

SOCIETÀ ITALIANA

Atti del convegno

La Statistica per le imprese

Sessioni plenarie
Sessioni specializzate
Tavole rotonde
volume 1

Convegno SIS
Torino 2-4 aprile 1997



UNA METODOLOGIA STATISTICA PER LA PERFORMANCE EVALUATION DEI SERVIZI ALLA PERSONA DI PUBBLICA UTILITÀ

*A statistical methodology for Performance Evaluation of Public Utility
Service to People*

GIORGIO VITTADINI

Università degli Studi di Milano

Facoltà di Scienze Politiche - Istituto di Scienze Statistiche e Matematiche "M. Boldrini"

1. Introduzione

Si definiscono (Vittadini, 1996) come *servizi alla persona di pubblica utilità* (SPPU) particolari servizi, quali quelli prestati nel campo dell'istruzione, della sanità, dell'assistenza e della cultura, che presentano le seguenti caratteristiche (Erba e Martini, 1989): a) rispondono all'esigenza di un incremento del benessere degli utenti (⁽¹⁾); b) sono personalizzati (in quanto l'utente determina la modalità dell'erogazione); c) necessitano dell'interazione dell'agente con l'utente per tutta la durata del processo attraverso cui si esplica il servizio.

Tali caratteristiche peculiari comportano una differenziazione tra *output* riferiti ai prodotti dei SPPU e *outcome* intesi come risultati sintetici, durevoli e personalizzati (Aitkin e Longford, 1986) dell'erogazione dei SPPU su un aspetto del benessere dell'utente (della salute, dell'istruzione, della cultura, ecc.).

Si definisce, inoltre, sempre a riguardo di particolari aspetti latenti di benessere dell'utente, lo *stato iniziale* dell'utente, costruito teorico dalle caratteristiche analoghe a quelle dell'*outcome*, ma inerente la situazione prima dell'erogazione del SPPU.

Perciò la qualità dei SPPU (Vittadini, 1996) non può essere studiata esclusivamente mediante tecniche di controllo di qualità in un'ottica di efficienza (rapporto tra mezzi impiegati e scopi aziendali raggiunti) e di qualità negativa (definita come assenza di difetti), deve essere anche analizzata in un'ottica di *performance evaluation* e di qualità positiva (efficacia dei risultati rispetto agli utenti) (Martini A., 1996).

Tale valutazione coincide con la misura dell'incremento di benessere avvenuto tra lo *stato iniziale* e lo *stato finale* dell'utente (*outcome*) e consente lo studio dell'impatto complessivo dei servizi su un collettivo sociale. Inoltre rende possibile la verifica dell'efficacia relativa di differenti strutture eroganti SPPU o di differenti modalità secondo cui sono erogati (Gori 1992a, 1992b, 1993; Goldstein e Spiegelhalter, 1996).

Il presente contributo si pone in questa seconda ottica proponendo una nuova metodologia statistica basata sulle variabili latenti, che si aggiunge ad altre utilizzate nella valutazione della qualità in numerose discipline applicate (Vittadini, 1996), in particolare a quelle che fanno capo al "*multilevel model*" (Gori, 1996; Goldstein e Spiegelhalter, 1996).

2. La metodologia prescelta per l'analisi

Per lo studio della *performance evaluation*, le caratteristiche dei SPPU suggeriscono di esprimere gli *stati iniziali* dell'utente e gli *stati finali* corrispondenti come variabili latenti misurate prima e dopo l'erogazione del SPPU, a meno di errori, attraverso indicatori longitudinali osservati (Vittadini, 1996).

Le ragioni di tale scelta sono le seguenti:

- A) La definizione statistica del "costrutto teorico" variabile latente è particolarmente conflante a quella di *stato iniziale* dell'utente e di *outcome*.
- B) I punteggi delle variabili latenti sono definiti sulle singole persone e perciò tengono conto del fatto che gli *outcome* sono necessariamente personalizzati e difficilmente riconducibili a *standard* prefissati.
- C) La longitudinalità rispetta il criterio di valutare gli aspetti latenti del benessere di un individuo lungo un arco temporale.
- D) E' possibile eliminare l'effetto del *bias* individuale costruendo misure basate sulle differenze fra *outcome* e corrispondenti *stati iniziali* dell'utente.

Quanto detto, offrendo indicazioni di massima per una analisi dell'efficacia relativa dei SPPU, apre anche ad un suo affronto operativo.

- a) Per ricavare variabili latenti quantitative che descrivano aspetti latenti di benessere, occorre innanzitutto scegliere opportuni insiemi di indicatori osservati, che devono essere rilevati prima e dopo l'erogazione del SPPU e possono essere sia quantitativi che qualitativi. Per rendere questi ultimi quantitativi, si utilizza la *Rasch Analysis* particolare versione di *latent trait analysis* (Rasch 1960; Wright e Masters, 1982). Attraverso essa si ricavano punteggi *logit* che permettono di superare i problemi di non confrontabilità delle scale ordinali.

- b) Secondariamente, in modo originale rispetto alla *Rasch Analysis*, si ricavano gli aspetti latenti del benessere come trasformate lineari degli indicatori osservati (Haagen e Vittadini, 1991; Vittadini e Haagen, 1997). Tali combinazioni lineari hanno proprietà analoghe alle "cause latenti" ottenute secondo la tradizionale teoria dei modelli strutturali (Leveritt, 1984) ma, a differenza di esse, non sono affette da problemi di non identificabilità dei parametri e di indeterminazione degli *scores*.

Le combinazioni lineari sono ricavate in modo da assegnare ai singoli indicatori una ponderazione coerente con gli scopi dell'analisi.

Nel caso degli *stati iniziali*, ove sono rilevanti le peculiarità individuali dei singoli utenti e le caratteristiche del benessere descritte dagli indicatori, si utilizzano le componenti principali in cui ogni indicatore osservato ha una ponderazione proporzionale al suo contributo alla variabilità complessiva.

Nel caso degli *outcome*, dato che lo scopo dell'analisi è quello di esaminare l'efficacia relativa di diversi trattamenti o strutture ove sono erogati i SPPU, si tiene conto anche della suddivisione degli individui in differenti gruppi. Perciò si utilizzano funzioni discriminanti, ove ogni indicatore assume una ponderazione proporzionale al suo contributo al loro potere discriminante.

- c) Al fine di eliminare il *bias*, dovuto ad una distribuzione non casuale di utenti con diversi *stati iniziali* del benessere tra differenti strutture eroganti SPPU (ad esempio, ragazzi più dotati che accedono a determinate scuole, persone in migliori condizioni fisiche che accedono a certi ospedali), si costruisce una nuova variabile definita come differenza standardizzata tra *stati iniziali* e *finali* riferiti allo stesso aspetto del benessere.

- d) Infine, per verificare l'efficacia relativa delle diverse strutture o trattamenti, si effettuano analisi delle varianze o delle covarianze ove le variabili latenti depurate dal *bias* sono le variabili dipendenti, e le differenti tipologie di trattamenti o strutture da cui è erogato il SPPU, le variabili indipendenti.

3. Indicatori osservati e aspetti latenti del benessere

3.1. Gli effetti degli output sullo stato di benessere dell'utente

Gli indicatori osservati, che forniscono informazioni utili a riguardo delle variabili latenti inerenti aspetti particolari del benessere, possono essere classificate secondo differenti criteri quali: la tipologia; le informazioni di base o di particolari aspetti del benessere, come, ad esempio, indicatori "*Activity for Day Living*"; il carattere qualitativo o quantitativo; l'origine oggettiva o soggettiva (ad esempio, misure di *customer satisfaction*); la natura semplice o composta; la fonte (ad esempio, questionari o osservazioni del rilevatore); la disciplina di appartenenza (ad esempio, sanitaria, psicologica, psichiatrica, demografica, economica, ecc.).

Pur nella loro differente natura, tutti gli indicatori osservati descrivono caratteristiche dello stato di benessere di un individuo e sono raggruppabili per insiemi che sottendono ognuno un diverso aspetto latente del benessere stesso.

Tali aspetti si possono modificare in seguito all'erogazione del SPPU; perciò ognuno di tali indicatori, se misurati sugli utenti dopo l'erogazione del SPPU, mette in luce gli effetti degli *output* su un particolare aspetto dello *stato finale* dell'utente. Gli indicatori osservati di benessere costituiscono perciò la fonte informativa attraverso cui quantificare gli *stati iniziali* dell'utente e gli *outcome* in legame intrinseco con gli *output*.

3.2. Proprietà di indicatori osservati di benessere quantitativi

- a) Un indicatore osservato di benessere quantitativo deve essere crescente all'aumentare del livello di benessere considerato (ad esempio, decimi di vista). Se ciò non avviene in via naturale (ad esempio, pressione sanguigna il cui valore inerente il maggior benessere non è il valore massimo), si devono operare opportune trasformazioni⁽²⁾. Inoltre, essendo spesso gli indicatori espressi in unità di misura e ordini di grandezza differenti, per poter essere utilizzati simultaneamente devono essere standardizzati.

- b) Per ciò che concerne il legame esistente tra differenti indicatori inerenti un medesimo aspetto latente di benessere risulta utile rifarsi al concetto di "unidimensionalità", espresso come elemento base della *Rasch Analysis* (Rasch, 1960),

tecnica statistica avente come scopo quello di rendere confrontabili scale ordinali provenienti da prove attitudinali e questionari, ecc.. In tale impostazione "prescrittiva" e non descrittiva (Wright e Masters, 1982), ogni insieme di indicatori sottende uno e un solo aspetto latente di benessere e, perciò, tra ogni coppia di indicatori, prima dell'erogazione del SPPU, può esistere solo una relazione diretta.

Ciò non significa che non possano essere misurati effetti negativi dell'erogazione del SPPU sugli utenti, per due ragioni. Innanzitutto possono risultare relazioni di tipo inverso tra variabili latenti descrittivi aspetti diversi del benessere (ad esempio, grado di preparazione complessivo precedente di uno studente e propensione attuale ad imparare). Secondariamente, se a causa dell'unidimensionalità, prima dell'erogazione del SPPU, è necessaria l'esistenza di un legame diretto tra ogni coppia di indicatori osservati che sottendono il medesimo aspetto latente, dopo l'erogazione possono risultare anche legami inversi tra essi dovuti a effetti non desiderati e perversi del SPPU.

3.3. Criteri per la costruzione di indicatori quantitativi e oggettivi

Fra gli indicatori osservati di benessere, particolare importanza assumono gli indicatori qualitativi che descrivono riuscita in prove, gradazioni di preferenze, soddisfazioni, propensioni verso determinate attività o situazioni (rilevate in prove attitudinali, test, questionari, ecc.). Per utilizzarli nell'analisi dell'efficacia relativa occorre trasformarli in indicatori quantitativi, definiti su numeri naturali o reali. A tale scopo è stata proposta la già citata *Rasch Analysis* (Rasch, 1960), tecnica appartenente alla più ampia famiglia dei *latent trait models*. In particolare, la *Rasch Analysis* si propone di interpretare i risultati in prove attitudinali (o questionari) di un insieme di utenti esclusivamente sulla base dell'abilità degli individui, della difficoltà delle prove, dei diversi livelli di riuscita. Tale risultato è raggiunto innanzitutto attraverso la scelta di prove attitudinali o questionari e, quindi, di indicatori di risultati che sottendono un solo aspetto latente (ad esempio, capacità motoria o capacità intellettuale di disabili), e, secondariamente, attraverso l'utilizzo di un modello logistico che rende additive e quantitative le scale ordinali attraverso cui sono espressi i risultati delle prove o dei questionari (Tesi et al., 1996).

Il modello logistico è qui presentato in una delle versioni più comuni e complete, denominata *partial credit model* (Andrich, 1979; Wright e Masters, 1982). Si considerino n individui sottoposti a m prove attitudinali. Ognuno degli individui in ogni prova può avere un livello di riuscita esprimibile attraverso una scala ordinale che va da 0 a s_k .

Si ipotizza che in una prova k ($k=1, 2, \dots, m$) che sottende l'aspetto latente α ($\alpha=1, 2, \dots, n_2$), la probabilità di riuscita π_{ijk} ($w=0, \dots, j, \dots, s_k$) di un individuo i ($i=1, 2, \dots, n$) dipenda esclusivamente da tre parametri incogniti: abilità dell'individuo θ_i , difficoltà della prova d_k , difficoltà nel passare da un livello di riuscita $j-1$ a un livello di riuscita j (z_j) (con difficoltà totale nel passare da un livello di riuscita 0 ad un livello di riuscita j data da

$(z_j^* = z_0 + z_1 + \dots + z_{j-1} + z_j)$. Il modello ha la seguente configurazione (Andrich, 1979; Wright e Masters, 1982):

$$\pi_{ijk} = \exp \sum_{w=0}^j [\theta_i - (d_k + z_w)] / \sum_{w=0}^{s_k} \exp \sum_{w=0}^w [\theta_i - (d_k + z_w)] \quad [1]$$

ove i è il generico livello di riuscita $w=l$ mentre $z_0 = 0$ cosicché $\exp \sum_{w=0}^0 [\theta_i - (d_k + z_w)] = 1$ ed è riesprimibile in termini di punteggi *logit*:

$$x_{ijk} = \text{logit} (\pi_{ijk} / 1 - \pi_{ijk}) = \sum_{w=0}^j [\theta_i - (d_k + z_w)] \quad [2]$$

Utilizzando come informazioni iniziali per i parametri θ_i , d_k , z_j la quota di risposte fornite per ogni livello di difficoltà j di ogni prova k da ogni soggetto i e, per ogni prova k e ogni livello di difficoltà j , la quota di risposte fornita da tutti gli individui, si ottengono, attraverso opportuni algoritmi, le stime di massima verosimiglianza $\hat{\theta}_i$, $\hat{d}_k$, $\hat{z}_j$ che rendono più probabile l'insieme di punteggi ottenuti dagli n individui nelle m prove.

A questo punto si stimano le probabilità [1] e i punteggi *logit* [2].

Grazie ad un teorema (Rasch, 1960), che dimostra la possibilità di stimare indipendentemente gli uni dagli altri i parametri inerenti l'abilità dei soggetti e le difficoltà delle prove, la *Rasch Analysis* consente di considerare i punteggi [2] quali "objective measures", indipendenti dal campione di individui e dalla tipologia della prova, superando così i problemi della non confrontabilità delle valutazioni espresse con scale ordinali.

La *Rasch Analysis* si conclude verificando i casi di *misfit*, in cui le risposte prodotte da ogni soggetto o ricevute da ogni prova si discostano in modo significativo dalle probabilità ottenute con il metodo descritto (Tesi et al., 1996).

3.4. Gli indicatori di benessere come combinazioni lineari delle variabili osservate

A differenza di quanto accade nella *Rasch Analysis*, nel caso in questione il problema non è solo quello di ottenere misure oggettive dell'abilità del soggetto e delle difficoltà delle prove. Si vogliono anche ricavare, secondo la tradizionale accezione (Everitt, 1984), variabili latenti che permettano di verificare, attraverso i loro *scores* definiti su ogni utente, l'efficacia relativa delle diverse strutture.

Tuttavia, nella letteratura statistica è stato dimostrato che quando si ricavano "cause latenti" dagli indicatori osservati non si ottengono soluzioni uniche. Infatti, in questo caso, data una distribuzione delle variabili osservate, i modelli non sono, sotto condizioni generali, identificabili nei parametri (vedi per il modello LISREL lo stesso Joreskog, 1981)

e, anche se identificati nei parametri, sono comunque indeterminati negli *scores* (Vittadini, 1988).

Il problema si aggrava in contesti in cui si dispone di informazioni a priori e di vincoli desunti dalla realtà e da precedenti studi. Non esistendo condizioni sufficienti per l'identificabilità dei modelli, generalmente non solo non si modificano le soluzioni in funzione delle informazioni a priori e dei vincoli posti dalla realtà, ma anzi si introduce un insieme di vincoli diversi da essi all'unico scopo di rendere identificabili i parametri (vedi per il modello LISREL ancora lo stesso Joreskog, 1981). La ricerca di cause latenti di variabili osservate nei modelli strutturali è quindi suggestiva, ma non fornisce adeguate garanzie di tipo statistico-matematico. È stato perciò proposto di ricavare le variabili latenti come combinazioni lineari delle variabili osservate. Tra i diversi metodi atti a tale scopo, quali il *Partial Least Squares* (P.L.S.) (Wold, 1980; Apel e Wold, 1982), la *Regression Component Decomposition* (R.C.D.) (Haagen e Vittadini, 1991), si opta per la *Regression Component Decomposition Restricted* (R.C.D.R.) (Vittadini e Haagen, 1997), estensione della R.C.D., che, attraverso una decomposizione dello spazio generato dagli indicatori, permette di ricavare le variabili latenti come loro combinazioni lineari, dalle proprietà analoghe alle "cause latenti".

Ottenuti punteggi *logit* sintetici per ogni individuo i ed ogni prova k relativa ad un aspetto del benessere α

$$X_{i(k)\alpha} = \sum_{j=0}^n x_{i(jk)\alpha} \quad [3]$$

e considerando anche gli indicatori osservati quantitativi dopo averli standardizzati, si definisce il seguente modello di misura:

$$X_{i\alpha} = L_{X_{i\alpha}} Z_{i\alpha} + D_{i\alpha} \quad (X_{i\alpha} = [X_{0_{i\alpha}}, X_{1_{i\alpha}}, \dots, X_{n_{i\alpha}}]') \quad [4]$$

ove: $t=0$ e $t=1$ sono rispettivamente il tempo iniziale prima dell'erogazione del SPPU e il tempo finale dopo l'erogazione del SPPU; $X_{i\alpha}$ le matrici degli indicatori quantitativi al tempo t ; $Z_{i\alpha}$ le matrici dello stato del benessere α al tempo t ; $D_{i\alpha}$ le matrici degli errori di misura al tempo t ; $L_{X_{i\alpha}}$ le matrici dei parametri.

Si ipotizza incorrelazione fra variabili latenti Z_i ed errori D_i e valore atteso nullo sia per le variabili osservate che per le variabili latenti ed errori. Inoltre, come già avviene nel P.L.S. di Wold (1980), si ricava ogni variabile latente $Z_{i\alpha}$ esclusivamente dai suoi indicatori osservati $X_{i\alpha}$, in luogo di ottenere simultaneamente tutte le variabili latenti Z_i da tutti gli indicatori X_i . Il modello può essere, inoltre, raffinato in modo da tenere conto di vincoli e informazioni a priori (Vittadini e Haagen, 1997).

Per ciò che concerne i metodi attraverso cui ricavare le n_z combinazioni lineari $Z_{0\alpha}$ e $Z_{1\alpha}$ ($\alpha=1, \dots, n_z$), occorre ricordare che uno dei problemi da affrontare nella valutazione della qualità è quello della ponderazione degli indicatori osservati utilizzati. In altri casi, come nel *multilevel*, tale problema è risolto attraverso opportuni e raffinati metodi di ponderazione (Gori, 1996). Nel caso in questione, per evitare arbitrarietà e

manenza di oggettività si attribuiscono le ponderazioni tenendo conto dell'obiettivo di effettuare un'analisi dell'efficacia relativa di strutture e trattamenti.

Per quanto riguarda gli *stati iniziali* dell'utente è ragionevole ipotizzare che gli individui siano distribuiti, rispetto agli indicatori osservati di benessere, in modo casuale tra varie strutture o modalità eroganti il SPPU. In questa situazione appare opportuno costruire la combinazione lineare con l'obiettivo di massimizzare la sua variabilità complessiva, che permette di cogliere le differenze interpersonali fra gli utenti del SPPU. Ciò equivale a ricavare ogni *stato iniziale* dell'utente dalla matrice di correlazione degli indicatori osservati al tempo $t=0$ $S_{X_{0\alpha}}$ come componente principale $Z_{i\alpha}$ ($\alpha = 1, \dots, n_z$), in cui ogni indicatore osservato ha una ponderazione proporzionale al contributo alla variabilità complessiva.

Inoltre, poiché per l'impostazione "prescrittiva" della *Rasch Analysis* descritta al paragrafo 3.2, le covarianze tra gli indicatori $X_{0\alpha}$ sono tutte positive e, quindi, $S_{X_{0\alpha}}$ ha tutti elementi positivi, per un noto teorema di algebra lineare (Manara e Nicola, 1970) lo sono anche i coefficienti della prima componente principale ottenuti quali autovettori dell'equazione al primo autovalore. Ogni variabile latente che esprime lo *stato iniziale* dell'utente è perciò funzione crescente di tutti i suoi indicatori e di ogni loro covarianza.

Per ciò che concerne gli *outcome*, dato l'obiettivo di studiare l'efficacia relativa, occorre considerare ai fini della costruzione della combinazione lineare anche la suddivisione degli utenti tra gruppi inerenti diversi trattamenti e strutture da cui è erogato il SPPU.

Si definiscono perciò con $X_{1_{\alpha g}} = [X_{1_{\alpha g_1}}, X_{1_{\alpha g_2}}, \dots, X_{1_{\alpha g_{n_g}}}]'$ ($g = 1, \dots, n_g$) gli indicatori osservati relativi all'aspetto latente α misurati dopo l'erogazione del SPPU sugli individui appartenenti al gruppo g . Dalle matrici di covarianze fra gruppi $S_{X_{1_{\alpha g}}}$ ($g = 1, \dots, n_g$) si ricava una combinazione lineare $Z_{1\alpha}$ ($\alpha = 1, \dots, n_z$) attraverso l'analisi discriminante in modo da massimizzare la varianza spiegata fra i gruppi rispetto alla varianza residua.⁽³⁾

In questo modo ogni indicatore ha una ponderazione proporzionale al contributo al potere discriminante della funzione $Z_{1\alpha}$.

4. Costruzione di outcome depurati dal bias.

Come già accennato, può accadere frequentemente per le più svariate ragioni (ad esempio, costo, qualità, differenti possibilità di accesso) che gli individui non si distribuiscano fra diverse strutture e trattamenti in modo casuale rispetto allo *stato iniziale* dell'utente, ma che, a differenti *stati iniziali*, corrispondano differenti strutture o trattamenti. In questo caso di presenza di *bias*, la semplice scelta della funzione discriminante come metodo per ricavare gli *outcome* finirebbe per essere vanificata in quanto il valore maggiore o minore di alcuni *outcome* relativi a determinati aspetti del benessere in certi gruppi potrebbe essere dovuto a un migliore o peggiore

corrispondente *stato iniziale* dell'utente. Per evitare questo problema si definisce una nuova variabile, differenza tra ogni *outcome* e il corrispondente *stato iniziale* dell'utente:

$$W_{\alpha} = (Z_{\alpha_1} - Z_{\alpha_0}) / \{2[1 - \text{cov}(Z_{\alpha_1} - Z_{\alpha_0})]\} \quad (\alpha = 1, \dots, n_2) \quad [5]$$

Le W_{α} sono standardizzate in quanto, indicando con $E[W_{\alpha}]$ e $\sigma_{W_{\alpha}}^2$ valore atteso e varianza di W_{α} si ha:

$$E[W_{\alpha}] = E[Z_{\alpha_1}] - E[Z_{\alpha_0}] = \frac{1}{n} \sum_{i=1}^n (Z_{\alpha_1} - Z_{\alpha_0})^2 = 2[1 - \text{cov}(Z_{\alpha_1}, Z_{\alpha_0})] [6]$$

Le correlazioni fra le diverse W_{α} ($\alpha = 1, \dots, n_2$) mettono in luce se i differenti outcome depurati dal *bias* hanno legami diretti o inversi⁽⁴⁾, i valori attesi di ogni gruppo $E[W_{\alpha_g}]$ ($g = 1, \dots, n_g$) segnalano per quali gruppi di utenti i diversi *outcome* depurati dal *bias* sono maggiori o minori.

5. Verifica dell'efficacia relativa del SPPU.

Per verificare l'efficacia relativa del SPPU tra diverse strutture, oltre ad esaminare i risultati delle analisi delle componenti principali e discriminanti, si possono effettuare una serie di *analisi delle varianze* rispetto alle singole variabili W_{α} o *analisi delle varianze multivariate* tra le variabili W , considerate simultaneamente, ove i fattori qualitativi sono i gruppi degli utenti cui sia stato erogato il SPPU. Si ha quindi:

$$W = BH + G \quad [7]$$

con: H (n_g, n) matrice di *factor design* che segnala a quale gruppo appartiene l'individuo, $W = [W_1, W_2, \dots, W_{n_2}]$, $B = (n_2, n_g)$ matrice di parametri, $G = (n_2, n)$ matrice di errori.

Qualora si pongano opportuni vincoli sui parametri (Vitali, 1993), ogni equazione della [7] può essere espressa e interpretata come un modello lineare; il modello [7] nel suo complesso è un sistema lineare che considera simultaneamente l'effetto dell'erogazione del SPPU sulle dimensioni latenti W .

Si può infine ricercare quali siano i fattori qualitativi e le variabili quantitative che determinano nel complesso l'efficacia del SPPU mediante un'analisi delle covarianze delle variabili W :

$$W = AW + CH + G \quad [8]$$

con $A(n_2, n_2)$ matrice di parametri dalla diagonale principale nulla, $C(n_2, n_g)$ matrice di parametri, $H(n_g, n)$ ($n_g = n_2 + n_g$) matrice di variabili esplicative di W , composta da una sottomatrice di variabili quantitative (n_g, n) e da una matrice di *factor design* (n_g, n).

6. Analisi empirica

Si vuole verificare l'efficacia relativa di nove strutture sanitarie dislocate in cinque regioni italiane (Lombardia, Liguria, Emilia, Abruzzo e Piemonte) nell'erogazione di un particolare SPPU consistente nel trattamento riabilitativo dell'emiparesi su una popolazione totale di 912 persone colpite da ictus nel 1996. La patologia influisce sul benessere dell'individuo modificando gli aspetti non direttamente osservabili, motori e cognitivi. Perciò, al fine di ottenere indicatori osservati che, a meno di errori, misurino tali aspetti latenti, si considerano i risultati di 18 prove attitudinali (13 relative all'aspetto motorio e 5 all'aspetto cognitivo) a cui sono sottoposti gli utenti prima e dopo l'erogazione del SPPU. I risultati sono misurati secondo la scala qualitativa ordinale FIM (Functional Independence Measure) con punteggi che vanno da 1 a 7.

I dati inerenti gli indicatori qualitativi, unitamente a informazioni riguardanti la provenienza degli utenti, la durata dell'erogazione del SPPU, il tempo intercorrente tra l'insorgere dell'ictus e l'inizio della cura, sono stati forniti dal prof. Tesio della Fondazione Maugeri Clinica del Lavoro di Pavia⁽⁵⁾. I due insiemi di punteggi ordinali, che sottendono ognuno un unico aspetto latente, nel senso indicato nel paragrafo 3.3, sono trasformati mediante la *Rasch Analysis* in punteggi *logit* quantitativi. I parametri inerenti le difficoltà delle prove e le difficoltà nel passare da un livello di riuscita a un altro livello (identiche per tutte le prove) sono espressi nella Tabella 1.

A questo punto, dopo aver standardizzato i punteggi quantitativi così ottenuti, si ricavano gli *stati iniziali* dell'utente e i relativi *outcome*. Pur non essendo la distribuzione dei punteggi multinormale, essendo i dati relativi ad una popolazione di individui e non a un campione, si può adottare l'analisi discriminante secondo la metodologia proposta nella nota 3. Essendo positive le matrici di correlazioni al tempo $t=0$ $S_{X_{0\alpha}}$ e $S_{X_{0\beta}}$ e le matrici di correlazioni fra gruppi al tempo $t=1$ $S_{X_{1\alpha}}$ e $S_{X_{1\beta}}$, le variabili latenti ($Z_{0\alpha}, Z_{0\beta}$) e ($Z_{1\alpha}, Z_{1\beta}$), ricavate mediante l'analisi delle componenti principali e l'analisi discriminante, sono correlate positivamente con ogni indicatore osservato e quindi esprimono in modo opportuno i due aspetti latenti del benessere.

Buona è la quota di varianza iniziale associata alle variabili latenti inerenti la dimensione cognitiva ($\cong 0.75$); discretamente elevata quella relativa alla dimensione motoria ($\cong 0.60$).

A questo punto, ottenute le variabili risultato depurate dal *bias* iniziale W_{α} , W_{β} , si effettua un'analisi delle covarianze utilizzando come variabili esplicative l'altra dimensione latente, le strutture suddivise tra quelle inserite in ospedali generali

(1/4/5/7/9) e quelle a sé stanti (2/3/6/8), l'età, la durata della degenza, il tempo intercorso tra l'ictus e l'inizio della cura e la provenienza. Per poter interpretare tale analisi come un modello lineare si introducono opportuni vincoli.

Si osserva che la bontà di accostamento complessiva delle variabili, pur non altissima, è sufficientemente elevata ($\approx 0.40, \approx 0.50$). Tra le variabili esplicative è risultata rilevante, oltre all'altro *outcome* depurato dal *bias* (W_a per W_p e W_p per W_a), la suddivisione dei pazienti tra quelli ricoverati in strutture di cura inserite in ospedali generali e quelli assistiti in strutture a sé stanti. Questo risultato porta a concludere che si può rilevare una diversa efficacia relativa del SPPU a secondo dell'inserimento o meno delle strutture di cura dell'emiparesi in strutture ospedaliere più complesse. In particolare, l'osservazione delle medie dei gruppi di W_a e W_p porta a concludere che l'efficacia relativa è superiore nel sottoinsieme di strutture di cura inserite in ospedali generali.

Note

(1) Con il termine *benessere* si prescinde da ogni significato filosofico ed economico inerente un aspetto ultimo dell'individuo per identificare, nella tradizione dello studio delle variabili latenti, una dimensione latente particolare di natura psico-fisico-relazionale, quali ad esempio la salute, l'istruzione, la cultura, non coinvolgente l'essenza ultima della persona.

(2) Ad esempio definito l'indicatore X_a , decrescente all'aumentare delle caratteristiche di benessere α si pone:

$$X_a^b = (X_a^a)^{-1} \quad (N21)$$

o definito l'indicatore X_b , che raggiunge il suo massimo per un valore k che non è né minimo né massimo si pone:

$$X_b^b = (|X_b - k|)^{-1} \quad (N22)$$

(3) Un *outcome* può avere correlazioni negative con alcuni dei suoi indicatori senza perdere significato perché è un misuratore sintetico della loro capacità discriminante fra i gruppi. Tuttavia se tutte le correlazioni sono positive un *outcome* ha la proprietà ottimale di essere funzione crescente di tutti gli indicatori e delle loro covarianze. Ciò è possibile per il teorema citato (Manara e Nicola, 1970), quando la matrice di covarianze fra i gruppi $S_{\bar{X}_p, \bar{X}_p}$ è positiva in ogni suo elemento ed, essendo la distribuzione degli indicatori multinormale o avendo a disposizione una popolazione di individui, si può costruire l'analisi discriminante attraverso il seguente metodo alternativo a quello classico non parametrico:

W_a	W_p	$W_a(W_p)$	Tipo centro	Durata degenza	Provenienza	ETA	P
-534	-471	440		337	030	003	4089
249				482	025	002	5018

ANALISI DELLA COVARIANZA									
COEFFICIENTI VARIABILI INDIPENDENTI									
Media Gruppi									
	1	2	3	4	5	6	7	8	9
STRUTTURE INSERITE IN OSPEDALI GENERALI		.44	.43			.40		.42	0.42
STRUTTURE NON INSERITE IN OSPEDALI GENERALI				.48	.46		.55	.58	0.50

Z ₁ Difficoltà nel passare da riuscita a riuscita j in Rasch Analysis.	Z ₂ Difficoltà delle prove nella Rasch Analysis.	Dimensione latente motoria indicatori qualitativi	Analisi componenti principali coeff.	Analisi discriminante coefficienti	Quota 1° autorevole sul totale	RASCII ANALYSIS									
						Z ₁₀	Z ₀₉	Z ₁₀	Z ₀₉	Z ₁₀	Z ₀₉	Z ₁₀	Z ₀₉	Z ₁₀	Z ₀₉
-1.69	1.12	Nutritsi	0.78	2.68	.655										
-1.86	1.05	Rassettarsi	1.16	2.78	.596										
-1.44	.83	Lavarsi	1.22	434											
-1.11	.75	Vestirsi dalla vita in su	1.42	387											
-1.11	.71	Vestirsi dalla vita in giù	1.43	117											
.27	.52	Igiene perineale	1.51	115											
.89	.33	Vescica	1.02	102											
2.10	.26	Alvo	.088	153											
	.13	Trasferimenti letto/sedia/carrozzina	.152	.076											
	-1.12	Trasferimenti su wc	.158	.148											
	-.20	Trasferimenti vasca/doccia	.109	.649											
	-.36	Cammino	.05	.047											
		Scale		.07											
		Dimensione latente cognitiva indicatori qualitativi													
	-.64	Comprensione	.238	.455	.733										
	-.76	Espressione	.225	.428											
	-.80	Rapporto con gli altri	.238	.430											
	-1.03	Soluzione di problemi	.241	.462											
	-1.15	Memoria	.258	.460											

Tabella 2 Risultati dell'analisi empirica

Tab. I.

	Metodo Classico Non Parametrico [7]	Metodo Alternativo
Massimizzazione [funz. obiettivo/vincolo]	$[v_1 S_{X_{ip}} v_1] / [v_1 S_{X_{ip}}^* v_1]$	$\left[\frac{\tau_1 (S_{X_{ip}} - S_{X_{ip}}^*)}{\tau_1 \tau_1'} \right] / \tau_1 \tau_1' = \tau S_{X_{ip}} \tau_1' / \tau_1 \tau_1'$
Lagrangiana	$v_1 S_{X_{ip}} v_1' - \lambda (v_1 S_{X_{ip}}^* v_1' - 1)$	$\tau_1 S_{X_{ip}} \tau_1' - \mu (\tau_1 \tau_1' - 1)$
Soluzione per la 1° funzione discriminante	$S_{X_{ip}}^* S_{X_{ip}} v_1 = \lambda_1 v_1$	$S_{X_{ip}} \tau_1 = \mu_1 \tau_1$

Come si vede nella prima riga l'unico cambiamento rispetto all'approccio classico non parametrico sta nella differente grandezza utilizzata per la ponderazione $[(\tau_1 \tau_1')]$ in luogo di $(v_1 S_{X_{ip}}^* v_1')$.

(4) L'utilizzo di un gruppo di controllo e costituito da individui, cui non sia stato somministrato il SPPU permette di verificare se vi sono stati cambiamenti di aspetti di benessere negli utenti non imputabili all'erogazione del SPPU ma dovuto a variabili non considerate nell'analisi. In questo caso si ha $E[W_{\alpha}^*] = E[W_{\alpha}] - E[W_{\alpha}]$. Coerentemente anche nell'analisi delle varianze [7] e nell'analisi delle covarianze [8] occorre tener conto della presenza del gruppo di controllo.

(5) Si ringrazia il prof. Tesio e la Fondazione CARIPLO per la Ricerca Scientifica, che hanno permesso lo svolgimento dell'indagine empirica.

Riferimenti bibliografici

- AITKIN M., LONGFORD N., (1986), "Statistical modeling issues in school effectiveness studies (with discussion)", *Journal of Royal Statistical Society, A*, vol. 149, n° 1, pp. 1-43.
- ANDRICH, D., (1979), "A model for contingency tables having an ordered response classification", *Biometrics*, 1979, vol. 35, pp. 403-415
- APEL H., WOLD H., (1982), *Soft modeling with latent variables in two or more dimensions: PLS estimation and testing for predictive relevance*, in K.G. Joreskog e H. Wold (eds), *Systems under indirect observation, causality, structure, prediction*, North Holland, Amsterdam 2, pp. 209-247.

ERBA A., MARTINI M., (1989), "Le attività terziarie: definizione e problemi di classificazione ed economie locali", in *Il sistema terziario in Italia*, AA.VV., ISCOM, Roma.

EVERITT B.S., (1984), *An introduction to latent variable models*, Chapman and Hall, Londra.

GOLDSTEIN H., SPIEGELHALTER D.J., (1996) "League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance", *Journal of Royal Statistical Society, A*, vol. 159, pp. 385-443.

GORI E., (1992a), "Variabili latenti e 'self-selection' nella valutazione dei processi formativi", *Statistica*, LII, 5, pp. 463-480.

GORI E., (1992b), "La valutazione dell'efficienza e dell'efficacia nell'istruzione", *Atti della XXXVI Riunione Scientifica Pescara 21-24 Aprile 1992*, vol. 1, pp. 219-229.

GORI E., MEALLI F., RAMPICHINI C., (1993) "Indicatori di efficienza ed efficacia per la valutazione dell'attività di formazione professionale", *Statistica*, LIII, n° 3, pp. 501-533.

GORI E., (1996), *La valutazione dei servizi ed interventi pubblici*, Relazione presentata alle Giornate di Statistica Economica, 26-27 Settembre, Siena.

HAAAGEN K., VITTADINI G., (1991), "Regression Component Decomposition in structural analysis", *Communications in statistics*, vol. 20, n° 4, New York, pp. 1153-1161.

HAAAGEN K., VITTADINI G., (1997), *Regression Component Decomposition Restricted: un'alternativa al Lisrel model*, Milano, pro manuscripto.

JORESKOG K.G., (1981), "Analysis of covariance structures", *Scandinavian Journal of Statistics*, vol. 8, pp. 65-92.

MANARA C.E., NICOLA E., (1970), *Elementi di economia matematica*, Viscontea, Milano.

MARTINI A., (1996), *Valutazione dell'efficacia di interventi pubblici contro la povertà: questioni di metodo ed esperienze di altri paesi*, Relazione presentata alle Giornate di Statistica Economica, 26-27 Settembre, Siena.

RASCH G., (1960), *Probabilistic Models for Some Intelligence and Attainment Tests*, Danmarks Paedagogiske Istitut, Copenhagen.

Summary

A statistical methodology for Performance Evaluation of Public Utility Service to People

We define "public utility service to people" (PUSP), those particular services to people (for example, Education, Health, etc.) for which the output is different from the outcome on the user, defined as betterment in a particular aspect of his well-being. It is possible to quantify the dimension of such well-being before and after the supply of the PUSP through latent variables obtained through proper method of linear transformation from particular quantitative indicators. Said indicators may also be obtained from ordinal qualitative indicators transformed through the Rasch Analysis in quantitative scores. Through adequate indices and covariance analysis of these latent variables we verify the efficacy of PUSP.

Keywords.

Outcome, Rasch Analysis, Regression Component Decomposition Restricted, Covariance Analysis.

SADOCCHII S., (1982), "Sull'interpretazione di alcuni strumenti di analisi statistica multivariata e sulle loro connessioni", *Serie Ricerche Teoriche*, vol. 6.

TESIO L., BATTAGLIA M.A., PERRUCA L., FRANCHIGNONI F.P., (1996), *Il controllo dei qualità delle misure FIM (Functional Independence Measure): un esempio di applicazione dell'analisi di Rasch*, pro manuscripto, in corso di pubblicazione.

TESIO L., GRANGER C.V., FIEDLER R.C., (1996), *A unidimensional pain/disability measure for low back pain syndromes*, International Association for the Study of Pain, pp. 1-10.

VITALI O., (1993), *Statistica per le Scienze applicate*, Cacucci, Bari.

VITTADINI G., (1988), "On the validity of the indeterminate latent variables in the LISREL model", *Communications in Statistics*, New York vol. 17, n° 3, pp. 861-876.

VITTADINI G., (1992), "Un confronto fra un modello di analisi causale con variabili latenti e metodi alternativi", *Statistica*, LII, n° 3, pp. 379-395.

VITTADINI G., (1996), *Modelli statistici per l'analisi della qualità nei servizi alla persona di pubblica utilità*, Relazione presentata alla Giornata di Statistica Economica, 26-27 Settembre, Siena.

WOLD H., (1980), "Model construction and evaluation when theoretical knowledge is scarce: Theory and application of Partial Least Squares" in J. Kmenta and J. Ramsey, eds. *Model Evaluation in Econometrics*, Academic Press, New York, pp. 47-74.

WRIGHT B.D., MASTERS G.N., (1982), *Rating scale analysis*, MESA Press, Chicago.