

## L'INTERPOLAZIONE RAZIONALE E LA SUA VALIDITÀ (\*)

M. Passaquindici, G. Barbieri

## 1. PREMESSA

Con questo metodo interpolatorio si ricerca una funzione  $f(x)=S(x)/T(x)$ , dove  $S(x)$  e  $T(x)$  sono due polinomi razionali in  $x$ , che passi per un numero dato di punti  $P_i(x_i, y_i)$ , con le ascisse diverse fra loro e comunque disposti.

Se tale funzione non fosse continua sull'intervallo che contiene le ascisse dei punti dati, si cercherà di renderla continua con metodi opportuni che saranno descritti in questo lavoro.

Questo metodo interpolatorio è utile unicamente per ottenere dati mancanti nella distribuzione dei punti dati o per correggere dati ritenuti anomali. La validità di tale metodo è stata testata su dati economici, clinici ed ambientali e ha dato risultati, come vedremo, positivi<sup>1</sup>.

## 2. LA FUNZIONE RAZIONALE INTERPOLANTE

Come è noto, dato un numero  $n$  qualsiasi di punti, comunque disposti,  $x_i \neq x_j$ ,  $i \neq j$ , è sempre possibile determinare una funzione razionale:

$$f_n(x) = S_n(x) / T_n(x)$$

che passa per i punti dati, dove  $S_n(x)$  e  $T_n(x)$  sono due polinomi razionali in  $x$  i cui gradi dipendono da  $n$ :

se  $n=2k$  sono rispettivamente di grado  $k$  e  $k-1$

se  $n=2k+1$  sono entrambi di grado  $k$ .

Indicando i punti con  $P_i(x_i, y_i)$ ,  $i=1, 2, \dots, n$ , e ponendo:

---

(\*) Lavoro eseguito con il contributo del C.N.R; i par. da 1 a 5 sono da attribuire a Maria Passaquindici, il par. 6 e l'appendice a Gianna Barbieri.

<sup>1</sup> I dati economici sono stati ripresi dalle quotazioni del dollaro del "SOLE-24 ORE". I dati clinici e quelli ambientali sono stati forniti rispettivamente dall'Istituto "Regina Elena" di Roma e dalle Regioni Campania, Lazio e Lombardia, che si ringraziano.

$$y_1(x_i) = y_i \quad ; \quad i=1,2,\dots,n \quad (1)$$

$$y_j(x_i) = \frac{x_i - x_{j-1}}{y_{j-1}(x_i) - y_{j-1}(x_{j-1})} \quad ; \quad j=2, \dots, n; \quad i=j, \dots, n$$

la funzione razionale interpolante, espressa come una funzione continua finita, è data da:

$$f_n(x) = y_1(x_1) + \frac{x - x_1}{y_2(x_2) + \frac{x - x_2}{y_3(x_3) + \frac{x - x_3}{\dots \dots \dots y_{n-1}(x_{n-1}) + \frac{x - x_{n-1}}{y_n(x_n)}}} \quad (2)$$

Sviluppando la (2), facendo riferimento allo sviluppo in frazione continua di un numero razionale, le ridotte  $A_i/B_i$  sono espresse da:

$$\begin{aligned} A_1 &= y_1(x_1) = y_1 \quad ; \quad B_1 = 1 \\ A_2 &= y_1 y_2(x_2) + (x - x_1) \quad ; \quad B_2 = y_2(x_2) \quad \quad i=3, \dots, n \\ A_i &= A_{i-1} y_i(x_i) + A_{i-2}(x - x_{i-1}) \\ B_i &= B_{i-1} y_i(x_i) + B_{i-2}(x - x_{i-1}) \end{aligned} \quad (3)$$

Indicando con  $S_l(i)$  il coefficiente di  $x^i$  nel polinomio  $A_l$  e con  $T_l(i)$  quello di  $x^i$  nel polinomio  $B_l$ , e ponendo:

$$S_l(-1) = T_l(-1) = 0 \quad ; \quad l=1,2,\dots,n-2$$

$$S_l(i) = 0 \quad ; \quad l=1,2,3,\dots,2 \left[ \frac{n}{2} \right] - 1 \quad ; \quad i = \left[ \frac{l}{2} \right] + 1$$

$$T_l(i) = 0 \quad ; \quad l=1,2,\dots,2 \left[ \frac{n-1}{2} \right] \quad ; \quad i = \left[ \frac{l+1}{2} \right]$$

si ha (Passaquindici, 1988):

$$S_l(i) = S_{l-1}(i)y_l(x_i) - S_{l-2}(i)x_{l-1} + S_{l-2}(i-1) \text{ con } l=3,4,\dots,n \text{ ed } i=0,1,\dots, \left[ \frac{l}{2} \right] \quad (4)$$

$$T_l(i) = T_{l-1}(i)y_l(x_i) - T_{l-2}(i)x_{l-1} + T_{l-2}(i-1) \text{ con } l=3,4,\dots,n \text{ ed } i=0,1,\dots, \left[ \frac{l-1}{2} \right] \quad (5)$$

Osserviamo che per calcolare gli  $y_j(x_i)$  deve essere, per le (1):

$$y_{j-1}(x_i) \neq y_{j-1}(x_{j-1}) \quad (6)$$

altrimenti si avrebbe qualche denominatore nullo.

Ora osserviamo che l'espressione della  $f_n(x)$  non dipende dall'ordine in cui sono considerati i punti  $P_i(x_i, y_i)$ . Per non ottenere dei denominatori nulli basta (Passaquindici, 1988) poter ordinare gli stessi punti in modo che i punti successivi ai primi  $i$  ( $i=2,3,\dots,n-1$ ) non appartengano alla curva determinata dagli stessi primi  $i$  punti. Per questo è necessario poter ordinare i punti in modo che l'ordinata del primo punto sia diversa da tutte le altre e che i punti successivi ai primi  $i$  non appartengano alla curva  $f_i(x) = A_i(x)/B_i(x)$  definita dai precedenti  $i$  punti. Se dunque per una certa coppia  $(i, j)$  non vale la (6), significa che  $P_j$  appartiene alla curva determinata dai primi  $i$  punti. Occorre allora modificare la scelta dei primi  $i$  punti, scambiando – per esempio – uno di loro con uno dei successivi, escluso l' $j^{\text{mo}}$ . Se non è possibile nessuno scambio, il punto  $P_j$  va eliminato insieme ai punti successivi ai primi  $i$ . Se  $y_{n-1}(x_{n-1}) = y_{n-1}(x_n)$  va eliminato solo il punto  $P_n$ . È evidente che se si aggiunge un altro punto agli  $n$  punti dati, i polinomi  $S_{n+1}(x)$  e  $T_{n+1}(x)$  si ricavano da quelli di  $S_{n-1}(x)$ ,  $S_n(x)$ ,  $T_{n-1}(x)$  e  $T_n(x)$  in base alle formule ricorrenti (4) e (5).

### 3. RICERCA DELLA FUNZIONE INTERPOLANTE CONTINUA

La funzione  $f_n(x)$  non è continua sull'intervallo  $(x_1, x_n)$  negli eventuali punti in cui  $T_n(x)$  si annulla su  $(x_1, x_n)$ .

Vediamo se è possibile ottenere in questo caso una funzione continua che passi sempre per gli  $n$  punti dati.

Supponiamo dunque che  $T_n(x)$  si annulli almeno una volta su  $(x_1, x_n)$ . Consideriamo un  $(n+1)^{\text{mo}}$  punto  $P_{n+1}(x_{n+1}, y_{n+1})$ , per altro generico, e ponendo  $y_1(x_{n+1}) = y_{n+1}$ , calcoliamo:

$$y_i(x_{n+1}) \quad ; \quad i=2,3,\dots,n+1$$

È possibile determinare un  $y_{n+1}(x_{n+1})$  in modo che  $T_{n+1}(x)$  non si annulli su  $(x_1, x_n)$ ?

Ricordiamo all'uopo un teorema di Boudan-Fourier, che afferma: data una funzione  $f(x)$  razionale di grado  $n$ , siano  $a$  e  $b$  due valori reali, con  $b > a$ , tali che non annullino  $f^{(1)}(x)$ ,  $f^{(2)}(x)$ , ...,  $f^{(n)}(x)$ . Il numero delle variazioni di:

$$f^{(0)}(a)=f(a), f^{(1)}(a), \dots, f^{(n)}(a)$$

diminuito del numero di variazioni di:

$$f^{(0)}(b)=f(b), f^{(1)}(b), \dots, f^{(n)}(b)$$

dà il numero degli zeri di tale funzione in  $(a, b)$  o lo supera di un numero pari.

Calcoliamo dunque:

$$T_{n+1}^{(0)}(x_1)=T_{n+1}(x_1), T_{n+1}^{(1)}(x_1), \dots, T_{n+1}^{(i)}(x_1), \dots$$

$$T_{n+1}^{(0)}(x_n)=T_{n+1}(x_n), T_{n+1}^{(1)}(x_n), \dots, T_{n+1}^{(i)}(x_n), \dots$$

e consideriamo, per ognuno di essi, il valore di  $y_{n+1}(x_{n+1})$  che l'annulla. Riportando tali valori in ordine crescente, otterremo un certo numero di intervalli, nei quali considereremo i segni di  $T_{n+1}^{(i)}(x_1)$  e  $T_{n+1}^{(i)}(x_n)$ , per  $i=0, 1, 2, \dots$ . Negli intervalli in cui la differenza del numero delle variazioni è zero esistono valori di  $y_{n+1}(x_{n+1})$  validi per la nostra ricerca. Tra gli  $y_{n+1}(x_{n+1})$  validi, poi, va scelto quello per il quale la  $S_{n+1}(x)/T_{n+1}(x)$  approssima meglio la  $S_n(x)/T_n(x)$ , per esempio, col metodo dei minimi quadrati.

Per la ricerca di  $T_{n+1}(x)$  che non si annulli su  $(x_1, x_n)$  si può ricorrere anche a un metodo numerico-grafico.

Supponiamo sempre dunque che  $T_n(x)$  abbia almeno uno zero su  $(x_1, x_n)$ . Sia, al solito, un  $(n+1)^{mo}$  punto  $P_{n+1}(x_{n+1}, y_{n+1})$ , per altro qualsiasi. Si ha:

$$T_{n+1}(x)=T_n(x) y_{n+1}(x_{n+1}) + T_{n-1}(x) (x-x_n)=T_n(x) y_{n+1}(x_{n+1}) - N(x)$$

Ora  $T_{n+1}(x)$  non ha zeri in  $(x_1, x_n)$  se e solo se non è mai:

$$T_n(x) y_{n+1}(x_{n+1}) - N(x)=0$$

Disegnando il grafico

$$y_{n+1}(x_{n+1})=N(x)/T_n(x)$$

si può concludere che se sull'asse delle ordinate esiste un intervallo per il quale nessun punto di  $N(x)/T_n(x)$  assume una di quelle ordinate, ogni  $y_{n+1}(x_{n+1})$  di tale intervallo rende  $S_{n+1}(x)/T_{n+1}(x)$  continua e passante per gli  $n$  punti dati. Anche ora va scelto l' $y_{n+1}(x_{n+1})$  più opportuno.

Per le applicazioni che verranno effettuate indichiamo con  $F(x)=S(x)/T(x)$  la funzione interpolante degli  $n$  punti dati, ed eventualmente passante per un  $(n+1)^{mo}$  necessario per la continuità della funzione (se sussistessero discontinuità).

#### 4. CASO DI NON ESISTENZA DELLA FUNZIONE RAZIONALE CONTINUA INTERPOLANTE

Facendo riferimento al teorema di Boudan-Fourier per la ricerca della funzione continua che interpola gli  $n$  punti dati, osserviamo che se esistono solo intervalli in cui la differenza del numero delle variazioni è dispari non è possibile trovare la funzione razionale interpolante continua su  $(x_1, x_n)$  passante per gli  $n$  punti

$P_i(x_i, y_i)$ . Se esistono intervalli in cui tale differenza è zero, è possibile che esistano in essi degli  $y_{n+1}(x_{n+1})$  validi. Qui, comunque, va fatta una ricerca punto per punto.

Nel caso del metodo numerico-grafico, se sulle ordinate tutti i valori sono assunti da almeno un punto di  $N(x)/T_n(x)$ , la curva per gli  $n$  punti dati è giocoforza discontinua.

Ci si può chiedere perché talvolta non sia possibile trovare una curva siffatta. Indicando con  $x^0$  uno degli zeri di  $T_n(x)$  su  $(x_1, x_n)$ , si può osservare che, per qualsiasi valore di  $y_{n+1}(x_{n+1})$ ,  $T_{n+1}(x)$  passa sempre per il punto:

$$P^0(x^0, -N(x^0))$$

Se i valori  $N(x^0)$  hanno segni diversi è chiaro che  $T_{n+1}(x)$  si deve annullare in  $(x_1, x_n)$ . Se invece  $N(x^0)$  ha sempre lo stesso segno, può darsi che  $T_{n+1}(x)$  non si annulli. Solo in tal caso può esistere la funzione razionale continua interpolante.

## 5. SULLA SUA VALIDITÀ

Per esaminare la validità di tale metodo interpolatorio, occorre distinguere due casi:

- 1 - non esistenza della funzione razionale continua interpolante;
- 2 - esistenza di tale funzione.

Nel primo caso si considera l'intervallo  $(x_1, x_n)$  ad esclusione di opportuni interni degli zero di  $T_n(x)$ . Nel secondo caso si prende in considerazione tutto l'intervallo.

Sia ora  $i=2, 3, \dots, n-1$ . Si considerino le funzioni  $S_{n-1}(x)^{(i)}/T_{n-1}(x)^{(i)}$  che passano per gli  $n-1$  punti  $P_1(x_1, y_1), P_2(x_2, y_2), \dots, P_{i-1}(x_{i-1}, y_{i-1}), P_{i+1}(x_{i+1}, y_{i+1}), \dots, P_n(x_n, y_n)$ . Tali curve, in numero di  $n-2$ , vengono via via confrontate con la curva originaria. Anche queste  $n-2$  curve, se non sono già continue, si cerca – se possibile – di renderle continue con uno dei due metodi sopra esposti.

Se ciò non fosse possibile, il confronto va fatto su  $(x_1, x_n)$  escludendo opportuni interni degli zeri di  $T_{n-1}(x)^{(i)}$ .

La valutazione degli scarti fra le ordinate del punto reale e quelle del punto stimato ci darà la risposta sulla validità del metodo.

## 6. VERIFICA DEL METODO

Nelle analisi di dati in serie storica sono frequenti i casi in cui la serie in esame presenta dati mancanti e/o outliers. Queste anomalie, riscontrabili nell'intervallo per i quali sono noti i valori della serie, vengono in generale rimosse e corrette ricorrendo a procedure di interpolazione la cui finalità è quella di dare una regolarità alla sequenza dei dati.

La valutazione del metodo interpolatorio presentato in questo lavoro è stata effettuata utilizzando dati relativi a diversi campi di applicazione. In un primo caso l'applicazione del metodo ha riguardato dati di natura economica. Sono state

considerate le rilevazioni del dollaro nell'anno 1990 effettuando raggruppamenti di 7 osservazioni, ciascuno di essi denominato *settupla*. In un secondo caso l'esperimento è stato effettuato con dati medici. Si è trattato di analizzare l'andamento del numero dei globuli rossi di un certo numero di pazienti del Regina Elena di Roma sottoposti ad una determinata terapia. Per ciascun paziente è stato possibile effettuare un diverso numero di rilevazioni nell'arco di un periodo la cui unità di tempo è rappresentata dal mese. Infine, un terzo esperimento ha riguardato l'analisi del metodo relativamente a dati ambientali. In questo caso sono stati esaminati i dati relativi alla presenza di ossido di carbonio rilevata in tre centraline di Milano nell'anno 1997, di Roma e di Napoli nell'anno 1998. Per tale verifica l'intervallo temporale scelto è stato tra le ore 5 e le ore 10. Le osservazioni rilevate a frequenza oraria sono state considerate a gruppi di 6.

La validità del metodo è stata verificata eliminando di volta in volta un punto interno all'intervallo dei dati (esclusi quindi gli estremi). La verifica del metodo di interpolazione è stata quindi effettuata esaminando per ciascun esperimento tutti i gruppi di osservazioni di  $n-1$  punti che si ottengono eliminando via via il 2°, il 3°, il 4°, ..., l' $(n-1)^{mo}$  punto. Per ciascuno dei tre esperimenti sono stati ricavati i polinomi  $S(x)$  e  $T(x)$  e si è rappresentata graficamente la funzione  $F(x)=S(x)/T(x)$ . Per ciascuno dei gruppi di  $n-1$  osservazioni viene determinata la  $F_i(x)$ ,  $i=2,3,\dots,n-1$  che è possibile mettere a confronto, attraverso opportuni metodi statistici, con la  $F(x)$ .

Dall'analisi dei risultati relativi ai tre diversi set di dati è stato possibile dedurre uno schema che sintetizza le possibili situazioni che in generale si riscontrano nell'applicazione del metodo.

Questo si deduce considerando le varie combinazioni ottenibili tra le caratteristiche della curva originaria  $F(x)$  e la  $F_i(x)$  ottenuta escludendo un punto. Di ciascun caso riportato nello schema ne viene data un'illustrazione attraverso i risultati dei tre esperimenti.

TABELLA 1  
*Curva originale*

	Continua	Con discontinuità eliminabili	Con discontinuità non eliminabili
<i>Curva escludendo un punto</i>	Continua	Continua	Continua
	Con discontinuità eliminabili	Con discontinuità eliminabili	Con discontinuità eliminabili
	Con discontinuità non eliminabili (caso impossibile)	Con discontinuità non eliminabili	Con discontinuità non eliminabili

La funzione razionale che interpola gli  $n$  punti dati,  $F(x)$ , può risultare 1) continua; 2) discontinua, con discontinuità eliminabili; 3) discontinua, con discontinuità non eliminabili. L'eliminazione di un punto dalla  $F(x)$  di partenza può dare origine a diverse  $F_i(x)$  che vengono illustrate in tutti i casi verificabili negli esempi qui di seguito. In appendice vengono riportati i grafici delle curve ottenute per le quali vengono riportati i coefficienti dei polinomi  $S(x)$  e  $T(x)$  rispettivamente della curva  $F(x)$  e della  $F_i(x)$ . Per una corretta lettura di tali grafici è utile considerare la legenda qui di seguito riportata:

1. curva passante per gli  $n$  punti dati (continua o discontinua): linea continua;
2. curva continua passante per gli  $n$  punti dati e un altro punto opportunamente scelto nel caso la curva di cui al punto 1 sia discontinua: linea continua in grassetto;
3. curva passante per gli  $n-1$  punti dati (continua o discontinua): linea tratteggiata;
4. curva continua passante per gli  $n-1$  punti dati e un altro punto opportunamente scelto nel caso la curva di cui al punto 3 sia discontinua: linea a crocette.

Nella prima situazione di  $F(x)$  continua sono stati considerati due esperimenti in cui l'eliminazione di un punto può dare origine ad una funzione  $F_i(x)$  a) continua o b) discontinua con punti di discontinuità eliminabili.

Il primo di questi riguarda la rilevazione del dollaro nel periodo 11/2/90–19/2/90 (5<sup>a</sup> *settopla*):

TABELLA 2

$x_i$	Valore \$	Data
0	1112,95	11/2/90
1	1119,40	12/2/90
2	1118,85	13/2/90
3	1115,75	14/2/90
4	1116,25	15/2/90
7	1130,90	18/2/90
8	1136,20	19/2/90

L'eliminazione del 4° punto dà origine ad una  $F_4(x)$  anch'essa continua (Fig. 1).

Il secondo esperimento è riferito al campo medico. Si dispone delle seguenti rilevazioni relative al paziente X:

TABELLA 3

$x_i$	Valore globuli rossi
0,0	4,20
0,4	3,59
1,0	4,27
1,4	5,60
2,2	4,70
2,6	4,30

Eliminando il 2° punto si ottiene una  $F_2(x)$  che presenta due discontinuità in  $x_1=1,448$  e  $x_2=2,044$  entrambe appartenenti all'intervallo di interpolazione (0, 2,60), ma comunque eliminabili (Fig. 2).

Nella situazione di  $F(x)$  discontinua con punti di discontinuità eliminabili la  $F_i(x)$  può risultare a) continua, b) discontinua con punti di discontinuità eliminabili o anche c) discontinua con punti di discontinuità non eliminabili.

L'esempio di  $F_i(x)$  continua è relativo alla rilevazione della centralina MI<sub>1</sub> 17-1-97. L'interpolazione dei punti:

TABELLA 4

Tempo	Valore di CO2
5	9,2
6	6,2
7	5,4
8	7,7
9	11,8
10	18,6

dà origine ad una  $F(x)$  discontinua nei punti  $x_1=7,47$  e  $x_2=112,03$  (esterno all'intervallo (5, 10)). La discontinuità in  $x_1$  viene eliminata. Escludendo il 5° punto la curva  $F_5(x)$  che si ottiene risulta come anticipato continua (Fig. 3).

Il secondo esempio è anch'esso riferito a dati ambientali relativi alla centralina MI<sub>3</sub> 26-3-97:

TABELLA 5

Tempo	Valore di CO2
5	2,9
6	2,8
7	3,0
8	3,4
9	3,8
10	4,5

La curva  $F(x)$  passante per tali punti presenta due discontinuità in  $x_1=8,83$  e  $x_2=-4,43$  (la seconda esterna all'intervallo di interesse) eliminabili. Dall'esclusione del terzo punto si ottiene una curva  $F_3(x)$  con due discontinuità in  $x_1=5,8$  e  $x_2=12,61$  (la seconda esterna all'intervallo) che sono anch'esse eliminabili (Fig. 4).

Il terzo esempio riguarda la 4<sup>a</sup> *settopla* (periodo 29/1/90–8/2/90) di rilevazione del dollaro con le seguenti osservazioni:

TABELLA 6

$x_i$	Valore \$	Data
0	1099,20	29/1/90
1	1095,50	30/1/90
2	1094,10	31/1/90
3	1097,75	1/2/90
6	1094,90	4/2/90
7	1094,50	5/2/90
10	1107,40	8/2/90

Come nei due esempi precedenti la  $F(x)$  presenta due discontinuità eliminabili mentre la curva ottenuta eliminando il 2° punto presenta due punti di discontinuità in  $x_1=2,2$  e  $x_2=10,6$ , essendo il primo dei due interno all'intervallo di interpolazione. In questo caso la discontinuità non risulta eliminabile (Fig. 5).

Infine, si passa ad analizzare la situazione della  $F(x)$  discontinua con discontinuità non eliminabili che può dare origine ad una  $F_i(x)$  a) continua, b) discontinua con punti di discontinuità eliminabili o anche c) discontinua con punti di discontinuità non eliminabili.

Per illustrare il caso a) si consideri le rilevazioni relative alla centralina MI<sub>1</sub> 4-1-97:

TABELLA 7

Tempo	Valore di CO2
5	1,0
6	1,4
7	1,7
8	3,0
9	3,8
10	2,2

La  $F(x)$  presenta due discontinuità ( $x_1=7,64$  e  $x_2=9,44$ ) che non è possibile eliminare. Modificando la curva si ottiene una nuova  $F(x)$  che ha tre discontinuità in  $x_1=3,64$  e  $x_2=8,53$  (doppia); eliminando il 2° punto si ottiene invece una  $F_2(x)$  continua (Fig. 6).

Nel caso b) l'esame della centralina MI<sub>3</sub> 10-10-97 è relativo ai seguenti punti:

TABELLA 8

Tempo	Valore di CO2
5	2,1
6	2,0
7	2,2
8	2,5
9	3,5
10	5,6

La  $F(x)$  interpolante ha due discontinuità in  $x_1=7,12$  e  $x_2=13,12$  (la seconda esterna all'intervallo) che non possono essere eliminate mentre la  $F_5(x)$  presenta due discontinuità di cui una interna all'intervallo di interpolazione che è possibile eliminare (Fig. 7).

L'ultimo esempio relativo al caso c) riguarda le seguenti osservazioni riferite al paziente Y:

TABELLA 9

$x_i$	Valore globuli rossi
0,0	3,90
0,4	3,45
1,0	3,80
1,2	3,66
1,6	4,23
2,2	3,95

Il metodo interpolatorio dà luogo a una  $F(x)$  discontinua che non è possibile rendere continua. Le discontinuità sono in  $x_1=0,42$  e  $x_2=1,42$ . Eliminando il 5° punto si ottiene una  $F_5(x)$  che presenta due discontinuità non eliminabili (Fig. 8).

Con questa serie di esempi è stato possibile illustrare il quadro completo delle situazioni che si sono realizzate in ciascun esperimento. A questo punto, va sottolineato che la scelta di un metodo di interpolazione non è scindibile dalla garanzia che può offrire in termini di errore che si commette sostituendo alla curva originale la curva ottenuta eliminando l' $i^{\text{mo}}$  punto. A tal fine, per analizzare l'affidabilità del metodo utilizzato in questo studio si è cercato di effettuare una valutazione globale utilizzando due statistiche. La prima è ottenuta considerando gli scarti assoluti tra punto reale e punto stimato mentre la seconda si riferisce agli scarti relativi percentuali sul totale dei dati.

Andando a considerare i risultati nel caso della rilevazione del dollaro si evince che solo nel 7% dei casi lo scarto tra il punto e la sua stima è maggiore di 1 mentre nella maggior parte dei casi (65,5%) lo scarto risulta compreso tra 0,01 e 0,1. Per quanto riguarda gli scarti relativi percentuali sulla stima, nell'11,03% dei casi l'errore è al di sotto dello 0,1%, nel 68,3% dei casi è compreso tra lo 0,1% e l'1% e nel 20,7% dei casi è superiore all'1%.

TABELLA 10

*Scarto tra il punto reale e la sua stima*

Scarto	Casi (val. cumulati)
<0,01	8,97 (8,97%)
0,01-0,1	65,52 (74,49%)
0,1-1	24,83 (99,32%)
>1	0,69 (100%)

TABELLA 11

*Errore percentuale nella stima del punto reale*

Errore %	Casi (val. cumulati)
<0,1	11,03 (11,03%)
0,1-0,5	46,90 (57,93%)
0,5-1	21,38 (79,31%)
1-5	17,24 (96,55%)
>5	3,44 (100%)

I risultati ottenuti relativamente ai pazienti esaminati sono anch'essi soddisfacenti. Lo scarto tra il punto e la sua stima è per il 69% dei casi compreso tra 0,1 e 1 mentre per gli scarti relativi percentuali sulla stima l'errore che si commette nel 97% dei casi è inferiore allo 0,5%.

TABELLA 12

*Scarto tra il punto reale e la sua stima*

Scarto	Casi (val. cumulati)
<0,05	17 (17,00%)
0,05-0,1	5 (22%)
0,1-1	69 (91%)
>1	9 (100%)

TABELLA 13

*Errore percentuale nella stima del punto reale*

Errore %	Casi (val. cumulati)
<0,1	60 (60%)
0,1-0,5	37 (97%)
0,5-1	2 (99%)
>1	1 (100%)

Analogamente, i risultati ottenuti in riferimento alle rilevazioni di  $\text{CO}_2$  in termini di scarti assoluti tra punto reale e stima sono abbastanza soddisfacenti, risultando pari al 67,56% il numero di casi in cui lo scarto è inferiore allo 0,1. È invece solo pari all'8,8% il numero delle volte in cui lo scarto è superiore a 1. L'analisi degli scarti relativi percentuali mostra che nel 10,03% dei casi si commette un errore inferiore allo 0,1%, mentre risultano pari al 72,23% i casi in cui

si effettua un errore compreso tra lo 0,1% e l'1%. In media l'errore di stima risulta generalmente al di sotto del 5% (92,5%).

TABELLA 14

*Scarto tra il punto reale e la sua stima*

Scarto	Casi (val. cumulati)
<0,01	7,05 (7,05%)
0,01-0,1	60,51 (67,56%)
0,1-1	23,62 (91,18%)
>1	8,82 (100%)

TABELLA 15

*Errore percentuale nella stima del punto reale*

Errore %	Casi (val. cumulati)
<0,1	10,03 (10,03%)
0,1-0,5	48,05 (58,08%)
0,5-1	24,18 (82,26%)
1-5	10,24 (92,50%)
>5	7,50 (100%)

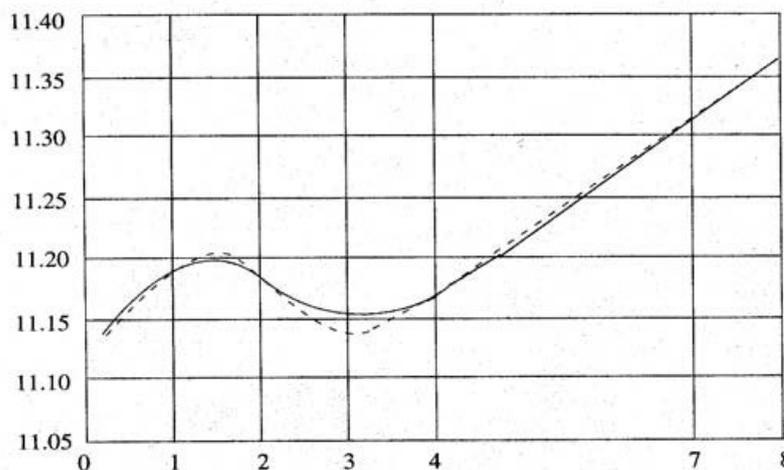
*Dipartimento di Statistica, Probabilità e Statistiche Applicate,  
Università degli Studi di Roma "La Sapienza".*

*Ministero del Lavoro e delle Politiche Sociali.*

MARIA PASSAQUINDICI

GIANNA BARBIERI

## APPENDICE



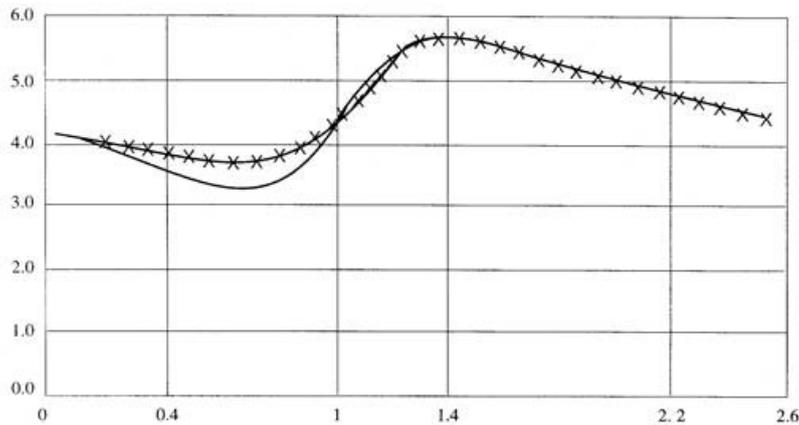
Coefficienti curva 7 p.ti  
 $S(3)=11.663$   
 $S(2)=-54.843$   
 $S(1)=116.279$   
 $S(0)=705.254$

$T(3)=1$   
 $T(2)=-4.573$   
 $T(1)=9.738$   
 $T(0)=63.368$

Coefficienti curva 6 p.ti  
 $S(3)=1$   
 $S(2)=224.218$   
 $S(1)=-904.891$   
 $S(0)=1568.969$   
 $T(2)=20.719$   
 $T(1)=-82.249$   
 $T(0)=140.974$

*Figura 1 – Settopla 5<sup>a</sup> di rilevazione del dollaro*

La curva in linea continua passa per i 7 punti dati, quella tratteggiata passa per gli stessi punti escluso il quarto. Entrambe non presentano discontinuità.



Coefficienti curva 6 p.ti

$S(3)=1$   
 $S(2)=-10.889$   
 $S(1)=17.108$   
 $S(0)=-8.369$

$T(2)=-1.7112$   
 $T(1)=3.435$   
 $T(0)=-1.993$

Coefficienti curva 5 pt.i+1

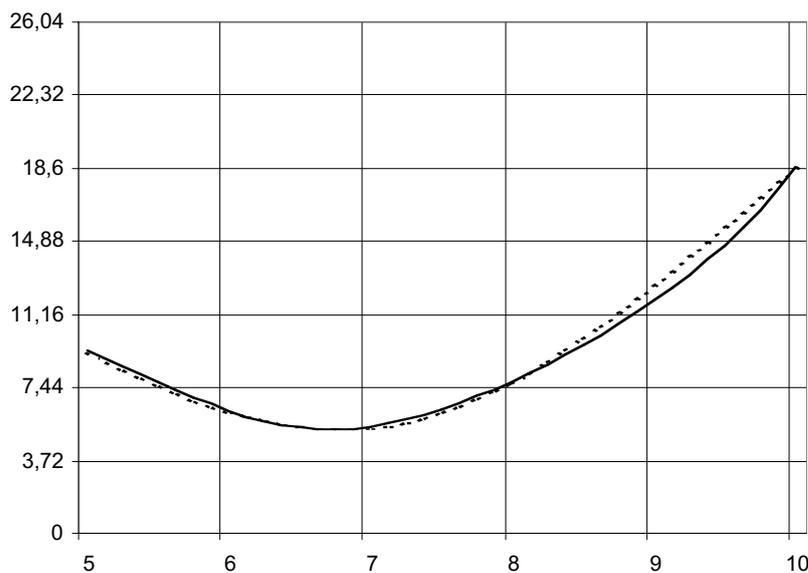
$S(3)=1$   
 $S(2)=-12.260$   
 $S(1)=21.886$   
 $S(0)=-12.428$

$T(2)=-2.039$   
 $T(1)=4.576$   
 $T(0)=-2.959$

Figura 2 – Paziente X

La curva in linea continua passa per i 6 punti dati e non presenta discontinuità, quella a crocette passa per gli stessi punti escluso il secondo.

Questa seconda curva è stata ottenuta eliminando le discontinuità citate nel testo.



Coefficienti curva 6 p.ti+1

$S(3) = 6.637$   
 $S(2) = -209.315$   
 $S(1) = 1905.174$   
 $S(0) = -5520.950$

$T(3) = 1$   
 $T(2) = -23.984$   
 $T(1) = 185.210$   
 $T(0) = -494.742$

Coefficienti curva 5 p.ti

$S(2) = 40.493$   
 $S(1) = -558.295$   
 $S(0) = 2044.075$

$T(2) = 1$   
 $T(1) = -15.271$   
 $T(0) = 80.150$

Figura 3 – Rilevazione di  $CO_2$  il 17-1-97 in  $MI_1$

La curva in linea continua passa per i 6 punti dati e non ha discontinuità, quella tratteggiata passa per gli stessi punti escluso il quinto ed è continua.

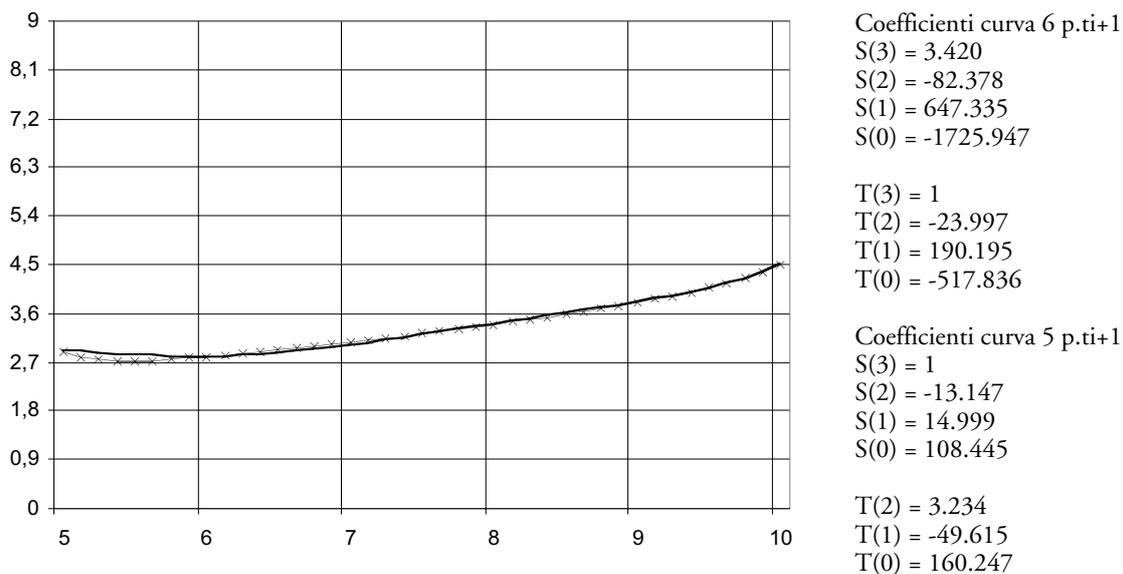


Figura 4 – Rilevazione di CO<sub>2</sub> il 26-3-97 in MI<sub>3</sub>

La curva a linea continua passa per i 6 punti e non presenta discontinuità, la curva a crocette passa per gli stessi punti escluso il terzo e non presenta discontinuità.

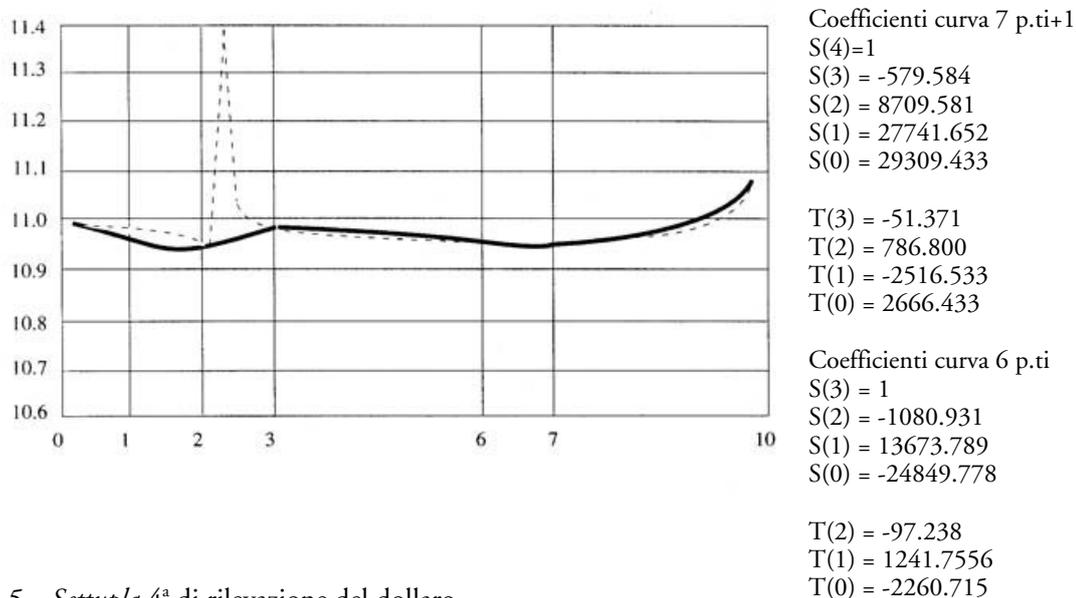


Figura 5 – Settopla 4<sup>a</sup> di rilevazione del dollaro

La curva in linea continua passa per i 7 punti dati e non presenta discontinuità, quella tratteggiata passa per gli stessi punti escluso il secondo. Le discontinuità non possono essere eliminate.

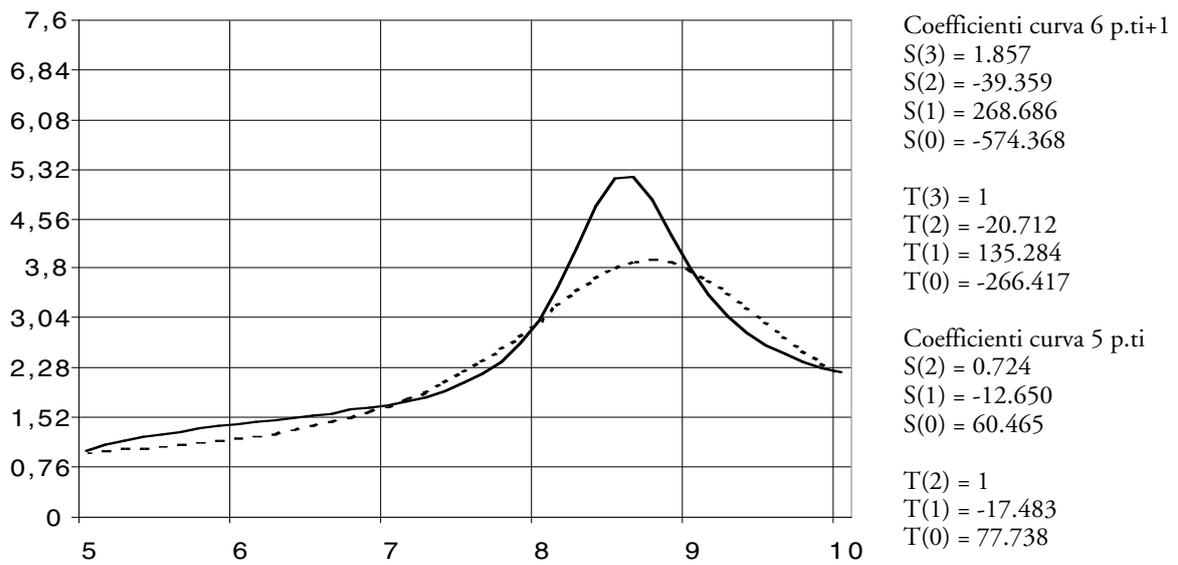


Figura 6 – Rilevazione di CO<sub>2</sub> il 4-1-97 in MI<sub>1</sub>

La curva in linea continua passa per i 6 punti dati ma presenta discontinuità non eliminabili.

La curva tratteggiata passa per gli stessi punti escluso il secondo e non presenta discontinuità.

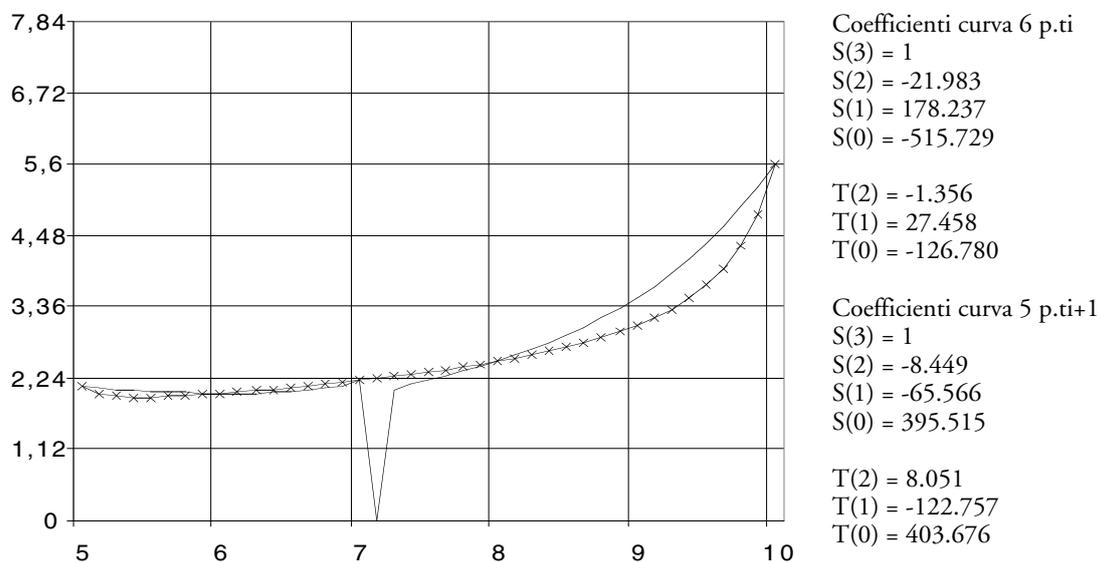
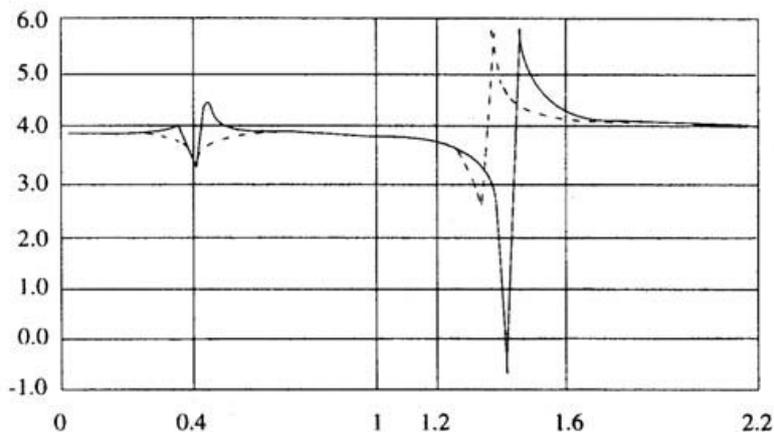


Figura 7 – Rilevazione di CO<sub>2</sub> il 10-10-97 in MI<sub>3</sub>

La curva in linea continua passa per i 6 punti dati, con discontinuità non eliminabili; quella a crocette passa per gli stessi punti escluso il quinto ed è continua, avendo eliminato le discontinuità.



Coefficienti curva 6 p.ti

$S(3)=1$   
 $S(2)=-92.337$   
 $S(1)=165.853$   
 $S(0)=-53.501$

$T(2)=-22.820$   
 $T(1)=42.069$   
 $T(0)=-13.718$

Coefficienti curva 5 pt.i

$S(2)=3.912$   
 $S(1)=-6.742$   
 $S(0)=2.028$

$T(2)=1$   
 $T(1)=-7.255$   
 $T(0)=7.024$

Figura 8 – Paziente Y

La curva in linea continua passa per i 6 punti dati e presenta discontinuità non eliminabili; quella tratteggiata passa per gli stessi punti escluso il quinto. Questa seconda curva ha due discontinuità nei punti:

$$x_1 = 0.387 \quad x_2 = 1.344$$

Tali discontinuità non possono essere eliminate.

#### RIFERIMENTI BIBLIOGRAFICI

- G. BARBIERI, N.M. LO BELLO (2000), *L'interpolazione razionale. Verifica della sua validità statistica relativamente a dati ambientali*, a cura di M. Passaquindici, Quaderni n. 9, Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università degli Studi di Roma "La Sapienza".
- N. M. BESKIN (1986), *Fascinating fractions*, MIR PUBLISHER Moscow, english translation.
- M. CUGIANI (1967), *Metodi numerico statistici*, UTET.
- M. PASSAQUINDICI (1988), *L'interpolazione per mezzo della frazione continua*, "Metron" vol. XLVI, n. 1-4, pp. 175-194.
- M. PASSAQUINDICI (1993), *Interpolazione per mezzo di una frazione razionale continua*, "Metron", vol. LI, n. 1-2, pag. 98-118.
- M. PASSAQUINDICI, N.M. LO BELLO, F. BILLI (1996), *L'interpolazione razionale e sua validità statistica*, a cura di M. Passaquindici, Quaderni n. 7, Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università degli Studi di Roma "La Sapienza".

#### RIASSUNTO

##### *L'interpolazione razionale e la sua validità*

L'interpolazione razionale può essere applicata per correggere valori anomali o per aggiungere valori mancanti in una distribuzione di  $n$  punti. Con questo metodo è possibile

determinare una funzione passante per  $n$  punti dati, comunque disposti. Tale funzione è data da  $F(x)=S(x)/T(x)$ , dove  $S(x)$  e  $T(x)$  sono due polinomi razionali in  $x$  i cui gradi dipendono da  $n$ . In questo lavoro presentiamo il metodo di interpolazione e valutiamo la sua validità statistica con riferimento a dati economici, medici e ambientali. I risultati mostrano che la discrepanza tra il valore reale e la sua stima ottenuta applicando l'interpolazione razionale è inferiore di 0,1 nel 74,5% e 67,6% dei casi utilizzando rispettivamente dati economici e ambientali, mentre nel caso dei dati medici risulta pari al 91% il numero dei casi con scarti inferiori a 1.

#### SUMMARY

##### *On the validity of rational interpolation*

The rational interpolation can be used to correct anomalous values or to add missing values in a distribution of  $n$  points. Applying this method it is possible to identify a function passing by any  $n$  given points. The function is given by  $F(x)=S(x)/T(x)$  where  $S(x)$  and  $T(x)$  are two rational polynomials in  $x$  whose degrees depend on  $n$ . In this paper we present an interpolating method and we evaluate the statistical validity of the method here developed with specific application to economical, clinical and environmental data. The results show that the discrepancy between the actual value and the estimate of the same value obtained applying the rational interpolation is lower than 0,1 in 74,5% and 67,6% of cases using respectively economical and environmental data, while using clinical data the discrepancy is lower than 1 in 91% of cases.