

# INDICATORI STATISTICI PER LA VALUTAZIONE DELLA QUALITÀ DELLA DIDATTICA UNIVERSITARIA: UNA PROPOSTA METODOLOGICA

Massimo Alfonso Russo

## 1. INTRODUZIONE

In Italia è stato di recente istituito un articolato sistema per il controllo delle attività universitarie nazionali e locali, tra i cui obiettivi principali indubbiamente rientra la *valutazione* dell'attività didattica. Questo riferimento operativo ha trovato il suo consolidamento normativo nella legge 370/99. Tuttavia in questi anni ogni Organo periferico di controllo (Nucleo di valutazione di Ateneo) ha perseguito tale nuovo obiettivo con procedure, tecniche e metodi di analisi statistiche non omogenee. Pertanto, emerge fortemente una nuova esigenza, tipica delle situazioni transitorie, rappresentata dalla necessità di standardizzare metodologicamente la procedura di valutazione. In questo contesto si inquadra il presente lavoro che ha quale scopo principale quello di delineare una nuova metodologia, statisticamente coerente con la natura dei dati elaborati, capace di costruire un indicatore lordo multidimensionale della qualità della didattica e ciò indipendente sia da preventive assunzioni teoriche che dalla soggettività delle scelte individuali. A tal fine sono preliminarmente riesaminate le proposte ad oggi già formulate evidenziando, oltre alle caratteristiche di ognuna, anche le possibili differenze nella valutazione della qualità che per il loro tramite possono ricavarsi. Un'applicazione empirica ai dati rilevati presso l'Università di Foggia nell'anno accademico 2000-2001 viene infine presentata.

È opportuno precisare fin d'ora che nel presente lavoro con la dizione *qualità della didattica* si considera il grado di soddisfazione dello studente frequentante in funzione di svariati aspetti quali: l'adeguatezza delle aule, l'organizzazione e la conduzione delle lezioni, ecc.... La sua misurazione avviene attraverso l'impostazione di un *modello giudicatorio*, nel quale lo studente viene posto nella condizione di giudice del servizio didattico fruito. Questo modello è attualmente ritenuto il riferimento standard per la rilevazione della qualità della didattica universitaria con i moderni sistemi *computer-assisted* (Fabbris e Gasparotto, 2001).

## 2. I DATI

Per esemplificare le tecniche analizzate e la nuova proposta formulata in questo lavoro sono stati utilizzati i dati raccolti (attraverso l'elaborazione di 1.148 questionari) in occasione della rilevazione sulla qualità della didattica svolta dall'Università di Foggia, con riferimento particolare alla Facoltà di Economia ed ai 28 insegnamenti attivati nel secondo semestre dell'anno accademico 2000-2001. La rilevazione delle informazioni ritenute necessarie si fonda sulla somministrazione di un questionario agli studenti frequentanti e presenti in aula, una sola volta e verso il periodo terminale delle lezioni di ogni singolo corso universitario.

In particolare, il questionario utilizzato ricalca sostanzialmente quello proposto quale base minima di riferimento dall'Osservatorio Nazionale per la Valutazione del Sistema Universitario (ora diventato Comitato Nazionale per la Valutazione del Sistema Universitario - C.N.V.S.U.) a seguito delle indicazioni fornite dalla *relazione finale del gruppo di ricerca Ministeriale* (Chiandotto e Gola, 1999) appositamente costituito. Le domande in esso inserite sono 27: le prime 8 riguardano informazioni anagrafiche sullo studente intervistato o generiche caratteristiche del corso di laurea frequentato; ognuna delle restanti 19 analizza una particolare variabile del più ampio fenomeno della qualità della didattica offerta. Queste ultime 19 domande possono essere accorpate, per omogeneità di argomento trattato, in cinque grandi categorie: *Categoria A* – domande riguardanti le condizioni e l'adeguatezza delle Aule e delle attrezzature didattiche (4 domande); *Categoria B* – domande riguardanti il Carico di lavoro dell'insegnamento, gli orari e il calendario degli esami (5 domande); *Categoria C* – domande riguardanti le caratteristiche delle Lezioni impartite (5 domande); *Categoria D* – domande riguardanti le Esercitazioni pratiche (3 domande); *Categoria E* – domande di chiusura per una sintetica valutazione finale del grado di soddisfazione nei confronti della disciplina esaminata (2 domande). Nella banca dati utilizzata per le applicazioni empiriche ogni insegnamento è stato identificato, per ovvi motivi di *privacy*, con un codice univoco numerico di tre cifre.

## 3. GLI INDICATORI

Gli indicatori di qualità della didattica sono inquadrabili in due categorie: quelli *lordi* e quelli *netti*. Mentre i primi si limitano a sintetizzare i dati rilevati in un'unica misura, i secondi tentano anche di spiegare gli eventuali fattori che possono aver inciso sull'entità dei risultati ottenuti. Infatti, occorre dire che le valutazioni espresse direttamente dagli studenti possono essere influenzate, oltre che dal reale livello qualitativo dei singoli insegnamenti, anche e soprattutto da altri fattori:

a) di natura *macro-contestuale* o di contesto esterno quali, ad esempio, gli aspetti demografici, culturali, politico-istituzionali che caratterizzano l'ambiente socio-economico in cui il processo formativo universitario dell'Ateneo è inserito;

b) di natura *micro-contestuale* o di contesto interno quali, ad esempio, gli aspetti organizzativi, gestionali, politico-istituzionali che caratterizzano l'ambiente universitario;

c) di natura *individuale* quali il sesso, l'età, la formazione scolastica secondaria superiore, l'interesse per l'area disciplinare ecc..., che evidenziano le peculiarità della popolazione studentesca destinataria del servizio.

Indicatori netti si possono ottenere, com'è noto (Gori e Vittadini, 1999), utilizzando la generica classe dei modelli multilivello (Goldstein, 1995; Snijders e Bosker, 1999), i quali consentono di ricavare risultati privi dell'influenza di specifici fattori di contesto o individuali.

L'approfondimento e l'affinamento delle metodologie relative agli indicatori lordi risulta comunque preliminarmente molto proficua. Infatti, la quasi totalità delle elaborazioni dei Nuclei locali di valutazione sottoposte all'attenzione del C.N.V.S.U. è fondata proprio su tali indicatori. D'altronde, in condizioni di sufficiente omogeneità del collettivo analizzato<sup>1</sup>, è stata empiricamente verificata (Rampichini *et al.*, 2002) la similarità tra le valutazioni ottenute da indicatori lordi e netti. Non va trascurato, inoltre, il valore informativo sullo stato dei fatti che i primi possono offrire in ogni situazione. In conclusione, gli indicatori lordi rappresentano un importante punto di partenza<sup>2</sup> nella valutazione della qualità della didattica universitaria, mentre la successiva comprensione delle dinamiche di tali valutazioni implica il ricorso ad indicatori netti.

#### 4. INDICATORI LORDI UNIDIMENSIONALI

Gli indicatori lordi possono essere ricavati considerando esclusivamente un aspetto della qualità percetiva (adeguatezza delle aule, chiarezza del docente, ecc...) o, in alternativa, considerando più aspetti contemporaneamente. Nella prima ipotesi, essi sono *unidimensionali*, mentre nella seconda *multidimensionali*. È immediato rilevare la parzialità dell'informazione sulla qualità della didattica che i primi possono offrire rispetto ai secondi. La costruzione di questi ultimi, tuttavia, è diretta conseguenza delle scelte metodologiche preliminarmente effettuate sui primi.

Va chiarito innanzi tutto che ognuno dei quesiti presenti nel questionario-tipo utilizzato nella rilevazione statistica prevede sempre le stesse possibilità di risposta su scala ordinale a quattro modalità bilanciate: a) decisamente no; b) più no che sì; c) più sì che no; d) decisamente sì. Dunque, per ogni *item*<sup>3</sup> la prima elaborazione dovrebbe essere rappresentata dalla sintesi dei dati per il suo tramite rilevati. In tal senso, tenuto conto della natura ordinale dei caratteri qualitativi analizzati, è

<sup>1</sup> Come spesso accade con riferimento alle caratteristiche dei gruppi di studenti frequentanti.

<sup>2</sup> Soprattutto per evidenziare le situazioni più critiche per le quali si impone un approfondimento sulle eventuali cause e, dunque, sugli interventi da adottare.

<sup>3</sup> Da ora in poi, per brevità, ogni mutabile osservata con le singole domande del questionario è identificata come *item*.

ovvia l'associazione ad ogni *item* di una media lasca quale la mediana  $M$  e di misure assolute e relative di dispersione (Leti, 1983; Piccarreta, 2001) quali:

$$D_{jq} = \sum_{k=1}^{K-1} F_{kjq} (1 - F_{kjq}), \quad 0 \leq D_{jq} \leq \frac{K-1}{4}; \quad (1)$$

$$d_{jq} = \frac{D_{jq}}{(K-1)/4}, \quad 0 \leq d_{jq} \leq 1. \quad (2)$$

Nelle precedenti espressioni:  $k = 1, 2, \dots, K$  identifica l'indice delle modalità in cui può articolarsi il carattere  $Y_q$  osservato (nel caso in esame  $k = 1, 2, 3, 4$ );  $q = 1, 2, \dots, Q$  identifica lo specifico *item* del questionario (nel caso in esame  $q = 1, 2, \dots, 27$ );  $j = 1, 2, \dots, J$  identifica il singolo insegnamento (nel caso in esame  $j = 1, 2, \dots, 28$ ); mentre  $F_{kq} = \text{prop}(Y_q \leq y_{kq})$  è la funzione di ripartizione relativa della generica mutabile ordinale  $Y_q$  e  $y_{kq}$  rappresenta la  $k$ -esima modalità in cui il  $q$ -esimo carattere qualitativo può articolarsi. Anche se l'estremo superiore dell'indice di dispersione identificato nella (1) vale solo per collettivi con numerosità pari, è possibile estendere la sua validità anche quando la numerosità è dispari purché, ovviamente, la stessa sia sufficientemente elevata. Tale condizione è quasi sempre rispettata nelle indagini condotte sulla qualità della didattica e, pertanto, il limite superiore dell'indice  $D_{jq}$  può essere posto pari a  $0,75^4$ . Di conseguenza, la costruzione dell'indice relativo nella (2) è immediata.

Se per tutti gli *item* osservati, con riferimento al  $j$ -esimo insegnamento, si calcolassero i rispettivi valori di  $M_{jq}$  e  $d_{jq}$  si otterrebbero  $Q$  misurazioni unidimensionali del livello di qualità soggettivamente percepito dagli studenti frequentanti quel determinato insegnamento. L'accoppiamento tra valori medi e indici di dispersione risulta quanto mai necessario, tenuto conto sia della natura qualitativa dei caratteri sia, soprattutto, del ridotto numero di modalità in cui gli stessi possono manifestarsi.

La sintesi unidimensionale dei dati, così come prospettata fino ad ora, è quella statisticamente più corretta e ciò in funzione della natura qualitativa ed ordinale delle variabili considerate. Tuttavia, il ricorso ad una preventiva quantificazione delle modalità della scala ordinale potrebbe rappresentare un costo dal punto di vista concettuale piuttosto elevato, ma operativamente molto utile. Infatti, la trasformazione del carattere qualitativo consente l'utilizzo, per la relativa sintesi ed analisi dei dati, degli usuali strumenti statistici riservati alle variabili di tipo quantitativo. In tal senso, la stessa relazione finale del gruppo di ricerca ministeriale in precedenza citata, nell'indicare le modalità di elaborazione dei dati provenienti dai

<sup>4</sup> Infatti, il valore massimo di  $D_{jq}$ , pur variando in funzione della diversa mediana rilevata per singola  $q$ -esima distribuzione di frequenza, può essere approssimato sempre dal valore 0,75. Ciò deriva dal fatto che al diminuire del numero della  $K$  modalità in cui il carattere può esplicitarsi e all'aumentare della numerosità del collettivo statistico osservato, il massimo della mutabilità differisce per difetto da 0,75 di una quantità tanto piccola da non essere ritenuta influente.

questionari somministrati agli studenti frequentanti, effettua proprio tale opzione per il tramite di una *quantificazione determinata diretta*<sup>5</sup> delle modalità qualitative, ponendo: decisamente no = 2; più no che sì = 5; più sì che no = 7; decisamente sì = 10.

Nel caso specifico, inoltre, il mancato ricorso alla quantificazione tradizionale priva di salti tra valori (ad esempio: 0, 1, 2, 3 oppure 1, 2, 3, 4) si rivela non solo la scelta più adatta empiricamente, ma anche quella più aderente sotto il profilo teorico (Marbach, 1974). Premesso che  $x_k$  indica la  $k$ -esima modalità quantitativa in cui può esprimersi il generico carattere osservato e  $n_{kj}$  la corrispondente frequenza assoluta rilevata nel  $j$ -esimo insegnamento per la  $q$ -esima domanda, è immediato ricavare:

$$\mu_{jq} = \frac{\sum_{k=1}^K x_k n_{kj}}{N_{jq}}, \quad D'_{jq} = \sigma_{jq} = \sqrt{\frac{\sum_{k=1}^K (x_k - \mu_{jq})^2 n_{kj}}{N_{jq}}}; \quad (3)$$

$$d'_{jq} = \frac{D'_{jq}}{\max D'_{jq}}, \quad 0 \leq d'_{jq} \leq 1. \quad (4)$$

È opportuno precisare inoltre che:

$N$  = totale studenti frequentanti rispondenti al questionario;

$N_j$  = totale studenti frequentanti il  $j$ -esimo insegnamento rispondenti al questionario;

$N_{jq} = \sum_{k=1}^K n_{kj} =$  totale studenti frequentanti il  $j$ -esimo insegnamento e rispondenti al  $q$ -esimo *item*;

$N_{.q} = \sum_{j=1}^J N_{jq} =$  totale studenti rispondenti al  $q$ -esimo *item* in tutti gli insegnamenti;

$N_{j.} = \sum_{q=1}^Q N_{jq} =$  totale risposte sul  $j$ -esimo insegnamento;

$N_{..} = \sum_{j=1}^J \sum_{q=1}^Q N_{jq} =$  totale generale delle risposte.

Analogamente, per le frequenze assolute ed i valori medi:

<sup>5</sup> Tale modo di procedere si fonda sull'ipotesi che alle  $K$  modalità qualitative, in cui si manifesta il generico carattere  $Y$  analizzato, si sottenda un *continuum* di valori numerici, il quale è in stretta relazione sia con il diverso grado di intensità del concetto espresso, che con le caratteristiche individuali dei rispondenti (Young, 1995). Tuttavia, l'individuazione puntuale dei valori numerici da sostituire alle modalità ordinali avviene in modo sostanzialmente soggettivo o addirittura arbitrario.

$$n_{k \cdot q} = \sum_{j=1}^J n_{kj q}; \quad n_{kj \cdot} = \sum_{q=1}^Q n_{kj q}; \quad n_{k \cdot \cdot} = \sum_{j=1}^J \sum_{q=1}^Q n_{kj q}, \quad 1 \leq k \leq K;$$

$$\mu_q = \frac{\sum_{k=1}^K x_k n_{k \cdot q}}{N_{\cdot q}}; \quad \mu_j = \frac{\sum_{k=1}^K x_k n_{kj \cdot}}{N_{j \cdot}}; \quad \mu = \frac{\sum_{k=1}^K x_k n_{k \cdot \cdot}}{N_{\cdot \cdot}}.$$

Nell'ipotesi di assenza di mancate risposte, valgono le seguenti relazioni:

$$N_{jq} = N_j; \quad N_{\cdot q} = \sum_{j=1}^J N_j = N; \quad N_{j \cdot} = Q N_j; \quad N_{\cdot \cdot} = Q \sum_{j=1}^J N_j.$$

Come è ovvio, il limite di questo tipo di impostazione è l'estrema soggettività della tecnica di quantificazione utilizzata e la conseguente divergenza nei risultati a seconda della scala di misura predeterminata (De Luca, 2001). Senza alcuna pretesa di risolvere uno dei problemi statistici tra i più dibattuti in letteratura e pur condividendo i vantaggi operativi che si otterrebbero dalla quantificazione delle modalità dei caratteri qualitativi osservati, si ritiene tale preventiva trasformazione una scelta non obbligata e, al limite, meglio realizzabile per mezzo della tecnica della *quantificazione determinata indiretta*<sup>6</sup> (Fielding, 1997; Zanella e Cerri, 2000) o ancora con il ricorso alla logica degli insiemi sfocati (Lalla e Facchinetti, 2001). Queste ultime soluzioni non solo possono essere ritenute più rispondenti alle condizioni reali, ma sembrano quelle più conformi ai recenti sviluppi metodologici nella ricerca di indicatori netti della qualità della didattica.

## 5. INDICATORI LORDI MULTIDIMENSIONALI DESCRITTIVI

### 5.1. *Analisi grafica*

Gli indicatori unidimensionali sono in grado di offrire una visione parziale del livello complessivo della qualità percepita dal fruitore (studente) del servizio (insegnamento), in quanto ognuno di essi riesce a misurare il gradimento espresso dagli studenti soltanto con riferimento al singolo aspetto qualitativo esaminato tramite il  $q$ -esimo *item* del questionario. Pertanto, è quanto mai opportuno sintetizzare tali indicatori ottenuti in corrispondenza delle  $Q$  domande sommini-

<sup>6</sup> Questa procedura di quantificazione prevede di assegnare alle modalità della mutabile ordinale numeri reali convenzionali, non necessariamente equidistanti tra loro, ricorrendo ad una variabile continua latente (ad esempio la curva normale) in grado di rappresentare matematicamente il *continuum* di valori sotteso al carattere qualitativo osservato (Torgerson, 1958). Ciò al fine di misurare quantitativamente la diversa intensità delle iniziali risposte qualitative, ricavando i singoli valori numerici in via indiretta per il tramite della distribuzione di frequenze empiriche. In questo caso, la validità della procedura di individuazione dei puntuali valori numerici da utilizzare non dipende più dalla trasformazione soggettiva del ricercatore che la attua, ma dalla sussistenza dell'ipotesi teorica su cui la stessa procedura si fonda, ovvero la rappresentatività della variabile latente prescelta in rapporto al fenomeno osservato.

strate. Risulta dunque necessario determinare una metodologia per l'individuazione di una misura multidimensionale da associare ad ogni insegnamento<sup>7</sup> che consenta, una volta calcolata, anche la comparabilità dei risultati tra insegnamenti. La prima elementare forma di sintesi dei dati unidimensionali è costituita dalla rappresentazione grafica di tutti i valori medi calcolati per ogni *item*, ottenendo quello che viene definito il *profilo* del *j*-esimo insegnamento. Alcuni esempi sono riportati nei grafici della successiva figura 1, in ognuno dei quali vengono rappresentati due insegnamenti della Facoltà di Economia dell'Università di Foggia: il primo grafico è ottenuto sintetizzando i dati rilevati per ogni *item* con l'ausilio della mediana e, quindi, mantenendo inalterata la natura qualitativa ed ordinale dei caratteri osservati; il secondo, invece, è ottenuto con l'ausilio della media aritmetica, attraverso la preventiva trasformazione quantitativa diretta delle modalità dei caratteri qualitativi.

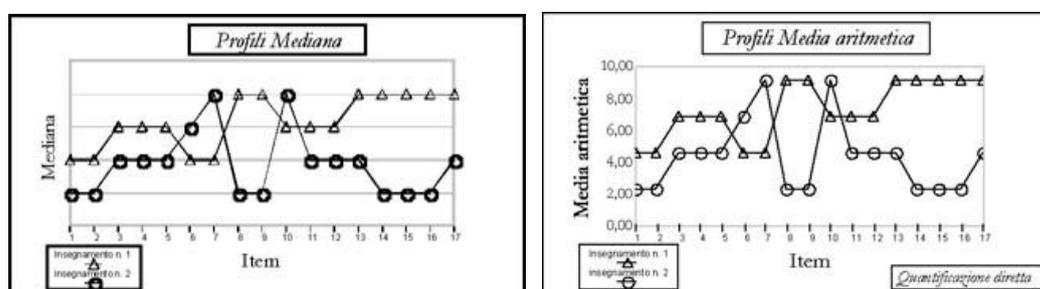


Figura 1 – Rappresentazione grafica dei profili di due insegnamenti – (le 17 domande considerate sono quelle inserite nelle *Categorie A, B, C e D* in precedenza specificate).

Attraverso l'analisi dei profili appare evidente la duplice possibilità di effettuare da un lato valutazioni comparative tra insegnamenti<sup>8</sup> e, dall'altro, di individuare eventuali *item* o gruppi di *item* limite<sup>9</sup>. Pur tuttavia, tale procedura costituisce statisticamente un approccio debole alla sintesi dei dati. Per questo un indicatore multidimensionale dovrebbe prevedere una vera e propria elaborazione quantitativa dei valori medi unidimensionali ottenuti nella prima fase, con il problema aggiuntivo rappresentato dalla convenienza di ricorrere a sintesi *semplici*, ovvero *ponderate*. Com'è ovvio, l'uso di indicatori ponderati appare in questo contesto un percorso obbligato.

<sup>7</sup> Se l'insegnamento è collegato univocamente ad un docente, la valutazione dell'insegnamento si può estendere a quella dello stesso docente, eliminando però l'incidenza delle variabili ad esso estranee quali, ad esempio, l'adeguatezza delle aule e delle strutture, la disponibilità di laboratori e di biblioteche, ecc....

<sup>8</sup> Preferibilmente un numero contenuto di insegnamenti, tale da garantire la chiarezza del grafico stesso.

<sup>9</sup> Ovvero oltre quelle soglie minime (di attenzione) e massime (di eccellenza) individuate dai Nuclei di valutazione locali quali estremi del *range* di sostenibilità per ogni *item*.

## 5.2. Indicatori ponderati

Una delle tecniche di ponderazione più oggettive si fonda sul diverso livello di variabilità associabile ad ogni *item*<sup>10</sup>. Premesso che il ricorso eventuale ad altre forme di ponderazione tra i diversi *item* non muta la sostanza di quanto verrà detto in seguito, la sintesi quantitativa dei  $Q$  valori medi rilevati per il  $j$ -esimo insegnamento può avvenire esclusivamente in modo analitico e, quindi, essa è praticabile soltanto se si è preventivamente optato per la quantificazione delle modalità ordinali. Di converso, essa risulta impossibile se si è ritenuto di mantenere immutata la natura qualitativa delle stesse.

Quanto appena argomentato induce a concludere che:

I) nel caso di mantenimento della natura qualitativa delle  $Q$  mutabili ordinali osservate, l'indicatore lordo multidimensionale ponderato non potrebbe essere costruito;

II) nel caso di preventiva trasformazione quantitativa delle  $Q$  mutabili ordinali osservate, l'indicatore lordo multidimensionale può essere ottenuto attraverso:

$$I_j^l = \frac{\sum_{q=1}^Q \mu_{jq} (1 - d'_{jq})}{\sum_{q=1}^Q (1 - d'_{jq})}. \quad (5)$$

Il superamento dell'ostacolo operativo di cui al precedente punto I) è stato di recente oggetto di approfondimento in alcuni studi (Capursi e Porcu, 2001; Civardi, 2002). Con il medesimo obiettivo, questo lavoro intende proporre una procedura alternativa fondata sulla preventiva rappresentazione grafica dei valori mediani e degli associati indici di dispersione di cui alla (2), utilizzando un diagramma nel quale sull'asse delle ascisse viene riportato il campo di variazione dell'indice  $d'_{jq}$ , mentre sull'asse delle ordinate le  $K$  modalità associate alle  $Q$  mutabili ordinali osservate. La rappresentazione che ne scaturisce (contenuta nel primo quadrante del diagramma) viene suddivisa in  $V$  sotto-quadranti con  $V = 1, 2, 4, 6, 8, \dots$ . Il numero  $V$  dei sotto-quadranti in cui si intende suddividere il grafico dipende essenzialmente da due elementi: il numero  $K$  di modalità; l'importanza che si intende dare alla variabilità (o in generale al fattore di ponderazione) quale fenomeno inversamente discriminante in termini di qualità percepita. L'unico vincolo nella costruzione di tutti i sotto-quadranti è rappresentato dal fatto che gli stessi devono avere la medesima dimensione, ovvero al loro interno deve poter ricadere un numero uguale di casi possibili. Il grafico di figura 2, con riferimento al generico  $j$ -esimo insegnamento e con  $V = 4$ , riporta un esempio della rappresentazione grafica in esame.

<sup>10</sup> Tanto maggiore sarà la variabilità tanto meno affidabile sarà il valore informativo, in termini di qualità, del valore medio unidimensionale considerato e, pertanto, minore dovrà essere il suo peso nella sintesi finale, e viceversa.

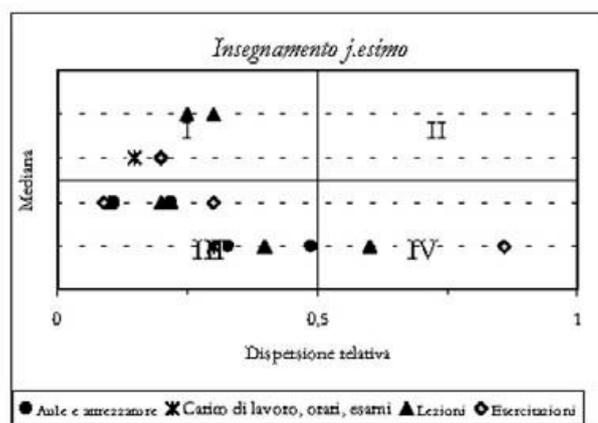


Figura 2 – Rappresentazione grafica della valutazione dell'insegnamento  $j$ -esimo - Valore mediano e indice relativo di dispersione per tutti i 17 *item*.

In tal modo per il  $j$ -esimo insegnamento, ogni *item* rilevato è indentificato graficamente da un punto sul diagramma cartesiano. Tenuto conto della *trasferibilità* del carattere quantitativo *numero di item rilevabili nei diversi quadranti*, è dunque possibile considerare che:

a) la rappresentazione grafica non comporta concettualmente alcuna quantificazione del carattere qualitativo<sup>11</sup>;

b) la presenza di un generico *item* nel I sotto-quadrante (alto valore mediano e bassa dispersione) rappresenta una situazione estremamente positiva in termini di qualità, che tende a diminuire progressivamente con il passaggio ideale dello stesso *item* al sotto-quadrante II (alto valore mediano e alta dispersione), poi al III (basso valore mediano e bassa dispersione), ed infine al IV (basso valore mediano e alta dispersione);

c) da una opportuna misura dell'addensamento nei  $V$  sotto-quadranti dei  $Q$  *item* osservati per un dato insegnamento è plausibile desumere il relativo indicatore multidimensionale di qualità del medesimo.

Quest'ultimo indicatore non può ricavarsi immediatamente tramite il noto *rapporto di concentrazione di Gini*, in quanto lo stesso misurerebbe soltanto il grado di addensamento dei  $Q$  punti, indipendentemente dalla distribuzione degli stessi nei diversi sotto-quadranti del grafico di riferimento<sup>12</sup>. Nel caso in esame, invece, l'eventuale addensamento dei  $Q$  punti in sotto-quadranti differenti ha un notevole valore discriminante in termini qualitativi. Per tenere conto anche di quest'ultima esigenza è sufficiente ordinare preventivamente i  $V$  sotto-quadranti,

<sup>11</sup> Va precisato che l'equidistanza delle quattro modalità qualitative (valori mediani) presente nella figura 2 è stata fissata solo al fine di rendere più chiara la lettura del grafico. Infatti, anche in caso di non equidistanza tra modalità (caso più reale) quanto verrà detto in seguito non perderà di significato.

<sup>12</sup> Infatti, il classico rapporto di concentrazione di Gini assume sempre il suo massimo valore (+1) nel caso in cui tutti i  $Q$  punti si dovessero addensare in uno qualunque dei  $V$  sotto-quadranti, e ciò indipendentemente dal fatto che tale massima concentrazione si realizzi nel primo, nel secondo o nell'ultimo degli stessi sotto-quadranti.

in virtù della crescente importanza in termini qualitativi e, successivamente, associare ad ogni sotto-quadrante la rispettiva frequenza assoluta di *item* in esso rilevate. Così facendo è possibile misurare il differente grado di addensamento applicando la seguente espressione<sup>13</sup> alla serie delle frequenze:

$$R_j = \frac{\sum_{s=1}^{V-1} (p_s - q_s)}{\sum_{s=1}^{V-1} p_s}, \quad -1 \leq R_j \leq +1. \quad (6)$$

Nella (6):  $p_s$  è il rapporto tra il numero dei primi  $s$  sotto-quadranti ordinati qualitativamente<sup>14</sup> sul totale degli stessi;  $q_s$  rappresenta la frazione di *item* presenti nella corrispondente frazione  $p_s$ . Il rapporto  $R_j$  assume il suo valore minimo nel caso di totale addensamento dei  $Q$  punti rappresentati nel grafico di figura 2 nel sotto-quadrante IV, ovvero nel peggiore dei sotto-quadranti in termini di qualità osservata. Viceversa, assume il suo valore massimo nel caso di totale addensamento dei medesimi punti nel sotto-quadrante I, ovvero nel migliore dei sotto-quadranti in termini di qualità osservata. Com'è facile notare il campo di variazione di  $R_j$   $[-1, +1]$  risulta diverso da quello del classico rapporto di concentrazione  $[0, +1]$  e tanto deriva dalla sua applicazione alle frequenze (non ordinate in modo crescente) dei punti ricadenti nei rispettivi  $V$  sotto-quadranti (ordinati in modo crescente). Pertanto le  $q_s$  possono essere anche maggiori delle  $p_s$  e, quindi, il rapporto può risultare negativo.

Quanto deriva dalla (6) si ottiene, ovviamente, anche attraverso il rapporto tra le aree di addensamento dalla seguente espressione:

$$\left( \frac{V}{V-1} \right) \left[ 1 - \sum_{s=1}^V (p_s - p_{s-1})(q_s + q_{s-1}) \right], \quad p_0 = q_0 = 0.$$

Il ricorso alle aree permette anche la misurazione visiva del livello di addensamento per mezzo della relativa rappresentazione grafica, così come si evidenzia di seguito nella figura 3. In quest'ultima sono riportati i casi estremi: tutti gli *item* sono situati nel sotto-quadrante I, in tal caso l'indice  $R_j = +1,00$  e la spezzata di addensamento è tutta al di sotto della retta di equi-addensamento; tutti gli *item* sono situati nel sotto-quadrante II, in tal caso l'indice  $R_j = +0,33$  e la spezzata di addensamento è in gran parte al di sotto, ma anche al di sopra della retta di equi-addensamento; tutti gli *item* sono situati nel sotto-quadrante III, in tal caso l'indice  $R_j = -0,33$  e la spezzata di addensamento è in gran parte al di sopra, ma anche al di sotto della retta di equi-addensamento; tutti gli *item* sono situati nel

<sup>13</sup> Identica all'originario rapporto di concentrazione di Gini.

<sup>14</sup> Nel caso di suddivisione del grafico in quattro sotto-quadranti (figura 2) l'ordine crescente degli stessi in termini di qualità è: sotto-quadrante IV; sotto-quadrante III; sotto-quadrante II e sotto-quadrante I.

sotto-quadrante IV, in tal caso l'indice  $R_j = -1,00$  e la spezzata di addensamento è tutta al di sopra della retta di equi-addensamento.

Nell'ipotesi in cui  $R_j = 0$  si verifica la situazione di equi-addensamento, ovvero quando i  $Q$  punti rappresentati nel grafico risultano equamente distribuiti nei  $V$  sotto-quadranti. Tuttavia è opportuno precisare che per  $R_j$  lo zero non rappresenta un valore estremo, infatti  $-1 \leq R_j \leq +1$ , pertanto esso può realizzarsi anche in situazioni differenti dal perfetto equi-addensamento. Quanto immediatamente precede, oltre ad essere vero per qualunque altro valore compreso nell'intervallo  $[-1, +1]$ , è comunque un costo operativo inevitabile per qualunque indicatore di sintesi.

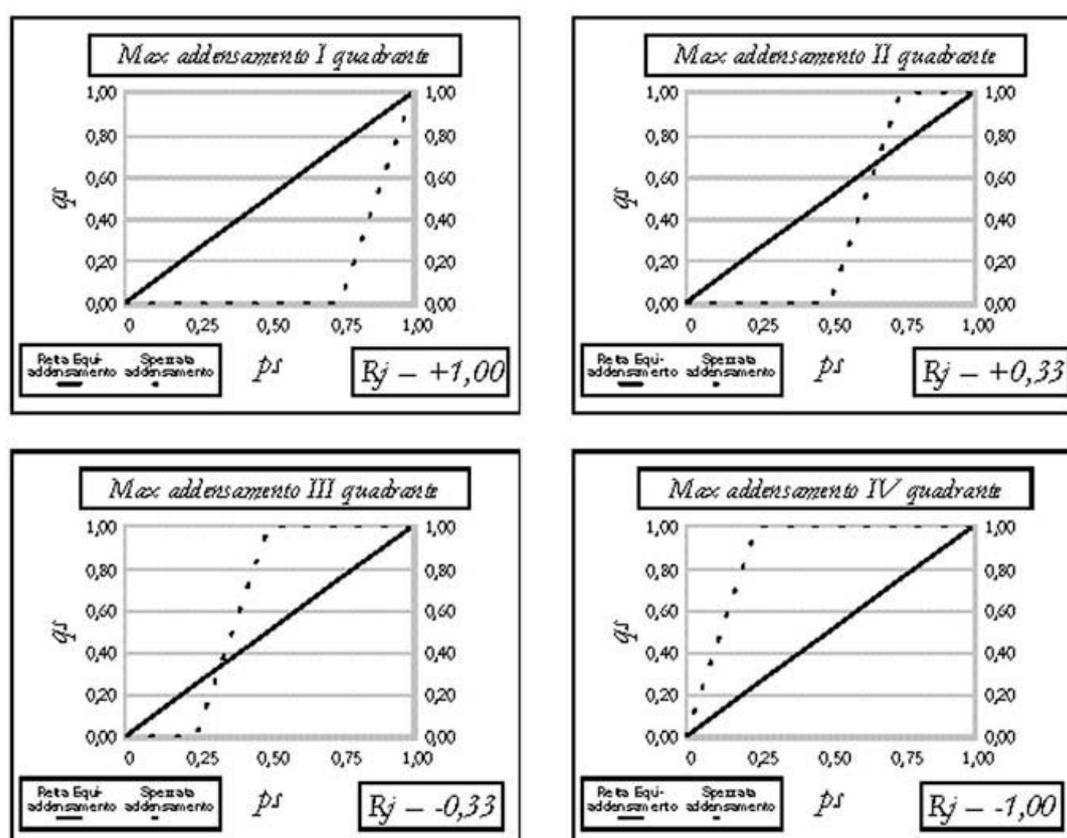


Figura 3 – Rappresentazione grafica delle ipotesi estreme di addensamento degli *item* nei quattro sotto-quadranti.

Le rappresentazioni grafiche di figura 3 chiariscono anche il perché il calcolo di  $R_j$ , tramite il ricorso alle aree, conduce ad un valore dell'indicatore compreso nell'intervallo  $[-1, +1]$ . Di fatto, le aree di addensamento<sup>15</sup> possono trovarsi: sia al di sotto della retta di equi-distribuzione - in questo caso vengono computate con

<sup>15</sup> Area compresa tra la spezzata di addensamento e la retta di equi-addensamento.

il segno positivo; sia al di sopra della medesima retta - in questo caso vengono computate con il segno negativo. Ipotizzando due insegnamenti con la seguente distribuzione di frequenze dei 17 *item* nei quattro sotto-quadranti in cui è stato suddiviso il grafico, si avrà:

Quadrante	Frequenze	$p_s$	$q_s$
IV	2	0,25	0,12
III	2	0,50	0,24
II	3	0,75	0,41
I	10	1,00	1,00

Quadrante	Frequenze	$p_s$	$q_s$
IV	10	0,25	0,59
III	3	0,50	0,76
II	3	0,75	0,94
I	1	1,00	1,00

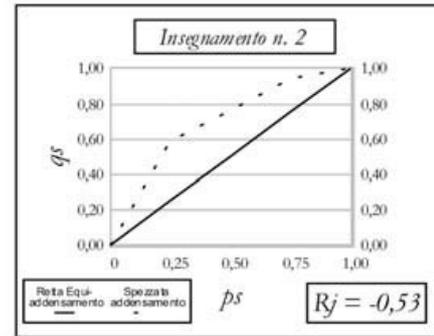
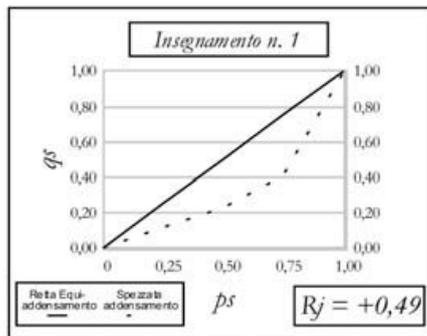


Figura 4 – Rappresentazione grafica della valutazione tramite  $R_j$  dei due insegnamenti.

Va rilevato che il valore ottenuto di  $R_j$  può essere normalizzato in modo da ottenere un indicatore espresso nella medesima scala di riferimento utilizzata nell'ipotesi di preventiva quantificazione della mutabile ordinale, ricavando così il corrispondente *indice lordo multidimensionale qualitativo*. Nello specifico se la scala di misurazione ha campo di variazione tra 0 e  $t$ , l'indice  $R_j$  andrebbe trasformato come segue:

$$I_j^l = \frac{R_j + 1}{2} \cdot t. \quad (7)$$

Se, ad esempio,  $t = 10$  e quindi la scala di misurazione è tra 0 e 10, si ha:  $I_1^l = 7,45$  per il primo insegnamento;  $I_2^l = 2,35$  per il secondo insegnamento.

È opportuno puntualizzare che la *sensibilità*, ovvero la capacità di discriminare situazioni differenti, dell'indicatore individuato con la (7) è in funzione crescente della numerosità dei  $V$  sotto-quadranti<sup>16</sup>. Tuttavia, le ulteriori suddivisioni dell'area del grafico in  $V$  sotto-quadranti (per esempio  $V = 16$ ) fanno sorgere altre problematiche. Ad esempio, esse possono effettuarsi secondo logiche diverse in funzione del peso che si intende assegnare alla variabilità quale fattore di ponderazione, così come si evince dalle seguenti figure 5a e 5b.

<sup>16</sup> Per assurdo, infatti, se  $V = 1$  qualunque insegnamento totalizzerebbe sempre la stessa valutazione qualitativa.

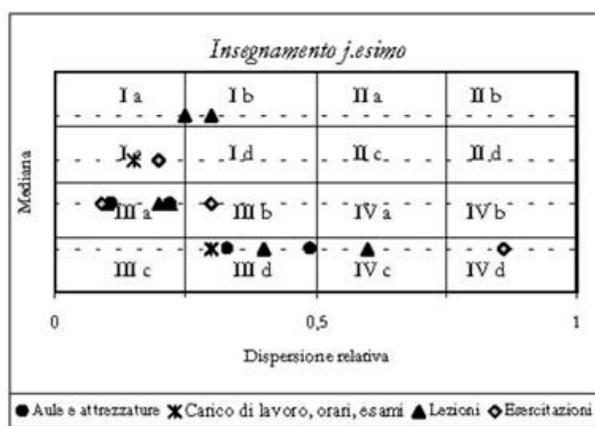


Figura 5a – Suddivisione dell'area del grafico in  $V = 16$  sotto quadranti (prima ipotesi). La variabilità ha un peso maggiore nella differenziazione qualitativa tra sotto-quadranti.

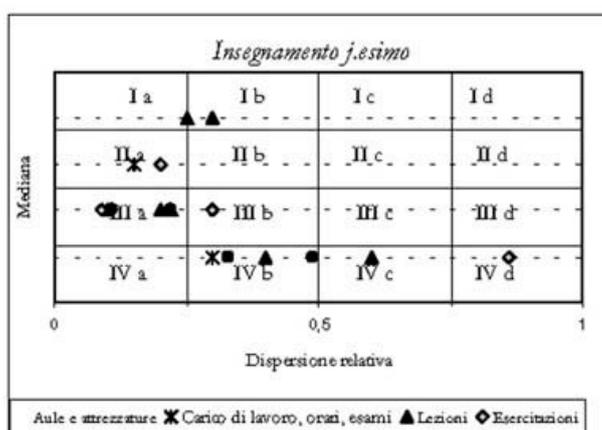


Figura 5b – Suddivisione dell'area del grafico in  $V = 16$  sotto quadranti (seconda ipotesi). La variabilità ha un peso minore nella differenziazione qualitativa tra sotto-quadranti.

Va aggiunto, infine, che in base a questa impostazione di tipo areale, sarebbe possibile anche spingere ulteriormente l'analisi<sup>17</sup> studiando, per esempio tramite l'analisi delle corrispondenze, le ipotetiche relazioni esistenti tra le quattro categorie di *item* che di solito compongono il questionario utilizzato nella rilevazione statistica.

### 5.3. La scomposizione della dispersione totale

Gli indicatori lordi multidimensionali illustrati focalizzano l'attenzione sul singolo insegnamento, fornendone una valutazione qualitativa mediante la sintesi delle informazioni provenienti da ogni *item* rilevato. Tuttavia, potrebbe essere

<sup>17</sup> Eventualmente per produrre indicazioni di corredo ai *reports* forniti al singolo docente dal Nucleo di valutazione interna locale.

opportuno non tralasciare la natura gerarchica delle stesse informazioni, considerando adeguatamente anche la rilevanza che la scomposizione della dispersione totale può avere sul singolo *item*. È possibile, quindi, scomporre la dispersione complessiva (*total*) di ogni *q*-esimo *item* nelle due quote entro (*within*) e tra (*between*) le classi o, meglio, entro e tra gli insegnamenti. In altri termini, così come risulta dimostrato per mutabili ordinali (Grilli e Rampichini, 2000):

$$D_{qT} = D_{qW} + D_{qB} ;$$

$$D_{qT} = \sum_{k=1}^{K-1} F_{kq} (1 - F_{kq}) ; \quad (8)$$

$$D_{qW} = \sum_{j=1}^J \pi_j D_{jq} ; \quad (9)$$

$$D_{qB} = \sum_{j=1}^J \pi_j Z_{2jq}^2 . \quad (10)$$

Nella (8)  $F_{kq}$  rappresenta la funzione di ripartizione marginale relativa, ottenuta come mistura delle distribuzioni relative condizionate  $F_{kjq}$ ; nella (9)  $\pi_j = N_j/N$  è la frazione di unità statistiche rilevate nel *j*-esimo insegnamento rispetto al totale; nella (10)  $Z_{2jq}^2$  rappresenta il quadrato dell'indice di dissomiglianza quadratico (Leti, 1983) tra la *j*-esima distribuzione condizionata e la distribuzione marginale della variabile ordinale  $Y_q$ . Più in particolare:

$$F_{kq} = \sum_{j=1}^J \pi_j F_{kjq} ; \quad Z_{2jq}^2 = \sum_{k=1}^{K-1} (F_{kjq} - F_{kq})^2 .$$

Nell'ipotesi di preventiva trasformazione quantitativa (diretta o indiretta) dei caratteri qualitativi analizzati, la scomposizione della dispersione totale, misurabile ad esempio per il tramite della varianza, è notoriamente uguale a:

$$(D'_{qT})^2 = \sigma_{qT}^2 = \sigma_{qW}^2 + \sigma_{qB}^2 ;$$

$$\sigma_{qT}^2 = \frac{\sum_{k=1}^K (x_k - \mu_q)^2 n_{k \cdot q}}{N_{\cdot q}} ;$$

$$\sigma_{qW}^2 = \sum_{j=1}^J \pi_j \sigma_{jq}^2 ;$$

$$\sigma_{qB}^2 = \sum_{j=1}^J \pi_j (\mu_{jq} - \mu_{\cdot q})^2.$$

Rapportando la dispersione tra le classi con quella totale si ottiene, per ognuno dei  $Q$  quesiti, un'utile interpretazione delle differenze nei giudizi formulati sull'insegnamento dagli studenti. Infatti, valori crescenti di tali rapporti evidenziano giudizi sempre più influenzati dalle caratteristiche con le quali l'insegnamento è stato impartito, piuttosto che da quelle individuali degli studenti frequentanti. Pertanto, calcolando:

$$\eta_q = \frac{D_{qB}}{D_{qT}}, \quad 0 \leq \eta_q \leq 1; \quad (11)$$

$$\eta'_q = \frac{\sigma_{qB}^2}{\sigma_{qT}^2}, \quad 0 \leq \eta'_q \leq 1; \quad (12)$$

è possibile costruire altri indicatori di qualità.

Infatti, mantenendo la natura qualitativa dei caratteri osservati, un ulteriore indice multidimensionale può essere ricavato sempre attraverso la (7) previa sostituzione, nella rappresentazione grafica da cui scaturisce, dell'indice relativo di dispersione  $d_{jq}$  con il rapporto  $\eta_q$  derivante dalla (11). Trasformando, invece, quantitativamente le modalità dei caratteri ordinali, risulta immediato ricavare l'indicatore cercato ponendo nella (5), quali fattori di ponderazione, i rapporti  $\eta'_q$  calcolati mediante la (12).

Naturalmente la ponderazione fondata sulla scomposizione della dispersione totale trascura l'entità delle differenze di variabilità entro le singole classi. Infatti, ogni *item* assume, per tutti gli insegnamenti, un'importanza proporzionale al rapporto  $\eta_q$  o  $\eta'_q$ . Tale caratteristica mentre appare sostenibile nella misurazione della qualità percepita da studenti frequentanti istituzioni formative di eccellenza, quali l'Università, essa potrebbe risultare non accettabile se riferita a gradi di istruzione inferiore.

In conclusione, il livello qualitativo di ogni singolo insegnamento universitario può essere misurato con uno qualunque dei metodi descritti fino a questo momento, ottenendo immediatamente la possibilità di ricavare anche:

- A) una graduatoria degli insegnamenti medesimi;
- B) una base di dati dalla quale effettuare una eventuale ulteriore sintesi quantitativa, utile per la valutazione di aggregati di ordine superiore (Facoltà o Ateneo).

Tuttavia la natura qualitativa ordinale delle mutabili osservate rende gli indici  $I_j^l$  statisticamente più coerenti.

## 6. ALTRI INDICATORI MULTIDIMENSIONALI

Va chiarito che gli indicatori ottenuti attraverso la sintesi descrittiva dei dati presentano il limite di non considerare la parzialità dell'indagine dalla quale i dati stessi provengono. Non va dimenticato, infatti, che gli studenti rispondenti al questionario rappresentano una parte dell'intera popolazione dei frequentanti. In considerazione di ciò è stato presentato (Fabbris, 2000) un indicatore, per misurare la qualità di un determinato servizio, costruito attraverso la misura del divario tra il reale livello di soddisfazione percepito da un campione di clienti (studenti) e quello ritenuto teoricamente ottimale. L'indice è ricavato come *distanza di Minkowski di parametro  $\lambda$*  (Leti, 1979), in base alla quale si esplicitano e si derivano le relative distribuzioni statistiche al variare del parametro medesimo. Oltre alla notevole duttilità dell'indicatore proposto, grazie soprattutto alle diverse forme che lo stesso può assumere al variare del parametro, esso risulta anche indipendente dalla scala di misura del fenomeno osservato e produce un numero puro, ossia delimitato dai valori zero e uno.

Tuttavia, l'eventuale successiva ricerca dei fattori che possono aver influito sul livello qualitativo misurato dovrebbe considerare:

1) la natura raggruppata (*clusters*) dei soggetti osservati, infatti gli studenti sono accorpati in classi;

2) la possibile struttura gerarchica dei gruppi osservati, infatti gli studenti sono accorpati in classi, le quali sono raggruppate in corsi di laurea i quali, a loro volta, sono raggruppati in Facoltà, ecc....

In quest'ottica risulta immediata l'utilità di ricorrere ai *modelli statistici di regressione ad effetti misti*. In tale ambito, molti lavori hanno inizialmente concentrato la loro attenzione su variabili risposta di tipo continuo assumendo, a seconda del particolare tipo di approccio, varie denominazioni quali: modelli ad effetti casuali; modelli a componenti di varianza; modelli lineari gerarchici; modelli multilivello; ecc... In seguito, un crescente numero di studi ha riguardato l'estensione di questi modelli al caso di variabili qualitative dicotomiche e, in genere, ordinali.

In quest'ultimo caso, se in corrispondenza di ogni generica variabile risposta ordinale osservata  $Y$  viene definita la variabile latente continua<sup>18</sup>  $Y^*$ , il seguente modello di regressione a componenti di varianza può essere formulato:

$$Y_{ij}^* = \alpha + \beta' X_{ij} + u_j + \varepsilon_{ij}, \quad (13)$$

sotto le ipotesi:

$$u_j \sim N(0, \sigma_u^2); \quad \varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2); \quad Cov(u_j, \varepsilon_{ij}) = 0. \quad (14)$$

<sup>18</sup> La quale rappresenta l'intensità del giudizio espresso dagli studenti.

Nelle (13) e (14):  $i = 1, 2, \dots, N_j$  rappresenta il generico rispondente per ogni  $j$ -esimo insegnamento;  $\mathbf{X}_{ij}$  è il vettore di dimensione  $\rho \times 1$  delle eventuali variabili esplicative osservate, misurate a livello individuale (primo livello) o di insegnamento (secondo livello);  $\boldsymbol{\beta}'$  è il vettore  $\rho \times 1$  dei coefficienti riferiti alle precedenti covariate; le v.c.  $\varepsilon_{ij}$  rappresentano i residui individuali o di primo livello, mentre le v.c.  $u_j$  indicano l'effetto casuale del  $j$ -esimo insegnamento o residui di secondo livello;  $\sigma_\varepsilon^2$  e  $\sigma_u^2$  sono le relative varianze, componenti il modello, rispettivamente di primo e di secondo livello.

Si deve rilevare che in base alla (14), condizionatamente alle variabili esplicative inserite nel modello, le valutazioni espresse dallo stesso studente su insegnamenti diversi sono considerate indipendenti, mentre le valutazioni di uno stesso insegnamento da parte di studenti diversi sono tra loro correlate. In altri termini:

$$\text{Corr}(Y_{ij}^*, Y_{lm}^*) = \rho = \begin{cases} 0 & j \neq m \\ \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\varepsilon^2} & j = m \end{cases} \quad (15)$$

Utilizzando l'identificazione del modello per variabili ordinali attraverso i cosiddetti *modelli soglia*, si ottiene che la variabile osservata può essere legata alla variabile latente dalle seguenti relazioni:

$$y_{ij} = k \quad \text{solo se} \quad \gamma_{k-1} < Y_{ij}^* \leq \gamma_k, \quad -\infty = \gamma_0 \leq \gamma_1 \leq \dots \leq \gamma_{k-1} \leq \gamma_k = +\infty; \quad (16)$$

$$P(Y_{ij} = k \mid u_j) = P(\gamma_{k-1} < Y_{ij}^* \leq \gamma_k \mid u_{jk}) = \Phi(\gamma_k - z_{ij}) - \Phi(\gamma_{k-1} - z_{ij}). \quad (17)$$

Nella (17):  $\Phi(\cdot)$  rappresenta la funzione di densità cumulata della normale standardizzata;  $z_{ij} = \alpha + \boldsymbol{\beta}' \mathbf{X}_{ij} + u_j$  è il predittore lineare. Ai fini dell'identificazione successiva del modello è necessario ulteriormente assumere che  $\gamma_1 = 0$  e  $\sigma_\varepsilon^2 = 1$ . Pertanto, le equazioni dalla (13) alla (17) definiscono un *modello probit ordinale a componenti di varianza* e, conseguentemente, il *coefficiente di correlazione intraclasse*  $\rho$  può essere interpretato anche come misura dell'entità della variabilità tra insegnamenti, in rapporto a quella complessiva. Quindi, come per gli indici di dispersione  $\eta$  identificati nella (11) e (12), anche valori crescenti del coefficiente  $\rho$  mettono in evidenza che le eventuali differenze nella valutazione espressa dagli studenti sono da imputare sempre maggiormente alle caratteristiche ed alle modalità dell'insegnamento impartito. A questo punto è naturale, per ciascun *item* del questionario, tentare inizialmente la stima di un modello senza covariate esplicative  $Y_{ijq}^* = \alpha_q + \mu_{jq} + \varepsilon_{ijq}$ , detto per l'appunto *modello nullo*, dal quale ricavare il relativo valore di  $\rho_q$ . In tale prospettiva si può costruire un altro indicatore lordo multivariato attraverso:

$$I''_j{}^l = \frac{\sum_{q=1}^Q \hat{\mu}_{jq} \rho_q}{\sum_{q=1}^Q \rho_q}. \quad (18)$$

## 7. RISULTATI

Per l'applicazione empirica dell'indicatore statistico qualitativo<sup>19</sup>  $I_j^l$  proposto nel presente lavoro sono stati elaborati tutti i 1.148 questionari somministrati nel secondo semestre dell'anno accademico 2000-2001 nella Facoltà di Economia dell'Università di Foggia, per un totale di 28 insegnamenti. Delle 27 domande di cui era composto ciascuno dei questionari, sono state considerate solo quelle direttamente riferibili alla qualità dell'insegnamento impartito dal docente<sup>20</sup>. Ai fini comparativi è stato anche calcolato l'indicatore quantitativo  $I_j^l$  sia nel caso di preventiva quantificazione determinata diretta<sup>21</sup> (in sintesi q.d.d.), sia nell'ipotesi di quantificazione determinata indiretta<sup>22</sup> (in sintesi q.d.i.). I valori ottenuti sono riportati nella successiva tavola 1.

Da quest'ultima è immediato verificare due importanti circostanze:

1) il livello di cograduazione dei risultati ottenuti è quasi massimo, come emerge dal calcolo degli indici di cograduazione di Gini, sempre superiori a 0,90. Ciò sottolinea la validità della proposta metodologica di sintesi dei dati qualitativi avanzata in questa sede.

2) lo scarto assoluto medio tra i risultati ricavati con i tre metodi è significativo soltanto rapportando il primo di questi ultimi rispettivamente con il secondo ed il terzo metodo (nello specifico 0,72 e 0,58 di voto), mentre lo stesso scarto è quasi inesistente tra il secondo ed il terzo metodo. Tanto evidenzia una valutazione tendenziosamente positiva del primo dei metodi testati (solo un docente totalizza in questo caso un voto finale < 7, contro i dieci degli altri due metodi).

Quanto precede mette in luce, infine, quanto già anticipato nella parte introduttiva del presente lavoro, ovvero l'assoluta necessità di standardizzare la tecnica di elaborazione statistica dei dati, tanto ai fini di un utilizzo comparativo delle informazioni raccolte.

<sup>19</sup> Utilizzando  $V = 16$  sotto-quadranti, così come suddivisi nel grafico di figura 5b.

<sup>20</sup> Escludendo, pertanto, sia le domande preliminari di natura anagrafica che quelle riguardanti il corso di Laurea in generale.

<sup>21</sup> Secondo le indicazioni fornite dal gruppo di ricerca ministeriale (Chiandotto e Gola, 1999) e, quindi, associando alle quattro modalità qualitative rispettivamente i valori numerici 2, 5, 7 e 10.

<sup>22</sup> Ipotizzando la curva normale (variabile latente) quale funzione rappresentativa del *continuum* di valori sotteso al giudizio qualitativo degli studenti.

TAVOLA 1  
*Valutazione e graduatoria di 28 insegnamenti della Facoltà di Economia*

Codice insegnamento	Voto Finale $I'_j$ q.d.d.	Voto Finale $I'_j$ q.d.i.	Voto Finale $I_j$	Graduatoria $I'_j$ q.d.d.	Graduatoria $I'_j$ q.d.i.	Graduatoria $I_j$
309	7,5928458	6,6757501	6,5000000	20	20	21
310	7,4721116	6,5370834	6,4166667	21	21	23
311	8,1756623	7,4496067	7,8333333	14	14	14
312	8,2592638	7,5761005	8,1666667	13	12	12
313	7,2677265	6,2520523	6,0000000	25	25	27
314	7,0452615	5,9578717	6,1666667	27	27	24
315	8,2963152	7,6419552	8,1666667	9	9	11
316	9,6173390	9,4825136	9,8333333	3	3	2
317	7,7532706	6,9076609	7,1666667	19	19	18
318	7,3848066	6,3692004	6,4166667	23	23	22
319	8,5070421	7,9222060	8,3333333	8	8	9
320	8,2814082	7,6367945	8,5000000	10	10	8
321	8,6349039	8,1157637	8,2500000	7	7	10
322	8,1278594	7,4280467	7,5000000	15	15	15
323	7,2761351	6,2783649	6,5000000	24	24	20
324	7,8672758	7,0519738	6,9166667	17	17	19
325	9,1464619	8,8196508	9,4166667	4	4	4
326	8,8563243	8,4174327	9,0000000	6	6	6
327	7,8132028	7,0064696	7,3333333	18	18	16
328	7,4674705	6,5072686	6,0833333	22	22	25
329	8,2651362	7,5637511	8,5833333	12	13	7
330	9,7529617	9,6672862	9,8333333	1	1	1
331	9,7248649	9,6486082	9,6666667	2	2	3
332	4,4404099	2,6970996	3,1666667	28	28	28
333	9,1037438	8,7549262	9,3333333	5	5	5
334	8,0484606	7,3116257	7,1666667	16	16	17
335	7,0942366	6,0338045	6,0000000	26	26	26
336	8,2669626	7,5874446	7,9166667	11	11	13

## 8. CONCLUSIONI

La preventiva quantificazione delle modalità dei caratteri qualitativi ordinali, allo scopo di una migliore trattazione statistica di questi ultimi, non sempre risulta necessaria. In tal senso, l'indicatore multidimensionale proposto  $I'_j$ , pur mantenendo la natura qualitativa dei caratteri osservati, produce ugualmente un valore numerico di sintesi utile sia per la valutazione complessiva del livello qualitativo del fenomeno osservato, sia per le successive elaborazioni quantitative. Questo risultato, tra l'altro, viene raggiunto senza fissare prioritariamente vincoli teorici e senza scontare l'arbitrarietà derivante da eventuali scelte individuali. Lo specifico problema della valutazione della qualità della didattica universitaria rappresenta una delle situazioni operative esemplari nelle quali l'indicatore presentato in questo lavoro trova la sua ideale collocazione.

Le applicazioni empiriche effettuate, oltre a confermare l'ovvia preferenza delle metodologie di quantificazione indiretta in luogo di quelle di quantificazione

diretta, mettono in evidenza la notevole concordanza dei risultati ottenuti con le suddette tecniche statistiche con quelli derivanti dalla metodologia in questa sede avanzata. Infatti, come riscontrato nel precedente paragrafo elaborando i dati dell'Università di Foggia, se il Nucleo locale di valutazione avesse avuto l'intenzione di incentivare soltanto i primi dieci docenti della graduatoria di merito (attuando quindi una politica premiante), l'utilizzo di uno qualunque degli indicatori testati avrebbe indotto alle medesime decisioni finali. Di converso, l'elevatissima concordanza tra risultati (graduatorie) non sempre risulta accompagnata dall'identità dei valori assoluti da cui le medesime graduatorie derivano. Pertanto, se gli stessi organi di controllo avessero invece voluto incentivare tutti quei docenti con un punteggio finale superiore ad una certa soglia minima in termini qualitativi (ad esempio voto finale  $> 7$ ), così come tra l'altro prescrive la stessa legge 370/99, è facile verificare la netta tendenziosità positiva della tecnica di quantificazione determinata diretta, il cui utilizzo condurrebbe molti insegnamenti ad essere erroneamente premiati. Dunque, a indicatori differenti corrispondono inevitabilmente altrettante valutazioni differenti le quali non consentono, ovviamente, la realizzazione di oggettivi confronti (per esempio tra Facoltà, tra Atenei, ecc...), a meno che tutti gli organi delegati alla valutazione non adottino a priori uno standard metodologico uniforme e rigidamente esplicitato.

*Dipartimento di Scienze economiche  
matematiche e statistiche  
Università di Foggia*

MASSIMO ALFONSO RUSSO

#### RINGRAZIAMENTI

L'autore desidera ringraziare il Prof. Bruno Chiandotto, la Prof.ssa Carla Rampichini, il Dott. Gabriele Soffritti per i preziosissimi miglioramenti apportati al presente lavoro e, in particolare, il Prof. Italo Scardovi per la paziente opera di incoraggiamento scientifico. In ogni caso, il contenuto di quanto riportato nel presente lavoro è da attribuire esclusivamente alla personale responsabilità dell'autore stesso.

#### RIFERIMENTI BIBLIOGRAFICI

- A.A. V.V. (1998), *Valutazione della didattica da parte degli studenti*, "Rapporto Osservatorio Nazionale per la Valutazione del Sistema Universitario", MURST, Rapporto di Ricerca n. 1, Roma.
- V. CAPURSI, M. PORCU (2001), *La didattica universitaria valutata dagli studenti: un indicatore basato su misure di distanza fra distribuzioni di giudizi*, Atti del Convegno intermedio della Società Italiana di Statistica "Processi e metodi statistici di valutazione", Roma.
- B. CHIANDOTTO, M.M. GOLA (1999), Questionario di base da utilizzare per l'attuazione di un programma per la valutazione della didattica da parte degli studenti, Osservatorio Nazionale per la Valutazione del Sistema Universitario, Roma.
- M. CIVARDI (2002), *Le opinioni degli studenti sull'attività didattica dell'Ateneo nell'anno accademico 2000-2001*, "Relazione del Nucleo di valutazione dell'Università degli studi di Milano-Bicocca", Milano.

- A. DE LUCA (2001), *Confronto tra scale verbali, simboliche e numeriche per la rilevazione della qualità percepita*, in L. Fabbri, "CAPTOR 2000: Qualità della didattica e sistemi computer assisted", Cleup, Padova, pp. 132-143.
- F. DELVECCHIO (2000), *Statistica per la ricerca sociale*, Cacucci Editore, Bari.
- F. DELVECCHIO (2000), *Scale di misura e indicatori sociali*, Cacucci Editore, Bari.
- L. FABBRIS (2000), *Un indice per misurare il divario tra qualità attuale e quella ottimale di un servizio*, in M. Civardi e L. Fabbri, "Valutazione della didattica con sistemi computer-assisted", Cleup, Padova, pp. 169-178.
- L. FABBRIS (2002), *La misura della student satisfaction per la valutazione della qualità della didattica*, in F. Delvecchio e L. Carli Sardi, "Indicatori e metodi per l'analisi dei percorsi universitari e post-universitari", Cleup, Padova, pp. 1-20.
- L. FABBRIS, C. GASPAROTTO (2001), *Modelli di valutazione della qualità della didattica universitaria*, in L. Fabbri, "CAPTOR 2000: Qualità della didattica e sistemi computer assisted", Cleup, Padova, pp. 27-44.
- A. FIELDING (1997), *On scoring ordered classifications*, "British Journal of Mathematical and Statistical Psychology", n. 50, pp. 285-307.
- H. GOLDSTEIN (1995), *Multilevel Statistical Models*, Edward Arnold, Londra.
- E. GORI, G. VITTADINI (1999), *La valutazione dell'efficienza ed efficacia dei servizi alla persona. Impostazione e metodi* in E. Gori e G. Vittadini, "Qualità e valutazione nei servizi di pubblica utilità", ETAS, pp. 121-241.
- L. GRILLI, C. RAMPICHINI (2000), *Scomposizione della dispersione per variabili statistiche ordinali*, in corso di pubblicazione su "Statistica", Clueb, Bologna.
- J. HOX (2002) *Multilevel Analysis, Techniques and Applications*, Lawrence Erlbaum Associates Inc. Publishers, New Jersey USA.
- M. LALLA, G. FACCHINETTI (2001), *Scale e insiemi sfocati per valutare la didattica universitaria*, Atti del Convegno intermedio della Società Italiana di Statistica "Processi e metodi statistici di valutazione", Roma.
- G. LETI (1979), *Distanze e indici statistici*, La Goliardica, Roma.
- G. LETI (1983), *Statistica Descrittiva*, Il Mulino, Bologna.
- G. MARBACH (1974), Sulla presunta equidistanza degli intervalli nelle scale di valutazione, "Metron", Vol. XXXII n. 1-4.
- L. MODICA, E. STEFANI (1997), *Valutazione delle attività didattiche universitarie. Le esperienze condotte dalla CRUI*, CRUI, Roma.
- R. PICCARRETA (2001), *A new measure of nominal-ordinal association*, "Journal of Applied Statistics", n. 28, pp. 107-120.
- C. RAMPICHINI, A. PETRUCCI (2000), *Indicatori statistici per la valutazione della didattica universitaria*, in M. Civardi e L. Fabbri, "Valutazione della didattica con sistemi computer-assisted", Cleup, Padova, pp. 47-72.
- C. RAMPICHINI, L. GRILLI, A. PETRUCCI (2002), *Analysis of university course evaluation: from descriptive measures to multilevel models*, Proposto per la pubblicazione.
- T.A.B. SNIJDERS, R.J. BOSKER (1999), *Multilevel Analysis. An introduction to basic and advanced multilevel modelling*, Sage, Londra.
- W.S. TORGERSON (1958), *Theory and Methods of Scaling*, Wiley and Sons, New York.
- M.R. YOUNG (1995), A parametric procedure for ultrametric tree estimation from conditional rank order proximity data, "Psychometrika", Vol. 60, n. 1, pp. 45-75.
- A. ZANELLA, M. CERRI (2000), *La misura di customer satisfaction: Qualche riflessione sulla scelta delle scale di punteggio* in: Valutazione della Qualità e Customer Satisfaction, il Ruolo della Statistica-Misure e Modelli interpretativi, "Vita e Pensiero", Milano, pp. 217-231.

## RIASSUNTO

*Indicatori statistici per la valutazione della qualità della didattica universitaria: una proposta metodologica*

La legge 370/99 ha impresso un forte impulso alla ricerca di tecniche utili all'individuazione di indicatori capaci di misurare la qualità della didattica universitaria, ormai diventata un obiettivo fondamentale per l'ottimizzazione dell'offerta formativa. Tuttavia, esistono numerose metodologie statistiche per l'identificazione di appropriati indicatori. In questo lavoro vengono dapprima analizzate le principali tecniche quantitative in grado di fornire indicatori lordi multidimensionali, per poi proporre una procedura alternativa. Quest'ultima si caratterizza per il fatto che il valore di sintesi da essa ottenuto non richiede la preventiva trasformazione numerica delle modalità delle variabili qualitative ordinali osservate. Inoltre, essa risulta indipendente sia da vincoli di natura teorica che dall'arbitrio delle scelte individuali. Un'applicazione empirica del nuovo indicatore qualitativo viene infine effettuata, a scopi comparativi, utilizzando i dati derivanti dallo spoglio dei questionari sottoposti agli studenti frequentanti la Facoltà di Economia dell'Università di Foggia per la periodica valutazione della qualità della didattica, relativamente all'anno accademico 2000-2001.

## SUMMARY

*Statistical indicators for the evaluation of university teaching quality: a methodological proposal*

The measurement of the teaching quality is nowadays one of the basic target if the university training offer wants to be optimised. With respect to this, the law n. 370/99 represented a strong incentive to the search for techniques aimed to the identification of indicators which are capable to adequately measure the university teaching quality. However, numerous statistical methods exist which attempt to identify appropriate indicators. In this work we firstly analyse the main quantitative techniques from which it is possible to achieve multidimensional gross indicators. Afterwards, we propose an alternative methodological procedure. Its main feature is that the resulting quantitative measure of synthesis does not require any preventive numerical transformation of the modalities of the ordinal qualitative variables observed. Furthermore, it is independent from both theoretical constraints and the arbitrary effect of individual choices. Lastly, as an empirical application and for comparative purposes, we carry out an analysis while using the new qualitative indicator. To this end, we use data from questionnaires submitted to the students of the University of Foggia during the academic year 2000-2001.