

SULLA STIMA DELL'EFFETTO DI TRATTAMENTO IN SPERIMENTAZIONI CLINICHE CON PIANI STRATIFICATI NON BILANCIATI

Francesco Grigoletto, Sonia Masiero

Istituto di Igiene, Università di Padova.

RIASSUNTO

I dati provenienti da sperimentazioni cliniche sono generalmente categorizzati, oltre che per i livelli di trattamento, per le modalità di una o più variabili di stratificazione, e quasi inevitabilmente si presentano all'analisi non bilanciati. La stima complessiva dell'effetto di trattamento, derivata da una combinazione lineare degli effetti negli strati, risulta perciò ambigua perché non univoca. L'eventuale presenza di interazione trattamento per strato rende più complessa l'interpretazione dei risultati.

In questo lavoro, dopo avere esaminato la stimabilità di funzioni di parametri in presenza di un modello sovrapparametrizzato, consideriamo tre stimatori alternativi dell'effetto di trattamento, discutendone opportunità di applicazione ed interpretazione soprattutto in relazione al tipo di variabili di stratificazione. Quegli stimatori inoltre vengono confrontati con le soluzioni proposte da alcuni pacchetti di software statistico.

Parole chiave: trial clinici, modelli lineari, stratificazione, piani non bilanciati.

1. INTRODUZIONE

L'analisi dei dati relativi ad una sperimentazione clinica può trarre vantaggi da una pre-stratificazione dei pazienti basata su una o più variabili legate alla risposta sperimentale. Il guadagno in efficienza statistica è funzione della strettezza di quel legame ed è più rilevante in sperimentazioni con meno di 50 pazienti per gruppo di trattamento (Meinert, 1986, pag. 93). Le variabili di stratificazione riguardano caratteristiche dei pazienti quali, ad esempio, sesso, età e livello di gravità. Negli studi multicentrici, realizzati al fine di assicurare il numero necessario di pazienti in tempi più brevi, è utile stratificare i pazienti secondo la variabile centro. Questa servirà a spiegare differenze nella risposta sperimentale dovute sia

alla diversità da centro a centro delle caratteristiche demografiche e cliniche dei pazienti, sia al differente grado di adesione dei centri al protocollo di studio (Fleiss, 1986 a). L'assegnazione casuale dei trattamenti ai pazienti, effettuata separatamente ed indipendentemente all'interno degli strati, tende a rendere i gruppi di trattamento più comparabili.

Anche se la pre-stratificazione consente di programmare un piano bilanciato, ossia con lo stesso numero di pazienti nelle classi definite dai livelli dei fattori di stratificazione e di trattamento, quasi inevitabilmente i dati che si presentano all'analisi non sono bilanciati. Ciò può essere dovuto, tra l'altro, ad una incidenza della mortalità e ad un numero di uscite dallo studio diversi da strato a strato. Per quanto riguarda il fattore centro, vi può essere una diversa potenzialità dei centri a reclutare pazienti. Le cause di sbilanciamento possono essere in relazione diversa con la risposta sperimentale: tipicamente, la mortalità è in relazione stretta mentre la capacità dei centri a reclutare pazienti può essere in nessuna relazione con la risposta. Nelle considerazioni che seguono, assumeremo che la causa di sbilanciamento sia indipendente dalla risposta sperimentale, cosicché l'interpretazione dei risultati abbia pieno significato.

A complicare l'analisi si può aggiungere la presenza di interazione tra trattamento e fattori di stratificazione, ossia una differenza tra i livelli di trattamento che varia da strato a strato. Lo sbilanciamento dei dati e la presenza di interazione rendono la stima complessiva dell'effetto di trattamento ambigua, in quanto non univoca. Così l'analisi diventa problematica sia da un punto di vista metodologico che interpretativo. Pur adottando il medesimo modello lineare per esprimere gli stessi dati, due ricercatori possono giungere a conclusioni opposte circa la significatività statistica di un effetto di trattamento.

In questo articolo, dopo avere esaminato la stimabilità e la verifica di ipotesi relative a funzioni di parametri in presenza di un modello sovrapparametrizzato (paragrafo 2), vengono considerati tre stimatori alternativi dell'effetto di trattamento (paragrafo 3). Il paragrafo 4 è dedicato ad una discussione sulle opportunità di applicazione e sulla interpretazione dei tre stimatori. Questi, infine, vengono messi in relazione con le soluzioni proposte da tre comuni pacchetti di software statistico (paragrafo 5).

2. STIMABILITÀ CON MODELLI SOVRAPARAMETRIZZATI

Si consideri, per comodità di trattazione, un campione di soggetti classificati secondo i due livelli di un fattore A "Trattamento" e J livelli di un fattore B di stratificazione. Sia y_{ijk} la k -esima osservazione relativa al livello i di A e j di B.

Si presuppongano gli y_{ijk} espressi dal modello lineare:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \Phi_{ij} + e_{ijk}, \quad (1)$$

essendo gli e_{ijk} indipendentemente distribuiti come $N(0, \sigma^2)$.

Indicando con \mathbf{y} il vettore $n \times 1$ delle osservazioni, con \mathbf{X} la matrice $n \times p$ del piano sperimentale, con β il vettore dei $p = 2 + J + 2J$ parametri del modello e con \mathbf{e} il vettore $n \times 1$ dei termini di errore, il modello può essere così riscritto:

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{e}.$$

In generale siamo interessati alla stima di (o alla verifica d'ipotesi su) effetti espressi da combinazioni lineari $\mathbf{L}\beta$ dei parametri. Tuttavia, un effetto $\mathbf{L}\beta$ può non essere stimabile, e ciò accade ogniqualvolta la combinazione lineare include termini che rendono quell'effetto non univoco. Possibili ragioni della non stimabilità sono la sovraspecificazione del modello e la mancanza di osservazioni in corrispondenza di alcune celle.

Il problema della sovrapparametrizzazione può avere tre soluzioni: la riparametrizzazione del modello, l'applicazione di vincoli sui parametri, o il ricorso ad una inversa generalizzata per invertire la matrice $\mathbf{X}'\mathbf{X}$ e, quindi, ottenere la soluzione:

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y},$$

essendo $(\mathbf{X}'\mathbf{X})^{-1}$ una inversa generalizzata di $\mathbf{X}'\mathbf{X}$.

Questi tre metodi sono collegati tra di loro, nel senso che ad una determinata riparametrizzazione corrisponde la definizione di certi vincoli sui parametri oppure una inversa generalizzata specificata. Con riferimento a quest'ultimo metodo, per stimare la funzione lineare $\mathbf{L}\beta$ è necessario che sia verificata la condizione:

$$\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X}) = \mathbf{L},$$

essendo \mathbf{L} una combinazione lineare delle righe di \mathbf{X} . Si noti che il vettore stima \mathbf{Lb} e la sua varianza sono invarianti rispetto a tutte le possibili inverse generalizzate (Freund, Littell e Spector, 1986, pag. 6).

Se la funzione \mathbf{Lb} è stimabile, allora si può verificare l'ipotesi:

$$H_0: \mathbf{L}\beta = 0. \quad (2)$$

La somma dei quadrati per la verifica dell'ipotesi (2) è data da:

$$SSL = \mathbf{b}'\mathbf{L}'(\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}')\mathbf{Lb},$$

e la relativa funzione test è:

$$F = \frac{SSL}{MSE},$$

essendo MSE la stima della varianza σ^2 .

3. STIMATORI DELL'EFFETTO DI TRATTAMENTO

Con riferimento alla stima dell'effetto di trattamento, rappresentato dal fattore A, si considerino i seguenti tre stimatori.

Stimatore I

L'effetto di trattamento è espresso dalla differenza tra le medie $y_{i.}$ ($i=1,2$) marginali rispetto al fattore B:

$$d_I = y_{1.} - y_{2.} = \frac{\sum_j n_{1j} y_{1j.}}{n_{1.}} - \frac{\sum_j n_{2j} y_{2j.}}{n_{2.}},$$

dove $n_{1.}$ e $n_{2.}$ rappresentano il numero di pazienti destinati rispettivamente ai livelli 1 e 2 di trattamento. Tenuto conto del modello (1), il valore atteso di d_I è dato da:

$$\begin{aligned} E(d_I) &= \frac{\sum_j n_{1j} (\alpha_1 + \beta_j + \Phi_{1j})}{n_{1.}} - \frac{\sum_j n_{2j} (\alpha_2 + \beta_j + \Phi_{2j})}{n_{2.}} \\ &= \alpha_1 - \alpha_2 + \sum_j \beta_j \left(\frac{n_{1j}}{n_{1.}} - \frac{n_{2j}}{n_{2.}} \right) + \frac{1}{n_{1.}} \sum_j n_{1j} \Phi_{1j} - \frac{1}{n_{2.}} \sum_j n_{2j} \Phi_{2j}. \end{aligned} \quad (3)$$

I coefficienti dei parametri di quest'ultima espressione rappresentano gli elementi del vettore L_I che specificano il particolare confronto. Si noti che l'effetto di trattamento così stimato, oltre che considerare i parametri α_1 e α_2 , include una combinazione lineare sia degli effetti di strato β_j che di quelli di interazione Φ_{ij} . Si può pertanto dire che esso non è al netto dell'effetto di strato né di quello di interazione. La relativa somma dei quadrati può essere indicata con $SSL_I = SS(A|\mu)$, ossia come condizionata al solo effetto medio generale.

Stimatore II

L'effetto di trattamento è espresso da una media ponderata delle differenze di trattamento di ciascun livello del fattore B:

$$d_{II} = \frac{\sum_j w_j (y_{1j.} - y_{2j.})}{\sum_j w_j}. \quad (4)$$

Essendo la varianza di $y_{1j.} - y_{2j.}$ direttamente proporzionale a $\frac{n_{1j} + n_{2j}}{n_{1j} n_{2j}}$, se si

pone $w_j = \frac{n_{1j}n_{2j}}{n_{1j} + n_{2j}}$ con $j = 1, \dots, J$, allora lo stimatore d_{II} ha minima varianza in

assenza di interazione trattamento per strato (Fleiss, 1986 a).

Il suo valore atteso è dato da:

$$\begin{aligned} E(d_{II}) &= \frac{1}{\sum_j w_j} \sum_j w_j \left[(\alpha_1 + \beta_j + \Phi_{1j}) - (\alpha_2 + \beta_j + \Phi_{2j}) \right] \\ &= \alpha_1 - \alpha_2 + \frac{1}{\sum_j w_j} \sum_j w_j (\Phi_{1j} - \Phi_{2j}). \end{aligned} \quad (5)$$

Se ne deduce che i coefficienti del vettore \mathbf{L}_{II} sono tali da esprimere l'effetto di trattamento al netto dell'effetto di strato ma non di quello di interazione. Così la somma dei quadrati relativa può essere espressa da $SSL_{II} = SS(A|\mu, B)$.

Stimatore III

Attribuendo allo stimatore d_{II} , espresso dalla (4), pesi w_j costanti e uguali a 1, si ottiene lo stimatore:

$$d_{III} = \frac{1}{J} \sum_j (y_{1j} - y_{2j}), \quad (6)$$

il cui valore atteso è:

$$\begin{aligned} E(d_{III}) &= \frac{\sum_j \left[(\alpha_1 + \beta_j + \Phi_{1j}) - (\alpha_2 + \beta_j + \Phi_{2j}) \right]}{J} \\ &= \alpha_1 - \alpha_2 + \frac{1}{J} \sum_j (\Phi_{1j} - \Phi_{2j}). \end{aligned} \quad (7)$$

Si noti che qui la stima dell'effetto di trattamento è al netto dell'effetto di strato, pur non essendo al netto degli effetti di interazione. Tuttavia, se si assumono gli usuali vincoli $\sum_j \Phi_{1j} = \sum_j \Phi_{2j} = 0$, lo stimatore d_{III} dipende unicamente dall'effetto di trattamento, per cui la somma dei quadrati diventerà $SSL_{III} = SS(A|\mu, B, A*B)$, intendendo con $A*B$ l'effetto di interazione.

4. PROPRIETÀ DEGLI STIMATORI d_I, d_{II} E d_{III}

Gli stimatori introdotti nel paragrafo 3 derivano da differenti sistemi di ponderazione utilizzati nel calcolare la media degli effetti di trattamento dei singoli strati. Ai sistemi di ponderazione corrispondono diversi tipi di condizionamento per l'effetto del fattore trattamento, e diverse somme dei quadrati da utilizzare per la verifica di ipotesi formulate sugli stessi. Comparando le espressioni relative ai tre operatori si deducono le seguenti identità:

$$\text{Con dati bilanciati: } d_I = d_{II} = d_{III}$$

$$\text{Con interazione assente: } d_{II} = d_{III}$$

Le differenziazioni tra i tre stimatori si evidenziano, quindi, con lo sbilanciamento delle numerosità delle celle, il cui effetto sullo stimatore è diverso a seconda che l'interazione trattamento per strato sia assente o presente. L'esistenza di interazione può essere indagata mediante la verifica di ipotesi formulate su definite combinazioni lineari Lb . Ma l'interpretazione dell'effetto di trattamento basata sui risultati di tali verifiche può essere inficiata dai possibili errori di I e II tipo. Perciò riteniamo sia più opportuno basare la presunzione di presenza o no di interazione sulla conoscenza a priori, derivata da studi precedenti, del possibile ruolo delle variabili di stratificazione. Ad esempio, se la formazione degli strati è basata su caratteristiche dei pazienti, quali lo stadio di sviluppo della malattia o l'età, può essere che l'effetto di trattamento sia minore nei soggetti meno gravi o nei soggetti più anziani. Quando, invece, gli strati sono rappresentati da cliniche, e queste sono strettamente controllate nell'applicazione del protocollo di studio che tende a reclutare pazienti con le medesime caratteristiche e a valutarli con criteri omogenei, allora può non esservi ragione di pensare che la differenza tra i trattamenti vari da centro a centro, ossia può essere assunta a priori l'assenza di interazione. Per quanto detto, nelle considerazioni che seguono conviene distinguere il caso di interazione assente da quello di interazione presente.

Modello senza interazione

Lo stimatore d_I dipende, oltre che dall'effetto di trattamento (fattore A) anche dagli effetti del fattore B: l'espressione (3), infatti, mostra che il suo valore atteso è funzione, oltre che degli α_j , anche dei β_j . Una verifica di ipotesi basata su d_I può portare ad affermare che $\alpha_1 \neq \alpha_2$ anche se $\alpha_1 = \alpha_2$ ma vi è una differenza tra i β_j , ossia l'effetto del fattore A è apparente e dovuto invece al fattore B. D'altra parte, è possibile anche concludere che $\alpha_1 = \alpha_2$ quando $\alpha_1 \neq \alpha_2$, ma

$\alpha_1 - \alpha_2 \cong \sum_j \beta_j \left(\frac{n_{1j}}{n_{1.}} - \frac{n_{2j}}{n_{2.}} \right)$: qui B nasconde l'effetto di A. Perciò, la stima

prodotta da d_I è difficilmente interpretabile. Ciò nonostante, alcuni autori (Collet, Roger e Whitehead, 1991) ritengono questo stimatore il più idoneo tra quelli proposti, dal momento che risulta essere il meno condizionato a vincoli comunque arbitrari. A nostro avviso, tuttavia, non conviene stimare l'effetto di A sulla base di medie di dati aggregati, senza tenere conto degli strati.

Lo stimatore d_{II} , espresso dalla (4), opera invece una media ponderata degli effetti di A all'interno degli strati. In assenza di interazione, il suo valore atteso (che si identifica con quello di d_{III}) dipende dai soli parametri α_i ed il sistema di ponderazione da cui viene definito lo fa risultare a varianza minima. Per quanto detto, quando è possibile assumere a priori l'assenza di interazione, d_{II} è lo stimatore da preferire.

Modello con interazione

Può essere opportuno assumere una presenza di interazione non solo quando la stratificazione è basata su variabili prognostiche, ma anche quando i pazienti sono classificati per centro. I centri, infatti, possono reclutare tipi di pazienti diversi per caratteristiche demografiche e prognostiche e possono valutare la risposta sperimentale con modalità disomogenee, influenzando così sulla differenza tra i trattamenti. In questo caso, anche se il sistema di ponderazione che utilizza non è più ottimale, d_{II} mantiene la sua validità in quanto $w_j / \sum w_j$ è una stima consistente della proporzione di pazienti nella popolazione sotto studio che provengono dallo strato j (Elston e Bush, 1964). Di conseguenza, d_{II} è uno stimatore consistente della differenza tra le medie dei trattamenti nella popolazione.

Può essere, diversamente, che la numerosità degli strati non abbia alcuna relazione con definite categorie della popolazione. Questo accade, in particolare, quando i pesi w_j dipendono essenzialmente dalla capacità di reclutamento delle cliniche. In tal caso, il sistema di pesi utilizzato in d_{II} perde il suo significato, e risulta inapplicabile. In questa situazione si potrebbe affermare che, a prescindere dalle caratteristiche di ottimalità statistica, nessuno stimatore dell'effetto medio generale può aver senso dal punto di vista interpretativo. Volendo comunque ricavare una stima, si può attribuire agli effetti specificati per strato un peso uguale, come fa lo stimatore d_{III} espresso dalla (6).

Esso non gode delle proprietà di ottimalità dello stimatore d_{II} , essendo $Var(d_{III}) \geq Var(d_{II})$. Inoltre si osserva come, utilizzando d_{III} , un risultato positivo in pochi strati con bassa numerosità possa controbilanciare risultati negativi in strati

molto numerosi (o il contrario). Di qui l'opportunità di organizzare il disegno sperimentale in strati (in particolare: centri) che possono fornire numerosità campionarie confrontabili. Lo stimatore d_{III} , tuttavia, risulta essere assai frequentemente utilizzato per la sua proprietà di fornire stime dell'effetto di trattamento al netto dell'effetto di strato e di interazione.

5. LE SOLUZIONI DI ALCUNI SOFTWARE STATISTICI

Ci proponiamo di mettere in relazione, in questo paragrafo, gli stimatori d_p , d_{II} e d_{III} con le soluzioni proposte da tre comuni pacchetti di software statistico, ossia SAS, GLIM e BMDP. Faremo riferimento, per maggiore chiarezza espositiva, all'esempio di Tab. I, dove i dati sono classificati secondo due fattori rispettivamente di 2 e 3 livelli.

Tab. I: Dati classificati secondo un fattore A di trattamento ed uno B di stratificazione.

		STRATO (B)		
		1	2	3
TRATTAMENTO (A)	1	27	9	14
		15	17	22
			6	15
	2	31	18	
		14	8	41
		27	24	

Volendo calcolare gli elementi dei vettori di coefficienti L_I , L_{II} e L_{III} , riferiti rispettivamente agli stimatori d_p , d_{II} e d_{III} , si considerino le espressioni (3), (5) e (7) inserendovi gli appropriati valori n_{ij} . Si troveranno i seguenti risultati:

$$\begin{aligned}
 E(d_I) &= \alpha_1 - \alpha_2 + 0,2\bar{7}\beta_i + 0,2\bar{1}\beta_2 + 0,1\bar{6}\beta_3 \\
 &+ 0,2\bar{2}\Phi_{11} + 0,4\bar{4}\Phi_{12} + 0,3\bar{3}\Phi_{13} \\
 &- 0,5\bar{5}\Phi_{21} - 0,3\bar{3}\Phi_{22} - 0,1\bar{6}\Phi_{23},
 \end{aligned}$$

$$\begin{aligned}
 E(d_{II}) &= \alpha_1 - \alpha_2 \\
 &+ 0,3655\Phi_{11} + 0,4061\Phi_{12} + 0,2284\Phi_{13} \\
 &- 0,3655\Phi_{21} - 0,4061\Phi_{22} - 0,2284\Phi_{23},
 \end{aligned}$$

$$\begin{aligned}
 E(d_{III}) &= \alpha_1 - \alpha_2 \\
 &+ 0,3\bar{3}\Phi_{11} + 0,3\bar{3}\Phi_{12} + 0,3\bar{3}\Phi_{13} \\
 &- 0,3\bar{3}\Phi_{21} - 0,3\bar{3}\Phi_{22} - 0,3\bar{3}\Phi_{23}.
 \end{aligned}$$

La conoscenza di tali coefficienti permette di effettuare la stima e la verifica di ipotesi sulle funzioni da essi definite.

La procedura SAS GLM (SAS, 1982, pag. 892–996) esplicita quattro tipi di vettori L , cui corrispondono altrettanti tipi di somme dei quadrati.

Il Tipo I identifica gli effetti secondo l'ordine con cui vengono dichiarati nella specificazione del modello: (A, B, A*B). Ciascun effetto è stimato al netto di quelli dichiarati precedentemente, per cui il vettore L , che rende stimabile l'effetto di trattamento, è del tipo L_j , esplicitato dal valore atteso $E(d_j)$. È possibile ottenere tale vettore dalla procedura, come si legge in Tab. II, nella colonna Type I. La somma dei quadrati relativa all'effetto A di trattamento, dichiarato per primo, non è condizionato ad alcuno degli altri due: quindi si può indicare come $SS(A|\mu)$. Il suo valore si legge nella Tab. III.

Tab. II: Coefficienti delle funzioni stimabili dell'effetto di Trattamento esplicitati da SAS GLM.

General Linear Model Procedure Estimable Function for : TRATTAMENTO					
Effect		Type I	Type II	Type III	Type IV
TRATTAMENTO	1	L2	L2	L2	L2
	2	-L2	-L2	-L2	-L2
STRATO	1	-0.2778*L2	0	0	0
	2	0.1111*L2	0	0	0
	3	0.1667*L2	0	0	0
TRATTAMENTO*	1 1	0.2222*L2	0.3655*L2	0.333*L2	0.333*L2
	1 2	0.4444*L2	0.4061*L2	0.333*L2	0.333*L2
STRATO	1 3	0.3333*L2	0.2284*L2	0.333*L2	0.333*L2
	2 1	-0.5*L2	-0.3655*L2	-0.333*L2	-0.333*L2
	2 2	-0.3333*L2	-0.4061*L2	-0.333*L2	-0.333*L2
	2 3	-0.1667*L2	-0.2284*L2	-0.333*L2	-0.333*L2

Il Tipo II della procedura SAS GLM condiziona ogni effetto principale del modello ad ogni altro, escluse le sue interazioni con gli altri effetti. Così il vettore L di coefficienti si identifica con L_{II} , esplicitato dall'espressione $E(d_{II})$ e in colonna Type II nella Tab. II. La somma dei quadrati dell'effetto di trattamento del Tipo II si può esprimere come $SS(A|\mu, B)$ (vedi Tab. III).

Il Tipo III condiziona ogni effetto ad ogni altro nel modello, con il vincolo: $\sum_j \Phi_{ij} = \sum_i \Phi_{ij} = 0$ (Searle, 1987, pag. 463), ottenendo un vettore L_{III} identificato da $E(d_{III})$ ed esplicitato dal SAS come in colonna Tipo III della Tab. II. La somma dei quadrati dell'effetto di trattamento condizionato agli effetti di stratificazione e di interazione sarà del tipo: $SS(A|\mu, B, A*B)$ (vedi Tab. III).

Il Tipo IV costruisce delle ipotesi sulle celle piene e vuote, ipotesi che dipendono dall'ordine in cui le celle sono disposte. Le stime dei vari effetti e le

somme dei quadrati di questo tipo varieranno, perciò, al variare della disposizione delle osservazioni negli strati. Se non ci sono celle vuote, come nel caso particolare della Tab. I, il vettore \mathbf{L} identificato dal Tipo IV è identico a \mathbf{L}_{III} (vedi Tab. II) e la sua somma dei quadrati è uguale a quella ottenuta con il Tipo III (vedi Tab. III).

La procedura P2V del BMDP (BMDP, 1979, pag. 540–580) prevede la stessa stima dell'effetto di trattamento in presenza di uguale o diversa dimensione dei vari strati. Così lo stimatore dell'effetto di trattamento ottenibile da questa procedura è indipendente dalle numerosità delle celle. Esso, perciò, si identifica con d_{III} . Dalla BMDP2V si ottengono, di conseguenza, somme dei quadrati per la verifica d'ipotesi (vedi Tab. III) esprimibili come $SS(A|\mu, B, A*B)$, che si identificano con quelle del Tipo III della SAS GLM.

Il GLIM (Aitkin ed altri, 1990, pag. 72–73), invece, offre la possibilità di ottenere più di un tipo di somma dei quadrati per la verifica d'ipotesi. In particolare, si ricavano con facilità le somme dei quadrati del tipo $SS(A|\mu)$ e $SS(A|\mu, B)$, come si vede dalla Tab. III, corrispondenti al Tipo I e Tipo II della SAS GLM. La difficoltà di calcolo del terzo genere di somma dei quadrati deriva dalla mancanza nel pacchetto di operatori matriciali, chiamati in causa quando il fattore di stratificazione a cui si vuole condizionare l'effetto di trattamento presenta più di 2 livelli, come nel caso particolare dell'esempio in Tab. I.

Tab. III: Somme dei quadrati relative al fattore A (Trattamento) esplicitate da SAS, BMDP e GLIM, e relativo livello di significatività osservato P.

Stimatore	Tipo SSL	Valore SSL	Valore P	Software utilizzabile
d_I	$SS(A \mu)$	246.677778	0.0646	SAS GLM Tipo I, GLIM
d_{II}	$SS(A \mu, B)$	210.133333	0.0839	SAS GLM Tipo II, GLIM
d_{III}	$SS(A \mu, B, A*B)$	318.942857	0.0403	SAS GLM Tipo III e Tipo IV, BMDP2V

Si osservi, infine, che un ricercatore che avesse definito il convenzionale livello di significatività pari al 5% sulla base dei dati qui esemplificati affermerebbe la non significatività statistica del trattamento se utilizzasse gli stimatori d_I oppure d_{II} , ma arriverebbe ad una conclusione opposta utilizzando d_{III} .

RINGRAZIAMENTO

Gli autori esprimono profonda gratitudine a Michael Gent, Robin Roberts e Francesca Patarnello per le assai utili discussioni durante la preparazione di questo articolo.

RIFERIMENTI BIBLIOGRAFICI

- AITKIN M., ANDERSON D. Francis B. e HIND J., (1990), *Statistical Modelling in GLIM*, Oxford University Press, Oxford.
- BMDP, (1979), *Biomedical Computer Programs – Series P*, University of California Press, Berkeley, CA.
- COLLETT D., ROGER J. e WHITEHEAD A., (1991), *Using Proc GLM in SAS*, Unpublished Course Notes, University of Reading.
- ELSTON R.C. e BUSH N., (1964), The Hypotheses that can be tested when there are Interaction in an Analysis of Variance Model, *Biometrics*, 20, 681–698.
- FLEISS J.L., (1986 a), Analysis of Data from Multiclinic Trials, *Controlled Clinical Trials*, 7, 267–275.
- FLEISS J.L., (1986 b), *The Design and Analysis of Clinical Experiments*, John Wiley & Sons, New York.
- FREUND R.J., LITTELL R.C. e SPECTOR P.C., (1986), *SAS System for Linear Model*, SAS Institute, Inc., Cary, N.C.
- MEINERT C.L., (1986), *Clinical Trials – Design, Conduct and Analysis*, Oxford University Press, New York.
- SAS Institute, Inc., (1990), *SAS/STAT User's Guide, Version 6*, Fourth Edition, Vol. 2, SAS Institute, Inc, Cary, N.C.
- SEARLE S.R., (1987), *Linear Models for Unbalanced Data*, John Wiley & Sons, New York.

**ON THE TREATMENT EFFECT ESTIMATE IN STRATIFIED
CLINICAL TRIALS WITH UNBALANCED DATA****SUMMARY**

Data from clinical trials are generally categorized by the levels of one or more stratification variables, and most often the number of observations is different among strata. The overall estimate of the treatment effect is derived from treatment differences within strata by means of a weighted average for which different sets of weights can be chosen. This makes the estimate not uniquely defined and therefore ambiguous. The possible presence of treatment-by-stratum interaction makes the results' interpretation furthermore complex.

In this paper we first approached an over-specified model and examined the estimability of a function of its parameters. Then we considered three different estimators of the overall treatment effect and discussed their interpretation and application opportunities in relation to different types of stratification variables. Finally, the estimators were compared to the solutions provided by some common statistical packages.

Keywords: clinical trials, linear models, stratification, unbalanced data.