quali lo studente abbia annerito più caselle o apportato correzioni
- D) Non piegare, sguaiare o macchiare il questionario

SI RICORDA CHE I QUESTIONARI COMPILATI SONO RIGOROSAMENTE ANONIMI E LE INFORMAZIONI CONTENUTE SARANNO ELABORATE E DIFFUSE SOLO IN FORMA AGGREGATA

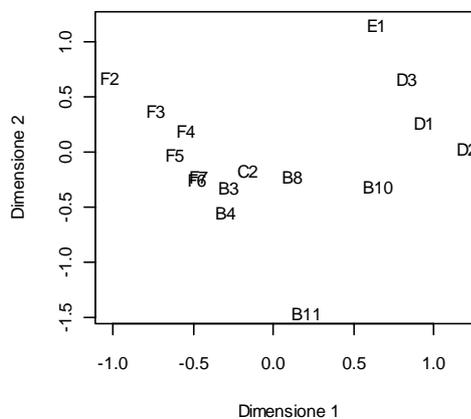
APPENDICE D

**Scaling multidimensionale non metrico:
configurazione degli item nelle
diverse facoltà**

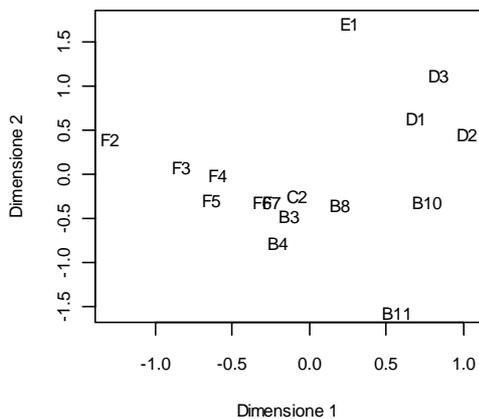
Facoltà 2



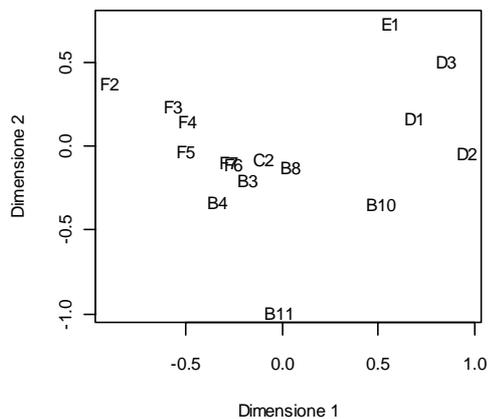
Facoltà 3



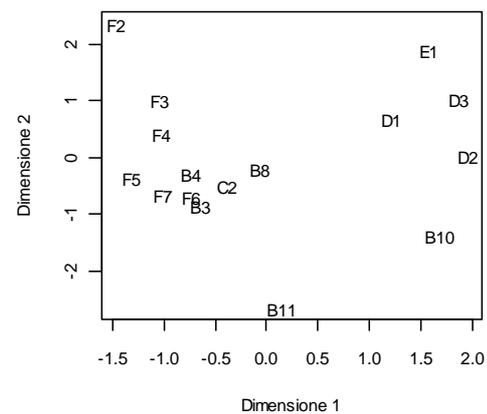
Facoltà 4



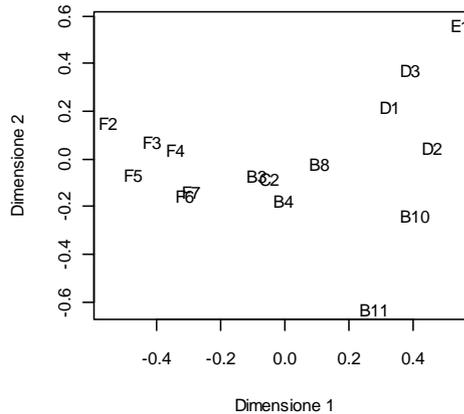
Facoltà 5



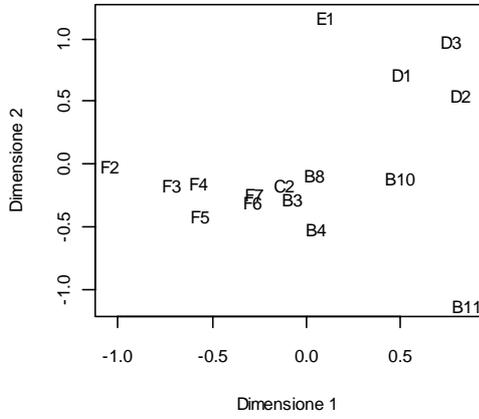
Facoltà 6



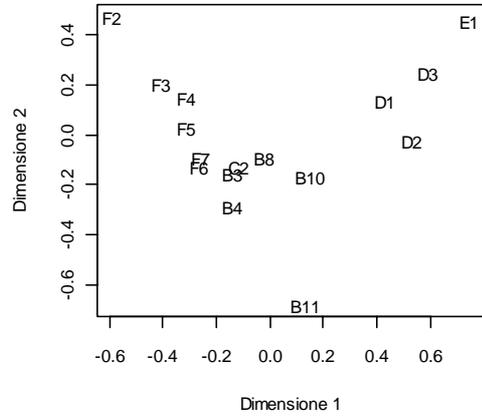
Facoltà 7



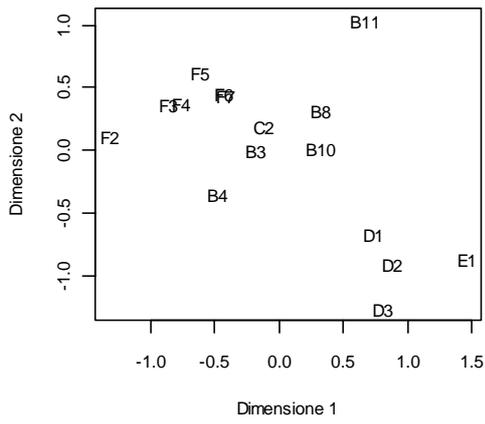
Facoltà 8



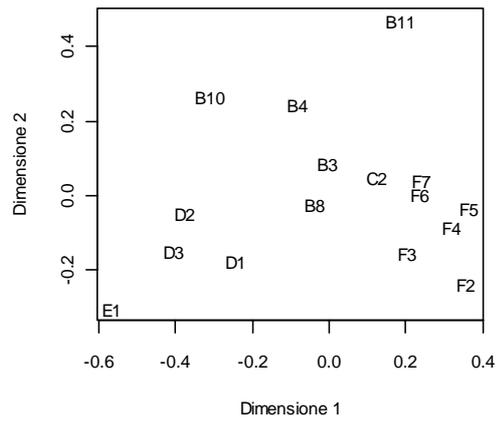
Facoltà 9



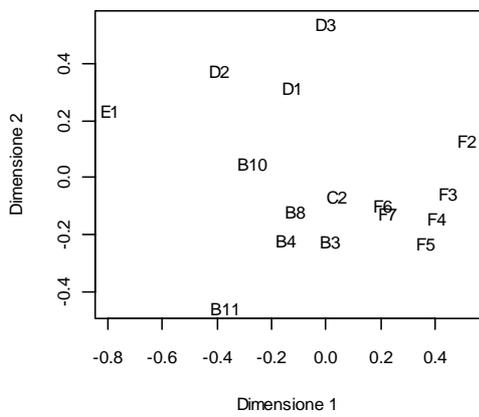
Facoltà 10



Facoltà 11



Facoltà 12



BIBLIOGRAFIA

- Bertin G. (1995), *“Valutazione e sapere sociologico. Metodi e tecniche di gestione dei processi decisionali”*, Franco Angeli, Milano.
- Briglia A. (2001), *“La valutazione della didattica dell’ateneo Modenese”*, Tesi di Laurea in Economia Aziendale, Università degli Studi di Modena e Reggio Emilia, A.A. 2000/2001.
- Capursi V. (2002), *“La valutazione della didattica e dei servizi: un punto di forza di una Politica della Qualità di un ateneo”*, Atti della Giornata di Studio su Valutazione della didattica e dei servizi nel Sistema Università, Università degli Studi di Salerno, 31 maggio, Fisciano, 87-103.
- Capursi V., D’Agata R., Librizzi L. (2008), *“Un indicatore composto della qualità della didattica: dalle scale ordinali alle metriche di importanza relativa”*.
- Capursi V, Librizzi L. (2008), *“Quando è opportuno misurare la qualità della didattica con un indicatore composto?”*
- Cox T. F., Cox M. A. A. (1994) *“Multidimensional scaling”*, Chapman & Hall.
- Davison M. L. (1983), *“Multidimensional scaling”*, John Wiley & Sons, Toronto.
- Dell’Omodarme M. (2007), *“Esercitazioni di Statistica Biomedica, alcune note su R”*, A.A. 2006/2007.
- Delvecchio F. (1992), *“Analisi Statistica di dati multidimensionali”*, Cacucci Editore, Bari.
- Delvecchio F. (1992), *“Statistica per la ricerca sociale”*, Cacucci Editore, Bari.
- Di Ciaccio A., Borra S. (1996), *“Introduzione alla Statistica descrittiva”*, McGraw-Hill, Milano
- Everitt. B. S., Torsten H. (2006), *“Multidimensional Scaling: British Water Voles and Voting in US Congress”*, A Handbook of Statistical Analyses Using R.
- Fabbris L. (1997) *“Statistica multivariata. Analisi esplorativa dei dati”*, McGraw-Hill, Milano.

- Fanciullacci D., Guelfi C., Pennini G. (1991), “*Valutare lo sviluppo. Volume primo. Metodologia e tecnica della valutazione retrospettiva dei progetti*”, Franco Angeli, Milano.
- Feldman B. (2005). “*Relative Importance and Value*” Manuscript version 1.1, 2005-03-19, <http://www.prismanalytics.com/docs/RelativeImportance.pdf>.
- Grömping, U. (2006), “*Relative Importance for Linear Regression in R: The Package relaimpo*”, Journal of Statistical Software 17, jstatsoft.org/v17/i0i/v17i0i.pdf.
- Grömping, U. (2007), “*Estimators of Relative Importance in Linear Regression Based on Variance Decomposition*”, The American Statistician 61, 139-147.
- Kendall M., Gibbons D. (1990), “*Rank correlation Methods*”, Oxford University Press, New York (pag 118).
- L. Kish (1987), “*Statistical Design for research*”, John Wiley & Sons.
- Legge 19 Ottobre 1999, n. 370, “*Disposizioni in materia di università e di ricerca scientifica e tecnologica*”, pubblicata nella *Gazzetta ufficiale* n. 252 del 26 Ottobre 1999, CAPO 1 – DISPOSIZIONI IN MATERIA UNIVERSITARIA, Artt. 1,2.
- Librizzi L. (2003), “*Valutazione della ‘qualità della didattica’: quali misure per quali obiettivi?*”, Tesi di Laurea in scienze Statistiche ed Economiche, Università degli Studi di Palermo, A.A. 2002/2003. London.
- Mardia K. V., Kent J. T., Bibby J. M. (1979), “*Multivariate Analysis*”, Academic Press, Mcgraw-Hill, London.
- Margheri A. (2002), “*Qualità e Customer Satisfaction negli enti locali*”, Edizione Cedam, Padova.
- Mistro D. (2004), “*La customer satisfaction per alcuni enti di formazione professionale del Veneto: analisi, proposte, redazione di un questionario unico*”, Tesi di Laurea in Scienze Statiche ed Economiche, Università degli Studi di Padova, A.A. 2003/2004.
- MIUR – Comitato nazionale per la valutazione del sistema universitario (febbraio 2004), “*Analisi delle relazioni nuclei di valutazione sulle opinioni degli studenti frequentanti in merito alle attività didattiche*”, Doc 05/04.
- MURST – Osservatorio per la valutazione del sistema universitario (luglio 1998), “*Valutazione della didattica da parte degli studenti*”, RdR 1/98.

NdV. – Nucleo di Valutazione (2007), “*Rilevazione sull’opinione degli studenti sulla didattica, A.A. 2006/2007*”, Università degli Studi di Palermo.

Stame M. (2001), “*Approcci e problemi nella valutazione di politiche e di programmi nella realtà del nuovo millennio*”, Atti del Convegno SIS 2001 Processi e Metodi Statistici di Valutazione.

Stame N. (1998), “*La valutazione in Italia: esperienze e prospettive*”, IRRES Gramma, Perugia.

Sulis I. (2007), “*Measuring students’ assessment of ‘university course quality’ using mixed-effects models*”, Tesi di Dottorato di Ricerca in Statistica Applicata, Università degli Studi di Palermo.