

Penna Model: Mutation Accumulation and the Catastrophic Senescence of Pacific Salmon

Esame di Fisica Statistica e Teoria dell'informazione
A.A. 2013/2014

Bonaiuti Simone
simone.bonaiuti@stud.unifi.it

Cambi Dario
dario.cambi@stud.unifi.it

Introduzione

L'invecchiamento di un qualsiasi essere vivente può essere descritto come il decremento della probabilità di sopravvivenza con l'avanzare dell'età. Esistono almeno 2 teorie di invecchiamento basate sull'evoluzione [1, 6]: la teoria di Antagonistic Pleiotropy (basata sul concetto di privilegiare le mutazioni che apportano benefici ad un individuo nei primi anni di vita, anche se potrebbero avere un effetto negativo piuttosto catastrofico in un individuo di età superiore) e la teoria di Mutation Accumulation (basata su un aumento dell'accumulo di mutazioni ad età avanzate rispetto ai primi anni di vita). Tali teorie, oltre a spiegare molti aspetti dell'invecchiamento, permettono di utilizzare tecniche di Fisica Statistica per trattare tale fenomeno in relazione ad una popolazione di individui.

L'obiettivo di questo elaborato consiste quindi nel modellare l'evoluzione temporale della popolazione di un insieme di Salmoni del Pacifico, basandoci su un modello detto Penna Model [3, 4].

Descrizione del problema

Consideriamo una popolazione di \mathbf{N} salmoni che si evolve nel tempo a intervalli di tempo discreti $t = 1, 2, 3, \dots$ (andamento annuale). Il corredo genetico di ciascun salmone è rappresentato come un vettore binario di \mathbf{M} gruppi di geni; tali gruppi di geni possono essere positivi (uguali a 0) o negativi (uguali a 1). Ogni anno ai salmoni si attiva, in ordine a partire dal primo, un gruppo di geni: all'attivazione di \mathbf{T} gruppi di geni negativi il salmone muore. Oltre al fattore genetico, per simulare le condizioni di morte dovute a malattie o fattori ambientali (scarsità di cibo in relazione alla dimensione totale della popolazione), utilizziamo il **Verhulst Factor** [5], che è calcolato come:

$$V_F = 1 - N(t)/S.$$

Tale fattore indica la probabilità che ha un salmone di rimanere vivo a prescindere dal fattore genetico e dipende dal numero totale di salmoni $N(t)$ al tempo t , e dal numero massimo di salmoni che l'ambiente è in grado di mantenere S .

In ogni caso, l'età massima di un salmone è M : superata tale età il salmone muore indipendentemente dal suo corredo genetico e dal Verhulst Factor (morte per vecchiaia). Ogni salmone inizia a fare figli da una certa età R in poi, facendo b figli ogni anno. I figli ereditano il genoma del genitore, ma hanno una determinata probabilità di avere una o più mutazioni negative (gruppi di geni positivi che diventano negativi).

Implementazione

Presentiamo di seguito i parametri principali del problema:

- **Numero di iterazioni:** massimo numero di iterazioni che verranno effettuate. Ogni iterazione è da considerarsi come il passare di un anno di tempo. Se la popolazione si estingue, le iterazioni vengono immediatamente fermate.
- **Numero iniziale di salmoni ($N_{t=0}$):** dimensione della popolazione iniziale.
- **Numero massimo di salmoni (S):** soglia di popolazione oltre la quale la probabilità dei salmoni di fare figli è nulla.
- **Lunghezza del genoma:** dimensione del vettore che rappresenta il genoma, nonché l'età massima raggiungibile, di un salmone (dimensione standard = 32).
- **Numero di mutazioni negative (N_{Mut}):** numero di possibili mutazioni del genoma dei figli (valore standard = 1).
- **Numero di figli (b):** numero di nascite per ciascun salmone ad ogni anno di riproduzione.
- **Soglia di Morte (T):** numero di genomi negativi attivati i quali il salmone muore.
- **Età di riproduzione (R):** età dalla quale i salmoni possono riprodursi.

Nella nostra implementazione l'utente ha inoltre la possibilità di modificare le simulazioni impostando 3 parametri booleani:

- **Reproduce at Exactly Age:** se è uguale a 0 i salmoni possono riprodursi a partire dall'età R in poi; se è uguale a 1 possono riprodursi soltanto quando hanno età R .
- **Always Bad Mutation:** se è uguale a 1 i figli avranno sempre N_{Mut} mutazioni negative; se è uguale a 0 per ciascun figlio vengono scelti in modo casuale N_{Mut} gruppi di geni che, se positivi, verranno mutati a negativi, altrimenti resteranno negativi.

- **Random Genomas:** se uguale a 1 la popolazione iniziale avrà un corredo genetico casuale; se uguale a 0 la popolazione iniziale avrà un corredo genetico completamente sano (tutti i gruppi di geni posti a zero).

Output

Una volta impostati i necessari parametri, avviando lo script verrà eseguita la simulazione desiderata, la quale restituirà all'utente:

- Un **Grafico Demografico** che mostra l'andamento della popolazione nel tempo (numero totale di salmoni nelle varie iterazioni eseguite).
- Un **Grafico del Survival Rate** diviso per fasce di età: tasso di sopravvivenza dei salmoni all'età successiva, calcolato come

$$SR = N_k(t)/N_{k-1}(t-1),$$

dove k è l'età e t è il tempo. In questo caso l'utente deve specificare una iterazione iniziale ed una finale: il risultato sarà mediato con i dati compresi fra tali iterazioni, in modo da avere un grafico che presenta un andamento con ridotte fluttuazioni.

- La **Matrice di Leslie** [2] calcolata ad una specifica iterazione, il suo autovalore massimo ed il relativo autovettore, oltre al grafico di tale autovettore.

Tramite il *Grafico Demografico* l'utente può vedere l'andamento della popolazione al passare degli anni: solitamente tale grafico presenta una prima fase di crescita esponenziale della popolazione, seguita da una fase di stabilità nella quale la popolazione si mantiene circa costante (eccetto per piccole fluttuazioni) ed infine da una fase di decrescita abbastanza rapida che porta all'estinzione della popolazione (l'incidenza delle mutazioni negative diventa il fattore dominante che influisce negativamente sulla sopravvivenza della popolazione).

Il *Grafico del Survival Rate* offre invece informazioni circa l'aspettativa di vita media di un salmone, mostrando la probabilità che ha un salmone di età k di sopravvivere fino all'età $k+1$. Chiaramente, la probabilità di sopravvivenza è maggiore nei primi anni di vita e decresce con l'aumentare dell'età.

La *Matrice di Leslie* è utilizzata nei modelli con struttura di età per la valutazione di due problemi di stabilità: l'evoluzione globale della popolazione nel tempo e la ripartizione della popolazione fra le varie classi di età. L'autovalore dominante λ_1 della matrice di Leslie può essere utilizzato per valutare l'andamento asintotico della popolazione (quando essa è nella fase stazionaria): se $\lambda_1 \geq 1$ la popolazione tende a crescere; viceversa, se

$\lambda_1 < 1$ tende a diminuire. L'autovettore v relativo all'autovalore λ_1 esprime la distribuzione asintotica delle varie classi di età.

Esperimenti e risultati

Per tutti i seguenti esperimenti la lunghezza del genoma è 32, il numero di salmoni iniziale è $N(0) = 2 \cdot 10^5$, $S = 10 \cdot N(0)$, $N_{Mut} = 1$ e Random Genomas = 0 (la popolazione iniziale è completamente sana).

Nel primo esperimento vogliamo valutare l'andamento del Survival Rate facendo riprodurre i salmoni esclusivamente ad una specifica età, oppure da quell'età in poi. Impostiamo quindi $R = 27$, $b = 30$ e $T = 2$ in entrambi i casi. Otteniamo i seguenti grafici:

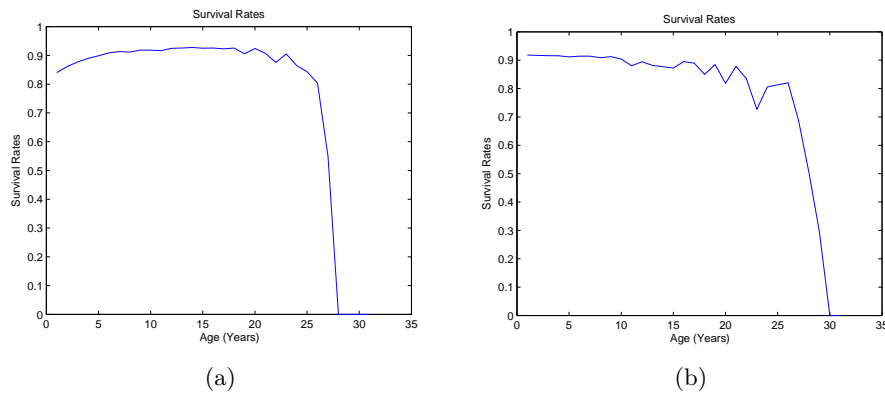


Figura 1: Survival Rate al variare dell'età riproduttiva. Caso (a) $R = 27$, caso (b) $R \geq 27$.

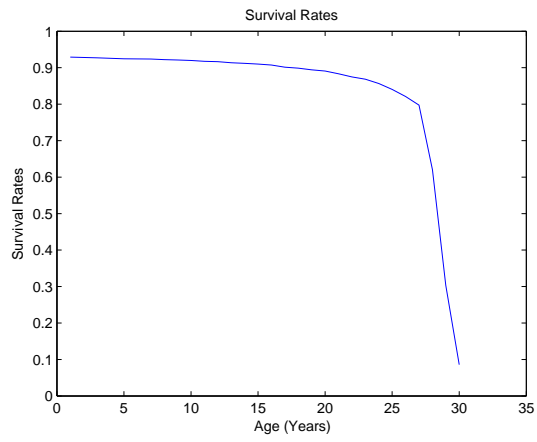
Dai grafici si nota un andamento simile del Survival Rate fino alla minima età di riproduzione; successivamente, nel caso (a) si nota un decremento più netto rispetto al caso (b).

Nel secondo esperimento vogliamo mostrare l'andamento del Survival Rate al variare della soglia di morte T . In Figura 2 sono mostrati i grafici con $R = 27$, $b = 30$ e rispettivamente con $T = 1$, $T = 2$ e $T = 4$.

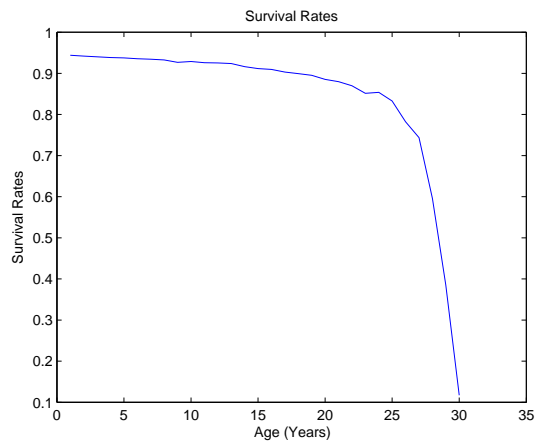
Come si nota dai grafici, l'aumentare della soglia di morte T non provoca sostanziali cambiamenti nell'andamento del Survival Rate.

Il valore di T influenza invece la dimensione della popolazione: in Figura 3 vediamo che all'aumentare di T aumenta la dimensione media della popolazione.

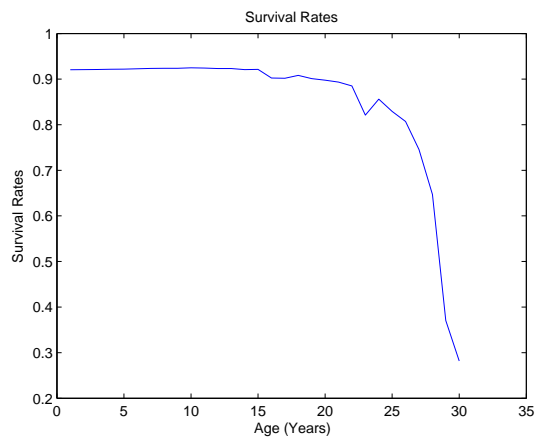
Un altro fattore che influisce sulla dimensione della popolazione media è l'età di riproduzione: diminuendo R l'andamento del Survival Rate rimane pressochè immutato, mentre la dimensione della popolazione è maggiore, come si nota dalla Figura 4.



(a)

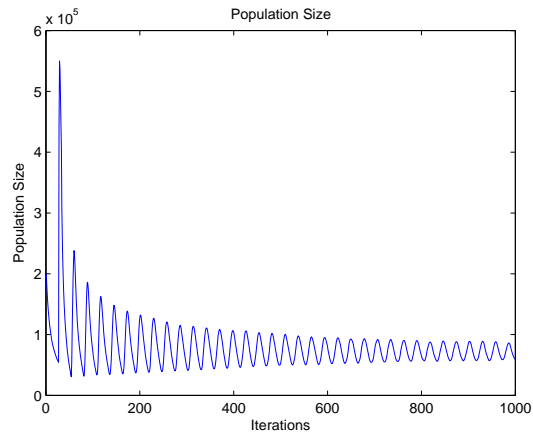


(b)

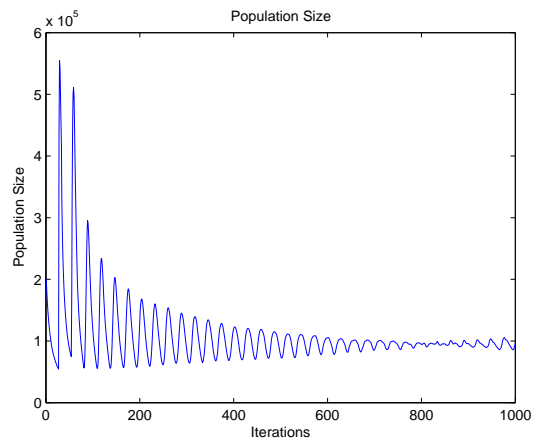


(c)

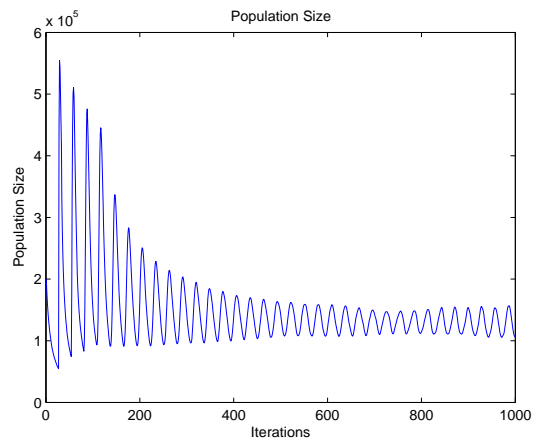
Figura 2: Survival Rate al variare della soglia T . Caso (a) $T = 1$, caso (b) $T = 2$, caso (c) $T = 4$.



(a)



(b)



(c)

Figura 3: Popolazione al variare della soglia T . Caso (a) $T = 1$, caso (b) $T = 2$, caso (c) $T = 4$.

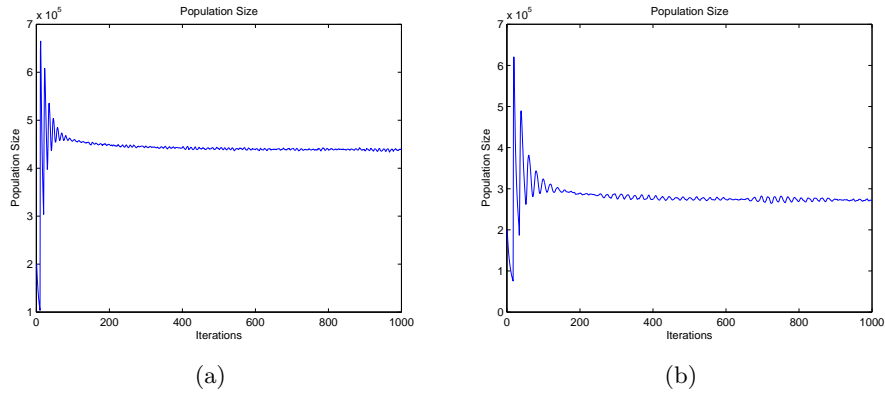


Figura 4: Popolazione al variare di R (con $T = 2$ e $b = 10$). Caso (a) $R = 10$, caso (b) $R = 17$.

Per concludere, mostriamo in Figura 5 (a) l'andamento asintotico della popolazione per le diverse fasce di età attraverso il grafico dell'autovettore v , relativo all'autovalore dominante λ_1 della matrice di Leslie. Tale autovettore è stato ottenuto dopo 2600 iterazioni (situazione di popolazione stazionaria) a partire dai valori standard usati per i test con $N_{Mut} = 1$, $b = 30$ e $R \geq 27$.

Possiamo notare che l'andamento dell'autovettore rispecchia opportunamente la distribuzione reale della popolazione nelle diverse fasce di età ottenuta tramite simulazione (Figura 5 (b)).

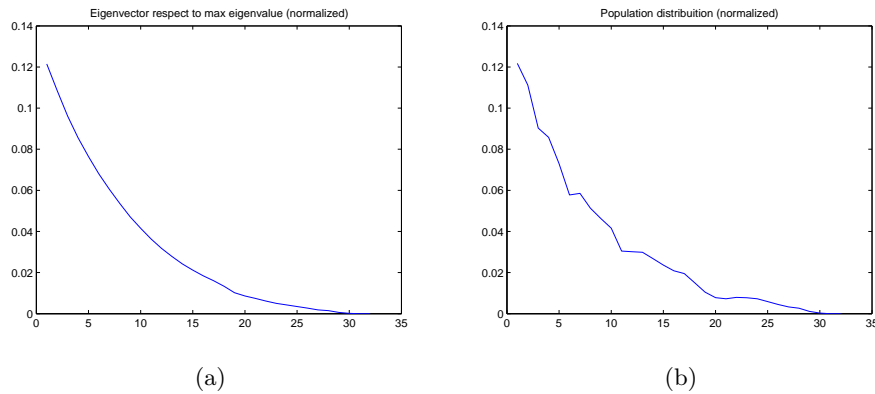


Figura 5: (a) Autovettore v relativo all'autovalore dominante λ_1 corrispondente alla matrice di Leslie. (b) Distribuzione della popolazione nelle diverse fasce d'età, ottenuta dalla simulazione. Entrambi i grafici sono normalizzati.

Conclusioni

In conclusione, lo sviluppo di questo lavoro ci ha permesso di modellare l'andamento della popolazione dei Salmoni del Pacifico attraverso tecniche di Fisica Statistica e di studiare la dinamica di evoluzione della popolazione nel tempo attraverso la matrice di Leslie.

Basandoci sugli esperimenti fatti in [4] e replicandoli, abbiamo notato come il Survival Rates dei salmoni mantenga il medesimo andamento al variare di alcuni parametri fondamentali come il tasso di riproduzione b o il numero massimo di mutazioni consentite T . Anche la variazione dell'età di riproduzione mantiene il risultato immutato.

Abbiamo inoltre mostrato come la popolazione tenda a stabilizzarsi dopo un sufficiente numero di iterazioni, valutando il comportamento asintotico tramite modello di Leslie. Abbiamo infine evidenziato come tale modello rispetti la distribuzione effettiva della popolazione, ripartita nelle diverse fasce d'età.

Bibliografia

- [1] Brian. Charlesworth. *Evolution in age-structured populations / Brian Charlesworth*. English. Cambridge University Press Cambridge [Eng.] ; New York, 1980, xiii, 300 p. ; ISBN: 0521230454 0521297869 0521297869 0521230454.
- [2] D.L. Deangelis et al. «Stability and return times of Leslie matrices with density-dependent survival: applications to fish populations». In: *Ecological Modelling* 8.0 (1980), pp. 149 –163. ISSN: 0304-3800. DOI: [http://dx.doi.org/10.1016/0304-3800\(80\)90034-4](http://dx.doi.org/10.1016/0304-3800(80)90034-4). URL: <http://www.sciencedirect.com/science/article/pii/0304380080900344>.
- [3] Zhi Feng Huang e Dietrich Stauffer. «Stochastic Penna model for biological aging». In: *Theory in Biosciences* 120.1 (2001), pp. 21–28.
- [4] T.J.P. Penna, S. Moss de Oliveira e Dietrich Stauffer. «Mutation accumulation and the catastrophic senescence of the Pacific salmon». In: *Phys. Rev. E* 52 (4 1995), R3309–R3312. DOI: 10.1103/PhysRevE.52.R3309. URL: <http://link.aps.org/doi/10.1103/PhysRevE.52.R3309>.
- [5] Chryslie Margus N. Piñol e Ronald S. Banzon. «Stability in a population model without random deaths by the Verhulst factor». In: *Physica A: Statistical Mechanics and its Applications* 390.7 (2011), pp. 1295 –1299. ISSN: 0378-4371. DOI: <http://dx.doi.org/10.1016/j.physa.2010.11.046>. URL: <http://www.sciencedirect.com/science/article/pii/S0378437110010113>.
- [6] Michael R. Rose. *Evolutionary biology of aging / Michael R. Rose*. English. <http://www.loc.gov/catdir/enhancements/fy0640/90006787-t.html>. Oxford University Press New York, 1991, ix, 221 p. : ISBN: 0195061330 0195095308.