



Department of Statistical Sciences
University of Padua
Italy

UNIVERSITÀ
DEGLI STUDI
DI PADOVA
DIPARTIMENTO
DI SCIENZE
STATISTICHE

Analisi multivariata per osservazioni appaiate con dati mancanti: un caso studio

Chiara Brombin, Fortunato Pesarin

Department of Statistical Sciences
University of Padua
Italy

Giovanni Fava

Department of Psychology
University of Bologna
Italy

Abstract: All parametric approaches require that analysis should be done on complete data sets and so, in presence of missing data, parametric solutions are based either on the so-called deletion principle or imputation methods. But when we delete incomplete vectors we also remove all information they contain, which may be valuable and useful for analysis. And when we replace missing data by suitable functions of actually observed data, that is imputing method, we may introduce biased information which may negatively influence the analysis. On the other hand, non-parametric solutions in a permutation framework consider data as they are, and units with missing data participate in the permutation mechanism as well as all other units, without deletion or imputing. In this paper we provide a comparison between a parametric solution, represented by ITT principle, and a non parametric one, in a testing problem with multivariate paired observations.

Keywords: missing data, permutation tests, ITT analysis.

Indice

1	Premessa	1
2	Introduzione allo studio	2
2.1	Il metodo	3
2.2	I limiti dello studio	4
3	Il problema dei missing data	4
4	Approccio di permutazione	5
5	Il sistema d'ipotesi nei modelli non-MAR	7
5.1	Combinazione non parametrica	7
5.2	Il sistema d'ipotesi nei modelli MCAR	8
6	Osservazioni appaiate e missing data	8
7	Intention to treat analysis	9
8	Risultati dello studio: soluzione non parametrica	10
9	Risultati dello studio: soluzione parametrica	13
10	Conclusioni	13
	Riferimenti bibliografici	17

Analisi multivariata per osservazioni appaiate con dati mancanti: un caso studio

Chiara Brombin, Fortunato Pesarin

Department of Statistical Sciences
University of Padua
Italy

Giovanni Fava

Department of Psychology
University of Bologna
Italy

Abstract: All parametric approaches require that analysis should be done on complete data sets and so, in presence of missing data, parametric solutions are based either on the so-called deletion principle or imputation methods. But when we delete incomplete vectors we also remove all information they contain, which may be valuable and useful for analysis. And when we replace missing data by suitable functions of actually observed data, that is imputing method, we may introduce biased information which may negatively influence the analysis. On the other hand, non-parametric solutions in a permutation framework consider data as they are, and units with missing data participate in the permutation mechanism as well as all other units, without deletion or imputing. In this paper we provide a comparison between a parametric solution, represented by ITT principle, and a non parametric one, in a testing problem with multivariate paired observations.

Keywords: missing data, permutation tests, ITT analysis.

1 Premessa

I metodi parametrici utilizzati per la soluzione di problemi di carattere univariato e multivariato hanno, come limitazione, la necessità di dover ricorrere all'introduzione di ipotesi restrittive, spesso ingiustificate se non impossibili da giustificare, talvolta irrealistiche, non sempre chiare, difficilmente interpretabili o formulate *ad hoc* per poter fare inferenza. A questo si deve aggiungere che le assunzioni che rendono valida l'applicazione di tali metodi (normalità, omoschedasticità, indipendenza e identica distribuzione della componente stocastica erratica) sono di norma raramente soddisfatte e, quand'anche soddisfatte, i risultati sono spesso ottenuti tramite approssimazione.

Ci sono casi in cui, anche assumendo la normalità della distribuzione degli errori, le soluzioni parametriche non forniscono risultati esatti: ci si riferisce per esempio ai test per campioni con osservazioni appaiate quando i parametri di scala dipendono

dalle unità o ci sono dati mancanti, allo studio dell'analisi della varianza a blocchi (two-way ANOVA), a problemi derivanti dal fatto che il numero delle variabili rilevate è superiore alla dimensione del campione, oppure nei test a due campioni quando si vuole fare inferenza congiuntamente su media e varianze e le risposte sono positive.

Sempre più spesso, per problemi multivariati più complessi studiati in ambito biomedico, ingegneristico, psicologico, farmacologico, negli esperimenti clinici, nel controllo della qualità, quando non è noto il modello distributivo o ancora quando l'inferenza riguarda variabili di tipo qualitativo, si passa da un approccio parametrico ad uno non parametrico, ovviando così, senza perdita sostanziale di efficienza, alle limitazioni sopra accennate. E tra i metodi non parametrici, un ruolo di grande rilievo per le proprietà generali di cui godono, vi sono i metodi di permutazione.

Non bisogna però lasciarsi trarre in inganno, ovvero non concludere erroneamente che i metodi non parametrici di permutazione siano una *panacea*, per tutti i problemi inferenziali di interesse: se, sotto H_0 , non ci si condiziona ad un insieme di statistiche sufficienti e non si assume l'ipotesi di scambiabilità dei dati le soluzioni ottenute potranno essere non esatte [1].

Nei problemi multivariati, la soluzione di permutazione si ottiene applicando un preciso metodo di combinazione non parametrica di test dipendenti [2]. Tale metodo si basa sulla decomposizione delle ipotesi da testare in k , con $k > 1$, sottoipotesi per ciascuna delle quali si calcola un appropriato test di permutazione parziale, utilizzando poi un processo di simulazione consistente in B ricampionamenti casuali senza reinserimento, tramite condizionamento all'insieme dei dati osservati, si ottiene una stima della distribuzione di permutazione della statistica test. Infine i risultati parziali della simulazione condizionata vengono combinati in una statistica test del secondo ordine tramite una funzione reale (misurabile) continua, monotona decrescente e non degenera [3].

2 Introduzione allo studio

Recenti studi in psicologia e neuropsichiatria infantile hanno messo in luce un aumento del numero di bambini che presentano disturbi comportamentali, disagi di tipo psicologico come ansia, depressione, stress, incapacità di concentrazione, disturbi alimentari (anoressia, bulimia, obesità) che si traducono in una difficoltà nella gestione dei rapporti interpersonali e si riflettono nei processi evolutivi, d'apprendimento e di sviluppo.

Nell'intento di prevenire abusi e sofferenze nei bambini e negli adolescenti, associazioni ed enti no-profit per la tutela e protezione dei diritti d'infanzia si adoperano sempre più per promuovere interventi scolastici mirati a ristabilire l'equilibrio affettivo nel bambino. Accanto ad interventi della tradizione psicologica più consolidata, come la terapia cognitivo-comportamentale basata sulla correzione delle distorsioni cognitive e sull'insegnamento di una serie di abilità coping e problem solving per migliorare l'adattamento psicologico in situazioni di stress, si sta diffondendo un filone di ricerca d'impronta americana noto come *Positive Psychology*.

Il concetto cardine della *Positive Psychology* è la necessità di distogliere lo sguardo

dai deficits psicologici e comportamentali che caratterizzano lo sviluppo del bambino e di concentrarsi piuttosto sulle competenze e i punti forza del bambino stesso, ma anche della sua famiglia, per promuovere una crescita e un rafforzamento da un punto di vista psicologico.

Su questa scia è stato sviluppato un protocollo d'intervento per l'applicazione della *Well-Being Therapy (WBT)*, un aspetto della *Positive Psychology*, in ambito scolastico, articolato in una serie d'incontri di gruppo, supportato da metodologie laboratoriali ed esperienziali, come giochi, simulazioni e discussioni di gruppo, e avente come obiettivo quello di educare i ragazzi a migliorare l'espressione di emozioni e sentimenti e a favorire i processi relazionali e comunicativi.

2.1 Il metodo

Nel corso dell'anno scolastico 2003/2004 è stato effettuato uno studio su una popolazione scolastica costituita da studenti delle scuole medie inferiori. Sono state esaminate 6 classi, per un totale di 120 partecipanti, di 4 scuole medie di Padova. Oggetto dell'intervento in aula era un breve corso di educazione all'affettività da parte di due psicologhe cliniche (4 incontri per classe di 2 ore ciascuno). I primi due incontri, uguali in tutte le 6 classi, si sono incentrati in maniera precipua sul riconoscimento delle emozioni e sul miglioramento del clima emotivo nelle classi.

Nei successivi due incontri l'insieme delle 6 classi è stato suddiviso, in maniera casuale, in due gruppi, gruppo *A* e gruppo *B*. Ogni gruppo comprendeva quindi 3 classi di 60 soggetti ciascuna. Nel gruppo *A* si è cercato di migliorare il benessere e favorire le esperienze positive applicando il modello di benessere psicologico (*Well Being Therapy, WBT*), nel gruppo *B*, seguendo esclusivamente le teorie cognitive (*Cognitive Behaviour Therapy, CBT*) l'attenzione è stata rivolta alla risoluzione dei problemi e dei conflitti nei ragazzi. Prima e dopo gli incontri, i ragazzi hanno compilato due scale di autovalutazione del proprio stato psicologico: il *Kellner's Symptom Questionnaire (SQ)* [4] e la forma breve per adolescenti del *Psychological Well-Being (PWB)* [5]. Il *Kellner's Symptom Questionnaire (SQ)* è una scala di autovalutazione con 92 item che comprende 8 sottoscale, 4 di disagio psicologico (ansia, depressione, sintomi somatici, ostilità) con un punteggio che va da 0 a 17, per cui maggiore è il punteggio maggiore è il disagio, e 4 di benessere psicologico (rilassamento, contentezza, benessere fisico e buona disponibilità verso gli altri) con un punteggio che va da 0 a 6, per cui maggiore è il punteggio minore è il benessere.

La forma breve del PWB è un questionario con 18 item per la valutazione del benessere psicologico validato per l'Italia dall'équipe del prof. Fava, che ha messo a punto anche la versione breve per adolescenti utilizzata in questo protocollo. Il questionario fornisce una valutazione di 6 aree del benessere (autonomia, padronanza dell'ambiente, crescita personale, relazioni positive, scopo nella vita e accettazione di sé). I soggetti dovevano fornire risposte SI/NO agli item indicati.

Il punteggio assegnato alle risposte va da 0 a 3, per cui maggiore è il punteggio, maggiore è il benessere. L'obiettivo dello studio è di scoprire se ci sono differenze significative tra i due gruppi e l'ipotesi che si vuole testare è se l'intervento focalizzato sul benessere, caratterizzato da attività più divertenti e piacevoli, è stato più efficace.

2.2 I limiti dello studio

Le limitazioni sono legate alla natura preliminare dello studio. Innanzitutto, il numero esiguo di incontri (4 incontri di 2 ore ciascuno) e il campione di soggetti prevalentemente sani non hanno di certo agevolato la rilevazione di differenze significative tra tempo 0 e tempo 1 (pre e post trattamento). I partecipanti, poi, non sono stati sottoposti a delle terapie standard (WBT, CBT), ma ad un protocollo d'intervento assolutamente preliminare, soprattutto nel caso del trattamento WBT, adattato in un contesto scolastico, avente in comune con la terapia classica completa solo alcuni degli aspetti salienti.

Da non sottovalutare l'autoselezione delle scuole che non sono state randomizzate per ricevere il trattamento, ma si sono spontaneamente proposte per essere incluse nello studio (hanno aderito spontaneamente allo studio). Il piano sperimentale più che progettato è stato in parte subito: i risultati finali infatti potrebbero essere condizionati da alcuni fattori quali la mancanza di omogeneità nella composizione delle classi esaminate, sia per quanto riguarda il sesso sia per quanto riguarda l'età degli studenti, che potrebbe mascherare un effetto *leader*, la tendenza negli adolescenti alla *desiderabilità sociale*, il fattore *dislocazione delle scuole*, se queste si trovassero in zone residenziali o più in periferia, e legato a questo il fattore *condizione sociale* dei partecipanti. Altre caratteristiche dello studio che hanno costituito una limitazione alla sua piena riuscita sono la mancanza di un follow-up in grado di rilevare gli effetti del protocollo WBT a lungo termine e il fatto che entrambi gli interventi condividono tecniche e aspetti in comune, la WBT è infatti una tecnica specifica all'interno dell'ampio insieme delle strategie cognitivo-comportamentali [6].

Tra tutti questi aspetti, a rendere un po' più complicata l'analisi dei dati è stata anche la presenza di *missing values* e in questo lavoro il problema viene affrontato seguendo due diversi approcci: tramite l'applicazione del principio dell'*Intention To Treat* (ITT) *analysis* e utilizzando i metodi non parametrici di permutazione.

3 Il problema dei missing data

Il problema dei *missing data* è assai frequente ([7];[8];[9];[10];[11]) specie nelle indagini demografiche quando una popolazione sotto studio in un certo anno, al momento del *follow-up* effettuato dopo qualche anno si presenta in numero ridotto per ragioni tra le più svariate. Spesso le analisi di tali dati procedono con l'assunzione, esplicita o implicita, che il processo generatore di *missing-data* può essere ignorato. E' naturale chiedersi quando questa procedura è appropriata. Si tratta di un problema assai discusso in ambito statistico cui sono state proposte parecchie soluzioni. Un lavoro molto importante a riguardo è quello di Rubin (1976) [12]. L'autore fornisce questa definizione:

“The missing data are missing at random if for each possible value of the parameter ϕ , the conditional probability of the observed pattern of missing data, given the missing data and the value of the observed data, is the same for all possible values of the missing data. The observed data are observed at random if for each possible value of the missing data and the parameter ϕ , the conditional probability of the observed pattern of missing data, given the missing data and the observed data, is the

same for all possible values of the observed data. The parameter ϕ is distinct from θ if there are no a priori ties, via parameter space restrictions or prior distributions, between ϕ and θ .”

Se i *missing data* sono MAR e gli *observed data* sono OAR, i *missing data* sono *missing completely at random* (MCAR). In questo caso, la mancanza non dipende dai valori osservati o non osservati, e i valori osservati possono essere considerati come un sotto campione casuale dell'intero insieme dei dati. In questo caso è possibile ignorare il processo che genera i dati mancanti e far inferenza su ϑ . Se i dati sono MAR e ϕ e ϑ sono distinti, allora l'inferenza di verosimiglianza su ϑ può essere basata sulla funzione di verosimiglianza senza includere un termine per il processo dei *missing data*, in questo caso il processo che genera i *missing data* è definito *ignorable* e l'analisi può essere condotta via condizionamento ai dati effettivamente osservati. Se i *missing data* sono *missing not at random* (non-MAR), per fare una corretta inferenza, in ambito parametrico, il processo generatore deve essere specificato. Dal momento che in molti casi i *missing data* provengono da indagini campionarie e le non risposte sono dovute ai più svariati motivi, in generale è molto complesso definire un modello che, tra tutti, meglio descriva i dati, siano essi osservati o mancanti. E' per questo che un'analisi condotta su dati non-MAR è di gran lunga più difficile rispetto a quella condotta su dati MCAR.

4 Approccio di permutazione

In situazioni di MCAR, le soluzioni parametriche sfruttano il *deletion principle* o un metodo di imputazione: il primo analizza i dati restanti dopo la rimozione di tutti i vettori incompleti, il secondo consiste nella sostituzione dei dati mancanti con appropriate funzioni, in genere la media, dei dati effettivamente osservati. Tutto questo perché, un approccio parametrico richiede che vi sia a disposizione un insieme completo di dati.

Un approccio di permutazione, invece, considera i dati come sono senza invocare il principio di eliminazione o imputazione: quando infatti si eliminano i vettori incompleti, si eliminano anche quelle informazioni che sono utili per l'analisi e, viceversa, quando si fanno delle imputazioni si possono introdurre informazioni distorte che possono influenzare negativamente l'analisi. Senza perdita di generalità, supponiamo di voler testare l'uguaglianza in distribuzione di C campioni, in cui alcuni dati sono mancanti .

Dati C gruppi di variabili risposta scambiabili q -dimensionali

$$\mathbf{X}_j = \{\mathbf{X}_{ji} = (X_{hji}, h = 1, \dots, q), i = 1, \dots, n_j, j = 1, \dots, C\}$$

con funzione di distribuzione P_j , $\mathbf{X}_{ji} \in \mathcal{R}^q$, dove $n = \sum_j n_j$ è l'ampiezza del campione, e il sistema d'ipotesi è dato da

$$\begin{aligned} H_0 &= \{P_1 = \dots = P_C = P\} = \{\mathbf{X}_1 \stackrel{d}{=} \dots \stackrel{d}{=} \mathbf{X}_C\} \\ H_1 &= \{H_0 \text{ non è vera}\} \end{aligned} \quad (1)$$

Assumiamo che i dati, siano essi effettivamente osservati o mancanti, siano scambiabili rispetto ai C gruppi sotto l'ipotesi nulla. Consideriamo appropriate trasfor-

mazioni dei dati, che indichiamo con φ_h , $h = 1, \dots, q$, che siano specifiche per ogni variabile. In modo particolare utilizziamo la trasformazione media campionaria come indicatore appropriato degli effetti del trattamento. L'insieme dei dati può essere così rappresentato:

$$\mathbf{Y} = \{Y_{hji} = \varphi_h(X_{hji}), i = 1, \dots, n_j, j = 1, \dots, C, h = 1, \dots, q\} \quad (2)$$

e i test di permutazione parziali possono essere costruiti utilizzando appropriate funzioni del totale del campione, ossia $S_{hj}^* = \sum_i Y_{hji}^*$, con $j = 1, \dots, C$ e $h = 1, \dots, q$. Consideriamo anche un vettore di variabili indicatrici della presenza o meno dell'osservazione, ovvero

$$\mathbf{O} = \{O_{hji}, i = 1, \dots, n_j, j = 1, \dots, C, h = 1, \dots, q\} \quad (3)$$

in cui $O_{hji} = 1$ se X_{hji} stato osservato e $O_{hji} = 0$ altrimenti. Il vettore \mathbf{O} rappresenta la configurazione osservata del data set. L'insieme dei dati osservati può dunque essere riscritto come (\mathbf{Y}, \mathbf{O}) . Definiamo dunque l'effettiva dimensione del campione con $\nu_{hj} = \sum_i O_{hji}$, dove $j = 1, \dots, C$ e $h = 1, \dots, q$. Se si suppone dunque che, sotto H_0 , i dati siano congiuntamente scambiabili rispetto ai gruppi su entrambe le variabili \mathbf{Y} e \mathbf{O} , il sistema d'ipotesi (1) può essere così riformulato:

$$\begin{aligned} H_0 &= \left\{ (\mathbf{Y}_1, \mathbf{O}_1) \stackrel{d}{=} \dots \stackrel{d}{=} (\mathbf{Y}_C, \mathbf{O}_C) \right\} \\ H_1 &= \{H_0 \text{ non è vera} \} \end{aligned} \quad (4)$$

In questo contesto, sempre sotto H_0 , P diventa la distribuzione congiunta multivariata $(\mathbf{Y}_j, \mathbf{O}_j)$ con $j = 1, \dots, C$ e si può anche scrivere che $P = P_{\mathbf{O}} \cdot P_{\mathbf{Y}|\mathbf{O}}$ enfatizzando così la variabile indicatrice dell'inclusione nel campione e la variabile trasformazione dei dati osservati. La coppia (\mathbf{Y}, \mathbf{O}) rappresentante l'intero insieme dei dati osservati, sotto l'ipotesi nulla, è congiuntamente un insieme di statistiche sufficienti per la distribuzione sottostante i dati osservati e il processo generatori dei dati mancanti. Se ci sono q differenti aspetti da porre sotto ipotesi, ci saranno anche q test parziali con la loro relativa combinazione non parametrica. Fatta questa considerazione, il sistema d'ipotesi (1) può essere riscritto come

$$\begin{aligned} H_0 &: \left\{ \bigcap_{h=1}^q \left[(Y_{h1}, O_{h1}) \stackrel{d}{=} \dots \stackrel{d}{=} (Y_{hC}, O_{hC}) \right] \right\} = \left\{ \bigcap_{h=1}^q H_{0h} \right\} \\ H_1 &= \{ \cup_h H_{1h} \} \end{aligned} \quad (5)$$

in cui H_0 è vera se tutte le H_{0h} sono congiuntamente vere, mentre H_1 è vera se almeno una H_{1h} è vera. Per testare H_0 contro H_1 , consideriamo un vettore q -dimensionale di statistiche test $\mathbf{T} = \{T_1, \dots, T_C\}$, la h -esima componente del quale è la statistica test per verificare l' h -esima sottoipotesi H_{0h} . Assumiamo che i test parziali siano non degeneri, marginalmente non distorti, consistenti, significativi per valori grandi. Di qui il test combinato è funzione dei q test parziali dipendenti. Proprio per la presenza di relazioni di dipendenza sottostanti, la combinazione deve essere condotta via non parametrica.

5 Il sistema d'ipotesi nei modelli non-MAR

Se i *missing data* sono non-MAR (*Missing Not at Random*), allora l'ipotesi nulla H_0 nel verificare l'uguaglianza in distribuzione dei C gruppi deve prendere in considerazione il vettore dei dati osservati e raccolti \mathbf{Y} congiuntamente al vettore associato d'inclusione \mathbf{O} , in quanto sotto l'ipotesi alternativa il trattamento può influenzare anche il processo che causa i *missing data*. Di qui l'ipotesi nulla richiede l'uguaglianza in distribuzione del processo generatore dei dati mancanti nei C gruppi, espresso da \mathbf{O} , e delle variabili risposta \mathbf{Y} condizionatamente ad \mathbf{O} , cioè:

$$H_0 : \left\{ \left[\mathbf{O}_1 \stackrel{d}{=} \dots \stackrel{d}{=} \mathbf{O}_C \right] \cap \left[\left(\mathbf{Y}_1 \stackrel{d}{=} \dots \stackrel{d}{=} \mathbf{Y}_C \right) \mid \mathbf{O} \right] \right\} \quad (6)$$

L'assunzione di scambiabilità, sotto l'ipotesi nulla, degli n vettori (\mathbf{Y}, \mathbf{O}) , con rispetto ai C gruppi, implica che gli effetti del trattamento sono nulli per tutte le variabili, osservate e non osservate. In altre parole, si assume che non vi sia differenza in distribuzione per la variabile d'inclusione multivariata \mathbf{O}_j , $j = 1, \dots, C$, e condizionatamente a \mathbf{O} , per le variabili osservate presenti in \mathbf{Y} . Di conseguenza non è necessario specificare il processo generatore dei *missing data* e la distribuzione delle osservazioni, purchè siano a disposizione test parziali marginalmente non distorti. In particolare dato che la soluzione è non parametrica, non è necessario specificare le relazioni di dipendenza sottostanti la distribuzione congiunta di (\mathbf{Y}, \mathbf{O}) . L'ipotesi nulla globale può essere scomposta in $2q$ sottoipotesi, ovvero:

$$\begin{aligned} H_0 & : \left\{ \left[\cap_h \left(O_{h1} \stackrel{d}{=} \dots \stackrel{d}{=} O_{hC} \right) \right] \cap \left[\cap_h \left(Y_{h1} \stackrel{d}{=} \dots \stackrel{d}{=} Y_{hC} \right) \mid \mathbf{O} \right] \right\} \\ & = \left\{ H_0^{\mathbf{O}} \cap H_0^{\mathbf{Y} \mid \mathbf{O}} \right\} = \left\{ \left(\cap_h H_{0h}^{\mathbf{O}} \right) \cap \left(\cap_h H_{0h}^{\mathbf{Y} \mid \mathbf{O}} \right) \right\} \end{aligned} \quad (7)$$

contro

$$H_1 = \left\{ \left(\cup_h H_{1h}^{\mathbf{O}} \right) \cup \left(\cup_h H_{1h}^{\mathbf{Y} \mid \mathbf{O}} \right) \right\}$$

dove $H_{0h}^{\mathbf{O}}$ indica l'uguaglianza in distribuzione, rispetto ai C livelli, dell' h -esima componente marginale del processo d'inclusione nel campione, ossia \mathbf{O} , e $H_{0h}^{\mathbf{Y} \mid \mathbf{O}}$ indica l'uguaglianza in distribuzione dell' h -esima componente di \mathbf{Y} , condizionatamente ad \mathbf{O} . Una statistica test di permutazione per ciascuna delle q sottoipotesi $H_{0h}^{\mathbf{O}}$ può essere la statistica X^2 . In alternativa si possono utilizzare altre statistiche appropriate per variabili categoriali, come quella di Anderson-Darling, oppure il test esatto di Fisher multivariato (*Fisher's exact multivariate probability test*), se le variabili categoriali sono ordinate.

5.1 Combinazione non parametrica

La combinazione non parametrica dei $2q$ test parziali può essere fatta in almeno tre modi diversi:

1. considerare un'unica funzione di combinazione su tutti i $2q$ test parziali, del tipo $T_a'' = \psi(\lambda_1^{\mathbf{O}}, \dots, \lambda_q^{\mathbf{O}}; \lambda_1^{\mathbf{Y} \mid \mathbf{O}}, \dots, \lambda_q^{\mathbf{Y} \mid \mathbf{O}})$;

2. considerare q combinazioni, per ciascuna variabile componente, del secondo ordine $T''_{bh} = \psi_h(\lambda_h^{\mathbf{O}}; \lambda_h^{\mathbf{Y}|\mathbf{O}})$, $h = 1, \dots, q$ e poi una combinazione del terzo ordine $T'''_b = \psi(\lambda''_{b1}, \dots, \lambda''_{bq})$;
3. considerare due combinazioni del secondo ordine, $T''_{c\mathbf{O}} = \psi_{\mathbf{O}}(\lambda_1^{\mathbf{O}}, \dots, \lambda_q^{\mathbf{O}})$ e $T''_{c\mathbf{Y}|\mathbf{O}} = \psi_{\mathbf{Y}|\mathbf{O}}(\lambda_1^{\mathbf{Y}|\mathbf{O}}, \dots, \lambda_q^{\mathbf{Y}|\mathbf{O}})$, rispettivamente sul vettore delle variabili indicatrici dell'inclusione, e sul vettore di dati effettivamente osservati (\mathbf{Y}, \mathbf{O}) , seguite da una combinazione del terzo ordine $T'''_c = \psi(T''_{c\mathbf{O}}; T''_{c\mathbf{Y}|\mathbf{O}})$

E' da notare che in tutte le fasi e in ciascuno dei tre metodi viene utilizzata la stessa funzione ψ e T''_a, T''_b, T''_c sono asintoticamente equivalenti.

5.2 Il sistema d'ipotesi nei modelli MCAR

Se i *missing data* sono MCAR (*Missing Completely at Random*) si procede condizionandosi al vettore indicante l'inclusione osservato, cioè \mathbf{O} , e si ignora $H_0^{\mathbf{O}}$ perché si assume che \mathbf{O} non fornisca alcuna informazione sugli effetti del trattamento. Essendo, per assunto, vere le sottoipotesi su \mathbf{O} , l'ipotesi $H_0^{\mathbf{O}} : \{\mathbf{O}_1 \stackrel{d}{=} \dots \stackrel{d}{=} \mathbf{O}_C\}$ può essere ignorata. L'ipotesi nulla può essere riscritta in modo più semplice nella forma

$$H_0 = H_0^{\mathbf{Y}|\mathbf{O}} : \left\{ \bigcap_h \left[(\mathbf{Y}_{h1} \stackrel{d}{=} \dots \stackrel{d}{=} \mathbf{Y}_{hC}) | \mathbf{O} \right] \right\} = \left\{ \bigcap_h H_{0h}^{\mathbf{Y}|\mathbf{O}} \right\} \quad (8)$$

contro

$$H_1^{\mathbf{Y}|\mathbf{O}} = \left\{ \bigcup_h H_{1h}^{\mathbf{Y}|\mathbf{O}} \right\}$$

in cui la notazione è la stessa di quella utilizzata nel paragrafo precedente. Questo problema può essere risolto da una combinazione non parametrica del tipo $T''' = T''_{c\mathbf{Y}|\mathbf{O}} = \psi_{\mathbf{Y}}(\lambda_1^{\mathbf{Y}|\mathbf{O}}, \dots, \lambda_q^{\mathbf{Y}|\mathbf{O}})$.

6 Osservazioni appaiate e missing data

Una variabile reale \mathbf{X} , q -dimensionale e non degenera, viene osservata in k occasioni temporali diverse sulle stesse n unità in due situazioni sperimentali, corrispondenti a due livelli di un simbolico trattamento. I dati vengono così rappresentati:

$$\begin{aligned} \mathbf{X} &= \{X_{hjit}, t = 1, \dots, k, i = 1, \dots, n, j = 1, 2, h = 1, \dots, q\} \\ &= \{\mathbf{X}_{jit}, t = 1, \dots, k, i = 1, \dots, n, j = 1, 2\} \\ &= \{\mathbf{X}_{ji}, i = 1, \dots, n, j = 1, 2\} \end{aligned} \quad (9)$$

dove \mathbf{X} rappresenta la risposta q -variata (X_1, \dots, X_q) , l'andamento temporale q -variato $(\mathbf{X}_1, \dots, \mathbf{X}_k)$ e l'intero insieme dei dati. Supponiamo che ci siano alcuni *missing data* e quindi che ci sia anche un vettore indicatore d'inclusione

$$\mathbf{O} = \{O_{hjit}, t = 1, \dots, k, i = 1, \dots, n, j = 1, 2, h = 1, \dots, q\}$$

in cui $O_{hjit} = 1$ se X_{hjit} è osservato e $O_{hjit} = 0$ se X_{hjit} è un *missing value*. Supponiamo di voler verificare se il trattamento ha effetto, senza curarci del tempo, delle

dipendenze sottostanti e dell'ignota distribuzione. Se i dati sono *missing* secondo un modello MCAR, le ipotesi da testare sono formalmente espresse con

$$H_0 : \left\{ \bigcap_{t=1}^k (\mathbf{X}_{1t} \stackrel{d}{=} \mathbf{X}_{2t}) \mid \mathbf{O} \right\} = \left\{ \bigcap_{t=1}^k \bigcap_{q=1}^h H_{0ht} \right\}$$

verso

$$H_1 : \left\{ \bigcup_{t=1}^k (\mathbf{X}_{1t} <\neq> \mathbf{X}_{2t}) \mid \mathbf{O} \right\} = \left\{ \bigcup_{t=1}^k \bigcup_{q=1}^h H_{1ht} \right\}$$

La statistica test di permutazione assume la forma

$$T_{ht}^* = \varphi_{ht} \left[\frac{\sum_i Y_{hit} \cdot O_{hit} \cdot S_i^*}{(\sum_i Y_{hit}^2 \cdot O_{hit}^2)^{1/2}} \right], \quad t = 1, \dots, k, \quad h = 1, \dots, q \quad (10)$$

dove $Y_{hit} = X_{h1it} - X_{h2it}$ sono le differenze osservate unità per unità, tempo per tempo e variabile per variabile, $O_{hit} = O_{h1it} \cdot O_{h2it}$ vale 1 se e soltanto se X_{h1it} e X_{h2it} sono entrambe osservate, S_i^* sono i segni di permutazione e φ_{ht} sono le funzioni valore assoluto o segno $+/-$ a seconda che l' h -esima alternativa H_{1ht} di interesse sia a due o ad una coda.

Tutti i test parziali T_{ht}^* , $t = 1, \dots, k$, $h = 1, \dots, q$ sono permutazionalmente esatti, purchè le dimensioni effettive del campione $\nu_{ht} = \sum_i O_{hit}$ siano positive, in quanto i ν_{ht} sono quantità permutazionalmente invarianti e la distribuzione nulla dipende solo dalla scambiabilità degli errori. I test parziali T_{ht}^* hanno media di permutazione nulla e varianza di permutazione unitaria e sono marginalmente non distorti. Di qui la combinazione non parametrica è una soluzione adeguata per i test multivariati. Tutti i test sono espressi in forma standardizzata così è possibile applicare una funzione di combinazione diretta e si ha $T_h^{**} = \sum_t T_{ht}^*$ per i test entro ciascuna variabile senza curarsi del tempo, e poi un test globale $T^{***} = \sum_h T_h^{**}$. L'estensione ai modelli non-MAR è semplice, purchè i test parziali per le componenti del vettore indicatore d'inclusione \mathbf{O} siano marginalmente non distorti. Se i test parziali per $H_0^{\mathbf{O}} : \{\cap_h H_{0h}^{\mathbf{O}}\}$ contro $H_1^{\mathbf{O}} : \{\cup_h H_{1h}^{\mathbf{O}}\}$ sono esatti, allora la combinazione non parametrica per l'ipotesi

$$H_0^{\mathbf{O}} : \left\{ [\cap_h H_{0h}^{\mathbf{O}}] \cap [\cap_h H_{0h}^{\mathbf{Y}|\mathbf{O}}] \right\}$$

contro

$$H_1^{\mathbf{O}} : \left\{ [\cup_h H_{1h}^{\mathbf{O}}] \cup [\cup_h H_{1h}^{\mathbf{Y}|\mathbf{O}}] \right\}$$

è pure esatta.

7 Intention to treat analysis

Una soluzione al problema dei dati appaiati può essere ottenuta in un contesto parametrico se si assume che le variabili siano normalmente distribuite e abbiano varianza ignota. La statistica test più comunemente usata è la statistica test t di Student per osservazioni appaiate, distribuita sotto H_0 come una t di Student centrale con $(n - 1)$ *g.d.l* e sotto l'alternativa H_1 come una t di Student non centrale

con un parametro di non centralità positivo così che valori grandi diventano significativi. Sempre rimanendo nell'ambito parametrico, altra soluzione è rappresentata dall'analisi della covarianza (*ANCOVA*), che può essere utilizzata a condizione che la variabile rilevata prima del trattamento (assunta con il ruolo di covariata del problema) e quella rilevata dopo il trattamento abbiano distribuzione normale bivariata. Quando però siamo in presenza di missing data, come già si è detto, si deve ricorrere al *deletion principle* o ad un metodo di imputazione. In questo lavoro i dati mancanti sono stati trattati utilizzando il principio dell'ITT.

Secondo tale principio la popolazione di analisi è costituita da tutti i pazienti randomizzati, inclusi quelli che hanno interrotto lo studio prematuramente (e quindi hanno alcune valutazioni mancanti o, addirittura, sono usciti dallo studio senza aver mai ricevuto il trattamento) e quelli che presentano delle violazioni ai criteri di selezione (per cui non sarebbero dovuti entrare nello studio).[13]

Si tratta di un'analisi conservativa, infatti non altera la randomizzazione iniziale, e prudenziale, perché prende in considerazione la possibilità che non tutti i pazienti seguano completamente le procedure previste dal protocollo specifico del gruppo cui sono stati assegnati, ma può fornire informazioni spurie, sopra o sotto stimando l'effetto del trattamento con una conseguente perdita di potenza del trial clinico. Nel nostro caso, il campione "costruito" sulla base del principio dell'*Intention To Treat* include tutti i soggetti che hanno effettuato entrambe le misurazioni o che non hanno effettuato il retest, e che, invece, esclude quei soggetti per i quali non sono disponibili i punteggi relativi alla prima misurazione. Per quanto riguarda i soggetti che non hanno effettuato il retest, applicando il principio dell'ITT, nella seconda misurazione vengono ripetuti i punteggi ottenuti, in ciascuna variabile, nella prima rilevazione. Si può subito constatare che trattare via parametrica il problema dei *missing data* comporta una perdita di informazione.

8 Risultati dello studio: soluzione non parametrica

Oggetto dell'analisi è l'intero campione, compresi i vettori contenenti *missing values*. Il primo obiettivo di uno studio basato sul confronto prima-dopo è quello di verificare se vi è un effetto del trattamento all'interno del gruppo: si tratta quindi di fare un'analisi per campioni con osservazioni appaiate e missing values (campioni dipendenti). Nella tabella (1) sono riportati i *p-value* dei test parziali non parametrici, ottenuti con $B = 10000$ permutazioni. L'efficacia del trattamento WBT si traduce in un aumento significativo dell'accettazione di sé ($p = 0.034$) e in una riduzione significativa degli atteggiamenti di ostilità ($p = 0.020$), mentre, per quanto riguarda il trattamento CBT, sono da evidenziare un aumento significativo nei punteggi delle variabili "scopo vita" ($p = 0.004$) e "accettazione di sé" ($p = 0.038$) e una riduzione dei punteggi nelle variabili "ansia" ($p = 0.001$), "depressione" ($p = 0.024$) e "benessere fisico" ($p = 0.009$). Da ricordare che quest'ultima variabile, considerata dal *Kellner's Symptom Questionnaire* (SQ), ha una valenza negativa: pur appartenendo alla sottoscala di benessere psicologico, è tale che maggiore è il punteggio ottenuto minore è il benessere. Per quanto riguarda la combinazione non parametrica dei test parziali (vedi tabella 2), si può vedere che il

	WBT	CBT
	p-value	p-value
autonomia	0.156	0.089
padronanza ambiente	0.100	0.193
crescita personale	0.109	0.090
relazioni	0.135	0.220
scopo vita	0.230	0.004
accettazione	0.034	0.038
ansia	0.106	0.001
depressione	0.518	0.024
sintomi somatici	0.122	0.115
ostilità	0.020	0.132
relax	0.691	0.163
contentezza	0.684	0.239
benessere fisico	0.904	0.009
buona disponibilità	0.612	0.420

Tabella 1: WBT *vs.* CBT (campioni dipendenti)

trattamento WBT ha un effetto sulla combinazione di tutte le variabili della scala PWB ($p= 0.012$). Il trattamento CBT, invece, ha effetto sulla combinazione delle variabili “ansia-relax” ($p= 0.001$), “depressione-contentezza” ($p= 0.042$), “sintomi somatici-benessere fisico” ($p= 0.014$), sulle variabili, considerate in maniera congiunta, della scala SQ ($p= 0.006$) e della scala PWB ($p= 0.002$) e infine un effetto globale su tutte le variabili considerate congiuntamente ($p= 0.001$). Per le variabili “relax” e “contentezza” vale lo stesso discorso fatto precedentemente per la variabile “benessere fisico”: ad esse viene associato un significato negativo. Quanto alla combinazione non parametrica dei p -value è stata impiegata la funzione di *Fisher* e, come sopra, un numero di permutazioni pari a 10000. Il secondo obiettivo consiste

combinazioni	WBT	CBT
	p-value	p-value
ansia-relax	0.258	0.001
depressione-contentezza	0.612	0.042
sintomi somatici-benessere fisico	0.335	0.014
ostilità buona-disponibilità	0.088	0.215
variabili scala SQ	0.267	0.006
variabili scala PWB	0.012	0.002
globale	0.084	0.001

Tabella 2: Combinazione non parametrica dei p -value (campioni dipendenti)

nel verificare quali dei due trattamenti è migliore. Si tratta quindi di fare test per campioni indipendenti sempre sulle variabili differenza con *missing values*. Nella

tabella 3 accanto alle variabili considerate e rilevate nello studio sono riportati, per ciascun gruppo, il numero di *missing values*. E' emerso che il trattamento WBT è più efficace del CBT nella riduzione degli stati d'ansia ($p= 0.021$) e nell'aumento delle sensazioni di benessere fisico ($p= 0.007$). Per quanto riguarda la combinazione

	non parametrico	N.missing	
	p-value	WBT	CBT
autonomia	0.270	4	14
padronanza ambiente	0.498	4	14
crescita personale	0.427	4	14
relazioni	0.564	4	14
scopo vita	0.120	4	14
accettazione	0.390	4	14
ansia	0.021	3	13
depressione	0.083	3	13
sintomi somatici	0.511	3	13
ostilità	0.783	3	13
relax	0.138	3	13
contentezza	0.246	3	13
benessere fisico	0.007	3	13
buona disponibilità	0.342	3	13

Tabella 3: WBT vs. CBT (campioni indipendenti)

non parametrica dei *p-value*, il WBT ha un'azione maggiore sulla combinazione delle variabili "ansia-relax" ($p= 0.030$) e "sintomi somatici-benessere fisico" ($p= 0.038$), ma non è stata riscontrata una maggiore efficacia a livello globale (vedi tabella 4). Per ottenere i risultati riportati in tabella, sono state effettuate $B = 10000$ permutazioni e, per quanto riguarda la combinazione non parametrica dei *p-value*, è stata utilizzata la funzione di combinazione di Fisher.

combinazioni	non parametrico
	p-value
ansia-relax	0.030
depressione-contentezza	0.126
sintomi somatici-benessere fisico	0.038
ostilità buona-disponibilità	0.566
variabili scala SQ	0.063
variabili scala PWB	0.355
globale	0.087

Tabella 4: Combinazione non parametrica dei *p-value* (campioni indipendenti)

9 Risultati dello studio: soluzione parametrica

Causa l'applicazione del principio dell'*Intention To Treat*, escono dallo studio 9 soggetti, 3 appartenenti al gruppo WBT, che quindi è composto da 57 soggetti anzichè 60, e 6 appartenenti al gruppo trattato con CBT, costituito ora da 54 soggetti anzichè 60. 9 soggetti invece sono stati valutati secondo l'ITT, di questi 1 appartenente al gruppo WBT e 8 al gruppo CBT. Sul "nuovo" campione, sono state effettuate via parametrica le stesse analisi. L'analisi per campioni con dati appaiati (campioni dipendenti) ha messo in evidenza che il trattamento WBT aumenta in maniera significativa il livello di accettazione di sè ($p=0.022$) e riduce l'ostilità ($p=0.012$), mentre il trattamento CBT comporta un aumento significativo dei punteggi ottenuti nelle variabili "scopo vita" ($p=0.002$) e "accettazione di sè" ($p=0.025$) e una riduzione di questi nelle variabili "ansia" ($p=0.001$), "depressione" ($p=0.023$) e "benessere fisico" ($p=0.014$). Per quanto riguarda la combinazione parametrica dei test parziali risulta che il trattamento WBT ha un effetto sulla combinazione delle variabili "ostilità-buona disponibilità" ($p=0.027$) e sulle variabili, considerate in maniera congiunta, della scala PWB ($p=0.004$). Il trattamento CBT, invece, ha un effetto significativo sulla combinazione delle variabili "ansia-relax" ($p\approx 0$), "depressione-contentezza" ($p=0.021$), "sintomi somatici-benessere fisico" ($p=0.032$), sulla combinazione di tutte le variabili della scala SQ ($p=0.004$), su quella delle variabili PWB ($p=0.024$) e infine un effetto globale ($p=0.013$). Nel confronto diretto tra i due trattamenti (campioni indipendenti), è stata rilevata una maggior efficacia del trattamento WBT rispetto al trattamento CBT nell'aumento del benessere fisico ($p=0.007$). I risultati si trovano nelle tabelle 5 e 6. Per quanto riguarda la combinazione parametrica dei p -value, il WBT ha un'azione maggiore solo sulla combinazione delle variabili "ansia-relax" ($p=0.046$), ma non è migliore a livello globale. I risultati si trovano nelle tabelle 7 e 8.

10 Conclusioni

Miglioramenti significativi in alcune scale del PWB e dell'SQ sono stati apportati sia dal trattamento WBT, incentrato sul riconoscimento e la valorizzazione delle emozioni positive secondo i concetti proposti dal modello di benessere psicologico di Carol Ryff, sia dall'intervento CBT, focalizzato sul riconoscimento delle emozioni negative secondo i concetti proposti dalla teoria cognitiva. L'efficacia degli interventi scolastici ad orientamento cognitivo-comportamentale è già stata ampiamente dimostrata in letteratura. Il fatto che anche un protocollo basato sulla promozione del benessere psicologico, anzichè sulla rimozione della sintomatologia, sia risultato in qualche modo efficace sia nella riduzione della sintomatologia che nell'aumento del benessere psicologico dei soggetti, mette in evidenza l'applicabilità di tale protocollo in ambito scolastico e su una popolazione adolescenziale, sia in un'ottica terapeutica che preventiva.

Da un punto di vista prettamente statistico si possono fare le seguenti considerazioni. Come si può vedere, in questo caso, la soluzione non parametrica di permutazione e quella parametrica basata sul principio dell'ITT forniscono, salvo lievi differenze, gli stessi risultati, anche perchè non vi erano molti dati mancanti. Il

problema di fondo è che, mentre il metodo improntato sui metodi non parametrici di permutazione non elimina alcuna informazione e non solo rispecchia di più la realtà ma, come si è visto, fornisce risultati esatti, l'applicazione del principio dell'ITT comporta una riduzione della numerosità del campione, escono infatti dallo studio 9 soggetti, e falsa i punteggi ottenuti nella seconda rilevazione per 9 soggetti.

	WBT	CBT
	p-value	p-value
autonomia	0.127	0.070
padronanza ambiente	0.090	0.147
crescita personale	0.088	0.071
relazioni	0.110	0.174
scopo vita	0.200	0.002
accettazione	0.022	0.025
ansia	0.061	0.001
depressione	0.461	0.023
sintomi somatici	0.100	0.140
ostilità	0.012	0.090
relax	0.430	0.114
contentezza	0.440	0.264
benessere fisico	0.122	0.014
buona disponibilità	0.455	0.389

Tabella 5: WBT *vs.* CBT e applicazione dell'ITT (campioni dipendenti)

combinazioni	WBT	CBT
	p-value	p-value
ansia-relax	0.108	~ 0
depressione-contentezza	0.490	0.021
sintomi somatici-benessere fisico	0.292	0.032
ostilità buona-disponibilità	0.027	0.104
variabili scala SQ	0.140	0.004
variabili scala PWB	0.004	0.024
globale	0.243	0.013

Tabella 6: Combinazione parametrica dei *p-value* e applicazione dell'ITT (campioni dipendenti)

	parametrico
	p-value
autonomia	0.403
padronanza ambiente	0.368
crescita personale	0.441
relazioni	0.288
scopo vita	0.275
accettazione	0.436
ansia	0.061
depressione	0.098
sintomi somatici	0.382
ostilità	0.190
relax	0.128
contentezza	0.320
benessere fisico	0.007
buona disponibilità	0.385

Tabella 7: WBT *vs.* CBT e applicazione dell'ITT (campioni indipendenti)

combinazioni	parametrico
	p-value
ansia-relax	0.046
depressione-contentezza	0.120
sintomi somatici-benessere fisico	0.212
ostilità buona-disponibilità	0.269
variabili scala SQ	0.157
variabili scala PWB	0.453
globale	0.144

Tabella 8: Combinazione parametrica dei *p-value* e applicazione dell'ITT (campioni indipendenti)

Riferimenti bibliografici

- [1] Pesarin, F. (2001) *Multivariate Permutation tests: with application in Biostatistics*. John Wiley & Sons, Chichester-New York.
- [2] Celant, G., Pesarin, F., Salmaso L. (2000) Two sample permutation tests for repeated measures with missing values. *Journal of Applied Statistical Science*, 291-304.
- [3] Solari, A., Finos, L., Pesarin, F. (2003) Test di permutazione per la differenziazione genetica. Working Paper no. 2003.17, Dip. di Scienze Statistiche, Università di Padova.
- [4] Fava, G.A., Kellner, R., Perini, G., Fava, M., Michelacci, L., Munari, F., Evangelisti, L.P., Grandi, S., Bernardi, M. & Mastrogiacomo, I. (1993) Italian validation of the Symptom Rating Test (SRT) and Symptom Questionnaire (SQ). *Can J Psych*, 28, 177-123.
- [5] Ryff, C.D. (1989) Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *J Pers Soc Psychology*, 57, 1069-1081.
- [6] Fava, G.A. (2000) Cognitive behavioral therapy. In M. Fink (Ed), *Encyclopedia of stress*. Academic Press, San Diego, pp.484-497.
- [7] Wei, L.J., Lachin, J.M. (1984) Two-sample asymptotically distribution-free tests for incomplete multivariate observations. *Journal of the American Statistical Association*, 79, 653-661.
- [8] Wei, L.J., Johnson, W.E. (1985) Combining dependent tests with incomplete repeated measurements. *Biometrika*, 72, 359-364.
- [9] Barton, C.N., Cramer, E.C. (1989) Hypothesis testing in multivariate linear models with randomly missing data. *Communications in Statistics. Simulation and Computation*, 18, 875-895.
- [10] Maritz, J.S. (1995) A permutation test allowing for missing values. *Australian Journal of Statistics*, 37, 153-159.
- [11] Little, R.J.A., Rubin, D.B. (1987) *Statistical Analysis with Missing Data*. Wiley, New York.
- [12] Rubin, D.B. (1976) Inference and missing data. *Biometrika*, 63, 581-92.
- [13] Bacchieri, A., Della cioppa, G. (2004) *Fondamenti di ricerca clinica*. Springer-Verlag, Milano.

Working Paper Series
Department of Statistical Sciences, University of Padua

You may order paper copies of the working papers by emailing wp@stat.unipd.it

Most of the working papers can also be found at the following url: <http://wp.stat.unipd.it>

