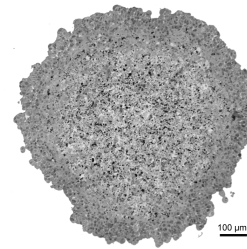


MatematicaMente



Publicazione mensile della sezione veronese della MATHESIS – Società Italiana di Scienze Matematiche e Fisiche – Fondata nel 1895 – Autorizzazione del Tribunale di Verona n. 1360 del 15 – 03 – 1999 – I diritti d'autore sono riservati. Direttore: Luciano Corso - Redazione: Luciano Corso, Elisabetta Capotosto, Arnaldo Vicentini - Via IV Novembre, 11/b – 37126 Verona – tel e fax (045) 8344785 – 338 6416432 e-mail: lcorso@iol.it – Stampa in proprio - Numero 102 – aprile 2006

Un'applicazione dell'equazione differenziale di Benjamin Gompertz

di Luciano Corso, Fabio Castelli, Gianfranco Pezzo

Nota: Il presente lavoro fa parte del Progetto Lauree Scientifiche del MIUR. È stato promosso dall'Università degli Studi di Verona, Facoltà Scienze. L'azione didattica ha coinvolto 18 studenti del 5° anno di Informatica dell'ITIS G. Marconi di Verona. Hanno collaborato alla realizzazione del progetto, oltre agli autori del presente articolo (docenti presso l'ITIS G. Marconi VR), i ricercatori Roberto Chignola (biologo) e Simone Zuccher (esperto di metodi numerici) dell'Università degli Studi di Verona. A questa breve presentazione strettamente matematico-statistica, seguirà il contributo di Zuccher. Il principio ispiratore dell'argomento è di natura biologica: (modello di crescita delle cellule di un Glioblastoma) e i dati su cui gli studenti hanno lavorato sono stati forniti da Roberto Chignola. Per una dettagliata presentazione di tutto il lavoro, anche informatico, si chiedi alla scuola la corrispondente pubblicazione della collana "Quaderni del Marconi" (VR).

1. Il sistema di equazioni differenziali di Gompertz

L'equazione differenziale di Gompertz (3) è la sintesi del seguente sistema di equazioni differenziali [B.5]:

$$\begin{cases} \frac{dX(t)}{dt} = Kp(t) \cdot X(t) \\ \frac{dKp(t)}{dt} = -\beta \cdot Kp(t) \end{cases} \quad (1)$$

Il sistema di equazioni differenziali (1) giustifica in sé, e forse meglio dell'equazione (3) (usualmente riportata nei testi come equazione differenziale di Gompertz), la descrizione del problema. La prima delle equazioni (1) dichiara che la variazione di una popolazione di individui al variare del tempo t dipende dal prodotto del moltiplicatore $Kp(t)$, che rappresenta il fattore di duplicazione della popolazione di individui $X(t)$ o, più in generale, il tasso di natalità, per la quantità di individui della popolazione $X(t)$. Il fattore $Kp(t)$, se il modello fosse di tipo esponenziale sarebbe una costante $1/T$, ove T è il tempo di duplicazione della popolazione; questo modello esponenziale, però, non descrive bene la crescita di una popolazione di individui, in quanto si nota sperimentalmente che all'aumentare degli individui del territorio di riferimento, diminuisce la crescita della popolazione al variare del tempo. Per tener conto di ciò si deve considerare $Kp(t)$ una funzione del tempo, non una costante. La seconda equazione di (1) tiene conto di questo effetto di smorzamento del fattore di crescita di $X(t)$. Tale smorzamento è dato dal parametro β che al variare del tempo agisce sul tasso di natalità $Kp(t)$ con effetto di farlo decrescere. Infatti se $\beta > 0$, allora $dKp(t)/dt$ è negativa al variare di t . Lo smorzamento di $Kp(t)$ si fa sentire sulla variazione della popolazione: la popolazione si incrementa sempre meno.

Con alcuni passaggi da (1) si passa a (2); Vediamoli:

$$\begin{cases} \frac{dX(t)}{dt} = Kp(t) \cdot X(t) \\ \frac{dKp(t)}{dt} = -\beta \cdot \frac{1}{X(t)} \cdot \frac{dX(t)}{dt} \end{cases} \quad \begin{cases} \frac{dX(t)}{dt} = Kp(t) \cdot X(t) \\ dKp(t) = -\beta \cdot \frac{1}{X(t)} \cdot dX(t) \end{cases}$$

$$\begin{cases} \frac{dX(t)}{dt} = -\beta \cdot X(t) \cdot (\ln X(t) + C) \\ Kp(t) = -\beta \cdot (\ln |X(t)| + C) \end{cases}$$

Se $C = -\ln K$ si ottiene

$$\begin{cases} \frac{dX(t)}{dt} = -\beta \cdot X(t) \cdot (\ln X(t) - \ln K) \\ Kp(t) = -\beta \cdot (\ln |X(t)| - \ln K) \end{cases}$$

Poiché $X(t) > 0$, per definizione, si ha:

$$\begin{cases} \frac{dX(t)}{dt} = \beta \cdot X(t) \cdot \ln \left(\frac{K}{X(t)} \right) \\ Kp(t) = -\beta \cdot \ln \left(\frac{X(t)}{K} \right) \end{cases} \quad (2)$$

La prima equazione del sistema (2) è esattamente la (3).

2. Soluzione generale della 1ª equazione di (2)

Sia data la seguente equazione differenziale:

$$\frac{dX(t)}{dt} = \beta \cdot X(t) \cdot \ln \left[\left(\frac{X(t)}{K} \right)^{-1} \right] \quad (3)$$

dove:

$X(t)$ è la quantità di una popolazione di individui al tempo t ;
 β è il tasso di crescita genetica della popolazione considerata;
 $X(t)/K$ è il tasso di occupazione dell'ambiente;
 $K = \text{Max}[X(t)]$ è la capacità portante di un territorio;
 $\beta \ln[(X(t)/K)^{-1}]$ è il tasso di crescita smorzata della popolazione considerata.

Questa equazione differenziale ordinaria del primo ordine venne proposta da Benjamin Gompertz come modello matematico di crescita di una popolazione di individui che occupano un territorio finito (nello spazio o nelle risorse) al tempo t . Recentemente essa è stata utilizzata per descrivere il comportamento dinamico di un particolare tipo di tumori sferoidali.

Prima di affrontare l'applicazione di questo modello al problema della crescita di questo tipo di tumori, determiniamo la soluzione generale di questa equazione differenziale. (3) può essere scritta anche nel modo seguente:

$$\frac{dX(t)}{X(t) \cdot \ln \left(\frac{X(t)}{K} \right)^{-1}} = \beta \cdot dt \quad (4)$$

Con successivi passaggi si ha:

$$\frac{dX(t)}{X(t) \cdot \ln \left(\frac{X(t)}{K} \right)} = -\beta \cdot dt \quad , \quad \int \frac{dX(t)}{X(t) \cdot \ln \left(\frac{X(t)}{K} \right)} = - \int \beta \cdot dt$$

Si pone:

$$u = \ln \left(\frac{X(t)}{K} \right) \quad ; \quad u = \ln X(t) - \ln K \quad ; \quad du = \frac{1}{X(t)} \cdot dX(t)$$

$$\int \frac{1}{u} du = - \int \beta \cdot dt \quad ; \quad \ln |u| = -\beta \cdot t + c_0$$

$$|u| = e^{-\beta t + c_0} \quad , \quad \left| \ln \left(\frac{X(t)}{K} \right) \right| = e^{-\beta t + c_0}$$

$$-\ln \frac{X(t)}{K} = C \cdot e^{-\beta t} \quad \text{dove} \quad C = e^{c_0}$$

$$\ln \left(\frac{X(t)}{K} \right)^{-1} = C \cdot e^{-\beta t} \quad , \quad \ln \left(\frac{K}{X(t)} \right) = C \cdot e^{-\beta t}$$

$$\frac{K}{X(t)} = e^{C \cdot e^{-\beta t}} \quad \text{e quindi } X(t) = K \cdot e^{-C \cdot e^{-\beta t}} \quad (5)$$

Questa relazione è la forma chiusa generale risolutiva di (3).

3. Discretizzazione del modello e accostamento

Il problema ora è quello di stimare i parametri del modello (3) o (5) adattandolo ai dati sperimentali mediante il principio dei minimi quadrati.

Il passaggio inevitabile è la discretizzazione del modello. Il metodo è quello di Eulero, ma più semplicemente basta muoversi per variazioni finite. Per stimare i parametri di (3), discretizziamo la (3):

$$\Delta X_t = \beta \cdot X_t \cdot \text{Ln} \left[\left(\frac{X_t}{K} \right)^{-1} \right]; \quad (6)$$

poniamo $\Delta t = 1$. Si prosegue con

$$X_{t+1} - X_t = \beta \cdot X_t \cdot \text{Ln} \left[\left(\frac{K}{X_t} \right) \right]$$

$$X_{t+1} = \left(1 + \beta \cdot \text{Ln} \left(\frac{K}{X_t} \right) \right) \cdot X_t \quad (7)$$

$$\frac{X_{t+1}}{X_t} = \left(1 + \beta \cdot \text{Ln} \left(\frac{K}{X_t} \right) \right) \quad (8)$$

Se si pone $X_{t+1} / X_t = Y_t$ si ha:

$$Y_t = 1 + \beta \cdot \text{Ln} K - \beta \cdot \text{Ln} X_t \quad (9)$$

Poniamo ora $A = 1 + \beta \cdot \text{Ln} K$ e $B = -\beta$. Si ottiene:

$$Y_t = A + B \cdot \text{Ln} X_t \quad (10)$$

Come si può vedere, abbiamo ottenuto un modello lineare nei parametri. Applicando il principio dei minimi quadrati, utilizzando i dati sperimentali, possiamo stimare i parametri A e B e quindi determinare, a ritroso, β , K e $Kp(t)$.

Vediamo come:

$$S(A, B) = \sum_{j=1}^n (Y_j - \hat{Y}_j)^2 = \sum_{j=1}^n (Y_j - A - B \cdot \text{Ln} X_j)^2 \quad (11)$$

$$\begin{cases} \frac{\partial S(A, B)}{\partial A} = \sum_{j=1}^n 2 \cdot (Y_j - A - B \cdot \text{Ln} X_j) \cdot (-1) \\ \frac{\partial S(A, B)}{\partial B} = \sum_{j=1}^n 2 \cdot (Y_j - A - B \cdot \text{Ln} X_j) \cdot (-\text{Ln} X_j) \end{cases}$$

$$\begin{cases} \sum_{j=1}^n (Y_j - A - B \cdot \text{Ln} X_j) \cdot (-1) = 0 \\ \sum_{j=1}^n (Y_j - A - B \cdot \text{Ln} X_j) \cdot (-\text{Ln} X_j) = 0 \end{cases} \quad (12)$$

(12) è equivalente alla seguente forma matriciale:

$$\begin{pmatrix} n & \sum_{j=1}^n \text{Ln} X_j \\ \sum_{j=1}^n \text{Ln} X_j & \sum_{j=1}^n (\text{Ln} X_j)^2 \end{pmatrix} \cdot \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n Y_j \\ \sum_{j=1}^n Y_j \cdot \text{Ln} X_j \end{pmatrix}$$

Con la risoluzione di questo sistema si trovano A e B . Da B si determina β e quindi K , e poi con (7) da X_t si determina X_{t+1} . Si può anche andare direttamente al modello generale (5). Qui esiste il problema della determinazione di C . Esso è legato a una condizione iniziale. Al tempo $t = 0$, $X(0) = K e^{-C}$. Sappiamo sia $X(0)$, numerosità degli individui all'inizio del processo, sia K . Quindi C è uguale a $C = \text{Ln} K - \text{Ln} X(0)$. Se studiamo il limite di (5) per $t \rightarrow \infty$, si ottiene che $X(t) \rightarrow K$ coerentemente ai significati dell'argomento.

4. Verifica statistica della bontà dell'accostamento

Abbiamo visto come si possono stimare i parametri del modello (3). Una volta che si è selezionato dalla famiglia di modelli (3) quello che passa tra i punti sperimentali rispettando la condizione dei minimi quadrati, si passa alla verifica del-

la bontà dell'accostamento fatto.

Si può procedere con 3 diversi livelli di approfondimento:

- 1) Calcolare l'indice di accostamento di Pearson; cioè il coefficiente di variazione relativo al modello stimato;
- 2) Studiare l'autocorrelazione dei residui e verificare se il suo andamento è smorzato all'aumentare del passo k .

Una stima sintetica della bontà dell'accostamento fatto tra modello teorico e dati sperimentali, si ottiene calcolando il seguente indice di accostamento quadratico relativo:

$$I_2 = \frac{1}{M(\hat{X}_j)} \cdot \sqrt{\frac{\sum_{j=1}^n (X_j - \hat{X}_j)^2}{n}} \quad (13)$$

Dove I_2 è il coefficiente di variazione e quindi il rapporto tra la deviazione standard tra dati sperimentali e valori teorici e la media aritmetica calcolata sui valori teorici.

Se questo risultato è inferiore a 0,1 (valore convenzionale generalmente accettato come buono) [B.3], il modello è pienamente accettabile; altrimenti possono sorgere dei dubbi sulla significatività dell'accostamento tra modello teorico e gruppo di dati sperimentali.

Una volta che abbiamo trovato il modello e abbiamo verificato la bontà dell'accostamento fatto mediante l'analisi di I_2 , rimane da analizzare il comportamento dei residui, della differenza

$$a_j = X_j - \hat{X}_j \quad (14)$$

Questi residui devono avere alcune importanti proprietà perché il modello stimato sia significativamente accettabile. Elenchiamo queste proprietà:

- 1) $M(a) = 0$; (15)
- 2) $V(a) = \sigma_a^2$; (16)
- 3) $c_k(a)$, che è la funzione di autocovarianza degli errori a_j , va a 0 rapidamente per k sufficientemente grande. La funzione di autocovarianza è data da:

$$c_k(a_j) = \frac{1}{n} \sum_{j=1}^{n-k} (a_{j+k} - M(a)) \cdot (a_j - M(a)) \quad \forall k \in \mathbb{N} \quad (17)$$

La relazione (17), se si dimostra vera la (15), può essere semplificata considerando che $M(a_j) = 0$ e $c_k(a_j) = c_k$. Si ottiene così:

$$c_k = \frac{1}{n} \cdot \sum_{j=1}^{n-k} a_{j+k} \cdot a_j \quad , \quad \forall k \in \mathbb{N}; \quad (18)$$

altrimenti si deve considerare il valore calcolato di $M(a_j)$.

Quando $k = 0$, $c_k = \sigma_a^2$.

Si deve verificare che la funzione di autocovarianza sperimentale converga a 0 rapidamente. Conviene lavorare con la funzione di autocorrelazione. In tal caso

$$r_k = c_k / c_0 \quad , \quad \forall k \in \mathbb{N} \quad (19)$$

Mediante la (19) si può costruire il correlogramma sperimentale al variare di k . Esso permette di valutare se il modello teorico stimato ha colto tutta la relazione lineare tra $X(t)$ e t oppure no. Un correlogramma non immediatamente degradante verso lo zero farebbe intendere che non tutto il fenomeno è stato colto dal modello teorico stimato. [Segue al numero 103]

Bibliografia: [B.1] Daisuke Satoh, *NTT Service Integration Laboratories* 3-9-11 Midori-cho Musashino-shi, Tokio 180-8585 Japan - email: satoh.daisuke@lab.ntt.co.jp; [B.2] R. Chignola, R. Foroni, A. Franceschi, M. Pasti, C. Candiani, C. Anselmi, G. Fracasso, G. Tridente and M. Colombatti, *Heterogeneous response of individual multicellular tumour spheroids to immunotoxins and ricin toxin*, British Journal of cancer (1995) 72, 607-614, Stockton Press; [B.3] Gambotto Manzone Susara Longo, *Probabilità e statistica 2*, ed. Tramontana, 1998, Milano; [B.4] George E. P. Box, Gwilym M. Jenkins, *Time series analysis, Forecasting and Control*, Holden-DAY, San Francisco, (1976); [B.5] Roberto Chignola, *Proposta di lavoro su "Matematica e crescita di cellule tumorali"*, Marzo 2006, Verona, e.mail roberto.chignola@univr.it; [B.6] R. Chignola, A. Schenetti, E. Chiesa, R. Foroni, S. Sartoris, A. Brendo-Ian, G. Tridente, G. andrighetto and D. Liberati, *Oscillating growth patterns of multicellular tumour spheroids*, 1999 Blackwell Science Ltd Cell Proliferation 32 39-48.