

Penna Model for biological aging

Esame di Fisica Statistica e Teoria dell'informazione

Giulia Pieralli
pieralli.g@gmail.com

Tommaso Resti
tommy88r@gmail.com

1 giugno 2015

1 Introduzione

Le ricerche sull'evoluzione dell'invecchiamento mirano a spiegare il motivo per cui tutti gli esseri viventi si indeboliscono e muiono con l'età, quindi in quest'ottica l'invecchiamento è visto come il decremento della probabilità di sopravvivenza con l'avanzare dell'età. Nella comunità scientifica esistono varie teorie sull'origine evolutiva della senescenza. In questo elaborato viene presa in considerazione la teoria di *Mutation Accumulation*, che consiste nell'accumulazione di mutazioni di un gene in età avanzata rispetto ai primi anni di vita, che porta alla morte di un individuo. In questo articolo studiamo l'evoluzione della popolazione di salmoni nel tempo mediante l'uso di un modello detto *Penna Model*.

2 Penna Model

L'analisi del modello parte considerando una popolazione di S salmoni che si evolve nel tempo ad intervalli discreti e che non può superare un limite superiore fissato S_{max} del numero di individui appartenenti alla popolazione. Ogni salmone ha un corredo genetico di G geni che si attivano con l'avanzare dell'età. A partire da una fissata età di procreazione, un salmone può riprodursi ed i figli avranno il solito corredo genetico del genitore con eventuali mutazioni negative (gruppi di geni positivi che diventano negativi). Quando i geni negativi di un salmone raggiungono una soglia fissata T rispetto all'età dell'esemplare, questo muore. Inoltre è fissata anche l'età massima di vita di un salmone oltre al quale indipendentemente da qualsiasi altro fattore sopraggiunge la morte per vecchiaia.

2.1 Verhulst Factor

In aggiunta alla morte dovuta all'accumulazione di geni negativi o dovuta alla vecchiaia, nell'elaborato è stata prevista la possibilità di simulare le condizioni di morte reali dovute a cause ambientali come la scarsità di cibo in relazione al numero di esemplari o malattie. Questa simulazione è raggiunta con l'utilizzo del Verhulst Factor calcolato come:

$$V = 1 - (S/S_{max}) \quad (1)$$

Tale fattore indica la possibilità che ha un salmone di rimanere vivo indipendentemente dal fattore genetico ed è applicata ad ogni salmone facente parte della popolazione senza limiti di età. Come si intuisce più la popolazione si avvicina al limite superiore S_{max} , più un generico salmone è probabile che muoia. Quindi non si arriverà mai ad avere la popolazione massima permessa.

2.2 Probabilità di procreazione

Nel modello si può considerare la possibilità che un salmone procrei sempre dal momento in cui raggiunge l'età riproduttiva, oppure che si riproduca solo con una certa probabilità data dalla formula:

$$AR = 1 - (S/S_{max}) \quad (2)$$

Questa formula tiene conto del numero di salmoni rispetto al numero massimo, risultando più alta quando la popolazione corrente è lontana da toccare il limite permesso, non impedendo comunque di raggiungere tale valore. In questo modo si calcola la probabilità che un salmone in età di procreazione abbia figli.

3 Implementazione

L'elaborato da noi svolto in Matlab prevede la possibilità di impostare diversi parametri su cui poter richiamare lo script del *Penna Model* in modo da confrontare i risultati al variare dei suddetti. Questi parametri sono i seguenti:

- **Numero iniziale di salmoni S :** grandezza della popolazione iniziale.
- **Numero massimo di salmoni S_{max} :** limite oltre il quale la popolazione di salmoni non può crescere.
- **Lunghezza del corredo genetico G :** dimensione del vettore binario che rappresenta il corredo genetico dei salmoni e corrisponde anche all'età massima che possono raggiungere.
- **Numero di mutazioni negative N_{mut} :** numero di mutazioni negative che all'atto della riproduzione possono incorrere tra il genitore e i figli.
- **Numero di figli F :** numero di figli che ogni salmone in età procreativa può avere. Il numero effettivo di figli che un genitore ha dipende dalla probabilità di riproduzione e deve rispettare la popolazione massima.
- **Soglia di morte T :** numero di geni negativi in relazione all'età del salmone oltre il quale l'esemplare muore.
- **Età riproduttiva R :** età in cui il salmone diventa riproduttivo.
- **Verhulst V :** se è impostato a 1 considera la probabilità di morte casuale ottenuta dal Verhulst Factor, se 0 i salmoni muoiono solo di vecchiaia o a causa del superamento della soglia dei geni negativi.

- **Always Reproduce *AR***: se è settato a 0 i salmoni in età riproduttiva si riproducono con una certa probabilità data dalla formula 2, se 1 la riproduzione è certa.
- **Reproduce Exactly Age *EA***: se è 0 i salmoni si riproducono a partire da una certa età in poi, se 1 si riproducono esclusivamente all'età predefinita.
- **Always Bad Mutation *BM***: se il parametro è uguale a 1 i figli procreati avranno sempre un numero di mutazioni negative pari a N_{mut} (i.e. nel vettore binario che rappresenta il corredo genetico dei salmoni si avrà una trasformazione di un elemento 0 in 1); se è settato a 0 per ciascun figlio vengono scelti in modo casuale N_{mut} gruppi di geni che se positivi, verranno mutati in negativi, se negativi restano tali (i.e. se l'elemento selezionato è 0 viene trasformato in 1, mentre se è 1 rimane tale).
- **Random Genoma *RG***: se è impostato a 1, il genoma dei salmoni è generato in modo casuale, altrimenti è completamente sano (cioè tutti 0).
- **Death After Birth *DB***: se è settato a 1, il salmone che ha figli muore dopo aver procreato, se 0 il salmone dopo la procreazione continua a vivere.

Il modello implementato non considera un numero massimo di iterazioni ma prevede l'arresto dell'algoritmo solo quando la popolazione dei salmoni si estingue, fatto che accade sicuramente prima o poi in quanto possono verificarsi solo mutazioni genetiche negative. Ognuna delle iterazioni fatte è da considerarsi inoltre come il passare di un anno di tempo.

4 Risultati

Eseguendo numerosi test è stato notato che la variazione di alcuni parametri non influisce in modo significativo sui risultati ottenuti. Sono elencati di seguito questi parametri con il valore di default fissato:

- Popolazione iniziale di salmoni S fissata a $2 * 10^4$ esemplari.
- Popolazione massima S_{max} di $2 * 10^6$ esemplari.
- Lunghezza del genoma G pari a 32.
- È sempre stato considerato che la popolazione iniziale sia sana quindi gli esemplari iniziali abbiano un corredo genetico formato da soli zeri ($RG = 0$).
- Numero di figli per salmone F pari a 5.
- Soglia di mutazioni negative osservate T uguale a 3.
- Età riproduttiva dei salmoni R pari a 5 anni.

Inoltre è stato previsto che lo script restituisca due grafici:

1. **Demografia della popolazione**: È mostrato il numero totale di salmoni nelle varie iterazioni effettuate.

2. **Aspettativa media di vita:** È mostrato il tasso di sopravvivenza dei salmoni all'età successiva calcolato come

$$\text{AspettativaVita} = S_k(t)/S_{k+1}(t+1)$$

dove k è l'età e t il tempo.

Gli esperimenti effettuati mostrano come si differenzia l'andamento della popolazione al variare dei parametri. In tutti i test si verifica una accentuata crescita iniziale della popolazione a cui segue una fase di decrescita che porta alla morte di tutti gli esemplari, che in base al settaggio dei parametri può avere diverso andamento.

Nella seguente tabella sono mostrati i valori che i parametri assumono ad ogni esperimento.

<i>Parametri</i>	<i>Test 1</i>	<i>Test 2</i>	<i>Test 3</i>	<i>Test 4</i>	<i>Test 5</i>	<i>Test 6</i>
N_{mut}	1	1	1	3	3	3
V	0	1	1	1	1	1
AR	1	1	0	1	1	0
EA	0	0	1	1	1	1
BM	1	1	1	0	0	1
DB	0	0	0	0	1	1

Nel primo esperimento, il cui risultato è riportato in Fig. 1, con i parametri settati come mostrato in tabella nella colonna "Test 1", si osserva come il numero di esemplari di salmoni abbia una notevole crescita iniziale che porta alla saturazione della popolazione finché questa non si estingue rapidamente. Nella fase di saturazione c'è un continuo ricambio della popolazione dovuto al fatto che muoiono gli esemplari a causa della vecchiaia o perché hanno un numero di geni negativi attivi che supera la soglia T , mentre nasce un numero di nuovi esemplari di salmoni pari ai deceduti. Infine il decadimento è repentino perché l'incidenza del numero di geni negativi accumulati e trasmessi ai figli diventa il fattore dominante.

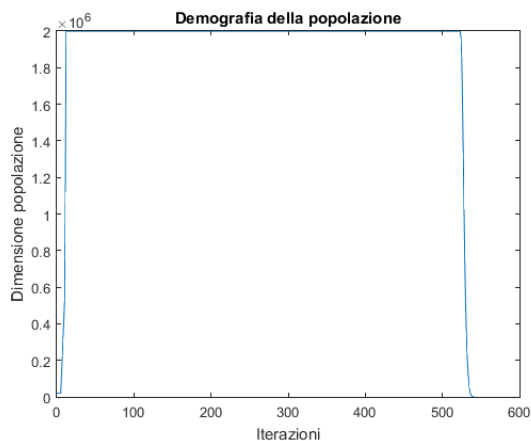


Figura 1: Andamento demografico della popolazione dei salmoni nel Test 1.

Rispetto al caso precedente, nel secondo test sono stati considerati i medesimi parametri, ma con il Verhulst Factor attivo. Questo impedisce alla popolazione di saturare portando all'estinzione in modo più dolce (Fig. 2).

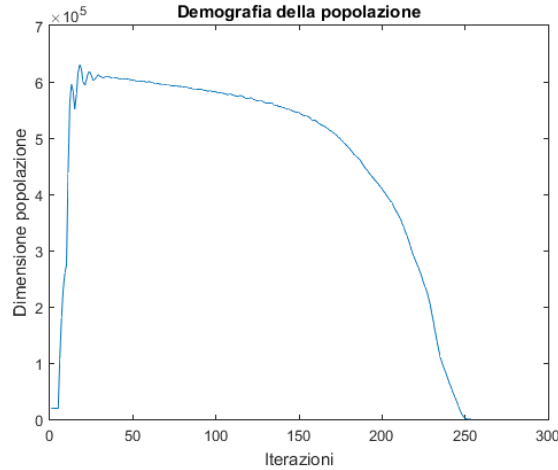


Figura 2: Andamento demografico della popolazione dei salmoni nel Test 2.

Nel terzo test è stato considerato il caso in cui il Verhulst Factor è attivo, i salmoni non si riproducono sempre ma solo con una certa probabilità (data dalla formula 2) ed esclusivamente all'età prestabilita. Come si nota dalla Fig. 3, l'andamento oscilla notevolmente senza mai arrivare alla saturazione. Differentemente dal caso precedente in cui la nascita dei salmoni poteva avvenire ogni anno e compensava il numero dei salmoni morti, qui i salmoni si riproducono solo ogni 5 anni e quindi tra due fasi di riproduzione c'è solo decadimento del numero di esemplari. Anche in questo esperimento il processo termina con l'estinzione di tutti i salmoni dovuta all'aumentare del numero dei geni negativi attivi.

Il parametro principale che cambia nel quarto esperimento è BM , cioè non è certo che si verifichi la mutazione negativa dal padre al figlio. Come si vede in Fig. 4, l'andamento oscillante è simile al test 3. Qui il grafico non mostra l'intera evoluzione della popolazione, ma è stato necessario interrompere la computazione dopo 1000 iterazioni perché il sicuro decadimento degli esemplari avveniva in modo estremamente lento. La causa principale è che non sempre la mutazione è presente; questo, unito al fatto che inizialmente la popolazione è completamente sana e che gli esemplari morti a causa del Verhulst Factor possono essere proprio quelli con più geni negativi, porta ad una ritardata estinzione.

Come ultimo esperimento vediamo in Fig. 5 l'evolversi della popolazione nel caso in cui sia imposta la morte degli esemplari che si riproducono. Si è aumentato il numero di mutazioni negative generate da 1 a 3 e in modo analogo all'esperimento 3, la popolazione cresce solo ogni 5 anni, cioè nel momento in cui si riproducono i salmoni. Il declino della popolazione in questo test è più rapido a causa del numero di mutazione generate e perché i genitori muoiono

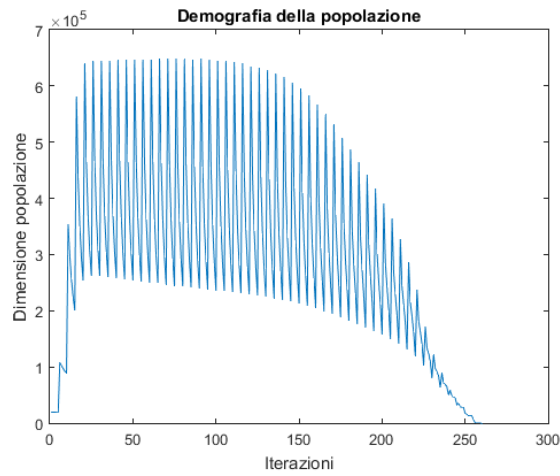


Figura 3: Andamento demografico della popolazione dei salmoni nel Test 3.

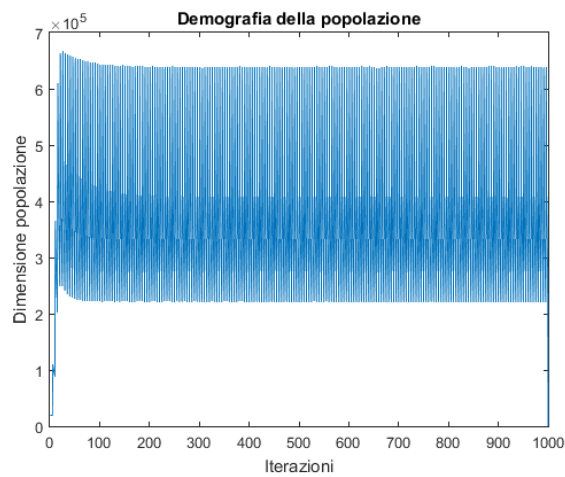


Figura 4: Andamento demografico della popolazione dei salmoni nel Test 4.

subito dopo aver creato i figli.

Altri risultati interessanti sono presentati nei grafici seguenti, che mostrano l'aspettativa di vita media dei salmoni che, come si intuisce dal nome, indica la probabilità che ha un salmone all'età k di sopravvivere fino all'età $k + 1$.

Come si vede in Fig. 6, in cui sono stati settati i parametri come nel test 2, è più alta la probabilità di sopravvivere all'età successiva dagli esemplari di salmoni giovani rispetto a quelli più anziani fino a al limite di età possibile. Questo rispecchia il grafico di Fig. 2 dove la popolazione ha una decrescita continua man mano che passa il tempo.

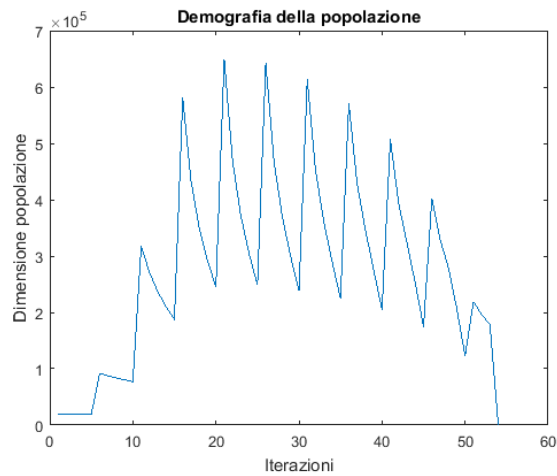


Figura 5: Andamento demografico della popolazione dei salmoni nel Test 5.

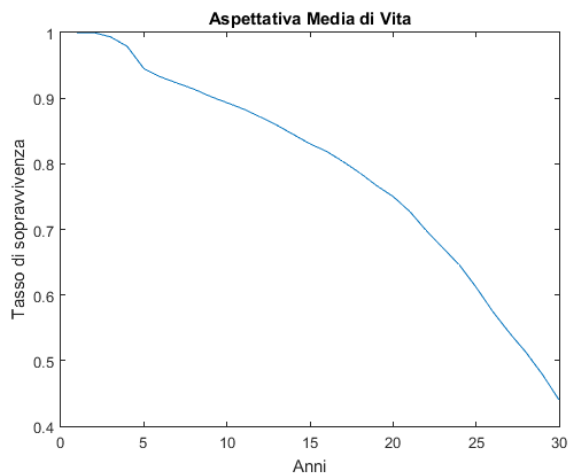


Figura 6: Aspettativa media di vita relativa al Test 1

Un caso completamente differente dal precedente è ottenuto nel test 5: il fattore principale della rapida decadenza della probabilità di sopravvivenza è la morte degli esemplari dopo la procreazione. In particolare si nota che, anche ipotizzando la presenza del Verhulst Factor, il fattore dominante resta DB che come si può vedere in Fig. 7 porta alla morte di tutti gli esemplari quando raggiungono l'età R .

Interessante è osservare il grafico dell'aspettativa di vita del Test 6, dove i parametri che cambiano rispetto al test 5 sono la BM e AR , cioè è possibile la presenza di sole mutazioni negative e i salmoni non è detto che riproducono appena raggiungono l'età R . Si nota che è molto simile al grafico precedente, con la differenza che qualche esemplare sopravvive all'età R per poi morire poco dopo a causa dei geni negativi attivi, casualmente per Verhulst oppure

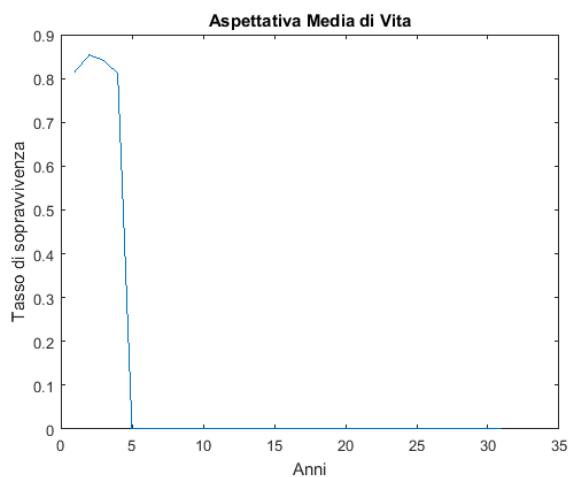


Figura 7: Aspettativa media di vita relativa al Test 5

procreando. In questo caso il grafico ha un andamento simile a quello di Fig. 7 riportato ma ripetuto come si può vedere dalla Fig. 8 in quando il parametro fondamentale resta *DB*.

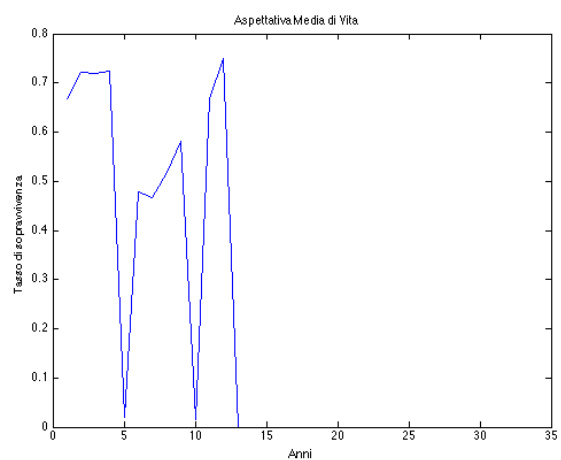


Figura 8: Aspettativa media di vita relativa al Test 6

Infine abbiamo analizzato il grafico dell'aspettativa di vita media relativo al Test 3, che presenta forti fluttuazioni di probabilità come si nota in Fig. 9. Le fluttuazioni sono dovute al fatto che il grafico è stato calcolato come media di tutti i salmoni che ci sono ad una certa età diviso quelli che sopravvivono all'età successiva, quindi queste fluttuazioni non tengono conto dell'effettivo numero di esemplari che ci sono nella popolazione (i.e. l'ultima fluttuazione dipende dal fatto che dei 4 esemplari sopravvissuti fino all'età di 27 anni ne sopravvivono 3 fino ad arrivare a 28 anni).

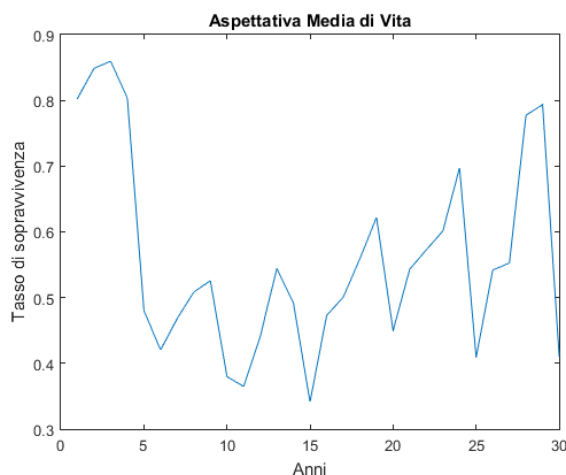


Figura 9: Aspettativa media di vita relativa al Test 3

5 Conclusioni

Lo sviluppo dell'elaborato ed i test effettuati ci hanno permesso di modellare il *Penna Model* in varie forme. I risultati ci hanno mostrato come gli andamenti diversi sul numero di esemplari della popolazione e sulla loro aspettativa media di vita dipendano fortemente dalla probabilità di morte casuale dovuta al Verhulst Factor e dalla presenza "forzata" di mutazioni negative oltre che alla soglia di morte T sui geni negativi. Questi tre fattori discriminano maggiormente i test effettuati e l'andamento complessivo della popolazione di salmoni osservata, basti pensare che considerare la presenza di mutazioni positive potrebbe anche mai portare, secondo questo modello, all'estinzione totale della specie. È stato visto come l'uso accorto di altri parametri come la probabilità di procreazione AR , la procreazione esclusiva ad una età o continua nel tempo EA e il numero di mutazioni che avvengono nella procreazione N_{mut} possono più o meno velocemente portare all'estinzione della specie. Infatti seppur la popolazione in alcuni casi tende a stabilizzarsi, l'assenza di mutazioni positive per come abbiamo costruito l'algoritmo porterà sempre alla morte dell'intera popolazione.

Riferimenti bibliografici

- [1] D. Stauffer, *The Penna Model of Biological Aging*, Bioinformatics and Biology Insights (2007), Pagg. 91-100
- [2] Z. F. Huang e D. Stauffer, *Stochastic Penna model for biological aging*, Theory in Biosciences 120.1 (2001), Pagg. 21-28
- [3] J. Dabkowski, M. Groth e D. Makowiec, *Verhulst Factor In The Penna Model Of Biological Aging*, CiteSeer (2000)