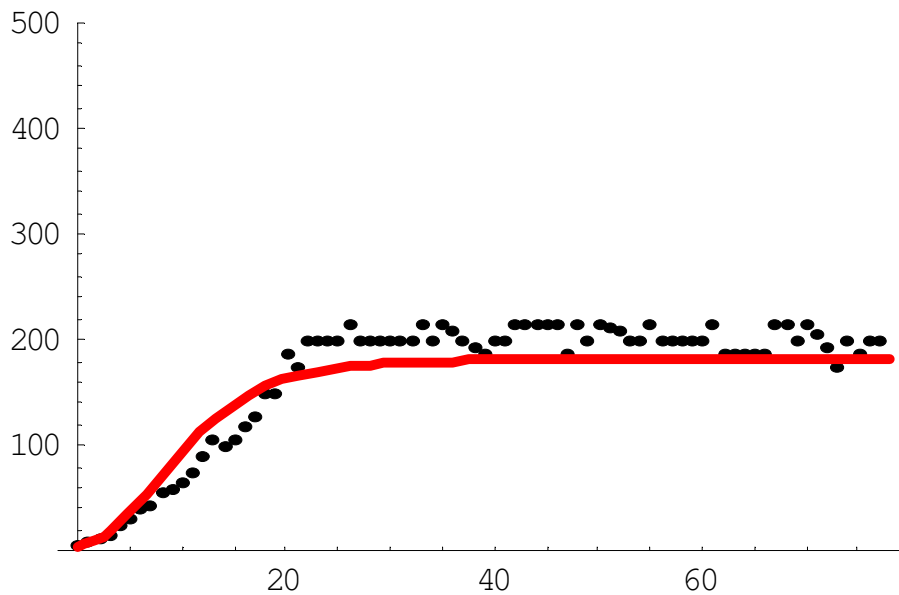


---

## I quaderni del Marconi

# La Biomatemematica in un problema di oncologia sperimentale



**a.s. 2005/2006**

**Laboratorio di matematica applicata  
dell'ITIS "G. Marconi"**

**per il progetto ministeriale "Lauree scientifiche"**

**a cura di**

**R. Chignola-F. Castelli-L. Corso-G. Pezzo-S. Zuccher**

---

---

## Presentazione

L'ITIS Marconi è stato scelto, per l'a.s. 2005/2006, come scuola Polo per l'attuazione del progetto ministeriale denominato "Lauree scientifiche", ideato per incentivare lo sviluppo delle discipline scientifiche nella scuola superiore.

Il progetto proponeva la creazione di un gruppo di lavoro con gli studenti per la soluzione di un problema di "Matematica Applicata" in modo da far comprendere l'importanza e la potenza della matematica nella soluzione di problemi reali.

In collaborazione con l'Università di Verona, la nostra scuola ha scelto un problema di Biologia (Biometria) che ha appassionato e affascinato i nostri alunni delle classi quinte informatica.

Visti i buoni risultati ottenuti si è pensato di rendere pubblico il lavoro nei "Quaderni del Marconi"

Hanno partecipato al progetto gli alunni:

Amorim Gabriela	Freddo Giorgia	Marchetto Giacomo
Banterle Marco	Furlani Dimitri	Ruffo Massimiliano
Battiston Mattia	Gaiga Michele	Segna Tommaso
Benasi Plinio	Longo Marco	Scandola Luca
Castellani Federico	Lorenzi Giovanni	Tenuti Mattia
Fraccaroli Francesco	Marazzi Mattia	Tinazzi Mattia

hanno collaborato i docenti:

- *Roberto Chignola* - Ricercatore di Patologia generale dell'Università degli studi di Verona  
per la presentazione del lavoro dal punto di vista biologico e per i dati sperimentali forniti
- *Fabio Castelli, Luciano Corso, Gianfranco Pezzo* - docenti ITIS Marconi di Verona  
per la presentazione dei modelli e delle procedure matematiche e statistiche e per l'elaborazione dei dati con il software applicativo "Mathematica" della Wolfram Research
- *Simone Zuccher* - collaboratore Università degli studi di Verona  
per l'integrazione sui metodi numerici

---

---

## Indice

<b>1. Il Progetto</b>	7
<b>2. Il problema biologico</b>	8
<b>3. L'elaborazione matematica e statistica</b>	25
<b>4. Una proposta di discretizzazione</b>	30
<b>5. L'elaborazione</b>	33
<b>6. Bibliografia</b>	45

---

---

## 1. Il progetto

L'applicazione nasce da una proposta del dottor Roberto Chignola, biologo, ricercatore di Patologia Generale presso l'Università degli Studi di Verona, che ha presentato un problema riguardante lo studio delle caratteristiche di crescita del volume degli aggregati di cellule tumorali (sferoidi) di un particolare istotipo.

Chignola, dopo aver introdotto i principi biologici del problema, ha presentato il modello di Gompertz come descrittore della crescita di questo tipo di tumori. Il modello si basa su un sistema di equazioni differenziali da cui poi si arriva all'equazione differenziale di Gompertz che è un'equazione differenziale del primo ordine (argomento già noto agli alunni) con soluzione generale in forma chiusa.

I docenti Castelli, Corso, Pezzo hanno poi trattato la parte matematica riguardante:

- la soluzione del sistema di equazioni differenziali del modello
- la discretizzazione del modello con il metodo di Eulero
- la stima dei parametri del modello mediante il metodo dei minimi quadrati (noto agli studenti perché in programma al 4° anno), analizzando la questione della linearizzazione del modello
- l'analisi dei residui e degli indici di accostamento.

La parte teorica è stata integrata da un intervento di Zuccher sul metodo numerico di Eulero implicito ed esplicito.

I dati forniti dal dott. Chignola riguardanti:

- tempi di osservazione in giorni
- volumi degli sferoidi in  $\text{mm}^3 \cdot 10^{-3}$
- volumi teorici calcolati con il modello di Gompertz
- residui,

sono stati elaborati dai ragazzi con il software Mathematica della Wolfram Research sotto la guida dei docenti Castelli, Pezzo, Corso.

---

## 2. Il problema biologico

di Roberto Chignola

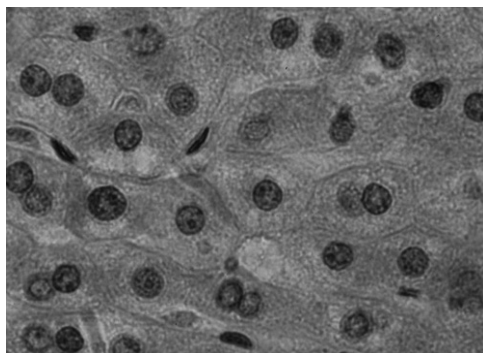
### 2.1. Introduzione

I tumori rappresentano ancora oggi una delle principali cause di morte nei paesi sviluppati. La prevalenza dei tumori in Italia varia tra 1.300-11.300 casi per 100.000 abitanti a seconda dell'età, del sesso e della regione geografica (stimata su 250.000 casi diagnosticati di tumore nel periodo 1978-1992, si veda F.Berrino, in Istituto Nazionale Tumori: l'attività clinico-scientifica del 2000 (disponibile online all'URL:

[www.istitutotumori.mi.it/menuistituto/struttura/pubblicazioni/pdf/epidemiologia.pdf](http://www.istitutotumori.mi.it/menuistituto/struttura/pubblicazioni/pdf/epidemiologia.pdf)).

Ma cosa sono i tumori?

Il nostro organismo è costituito da unità fondamentali che vengono chiamate cellule della dimensione media di circa  $10^{-5}$  m di diametro. Si stima che il corpo di un adulto contenga qualcosa come  $10^{14}$  cellule tutte derivanti da un'unica cellula, ovvero da una cellula uovo fecondata da uno spermatozoo. Le cellule non sono tutte uguali ma assumono diverse funzioni, aspetto e nome a seconda del tessuto a cui sono destinate. Ad esempio un neurone, la cellula fondamentale del sistema nervoso, è una cellula completamente diversa da un epatocita, la cellula fondamentale del fegato. "Completamente diversa" vuol dire che ha una struttura differente che, a sua volta, indica che è pure differente la composizione molecolare e quindi la funzione della cellula stessa. Le cellule formano i tessuti prendendo contatto fra loro e quindi gli organi come il fegato, il pancreas, la milza e così via. L'unica eccezione è rappresentata dalle cellule del sangue che non prendono contatto tra loro per poter scorrere nelle vene.



**Fig.1.** Un gruppo di epatociti (i cerchi più scuri sono i nuclei delle cellule) a sinistra e alcuni neuroni a destra. Da:

epatociti: [www.sci.sdsu.edu/histology/cy01.htm](http://www.sci.sdsu.edu/histology/cy01.htm)

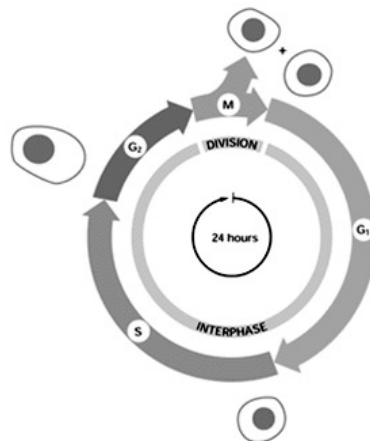
neuroni: [inc2.ucsd.edu/~leelab/\\_\\_\\_Projects/Bioinformatics/bioinformatics.html](http://inc2.ucsd.edu/~leelab/___Projects/Bioinformatics/bioinformatics.html)



---

Una caratteristica comune a "tutte" le cellule (questo vale in prima approssimazione. Per le cellule del sistema nervoso il discorso che segue diventa assai più complicato) è la capacità di moltiplicarsi mediante un meccanismo di divisione. Le cellule devono poter avere questa capacità per poter rigenerare i tessuti danneggiati a seguito di un trauma o a causa del normale rinnovo dei tessuti. Ad esempio, le cellule del sistema digerente sono sottoposte ad un notevole stress meccanico e chimico e quindi muoiono, e le cellule morte vengono rimpiazzate da nuove cellule che si originano per divisione cellulare. Inoltre ogni cellula ha un tempo di vita limitato (nell'ordine dei mesi) ben inferiore al tempo di vita media di un essere umano (circa 80 anni) e dunque nei tessuti vi è un continuo ricambio di nuove cellule giovani che vanno a sostituire le cellule morte. Il processo della divisione cellulare viene generalmente schematizzato in un "ciclo" della durata di circa 24 ore e suddiviso in quattro fasi: la fase G<sub>1</sub>, S, G<sub>2</sub> e M. Durante queste fasi la cellula madre deve raddoppiare le proprie componenti molecolari onde suddividerle nelle due cellule figlie. Dunque alla fine del ciclo cellulare si originano due copie "quasi identiche" della cellula madre.

Le 4 fasi del ciclo cellulare sono regolate in sequenza da complicati circuiti molecolari che però garantiscono che la duplicazione della cellula venga fatta in modo corretto. Ad esempio, il codice genetico della cellula madre viene duplicato interamente durante la fase S. La duplicazione del DNA richiede energia e dunque durante la fase G<sub>1</sub>, che precede la fase S, vi sono punti di controllo - dei veri e propri checkpoint molecolari - che monitorano lo stato energetico della cellula. Se non vi è abbastanza energia il ciclo cellulare viene arrestato e la cellula non inizia a duplicare il DNA. Inoltre, i sistemi molecolari per quanto finemente regolati possono commettere errori.



**Fig.2.** Schema del ciclo cellulare. Da:  
[www.geneticsebiologia.unito.it/duplicazione\\_dna/introduzione.html](http://www.geneticsebiologia.unito.it/duplicazione_dna/introduzione.html)

---

Così è possibile che vengano commessi errori nella duplicazione del DNA ed infatti nella stessa fase S e nella successiva fase G2 vi sono punti di controllo che "permettono alla cellula di verificare" che la duplicazione sia avvenuta correttamente. Dato che il DNA contiene l'informazione per la produzione di tutte le proteine - tra cui gli enzimi che permettono ad una cellula di vivere - e che le proteine conferiscono alla cellula la sua struttura e la sua funzione proprie di un dato tessuto, si capisce come i punti di controllo molecolare del ciclo cellulare siano fondamentali per garantire la corretta divisione cellulare e, in ultima analisi, l'omeostasi tissutale. Infine, vi sono circuiti molecolari che permettono alla cellula di arrestare il processo di divisione quando questo non sia necessario. Ad esempio, quando ci procuriamo una ferita e lesioniamo un tessuto come la pelle, le cellule iniziano a duplicarsi per rimarginare la ferita ma, quando la ferita è rimarginata, le cellule cessano di proliferare. In modo molto schematico possiamo chiamare questo meccanismo "inibizione da contatto": quando le cellule si trovano in stretto contatto tra loro vengono inviati segnali molecolari che arrestano il ciclo cellulare.

Da questa sintetica descrizione delle cellule si può intuire la complessità del mondo biologico a livello di scala microscopica: ogni cellula è in stretto contatto con le altre e con l'ambiente circostante ed è in grado di percepire le variazioni della composizione dell'ambiente e reagire ad esse regolando i propri circuiti molecolari. Questo è possibile grazie alla presenza di recettori che vengono espressi sulla superficie della cellula e che sono in grado di raccogliere i segnali provenienti dall'ambiente esterno e di veicarli dentro alla cellula stessa.

Per una descrizione più dettagliata delle cellule si può consultare gratuitamente il sito:

<http://it.wikipedia.org/wiki/Cellula>

Torniamo ora alla domanda originale: che cosa sono i tumori? Un tumore è costituito da un gruppo di cellule caratterizzate dal fatto di aver perduto il controllo del ciclo cellulare e di avere alterazioni nei circuiti molecolari che istruiscono la cellula su quando arrestare la proliferazione. Ad esempio, le cellule tumorali non sono più sensibili all'inibizione da contatto e continuano a proliferare, dividendosi, anche quando si trovano le une accanto alle altre. La loro capacità proliferativa è molto estesa nel tempo. Ad esempio, ancora oggi in laboratorio coltiviamo e facciamo proliferare cellule di tumori sviluppati negli anni 50 del secolo scorso. Per questo si dice che le cellule tumorali sono immortali.

La definizione di tumore data qui sopra è ovviamente molto semplificata dato che le alterazioni delle cellule tumorali coinvolgono anche i meccanismi che obbligano una determinata cellula ad appartenere ad un certo tessuto, i meccanismi di cattura ed utilizzo dei nutrienti, i meccanismi che permettono alle cellule di relazionarsi con l'ambiente, ecc... Ad esempio, una cellula normale dell'intestino è "programmata" per vivere solo nell'intestino, poiché in quell'ambiente riceve opportuni segnali di sopravvivenza; una cellula tumorale, invece, può vivere anche in altri tessuti e colonizzarli (le metastasi).

Informazioni più dettagliate sui tumori si possono consultare gratuitamente al sito: <http://it.wikipedia.org/wiki/Tumori>

---

## 2.2. La biomatematica per l'oncologia

L'oncologia è la scienza che studia i tumori sotto il profilo biologico e clinico, intendendo con quest'ultimo termine tutti quegli aspetti che riguardano la gestione dei pazienti ammalati di tumore. La biomatematica è quella parte della matematica che si occupa di costruire modelli formali dei processi biologici. Che cos'è che unisce queste due scienze così apparentemente distanti tra loro?

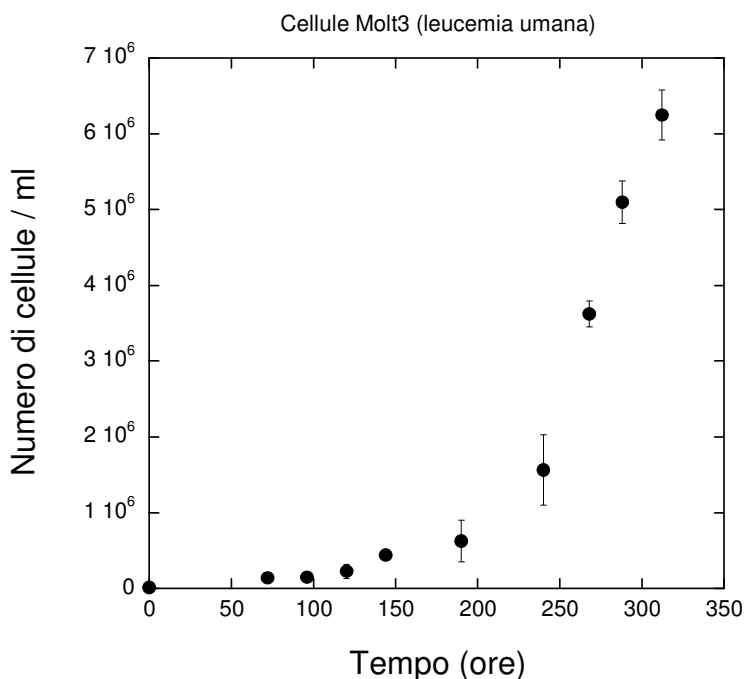
Consideriamo un esperimento biologico condotto con cellule tumorali. Per nostra fortuna, spesso le cellule tumorali possono essere prelevate da un paziente e coltivate in laboratorio. Questo permette di poterle studiare in dettaglio e, ad esempio, di verificare a quali farmaci esse siano più sensibili. Supponiamo allora di avere delle cellule tumorali in coltura e di farle proliferare in appositi contenitori di materiale plastico (le fiasche da coltura). La prima domanda che ci poniamo è: con quale velocità queste cellule si riproducono?

La risposta a questa domanda è fondamentale perché permette di tradurre in termini quantitativi una caratteristica osservabile delle cellule tumorali: la proliferazione. Questa quantificazione può servire a comparare tra loro cellule provenienti da tumori diversi al fine di capire quali tumori siano più aggressivi (se un tumore prolifera molto velocemente è più pericoloso perché compete in modo più aggressivo con l'ospite). Inoltre, da molti anni sappiamo che i trattamenti antitumorali hanno una efficacia che dipende anche dalla velocità di proliferazione delle cellule.

Quello riportato qui sopra è solo uno dei tanti esempi in cui la modellizzazione matematica è utile alla biomedicina. Come ulteriore esempio prendiamo i circuiti molecolari che presiedono al controllo del ciclo cellulare di cui si è accennato nell'introduzione. L'oncologia sperimentale ha compiuto passi da gigante nell'individuare le molecole di tali circuiti e le loro relazioni funzionali. La modellizzazione matematica in questo caso aiuta la comprensione di tali circuiti perché permette di analizzare in modo razionale la gran quantità di dati sperimentali raccolti e, ancora una volta, di trarre conclusioni quantitative.

### 2.3. Un modello semplice di proliferazione tumorale

Consideriamo i dati sperimentali rappresentati nella figura qui sotto.



**Fig.3.** Un esperimento di proliferazione di cellule tumorali umane.

Sull'asse delle ascisse è riportato il tempo misurato in ore, su quello delle ordinate il numero di cellule per unità di volume. Le barre verticali che attraversano i cerchi neri rappresentano una stima dell'errore sperimentale.

Si tratta di cellule tumorali di una leucemia umana coltivate in laboratorio. Queste sono state fatte crescere in condizioni controllate e, ai tempi indicati, le cellule sono state contate. Per contare le cellule si preleva un piccolo campione della soluzione nutritiva in cui le cellule vengono fatte crescere (ad esempio  $10^{-5}$  litri) e lo si deposita su un particolare vetrino da microscopia che contiene una camera di volume noto e su cui è disegnata una griglia. Utilizzando un microscopio si contano le cellule che finiscono nella griglia e con opportuni calcoli si ricava il numero di cellule per unità di volume. I dati mostrati in figura, dunque, descrivono come varia la densità cellulare in funzione del tempo. Perché i dati sperimentali si comportano proprio così e non in un altro modo? (Per esempio, la densità cellulare potrebbe aumentare e diminuire nel tempo seguendo un moto periodico).

Si intuisce l'analogia con i problemi tipici della fisica: per esempio descrivere la traiettoria di un punto materiale sparato da un cannone in un campo gravitazionale. Come in quel caso, anche noi ora vogliamo dare una descrizione matematica della traiettoria seguita dalle cellule nel tempo, ovvero della proliferazione di queste cellule. E lo facciamo cercando di partire dai principi primi.

---

Per quanto detto nell'introduzione ora sappiamo che le cellule si riproducono duplicandosi: da una cellula se ne formano due, da due quattro e così via. Supponiamo di fotografare le cellule ad intervalli regolari ogni volta che esse si dividono. Potremmo così ottenere i seguenti dati:

Tempo dello scatto	Numero di cellule
$t_1$	1
$t_2$	2
$t_3$	4
$t_4$	8
...	...
$t_n$	????

Che cosa succede all' $n$ -esimo scatto? Il problema che vogliamo risolvere è trovare una legge generale. Ci accorgiamo innanzitutto che le cellule osservate ad uno certo scatto sono il doppio delle cellule osservate allo scatto precedente. Ad esempio:

$$N(t_4) = 2N(t_3)$$

dove con  $N(t_3)$  si intende "numero di cellule contate al tempo 3".

Ma ciò vale anche per i tempi precedenti e successivi. Se chiamiamo  $T$  il tempo intercorso tra uno scatto e l'altro possiamo scrivere:

$$N(t + T) = 2N(t)$$

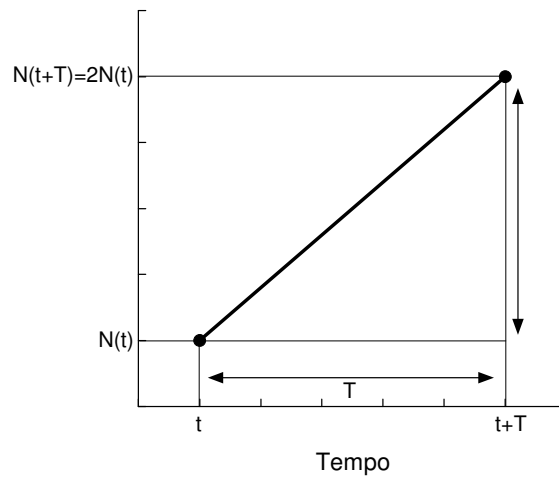
ovvero che le cellule contate al tempo  $t$  più un intervallo  $T$  sono il doppio delle cellule contate al tempo  $t$ . È chiaro che  $T$  assume il significato di "tempo di duplicazione" della popolazione cellulare, cioè il tempo necessario alla popolazione cellulare per raddoppiare la propria dimensione.

Dato che  $2N(t) = N(t) + N(t)$ , possiamo scrivere:

$$N(t + T) = N(t) + N(t)$$

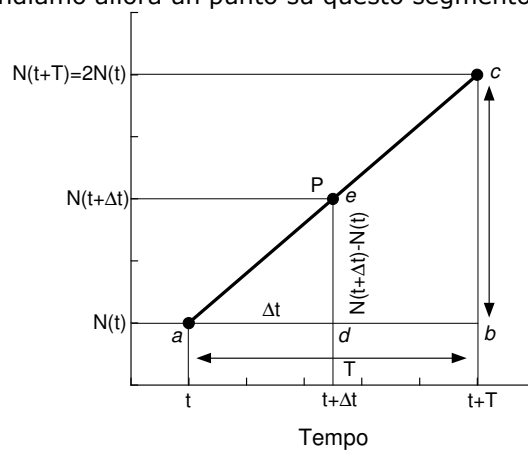
$$N(t + T) - N(t) = N(t)$$

Il membro di sinistra della formula qui sopra dice qual'è la differenza tra le cellule contate dopo un intervallo di tempo e le cellule contate all'inizio di quell'intervallo. Questa differenza è uguale al numero di cellule contate all'inizio dell'intervallo temporale. Se mettiamo in grafico questa formula apparirà subito chiaro il suo significato. Infatti:



**Fig.4.** Rappresentazione grafica della duplicazione cellulare. Il numero di cellule misurato al tempo  $t$ ,  $N(t)$ , raddoppia dopo che è trascorso un tempo  $T$ . Dunque le cellule in  $N(t+T) = 2N(t)$  se le due osservazioni, rappresentate dai cerchi neri, vengono interpolate con una linea retta (*interpolazione lineare*). Si assume che nel passaggio da  $N(t)$  a  $2N(t)$  le cellule abbiamo seguito un percorso regolare e rettilineo nello spazio  $Nxt$  qui rappresentato.

Questo grafico ci suggerisce anche come cercare di ottenere una legge generale del processo di crescita cellulare. Innanzitutto si noti che tra i due punti - che rappresentano il numero di cellule misurato al tempo  $t$  e al tempo  $t+T$  - abbiamo disegnato un segmento rettilineo che li collega. Nessuna osservazione ci autorizza a collegare i due punti con una retta. Si tratta di una approssimazione che facciamo per costruire il modello e tecnicamente questa approssimazione si chiama *interpolazione lineare*. Prendiamo allora un punto su questo segmento e chiamiamolo  $P$ :



**Fig.5.** La figura rappresenta il grafico di Fig.4. Qui, rispetto al grafico precedente, abbiamo considerato un punto  $P$ , lungo la linea retta con cui abbiamo interpolato le osservazioni, in corrispondenza del tempo  $t$  più un intervallo temporale  $\Delta t$  la cui durata è minore della durata di  $T$ .

Le coordinate di questo punto *qualsiasi* preso sul segmento che collega i punti che rappresentano le osservazioni sperimentali sono (si noti l'analogia con quanto definito più sopra):  $P=[t+\Delta t ; N(t+\Delta t)]$ , dove con  $\Delta t$  indichiamo un generico intervallo di tempo inferiore a  $T$ . Si noti a questo punto che abbiamo costruito due triangoli rettangoli tra loro simili, i triangoli  $abc$  e  $ade$ . Ricordandoci le regole geometriche elementari, possiamo allora scrivere:

$$\frac{\overline{bc}}{\overline{ab}} = \frac{\overline{de}}{\overline{ad}}$$

$$\frac{N(t)}{T} \approx \frac{N(t+\Delta t) - N(t)}{\Delta t}$$

dove il simbolo *circa-uguale* ci ricorda che tale espressione vale se accettiamo l'approssimazione ottenuta con l'interpolazione lineare dei dati osservati. A questo punto formuliamo anche l'ipotesi che il processo della crescita si comporti nel tempo in modo regolare, ovvero *assumiamo che  $N(t)$  sia una funzione continua (ovvero non faccia "salti") e derivabile (per esempio il suo grafico non deve avere "angoli") del tempo e che dunque esista il limite:*

$$\lim_{\Delta t \rightarrow 0} \frac{N(t+\Delta t) - N(t)}{\Delta t} = \frac{dN(t)}{dt}$$

$$\frac{dN(t)}{dt} = \frac{1}{T} N(t)$$

In questo modo otteniamo una equazione differenziale che sappiamo risolvere. Prima di cercare la soluzione di questa equazione osserviamo la quantità  $1/T$  nel membro di destra. Questa quantità ha il significato biofisico dell'inverso del tempo di duplicazione. È ragionevole pensare che cellule diverse - o cresciute in contesti ambientali diversi - possano avere tempi di duplicazione diversi. Dunque anche la quantità  $1/T$  potrà variare di conseguenza. Inoltre noi non abbiamo nessuna conoscenza *a priori* del tempo di duplicazione cellulare, vale a dire che non sappiamo se la tal popolazione cellulare si duplica in 24 ore piuttosto che in 36 o 12. In realtà la quantità  $1/T$  - a noi ignota - è proprio quella che vogliamo cercare - stimare dai dati sperimentali come si vedrà tra poco. È importante quindi notare che la quantità  $1/T$  è una quantità ignota, è caratteristica per ogni popolazione cellulare ma soprattutto non è una funzione del tempo. Riscriviamo l'equazione differenziale come:

$$\frac{dN(t)}{dt} = kp \cdot N(t)$$

dove  $kp$  è una costante con le dimensioni di  $1/\text{tempo}$ . Ricordiamo che la sua soluzione si calcola nel seguente modo:

---

$$\frac{1}{N(t)} dN(t) = kp \cdot dt$$

$$\int_{N(0)}^{N(t)} \frac{1}{N(s)} dN(s) = kp \int_0^t ds \rightarrow \log(N(s)) \Big|_{N(0)}^{N(t)} = kp \cdot s \Big|_0^t \rightarrow \log(N(t)) - \log(N(0)) = kp \cdot (t - 0)$$

$$\log \frac{N(t)}{N(0)} = kp \cdot t$$

$$N(t) = N(0)e^{kp \cdot t}$$

dove  $N(0)$  è la dimensione della popolazione iniziale.

Notiamo innanzitutto che abbiamo raggiunto un risultato importante: abbiamo cioè scoperto che una popolazione cellulare che prolifera mediante duplicazione DEVE crescere con legge esponenziale. Non può fare altro, e se fa altro allora vuol dire che la popolazione cellulare cresce con altri meccanismi. È proprio qui che sta la forza della modellizzazione matematica. Date certe ipotesi si ottengono le conclusioni che ne conseguono e se le conclusioni non aderiscono alla realtà allora vuol dire che le ipotesi sono incomplete.

Discutiamo ancora un po' la nostra procedura di modellizzazione. L'equazione differenziale dice che la velocità di crescita di una popolazione cellulare (membro di sinistra) è proporzionale alla dimensione della popolazione cellulare stessa ad un dato tempo e che  $kp$  è la costante di proporzionalità. Ma che informazioni ci dà  $kp$ ? Torniamo alla soluzione della equazione differenziale e chiediamoci *quando* la popolazione cellulare è raddoppiata, cioè quando  $N(t) = 2N(0)$ :

$$N(t) = N(0)e^{kp \cdot t}$$

$$2N(0) = N(0)e^{kp \cdot T}$$

$$2 \frac{N(0)}{N(0)} = e^{kp \cdot T} \quad \text{da cui} \quad \log 2 = kp \cdot T$$

$$T = \log 2 \frac{1}{kp} \approx 0.693 \frac{1}{kp}$$

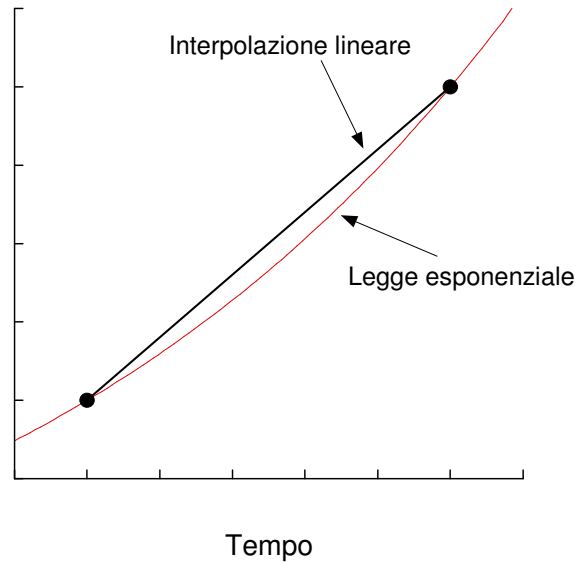
Dunque il *tempo di duplicazione determinato dal modello esponenziale* è dato dal rapporto tra un numero - il logaritmo di 2 - e la costante  $kp$ .

Il lettore attento si chiederà come mai il tempo di duplicazione  $T$  non risulta semplicemente uguale all'inverso della costante  $kp$ . Si ricordi che all'inizio avevamo



---

semplificato il percorso da  $N(t)$  a  $2N(t)$  mediante una interpolazione lineare. Poi abbiamo ridotto progressivamente gli intervalli di tempo intermedi fino a considerare intervalli piccolissimi (tendenti a zero). Infine, a seguito della risoluzione della equazione differenziale abbiamo scoperto che il vero percorso da  $N(t)$  a  $2N(t)$  nello spazio  $Nxt$  non è *lineare* - come nella prima grossolana approssimazione - ma *esponenziale*. Il logaritmo di 2 tiene proprio conto di questa differenza nelle due definizioni. Si noti, pertanto, che il tempo di duplicazione stimato dalla interpolazione lineare è superiore rispetto a quello stimato dalla legge esponenziale di crescita. La figura seguente mette in evidenza questa differenza:



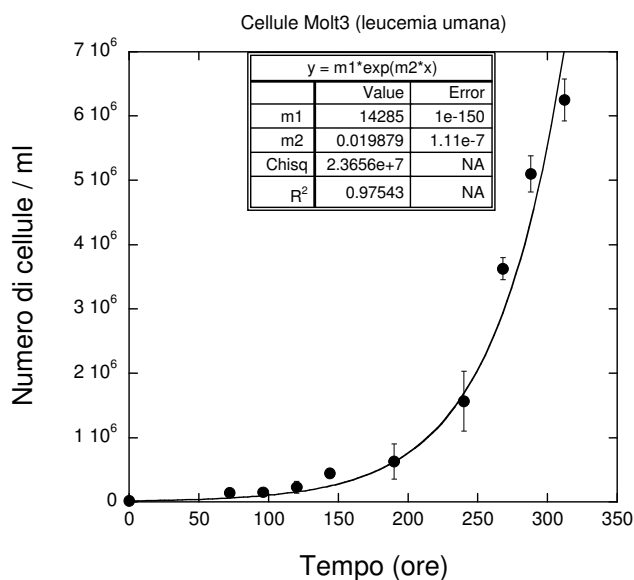
**Fig.6.** Differenza tra l'interpolazione lineare attraverso le osservazioni utilizzata come approssimazione nella costruzione del modello e la legge esponenziale di crescita trovata. Come si può notare il segmento rettilineo che collega i due punti sta sempre sopra alla curva esponenziale.

Il problema, ora, è capire come si fa a stimare il valore della costante  $kp$  a partire dai dati sperimentali e dal nostro modello esponenziale di crescita.

## 2.4. Uso del modello semplice

Qui sopra abbiamo dunque ricavato un modello semplice di proliferazione per le cellule tumorali. Che cosa ci possiamo fare? Da un lato abbiamo i nostri dati sperimentali e dall'altro il modello. Dobbiamo pertanto mettere assieme le due cose, e il modo per metterle assieme si chiama accostamento (=fit). Si tratta di fare un *fit* ai dati sperimentali, ovvero di trovare la curva che meglio approssima i dati tra tutte le curve originate dal nostro modello esponenziale. In pratica, cioè, dobbiamo considerare l'equazione esponenziale ottenuta come un fascio di curve di parametro  $kp$  e ricondurre il problema a quelli già visti di matematica analitica (es. trovare l'equazione della retta di equazione generale  $y=mx+q$  che passa dal punto  $P$  di coordinate  $x_p$  e  $y_p$  e parallela ad un'altra retta data). Il problema qui è che i dati sperimentali non sono MAI perfettamente precisi. Ogni misura sperimentale, infatti, contiene un errore (che viene compiuto dall'operatore, dalla macchina per fare le misure, da una fluttuazione del sistema sperimentale, ecc.) e dunque noi dobbiamo accontentarci di trovare l'equazione esponenziale non che passi *attraverso* i dati sperimentali ma che meglio *approssimi* i dati sperimentali.

Si intuisce così l'idea che sta dietro agli algoritmi di *fit*: la curva, che dipende dal parametro  $kp$ , passerà ad una certa distanza (euclidea) dai dati sperimentali. Possiamo allora calcolare la distanza della curva da ogni punto sperimentale e cercare il valore di  $kp$  che rende *minima* questa distanza. La descrizione di tali algoritmi, tuttavia, esula dagli scopi di questa nota. Accontentiamoci per ora di sapere che esistono algoritmi per fare i *fit* e usiamoli. Ecco il *fit* dei dati di Fig.3:



**Fig.7.** Fit dei dati sperimentali con il modello esponenziale.

---

La linea continua che passa tra i punti sperimentali è quella che meglio approssima i dati. Come si vede nel riquadro, si possono usare degli indici statistici per verificare la bontà del *fit*, ma comunque anche ad occhio si vede che la curva sembra rappresentare piuttosto bene i dati sperimentali. L'equazione che otteniamo dal *fit* è:

$$N(t) = 14285 \cdot e^{0.0198 \cdot t}$$

e pertanto  $N(0)=14285$  (cellule  $\text{ml}^{-1}$ ) e  $kp=0.0198$  (ore) $^{-1}$  (si noti che viene prestata la massima cura alle unità di misura).  $N(0)$  è il numero di cellule iniziale. Questo parametro lo conosciamo perché lo abbiamo stabilito noi all'inizio dell'esperimento quando abbiamo messo le cellule in coltura. Per questo esperimento erano state seminate circa 10000 cellule / ml, e quindi il *fit* ha dato un valore abbastanza buono di  $N(0)$ . Poi abbiamo il valore stimato di  $kp$  e, come abbiamo fatto al paragrafo precedente, possiamo usarlo per calcolare il tempo di duplicazione di questa popolazione cellulare:

$$T = \frac{\log 2}{kp} \approx \frac{0.693}{0.0198 \text{ (ore)}^{-1}} \approx 35 \text{ ore}$$

Dunque queste cellule, in questo esperimento, hanno mostrato un tempo di duplicazione di circa 35 ore. Un po' lentine per la verità dato che in altri esperimenti le stesse cellule si sono dimostrate assai più veloci con tempi di duplicazione di circa 24 ore. Ma questa è la variabilità del mondo biologico, mai perfettamente uguale a sé stesso. Ciò che importa notare è che ora abbiamo un modello che ci permette di quantificare la velocità di crescita delle cellule in coltura e grazie alla quantificazione possiamo confrontare, monitorare, ecc... Ad esempio ci possiamo chiedere: perché le cellule questa volta proliferano lentamente? C'è stato un errore commesso dallo sperimentatore, una variazione nella strumentazione (per esempio nella regolazione della temperatura dell'ambiente di coltura) oppure in coltura si è selezionata una popolazione di cellule più lente che ha preso il sopravvento?

È chiaro che ora abbiamo un modo formale per cercare di rispondere a queste domande ripetendo l'esperimento e analizzando i dati come fatto in questo paragrafo. Senza modello biomatematico questo non sarebbe possibile perché non avremmo un modo per ottenere una quantificazione oggettiva.

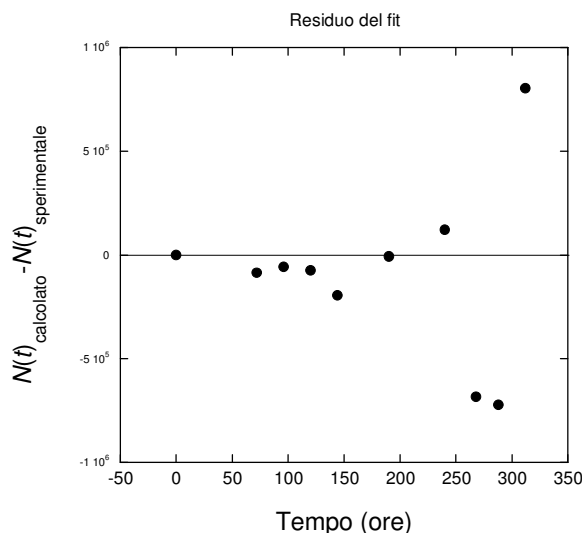
Infine affrontiamo un ulteriore problema. Ci chiediamo: il modello che abbiamo ottenuto descrive veramente bene il processo biologico di crescita o manca ancora qualcosa? Questa domanda è diversa dalla precedente se il *fit* sia buono o no. Come si vede nella figura precedente il *fit* è buono, ma il modello è completo? Per quanto detto qui sopra sappiamo che i dati sperimentali contengono sempre un certo errore, non se ne può fare proprio a meno, purtroppo. Allora, se questo errore è casuale l'errore stesso può essere rappresentato da un processo simile a quello del lancio di una moneta. Insomma, noi facciamo una misura e a questa sottraiamo o aggiungiamo una piccola quantità a seconda che la moneta (non truccata) dia testa o croce. Dunque ci attendiamo che i dati sperimentali stiano un po' sopra e un po' sotto ai valori teorici previsti dal *modello deterministico* (privo di rumore), ovvero stiano un po' sopra e un po' sotto alla curva del *fit*. Dato che noi

---

conosciamo questa curva possiamo calcolare i valori attesi dal modello e confrontarli con i valori sperimentali. Ci attendiamo cioè che la differenza:

$$N(t)_{\text{calcolato}} - N(t)_{\text{sperimentale}}$$

dia una successione di valori che sono un po' positivi e un po' negativi. Più esattamente ci attendiamo che la successione *fluttui casualmente attorno allo zero*. Questa successione ha un nome e si chiama *residuo del fit*. Esistono test statistici per determinare se il residuo del *fit* fluttua in modo casuale: se questo avviene allora il modello ha catturato tutta l'informazione contenuta nel processo biologico, perché ciò che rimane dalla differenza tra il modello deterministico e i dati sperimentali è solo rumore, cioè un processo casuale (=errore sperimentale); se questo non avviene allora manca qualcosa nel modello, cioè il modello non ha catturato tutta l'informazione del processo biologico analizzato.



**Fig.8.** Il residuo del *fit* di Fig.4.

La figura qui sopra mostra il residuo del *fit* e mette in luce un problema. Come si vede i dati sperimentali sono pochi. Con così pochi dati non è possibile applicare i test statistici per dimostrare che il residuo contiene solo una fluttuazione casuale. Ma ciò fornisce una informazione preziosa allo sperimentatore che al prossimo esperimento, se vorrà fare le cose per bene, sarà costretto a fare più misure.

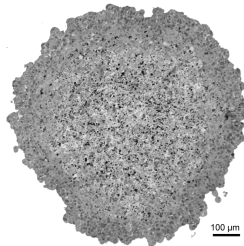
Quest'ultima considerazione fornisce un'ulteriore prova dell'importanza della modellizzazione nelle scienze sperimentali. La modellizzazione può dare importanti indicazioni su come effettuare esperimenti in modo che i risultati siano davvero interpretabili. Teoria e sperimentazione sono dunque strettamente legate tra loro (anche se questa conclusione non è ancora del tutto compresa nella scienza biomedica).

---

## 2. 5. Il problema

E veniamo ora al problema vero e proprio. L'esperimento che abbiamo analizzato fin qui è stato condotto con cellule tumorali separate le une dalle altre. In sostanza ogni cellula poteva crescere indipendentemente dalle altre. Tecnicamente, quando questo avviene, si parla di colture bidimensionali. Le colture bidimensionali sono importanti perché ci permettono di propagare le cellule tumorali e di ottenere una quantità di cellule sufficiente per eseguire varie analisi (es. risposta al trattamento antitumorale, espressione di particolari molecole, ecc...). Se però consideriamo un tumore solido - un aggregato di cellule tumorali - ci accorgiamo che le colture bidimensionali si discostano molto dalla realtà biologica. In un tumore solido, infatti, le cellule sono attaccate le une alle altre e formano una architettura tridimensionale. Questa architettura tridimensionale è responsabile di alcune proprietà biologiche importanti. Innanzitutto le forze di adesione tra una cellula e l'altra fanno acquisire alle cellule particolari proprietà molecolari (grazie all'espressione di determinati geni). In secondo luogo, la struttura tridimensionale determina un ostacolo alla diffusione di nutrienti ed ossigeno. La concentrazione di nutrienti ed ossigeno assume dunque per la cellula valori diversi a seconda della posizione della stessa cellula all'interno della massa tumorale, e le cellule più interne ricevono una minor quantità di nutrienti ed ossigeno. Quando questa quantità scende sotto ad un valore critico le cellule smettono di proliferare. Se scende ulteriormente le cellule muoiono. All'interno di una massa tumorale infatti troviamo cellule che proliferano poco, altre quiescenti, altre morte, mentre le cellule che proliferano attivamente le troviamo sul bordo della massa, ovvero nella zona più ricca di nutrienti ed ossigeno. Tutto questo non è osservabile con le colture bidimensionali poiché ogni cellula è irrorata in ugual modo dai nutrienti.

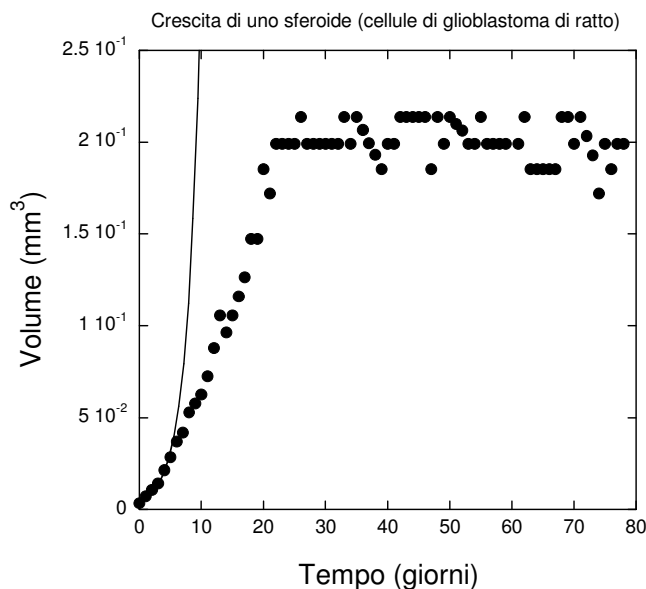
In laboratorio sono così state inventate - attorno agli anni 80 del secolo scorso - delle colture tridimensionali di cellule tumorali che si chiamano sferoidi. In pratica le cellule vengono costrette a stare assieme a formare un piccolo aggregato di poche cellule. Questi aggregati si possono poi isolare con tecniche di micromanipolazione al microscopio e lasciati crescere. Le cellule si dividono e il volume dello sferoide aumenta. Abbiamo numerose prove sperimentali che dimostrano che gli sferoidi hanno proprietà biologiche che li rendono molto simili alle piccole masse tumorali non irrorate da vasi sanguigni. Inoltre gli sferoidi hanno proprietà interessanti che sono ancora oggi oggetto di studio, come la resistenza a certi trattamenti antitumorali. Ciò che a noi interessa in questo momento è che gli sferoidi crescono aumentando il loro volume in funzione della proliferazione cellulare e che è possibile in laboratorio misurare con una certa precisione questo volume.



**Fig.9.** Uno sferoide tumorale costituito da diverse migliaia di cellule. I puntini più scuri evidenziano le cellule che stanno morendo.

Da: [www.vet.purdue.edu/cristal/oci-info.htm](http://www.vet.purdue.edu/cristal/oci-info.htm)

Come cresce uno sferoide? Ecco la curva di crescita misurata di uno sferoide:



**Fig.10.** Un esperimento di crescita con uno sferoide tumorale ottenuto con cellule di glioblastoma di ratto.

Il grafico mostra come cresce il volume di uno sferoide in funzione del tempo. La linea continua mostra il miglior *fit* ottenuto con il modello esponenziale di prima. Come si può vedere, però, in questo caso il modello non è in grado di approssimare i dati sperimentali. Dunque abbiamo bisogno di nuove ipotesi e dobbiamo costruire un nuovo modello. Come fare?

In genere, per costruire un modello si possono seguire due vie: la via *biofisica* e la via *fenomenologica*. Entrambi gli approcci usano tuttavia il linguaggio formale della matematica.

La prima via, quella *biofisica*, è la più difficile. Si tratta di pensare al problema biologico cercando di ricostruirne i tratti fondamentali. Per esempio ci si può chiedere perché la curva di crescita di uno sferoide devii dalla crescita esponenziale e raggiunga un limite (quasi) asintotico in alto. A cosa è dovuta la saturazione della crescita? Sappiamo che all'aumentare del volume si forma una stratificazione cellulare con cellule che proliferano sulla superficie, cellule quiescenti più all'interno e cellule morte al centro. Inoltre sappiamo che questa stratificazione è funzione del gradiente di nutrienti che si viene a formare e che a sua volta è funzione del volume dello sferoide. Potremmo quindi tentare di costruire un modello che tenga conto della diffusione dei nutrienti e di tutte queste cose, ma questo modello sarebbe davvero complicato. Oppure, più semplicemente, potremmo notare che il volume cresce con il cubo del raggio mentre la superficie, attraverso cui i nutrienti vengono assorbiti dall'ambiente esterno, cresce con il quadrato del raggio. Dunque la su-

---

perficie cresce più lentamente del volume. Anche in questo caso potremmo scrivere le opportune equazioni e verificare la correttezza delle nostre ipotesi confrontando il modello con i dati sperimentali. Ciò che importa notare, per il momento, è che l'approccio biofisico consiste nel pensare alla realtà biofisica cercando di individuarne le caratteristiche salienti e quindi traducendo il tutto in linguaggio matematico. È sicuramente l'approccio più complicato ma anche quello più informativo perché permette di legare un fenomeno - come la crescita di uno sferoide - a parametri biofisici importanti - come ad esempio la diffusione dei nutrienti nello sferoide.

La via *fenomenologica*, invece, consiste nel cercare di approssimare i dati sperimentali con una equazione senza alcuna pretesa che questa equazione rappresenti un qualche processo biofisico. A cosa serve? Avere un modello, anche se fenomenologico, serve a fare stime quantitative, stime che poi serviranno a trarre conclusioni biologiche importanti: ad esempio possiamo stimare il tempo di raddoppiamento del volume di una massa tumorale, parametro che sappiamo essere correlato all'efficacia di un trattamento farmacologico o radiologico.

Per semplicità cerchiamo di seguire questa via, ma prima facciamo un passo indietro e consideriamo nuovamente il modello esponenziale nella sua forma differenziale:

$$\frac{dN(t)}{dt} = kp \cdot N(t)$$

Ora sostituiamo le variabili ponendo  $y = dN(t)/dt$  e  $x = N(t)$ . Quello che otteniamo è:

$$y = \frac{dN(t)}{dt} ; \quad x = N(t)$$
$$y = kp \cdot x$$

l'equazione di una retta con coefficiente angolare uguale a  $kp$ . Dunque il modello esponenziale dice: il tasso ( $kp$ ) di crescita delle cellule è costante. Ma chi ha detto che deve essere sempre così? Il tasso di crescita potrebbe ad esempio variare nel tempo in funzione di una qualche variabile, anzi i dati mostrati in figura fanno pensare proprio al fatto che il tasso tenda a decrescere con il tempo facendo di conseguenza rallentare la crescita dello sferoide stesso. Ma quale equazione usare? Partiamo dalla più semplice e consideriamo un decadimento esponenziale del tasso di crescita. Si noti che non ci chiediamo perché il tasso decresca nel tempo ma che ci limitiamo a tentare una rappresentazione del suo decadimento. Questo è il senso dell'approccio fenomenologico.

Per quanto detto qui sopra, dunque, assumiamo che: 1. la crescita del volume sia esponenziale; 2. che il rateo di crescita decresca esponenzialmente nel tempo. Tradotto in formule ciò equivale a considerare il seguente sistema di equazioni differenziali:

---

$$\begin{cases} \frac{dVol(t)}{dt} = kp(t) \cdot Vol(t) \\ \frac{dkp(t)}{dt} = -\beta \cdot kp(t) \end{cases}$$

dove usiamo il simbolo  $Vol(t)$  per rappresentare la variabile "volume di uno sferoide al tempo  $t$ " e dove abbiamo introdotto il parametro  $\beta$  che deve assumere valori positivi e che descrive il rateo di decadimento esponenziale di  $kp(t)$ .

Per inciso questo modello lo dobbiamo al matematico Benjamin Gompertz che alla fine dell'ottocento lo ideò per descrivere e studiare le curve di mortalità della popolazione umana.

Si tratta ora di:

1. risolvere il sistema e studiarne le proprietà matematiche
2. cercare di *fittare* i dati sperimentali forniti e calcolare così parametri biologici interessanti
3. cercare di capire se il modello cattura tutta l'informazione contenuta nei dati
4. eventualmente cercare di modificare il modello
5. ecc....

Questo spetta a voi. Buon divertimento!



---

### 3. L'elaborazione matematica e statistica

a cura di L. Corso, F. Castelli e G. Pezzo

L'equazione di Gompertz si riferisce alla crescita di una popolazione di individui, mentre il sistema presentato da Roberto Chignola sostituisce il volume del tumore ( $Vol(t)$ ) alla numerosità delle cellule tumorali (individui).  $Vol(t)$  viene calcolato con formula empirica:  $Vol(t) \approx (4 \cdot \pi / 3) \cdot r^3$ , ove  $r$  è il raggio medio (media geometrica) dei due raggi massimo e minimo, rispettivamente  $D/2$  e  $d/2$  calcolati sperimentalmente sul tumore; cioè:

$$r = \frac{1}{2} \sqrt{D \cdot d}.$$

Nel seguito si considererà la popolazione delle cellule tumorali  $X(t)$ , non il volume del tumore. È una sostituzione che non produrrà alcun effetto sostanziale. I risultati finali, infatti, sono equivalenti a meno di una costante, in quanto:  
 $X(t) = Vol(t) / volume\ cellula$ .

Riprendiamo l'ultimo sistema di equazioni differenziali proposto, con l'accortezza di sostituire la funzione  $Vol(t)$ , con la funzione  $X(t)$ .

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

La prima delle equazioni di (1) afferma 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  $X(t)$ .

Come spiegato nel paragrafo precedente, la seconda equazione di (1) tiene conto dell'effetto di smorzamento del fattore di crescita di  $X(t)$ . Se  $\beta > 0$  allora  $dKp(t) / dt$  decresce al variare di  $t$ .

Sostituendo nella seconda equazione il valore di  $Kp(t)$  trovato nella prima, si ha:

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

Integrando entrambi i membri della seconda equazione e sostituendo nella prima:

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

Se  $C = -\ln K$  si ottiene

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

da cui:

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

La prima equazione del sistema (2) è esattamente l'equazione di Gompertz che ora andiamo a risolvere.

### 3.1 Ricerca della soluzione generale in forma chiusa

L'equazione:

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

dove

$X(t)$  è la quantità di una popolazione di individui al tempo  $t$ ,  
 $\beta$  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,

risulta essere una equazione differenziale del primo ordine a variabili separabili.  
 Risolviamo:

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

$$\frac{dX(t)}{X(t) \cdot \ln\left(\frac{X(t)}{K}\right)} = -\beta \cdot dt \quad , \text{ da cui, integrando entrambi i membri:}$$

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

Ponendo:

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

si ottiene:

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

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

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

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

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

$$X(t) = K \cdot e^{-C \cdot e^{-\beta t}} \quad (4)$$

Questa relazione è la forma chiusa generale risolutiva di (3) che, come si vede, dipende dai parametri  $K$ ,  $\beta$  e  $C$

### 3.2 Stima dei parametri con discretizzazione del modello

Il problema ora è quello di stimare i parametri del modello (3) o (4) 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. In pratica la discretizzazione consiste nella sostituzione della derivata con il rapporto incrementale:

$$\frac{\Delta X_t}{\Delta t} = \beta \cdot X_t \cdot \text{Ln} \left[ \left( \frac{X_t}{K} \right)^{-1} \right] \quad (5)$$

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

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

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

$$X_{t+1} = X_t \left( 1 + \beta \cdot \text{Ln} \left[ \left( \frac{K}{X_t} \right) \right] \right) \quad (6)$$

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

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

$$Y_t = 1 + \beta \cdot \text{Ln} K - \beta \cdot \text{Ln} X_t \quad (8)$$

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

$$Y_t = A + B \cdot \text{Ln} X_t \quad (9)$$

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,  $\beta$ ,  $K$  e  $Kp(t)$ . Vediamo come:

$$S(A, B) = \sum_{j=1}^n (Y_j - \hat{Y}_j)^2 = \sum_{j=1}^n (Y_j - A - B \cdot \text{Ln}X_j)^2 \quad (10)$$

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

$$\begin{cases} \sum_{j=1}^n (Y_j - A - B \cdot \text{Ln}X_j) \cdot (-1) = 0 \\ \sum_{j=1}^n (Y_j - A - B \cdot \text{Ln}X_j) \cdot (-\text{Ln}X_j) = 0 \end{cases} \quad (11)$$

(11) è equivalente alla seguente forma matriciale:

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

Risolto quest'ultimo sistema si trovano  $A$  e  $B$ .

Da  $B$  si determina  $\beta$  e quindi  $K$ ; successivamente è possibile determinare i valori di  $X_t$  o con la (7) oppure con il modello generale (4). In quest'ultimo caso esiste il problema della determinazione di  $C$  che viene risolto con la condizione iniziale: al tempo  $t = 0$ ,  $X(0) = K e^{-C}$ .

Sappiamo sia  $X(0)$ , numerosità degli individui all'inizio del processo, sia  $K$ . Sarà quindi  $C = \text{Ln} K - \text{Ln}X(0)$ .

Se studiamo il limite di (5) per  $t \rightarrow \infty$ , si ottiene che  $X(t) \rightarrow K$  coerentemente ai significati dell'argomento.

### 3.3 Verifica della bontà dell'accostamento

Una volta che si è selezionato dalla famiglia di modelli (3) quello il cui grafico si accosta ai punti sperimentali rispettando la condizione dei minimi quadrati, si passa alla verifica della bontà dell'accostamento fatto.

Si può procedere con due diversi livelli di approfondimento:

1. Calcolare l'indice di accostamento quadratico relativo, cioè il coefficiente di variazione relativo al modello stimato;
2. Studiare l'autocorrelazione dei residui e verificare se il suo andamento al variare del passo  $k$ , è smorzato oppure no.

---

### Indice di accostamento quadratico relativo [B. 3]

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

$$I_2 = \frac{1}{M(\hat{X}_j)} \cdot \sqrt{\frac{\sum (X_j - \hat{X}_j)^2}{n}} \quad (12)$$

che è il coefficiente di variazione relativo al modello e quindi il rapporto tra la deviazione standard dei dati sperimentali rispetto ai valori teorici e la media aritmetica calcolata sui valori teorici del modello.

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.

### Analisi dei residui

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, ovvero delle differenze:

$$a_j = X_j - \hat{X}_j. \quad (13)$$

Perché il modello stimato sia significativamente accettabile questi residui devono avere alcune importanti proprietà:

1.  $M(a_j) = 0$  ; (14)

2.  $V(a_j) = \sigma_a^2$  ; (15)

3.  $c_k(a_j)$ , 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:

$$c_k(a_j) = \frac{1}{n} \sum_{j=1}^{n-k} (a_{j+k} - M(a_j)) \cdot (a_j - M(a_j)) \quad \forall k \in N \quad (16)$$

La relazione (16), se si dimostra vera la (14), può essere semplificata considerando che  $M(a_j) = 0$  e  $c_k(a_j) = c_k$ .

Si ottiene così:

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

(Quando  $k = 0$ ,  $c_k = \sigma_a^2$ ). Se non è possibile dimostrare che  $M(a_j) = 0$ , allora si deve tenere il valore calcolato sperimentalmente come stima del valore incognito della popolazione.

Si deve verificare che la funzione di autocovarianza sperimentale converga a 0 rapidamente. Si può lavorare con la funzione di autocorrelazione.

$$\text{In tal caso } r_k = \frac{c_k}{c_0} \quad , \quad \forall k \in N \quad (18)$$

Mediante la (18) 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 esistente tra  $X(t)$  e  $t$ . Un correlogramma non immediatamente degradante verso lo zero farebbe intendere che non tutto il fenomeno è stato colto dal modello teorico stimato.

---

#### 4. Una proposta di discretizzazione

di Simone Zuccher

Nei paragrafi precedenti ci siamo occupati di vedere quanto un insieme di dati sperimentali possono essere compiutamente rappresentati da un modello di equazioni differenziali ordinarie. Dopo aver riscritto le due equazioni come un'unica equazione differenziale che ammette una soluzione in forma chiusa, ci siamo preoccupati di trovare i tre coefficienti  $C$ ,  $K$  e  $\beta$  tali per cui il modello linearizzato sia il *miglior fit* dei dati nel senso dei minimi quadrati.

Fatta l'analisi del residuo, però, alcuni hanno notato che forse il modello di Gompertz non è poi così appropriato per descrivere il fenomeno della crescita degli sferoidi tumorali.

Quindi ci si può chiedere:

- Esiste un modello migliore ?
- Che considerazioni posso fare per cambiare le equazioni differenziali del modello ?
- Se cambio le equazioni differenziali, esiste ancora una soluzione in forma chiusa del modello ? Se non esiste cosa faccio ?

In questo breve intervento si vuole proporre uno strumento numerico per ottenere la soluzione approssimata di un problema differenziale, nel caso non esista la soluzione in forma chiusa.

La scelta di tale strumento cade sulla risoluzione numerica delle equazioni differenziali lineari del primo ordine con il metodo di Eulero (esplicito ed implicito) il cui impiego prevede, in pratica, l'impostazione e la successiva risoluzione di un problema di Cauchy. Tale problema è definito dal sistema di equazioni

$$\begin{cases} \frac{dX(t)}{dt} = f(t, X) \\ X(t_0) = X_0 \end{cases} \quad (4.1)$$

in cui  $dX(t)/dt = f(t, X)$  è l'equazione differenziale non necessariamente lineare e  $X(t_0) = X_0$  è la condizione iniziale. Nel caso dell'equazione proposta da Gompertz il problema di Cauchy che si genera è il seguente:

$$\begin{cases} \frac{dX(t)}{dt} = \beta X(t) \ln\left(\frac{K}{X(t)}\right) \\ X(t_0) = X_0 \end{cases} \quad (4.2)$$

Se ora si introduce la variabile  $kp(t)$  legata a  $X(t)$  dalla relazione

$$kp(t) = \beta \ln\left(\frac{K}{X(t)}\right) \quad (4.3)$$

il precedente sistema diventa

$$\begin{cases} \frac{dX(t)}{dt} = Kp(t) X(t) \\ X(t_0) = X_0 \end{cases} \quad (4.4)$$

Questo ultimo sistema a sua volta potrà essere riscritto nella forma che segue:

$$\begin{cases} \frac{dKp(t)}{dt} = -\beta Kp(t) \\ Kp(t_0) = \beta \ln\left(\frac{K}{X_0}\right) \end{cases} \quad (4.5)$$

in cui la prima equazione si ottiene differenziando  $Kp(t)$  definito da (4.3) e sostituendo a  $\frac{dX(t)}{dt}$  il secondo membro della prima equazione di (4.4) e la seconda equazione tiene conto delle condizioni iniziali espresse in (4.4)

Al fine di ottenere delle equazioni algebriche a partire dalle equazioni differenziali, si esprimono le derivate prime sotto forma di rapporti incrementali, ossia:

$$\frac{dX(t)}{dt} \approx \frac{X_{i+1} - X_i}{\Delta t}, \quad \Delta t_i = t_{i+1} - t_i.$$

#### Metodo di Eulero esplicito

Rappresenta la via più immediata per la discretizzazione dei problemi (4.4) e (4.5). Relativamente al problema (4.4) si avrà :

$$\begin{cases} \frac{Kp_{i+1} - Kp_i}{\Delta t} = -\beta Kp_i \\ Kp_0 = \beta \ln\left(\frac{K}{X_0}\right) \end{cases} \quad (4.6)$$

da cui si ottiene

$$\begin{cases} Kp_{i+1} = Kp_i (1 - \beta \Delta t) \\ Kp_0 = \beta \ln\left(\frac{K}{X_0}\right) \end{cases}.$$

Analogamente, tenendo conto di (4.4), il problema (4.5) si può discretizzare nella forma seguente:

$$\begin{cases} X_{i+1} = X_i (1 + Kp_i \Delta t) \\ X_0 = X(t_0) \end{cases}$$

---

### Metodo di Eulero implicito

Talvolta il metodo sopra esposto può produrre soluzioni del tutto sbagliate con oscillazioni dovute a problemi di stabilità numerica.

Al contrario il metodo di Eulero Implicito assicura maggiore stabilità e minori problemi.

La differenza tra i due metodi sta nella discretizzazione della parte destra dell'equazione.

Nel primo caso è valutata la tempo  $t_i$ , nel secondo al tempo  $t_{i+1}$ .

In particolare la (4.6) diventa

$$\begin{cases} \frac{Kp_{i+1} - Kp_i}{\Delta t} = -\beta Kp_{i+1} \\ Kp_0 = \beta \ln\left(\frac{K}{X_0}\right) \end{cases}$$

da cui si ottiene

$$\begin{cases} Kp_{i+1} = \frac{Kp_i}{1 + \beta \Delta t} \\ Kp_0 = \beta \ln\left(\frac{K}{X_0}\right) \end{cases}$$

Analogamente per il problema (4.5) si ottiene

$$\begin{cases} X_{i+1} = \frac{X_i}{1 - Kp_{i+1} \Delta t} \\ X_0 = X(t_0) \end{cases}$$

### Proposta di lavoro

Si risolvano i problemi differenziali (4.4) e (4.5) utilizzando i metodi di Eulero esplicito e implicito. Si confrontino le due soluzioni ottenute numericamente con quella esatta (in forma chiusa). Sono simili/vicine? Quale delle due si comporta meglio? Come si può intervenire per migliorare la soluzione numerica?



---

## 5. L'elaborazione dei dati sperimentali con "Mathematica"

a cura degli alunni con il contributo dei docenti L. Corso, F. Castelli, G. Pezzo

L'elaborazione consiste in:

1. trovare i parametri del modello di Gompertz
  - importare i dati dai file
  - creare la lista dei punti di valore medio tra i valori di tutti gli esperimenti
  - trovare la curva interpolante per ottenere i parametri A e B
  - trovare i parametri  $\beta, k, c$tracciare il grafico dei punti e della curva interpolante.
2. trovato il modello, creare la lista dei dati teorici (ovvero i dati ottenuti utilizzando il modello)
3. trovare i residui (differenza tra i dati teorici e sperimentali) e studiarne media e deviazione standard
4. studiare la correlazione tra i residui per verificare, anche graficamente, che non esiste relazione tra gli scarti, e quindi che il modello è il migliore
5. stimare i dati teorici col metodo numerico di Eulero, sia esplicito che implicito

Viene ora presentato il lavoro eseguito dagli alunni Marco Banterle e Mattia Battiston, senza nulla togliere agli altri lavori che sono altrettanto meritevoli.

---

## Codice Programma e relativo output

### PARTE 1:Trovare i parametri del modello di Gompertz:

(\*Importare i dati da file e preparare le liste per studiarli\*)

```
datiSperimentali={};
scarti={};
datiInput=Import["C:\dati sperimentali\s9l30.txt","Table"];
For[i=1,i<=Length[datiInput],i++,
  datiSperimentali=Append[datiSperimentali,datiInput[[i,2]]];
  datiTeorici=Append[datiTeorici,datiInput[[i,3]]];
  scarti=Append[scarti,datiInput[[i,4]]];
];
listaY={};
For[i=1,i<Length[datiSperimentali],i++,
  listaY=Append[listaY,datiSperimentali[[i+1]]/datiSperimentali[[i]]];
];
listaY=Append[listaY,0];
listaLnX={};
For[i=1,i<=Length[datiSperimentali],i++,
  listaLnX=Append[listaLnX,Log[datiSperimentali[[i]]]];
];
tabella=Table[{i-1,datiSperimentali[[i]],
datiTeorici[[i]],scarti[[i]],listaLnX[[i]],(listaLnX[[i]])^2, listaY[[i]]},
{i,1,Length[datiSperimentali]};
TableForm[tabella,TableHeadings->{None,{"t", "Xt", "X̂t", "at = Xt - X̂t",
"ln Xt", "ln Xt2", "Y"}}}]
```

---

*Dati*

t	X <sub>t</sub>	X <sub>t</sub>	a <sub>t</sub> =X <sub>t</sub> - X <sub>t</sub>	ln X <sub>t</sub>	ln X <sub>t</sub> <sup>2</sup>	Y
0	3.57	0.811011	2.75899	1.27257	1.61942	2.06443
1	7.37	1.86082	5.50918	1.99742	3.98968	1.47931
2	10.9025	3.76842	7.13408	2.38899	5.70728	1.32401
3	14.435	6.8635	7.5715	2.66966	7.12706	1.48944
4	21.5	11.4232	10.0768	3.06805	9.41295	1.33023
5	28.6	17.6104	10.9896	3.35341	11.2453	1.2986
6	37.14	25.4385	11.7015	3.61469	13.066	1.13032
7	41.98	34.7698	7.21016	3.73719	13.9666	1.25989
8	52.89	45.3432	7.5468	3.96821	15.7467	1.09293
9	57.805	56.8181	0.986882	4.05708	16.4599	1.08503
10	62.72	68.8229	- 6.10288	4.13868	17.1287	1.15673
11	72.55	80.9961	- 8.44615	4.28428	18.355	1.21296
12	88.	93.0171	- 5.01708	4.47734	20.0465	1.2
13	105.6	104.623	0.9772	4.65966	21.7124	0.913826
14	96.5	115.615	- 19.1148	4.56954	20.8807	1.0943
15	105.6	125.857	- 20.2572	4.65966	21.7124	1.09896
16	116.05	135.27	- 19.2198	4.75402	22.6007	1.09005
17	126.5	143.819	- 17.3186	4.84024	23.4279	1.16522
18	147.4	151.505	- 4.10535	4.99315	24.9315	1.
19	147.4	158.359	- 10.9586	4.99315	24.9315	1.25645
20	185.2	164.425	20.7754	5.22144	27.2634	0.928726
21	172.	169.761	2.23908	5.14749	26.4967	1.15698
22	199.	174.431	24.5691	5.2933	28.0191	1.
23	199.	178.5	20.5003	5.2933	28.0191	1.
24	199.	182.031	16.9687	5.2933	28.0191	1.
25	199.	185.087	13.9132	5.2933	28.0191	1.07337
26	213.6	187.723	25.8768	5.36411	28.7736	0.931648
27	199.	189.993	9.00716	5.2933	28.0191	1.
28	199.	191.943	7.0572	5.2933	28.0191	1.
29	199.	193.615	5.38466	5.2933	28.0191	1.
30	199.	195.048	3.95211	5.2933	28.0191	1.
31	199.	196.273	2.7266	5.2933	28.0191	1.
32	199.	197.321	1.67926	5.2933	28.0191	1.07337
33	213.6	198.215	15.385	5.36411	28.7736	0.931648
34	199.	198.978	0.0219677	5.2933	28.0191	1.07337
35	213.6	199.629	13.9713	5.36411	28.7736	0.96676
36	206.5	200.183	6.31686	5.3303	28.4121	0.965617
37	199.4	200.655	- 1.25548	5.29531	28.0403	0.967904
38	193.	201.058	- 8.05769	5.26269	27.6959	0.959585
39	185.2	201.4	- 16.2001	5.22144	27.2634	1.07451
40	199.	201.691	- 2.69142	5.2933	28.0191	1.
41	199.	201.939	- 2.93932	5.2933	28.0191	1.07337
42	213.6	202.15	11.4498	5.36411	28.7736	1.
43	213.6	202.33	11.2705	5.36411	28.7736	1.
44	213.6	202.482	11.118	5.36411	28.7736	1.
45	213.6	202.612	10.9883	5.36411	28.7736	1.
46	213.6	202.722	10.8781	5.36411	28.7736	0.867041
47	185.2	202.816	- 17.6157	5.22144	27.2634	1.15335

---

t	$X_t$	$X_t$	$a_t = X_t - X_t$	$\ln X_t$	$\ln X_t^2$	Y
48	213.6	202.895	10.7047	5.36411	28.7736	0.931648
49	199.	202.963	- 3.96303	5.2933	28.0191	1.07337
50	213.6	203.021	10.5794	5.36411	28.7736	0.982912
51	209.95	203.069	6.88052	5.34687	28.589	0.982615
52	206.3	203.111	3.18895	5.32933	28.4018	0.964615
53	199.	203.146	- 4.14638	5.2933	28.0191	1.
54	199.	203.176	- 4.17639	5.2933	28.0191	1.07337
55	213.6	203.202	10.3981	5.36411	28.7736	0.931648
56	199.	203.224	- 4.22358	5.2933	28.0191	1.
57	199.	203.242	- 4.242	5.2933	28.0191	1.
58	199.	203.258	- 4.25765	5.2933	28.0191	1.
59	199.	203.271	- 4.27095	5.2933	28.0191	1.
60	199.	203.292	- 4.29186	5.2933	28.0191	1.07337
61	213.6	203.3	10.3	5.36411	28.7736	0.867041
62	185.2	203.307	- 18.107	5.22144	27.2634	1.
63	185.2	203.313	- 18.1128	5.22144	27.2634	1.
64	185.2	203.318	- 18.1178	5.22144	27.2634	1.
65	185.2	203.322	- 18.1221	5.22144	27.2634	1.
66	185.2	203.326	- 18.1257	5.22144	27.2634	1.15335
67	213.6	203.329	10.2712	5.36411	28.7736	1.
68	213.6	203.331	10.2686	5.36411	28.7736	0.931648
69	199.	203.334	- 4.33361	5.2933	28.0191	1.07337
70	213.6	203.335	10.2645	5.36411	28.7736	0.951311
71	203.2	203.337	- 0.137097	5.31419	28.2406	0.948819
72	192.8	203.338	- 10.5385	5.26165	27.685	0.892116
73	172.	203.34	- 31.3396	5.14749	26.4967	1.15698
74	199.	203.341	- 4.34059	5.2933	28.0191	0.930653
75	185.2	203.341	- 18.1414	5.22144	27.2634	1.07451
76	199.	203.342	- 4.34214	5.2933	28.0191	1.
77	199.	203.343	- 4.34274	5.2933	28.0191	0

---

**(\*Trovare la curva interpolante per ottenere i parametri A e B\*)**

```
listaPunti={};
For[i=1,i≤Length[listaLnX],i++,
  listaPunti = Append[listaPunti, {listaLnX[[i]], listaY[[i]]}];
];
n=Length[listaPunti];
soluzioni=
{a,b}/.NSolve[{n*a +

$$\left(\sum_{j=1}^n \text{listaPunti}[[j,1]]\right) * b = \sum_{j=1}^n \text{listaPunti}[[j,2]], \left(\sum_{j=1}^n \text{listaPunti}[[j,1]]\right) * a +$$


$$\left(\sum_{j=1}^n \text{listaPunti}[[j,1]]\right)^2 * b = \sum_{j=1}^n \text{listaPunti}[[j,2]] * \text{listaPunti}[[j,1]], (a,b)];
A=soluzioni[[1,1]];
B=soluzioni[[1,2]];$$

```

**(\*Trovare i parametri del modello ( $\beta$ , k, c) e la sua equazione\*)**

```
 $\beta$ =-B;
Print["Parametro  $\beta$ : ",  $\beta$ ];
 $k = e^{\frac{A-1}{\beta}}$ ;
Print["Parametro k: ",k];
c= - Log[datiSperimentali[[1]]/k];
Print["Parametro c: ",c];

Print["Eq. modello: "];

modello[t_] = k * e-c*e- $\beta$ *t
```

*Output*

Parametro  $\beta$ : 0.18076

Parametro k: 180.991

Parametro c: 3.92588

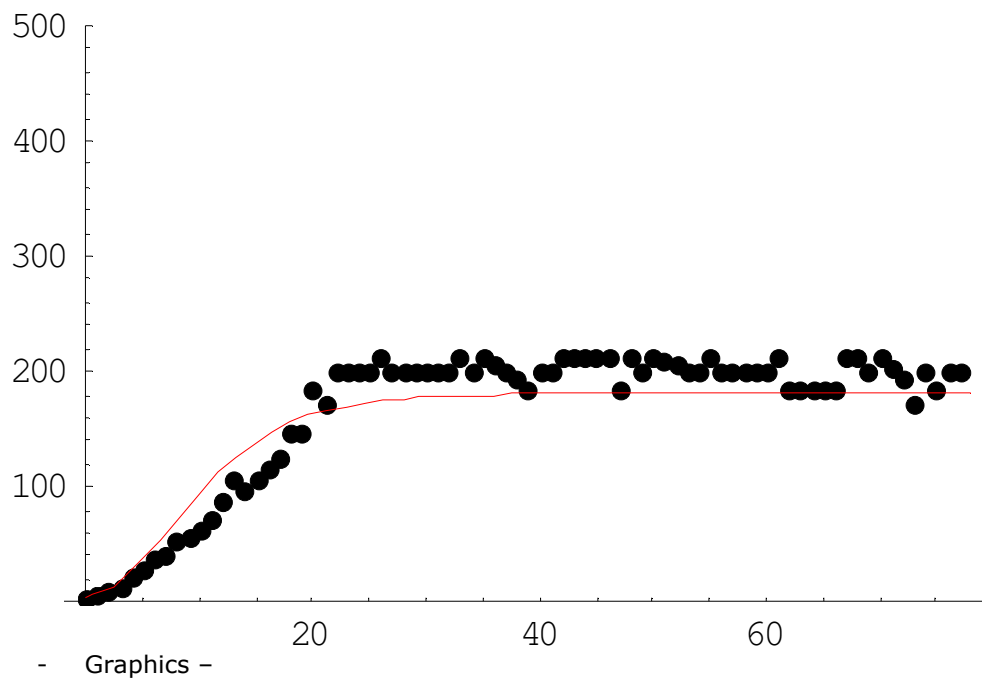
Eq. modello:  $180.991 \tilde{a}^{-3.92588 \tilde{a}^{0.18076 t}}$

---

**(\*Tracciare il grafico del modello\*)**

```
listaPuntiEsperimento={};  
For[i=1,i≤Length[datiSperimentali], i++,  
  listaPuntiEsperimento = Append[listaPuntiEsperimento,{i-1,  
  datiSperimentali[[i]]}]  
];  
graficoPunti =  
ListPlot[listaPuntiEsperimento,PlotStyle → PointSize[0.02],  
  DisplayFunction → Identity];  
graficomodello = Plot[modello[t],{t,0,Length[listaPuntiEsperimento]},  
  PlotStyle → RGBColor[1,0,0], DisplayFunction → Identity];  
Show[graficoPunti, graficomodello, DisplayFunction → $DisplayFunction,  
  PlotRange → {0,500}]
```

*Output*



---

**(\*Stima dell'errore mediante Indice di Pearson\*)**

```
listaPuntiModello={};  
For[i=1,i≤Length[datiSperimentali],i++,  
  listaPuntiModello=Append[listaPuntiModello,modello[i-1]];  
];  
Print["Indice di Pearson:"]
```

$$\frac{1}{\sum_{j=1}^{\text{Length}[\text{listaPuntiModello}]} \text{listaPuntiModello}[[j]]} * \sqrt{\frac{\sum_{j=1}^{\text{Length}[\text{listaPuntiEsperimento}]} (\text{datiSperimentali}[[j]] - \text{listaPuntiModello}[[j]])^2}{\text{Length}[\text{listaPuntiEsperimento}]}}$$

Indice di Pearson:

0.147581

**PARTE 2: trovato il modello, creare la lista dei dati teorici:**

**(\*Confronto dei dati sperimentali con quelli teorici\*)**

```
listaTabella={};  
datiMioModello={};  
listaScarti={};  
For[i=1,i≤Length[listaPunti],i++,  
  datiMioModello=Append[datiMioModello,modello[i-1]];  
  listaScarti=Append[listaScarti,datiSperimentali[[i]]-  
  datiMioModello[[i]];  
  listaTabella=Append[listaTabella,  
{i-1,datiMioModello[[i]],datiSperimentali[[i]],listaScarti[[i]]}];  
];  
TableForm[tabella,TableHeadings→{None,{"t", "Xt", "X̂t", "at = Xt - X̂t",  
"ln Xt", "ln Xt2", "Y"}}]
```

---

*Dati teorici e dati sperimentali*

t	X <sub>t</sub>	X <sub>t</sub>	a <sub>t</sub> =X <sub>t</sub> - X <sub>t</sub>	t	X <sub>t</sub>	X <sub>t</sub>	a <sub>t</sub> =X <sub>t</sub> - X <sub>t</sub>
0	3.57	3.57	- 4.44089 · 10 <sup>-16</sup>	48	180.87	213.6	32.7303
1	6.83301	7.37	0.536994	49	180.89	199.	18.1102
2	11.7471	10.9025	- 0.84464	50	180.906	213.6	32.6935
3	18.4646	14.435	- 4.02956	51	180.92	209.95	29.0296
4	26.9319	21.5	- 5.43191	52	180.932	206.3	25.3679
5	36.9052	28.6	- 8.3052	53	180.942	199.	18.0582
6	48.0046	37.14	- 10.8646	54	180.95	199.	18.0501
7	59.7853	41.98	- 17.8053	55	180.957	213.6	32.6433
8	71.8033	52.89	- 18.9133	56	180.962	199.	18.0377
9	83.6643	57.805	- 25.8593	57	180.967	199.	18.0329
10	95.0509	62.72	- 32.3309	58	180.971	199.	18.029
11	105.733	72.55	- 33.1826	59	180.974	199.	18.0257
12	115.561	88.	- 27.5614	60	180.977	199.	18.023
13	124.461	105.6	- 18.8609	61	180.979	213.6	32.6207
14	132.411	96.5	- 35.9113	62	180.981	185.2	4.21876
15	139.434	105.6	- 33.8344	63	180.983	185.2	4.21717
16	145.581	116.05	- 29.5306	64	180.984	185.2	4.21584
17	150.917	126.5	- 24.4174	65	180.985	185.2	4.21473
18	155.521	147.4	- 8.12112	66	180.986	185.2	4.2138
19	159.471	147.4	- 12.0709	67	180.987	213.6	32.613
20	162.844	185.2	22.3558	68	180.988	213.6	32.6124
21	165.714	172.	6.2857	69	180.988	199.	18.0118
22	168.148	199.	30.8515	70	180.989	213.6	32.6114
23	170.207	199.	28.7926	71	180.989	203.2	22.211
24	171.945	199.	27.0548	72	180.989	192.8	11.8107
25	173.409	199.	25.5908	73	180.99	172.	- 8.98956
26	174.641	213.6	38.9593	74	180.99	199.	18.0102
27	175.675	199.	23.3248	75	180.99	185.2	4.21004
28	176.543	199.	22.4567	76	180.99	199.	18.0099
29	177.271	199.	21.7288	77	180.99	199.	18.0098
30	177.881	199.	21.119				
31	178.392	199.	20.6085				
32	178.819	199.	20.1812				
33	179.176	213.6	34.4239				
34	179.475	199.	19.525				
35	179.725	213.6	33.8752				
36	179.934	206.5	26.5665				
37	180.108	199.4	19.292				
38	180.254	193.	12.7463				
39	180.375	185.2	4.82464				
40	180.477	199.	18.523				
41	180.562	199.	18.4381				
42	180.633	213.6	32.9672				
43	180.692	213.6	32.9081				
44	180.741	213.6	32.8587				
45	180.783	213.6	32.8174				
46	180.817	213.6	32.783				
47	180.846	185.2	4.35426				



---

### PARTE 3: Trovare i residui (la differenza) tra i dati teorici e sperimentali e analizzarli

#### (\*Trovare i residui\*)

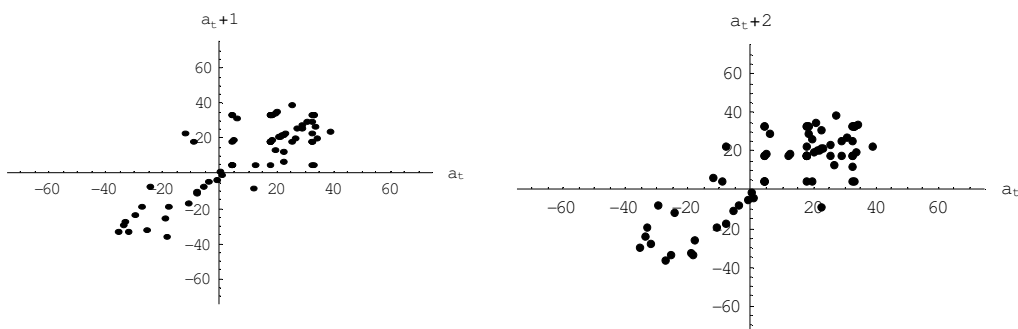
```
listaPuntiScarti = {};  
listaFinita = {};  
For[j=1, j<Length[listaScarti], j++,  
  listaPuntiScarti = {};  
  For[i=1, i≤Length[listaScarti]-j, i++,  
    listaPuntiScarti =  
    Append[listaPuntiScarti, {listaScarti[[i]], listaScarti[[i+j]]}];  
  ];  
  listaFinita = Append[listaFinita, listaPuntiScarti];  
];
```

### PARTE 4: Analisi dei Residui

#### (\*Tracciare i grafici delle relazioni tra gli scarti al passo $n$ e $n+k$ \*)

```
Do[ListPlot[listaFinita[[k]], PlotStyle→PointSize[0.018],  
  AxesLabel→{"at", "at+"<>ToString[k]},  
  PlotRange→{{-75, +75}, {-75, +75}}];, {k, 1, (Length[listaScarti]-1), 1}]
```

(Rappresentiamo solamente 2 dei 75 grafici)



---

**(\*Trovare la Media e verificare che sia Nulla\*)**

```
mediaScarti=Mean[listaScarti];  
Print["Media degli scarti =M(at) = ",mediaScarti]  
  
listaMedie={};  
For[i=1,i≤Length[listaScarti],i++,  
  listaTemp={};  
  For[j=i,j≤Length[listaScarti],j++,  
    listaTemp=Append[listaTemp,listaScarti[[j]]];  
  ];  
  listaMedie=Append[listaMedie,Mean[listaTemp]];  
];
```

Media degli scarti= $M(a_t)=11.135$

(la media non è nulla e quindi il modello non è 'buono')

**(\*Trovare l'Autocovarianza e rappresentarla graficamente\*)**

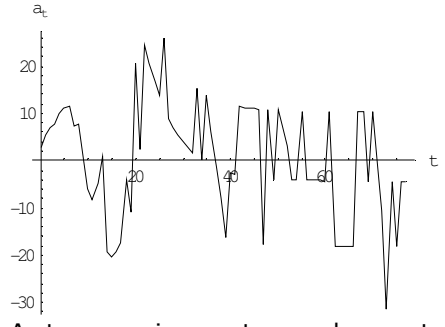
```
n=Length[listaScarti];  
listaAutoConMedie={};  
listaAuto={};  
CZeroConMedie=  
 $\frac{1}{n} * \sum_{j=1}^{n-0} (listaScarti[[j]] - mediaScarti) * (listaScarti[[j+0]]);$   
CZeroNoMedie  $\frac{1}{n} * \sum_{j=1}^{n-0} listaScarti[[j]] - listaScarti[[j]] * listaScarti[[j+0]];$   
Do[listaAutoConMedie=Append[listaAutoConMedie, [k,  
 $\frac{1}{n * CZeroConMedie} *$   
 $\sum_{j=1}^{n-k} (listaScarti[[j]] - mediaScarti) * (listaScarti[[j+k]] - listaMedie[[k]])$ ]];  
listaAuto = Append[listaAuto, {k,  $\frac{1}{n * CZeroConMedie} *$   
 $\sum_{j=1}^{n-k} listaScarti[[j]] * listaScarti[[j+k]]$  }];  
,{k, 1, n}];  
listaScarti={};  
For[i=1,i≤Length[scarti],i++,
```

```

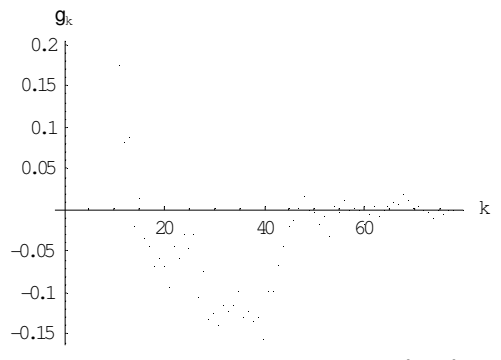
AppendTo[listaScarti,{i-1,scarti[[i]]}];
]
Print["Scarti:"];
ListPlot[listaScarti,AxesLabel→{"t","at"},PlotJoined→True];
Print["Autocovarianze tenendo conto delle medie degli scarti"]
ListPlot[listaAutoConMedie,AxesLabel→{"k","γk"};
Print["Autocovarianze in ipotesi che le medie degli scarti siano 0"]
ListPlot[listaAuto,AxesLabel→{"k","γk"};

```

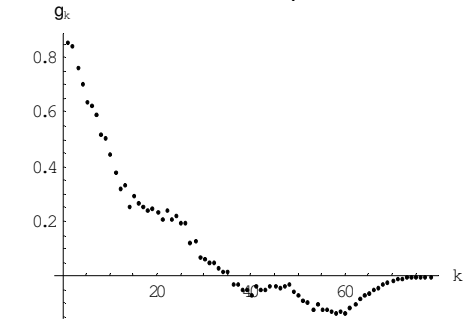
Scarti:



Autocovarianze tenendo conto delle medie degli scarti



Autocovarianze in ipotesi che le medie degli scarti siano 0

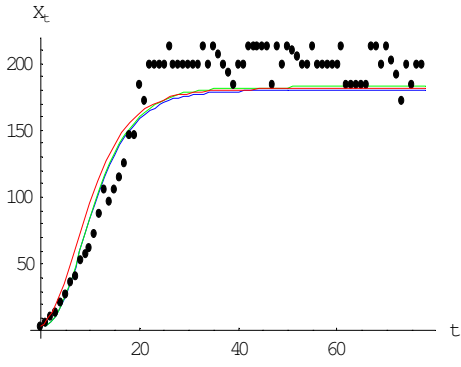


**PARTE 5: Approssimazione del modello tramite metodo numerico di Eulero Implicito ed Esplicito e rappresentazione di entrambi**

```

kpZero=β*Log[  $\frac{k}{\text{datiSperimentali}[[1]]}$  ];
For[i=2,i≤ (Length[datiSperimentali]*1/deltaT),i++,
  listaKpEsplicito=Append[listaKpEsplicito,
    listaKpEsplicito[[i-1]]*(1- β *deltaT)];
  listaKpImplicito=Append[listaKpImplicito,
    listaKpImplicito[[i-1]]/(1+ β*deltaT)];
];
listaModelloEsplicito={datiSperimentali[[1]]};
listaModelloImplicito={datiSperimentali[[1]]};
For[i=2,i≤ (Length[datiSperimentali]*1/deltaT),i++,
  listaModelloEsplicito=Append[listaModelloEsplicito,
    listaModelloEsplicito[[i-1]]*(1+listaKpEsplicito[[i-1]]*deltaT)];
  listaModelloImplicito=Append[listaModelloImplicito,
    listaModelloImplicito[[i-1]]/(1-listaKpImplicito[[i]]*deltaT)];
];
listaModelloEsplicitoG ={};
listaModelloImplicitoG={};
Do[AppendTo[listaModelloEsplicitoG,listaModelloEsplicito[[p]],
  {p,1,Length[listaModelloEsplicito],1/deltaT}];
Do[AppendTo[listaModelloImplicitoG,listaModelloImplicito[[p]],
  {p,1,Length[listaModelloImplicito],1/deltaT}];
graficoEuleroEsplicito=
  ListPlot[listaModelloEsplicitoG,DisplayFunction→Identity,
    PlotStyle→RGBColor[0,0,1],PlotJoined→True];
graficoEuleroImplicito=
  ListPlot[listaModelloImplicitoG,DisplayFunction→Identity,
    PlotStyle→RGBColor[0,1,0],PlotJoined→True];
Show[graficoEuleroEsplicito,graficoEuleroImplicito,graficoPunti,
  graficomodello,DisplayFunction→$DisplayFunction,
  AxesLabel→{"t", " $\hat{X}_t$ " }];
Print["Rosso=modello, Verde = appr. con Eulero implicito, Blu =
  appr. con Eulero esplicito"];

```



---

## Bibliografia

- [B 1] Daisusuke 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, A. Schenetti, E. Chiesa, R. Foroni, S. Sartoris, A. Brendolan, G. Tridente, G. Andrighetto and D.Liberati, Oscillating growth patterns of multicellular tumor spheroids, 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] F. Castelli, In laboratorio con Mathematica, Quaderni del Marconi