



**METODOLOGIE STATISTICHE UTILIZZATE PER
L'ANALISI DELLA SODDISFAZIONE DEGLI STUDENTI
FREQUENTANTI**

A.A. 2007-2008

NUCLEO DI VALUTAZIONE DI ATENEIO
MILANO • PIACENZA - CREMONA • ROMA • BRESCIA • CAMPOBASSO

LUGLIO 2008

Indice

INTRODUZIONE	pag. 3
I) Nota al § 3.1 e al § 4.1 Tabella 1 “Numero questionari compilati dagli studenti per anno di corso”	pag. 4
II) Nota al § 3.2 e al § 4.2 Tabella 2 “Profilo degli studenti iscritti e frequentanti”	pag. 4
III) Nota al § 3.3.1 e al § 4.3.1 Tabella 4 “Giudizio complessivo positivo dell’insegnamento” ripartito per raggruppamenti di studenti presenti in aula nel giorno della rilevazione e per anno di corso	pag. 5
IV) Nota al § 3.3.2 e al § 4.3.2 Tabella 5 “Distribuzione di frequenza percentuale degli items più rilevanti”	pag. 5
V) Nota al § 3.3.3 e al § 4.3.3 Tabella 6 “Posizionamento della Facoltà per sede rispetto ai valori medi delle Sedi Padane per gli item più rilevanti”	pag. 5
VI) Nota al § 3.3.3 e al § 4.3.3 Tabella 19 “Posizionamento delle Classi di laurea per tipo di laurea e numero medio di studenti frequentanti”	pag. 6
VII) Nota al § 3.4.1 e al § 4.4.1 Tabella 9 “Confronto tra la soddisfazione complessiva per gli insegnamenti e l’interesse agli argomenti trattati negli aa.aa. 2006/07 e 2007/08”	pag. 7
VIII) Nota al § 3.4.1 e al § 4.4.1 Tabella 10 “Correlazione di Spearman in ordine decrescente di importanza”	pag. 7
IX) Nota al § 3.4.3 e al § 4.4.3 Tabella 17 “Conoscenze preliminari e anno di corso”	pag. 9
X) Nota al § 3.4.4 e al § 4.4.4 Tabella 18 “Percezione degli studenti circa la corrispondenza fra i Crediti Formativi Universitari, attribuiti a ciascun insegnamento, e il carico di studio complessivo da sostenere”	pag. 9
XI) Nota al § 4.4.5 L’utilizzo delle nuove tecnologie	pag. 9

INTRODUZIONE

Il testo funge da corredo statistico alla relazione sulla valutazione della soddisfazione degli studenti frequentanti e ha lo scopo di illustrare le metodologie utilizzate per l'analisi dei dati.

In particolare viene descritto il contenuto e vengono esposti i criteri con i quali sono stati calcolati i dati delle tabelle contenute nella relazione. Inoltre viene presentato il coefficiente ρ di *Spearman* utilizzato per il calcolo della dipendenza tra gli items.

La relazione e questo fascicolo sulle metodologie statistiche utilizzate sono anche disponibili sulle pagine web della Facoltà dedicate alla "Valutazione della didattica".

Note sull'approfondimento a livello di Facoltà dei risultati emersi

I) Nota al § 3.1 e al § 4.1 Tabella 1 “Numero questionari compilati dagli studenti per anno di corso”

La tabella “Numero questionari compilati dagli studenti per anno di corso” che è stata prodotta avvalendosi dell’item del questionario “*Quanti questionari ha compilato, compreso il presente, nel corrente anno accademico?*”, riporta in funzione del N° di *questionari compilati* da ciascuno studente:

- il n° di *studenti* che hanno compilato i questionari, sia in valori assoluti che in percentuale
tale dato è stato calcolato con la seguente formula:

$$s_i = n_i - n_{i+1}$$

dove:

s_i è il numero di studenti che hanno compilato l’ i -esima quantità di questionari
 n_i è la frequenza assoluta dell’ i -esima modalità di risposta dell’item considerato

- il n° di *questionari compilati complessivi*, con la relativa *percentuale*, ottenuto moltiplicando il N° *questionari compilati* per il n° *studenti*
- il *numero di studenti* che hanno compilato i questionari, con la relativa *percentuale*, diviso per anno di corso

II) Nota al § 3.2 e al § 4.2 Tabella 2 “Profilo degli studenti iscritti e frequentanti”

In questo report sono state affiancate le distribuzioni di frequenza assoluta e relativa degli items del “Profilo studente” dei frequentanti con quelle degli iscritti; il profilo degli studenti frequentanti è stato compilato prevalentemente nei mesi di ottobre-dicembre per le facoltà che hanno corsi trimestrali o semestrali e nel mese di aprile per quelle che hanno corsi annuali, mentre il profilo degli studenti iscritti è aggiornato al 01-06-08.

Per effetto di tale sfasamento temporale, per gli items “Età” e “Media degli esami di profitto”, non vengono proposte le distribuzioni di frequenza assoluta degli iscritti e dei frequentanti in quanto risultano contrastanti.

Nella domanda, “Tipo di diploma” e “Anno di corso”, invece i casi in cui il numero dei frequentanti è maggiore di quello degli iscritti è causato dagli studenti che frequentano gli insegnamenti in anni successivi rispetto a quanto prevede il loro piano studi, ai fini del superamento dell’esame.

Per l’item «Attività lavorativa» invece sono presenti le sole percentuali in quanto i due dati (dei frequentanti e degli iscritti) non sono completamente paragonabili. Infatti gli studenti-lavoratori “Isritti” sono quelli che percepiscono un reddito annuo complessivo superiore a € 9.000, mentre nella categoria studenti-

lavoratori “Frequentanti” rientrano anche coloro che hanno svolto lavori di tipo saltuario.

III) Nota al § 3.3.1 e al § 4.3.1 Tabella 4 “Giudizio complessivo positivo dell’insegnamento” ripartito per raggruppamenti di studenti presenti in aula nel giorno della rilevazione e per anno di corso

Nella tabella “Giudizio complessivo positivo dell’insegnamento” ripartito per raggruppamenti di studenti presenti in aula nel giorno della rilevazione e per anno di corso viene riportato nell’ordine:

- *il numero di insegnamenti* di ciascun raggruppamento divisi in base alla frequenza riscontrata nel giorno della rilevazione
- *il totale di studenti* complessivamente presenti negli insegnamenti di ciascun raggruppamento
- *il totale rispondenti alla domanda* “E’ complessivamente soddisfatto di come è stato svolto questo insegnamento?”
- *la percentuale ed il relativo numero di studenti* che per ogni anno di corso hanno risposto positivamente (“Decisamente SI” e “Più Sì che no”)
- *il valore medio ponderato* delle precedenti percentuali di ciascun raggruppamento, della facoltà e delle classi (se ve ne fossero più di una).

IV) Nota al § 3.3.2 e al § 4.3.2 Tabella 5 “Distribuzione di frequenza percentuale degli item più rilevanti”

Nella tabella “Distribuzione di frequenza percentuale degli items più rilevanti” viene indicato, in termini percentuali, il giudizio espresso dagli studenti rispetto agli item del questionario maggiormente inerenti all’attività didattica, per ciascuna classe di laurea. Per ciascun Item è stato inoltre calcolato l’“Indice di sintesi” che varia tra 2 e 10 ed è ottenuto come media aritmetica delle risposte alle quali sono stati attribuiti i punteggi: 2 alla modalità “Decisamente NO”, 5 alla modalità “più NO che sì”, 7 alla modalità “Più SI che no” e 10 alla modalità “Decisamente SI”.

V) Nota al § 3.3.3 e al § 4.3.3 Tabella 6 “Posizionamento della Facoltà per sede rispetto ai valori medi delle Sedi Padane per gli item più rilevanti”

Nella tabella è stata calcolata, per gli “item più rilevanti”, la “frequenza relativa riscalata” sulle percentuali sommate delle modalità di risposta positive “Decisamente SI” e “Più Sì che no” cioè, tali percentuali sono state rapportate al minimo e al massimo applicando la seguente formula:

$$\text{Frequenza relativa riscalata} = \frac{f_{ij} - \text{Min}(f_j)}{\text{Max}(f_j) - \text{Min}(f_j)}$$

dove:

- f_{ij} è la percentuale di giudizio positivo dell' i -esimo item e della j -esima facoltà
- $\text{Max}(f)$ e $\text{Min}(f)$ sono le percentuali massime e minime degli items della j -esima facoltà.

Il valore in tal modo ottenuto è stato successivamente traslato, tramite una trasformazione lineare, tra 100 e 200 unicamente per facilitarne la lettura.

La frequenza relativa riscalata rapporta il giudizio di ciascun item a quello migliore e peggiore degli altri items della facoltà, cioè il grado di apprezzamento espresso, viene in tal modo giudicato in relazione agli altri aspetti indagati.

Inoltre è stato indicato il numero di facoltà per sede nelle quali la frequenza relativa riscalata di ciascun item assume un valore superiore uguale o inferiore rispetto a quella delle Sedi Padane.

VI) Nota al § 3.3.3 e al § 4.3.3 Tabella 19 “Posizionamento delle Classi di laurea per tipo di laurea e numero medio di studenti frequentanti”

Nella tabella viene indicato il posizionamento di ciascuna classe di laurea calcolato, tramite il metodo di scaling di Thurstone, sulla base della distribuzione di frequenza dell'item “*E' complessivamente soddisfatto di come è stato svolto questo insegnamento?*”.

Lo scaling è una tecnica di analisi multivariata che consiste nell'assegnazione di punteggi quantitativi attraverso opportune trasformazioni “ottimali” alle variabili categoriali.

Nel nostro caso la trasformazione applicata è l'inversa della *distribuzione normale standardizzata cumulata* $F(x)$ (v. **Grafico A**) applicata alla distribuzione di frequenza relativa dell'item.

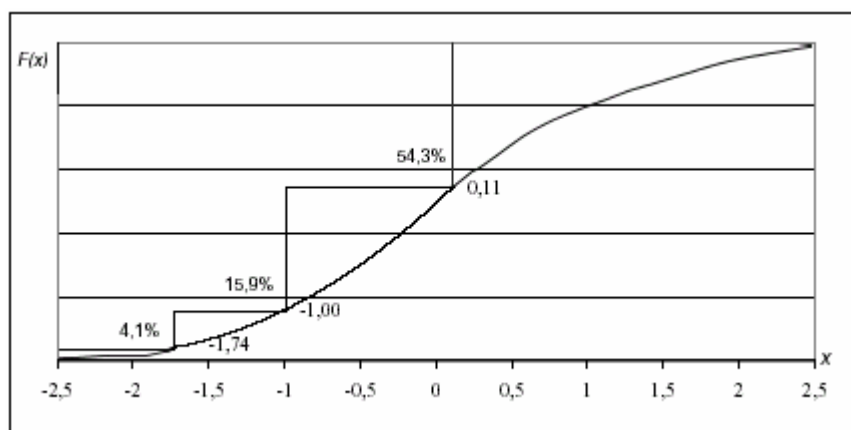


Grafico A

Infine per determinare il posizionamento delle classi di laurea è stato considerato il valore assunto dalla modalità che occupa la posizione centrale della nuova variabile (pari a -1 nel grafico di esempio).

Thurstone suppone infatti che la vera scala in base alla quale il soggetto effettua la sua valutazione è una scala distribuita secondo una legge normale; così propone una trasformazione che consente di ricondurre a scale ad intervalli le scale di punteggio,

solo di tipo ordinale, tipicamente utilizzate per descrivere le valutazioni sugli attributi di una dimensione di un servizio o di un prodotto.

Secondo il modello di Thurstone, proposto originariamente nell'ambito dello studio delle valutazioni soggettive conseguenti agli stimoli provenienti da un oggetto, si assume che allo stimolo corrisponda un processo discriminante aleatorio soggettivo che porta alla percezione dell'oggetto. Si assume che il risultato soggettivo di detto processo discriminatorio possa descriversi, per ogni aspetto, mediante una variabile casuale di tipo normale, "latente", cioè non direttamente osservata. Si suppone così che alla variabile categorica X (nel nostro caso il "giudizio complessivo") corrisponda una variabile casuale Z di tipo normale (o gaussiano) e quindi caratterizzata da valori misurati su una scala ad intervalli. Di tale variabile è possibile calcolare un valore medio, che vale come indice di posizione della distribuzione.

Alla luce, quindi, di tale metodologia di analisi è stato quindi riportato nella tabella il valore di posizionamento calcolato per ciascuna classe distinta per tipo di laurea e numero medio di studenti frequentanti per insegnamento. La scala adottata va da un massimo di + 300 a un minimo di - 300, che si identificano rispettivamente con il 100% del valore "Decisamente SI" ovvero del valore "Decisamente NO" nella somma delle frequenze di risposta. Quanto più il valore di posizionamento è spostato sullo zero tanto più il giudizio complessivo medio è neutro.

VII) Nota al § 3.4.1 e al § 4.4.1 Tabella 9 "Confronto tra la soddisfazione complessiva per gli insegnamenti e l'interesse agli argomenti trattati negli aa. aa. 2006/07 e 2007/08"

La tabella "Confronto tra la soddisfazione complessiva per gli insegnamenti e l'interesse agli argomenti trattati" propone un confronto tra la distribuzione di frequenza assoluta e percentuale delle risposte date agli items: "E' complessivamente soddisfatto di come è stato svolto questo insegnamento?" e "All'inizio dell'insegnamento, era interessato agli argomenti trattati dal corso?" negli aa.aa. 2006/2007 e 2007/2008. Il dato è dettagliato per anno di corso e proposto per le sole classi di laurea presenti in entrambi gli anni accademici.

VIII) Nota al § 3.4.1 e al § 4.4.1 Tabella 10 "Correlazione di Spearman in ordine decrescente di importanza"

Per determinare il grado di correlazione tra le variabili è stato utilizzato il *coefficiente di correlazione di Spearman* indicato con il simbolo greco ρ (rho) (proposto con l'articolo *The proof and measurement of association between two things* su *American Journal of Psychology* vol. 15, pp.72 - 101 e con l'articolo *A footnote for measuring correlation*, pubblicato nel 1906 su *British Journal of Psychology*).

Si tratta di un coefficiente di correlazione lineare, applicato a variabili ordinali, tra posizioni (*ranghi*), in grado di quantificare il grado di concordanza tra due variabili; dove per *rango* si intende la posizione assunta da ogni osservazione della variabile, dopo avere ordinato l'intera distribuzione. Ad esempio, in una variabile che assume le seguenti modalità: 2, 12, 7, 4, l'osservazione "2" ha rango 1, l'osservazione "4" ha rango 2, l'osservazione "7" ha rango 3, l'osservazione "12" ha rango 4.

Il metodo consiste nel calcolare la correlazione

$$r = \frac{\text{cov}(xy)}{\sqrt{\text{var}(x) * \text{var}(y)}}$$

non sui punteggi originali ma su quelli trasformati in ranghi, assegnando a ciascuna osservazione una posizione nella serie di dati preventivamente ordinati, pertanto la formula diviene:

$$\rho = \frac{\sum_{i=1}^N \left(Rx_i - \frac{N+1}{2} \right) * \left(Ry_i - \frac{N+1}{2} \right)}{\frac{N(N^2 - 1)}{12}}$$

dove:

Rx_i e Ry_i sono i ranghi delle 2 variabili

N è il numero complessivo delle osservazioni

Quando due o più valori delle variabili sono identici e quindi hanno lo stesso rango, l'attribuzione dei punteggi medi riduce il valore della varianza ($\text{var}(x)$, $\text{var}(y)$), in tal caso è opportuno calcolare un fattore di correzione T per entrambe le variabili x e y .

Dove:

$$T = \sum_{i=1}^g (t_i^3 - t_i)$$

g è il numero di valori con punteggi identici

t è il numero di ranghi identici per ogni valore

Con questa correzione la formula del ρ di *Spearman* diventa:

$$\rho = \frac{N^3 - N - 6 * d - (T_x + T_y) / 2}{\sqrt{(N^3 - N)^2 - (T_x + T_y) * (N^3 - N) + T_x * T_y}}$$

dove:

$$d = \sum_{i=1}^N (Rx_i - Ry_i)^2$$

Tale coefficiente assume valori compresi tra -1 e +1, dove gli estremi indicano la perfetta dipendenza (inversa o diretta) tra le variabili, anche se ciò è difficilmente riscontrabile nella realtà.

Se il coefficiente assume valori positivi c'è dipendenza diretta tra i giudizi dati ai singoli items e il giudizio complessivo (al crescere di un giudizio, cresce anche l'altro), altrimenti si manifesta una dipendenza inversa.

IX) Nota al § 3.4.3 e al § 4.4.3 Tabella 17 “*Conoscenze preliminari e anno di corso*”

Nella tabella “*Conoscenze preliminari e anno di corso*” viene analizzata la distribuzione di frequenza assoluta e percentuale dell'item “*Le conoscenze preliminari da lei possedute sono risultate sufficienti per la comprensione degli argomenti trattati?*” per ciascun anno di corso, è stato inoltre valutato il grado di correlazione tra le due variabili tramite il coefficiente ρ di Spearman, che è possibile calcolare solo quando nel raggruppamento sono presenti più anni di corso.

Nel calcolo del “*Totale*” sono stati omissi gli studenti che non hanno indicato l'anno di corso di iscrizione.

X) Nota al § 3.4.4 e al § 4.4.4 Tabella 18 “*Percezione degli studenti circa la corrispondenza fra i Crediti Formativi Universitari, attribuiti a ciascun insegnamento, e il carico di studio complessivo da sostenere*”

La tabella riporta, per ogni anno di corso, la distribuzione di frequenza assoluta e percentuale dell'item “*Rispetto ai crediti formativi assegnati la quantità di lavoro richiesta dall'insegnamento risulta essere*” e la distribuzione di frequenza ricavata paragonando le risposte date all'item “*Stimi quante ore di studio richiede complessivamente la preparazione dell'esame*” con le ore teoriche ottenute moltiplicando per 25, che è la quantità di lavoro corrispondente a un credito formativo, il numero di CFU attribuiti a ciascun insegnamento.

XI) Nota al § 4.4.5 L'utilizzo delle nuove tecnologie

(NB): le aule virtuali sono riferite anche a docenti non titolari di insegnamenti nell'anno accademico 2007-08, inoltre è opportuno sottolineare che la stessa aula virtuale può essere legata a più facoltà se il docente ha insegnamenti in diverse facoltà.