

UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI FISICA E ASTRONOMIA
"GALILEO GALILEI"

Corso di Laurea Triennale in Fisica

STUDIO DI MODELLI STATISTICI PER CORRELARE LA
PROBABILITÀ DI INSORGENZA DEL CANCRO CON LE
DIMENSIONI DEGLI ANIMALI: PARADOSSO DI PETO

Laureando

Simone Botticelli

Relatore

Prof. Flavio Seno

ANNO ACCADEMICO 2016/2017

Alla mia famiglia

Indice

1	Leggi di scala	3
1.1	Legge di Kleiber	5
2	Generalità sui tumori	7
2.1	Il paradosso di Peto	9
3	Modelli Matematici	11
3.1	Modello di Calabrese-Shibata	11
3.2	Modello di Wright-Fischer	13
3.2.1	Metodo computazionale	15
3.3	Confronti tra i due modelli	15
4	Dati sperimentali	17
5	Conclusioni	23

Introduzione

La biologia è la scienza delle cose viventi e se partiamo dall'ipotesi che tutte le cose sono fatte di atomi, allora la fisica può aiutarci a capire la vita [1]. La grande complessità dei fenomeni biologici ha spinto grandi fisici come Schrödinger e lo stesso Feynman ad affermare che la comprensione degli stessi avrebbe richiesto l'introduzione di principi fisici e chimici nuovi, estranei all'attuale formulazione di queste discipline. Infatti sia un organismo che tutti i processi importanti dal punto di vista biologico di cui esso è sede, debbono avere una struttura in cui interviene un gran numero di molecole, in modo tale che eventi casuali dovuti a particelle singole non possano acquistare una troppo grande importanza. E' pertanto impensabile scrivere l'equazione di Newton per un sistema che ha migliaia gradi di libertà (ad esempio le cellule di un organismo) e capire come questo sistema possa evolvere nel tempo. Bisogna quindi ragionare in maniera differente: stocastica, numerica o attraverso equazioni più complicate come quella di Stokes. Negli ultimi anni, i computer e i metodi della fisica computazionale hanno permesso di superare almeno in parte le difficoltà nello studio quantitativo dei fenomeni biologici e i successi della meccanica statistica nella trattazione di fenomeni complessi a molti corpi (transizioni di fase, fenomeni fuori dall'equilibrio, ecc.) fanno pensare che metodologie analoghe possano funzionare anche in biologia.

Un esempio di questa comparazione di leggi fisiche e biologiche può essere trovata in oncologia. Nell'ambito della scienza e della ricerca sul cancro esiste un'osservazione chiamata paradosso di Peto: sostanzialmente questo afferma che se l'incidenza del cancro è dovuta all'accumulo in una cellula di diverse mutazioni, gli animali di taglia maggiore e quindi con più cellule, dovrebbero avere un rischio maggiore di sviluppare il cancro durante la loro vita. Pertanto gli animali più grandi dovrebbero vivere di meno rispetto a quelli più piccoli. Eppure non è così e la vita dei mammiferi cresce con la massa. In questo lavoro si studieranno due modelli per predire l'insorgenza del cancro nei mammiferi e vedere se il paradosso possa essere risolto semplicemente modificando alcuni parametri che regolano i processi di divisione cellulare. Infine i risultati ottenuti saranno confrontati con dei dati sperimentali sulla probabilità di insorgenza del cancro raccolti dai ricercatori dell'Università di Padova.

Capitolo 1

Leggi di scala

Chi se non Galileo fu il primo a capire l'importanza che rivestono le leggi di scala nella Natura? Egli scrive così, nel 1863, in "Discorsi e dimostrazioni matematiche intorno a due nuove scienze" :

Or veggino [...] come anco non potrebbe la natura far alberi di smisurata grandezza, poich  i rami loro, gravati dal proprio peso, finalmente si fiaccherebbero; e parimente sarebbe impossibile far strutture di ossa per uomini, cavalli o altri animali, che potessero sussistere e far proporzionatamente gli uffizii loro, mentre tali animali si dovessero aumentare ad altezze immense, se gi  non si togliesse [=prendesse] materia molto pi  dura e resistente della consueta, o non si deformassero tali ossi, sproporzionatamente ingrossandogli, onde poi la figura ed aspetto dell'animale ne riuscisse mostruosamente grosso.

Galileo si interroga sull'effetto di una dilatazione di uno stesso fattore di tutte le dimensioni lineari di un essere vivente e, ad esempio, conclude che negli animali lo spessore e la lunghezza del femore sono tra loro proporzionali: $s \propto l^{3/2}$ [2]. Se ci  non succedesse allora un simile animale gigante non si reggerebbe sotto il suo peso e di nuovo conclude che lo scheletro degli animali di grossa taglia occupa, in proporzione, un maggior volume rispetto a quanto avviene in animali pi  piccoli.

Questo concetto viene oggi espresso tramite l'invarianza di scala, ossia la dilatazione di una determinata legge che continua a valere su svariati ordini di grandezza. Nell'esempio sopra citato, se scaliamo s e l di uno stesso fattore, la relazione di proporzionalit  continua ancora a valere. Tali leggi hanno giocato un ruolo non secondario nello sviluppo di teorie fisiche (ad esempio i fenomeni critici [3]) ed il loro emergere in ambito biologico fa ritenere che possano assumere un ruolo altrettanto importante poich  suggeriscono un principio comune alla base di fenomeni molto diversi i quali avvengono appunto su scale non paragonabili. Inoltre l'universalit  di certe relazioni   osservata anche in altri campi come geografia, economia, linguistica ecc. Da un punto di vista matematico, queste leggi vengono chiamate leggi di potenza e sono espresse dalla seguente formula

$$y(x) = y_0 x^k, \quad (1.1)$$

dove y rappresenta una qualche variabile osservabile, y_0   una costante e x   la variabile indipendente. In ambito biologico le variabili sono dette allometriche ¹ e per esempio x

¹Allometria: accrescimento relativo di un organo o di una parte di organismo rispetto a tutto il corpo.

può rappresentare la massa degli organismi. Per visualizzare queste leggi si fa spesso uso di un grafico "log-log" che permette di linearizzare la relazione e trovare il parametro k come pendenza della retta

$$\log(y(x)) = \log(y_0) + k \cdot \log(x). \quad (1.2)$$

Infine, l'invarianza di scala è espressa dalla seguente relazione:

$$y(cx) = y_0(cx)^k = c^k y(x), \quad (1.3)$$

secondo cui un incremento della variabile x produce semplicemente una variazione di scala della funzione stessa mantenendo inalterata la struttura matematica.

Alcuni esempi di variabili allometriche sono: tasso metabolico, lunghezza di vita, tasso di crescita, battito cardiaco, lunghezza dei genomi [4].

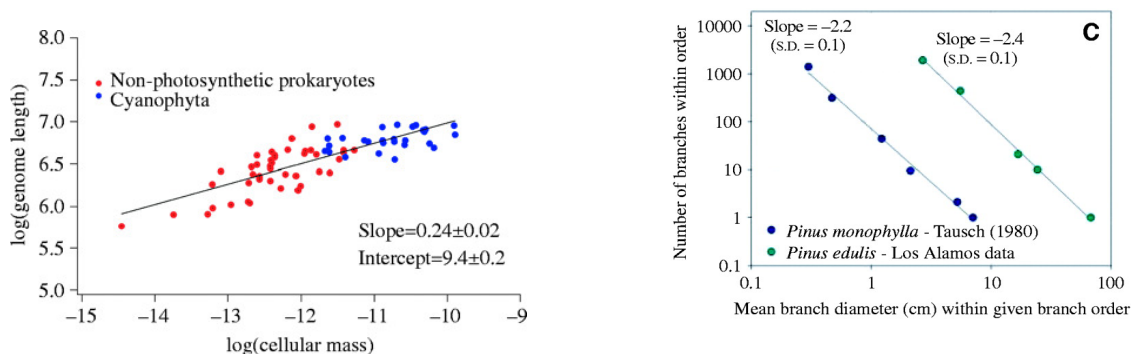


Figura 1.1: Esempi di leggi di scala in biologia [4].

	Tree	Animal
Metabolic rate	3/4	3/4
Characteristic length	1/4	1/3
Characteristic time	1/4	1/4
Characteristic rate	-1/4	-1/4
Organism volume	3/4	1
Transport velocity	0	1/12
Need for pump	No	Yes
Density	1/4	0
Surface area	3/4	2/3
Recirculation network	No	Yes
Fractality of organism	Yes	No
Service volume density	0	-1/4
Trunk radius	3/8	1/3

Figura 1.2: In questa tabella sono riportate alcune variabili allometriche e le rispettive dipendenze teoriche dalla massa per le piante e gli animali [5].

1.1 Legge di Kleiber

Tra tutte le leggi di scala che mettono in relazione le diverse variabili allometriche quella di Kleiber è sicuramente la più importante [6]. Tale legge mette in relazione la quantità di ossigeno consumata da un organismo, cioè il suo metabolismo², alla sua massa corporea:

$$P = B \cdot m^{3/4}. \quad (1.4)$$

P è la potenza metabolica di un determinato organismo, B è chiamata costante tassonomica ed è caratteristica di ogni specie mentre m è la massa. Questa legge, ricavata sperimentalmente per la prima volta dal biologo svizzero Max Kleiber (1893-1976), è così importante perchè esprime un concetto biologico fondamentale: gli esseri viventi, per svolgere le loro funzioni vitali e sopravvivere, minimizzano lo scarto di energia tra quella assorbita e quella spesa.

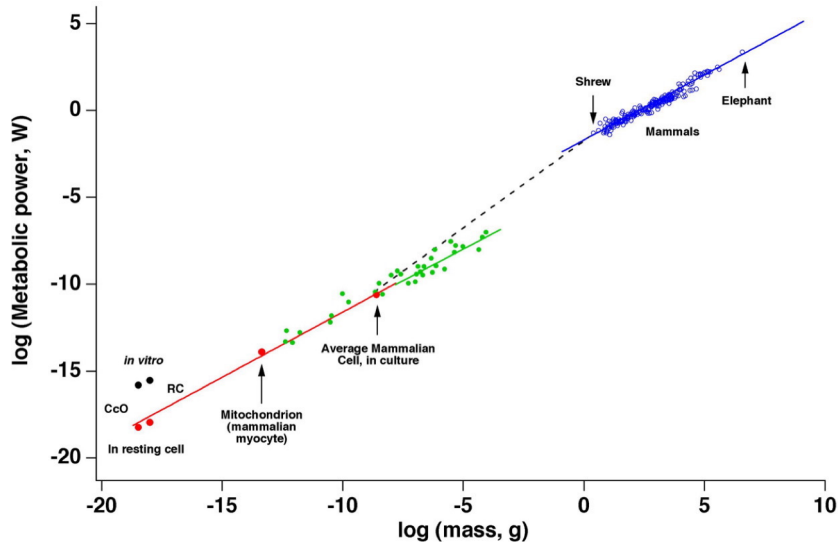


Figura 1.3: Legge di Kleiber [4].

Tale legge si può dimostrare utilizzando un principio di ottimizzazione per il trasporto di ossigeno, quindi di sangue, all'interno del corpo dell'organismo. Fissato il valore m della massa, P deve essere il più grande possibile e con considerazioni fisiche e geometriche si trova la legge 1.4. Una dimostrazione dettagliata si trova in [5] e anche a livello sperimentale la legge è stata largamente confermata.

Nel mondo vegetale e in quello animale rispettivamente, le piante hanno sviluppato le foglie che permettono di catturare più luce possibile mentre negli animali il flusso dei nutrienti e il calore che si crea quando li bruciano per produrre energia devono essere modificati in base alla loro massa. Per questo il cuore pompa il sangue in maniera differente a seconda dell'animale ($\nu_{cardiaca} \propto m^{-1/4}$). L'informazione più significativa che possiamo trarne è quindi che questa legge offre una spiegazione estremamente elegante

²Metabolismo: meccanismo attraverso il quale avviene il trasporto di energia all'interno di un organismo vivente.

di un fatto evolutivo: la forma e il metabolismo di tutti gli esseri viventi, animali e piante, sono stati plasmati dall'esigenza di raggiungere la massima efficienza energetica possibile. Una questione davvero intrigante e misteriosa, ancora non del tutto chiarita, è l'ubiquità della potenza $1/4$, o di un suo multiplo, la quale sembra regolare l'andamento di tutte queste variabili allometriche [4]. D'altronde la vita è tutt'oggi un grande mistero: perchè e come mai certe cose siano andate proprio così e tanti organismi viventi, compresi noi umani, siano qui oggi sono quesiti ai quali non sappiamo rispondere. Non è facile cercare un'ordine quando le variabili in gioco sono così tante, magari semplici principi chimico-fisici e ragionamenti di tipo geometrico, come le leggi di potenza, potrebbero rappresentare un valido strumento per indagare i fenomeni legati alla vita.

Capitolo 2

Generalità sui tumori

Una neoplasia o tumore indica "una massa anormale di tessuto che cresce in eccesso ed in modo sordinato rispetto ai tessuti normali, e che persiste in questo stato dopo la cessazione degli stimoli che hanno indotto il processo", questa è la definizione coniata dall'oncologo Rupert Allan Willis [7]. Occorre distinguere tra tumore e cancro: si parla di tumore quando le funzioni vitali delle cellule vengono meno ma non si ha una degenerazione, mentre nel cancro o tumore maligno, avviene il fenomeno della metastasi che può portare alla morte dell'organismo. Quando un organismo multicellulare si ammala di cancro le sue cellule non sono più in grado di dare il loro normale contributo ai tessuti o agli organi, e invece di concludere il loro normale ciclo vitale, invadono localmente i tessuti: questo fenomeno è detto appunto metastasi. Una volta che tale processo ha preso piede è veramente difficile da arrestare in quanto le cellule tumorali hanno genotipi¹ diversi e si espandono molto velocemente. Un'intera branca della fisica studia come è possibile curare, entro certi limiti, queste malattie attraverso l'uso delle radiazioni. Le ricerche sul cancro hanno messo in evidenza come questo fenomeno emerga e si sviluppi tramite un processo a multistadi: emergenza, sviluppo e metastasi [8].

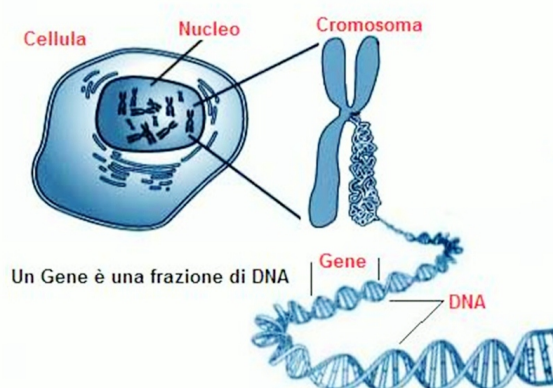


Figura 2.1: Il diagramma mostra in modo semplificato la relazione che intercorre tra gene, DNA e cromosoma.

Tipicamente il cancro ha origine dalla mutazione di alcuni geni che controllano la crescita delle cellule: quando c'è una crescita incontrollata si forma il tumore. Il gene

¹Per genotipo si intende il corredo genetico del DNA della cellula

è l'unità ereditaria fondamentale degli organismi viventi, ossia l'unità fisica e funzionale dell'eredità. I geni trasportano l'informazione da una generazione alla successiva e possiamo pensarli come dei cubi di lato 300 \AA , contenenti milioni di molecole (un numero grande, ma non abbastanza da poter soddisfare le leggi della termodinamica) [9]. Le mutazioni sono lo strumento attraverso il quale un organismo si evolve: tutti noi siamo il risultato di un lungo percorso genetico durato milioni di anni e che continua tutt'ora. Le mutazioni genetiche sono quindi un processo di sviluppo della vita e sono eventi rari, poichè se non lo fossero la probabilità di incontrare mutazioni nocive per l'organismo sarebbe abbastanza elevata e la specie invece di essere migliorata dalla selezione naturale non progredirebbe o addirittura scomparirebbe. I tumori possono svilupparsi anche nell'arco di poco tempo e pertanto devono esserci dei fattori microambientali (inquinamento, fumo, stili di vita ecc.) che aumentano l'errore di copiatura del DNA. Quando una cellula è al termine del suo ciclo vitale si duplica (mitosi o meiosi) e nel fare ciò c'è una certa probabilità che le nuove cellule abbiano ereditato dei geni alterati: in realtà quello che succede è ben più complicato, infatti ogni volta che il patrimonio genetico viene copiato e trasmesso ci sono specifici geni la cui funzione è quella di preservare l'informazione e se questa è danneggiata possono avere un'azione curativa o distruttiva. Questi geni sono detti oncorepressori e proteggono quindi la cellula dall'accumulo di mutazioni tumorali che sono portate dagli oncogeni, ossia quei geni che potenzialmente indirizzano la cellula verso lo sviluppo di un fenotipo² neoplastico. A seconda della gravità della situazione i geni oncorepressori sono in grado di ristabilire le funzioni iniziali del gene mutato oppure se questo è completamente compromesso lo eliminano. Negli articoli [10] e [11] si riportavano dati relativi al tumore al colon rettale e in particolare il numero di cellule presenti all'interno dell'organo era dato dal numero di cellule all'interno di ogni crypt moltiplicato per il numero di crypt totali nel colon stesso. I crypt non sono altro che piccole sacche (organelli) situate sulle pareti del colon all'interno delle quali si trovano le cellule staminali.

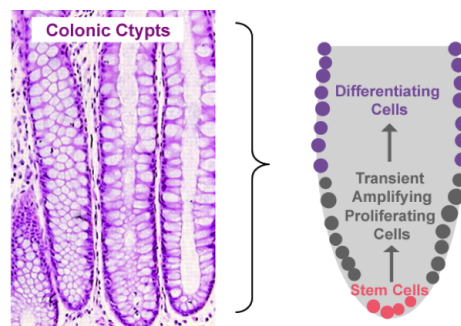


Figura 2.2: In fondo al crypt si trovano le cellule staminali, arrivate in cima ognuna ha la sua funzione specifica.

Non ci interessano altri dettagli specifici ma ci basta saper calcolare il numero totale di cellule nell'organo preso in considerazione e poi il ragionamento fatto per il cancro al colon può essere esteso a qualsiasi altro organo vitale.

²Per fenotipo si intende l'insieme di tutte le caratteristiche manifestate da un organismo vivente, quindi la sua morfologia, il suo sviluppo, le sue proprietà biochimiche e fisiologiche comprensive del comportamento

2.1 Il paradosso di Peto

Il Paradosso di Peto esprime la mancanza di una relazione statistica dell'incidenza del cancro con le dimensioni e la longevità di una determinata specie [8],[10]. L'epidemiologo Peto negli anni '70 si chiese come mai gli umani che sono circa mille volte più grandi di un topo e vivono all'incirca trent'anni di più non presentino un rischio al cancro maggiore. Il paradosso si basa sul fatto che ogni cellula ha una probabilità di mutare e di generare il cancro che aumenta con il tempo e il numero di cellule.

In realtà, quello che succede è che le probabilità di insorgenza dei tumori nelle varie specie sono simili tra loro. Alcune ipotesi per spiegare il paradosso sono le seguenti: in primo gli animali che vivono più a lungo o di taglia maggiore potrebbero aver sviluppato nel corso della loro storia meccanismi preventivi o di soppressione contro questa malattia, inoltre questi animali potrebbero avere tessuti più robusti e quindi il cancro per svilupparsi impiegherebbe più tempo. Un'altra ipotesi è che le dimensioni delle cellule siano diverse a seconda della taglia degli animali.

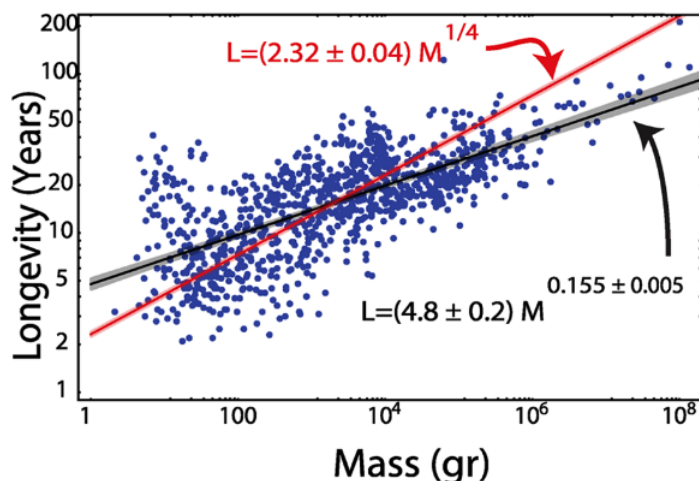


Figura 2.3: In questa figura si vede l'andamento della vita con la massa degli animali [12], che contraddice appunto il paradosso. In questo caso la legge di potenza non rispetta l'andamento teorico.

Nel seguito di questa tesi si cercherà di vedere come dei modelli statistici, introdotti per studiare l'insorgenza del cancro nei mammiferi, possano essere modificati per mostrare l'inconsistenza del paradosso di Peto.

Capitolo 3

Modelli Matematici

Quello che vogliamo fare ora è cercare di capire come i tumori possano originarsi all'interno degli organismi viventi e per fare ciò trascureremo tutti i fattori esterni che possano contribuire alla malattia, ma ci baseremo esclusivamente sul fatto che il tumore si generi a partire da un certo numero di mutazioni genetiche avvenute nelle cellule. I modelli che abbiamo studiato sono due: il modello di Calabrese-Shibata [13] e quello di Wright-Fischer [11]. Il tasso di mutazione cellulare, ossia la probabilità che una cellula abbia una mutazione genetica nella generazione successiva, è stato tenuto costante ma non ci sono ragioni per non supporre che vari di specie in specie. Entrambi i modelli, attraverso alcuni parametri, permettono di calcolare la probabilità d'insorgenza del cancro nei mammiferi.

3.1 Modello di Calabrese-Shibata

Il modello di Calabrese-Shibata [13] è un modello algebrico e afferma che se u è il tasso di mutazione per gene, la probabilità di contrarre il tumore è data da:

$$p = 1 - (1 - (1 - (1 - u)^d)^k)^{Nn}, \quad (3.1)$$

k è il numero di geni che devono mutare affinché il tumore si generi, N è il numero di cellule nei crypt e n rappresenta il numero di crypt per colon. In generale quindi Nn è il numero delle cellule staminali che costituisce l'intera popolazione. Infine, d è il numero di divisioni cellulari totali che si calcola tramite il rate di divisione cellulare pari a $1/4$, ossia una divisione ogni quattro giorni. Ogni termine all'interno delle parentesi rappresenta una determinata probabilità:

$z = 1 - u$	probabilità di non mutare
$y = 1 - z^d$	probabilità di mutare dopo d divisioni
$x = 1 - y^k$	probabilità di non mutare di k geni dopo d divisioni
$p = 1 - x^{Nn}$	probabilità di mutare delle cellule staminali k volte dopo d divisioni

In questo modo $0 \leq p \leq 1$ e arriviamo quindi alla probabilità di insorgenza del tumore dato che questo, per ipotesi, si origina a partire da almeno una cellula con k mutazioni.

I parametri che abbiamo usato per il calcolo sono presi dalla letteratura e sono riportati nella tabella 1.

parametri	valori
u	$3 \cdot 10^{-6}$
d	età/4 (rate=1 div./4 giorni)
k	6
N	8
n	$1.5 \cdot 10^3 - 1.5 \cdot 10^{10}$

Tabella 3.1: Valori dei parametri del modello [10].

Ogni specie avrà un determinato numero di cellule staminali che per l'uomo è stato preso pari a $1.5 \cdot 10^7$ mentre per gli altri animali abbiamo assunto dipendere dalla massa, ossia un qualunque mammifero avrà un numero di cellule pari a $1.5 \cdot 10^7 \cdot \frac{m_{animale}}{m_{uomo}}$. Il modello originale, come mostra il primo grafico nella figura 3.1, conferma il Paradosso di Peto, ossia gli animali più grandi hanno una probabilità maggiore di contrarre il cancro. Ciò non va bene e un'ipotesi plausibile per confutare il paradosso è la seguente: assumiamo che la proliferazione cellulare avvenga con una velocità dipendente dalla massa dell'animale, in particolare la velocità di divisione dipenda linearmente dal metabolismo secondo legge di Kleiber. Ricordando quello che è stato detto nel capitolo 2, possiamo ipotizzare che il numero di divisioni cellulari sia

$$d = v_{specie} \cdot t = 1/4 \cdot \left(\frac{m_{uomo}}{m_{animale}} \right)^{1/4} \cdot t, \quad (3.2)$$

con t numero di giorni trascorsi. Ipotizzando quindi che il tasso di divisione cellulare non sia costante ma proporzionale a $\left(\frac{m_{uomo}}{m_{animale}} \right)^{1/4}$, l'andamento delle probabilità diventa quello mostrato nel secondo grafico della figura 3.1. Gli animali più piccoli hanno quindi una velocità di divisione maggiore di quelli più grandi. Nei grafici l'uomo è indicato con h , l'elefante con e e la balena con w .

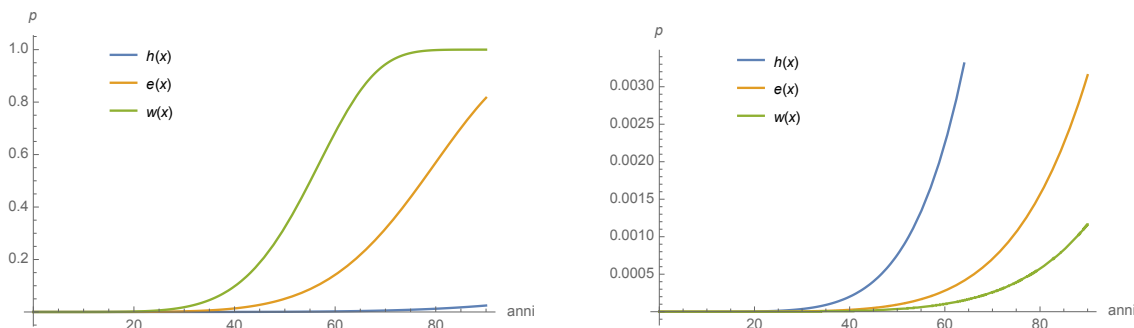


Figura 3.1: Calabrese-Shibata originale e Calabrese-Shibata modificato con la velocità di divisione cellulare dipendente dalla massa.

Con la correzione introdotta la situazione si capovolge e gli animali di taglia maggiore hanno una probabilità minore di contrarre il cancro ad una certa età rispetto a quelli più piccoli, come volevamo. In questo modo il paradosso di Peto non sussiste più.

Questa ovviamente è una prima spiegazione ma, potrebbe anche essere che altri parametri varino. Ad esempio gli animali più grandi potrebbero avere un k (numero di mutazioni richieste per l'insorgenza del cancro) maggiore oppure si potrebbe pensare che u (tasso di mutazione cellulare) scali, secondo una qualche legge di potenza, anch'esso con la massa e sia pertanto maggiore negli animali più piccoli. Proveremo a discutere quest'ultima ipotesi nella parte finale della tesi

3.2 Modello di Wright-Fischer

Analizziamo ora il modello di Wright-Fischer [11] che si appropria al problema in maniera differente rispetto a quello di Calabrese-Shibata. In questo caso non abbiamo una formula algebrica per calcolare la probabilità d'insorgenza dei tumori, ma per giungere al risultato finale dovremo utilizzare un metodo statistico.

Nel modello stocastico di Wright-Fischer ogni mutazione è un evento indipendente e la popolazione cellulare N è costante. In realtà si potrebbe estendere il modello con una popolazione crescente in maniera esponenziale in modo da rendere ancora più realistico il processo [11]. Tuttavia ciò non è stato fatto per non rendere troppo complesso il metodo computazionale da utilizzare per risolvere il modello. Con queste due ipotesi la probabilità di avere una certa configurazione di cellule con un numero di mutazioni tra 0 e k al tempo t può essere espressa tramite la distribuzione multinomiale, ossia:

$$[N_0(t+1), \dots, N_k(t+1)] \sim \frac{N(t)!}{N_0(t)! \dots N_k(t)!} \prod_{j=0}^k \theta_j^{N_j(t)}. \quad (3.3)$$

I termini fattoriali davanti la produttoria ci danno tutte le possibili configurazioni una volta scelti N_0, \dots, N_k , ossia i conteggi. I termini θ_j rappresentano invece la probabilità che una cellula possa avere una mutazione nella generazione successiva la quale deve essere proporzionale alla frazione di cellule mutate e alla probabilità che un gene muti, quindi si ha che

$$\theta_j = \sum_{i=0}^j \binom{d-i}{j-i} u^{j-i} (1-u)^{d-j} x_i(t) \quad (3.4)$$

dove u è il tasso di mutazione per gene, d il numero di geni che possono mutare e $x_i(t)$ è la frazione di cellule con i mutazioni al tempo t . Ad esempio nel caso di $k=2$ i termini della produttoria sono:

$$\begin{aligned} \theta_0^{N_0} &= \left[\binom{d}{0} (1-u)^d x_0(t) \right]^{N_0}, \\ \theta_1^{N_1} &= \left[\binom{d}{1} u (1-u)^{d-1} x_0(t) + \binom{d-1}{0} (1-u)^{d-1} x_1(t) \right]^{N_1}, \\ \theta_2^{N_2} &= \left[\binom{d}{2} u^2 (1-u)^{d-2} x_0(t) + \binom{d-1}{1} u (1-u)^{d-2} x_1(t) + \binom{d-2}{0} (1-u)^{d-2} x_2(t) \right]^{N_2}, \end{aligned}$$

ossia una cellula con 2 mutazioni ha una certa probabilità di essere prodotta da una cellula con 0 o 1 mutazione oppure che quella stessa cellula non muti nella generazione successiva.

Per fare un confronto con il modello di Calabrese-Shibata abbiamo usato gli stessi parametri che avevamo usato precedentemente e che sono riportati nella Tabella 3.1, quindi nel nostro caso $k = 6$ ($d = k$ ma in generale si può avere un numero abbastanza grande di geni che possono mutare e una volta che solo alcuni lo hanno fatto il tumore potrebbe già essere insorto [11]). Attraverso uno script in C++ abbiamo generato dei numeri casuali che seguono la distribuzione multinomiale e dalla simulazione abbiamo ricavato il seguente grafico che ci mostra come evolvono le cellule con un numero di mutazioni diverse.

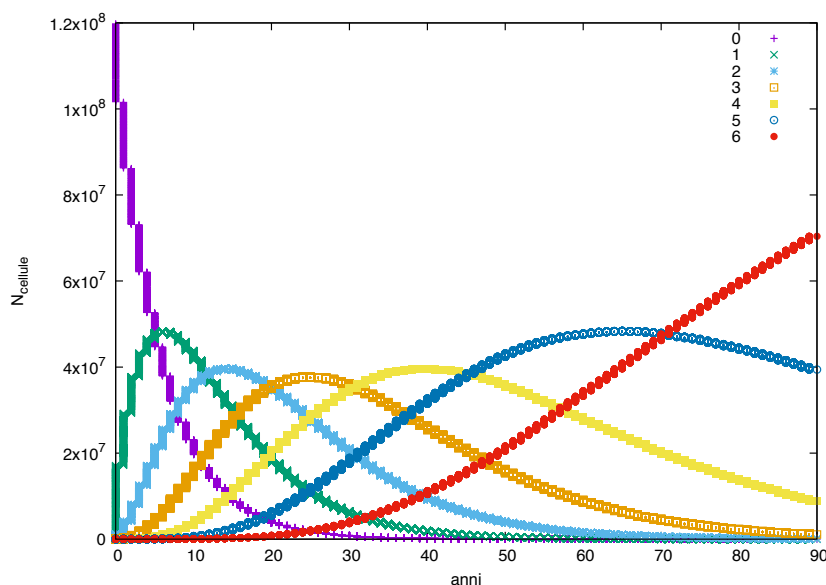


Figura 3.2: In questa simulazione abbiamo usato $N = 10^8$ e un rate di mutazione pari a $3 \cdot 10^{-4}$. Il numero totale di cellule rimane costante mentre con l'avanzare degli anni le sei classi variano: nello stato iniziale abbiamo solo cellule con 0 mutazioni (viola) che con il passare degli anni possono evolvere fino a 6 (rosso).

Dalla simulazione si possono ricavare informazione sulla dinamica del sistema, ad esempio ripetendola un numero sufficiente di volte si può controllare quando compare la prima cellula con 6 mutazioni, ossia l'organismo inizia ad ammalarsi, e ricavare il tempo di aspettazione per il tumore. Per quanto riguarda invece la probabilità che un organismo si ammali questa sarebbe data dalla somma di tutte le probabilità di tutte le possibili configurazioni compatibili con il tumore ma, calcolare questa probabilità richiederebbe un tempo astronomico. Abbiamo pertanto deciso di stimarla in maniera statistica facendo girare il programma 1000 volte e contando i casi in cui si aveva insorgenza del tumore ($N_{casifavorevoli}$). La probabilità diventa quindi $p = \frac{N_{casifavorevoli}}{1000}$.

3.2.1 Metodo computazionale

Per generare le N_0, \dots, N_6 popolazioni di cellule con le rispettive mutazioni è stata usata una funzione della libreria GNU Scientific Library, la quale genera numeri casuali secondo la distribuzione multinomiale di N eventi e di $\theta_0, \dots, \theta_6$ probabilità delle 6 classi. Queste vengono aggiornate per ogni nuova configurazione del sistema e una volta che $N_6 \neq 0$, oppure fissato un certo tempo nel quale si vuol vedere se il cancro è insorto, il programma viene ripetuto con una nuova generazione di numeri casuali.

3.3 Confronti tra i due modelli

In questo capitolo abbiamo visto due possibili modelli per stimare la probabilità di insorgenza del cancro nei mammiferi. Quello di Calabrese-Shibata ci fornisce una formula analitica che può anche essere esplicitata in base ai dati in nostro possesso. Il modello di Wright-Fisher, attraverso la distribuzione multinomiale, ci fornisce invece la probabilità di avere una certa configurazione cellulare ad un determinato istante. Per trovare la probabilità che l'animale si ammali abbiamo pertanto bisogno di un metodo numerico non banale. In entrambi i casi l'andamento delle probabilità è lo stesso, ossia gli animali più grandi hanno una probabilità maggiore di contrarre in cancro. Nel modello di Wright-Fisher però le probabilità di insorgenza del tumore sono simili tra le varie specie e questo risultato è importante per spiegare il paradosso di Peto. Infatti, confrontando i risultati trovati dai due modelli per un'età peri a 90 anni troviamo il seguente grafico.

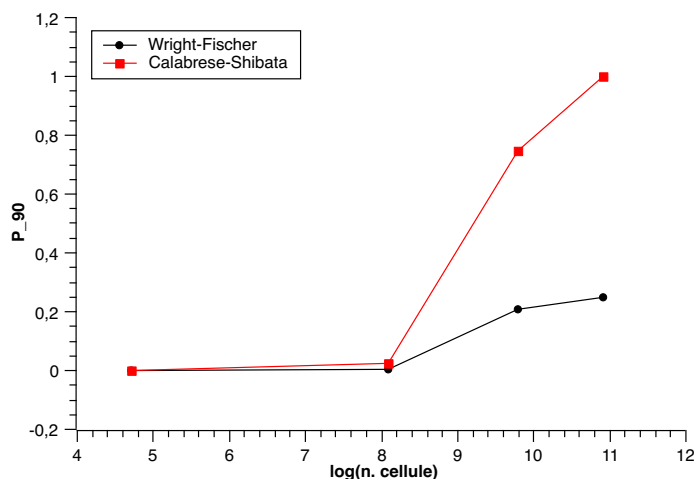


Figura 3.3: La morte casuale di una cellula è un ottimo deterrente al cancro: in nero le probabilità calcolate con il modello statistico di Wright-Fischer, in rosso quelle del modello algebrico di Calabrese-Shibata

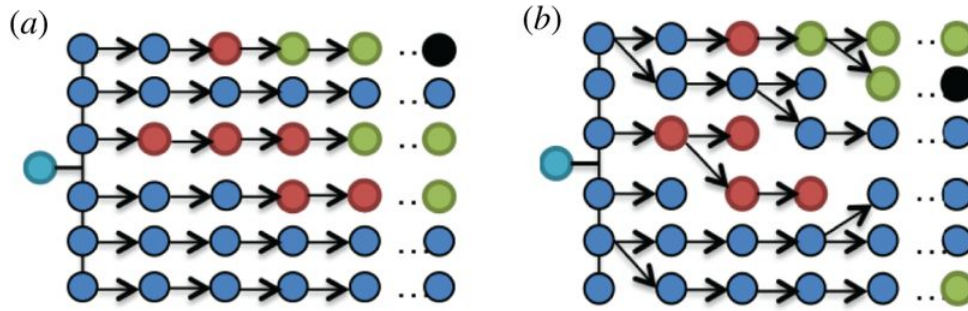


Figura 3.4: In (a) è rappresentato il modello di Calabrese-Shibata mentre in (b) quello di Wright-Fischer. Le cellule colorate in blu sono quelle staminali, in rosso quelle con una, in verde quelle con due e in nero quelle con k mutazioni.

La differenza sostanziale tra i due modelli sta infatti nel lignaggio cellulare, ossia lo sviluppo del sistema a livello spaziale e temporale. In Calabrese-Shibata, come si vede dalla figura 3.4, una cellula figlia può o meno ereditare il patrimonio genetico dalla cellula padre e lo stesso avviene in Wright-Fischer, ma in questo caso una cellula potrebbe anche morire e pertanto il suo patrimonio genetico andrebbe perso. Possiamo dire quindi che il modello di Wright-Fischer porta a risultati più ragionevoli. Consideriamo ad esempio una configurazione in cui si ha una sola cellula con 5 mutazioni: in questo stato siamo prossimi alla malattia ma per Wright-Fischer nello stato successivo questa cellula potrebbe non esistere più e quindi si potrebbe avere un organismo più sano. Il fatto che non si tiene conto di questo effetto nel modello di Calabrese-Shibata spiega il perché le probabilità trovate con questo modello siano maggiori.

Riassumendo il paradosso di Peto potrebbe venir risolto utilizzando il modello di Calabrese-Shibata ma ipotizzando che il numero di divisioni cellulari vari con la massa, oppure utilizzando un modello come quello di Wright-Fischer nel quale si tiene conto che le cellule possano morire e questo risulta essere un meccanismo di difesa contro l'insorgenza del cancro. Ovviamente anche se nel modello di Wright-Fischer introducessimo l'ipotesi che il numero di divisioni cellulari scali con la massa la probabilità di insorgenza dei tumori per animali grandi verrebbe ulteriormente ridotta.

Capitolo 4

Dati sperimentali

Lo scopo di questo capitolo è quello di vedere se i modelli precedentemente discussi possono essere consistenti con i dati sperimentali che ci sono stati forniti dai colleghi del Dipartimento di Medicina Molecolare [14]. I dati sono il frutto di una rapida analisi di banche dati e richiederebbero una più approfondita revisione per renderli completamente attendibili. Questi dati oltre ad indicare la massa e la vita media degli animali ci dicono quale è la probabilità che questi animali siano morti a causa di tumori. Per ogni singola specie la statistica è data dal numero di autopsie che sono state fatte. Per la nostra analisi abbiamo considerato solo i casi con numero di autopsie sufficientemente elevate. I dati utilizzati sono riportati nella tabella 4.1.

Nome	n° autopsie	n° tumori	massa (Kg)	m.l.s. (anni)	p %
Treeshrew	1145	64	0.2	12	5.59
Guinea pig	7000	100	1	6	1.40
Rabbit	2956	48	2	10	1.62
Dog	2002	460	30	10	23.0
Human			65	80	22.0
California Sea Lion	456	83	200	25	18.2
Horse	687	21	400	30	3.10
Elephant	644	20	4800	65	3.11

Tabella 4.1: Nella tabella le rispettive sigle stanno a significare: m.l.s.=maximum life span, massa=adult mass. Osservando i dati si può vedere come le probabilità di tumore siano abbastanza simili tra loro nonostante la variazione delle masse.

Prima di confrontare le probabilità di tumore sperimentali con i valori ottenuti dai due modelli abbiamo cercato se vi fosse qualche relazione tra la probabilità e le grandezze riportate nella tabella. In particolare come prima cosa abbiamo cercato una relazione tra il tempo di vita dei mammiferi e la loro massa, ipotizzando che $t_{vita} \propto m^{1/4}$. Se ciò fosse vero allora, ricordando che la $\nu_{cardiaca} \propto m^{-1/4}$, il tempo biologico, che viene misurato in numero di battiti cardiaci per specie, è costante: $t_{vita} \cdot \nu_{cardiaca} = k$ [15]. Abbiamo quindi

preso il tempo di vita di ogni specie e messa in relazione con la propria massa, il grafico log-log ottenuto è rappresentato nella Figura 4.1.

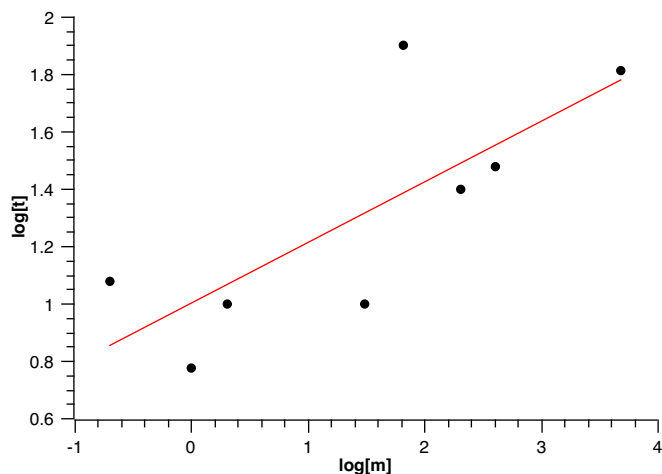


Figura 4.1: Andamento della vita con la massa.

Dal fit si trova una pendenza della retta pari a 0.21 ± 0.07 mentre il coefficiente di correlazione vale 0.6. Questo è consistente con quanto visto in letteratura [12] anche se i dati in nostro possesso non sono ottimali. Le specie a nostra disposizione sono poche e i campioni presi in considerazioni provengono da animali che non vivono nel loro ambiente naturale.

Di seguito riportiamo invece i grafici in cui la probabilità sperimentale d'insorgenza del tumore è stata messa in relazione con la massa della specie, il rapporto tra la massa e il tempo di vita ed infine il tempo di vita stesso. Alcune probabilità sono davvero piccole e per questo abbiamo fatto uno zoom per vedere meglio la distribuzione di quei valori molto vicini tra loro.

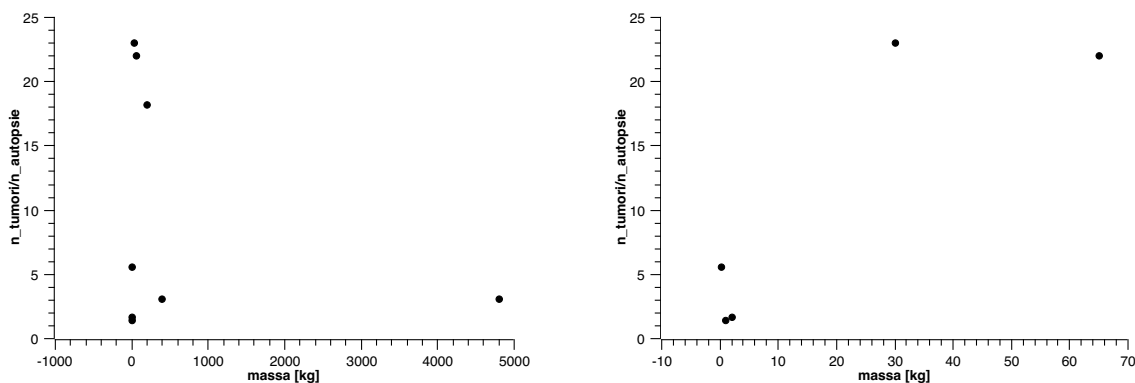


Figura 4.2: Probabilità in funzione della massa.

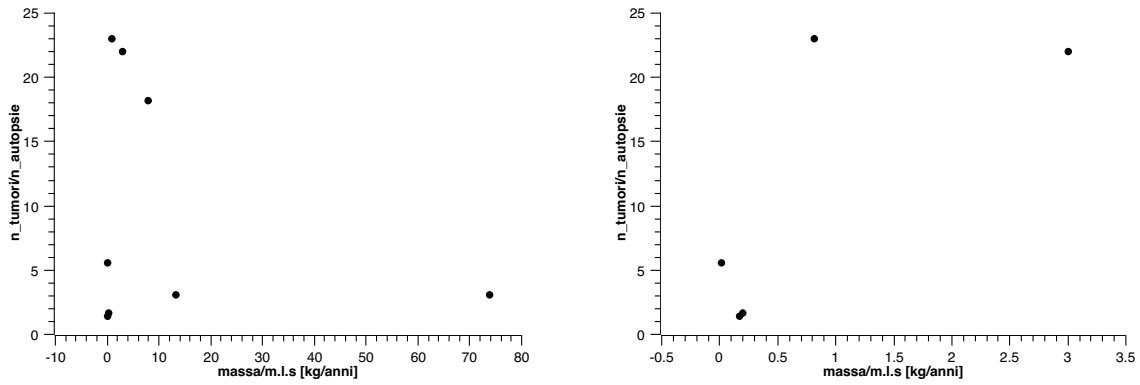


Figura 4.3: Probabilità in funzione del rapporto massa/tempo.

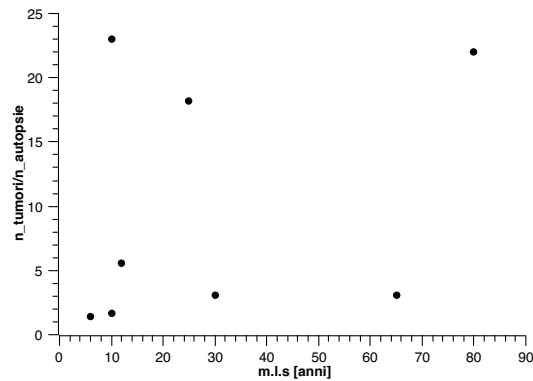


Figura 4.4: Probabilità in funzione del tempo di vita.

In tutti i grafici non si nota nessun andamento particolare e anche provando a classificare i mammiferi secondo la famiglia (primati, roditori, carnivori ...) alla quale appartengono non risulta nessuna relazione significativa. La specie più a rischio di tumori tra tutti i mammiferi è l'uomo e se pensiamo allo stile di vita che conduciamo non è difficile trovare una spiegazione. L'habitat dove ogni specie vive contribuisce sicuramente ad innalzare o abbassare il rischio verso questo tipo di malattie. Calcolando le probabilità di insorgenza di cancro con i modelli teorici precedentemente studiati otteniamo i valori nella tabella 4.2.

p_{spe}	p_{C-S}	p_{C-S}^K	p_{W-F}
5.59	$5.74 \cdot 10^{-8}$	$3.17 \cdot 10^{-4}$	0
1.40	0	$1.86 \cdot 10^{-6}$	0
1.62	$11.6 \cdot 10^{-7}$	$2.82 \cdot 10^{-5}$	0
23.0	2.46	$7.29 \cdot 10^{-6}$	0
22.0	1.23	1.23	10^{-3}
18.2	$3.71 \cdot 10^{-3}$	$6.92 \cdot 10^{-4}$	0
3.10	$1.47 \cdot 10^{-3}$	$1.46 \cdot 10^{-3}$	0
3.11	23.4	$4.35 \cdot 10^{-2}$	10^{-3}

Tabella 4.2: Le p^K indicano le probabilità calcolate con il modello modificato.

Guardando i dati si nota come non ci sia una grande relazione tra i valori dei modelli teorici e quelli dei dati sperimentali. I modelli andrebbero pertanto modificati per cercare di renderli più compatibili. Abbiamo comunque graficato le probabilità ottenute dai modelli di Calabrese-Shibata in funzione di quella sperimentale ottenendo i seguenti grafici.

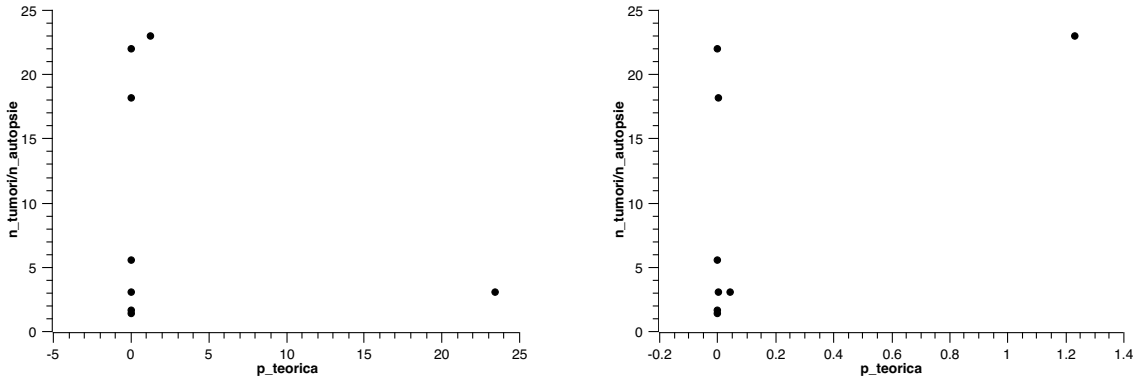


Figura 4.5: Calabrese-Shibata e Calabrese-Shibata modificato secondo la legge di Kleiber.

Abbiamo allora deciso di cercare di utilizzare i dati sperimentali, benchè non accurati, per cercare di ricavare i parametri dei modelli studiati precedentemente per ottenere una buona correlazione con gli esperimenti. Possiamo per esempio cercare di capire come il tasso di mutazione u vari da mammifero a mammifero. Infatti nella parte precedente della tesi abbiamo mantenuto costante il rate di mutazione u ma, un'ipotesi possibile è che anche questa quantità vari di specie in specie. Usando le probabilità sperimentali siamo in grado di ricavare u attraverso il modello di Calabrese-Shibata, infatti

$$u = 1 - (1 - (1 - (1 - p)^{1/(N \cdot m)})^{1/(k)})^{1/d}. \quad (4.1)$$

Inserendo le probabilità di insorgenza del tumore nell'uomo troviamo che u dovrebbe essere:

$$u_{uomo} = 0.6 \cdot 10^{-5}.$$

Per il modello di Wright-Fischer otteniamo un tasso di mutazione pari a

$$u = 0.8 \cdot 10^{-5}.$$

Entrambi i modelli ci danno un u leggermente maggiore rispetto a quello che avevano utilizzato. Per determinare u tramite il metodo numerico si è dovuto procedere per tentativi, inserendo manualmente un determinato valore del tasso di mutazione e andando a vedere quale approssimava meglio il dato sperimentale. Per questo motivo abbiamo calcolato u per tutte le specie solo con il modello analitico. I dati sono riportati nella tabella 4.3 e abbiamo poi graficato il $\log(u)$ in funzione del $\log(m)$ nella figura 4.6.

Nome	u_{C-S}
Treshreew	$6.9 \cdot 10^{-5}$
Guinae pig	$8.3 \cdot 10^{-5}$
Rabbit	$4.5 \cdot 10^{-5}$
Dog	$4.0 \cdot 10^{-5}$
Human	$0.6 \cdot 10^{-5}$
Californian Sea Lion	$1.3 \cdot 10^{-5}$
Horse	$0.7 \cdot 10^{-5}$
Elephant	$0.2 \cdot 10^{-5}$

Tabella 4.3: Valori di u ricavati dal modello di Calabrese-Shibata attraverso i dati sperimentali.

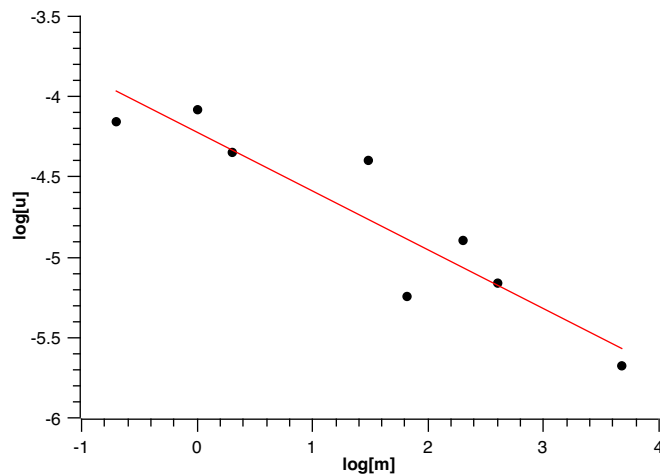


Figura 4.6: $\log(u) = A \cdot \log(m) + \log(u_0)$.

Dal fit si trova che:

$\log(u_0)$	$\sigma_{\log(u_0)}$	A	σ_A	R^2
-4.22	0.12	-0.36	0.06	0.85

Tabella 4.4: $\log(u) = -0.36 \cdot \log(m) - 4.22$.

Da questa semplice analisi concludiamo che u è maggiore negli animali più piccoli e troviamo una relazione precisa che lega queste due quantità. Possiamo introdurre questa correzione nel modello di Calabrese-Shibata ottenendo

$$p = 1 - [1 - [1 - [1 - u(m)]^d]^k]^{Nn}, \quad (4.2)$$

con

$$u(m) = u_0 m^{-0.36}. \quad (4.3)$$

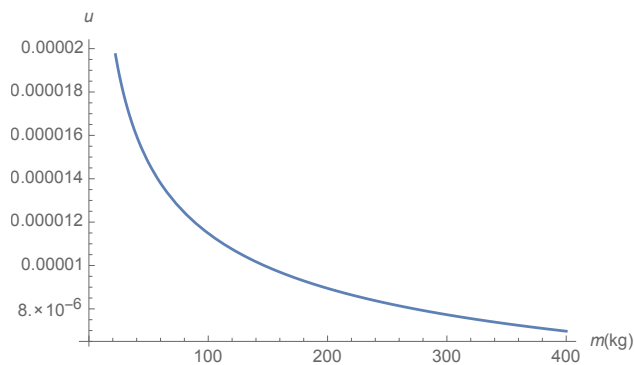


Figura 4.7: $u(m)$.

Quindi possiamo concludere che i dati sperimentali in nostro possesso potrebbero essere spiegati da un modello di Calabrese-Shibata in cui il tasso di mutazione varia con una legge di scala come 4.3. Il fatto che il fit risulti così buono è certamente stimolante ma per confermare questa ipotesi dovremmo sviluppare un'argomentazione fisica che possa giustificare questa legge di potenza, cosa che va al di là degli scopi di questa tesi.

Capitolo 5

Conclusioni

In questo lavoro di tesi sono stati studiati due modelli, uno algebrico e uno numerico, sviluppati per studiare l'insorgenza del cancro nei mammiferi. Lo scopo era quello di vedere se questi modelli permettono di spiegare il paradosso di Peto. Entrambi ci danno la probabilità che una specie possa ammalarsi in base a dei parametri fissati: il tasso di mutazione cellulare, la velocità con la quale le cellule si duplicano, il numero di mutazione cellulari necessarie a far insorgere la malattia e il numero di cellule. Sia il modello di Calabrese-Shibata che quello di Wright-Fischer ci dicono che la probabilità che un mammifero si ammali è maggiore per quelli più grandi confermando quindi il paradosso di Peto. Questo però avviene in maniera meno forte con il modello di Wright-Fischer dove si tiene conto che le cellule mutate possono morire prima di sviluppare la malattia. Se supponiamo che il tasso di divisione cellulare scali invece con la massa con una legge derivabile da quella di Kleiber, si vede che l'insorgenza dei tumori sarebbe molto minore negli animali più grandi rispetto a quelli più piccoli e questo spiegherebbe il paradosso di Peto. Tuttavia cercando di verificare se i nostri modelli possano riprodurre i dati sperimentali ci accorgiamo che non riusciamo ad ottenere una buona correlazione. Questo può essere derivato dal fatto che i dati sperimentali non siano completamente attendibili (per esempio animali in cattività e in condizioni ambientali diverse) oppure che i nostri modelli debbano essere ancora modificati per essere più realistici. Per esempio sarebbe opportuno correggerli inserendo una popolazione cellulare non costante, ma che aumenti con il passare del tempo [11], infatti come abbiamo detto precedentemente le cellule cancerose aumentano con la metastasi. Questa correzione è opportuna se si vuole calcolare il tempo di aspettazione del tumore tramite il modello di Wright-Fischer. Nell'ultima parte della tesi abbiamo visto che un modo ulteriore per spiegare il paradosso di Peto, tenendo presente i dati sperimentali, potrebbe essere quello di supporre che il tasso di mutazione u scali con la massa con una legge di potenza del tipo 4.3. Questa ipotesi assai interessante richiederebbe ulteriori analisi.

Possiamo concludere che studiare e confrontare il meccanismo attraverso il quale il cancro sopraggiunge nelle diverse specie di mammiferi può rivelarsi utile per ricavare informazioni importanti: gli studi comparativi sono nella loro infanzia, ma comparare i dati di diverse specie potrebbe fornire indicazioni terapeutiche. Come primo esempio di questa strategia, chiamata oncologia comparativa, vi sono alcuni risultati che hanno mostrato come in animali di grossa taglia siano presenti dei geni oncorepressori presenti anche nell'uomo [10]. Un'approfondita analisi a livello genomico, accompagnata da un'adeguata

modellizzazione fisica matematica dei vari livelli di espressione genica in mammiferi diversi, potrebbe quindi evidenziare i meccanismi che la natura ha escogitato per mitigare gli effetti del cancro nelle varie specie. Questi meccanismi potrebbero poi essere utilizzati per seguire nuove terapie di applicazione sull'uomo. Si auspica che questa tesi possa dare un piccolo contributo in un disegno scientifico di così ampia portata.

Bibliografia

- [1] Feynman R., Leighton R., Sands M., *The Feynman Lectures on physics, Volume 1: Mainly mechanics, radiation, and heat*, California Institute of Technology (1964).
- [2] Zumerle G., *Appunti di Fisica Generale 1* (2012/2103).
- [3] Huang K., *Statistical Mechanics*, Massachussets Institute of Technology (1987).
- [4] West G. and Brown J., *Life's Universal Scaling Law*, Physics Today (2004).
- [5] Banavar J., Cook T., Rinaldo A., and Maritan A., *Form, function, and evolution of living organisms*, PNAS (<http://www.pnas.org/content/111/9/3332>) (2014).
- [6] Kleiber M., *The Fire of Life: An Introduction to Animal Energetics*, Robert E. Krieger, Huntington, NY (1975).
- [7] Kumar V., Abbas A., Aster J., Robbins and Cotrain, *Pathology basis of disease*, 8° edizione Saunders/Elsevier, cap. 7 (2007).
- [8] Nunney L., Maley CC., Breen M., Hochberg ME., Schiffman JD., *Peto's Paradox and the promise of comparative oncology*, Phil. Trans. R. Soc. B 370:20140177 (2015).
- [9] Schrödinger E., *What is life? La cellula vivente dal punto di vista fisico*, Cambridge University (1947).
- [10] Caulin AF., Graham TA., Wang L-S., Maley CC., *Solution to Peto's paradox revealed by mathematical modelling and cross-species cancer gene analysis*, Tran. R. Soc. B 370:20140222 (2015).
- [11] Beerenwinkel N., Antal T., Dingli D., Traulsen A., Wkinzler K., Velculescu V., Vogelstein B., Nowak M., *Genetic Progression and the Waiting Time to Cancer*, Plos Computational Biology (2007).
- [12] Szekely P., Korem Y., Moran U., Mayo A., Alon U., *The Mass-Longevity Triangle: Pareto Optimality and the Geometry of Life-History Trait*, PLoS Comput Biol 11(10): e1004524. (doi:10.1371/journal.pcbi.1004524) (2015).
- [13] Calabrese P., Shibata D., *A simple algebraic cancer equation: calculating how cancers may arise with normal mutation rates*, BMC Cancer 10,3. (doi:10.1186/1471-2407-10-3) (2010).

- [14] Cordenonsi M., *Comunicazione privata* (2015).
- [15] Dattoli G., Ottaviani P.G., Pagnutti S., ENEA Tecnologie Fisiche e Nuovi Materiali, *Leggi di scala in biologia e dinamica dello sviluppo delle masse tumorali*, Centro Ricerche Frascati (2009).