STATISTICA, anno LXV, n. 4, 2005

UNA TECNICA DI DISAGGREGAZIONE FUZZY

Alessandro Polli

1. INTRODUZIONE

Nei problemi di decisione statistica l'incertezza è comunemente rappresentata tramite misure di probabilità. Tale schema, tuttavia, in presenza di un *information set* limitato, potrebbe rivelarsi inappropriato (Krätschmer, 2003). Da un punto di vista generale sarebbe quindi auspicabile adottare procedure più flessibili. Le misure del grado di incertezza, in altri termini, dovrebbero in alcuni casi soddisfare soltanto alcune proprietà generali.

Un'importante classe di misure è quella delle c.d. misure *fuzzy*, caratterizzate da monotonicità e da alcune condizioni di normalizzazione. I concetti di misura ed integrale *fuzzy* generalizzano la definizione comune di misura sostituendo la proprietà di additività con un requisito più debole, la monotonicità rispetto alla funzione d'insieme, strettamente legata alla nozione di *capacità* introdotta da Choquet (1954).

Nell'approccio di Walley (1991) l'incertezza è rappresentata da *previsioni minime*, intese come funzionali in $\Re$ e di cui le misure *fuzzy* rappresentano un caso particolare. Uno dei concetti chiave nella teoria delle previsioni minime è quello di *coerenza*, che esprime un requisito minimo di consistenza e la cui formulazione si deve originariamente a De Finetti (1974) che, come è noto, introduce il concetto di coerenza per fornire un fondamento comportamentale alla teoria della probabilità. Per garantire la proprietà di coerenza delle misure *fuzzy* Walley introduce il concetto di *estensione naturale*, che rappresenta una *previsione minima coerente*. Sotto certe condizioni l'estensione naturale è rappresentabile in termini di integrale di Choquet e la misura *fuzzy* assume un contenuto probabilistico.

2. AGGREGAZIONE DI INFORMAZIONI. CONCETTI GENERALI

Il principale motivo di interesse dell'integrale *fuzzy* è la sua capacità di rappresentare l'intera gamma di interazioni tra elementi componenti un *information set*, dalla ridondanza o *submodularità* (interazione negativa) alla sinergia o *supermodularità* (interazione positiva).

In generale sia $\Omega = \{\omega_1, \dots, \omega_p\}$ un insieme di *alternative* o *azioni* a ciascuna delle quali corrisponde un vettore $\mathbf{x}$ di attributi o criteri, che rappresentano il *pay off* associato alla scelta di uno degli elementi in Ω. Sia $\mathcal{X}$ l'insieme dei *pay off* $\mathbf{x}$. La valutazione delle alternative può essere effettuata attraverso la costruzione di una funzione d'utilità $u(\cdot)$ tale che

$$x_i \succ x_j \Leftrightarrow u(x_i) > u(x_j) \tag{1}$$

con $u(\cdot)$ funzione n-dimensionale. Un modo semplice di costruire $u(\cdot)$ è considerare idonee funzioni di utilità unidimensionali u_i per ciascun attributo, aggregate attraverso un operatore H:

$$u(x_1, \dots, x_n) = H[u_1(x_1), \dots, u_n(x_n)]\,. \tag{2}$$

La funzione $u(\cdot)$ è idonea se realizza la condizione (1). Una delle funzioni che soddisfano la (1) è la funzione somma:

$$u(x_1, \dots, x_n) = \sum_{i=1}^{n} u_i(x_i)\,. \tag{3}$$

3. PROPRIETÀ DEGLI OPERATORI DI AGGREGAZIONE

Per semplicità di notazione poniamo $a_i = u_i(x_i)$, dove a_i può essere considerato come un grado di preferenza per una coppia assegnata di alternative rispetto al criterio *i*. Senza perdita di generalità si assuma che $a_i \in [0,1]$. L'operatore H deve essere caratterizzato dalle seguenti proprietà:

1. *Proprietà degli estremi.*

$$H(0, \dots, 0) = 0, \qquad H(1, \dots, 1) = 1\,. \tag{4}$$

2. *Idempotenza.*

$$H(a, \dots, a) = a \qquad \forall a \in [0,1]\,. \tag{5}$$

3. *Continuità.*
4. *Monotonicità (usualmente non decrescente) rispetto a ciascun argomento.*
5. *Commutatività* (o *neutralità*), intesa come indifferenza rispetto ai criteri.
6. *Decomponibilità.*

$$H^{(n)}(a_1,\ldots,a_k,a_{k+1},\ldots,a_n)= \\ =H^{(n)}(a,\ldots,a,a_{k+1},\ldots,a_n)= \\ =a+H^{(n-k)}(a_{k+1},\ldots,a_n) \tag{6}$$

dove $a=H^{(k)}(a_1,\ldots,a_k)$.

7. *Stabilità in caso di trasformazione lineare positiva.*

$$H(ra_1+t,\ldots,ra_n+t)= \\ =rH(a_1,\ldots,a_n)+t \qquad \forall r\in\Re^+,\forall t\in\Re. \tag{7}$$

8. *Concatenabilità ordinale.*

$$H^{(n+1)}\{H^{(n)}[a_{(1)},\ldots,a_{(n)}],H^{(n)}[a_{(2)},\ldots,a_{(n+1)}];\ldots,H^{(n)}[a_{(n+1)},\ldots,a_{(2n)}]\}= \\ =H^{(n)}\{H^{(n+1)}[a_{(1)},\ldots,a_{(n+1)}],H^{(n+1)}[a_{(2)},\ldots,a_{(n+2)}];\ldots,H^{(n+1)}[a_{(n)},\ldots,a_{(2n)}]\}. \tag{8}$$

9. *Concatenabilità ordinale con permutazioni* (Grabisch, 1994)

$$H^{(n+1)}\{[H^{(n)}(a_{(1)},\ldots,a_{(n)}),H^{(n)}(a_{(2)},\ldots,a_{(n+1)});\ldots,H^{(n)}(a_{(n+1)},\ldots,a_{(2n)})]_\sigma\}= \\ H^{(n)}\{[\mathcal{H}^{(n+1)}(a_{(1)},\ldots,a_{(n+1)}),H^{(n+1)}(a_{(2)},\ldots,a_{(n+2)});\ldots,H^{(n+1)}(a_{(n)},\ldots,a_{(2n)})]_\sigma\}. \tag{9}$$

4. PRINCIPALI CLASSI DI OPERATORI DI AGGREGAZIONE

In via preliminare definiamo una misura *fuzzy* su **X** la funzione d'insieme $\mu:\mathcal{F}(\mathbf{X})\to[0,1]$ che soddisfa gli assiomi 1) $\mu(0)=0$, $\mu(\mathbf{X})=1$ e 2) $A\subset B\subseteq X\to\mu(A)<\mu(B)$.

Una misura *fuzzy* tale che $\mu(A\cup B)\ge\mu(A)+\mu(B)$ è detta *supermodulare*; *submodulare* se $\mu(A\cup B)\le\mu(A)+\mu(B)$; *modulare* nel caso in cui soddisfa la condizione $\mu(A\cup B)=\mu(A)+\mu(B)$, con $\mu(A\cap B)=\varnothing$. In quest'ultimo caso la misura *fuzzy* è definita esaustivamente dagli *n* pesi $\mu(x_1),\ldots,\mu(x_n)$, mentre negli altri è necessario definire i 2^n pesi corrispondenti ai 2^n sottoinsiemi di **X**.

Sia μ una misura *fuzzy* su **X**. Definiamo integrale di Sugeno della funzione $f:\mathbf{X}\to[0,1]$ rispetto a μ la classe di operatori

$$\mathcal{S}_\mu[f(x_1),\ldots,f(x_n)] \triangleq \bigvee_{i=1}^{n}[f(x_{(i)})\wedge\mu(A_{(i)})] \tag{10}$$

in cui $\cdot_{(i)}$ è una permutazione degli indici tale che $0\le f(x_{(1)})\le\ldots\le f(x_{(n)})\le 1$ e $A_{(i)}:=\{x_{(i)},\ldots,x_{(n)}\}$.

L'integrale di Sugeno calcola una media distorta delle misure $f(x_1),\dots f(x_n)$ utilizzando un operatore di aggregazione non lineare.

Sia μ una misura *fuzzy* su **X**. L'integrale di Choquet rispetto a μ di una funzione $f:\mathbf{X}\to[0,1]$ è definito da

$$\mathcal{C}_\mu[f(x_1),\dots,f(x_n)] \triangleq \sum_{i=1}^{n}[f(x_{(i)})-f(x_{(i-1)})]\ \mu(A_{(i)}) \qquad f(x_{(0)})=0 \tag{11}$$

in cui nuovamente $\cdot_{(i)}$ è una permutazione degli indici e $A_{(i)}:=\{x_{(i)},\dots,x_{(n)}\}$. L'integrale di Choquet genera medie 'distorte' delle misure $f(x_1),\dots,f(x_n)$ per mezzo di un operatore di aggregazione lineare. Krätschmer (2003) dimostra che una *previsione minima coerente* nel senso di Walley – cioè una previsione lineare nel senso di De Finetti – è rappresentabile tramite tale classe di operatori.

Definite le principali famiglie di operatori di aggregazione, analizziamo in particolare la classe degli operatori lineari OWA-*Ordered weighted average* (Yager, 1991), definita da

$$OWA_{(w_1,\dots,w_n)}(a_1,\dots,a_n)=\sum_{i=1}^{n}w_i a_i \tag{12}$$

dove $\sum_{i=1}^{n}w_i=1$ e $a_{(1)}\le a_{(2)}\le\ldots\le a_{(n)}$. La classe degli operatori OWA è caratterizzata da idempotenza, continuità e monotonicità (Fodor *et al.* 1995). Può mostrarsi che 1) gli operatori di classe OWA sono integrali di Choquet (Grabisch, 1996): infatti basta osservare che $\mu(A_{(i)})=\sum_{j=0}^{n}w_{n-j},\forall A\text{: } |\mathrm{A}|=i$ nella relazione (11); 2) ogni integrale di Choquet che goda anche della proprietà commutativa è tale che $\mu(A)$ dipende soltanto dalla cardinalità di $|A|$ e coincide con un operatore OWA. Da adesso in poi si assumerà l'esistenza di tale proprietà soltanto in alcuni ambiti di applicazione (quali ad esempio l'analisi di serie storiche).

5. UNA TECNICA DI DISAGGREGAZIONE FUZZY

I risultati presentati nelle precedenti sezioni rappresentano i fondamenti di un metodo di disaggregazione *fuzzy* da noi derivato sulla base della teoria degli operatori OWA.

Sia $\mathbf{w}^{\mathrm{T}}$ un vettore $(1\times k)$ di pesi $\{w_1\ldots w_k\}$ tale che $\sum_{i=1}^{k}w_i=1$ e sia

$$OWA_{w_1\ldots w_k}(a_1\ldots a_k)=\sum_{\mathrm{i}=1}^{\mathrm{k}}w_i a_i \tag{13}$$

la classe degli operatori OWA su $[0,\ 1]^{\mathrm{k}}$.

Nel problema in esame, in cui si dispone prevalentemente di informazione aggregata, l'osservazione al generico tempo *n* può essere considerata come il risultato dell'applicazione di un operatore OWA alla serie aggregata; infatti:

$$\overline{a}_n = \sum\nolimits_{i=1}^{k} a_{in} = \sum\nolimits_{i=1}^{k} w_{in}\overline{a}_n \equiv OWA_{w_{1n}\ldots w_{kn}}(\overline{a}_n) \tag{14}$$

in cui $w_{in} = a_{in}/\overline{a}_n$ – con $w_{in} \in [0,1]$ e $\sum\nolimits_{i=1}^{k} w_{in} = 1$ – rappresenta il peso della *i*-esima serie disaggregata a_{in} rispetto alla serie aggregata $\overline{a}_n$.

In maniera del tutto ovvia, considerando simultaneamente le *t* osservazioni aggregate, in notazione vettoriale potremo scrivere:

$$\underset{(t\times k)}{\begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{t1} & \cdots & a_{tk} \end{bmatrix}} = \underset{(t\times t)}{\begin{bmatrix} \overline{a}_1 & & \\ & \ddots & \\ & & \overline{a}_t \end{bmatrix}} \underset{(t\times k)}{\begin{bmatrix} w_{11} & \cdots & w_{1k} \\ \vdots & \ddots & \vdots \\ w_{t1} & \cdots & w_{tk} \end{bmatrix}} \tag{15}$$

e $\mathbf{A} = <\overline{\mathbf{a}}> \mathbf{W}$ in notazione compatta.

Nelle applicazioni pratiche, in cui una o più righe di $\mathbf{W}$ e le corrispondenti colonne di $\mathbf{A}$ non sono note, permutando e introducendo le partizioni *O* (contenente le informazioni disaggregate disponibili) e *U* (contenente soltanto osservazioni aggregate) si avrà che

$$\underset{\{[(t-s)+s]\times k]}{\begin{bmatrix} \mathbf{A_U} \\ \mathbf{A_O} \end{bmatrix}} = \underset{\{[(t-s)+s]\times[(t-s)+s]\}}{\left[\begin{array}{c|c} <\overline{\mathbf{a}}_\mathbf{U}> & \mathbf{0} \\ \hline \mathbf{0} & <\overline{\mathbf{a}}_\mathbf{O}> \end{array}\right]} \underset{\{[(t-s)+s]\times k\}}{\begin{bmatrix} \mathbf{W_U} \\ \mathbf{W_O} \end{bmatrix}}. \tag{16}$$

Nella (16) le componenti note sono, nell'ordine, $\mathbf{A_O}$, le due partizioni della matrice diagonale $<\overline{\mathbf{a}}>$ e $\mathbf{W_O}$. Uno stimatore *naïf* di $\mathbf{A_U}$ può essere quindi determinato ipotizzando un'evoluzione lineare dei pesi per l'intero intervallo di ricostruzione, a condizione che sia noto, con riferimento agli indici temporali $n \in O,\ O = \{t - s + 1 \quad \ldots \quad t\}$, il vettore di pesi $\{w_{n1} \quad \ldots \quad w_{nk}\}^T$ riferito ai *k* settori, per almeno due periodi di osservazione. Lo stimatore del generico elemento della partizione $\mathbf{W_U}$ può essere scritto come

$$\hat{w}_{ni} = E[w_{ni} | w_{t-s+1,i} \quad \ldots \quad w_{t,i}] \qquad \forall n \in U,\ U = \{1 \quad \ldots \quad t-s\} \tag{17}$$

da cui è immediato scrivere

$$\underset{\{[(t-s)+s]\times k]}{\begin{bmatrix} \hat{\mathbf{A}}_\mathbf{U} \\ \mathbf{A_O} \end{bmatrix}} = \underset{\{[(t-s)+s]\times[(t-s)+s]\}}{\left[\begin{array}{c|c} <\overline{\mathbf{a}}_\mathbf{U}> & \mathbf{0} \\ \hline \mathbf{0} & <\overline{\mathbf{a}}_\mathbf{O}> \end{array}\right]} \underset{\{[(t-s)+s]\times k\}}{\begin{bmatrix} \hat{\mathbf{W}}_\mathbf{U} \\ \mathbf{W_O} \end{bmatrix}}. \tag{18}$$

Lo stimatore *naïf* potrebbe risultare utile nel caso in cui la serie osservata non possa essere messa in relazione, a causa di *lack* informativi, con un *set* di covariate disaggregate, il che rende inapplicabili le metodologie derivate dall'algoritmo di Chow e Lin (1971), caso che si presenta non di rado – ad esempio nei tentativi di ricostruzione di serie centenarie di contabilità nazionale – laddove l'*information set* si riduca a dati disaggregati per pochi anni *benchmark*.

Naturalmente possono essere scelti anche profili evolutivi dei pesi diversi da quello lineare, ma per rispettare le proprietà dell'operatore OWA è di fondamentale importanza che tali profili siano caratterizzati da additività e sommabilità ad uno per ciascun periodo compreso nell'intervallo temporale di stima, il che ne restringe la scelta all'interno di un ristretto numero di specificazioni funzionali.

6. UN ESEMPIO DI APPLICAZIONE

La tecnica presentata nella sezione precedente è stata applicata ad un problema di disaggregazione della serie del valore aggiunto ottenuta sommando i valori aggiunti dei settori di classificazione DA (industrie alimentari, delle bevande e del tabacco), DB (industrie tessili e dell'abbigliamento), DC (pelle e prodotti in pelle), DD (legno e prodotti in legno), DE (carta e cartotecnica) pubblicati dall'Istat e relativi al periodo 1970-2003.

L'*information set* utilizzato per effettuare i *test* è composto da 1) la serie aggregata del valore aggiunto relativa ai settori citati per tutto l'intervallo temporale e 2) osservazioni disaggregate per quattro anni *benchmark* (1970, 1981, 1992, 2003).

La Tavola 1 riporta le principali statistiche descrittive relative rispettivamente alle serie originarie e a quelle stimate.

TAVOLA 1

Confronti tra serie Istat e serie stimate. Statistiche descrittive

		Media	Mediana	Max	Min	Dev. Std.	γ_1	γ_2	Jarque-Bera	ρ
DA	O	15.083,96	14.750,92	20.485,75	8.918,60	3.765,46	-0,1290	1,6093	2,8344	0,9921
	S	15.475,16	15.041,72	21.031,79	8.918,60	4.092,58	-0,0720	1,5418	3,0417	
DB	O	18.166,64	18.636,60	22.272,98	11.012,84	3.311,73	-0,7465	2,5760	3,4128	0,9811
	S	17.957,52	18.674,70	21.051,49	11.024,04	3.014,32	-1,0842	2,9341	6,6676	
DC	O	5.090,41	5.215,82	6.023,80	3.438,41	649,17	-0,9347	3,4365	5,2205	0,8992
	S	4.881,71	5.071,33	5.619,31	3.438,41	609,64	-1,0289	3,1250	6,0207	
DD	O	4.251,20	4.238,22	6.244,78	2.226,50	1.176,94	-0,0207	2,1913	0,9289	0,9884
	S	4.267,16	4.062,29	6.204,72	2.301,58	1.194,11	0,0894	1,9248	1,6830	
DE	O	9.602,64	9.814,84	14.144,77	3.883,90	3.298,76	-0,3099	1,8248	2,5010	0,9938
	S	9.613,29	9.417,83	14.252,23	3.883,90	3.332,31	-0,1737	1,7714	2,3091	

Nota: O serie originaria S serie stimata

Quanto ai parametri di posizione l'errore percentuale, in termini assoluti, è sempre inferiore al 5 per cento. Relativamente ai parametri di variabilità, l'errore è inferiore all'1,5 per cento in due serie ed è compreso tra il 5 e l'8 per cento nelle restanti. In quattro casi l'algoritmo di disaggregazione *fuzzy* consente di ottenere una rappresentazione corretta delle serie incognite, pur in presenza di un *information set* particolarmente contenuto.

Il confronto tra correlogrammi delle serie Istat e di quelle stimate, non riportato per motivi di brevità, evidenzia l'omogeneità negli andamenti delle serie stimate e di quelle Istat fino al sesto *lag*.

L'ipotesi nulla di uguaglianza delle mediane e delle varianze tra serie stimate e quelle Istat è stata sottoposta a *test* non parametrici, i cui risultati sono presentati nella successiva Tavola 2, dove si riporta il valore della statistica *test* e il relativo *p-value*.

TAVOLA 2

Test di posizione e di omoschedasticità

	DA	DB	DC	DD	DE
Location test					
Wilcoxon/Mann-Whitney	0.656207 (0.5117)	0.374099 (0.7083)	1.529327 (0.1237)	0.122656 (0.9024)	0.196249 (0.8444)
Kruskal-Wallis	0.438694 (0.5078)	0.144577 (0.7038)	2.388446 (0.1222)	0.086586 (0.8975)	0.040958 (0.8396)
Variance equality test					
Brown-Forsythe	0.642348 (0.4257)	0.585898 (0.4467)	0.212477 (0.6463)	1.029388 (0.7454)	0.002155 (0.9631)

Dall'esame della Tavola 2 risulta confermato quanto si evidenziava dall'analisi descrittiva dei risultati riportata nella Tavola 1. L'ipotesi nulla di uguaglianza delle mediane è accettata in quattro confronti, mentre è precauzionalmente rifiutata con riferimento al settore DC (pelle e prodotti in pelle). L'ipotesi nulla di omoschedasticità invece è accettata in tutti i settori.

Va osservato che i risultati dei *test* migliorano notevolmente con riferimento alla sottosequenza 1970-1992, caratterizzata da profili ciclici meno accentuati in tutti i settori analizzati.

In conclusione i fattori che influenzano in maniera determinante l'adattamento tra serie stimata e serie incognita sono il profilo ciclico di quest'ultima, la sua variabilità e il contributo rispetto al totale aggregato.

Le migliori *performance* dello stimatore *naïf* si riscontrano in corrispondenza di serie dalla contenuta dinamica ciclica e non caratterizzate da forte variabilità. Soltanto in un caso – rappresentato dal settore DC, il cui contributo alla serie aggregata, tuttavia, diminuisce nell'intervallo di osservazione dall'11,6 al 7 per cento del totale – la tecnica di disaggregazione *fuzzy* genera una serie dal profilo ciclico più *smooth* di quello che caratterizza la serie Istat, pur ricostruendone correttamente la maggior parte dei punti di svolta.

Tale risultato, se è estremamente confortante, necessita tuttavia di ulteriori approfondimenti teorici riguardanti da un lato le proprietà degli stimatori *fuzzy*, dall'altro l'analisi delle distribuzioni asintotiche, quesiti che al momento non saranno affrontati e rappresenteranno l'oggetto di future ricerche.

Facoltà di Scienze Politiche ALESSANDRO POLLI
Università di Roma La Sapienza

RIFERIMENTI BIBLIOGRAFICI

G. CHOQUET (1954), *Theory of capacities*, Annales de l'Institut Fourier, Grenoble, 5, pp. 131-295.

G.C. CHOW, A. LIN (1971), *Best linear unbiased interpolation, distribution, and extrapolation of time series by related series*, "Review of Economics and Statistics", 53 (4), pp. 372-375.

B. DE FINETTI (1974), *Theory of probability*, Vol. I, London, Wiley.

J. FODOR, J.L. MARICHAL, M. ROUBENS (1995), *Characterization of the ordered weighted averaging operators*, "IEEE Transactions on Fuzzy Systems", 3, pp. 236-240.

M. GRABISCH (1994), *Characterization of fuzzy integrals viewed as aggregation operators*, in "Proceedings of 3rd IEEE Conference on Fuzzy Systems", Orlando FL, giugno 1994, pp. 1927-1932.

M. GRABISCH (1996), *The application of fuzzy integral in multicriteria decision making*, "European Journal of Operational Research", 89, pp. 445-456.

V. KRÄTSCHMER (2003), *Coherent lower previsions and Choquet integrals*, "Fuzzy sets and systems", 138, pp. 469-484.

P. WALLEY (1991), *Statistical reasoning with imprecise probabilities*, London, Chapman&Hall.

R.R. YAGER (1991), *Connectives and quantifiers in fuzzy sets*, "Fuzzy sets and systems", 40, pp. 39-75.

RIASSUNTO

Una tecnica di disaggregazione fuzzy

Il *paper* affronta un problema di disaggregazione di serie storiche in presenza di estesi *lack* informativi. In contesti di questo tipo è impossibile mantenere gli assunti che giustificano l'applicazione di metodologie tradizionali, quali quelle derivate dall'algoritmo di Chow e Lin, basate sulle consuete ipotesi di normalità.

In termini generali, in presenza di *information set* limitati, il riferimento a misure probabilistiche potrebbe rivelarsi inappropriato e appare opportuno adottare misure del grado di incertezza che soddisfino soltanto alcune proprietà generali, quali ad esempio le c.d. misure *fuzzy*.

Dopo un esame sommario delle proprietà dei principali operatori di aggregazione, è quindi presentata una procedura di disaggregazione di informazioni in presenza di *lack* informativi, basata sulla nozione di integrale di Choquet, caratterizzata da coerenza nel senso di De Finetti.

SUMMARY

A fuzzy disaggregation technique

The aim of this paper is to analyze a problem of time series disaggregation in presence of broad information lack. In this framework it is not possible to follow standard methodologies, like those stemming from the Chow and Lin algorithm and based on probabilistic assumptions. In general terms, when information sets are limited, instead of referring to probabilistic measures it could be more appropriate to adopt an uncertainty measure satisfying only some general properties, like the fuzzy one. After a synthetic survey about fuzzy aggregation operators, we introduce a fuzzy disaggregation technique, based on Choquet capacity theory and characterized by De Finetti coherence.