HAL
open science

Una breve storia matematica della dinamica delle popolazioni

Nicolas Bacaër, Bruno Felice Filippo Flora, Yvan Castin, Mimmo Iannelli, Elisa Affili, Bruno Buonomo, Alessandro Margheri, Chiara Poletto, Luca Rossi, Ezio Venturino

To cite this version:

Nicolas Bacaër, Bruno Felice Filippo Flora, Yvan Castin, Mimmo Iannelli, Elisa Affili, et al.. Una breve storia matematica della dinamica delle popolazioni. [s.n.], 159 p., 2021, 9791034351398. hal-03313544

HAL Id: hal-03313544

https://hal.science/hal-03313544

Submitted on 27 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Nicolas Bacaër

con

Bruno Felice Filippo Flora

e con la partecipazione di

Yvan Castin, Mimmo Iannelli,
Elisa Affili, Bruno Buonomo, Alessandro Margheri,
Chiara Poletto, Luca Rossi, Ezio Venturino

Una breve storia matematica della dinamica delle popolazioni

Una breve storia matematica della dinamica delle popolazioni

Nicolas Bacaër

con

Bruno Felice Filippo Flora

e con la partecipazione di

Yvan Castin, Mimmo Iannelli,
Elisa Affili, Bruno Buonomo, Alessandro Margheri,
Chiara Poletto, Luca Rossi, Ezio Venturino

Nicolas Bacaër
Institut de recherche pour le développement
nicolas.bacaer@ird.fr

Bruno Felice Filippo Flora
Studio d'Ingegneria Clinica
brunofelicefilippo.flora@ingpec.eu

I lettori che desiderano acquistare la versione cartacea di questo libro possono inviare una mail a nicolas.bacaer@ird.fr.

«Quanti prima di allora rifiutavano la lingua di Roma, adesso aspiravano all'eloquenza. Nella loro inesperienza chiamavano tutto questo civiltà, mentre non era che un aspetto del loro asservimento.»
Tacito, *Agricola*

«La lingua dell'Europa è la traduzione.»
Umberto Eco

Ad Alice Sinatra, che ha rifiutato di insegnare fisica in inglese all'*École normale supérieure* di Parigi.

Foto di copertina:
Maison de l'Italie alla Città internazionale universitaria di Parigi

Titre original : *Histoires de mathématiques et de populations*.
© Cassini, Paris, 2008

Pour l'édition italienne :
© Nicolas Bacaër, 2021
ISBN : 9791034351398
Dépôt légal : mai 2021

Prefazione

La dinamica delle popolazioni è l'area della scienza che cerca di spiegare in modo semplice e meccanicistico le variazioni nel tempo delle dimensioni e della composizione delle popolazioni biologiche, come quelle di uomini, animali, piante o microrganismi. È collegata, ma ancora ben distinta, dall'area più descrittiva della statistica della popolazione. Un punto comune è che fanno un ampio uso del linguaggio matematico.

La dinamica delle popolazioni si trova all'intersezione di vari campi: matematica, scienze sociali (demografia), biologia (genetica ed ecologia delle popolazioni) e medicina (epidemiologia). Di conseguenza, non è spesso presentata come un insieme, nonostante le somiglianze tra i problemi incontrati nelle varie applicazioni. Un'eccezione degna di nota in francese è il libro *Teorie matematiche delle popolazioni* di Alain Hillion[1]. Ma presenta l'argomento dal punto di vista del matematico, distinguendo vari tipi di modelli: modelli a tempo discreto ($t = 0, 1, 2...$) e modelli a tempo continuo (t è un numero reale), modelli deterministici (gli stati futuri sono noti esattamente se lo stato presente è noto esattamente) e modelli stocastici (dove le probabilità giocano un ruolo). Il libro considera poi modelli deterministici logicamente discreti, modelli deterministici continui, modelli stocastici discreti e modelli stocastici continui.

Nel presente libro ho cercato di discutere lo stesso argomento ma da un punto di vista storico. La ricerca è spiegata nel suo contesto. Sono incluse brevi biografie di scienziati. Questo dovrebbe rendere il libro più facile da leggere per coloro che hanno meno familiarità con la matematica e di solito può aiutare a capire l'origine dei problemi in studio. Ma questo libro non è solo storia. Può anche servire come introduzione alla modellazione matematica. Ci è sembrato importante includere i dettagli della maggior parte dei calcoli in modo che il lettore possa davvero vedere i limiti dei modelli. Le parti tecniche sono sottolineate in caselle grigie e possono essere saltate alla prima lettura. L'ultimo capitolo si concentra sui numerosi problemi contemporanei della dinamica delle popolazioni che si può cercare di analizzare da un punto di vista matematico.

Non era possibile, in un libro di questa lunghezza, dare un quadro completo di tutti i lavori sviluppati fino ad ora o parlare di tutti gli scienziati che hanno contribuito all'argomento. La scelta fatta contiene necessariamen-

[1] *Presses Universitaires de France*, Parigi, 1986.

te una componente arbitraria, in particolare per i decenni più recenti. Spero tuttavia che il campione scelto sia sufficientemente rappresentativo, e che le persone attive nel campo il cui lavoro non è menzionato non saranno turbate.

Questo libro è rivolto a:

- Studenti delle scuole superiori e dell'università che si chiedono quali legami possano esistere tra i corsi di matematica che devono frequentare e il mondo che li circonda, o studenti che preparano un lavoro personale su un tema legato alla dinamica della popolazione.

- Gli insegnanti di matematica che cercano di rendere il loro corso più attraente. La conoscenza delle quattro operazioni elementari è sufficiente per comprendere la maggior parte dei capitoli 1, 2 e 5. Il capitolo 3 può servire come introduzione alle applicazioni dei logaritmi. Questo libro copre anche: equazioni di ricorrenza nei capitoli 1, 3, 8, 11, 14, 21, 23, 24; equazioni differenziali nei capitoli 4, 6, 12, 13, 16; equazioni differenziali alle derivate parziali nei capitoli 20, 25; un'equazione integrale nel capitolo 10; e applicazioni della teoria della probabilità nei capitoli 2, 7, 8, 9, 15, 16, 17, 18, 19, 22.

- Studiosi che hanno già familiarità con la demografia, l'epidemiologia, la genetica o l'ecologia e che vogliono confrontare la loro area preferita con altre che possono coinvolgere modelli matematici simili; lettori interessati alla storia della scienza.

Questo libro è essenzialmente una traduzione dell'edizione francese pubblicata da Cassini Editori (Parigi) nel 2008 con il titolo *Histoires de mathématiques et de populations*. Le liste di riferimenti alla fine di ogni capitolo sono state ampliate e aggiornate. Queste liste includono siti web che mostrano i lavori originali. Una traduzione inglese è stata pubblicata con il titolo *A Short History of Mathematical Population Dynamics* (Springer, 2011).

Vorrei ringraziare Andrea Pugliese che, invitandomi alla difesa di una tesi all'Università di Trento nel 2014, mi ha involontariamente portato a interessarmi a problemi linguistici. Per la traduzione italiana, sono molto grato a Bruno Felice Filippo Flora (14 capitoli di libro), Yvan Castin (tre capitoli), Mimmo Iannelli e un correttore anonimo (due capitoli), Elisa Affili, Bruno Buonomo, Alessandro Margheri, Chiara Poletto, Luca Rossi e Ezio Venturino (un capitolo), che hanno rivisto e corretto la traduzione automatica del software DeepL. Non vedo l'ora che arrivi il giorno in cui mi chiederanno di correggere la traduzione francese dei loro libri.

Capitolo 1

La successione di Fibonacci (1202)

Nel 1202, Leonardo da Pisa, chiamato anche Fibonacci, pubblicò un libro che rese popolare in Europa il sistema numerico decimale indiano che era stato già adottato anche dai matematici arabi. Tra i molti esempi riportati nel libro, uno si riferisce alla crescita di una popolazione di conigli. Si tratta di uno dei più antichi esempi di modello matematico per la dinamica di una popolazione.

Leonardo di Pisa, chiamato Fibonacci molto tempo dopo la sua morte, nacque intorno al 1170 nella Repubblica di Pisa quando era all'apice della sua potenza commerciale e militare nel mondo mediterraneo. Intorno al 1192, il padre di Fibonacci fu inviato dalla Repubblica al porto di Bejaia, ora in Algeria, per dirigere un posto di commercio. Suo figlio lo raggiunse poco dopo per prepararsi ad essere un mercante. Leonardo cominciò a imparare il sistema numerico decimale che gli arabi avevano portato dall'India e che è ancora in uso oggi quasi nella stessa forma: 0, 1, 2, 3, 4, 5, 6, 7, 8 e 9. Mentre viaggiava per affari nel Mediterraneo, confrontò i diversi sistemi numerici e studiò la matematica araba. Tornato a Pisa, portò a termine la scrittura nel 1202 un libro in latino intitolato *Liber abaci* in cui spiegava il nuovo sistema numerico e mostrava come usarlo per la contabilità, le conversioni di peso e di valuta, i tassi di interesse e molte altre applicazioni. Raccolse anche la maggior parte dei risultati in algebra e aritmetica conosciuti dagli arabi.

Fibonacci presentò nel suo libro quello che oggi si chiamerebbe un problema di dinamica delle popolazioni. Ma appariva semplicemente come un esercizio di calcolo in mezzo ad altri argomenti non correlati: il paragrafo precedente del libro è dedicata ai numeri perfetti che sono la somma dei loro fattori, come $28 = 14 + 7 + 4 + 2 + 1$, e il paragrafo seguente tratta un problema sulla divisione del denaro tra quattro persone che è equivalente a un sistema lineare di quattro equazioni. Ecco una traduzione dal latino del problema di una popolazione di conigli:

> «Un tale mise una coppia di conigli in un luogo completamente circondato da un muro, per scoprire quante coppie di conigli discendessero da questa in un anno : per natura le coppie di conigli generano ogni mese un'altra coppia e cominciano a procreare a partire dal secondo mese dalla nascita.»

Se c'è una coppia di conigli appena nati all'inizio del primo mese, questa coppia non sarà ancora fertile dopo un mese e ci sarà ancora una sola coppia di conigli all'inizio del secondo mese. Questa coppia di conigli darà alla luce un'altra coppia all'inizio del terzo mese, quindi ci saranno due coppie in totale. La coppia iniziale di conigli darà di nuovo alla luce un'altra coppia all'inizio del quarto mese. Ma la seconda coppia di conigli non sarà ancora fertile. Ci saranno solo tre coppie di conigli.

Usando notazioni moderne, sia P_n il numero di coppie di conigli all'inizio del mese n. Il numero di coppie di conigli P_{n+1} nel mese $n+1$ è la somma del numero P_n di coppie nel mese n e del numero di coppie appena nate nel mese $n+1$. Ma solo le coppie di conigli che hanno almeno due mesi danno vita a nuove coppie di conigli nel mese $n+1$. Queste sono le coppie che erano già presenti nel mese $n-1$ e il loro numero è P_{n-1}. Quindi

$$P_{n+1} = P_n + P_{n-1}.$$

Questa è una relazione di ricorrenza fornisce la popolazione del mese $n+1$ in funzione della popolazione dei mesi precedenti. Quindi Fibonacci potrebbe facilmente costruire la seguente tabella, dove $1+1=2$, $1+2=3$, $2+3=5$, $3+5=8$, ecc.

n	1	2	3	4	5	6	7	8	9	10	11	12	13
P_n	1	1	2	3	5	8	13	21	34	55	89	144	233

Infatti, Fibonacci considerò come condizione iniziale la situazione nel mese $n=2$. Poiché $P_{14} = 144+233 = 377$, ottenne infine 377 coppie di conigli dodici mesi dopo il suo punto di partenza. Notò cominque questa successione di numeri poteva continuare all'infinito.

Dopo il 1202 Fibonacci scrisse molti altri libri, come *Practica geometriae* nel 1220 e *Liber quadratorum* nel 1225. La sua reputazione portò ad un incontro con l'imperatore Federico II, che apprezzava la scienza. Nel 1240 la Repubblica di Pisa assegnò a Fibonacci una pensione annuale. L'anno della sua morte è sconosciuto.

Durante i secoli successivi, il problema del coniglio di Fibonacci fu dimenticato e non ebbe alcuna influenza sullo sviluppo di modelli matematici per la dinamica delle popolazioni. Diversi scienziati incontrarono la stessa successione di numeri nei loro studi ma non fecero riferimento a Fibonacci o a nessuna popolazione. Diversi libri di Keplero contengono l'osservazione

che il rapporto P_{n+1}/P_n converge, quando n tende all'infinito, al numero aureo $\phi = (1+\sqrt{5})/2$. Questo è un caso particolare di una proprietà comune alla maggior parte dei modelli di popolazione: la tendenza ad aumentare geometricamente (vedi i capitoli 3 e 21). Nel 1728 Daniel Bernoulli ottenne la formula esatta

$$P_n = \frac{1}{\sqrt{5}}\left[\frac{1+\sqrt{5}}{2}\right]^n - \frac{1}{\sqrt{5}}\left[\frac{1-\sqrt{5}}{2}\right]^n$$

mentre studiava le serie ricorrenti generali. Le opere complete di Fibonacci furono pubblicate nel XIX secolo. Da allora, la successione (P_n) poteva essere trovata nei libri di matematica ricreativa sotto il nome di successione di Fibonacci.

È chiaro che, per modellizzare una popolazione di conigli, le ipotesi che portano alla successione di Fibonacci sono tutt'altro che realistiche: nessuna mortalità, nessuna separazione dei sessi ecc. Il nostro interesse per questa successione negli ultimi decenni in biologia è venuto dal fatto che diverse piante contengono strutture che coinvolgono alcuni dei numeri P_n, per esempio, 8 e 13 nelle pigne o 34 e 55 nei girasoli. Una rivista scientifica, *The Fibonacci Quarterly*, è addirittura interamente dedicata alle proprietà e alle applicazioni della successione di Fibonacci!

Bibliografia

1. Ancora, L.: *Il Liber Abaci di Leonardo Fibonacci tradotto in italiano* (2018) https://archive.org
2. Bernoulli, D.: Observationes de seriebus... *Comment. Acad. Sci. Imp. Petropolitanae* 3, 85–100 (1728/1732) → *Die Werke von Daniel Bernoulli*, Band 2, Birkhäuser, Basel, 1982, 49–64.
3. Vogel, K.: Leonardo Fibonacci. In: Gillespie, C.C. (ed.) *Dictionary of Scientific Biography*, vol. 4, 604–613. Scribner, New York (1971)

Capitolo 2

La tavola di mortalità di Halley (1693)

Nel 1693 il famoso astronomo inglese Edmond Halley studiò i registri di nascita e morte della città di Breslavia, che erano stati trasmessi alla *Royal Society* da Caspar Neumann. Halley ha prodotto una tavola di mortalità che mostra il numero di persone che sopravvivono fino a una età data da una coorte nata lo stesso anno. Usò anche la sua tavola per calcolare il prezzo delle rendite vitalizie. Questo capitolo richiama questo lavoro e lo colloca nel contesto della vita di Halley e dei primi sviluppi dell'«aritmetica politica» e della teoria delle probabilità, che interessarono persone come Graunt, Petty, De Witt, Hudde, Huygens, Leibniz e de Moivre.

Edmond Halley nacque vicino a Londra nel 1656. Suo padre era un ricco fabbricante di sapone. Edmond si interessò all'astronomia in giovane età. Iniziò a studiare al *Queen's College* dell'Università di Oxford. Quando l'Osservatorio di Greenwich fu inaugurato nel 1675, Halley poteva già incontrare Flamsteed, l'Astronomo Reale. Interruppe i suoi studi dal 1676 al 1678 per andare sull'isola di Sant'Elena e stabilire un catalogo delle stelle che possono essere viste dall'emisfero sud. Al suo ritorno in Inghilterra divenne membro della *Royal Society*. Pubblicò anche le osservazioni che aveva fatto sulla circolazione dei venti durante il suo viaggio a Sant'Elena. Nel 1684 visitò Newton a Cambridge per discutere il legame tra le leggi del moto planetario di Keplero e la forza di attrazione esercitata dal Sole. Incoraggiò Newton a scrivere il famoso *Philosophiae Naturalis Principia Mathematica*, un libro che alla fine pubblicò a sue spese. Halley lavorava allora come impiegato della *Royal Society*. Nel 1689 progettò una campana per le immersioni subacquee, che testò lui stesso.

Più o meno nello stesso periodo, Caspar Neumann, un teologo che viveva a Breslavia, stava raccogliendo dati sul numero di nascite e morti nella sua città. Breslavia apparteneva all'impero asburgico (ora è in Polonia e si chiama Wrocław). I dati includevano l'età in cui le persone erano morte. Potevano quindi essere usati per costruire una tavola di mortalità che mostrava la probabilità di sopravvivere fino ad una certa età.

La prima tavola di mortalità era stata pubblicata a Londra nel 1662 in un libro intitolato *Osservazioni naturali e politiche fatte sui bollettini di morta-*

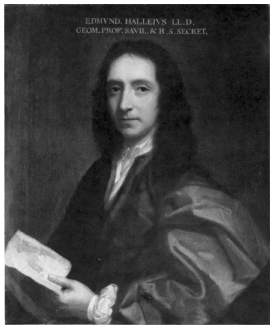

Figura 2.1:
Edmond Halley (1656–1742)

lità. Questo libro è solitamente considerato come il testo fondatore sia della statistica che della demografia e ha una strana particolarità: ancora oggi ci si chiede se sia stato scritto da John Graunt, un mercante londinese e autore indicato sulla copertina del libro, o dal suo amico William Petty, uno dei fondatori della *Royal Society*[1]. Comunque sia, la tavola di mortalità contenuta nel libro cercava di trarre vantaggio dai bollettini che riportavano regolarmente le sepolture e i battesimi a Londra dall'inizio del XVII secolo. Questi bollettini servivano principalmente per documentare le ricorrenti epidemie di peste. Questo è il motivo per il quale indicavano la causa della morte e non l'età in cui si moriva. Per ottenere una tavola di mortalità che desse la probabilità di sopravvivenza in funzione dell'età, Graunt o Petty dovevano indovinare come le diverse cause di morte fossero legate ai gruppi di età. Per ciò la loro tavola di mortalità poteva essere soggetta a grandi errori. Il libro ebbe comunque molto successo, con cinque edizioni tra il 1662 e il 1676. Diverse città in Europa avevano iniziato a pubblicare bollettini simili a quelli di Londra.

Fu quindi quasi trent'anni dopo questa prima tavola di mortalità che, seguendo il suggerimento di Leibniz, Neumann inviò a Henry Justel, il segretario della *Royal Society*, i suoi dati demografici della città di Breslavia per gli anni 1687–1691. Justel morì poco dopo, e Halley entrò in possesso dei dati, li analizzò e nel 1693 pubblicò le sue conclusioni nelle *Philosophical Transactions of the Royal Society*. Il suo articolo si chiama *Una stima dei gradi di mortalità del genere umano, ricavata da curiose tabelle delle nascite e dei funerali nella città di Breslavia, con un tentativo di accertare il prezzo delle rendite vitalizie*.

Per il periodo di cinque anni oggetto di studio, Halley notò che il numero di nascite a Breslavia era più o meno uguale al numero di morti, così che la popolazione totale era quasi costante. Per semplificare l'analisi, egli assunse

[1] Per una discussione dettagliata, vedere il libro di Hervé Le Bras nei riferimenti bibliografici.

che la popolazione fosse esattamente allo stato stazionario: il numero annuale di nascite (chiamato P_0), la popolazione totale, la popolazione di età k (P_k) e il numero annuale di morti all'età k (D_k) sono tutti costanti nel tempo. Questo sottolinea un'altra interessante proprietà dei dati di Breslavia, perché una tale semplificazione non sarebbe stata possibile per una città in rapida crescita come Londra, dove le statistiche erano anche distorte dal flusso di popolazione proveniente dalla campagna.

Tabella 2.1: La tavola di mortalità di Halley che mostra la popolazione P_k di età k.

k	P_k	k	P_k	k	P_k	k	P_k	k	P_k	k	P_k
1	1000	15	628	29	539	43	417	57	272	71	131
2	855	16	622	30	531	44	407	58	262	72	120
3	798	17	616	31	523	45	397	59	252	73	109
4	760	18	610	32	515	46	387	60	242	74	98
5	732	19	604	33	507	47	377	61	232	75	88
6	710	20	598	34	499	48	367	62	222	76	78
7	692	21	592	35	490	49	357	63	212	77	68
8	680	22	586	36	481	50	346	64	202	78	58
9	670	23	579	37	472	51	335	65	192	79	49
10	661	24	573	38	463	52	324	66	182	80	41
11	653	25	567	39	454	53	313	67	172	81	34
12	646	26	560	40	445	54	302	68	162	82	28
13	640	27	553	41	436	55	292	69	152	83	23
14	634	28	546	42	427	56	282	70	142	84	20

I dati di Breslavia avevano una media di 1 238 di nascite all'anno: questo è il valore che Halley prese per P_0. In linea di principio potrebbe anche calcolare dai dati la media annuale D_k del numero di morti tra le persone di età k per tutti $k \geqslant 0$. Usando la formula

$$P_{k+1} = P_k - D_k\,, \tag{2.1}$$

potrebbe costruire la tavola 2.1 che dà P_k. Viceversa, si possono trovare i valori di D_k da lui utilizzati dalla formula $D_k = P_k - P_{k+1}$: $D_0 = 238$, $D_1 = 145$, $D_2 = 57$, $D_3 = 38$ e così via. In realtà, Halley riorganizzò un po' i suoi risultati, sia per ottenere numeri tondi (è il caso di D_1, che è stato leggermente modificato in modo che $P_1 = 1\,000$) sia per smussare certe irregolarità dovute al piccolo numero di morti in età avanzata in uno studio quinquennale. Prendendo la somma di tutti i numeri P_k della tavola, Halley[2] ottenne una stima

[2] Per le persone di più di 84 anni, Halley ha soltanto detto che il loro numero è 107.

della popolazione totale di Breslavia vicina a 34 000. In sintesi, questo metodo aveva il grande vantaggio di non richiedere un censimento generale, ma solo la conoscenza durante alcuni anni del numero di nascite e morti e dell'età in cui le persone morivano.

La tavola di mortalità di Halley servì come riferimento per vari lavori nel XVIII secolo (vedi Capitolo 4). Infatti, anche se i valori di P_k erano specifici della città di Breslavia, si poteva considerare che il rapporto P_{k+1}/P_k era la probabilità di sopravvivere fino all'età $k+1$ sapendo di aver già raggiunto l'età k. Questa probabilità potrebbe essere ragionevolmente usata per le popolazioni di altre città europee dell'epoca. Per esempio, ci si potrebbe aspettare che un bambino di un anno abbia 661 possibilità su 1 000 di raggiungere i 10 anni o 598 possibilità su 1 000 di raggiungere i 20 anni.

Halley usò anche la sua tavola di mortalità per calcolare il prezzo delle rendite vitalizie. Durante il XVI e il XVII secolo, diverse città e stati avevano venduto tali rendite ai loro cittadini per raccogliere denaro. Gli acquirenti ricevevano ogni anno fino alla loro morte una quantità fissa di denaro, che era uguale ad una certa percentuale della somma inizialmente pagata, spesso il doppio del tasso di interesse del tempo, ma indipendentemente dall'età dell'acquirente. Naturalmente l'istituzione rischiava il fallimento se troppe persone con una speranza di vita molto lunga compravano queste rendite. Il problema non poteva essere affrontato correttamente senza una tavola di mortalità affidabile.

Nel 1671 Johan De Witt, primo ministro d'Olanda, e Johannes Hudde, uno dei sindaci della città di Amsterdam, avevano già pensato al problema di calcolare il prezzo delle rendite vitalizie. Temendo un'invasione delle truppe francesi, volevano raccogliere denaro per rafforzare l'esercito. Avevano dei dati riguardanti le persone che avevano comprato delle rendite vitalizie diversi decenni prima, in particolare l'età in cui le rendite erano state comprate e l'età in cui le persone erano morte. Erano riusciti a calcolare il prezzo delle rendite più o meno correttamente, ma il loro metodo fu poi dimenticato. L'Olanda fu invasa l'anno seguente e De Witt fu linciato dalla folla.

Halley considerò nuovamente il problema nel 1693 con la tavola di mortalità di Breslavia e assumendo un tasso di interesse del 6 %. Il metodo di calcolo è semplice. Sia i il tasso di interesse. Sia R_k il prezzo al quale una persona di età k può comprare una rendita di, diciamo, una sterlina all'anno. Questa persona ha una probabilità P_{k+n}/P_k di essere ancora in vita all'età $k+n$. La sterlina che lo Stato promette di pagare se raggiunge questa età può essere ottenuta mettendo $1/(1+i)^n$ sterline della somma iniziale al tasso di interesse i. Quindi, se si fa l'ipotesi semplificatrice che la somma iniziale sia

usata solo per pagare le rendite, allora il prezzo dovrebbe essere

$$R_k = \frac{1}{P_k}\left(\frac{P_{k+1}}{1+i} + \frac{P_{k+2}}{(1+i)^2} + \frac{P_{k+3}}{(1+i)^3} + \cdots\right). \tag{2.2}$$

Halley ha ottenuto in questo modo la tabella 2.2, che mostra il fattore R_k per il quale la rendita desiderata deve essere moltiplicata per ottenere la somma iniziale necessaria. Un uomo di 20 anni otterrebbe quindi ogni anno $1/12{,}78 \approx 7{,}8\,\%$ della somma iniziale. Ma un uomo di 50 anni otterrebbe $1/9{,}21 \approx 10{,}9\,\%$, perché avrebbe meno anni da vivere. Si noti che il doppio del tasso di interesse corrisponderebbe ad una rendita pari al $12\,\%$ della somma iniziale, o equivalentemente ad un prezzo pari a 8,33 volte la rendita.

Tabella 2.2: Fattore moltiplicatore che dà il prezzo delle rendite vitalizie.

Età k	Prezzo R_k	Età k	Prezzo R_k	Età k	Prezzo R_k
1	10,28	25	12,27	50	9,21
5	13,40	30	11,72	55	8,51
10	13,44	35	11,12	60	7,60
15	13,33	40	10,57	65	6,54
20	12,78	45	9,91	70	5,32

I calcoli sono ovviamente abbastanza fastidiosi. Halley poteva comunque usare le tabelle dei logaritmi per ottenere il termine generale $P_{k+n}/(1+i)^n$ più rapidamente. Poiché non ha mostrato valori di P_k oltre agli 84 anni, non è possibile verificare esattamente i suoi calcoli. Infine, il lavoro di Halley non ebbe un impatto immediato: per diversi decenni, le rendite vitalizie in Inghilterra e altrove continuarono ad essere vendute ad un prezzo indipendente dall'età dell'acquirente e ad un prezzo che era molto più basso di quanto avrebbe dovuto essere, per esempio 7 volte la rendita.

Le questioni legate alle tavole di mortalità interessarono molti scienziati durante l'epoca di Halley. L'olandese Christiaan Huygens, autore nel 1657 del primo opuscolo dedicato alla teoria delle probabilità, discusse nel 1669 nella sua corrispondenza con il fratello Graunt la tavola di mortalità e il calcolo della speranza di vita[3]. Qualche anno prima di mettere Neumann in contatto con la *Royal Society*, Leibniz scrisse anche sul calcolo della speranza di vita in un saggio che rimase inedito. Nel 1709 fu il turno di Nicolaus Bernoulli. Nel 1725 Abraham de Moivre pubblicò un intero *Trattato sulle rendite*. Egli notò in particolare che il prezzo R_k poteva essere facilmente

[3] La speranza di vita all'età k è data dalla formula (2.2) con $i = 0$.

calcolato per le persone anziane poiché la formula (2.2) conteneva solo pochi termini. Si può quindi utilizzare la formula di ricorrenza all'indietro

$$R_k = \frac{P_{k+1}}{P_k} \, \frac{1+R_{k+1}}{1+i} \, ,$$

che è facilmente dimostrabile partendo da (2.2). Usando il valore che Halley dà per il prezzo a 70 anni, si possono quindi controllare gli altri valori della tabella 2.2[4].

Dopo questa pausa dedicata alla demografia, Halley tornò ai suoi principali argomenti di ricerca. Tra il 1698 e il 1700 navigò nell'Oceano Atlantico per costruire una mappa del campo magnetico terrestre. Nel 1704 divenne professore all'Università di Oxford. L'anno seguente pubblicò un libro sulle comete e predisse che la cometa del 1682, che Keplero aveva osservato nel 1607, sarebbe tornata nel 1758: divenne nota come «cometa di Halley». Pubblicò anche una traduzione del libro di Apollonio di Perga sulle coniche. Nel 1720 sostituì Flamsteed come Astronomo Reale. Cercò di risolvere il problema di determinare la longitudine in mare con precisione dall'osservazione della Luna, un problema di grande importanza pratica per la navigazione. Morì a Greenwich nel 1742 all'età di 86 anni.

Bibliografia

1. Fox, M.V.: *Scheduling the Heavens: The Story of Edmond Halley*. Morgan Reynolds, Greensboro (2007)
2. Graunt, J.: *Natural and Political Observations Mentioned in a Following Index and Made upon the Bills of Mortality* (1665). echo.mpiwg-berlin.mpg.de
3. Hald, A.: *A History of Probability and Statistics and Their Applications before 1750*. Wiley, Hoboken (2003).
4. Halley, E.: An estimate of the degrees of the mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslaw. *Phil. Trans. Roy. Soc. London* 17, 596–610 (1693). gallica.bnf.fr
5. Heyde, C.C.: John Graunt. In: Heyde, C.C., Seneta, E. (eds.) *Statisticians of the Centuries*, 14–16. Springer, New York (2001)
6. Koch, P.: Caspar Neumann. In: Heyde, C.C., Seneta, E. (eds.) *Statisticians of the Centuries*, 29–32. Springer, New York (2001)
7. Le Bras, H.: *Naissance de la mortalité*. Gallimard, Paris (2000)

[4] Sembra che ci siano alcuni errori nella tabella, in particolare per le età 5 e 15.

Capitolo 3

Eulero e la crescita geometrica delle popolazioni (1748–1761)

Eulero ha scritto in diverse occasioni sulla dinamica della popolazione. Nel suo trattato del 1748, *Introductio in analysin infinitorum*, il capitolo dedicato alla funzione esponenziale conteneva quattro esempi sulla crescita esponenziale di una popolazione. Nel 1760 pubblicò un articolo che combinava questa crescita esponenziale con una struttura per età della popolazione. Questo lavoro è un precursore della teoria delle popolazioni stabili, che è stata sviluppata nel ventesimo secolo e gioca un ruolo importante nella demografia. Nel 1761 Eulero aiutò anche Süssmilch con la seconda edizione del suo trattato di demografia. Elaborò un modello interessante, che è una sorta di variante della successione di Fibonacci, ma non pubblicò la sua analisi dettagliata.

Leonhard Euler nacque nel 1707 a Basilea, in Svizzera. Suo padre era un pastore protestante. Nel 1720 Eulero iniziò a studiare all'università. Ricevette anche lezioni private di matematica da Johann Bernoulli, uno dei più famosi matematici della generazione dopo Leibniz e Newton. Fece amicizia con due dei figli di Johann Bernoulli: Nicolaus II e Daniel. Nel 1727 Eulero raggiunse Daniel alla neonata Accademia delle Scienze di San Pietroburgo. Oltre alla matematica si interessava anche alla fisica e a molte altre materie scientifiche e tecniche. Nel 1741 il re Federico II di Prussia lo invitò a diventare direttore della sezione di matematica dell'Accademia delle Scienze di Berlino. Eulero pubblicò un numero considerevole di articoli e libri su tutti gli aspetti della meccanica (astronomia, elasticità, fluidi, solidi) e della matematica (teoria dei numeri, algebra, serie infinite, funzioni elementari, numeri complessi, calcolo differenziale e integrale, equazioni differenziali e alle derivate parziali, ottimizzazione, geometria) ma anche sulla demografia. Fu il matematico più prolifico del suo tempo.

Nel 1748 Eulero pubblicò un trattato in latino intitolato *Introductio in analysin infinitorum*. Nel capitolo su esponenziali e logaritmi considera sei esempi: uno sulla teoria matematica delle scale musicali, un altro sulla restituzione di un prestito con interesse, e quattro sulla dinamica della popolazione. In quest'ultimo Eulero assunse che la popolazione P_n nell'anno n

Figura 3.1:
Eulero (1707–1783)

soddisfa $P_{n+1} = (1+x)\,P_n$ per tutti gli interi n. Il tasso di crescita x è un numero reale positivo. Partendo da una condizione iniziale P_0, la popolazione nell'anno n è data da $P_n = (1+x)^n\,P_0$. Questa è chiamata crescita geometrica o esponenziale. Il primo esempio chiede:

> «Se la popolazione di una certa regione aumenta annualmente di un trentesimo e un tempo c'erano 100 000 abitanti, vorremmo conoscere la popolazione dopo 100 anni.»

La risposta è $P_{100} = (1+1/30)^{100} \times 100\,000 \approx 2\,654\,874$. Per questo esempio Eulero si è ispirato dal censimento di Berlino che ha avuto luogo nel 1747 e che ha dato una stima di 107 224 abitanti. Il suo calcolo dimostra che una popolazione può aumentare più di dieci volte in un secolo. Questo è esattamente ciò che era stato osservato all'epoca per la città di Londra.

Bisogna notare che calcolare $(1+1/30)^{100}$ è molto facile con una moderna calcolatrice tascabile. Ma ai tempi di Eulero si dovevano usare i logaritmi per evitare numerose moltiplicazioni a mano e ottenere rapidamente il risultato. Si calcola prima il logaritmo decimale (in base 10) di P_{100}. La proprietà fondamentale del logaritmo $\log(a\,b) = \log a + \log b$ mostra che $\log P_{100} = 100\log(31/30) + \log(100\,000) = 100\,(\log 31 - \log 30) + 5$. I logaritmi erano stati introdotti nel 1614 dallo scozzese John Napier. Il suo amico Henry Briggs aveva pubblicato la prima tavola dei logaritmi decimali nel 1617. Nel 1628 l'olandese Adriaan Vlacq aveva completato il lavoro di Briggs pubblicando una tavola dei logaritmi decimali dei numeri interi da 1 a 100 000 con precisione a dieci cifre. Questo è il tipo di tavola che Eulero ha usato per ottenere $\log 30 \approx 1{,}477121255$, $\log 31 \approx 1{,}491361694$, e infine $\log P_{100} \approx 6{,}4240439$. Resta da trovare il numero P_{100} di cui si co-

nosce il logaritmo. Poiché i logaritmi decimali degli interi da 1 a 100.000 vanno da 0 a 5, si cerca invece il logaritmo di $P_{100}/100$, che è 4,4240439. Si può verificare nella tavola dei logaritmi che $\log 26\,548 \approx 4{,}424031809$ e $\log 26\,549 \approx 4{,}424048168$. Sostituendo la funzione logaritmica con una linea retta tra 26 548 e 26 549, Eulero ottiene che

$$\frac{P_{100}}{100} \approx 26\,548 + \frac{4{,}4240439 - 4{,}424031809}{4{,}424048168 - 4{,}424031809} \approx 26\,548{,}74\,.$$

Quindi $P_{100} \approx 2\,654\,874$.

Il secondo esempio riguardante la dinamica della popolazione nel libro di Eulero è il seguente:

> «Poiché dopo il Diluvio tutti gli uomini discendono da una popolazione di sei persone, vorremmo trovare il tasso di crescita annuale supponendo che la popolazione dopo duecento anni fosse di 1 000 000.»

Poiché $10^6 = (1+x)^{200} \times 6$, otteniamo con una calcolatrice tascabile $x = (10^6/6)^{1/200} - 1 \approx 0{,}061963$. Con le tavole dei logaritmi si deve passare attraverso $\log(10^6) = 200\log(1+x) + \log 6$ per ottenere $\log(1+x) = (6 - \log 6)/200 \approx 0{,}0261092$ e $1+x \approx 1{,}061963$. Così Eulero poteva concludere che la popolazione sarebbe aumentata di $x \approx 1/16$ all'anno. Per capire l'origine di questo esempio, bisogna ricordare che i filosofi contemporanei di Eulero stavano iniziando a negare la veridicità dei racconti biblici. Una lettura letterale fisserebbe il tempo del Diluvio intorno al 2350 a.C. con i seguenti sopravvissuti: Noè, i suoi tre figli e le loro mogli. Il libro della Genesi dice:

> «Questi sono i tre figli di Noè; da loro fu popolata tutta la terra.»

Un tasso di crescita della popolazione di $1/16$ (o 6,25 %) all'anno dopo il Diluvio non sembrava troppo irrealistico per Eulero. Essendo il figlio di un pastore protestante ed essendo rimasto religioso per tutta la vita, concluse:

> «Per questo motivo è abbastanza ridicolo per gli increduli obiettare che in un così breve spazio di tempo l'intera Terra non potrebbe essere popolata a partire da un solo uomo[1].»

[1] Nel libro pubblicato da Graunt nel 1662 (si veda il capitolo 2), si trova un'osservazione simile:

> «Una coppia, cioè Adamo ed Eva, raddoppiandosi ogni 64 anni dei 5 160 anni, che è l'età del mondo secondo le Scritture, produrrà molte più persone di quante ce ne siano ora. Perciò il mondo non ha più di 100 000 anni, come alcuni vanamente immaginano, né più di quanto la Scrittura ne faccia.»

Eulero notò anche che se la crescita fosse continuata allo stesso ritmo fino a 400 anni dopo il Diluvio, la popolazione sarebbe stata di $(1+x)^{400} \times 6 = (10^6/6)^2 \times 6 \approx 166$ miliardi:

> «Tuttavia, l'intera Terra non sarebbe mai in grado di sostenere quella popolazione.»

Questa idea sarebbe stata notevolmente sviluppata da Malthus mezzo secolo più tardi (vedi Capitolo 5).

Il terzo esempio di Eulero chiede:

> «Se ogni secolo la popolazione umana raddoppia, qual è il tasso di crescita annuale?»

Poiché $(1+x)^{100} = 2$, otteniamo con una calcolatrice tascabile $x = 2^{1/100} - 1 \approx 0{,}00695$. Con le tavole dei logaritmi, $100\log(1+x) = \log 2$. Quindi $\log(1+x) \approx 0{,}0030103$ e $1+x \approx 1{,}00695$. Quindi la popolazione cresce di $x \approx 1/144$ ogni anno. Il quarto e ultimo esempio si chiede nello stesso modo:

> «Se la popolazione umana aumenta annualmente di 1/100, vorremmo sapere quanto tempo ci vorrà perché la popolazione diventi dieci volte più grande.»

Con $(1+1/100)^n = 10$, troviamo $n\log(101/100) = 1$. Quindi $n = 1/(\log 101 - 2) \approx 231$ anni. Questo è tutto ciò che si può trovare nell'*Introduzione all'analisi dell'infinito* del 1748 riguardante la dinamica della popolazione. Eulero sarebbe tornato su questo argomento in modo più approfondito alcuni anni dopo.

Nel 1760 pubblicò negli atti dell'Accademia delle Scienze di Berlino un lavoro intitolato *Un'indagine generale sulla mortalità e la moltiplicazione della specie umana*. Questo lavoro era una sorta di sintesi tra la sua precedente analisi della crescita geometrica delle popolazioni e i precedenti studi sulle tavole di mortalità (vedi Capitolo 2). Eulero considerò per esempio il problema:

> «Conoscere il numero di nascite e di sepolture che avvengono nel corso di un anno, per trovare il numero di tutti i vivi e il loro aumento annuale, per una data ipotesi di mortalità.»

Eulero assume qui che i seguenti numeri siano noti:

- il numero di nascite B_n durante l'anno n;
- il numero di morti D_n durante l'anno n;

- la proporzione q_k di neonati che raggiungono l'età $k \geqslant 1$.

Sia P_n la popolazione nell'anno n. Eulero ha fatto due supposizioni implicite in più:

- la popolazione aumenta geometricamente: $P_{n+1} = rP_n$ (poniamo $r = 1+x$);
- il rapporto tra nascite e popolazione è costante: $B_n/P_n = m$.

Queste due ipotesi implicano che il numero di nascite aumenta geometricamente e allo stesso ritmo: $B_{n+1} = rB_n$. Eulero ha poi considerato lo stato della popolazione ad intervalli di cento anni, diciamo tra gli anni $n = 0$ e $n = 100$, assumendo che nessuno sopravviva oltre i cento anni. Per chiarire la presentazione, chiamiamo $P_{k,n}$ ($k \geqslant 1$) la popolazione all'inizio dell'anno n, che è nata nell'anno $n-k$. Chiamiamo $P_{0,n} = B_n$ il numero di nascite durante l'anno n. Dalla definizione del coefficiente di sopravvivenza q_k, abbiamo $P_{k,n} = q_k P_{0,n-k} = q_k B_{n-k}$. Quindi

$$\begin{aligned} r^{100} P_0 = P_{100} &= P_{0,100} + P_{1,100} + \cdots + P_{100,100} \\ &= B_{100} + q_1 B_{99} + \cdots + q_{100} B_0 \\ &= (r^{100} + r^{99} q_1 + \cdots + q_{100}) B_0. \end{aligned}$$

Dividendo questa equazione per $r^{100} P_0$, otteniamo

$$1 = m \left(1 + \frac{q_1}{r} + \frac{q_2}{r^2} + \cdots + \frac{q_{100}}{r^{100}}\right). \tag{3.1}$$

Questa è l'equazione che a volte viene chiamata «equazione di Eulero» in demografia. Contando separatamente le nascite e le morti, otteniamo

$$rP_n = P_{n+1} = P_n - D_n + B_{n+1} = P_n - D_n + rB_n\,. \tag{3.2}$$

Quindi anche il numero di morti aumenta geometricamente: $D_{n+1} = rD_n$. Inoltre,

$$\frac{1}{m} = \frac{P_n}{B_n} = \frac{D_n/B_n - r}{1-r}\,. \tag{3.3}$$

Sostituendo questo nell'equazione (3.1), si arriva infine all'equazione

$$\frac{D_n/B_n - 1}{1-r} = \frac{q_1}{r} + \frac{q_2}{r^2} + \cdots + \frac{q_{100}}{r^{100}}\,, \tag{3.4}$$

dove rimane una sola incognita: r. Questa è quella che di solito viene chiamata un'equazione implicita perché non possiamo estrarre r in funzione degli

altri parametri. Ma possiamo calcolare il primo e il secondo membro dell'equazione (3.4) per un valore fisso di r e far variare r fino a che i due membri non sono uguali. Il valore di r così ottenuto dà il tasso di crescita $x = r - 1$ della popolazione. Si noti che dalle equazioni (3.1) e (3.3), si ottiene per la popolazione P_n la seguente espressione:

$$P_n = B_n \left(1 + \frac{q_1}{r} + \frac{q_2}{r^2} + \cdots + \frac{q_{100}}{r^{100}}\right) .$$

Quando la popolazione è stazionaria ($r = 1$), questa espressione è la stessa di quella usata da Halley per stimare la popolazione della città di Breslavia (vedi il capitolo 2).

Eulero considerò anche la seguente domanda:

> «Essendo date le ipotesi di mortalità e fecondità, se si conosce il numero di tutti i viventi, trovare quanti saranno ad ogni età.»

Poiché i coefficienti di sopravvivenza q_k e il coefficiente di fertilità m sono noti, il tasso di crescita r può essere calcolato dall'equazione (3.1). Durante l'anno n, il numero di persone nate nell'anno $n-k$ è $q_k B_{n-k} = q_k B_n / r^k$ (con $q_0 = 1$). Quindi la proporzione della popolazione totale che ha l'età k è

$$\frac{q_k/r^k}{1 + q_1/r + q_2/r^2 + \cdots + q_{100}/r^{100}} .$$

Questa proporzione è costante. Usando la terminologia di Lotka (si veda il capitolo 10), si dice che la popolazione è «stabile»: la piramide delle età mantiene la stessa forma nel tempo.

Eulero ha poi riesaminato il problema di costruire una tavola di mortalità quando la popolazione non è stazionaria ma aumenta geometricamente:

> «Conoscendo il numero di tutti i viventi, si fa la stessa cosa con il numero di nascite e il numero di morti ad ogni età nel corso di un anno, per trovare la legge della mortalità.»

Per legge di mortalità, Eulero intendeva l'insieme dei coefficienti di sopravvivenza q_k. Si suppone ora che la popolazione totale sia nota attraverso un censimento, cosa che non era il caso di Halley (vedi il capitolo 2). L'equazione (3.2) mostra che il tasso di crescita è $r = (P_n - D_n)/(P_n - B_n)$. Sia $D_{k,n}$ il numero di persone che muoiono all'età k nell'anno n: queste persone sono nate nell'anno $n-k$. Quindi $D_{k,n} = (q_k - q_{k+1}) B_{n-k}$. Ma $B_{n-k} = B_n/r^k$. I coefficienti di sopravvivenza q_k possono quindi essere calcolati con la formula di ricorrenza $q_{k+1} = q_k - r^k D_{k,n}/B_n$ per tutti i $k \geqslant 0$, con $q_0 = 1$. Questa

formula moltiplicata per B_n restituisce la formula (2.1) usata da Halley per il caso stazionario $r = 1$. Eulero insisteva tuttavia sul fatto che il suo metodo di calcolo dei coefficienti di sopravvivenza q_k presuppone che la popolazione aumenti regolarmente, escludendo incidenti come epidemie di peste, guerre, carestie ecc. Se i censimenti al tempo di Eulero avessero registrato l'età delle persone (come in Svezia), questa ipotesi non sarebbe stata necessaria e i coefficienti q_k avrebbero potuto essere calcolati più facilmente.

Dati i coefficienti di sopravvivenza q_k, Eulero mostrò anche come calcolare il prezzo delle rendite vitalizie. Non ha menzionato i lavori di Halley o de Moivre su questo argomento. Eulero usò un tasso di interesse del 5 % e la tavola di mortalità pubblicata nel 1742 dall'olandese Willem Kersseboom.

Eulero non era l'unico scienziato interessato alla demografia all'Accademia di Berlino. Il suo collega Johann Peter Süssmilch aveva pubblicato nel 1741 un trattato in tedesco intitolato *L'ordine divino nei cambiamenti della generazione umana, stabilito attraverso la nascita, la morte e la procreazione*, che è considerato oggi come il primo trattato interamente dedicato alla demografia. Süssmilch aveva anche scritto un libro *Sulla rapida crescita della città di Berlino* nel 1752.

Figura 3.2:
Süssmilch (1707–1767)

Nel 1761 Süssmilch pubblicò una seconda edizione del suo trattato. Nel capitolo intitolato *Sul tasso di crescita e sul tempo di raddoppio delle popolazioni*, incluse un interessante modello matematico che Eulero aveva elaborato per lui. Il modello era simile a quello di Fibonacci (vedi Capitolo 1) ma per una popolazione umana. Partendo da una coppia (un uomo e una donna) entrambi di 20 anni nell'anno 0, Eulero assunse che le persone muoiono a 40 anni e si sposano a 20 anni, mentre ogni coppia ha sei figli: due figli (un maschio e una femmina) a 22 anni, altri due a 24 anni e gli ultimi due a 26 anni. Contando gli anni a due a due in modo che B_i sia il numero di nascite durante

l'anno $2i$, Eulero conclude che

$$B_i = B_{i-11} + B_{i-12} + B_{i-13} \tag{3.5}$$

per tutti gli $i \geqslant 1$. Le condizioni iniziali corrispondono a $B_{-12} = 0$, $B_{-11} = 0$, $B_{-10} = 2$ e $B_i = 0$ per $-9 \leqslant i \leqslant 0$. Eulero poteva quindi calcolare il numero di nascite, come mostrato nella seconda colonna della Tabella 3.1. Il numero di morti D_i nell'anno $2i$ è quindi uguale al numero di nascite nell'anno $2i-40$: $D_i = B_{i-20}$ per $i \geqslant 10$ mentre $D_i = 0$ per $i \leqslant 9$. Per quanto riguarda il numero P_i di persone in vita nell'anno $2i$, esso è uguale al numero di persone in vita nell'anno $2i-2$, più il numero di nascite nell'anno $2i$, meno il numero di morti nell'anno $2i$: $P_i = P_{i-1} + B_i - D_i$.

Questo capitolo del libro di Süssmilch finisce con un'osservazione che potrebbe essere già stata fatta sulla successione di Fibonacci:

> «Il grande disordine che sembra prevalere nella tabella di Eulero non impedisce che il numero delle nascite segua una specie di progressione che si chiama serie ricorrente [...] Qualunque sia il disordine iniziale di queste progressioni, esse si trasformano in una progressione geometrica se non vengono interrotte e i disordini dell'inizio si attenuano poco a poco e svaniscono quasi completamente.»

Il libro non dice di più sulla matematica di questo modello di popolazione. Tuttavia, Eulero spinse lo studio molto più in là in un manoscritto intitolato *Sulla moltiplicazione della razza umana*, che rimase inedito durante la sua vita. Cercando una soluzione dell'equazione (3.5) della forma $B_i = c\,r^i$, cioè della forma di una progressione geometrica, ottenne dopo una semplificazione un'equazione polinomiale di grado 13: $r^{13} = r^2 + r + 1$. Cercò una soluzione vicina a $r = 1$ e notò, usando una tavola di logaritmi per il calcolo di r^{13}, che

$$1 + r + r^2 - r^{13} \approx \begin{cases} 0{,}212 & \text{se} \quad r = 1{,}09\,, \\ -0{,}142 & \text{se} \quad r = 1{,}10\,. \end{cases}$$

Quindi l'equazione ha una radice tra 1,09 e 1,10. Approssimando la funzione $1 + r + r^2 - r^{13}$ con un segmento di retta su questo intervallo, Eulero ottenne

$$r \approx \frac{0{,}142 \times 1{,}09 + 0{,}212 \times 1{,}10}{0{,}142 + 0{,}212} \approx 1{,}0960\,.$$

Contando gli anni a due a due, il numero di nascite tende ad essere moltiplicato per $\sqrt{r}$ ogni anno. Questo numero raddoppia ogni n anni se $(\sqrt{r})^n = 2$, cioè

Tabella 3.1: La tabella di Eulero.

i	Nascite	Morti	Vivi	i	Nascite	Morti	Vivi
0	0	0	2	40	20	0	206
1	2	0	4	41	8	0	214
2	2	0	6	42	2	0	216
3	2	0	8	43	0	2	214
4	0	0	8	44	0	6	208
5	0	0	8	45	2	12	198
6	0	0	8	46	10	14	194
7	0	0	8	47	30	12	212
8	0	0	8	48	60	6	266
9	0	0	8	49	90	2	354
10	0	2	6	50	102	0	456
11	0	0	6	51	90	0	546
12	2	0	8	52	60	0	606
13	4	0	12	53	30	0	636
14	6	0	18	54	10	2	644
15	4	0	22	55	2	8	638
16	2	0	24	56	2	20	620
17	0	0	24	57	12	32	600
18	0	0	24	58	42	38	604
19	0	0	24	59	100	32	672
20	0	0	24	60	180	20	832
21	0	2	22	61	252	8	1 076
22	0	2	20	62	282	2	1 356
23	2	2	20	63	252	0	1 608
24	6	0	26	64	180	0	1 788
25	12	0	38	65	100	2	1 886
26	14	0	52	66	42	10	1 918
27	12	0	64	67	14	30	1 902
28	6	0	70	68	16	60	1 858
29	2	0	72	69	56	90	1 824
30	0	0	72	70	154	102	1 876
31	0	0	72	71	322	90	2 108
32	0	2	70	72	532	60	2 580
33	0	4	66	73	714	30	3 264
34	2	6	62	74	786	10	4 040
35	8	4	66	75	714	2	4 752
36	20	2	84	76	532	2	5 282
37	32	0	116	77	322	12	5 592
38	38	0	154	78	156	42	5 706
39	32	0	186	79	72	100	5 678

ogni $n = 2\log 2/\log r \approx 15$ anni. Poiché asintoticamente $B_i \approx c\,r^i$ e poiché il numero D_i di morti nell'anno $2i$ è uguale a B_{i-20}, otteniamo $D_i \approx B_i/r^{20}$ con $r^{20} \approx 6{,}25$. Il numero di nascite è circa sei volte il numero di morti. Il numero P_i di persone in vita nell'anno $2i$ è uguale a $B_i + B_{i-1} + \cdots + B_{i-19}$, otteniamo anche che

$$P_i \approx B_i \left(1 + \frac{1}{r} + \cdots + \frac{1}{r^{19}}\right) = B_i \frac{1 - r^{20}}{r^{19} - r^{20}} \approx 9{,}59\, B_i\,.$$

La popolazione totale è circa dieci volte il numero delle nascite.

La prova che la successione (B_i) mostrata nella Tabella 3.1 cresce davvero asintoticamente come r^i è più complicata. Era noto fin dal lavoro di Abraham de Moivre sulle serie ricorrenti che, introducendo la funzione generatrice $f(x) = \sum_{i=0}^{+\infty} B_i x^i$, si potrebbe esprimere $f(x)$ come una funzione razionale. Eulero aveva spiegato il metodo nella sua *Introductio in analysin infinitorum* nel 1748: la relazione di ricorrenza (3.5) dà infatti

$$\begin{aligned} f(x) &= \sum_{i=0}^{12} B_i x^i + \sum_{i=13}^{+\infty} (B_{i-11} + B_{i-12} + B_{i-13})\, x^i \\ &= 2x + 2x^2 + 2x^3 + 2x^{12} + f(x)\,(x^{11} + x^{12} + x^{13})\,. \end{aligned}$$

Quindi

$$f(x) = \frac{2x + 2x^2 + 2x^3 + 2x^{12}}{1 - x^{11} - x^{12} - x^{13}}\,.$$

Eulero sapeva che una tale funzione razionale poteva essere decomposta nella forma

$$f(x) = \frac{a_1}{1 - \frac{x}{x_1}} + \cdots + \frac{a_{13}}{1 - \frac{x}{x_{13}}}\,,$$

i numeri $x_1, \ldots, x_{13}$ sono le radici reali o complesse dell'equazione $1 - x^{11} - x^{12} - x^{13} = 0$. Quindi

$$f(x) = \sum_{i \geqslant 0} a_1 \left(\frac{x}{x_1}\right)^i + \cdots + a_{13} \left(\frac{x}{x_{13}}\right)^i\,.$$

Poiché B_i è il coefficiente di x^i in $f(x)$, Eulero ottenne che

$$B_i = \frac{a_1}{(x_1)^i} + \cdots + \frac{a_{13}}{(x_{13})^i} \approx \frac{a_k}{(x_k)^i}$$

quando $i \to +\infty$, dove x_k è la radice con il modulo più piccolo. In altre parole, B_i tende a crescere geometricamente come $(1/x_k)^i$. Restava da notare che x_k è una radice dell'equazione $1-x^{11}-x^{12}-x^{13}=0$ se e solo se $r=1/x_k$ è una radice dell'equazione $r^{13}=r^2+r+1$. Alcuni dettagli della dimostrazione furono finalmente chiariti da Gumbel nel 1916.

Süssmilch pubblicò una terza edizione del suo trattato nel 1765 e morì a Berlino nel 1767. In cattivi rapporti con il re di Prussia, Eulero tornò a San Pietroburgo nel 1766. Nonostante avesse perso la vista, continuò a pubblicare un gran numero di lavori con l'aiuto dei suoi figli e colleghi, specialmente sull'algebra, il calcolo integrale, l'ottica e le costruzioni navali. Le sue *Lettere su diversi argomenti di filosofia naturale indirizzate a una principessa tedesca*, scritte a Berlino tra il 1760 e il 1762, furono pubblicate tra il 1768 e il 1772 e diventarono un bestseller in tutta Europa. Eulero morì a San Pietroburgo nel 1783. Il suo contributo alla demografia matematica, in particolare la sua analisi della piramide di età stabile in una popolazione in crescita esponenziale, sarà riscoperto solo nel ventesimo secolo (si vedano i capitoli 10 e 21).

Bibliografia

1. Euler, L.: Recherches générales sur la mortalité et la multiplication du genre humain. *Hist. Acad. R. Sci. B.-Lett. Berl.* 16, 144–164 (1760/1767). eulerarchive.maa.org
2. Euler, L.: Sur la multiplication du genre humain. In: *Leonhardi Euleri Opera omnia*, Ser. I, vol. 7, 545–552. Teubner, Leipzig (1923)
3. Euler, L.: *Introductio in analysin infinitorum*, Tomus primus. Bousquet, Lausanne (1748). → Leonhardi Euleri *Opera omnia*, Ser. I, vol. 8, Teubner, Leizig (1922). gallica.bnf.fr
4. Fellmann, E.A.: *Leonhard Euler*. Birkhäuser, Basel (2007)
5. Gumbel, E.J.: Eine Darstellung statistischer Reihen durch Euler. *Jahresber. dtsch. Math. Ver.* 25, 251–264 (1917). digizeitschriften.de
6. Reimer, K.F.: Johann Peter Süssmilch, seine Abstammung und Biographie. *Arch. soz. Hyg. Demogr.* 7, 20–28 (1932)
7. Rohrbasser, J.M.: Johann Peter Süssmilch. In: Heyde, C.C., Seneta, E. (eds.) *Statisticians of the Centuries*, 72–76. Springer, New York (2001)
8. Süssmilch, J.P.: *Die göttliche Ordnung*. Berlin (1761). mpiwg-berlin.mpg.de
9. Warusfel, A.: *Euler, les mathématiques et la vie*. Vuibert, Paris (2009)

Capitolo 4

Daniel Bernoulli, d'Alembert e l'inoculazione del vaiolo (1760)

Nel 1760 Daniel Bernoulli scrisse un articolo sul modello del vaiolo. Ai suoi tempi c'era molta controversia intorno all'inoculazione, una pratica che poteva proteggere le persone ma poteva anche essere mortale. Usò la tavola di mortalità di Halley e alcuni dati riguardanti il vaiolo per dimostrare che l'inoculazione era vantaggiosa se il rischio associato di morire era inferiore all'11 %. L'inoculazione poteva aumentare la speranza di vita alla nascita fino a tre anni. D'Alembert criticò il lavoro di Bernoulli, che fu il primo modello matematico in epidemiologia.

Daniel Bernoulli è nato nel 1700 a Groningen nei Paesi Bassi. La sua famiglia comprendeva già due famosi matematici: suo padre Johann Bernoulli e suo zio Jakob Bernoulli. Nel 1705 Johann si trasferì a Basilea in Svizzera dove prese la cattedra lasciata libera dalla morte di Jakob. Johann non voleva che suo figlio studiasse matematica. Così Daniel si rivolse alla medicina, ottenendo il suo dottorato nel 1721 con una tesi sulla respirazione. Si trasferì a Venezia ed iniziò a concentrarsi sulla matematica, pubblicando un libro nel 1724. Avendo vinto un premio dall'Accademia delle Scienze di Parigi lo stesso anno per un saggio *Sul perfezionamento della clessidra su una nave in mare*, ottenne una cattedra alla nuova Accademia di San Pietroburgo. Durante questi anni, lavorò soprattutto sulle serie ricorrenti o sul «paradosso di San Pietroburgo» nella teoria della probabilità. Nel 1733 Daniel Bernoulli tornò all'Università di Basilea, dove insegnò successivamente botanica, fisiologia e fisica. Nel 1738 pubblicò un libro sulla dinamica dei fluidi che è rimasto famoso nella storia della fisica. Intorno al 1753 si interessò contemporaneamente a Eulero e d'Alembert al problema delle corde vibranti, che causò un'importante controversia matematica.

Nel 1760 presentò all'Accademia delle Scienze di Parigi un lavoro intitolato *Un tentativo di una nuova analisi della mortalità causata dal vaiolo e dei vantaggi dell'inoculazione per prevenirlo*. La questione era se l'inoculazione (l'introduzione volontaria di una piccola quantità di vaiolo meno virulento nel corpo per proteggerlo contro le infezioni successive) dovesse essere incoraggiata anche se a volte è un' operazione mortale. Questa tecnica era

Figura 4.1:
Daniel Bernoulli
(1700–1782)

conosciuta da molto tempo in Asia ed era stata introdotta nel 1718 in Inghilterra da Lady Montagu, moglie dell'ambasciatore britannico presso l'Impero Ottomano. In Francia, nonostante la morte del figlio maggiore di Luigi XIV per vaiolo nel 1711, l'inoculazione fu presa in considerazione con riluttanza. Voltaire, che era sopravvissuto al vaiolo nel 1723 e che aveva vissuto diversi anni in esilio in Inghilterra osservando le ultime innovazioni, perorò la causa dell'inoculazione nelle sue *Lettere filosofiche* nel 1734. Lo scienziato francese La Condamine, che era anche sopravvissuto al vaiolo, perorò la causa dell'inoculazione all'Accademia delle Scienze di Parigi nel 1754.

Prima di morire a Basilea nel 1759, Maupertuis incoraggiò Daniel Bernoulli a studiare il problema dell'inoculazione da un punto di vista matematico. Più precisamente, la sfida era quella di trovare un modo per confrontare il beneficio a lungo termine dell'inoculazione con il rischio immediato di morire. A questo scopo, Bernoulli fece le seguenti ipotesi semplificative:

- le persone infettate dal vaiolo per la prima volta muoiono con una probabilità p (indipendente dall'età) e sopravvivono con una probabilità $1-p$;

- ognuno ha una probabilità q di essere infettato ogni anno; più precisamente, la probabilità per un individuo di essere infettato tra l'età x e l'età $x+dx$ è $q\,dx$, dove dx è un periodo di tempo infinitesimo;

- Le persone che sopravvivono al vaiolo sono protette da nuove infezioni per il resto della loro vita (sono state immunizzate).

Sia $m(x)$ la mortalità all'età x per cause diverse dal vaiolo: la probabilità per un individuo di morire in un periodo di tempo infinitesimo dx tra l'età x e

l'età $x+dx$ è $m(x)\,dx$. Considerando un gruppo di P_0 persone nate lo stesso anno, chiamiamo

- $S(x)$ il numero di persone suscettibili di essere infettate, cioè che sono ancora in vita all'età x senza essere mai state infettate dal vaiolo [1];
- $R(x)$ il numero di persone che sono in vita all'età x e che sono sopravvissute al vaiolo;
- $P(x) = S(x) + R(x)$ il numero totale di persone in vita all'età x.

La nascita corrisponde all'età $x = 0$. Quindi $S(0) = P(0) = P_0$ e $R(0) = 0$. Applicando i metodi di calcolo che erano stati sviluppati alla fine del XVII secolo da Newton, Leibniz e più tardi da suo padre, Daniel Bernoulli notò che, tra l'età x e l'età $x+dx$ (con dx infinitamente piccolo), ogni individuo suscettibile di essere infettato ha una probabilità $q\,dx$ di essere infettato dal vaiolo e una probabilità $m(x)\,dx$ di morire per altre cause. Quindi la variazione del numero di persone suscettibili di essere infettate è $dS = -Sq\,dx - Sm(x)\,dx$, portando all'equazione differenziale

$$\frac{dS}{dx} = -qS - m(x)\,S\,. \tag{4.1}$$

In questa equazione, dS/dx è chiamata la derivata della funzione $S(x)$. Durante lo stesso piccolo intervallo di tempo, il numero di persone che muoiono di vaiolo è $pSq\,dx$ e il numero di persone che sopravvivono al vaiolo è $(1-p)\,Sq\,dx$. Inoltre, ci sono anche $Rm(x)\,dx$ persone che muoiono per cause diverse dal vaiolo. Questo porta ad una seconda equazione differenziale:

$$\frac{dR}{dx} = q(1-p)\,S - m(x)\,R\,. \tag{4.2}$$

Sommando le due equazioni, otteniamo

$$\frac{dP}{dx} = -pqS - m(x)\,P\,. \tag{4.3}$$

Dalle equazioni (4.1) e (4.3), Bernoulli poté mostrare che la frazione di persone che sono ancora suscettibili di essere infettate all'età x è

$$\frac{S(x)}{P(x)} = \frac{1}{(1-p)\,\mathrm{e}^{qx} + p}\,. \tag{4.4}$$

[1] Più esattamente, è il valore atteso di questo numero, che può variare continuamente e non solo di una o più unità.

Per ottenere la formula (4.4), Bernoulli eliminò $m(x)$ dalle equazioni (4.1) e (4.3):

$$-m(x) = q + \frac{1}{S}\frac{dS}{dx} = pq\frac{S}{P} + \frac{1}{P}\frac{dP}{dx}\,.$$

Segue dopo un riordinamento dei termini che

$$\frac{1}{P}\frac{dS}{dx} - \frac{S}{P^2}\frac{dP}{dx} = -q\frac{S}{P} + pq\left[\frac{S}{P}\right]^2.$$

Notiamo che la parte sinistra è la derivata di $f(x) = S(x)/P(x)$, che è la frazione di persone suscettibili di essere infettate nella popolazione di età x. Quindi

$$\frac{df}{dx} = -qf + pqf^2. \tag{4.5}$$

La soluzione di questo tipo di equazione era nota da diversi decenni grazie al lavoro di Jakob Bernoulli, lo zio di Daniel. Dividendo l'equazione per f^2 e ponendo $g(x) = 1/f(x)$, vediamo che $dg/dx = qg - pq$ e che $g(0) = 1/f(0) = 1$. Impostando $h(x) = g(x) - p$, otteniamo $dh/dx = qh$. Quindi $h(x) = h(0)\,\mathrm{e}^{qx} = (1-p)\,\mathrm{e}^{qx}$. Infine $g(x) = (1-p)\,\mathrm{e}^{qx} + p$ e $f(x) = 1/g(x)$.

Per applicare la sua teoria, Bernoulli usò la tavola di mortalità di Halley (vedi Capitolo 2). Questa tabella dà il numero di persone ancora in vita all'inizio dell'anno x (con $x = 1, 2...$) su una coorte di 1 238 persone nate nell'anno 0. Ma nel quadro del suo modello, Bernoulli aveva bisogno del numero di persone $P(x)$ che effettivamente raggiungono l'età x, che è leggermente diverso. Poiché Bernoulli - come la maggior parte dei suoi contemporanei - non si rese conto della differenza (l'articolo di Halley non è molto esplicito), mantenne i numeri della tabella di Halley tranne il primo numero 1 238, che sostituì con 1 300 per ottenere una mortalità realistica durante il primo anno di vita. Questi numeri sono mostrati nella seconda colonna della tabella 4.1.

Bernoulli scelse per la probabilità di morire di vaiolo $p = 1/8 = 12,5\,\%$, che è in accordo con le osservazioni del suo tempo. La probabilità annuale di prendere il vaiolo q non poteva essere stimata direttamente. Così Bernoulli probabilmente provò diversi valori per q e alla fine scelse quello tale che il numero di morti dovute al vaiolo dopo tutti i calcoli seguenti è circa 1/13 del numero totale di morti, una proporzione che era stata osservata allora in diverse città europee. La scelta di $q = 1/8$ all'anno si è rivelata un buon

Tabella 4.1: La tavola di mortalità di Halley e i calcoli di Bernoulli.

Età x	Vivi $P(x)$	Suscettibili di essere infettati $S(x)$	Immuni $R(x)$	Morti di vaiolo	Niente vaiolo $P^*(x)$
0	1 300	1 300	0	17,2	1 300
1	1 000	896	104	12,3	1 015
2	855	685	170	9,8	879
3	798	571	227	8,2	830
4	760	485	275	7,0	799
5	732	416	316	6,1	777
6	710	359	351	5,2	760
7	692	311	381	4,6	746
8	680	272	408	4,0	738
9	670	238	432	3,5	732
10	661	208	453	3,0	726
11	653	182	471	2,7	720
12	646	160	486	2,3	715
13	640	140	500	2,1	711
14	634	123	511	1,8	707
15	628	108	520	1,6	702
16	622	94	528	1,4	697
17	616	83	533	1,2	692
18	610	72	538	1,1	687
19	604	63	541	0,9	681
20	598	55	543	0,8	676
21	592	49	543	0,7	670
22	586	42	544	0,6	664
23	579	37	542	0,5	656
24	572	32	540		649
⋮	⋮	⋮	⋮	⋮	⋮

risultato, come vedremo ora[2].

Con la formula (4.4) e i valori di $P(x)$ nella seconda colonna della tabella, possiamo calcolare il numero $S(x)$ di persone suscettibili di essere infettate di età x: questa è la terza colonna della tabella arrotondata al numero intero più vicino. La quarta colonna mostra il numero $R(x) = P(x) - S(x)$ di persone di età x sopravvissute al vaiolo. La quinta colonna mostra nella riga corrispondente all'età x il numero di morti per vaiolo tra l'età x e l'età $x+1$. In teoria questo numero dovrebbe essere l'integrale $p\,q\int_x^{x+1} S(t)\,dt$ ma la formula $p\,q\,[S(x)+S(x+1)]/2$ dà una buona approssimazione, come indicato nella figura 4.2: l'area del trapezio è vicina all'area sotto la curva, cioè all'integrale della funzione.

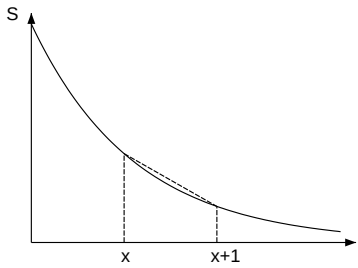

Figura 4.2: L'area del trapezio tratteggiato approssima l'integrale della funzione S tra x e $x+1$.

Bernoulli notò che la somma di tutti i numeri nella quinta colonna dà 98 morti per vaiolo prima dei 24 anni. Se continuassimo la tabella per età superiori, troveremmo solo tre morti in più per vaiolo tra le 32 persone che sono ancora suscettibili di essere infettate a 24 anni. In sintesi, partendo da 1 300 nascite, il destino di 101 persone è quello di morire di vaiolo. Questa è quasi esattamente la frazione attesa 1/13.

Bernoulli considerò allora la situazione in cui il vaiolo sarebbe stato inoculato a tutti alla nascita e non avrebbe causato alcuna morte. Il vaiolo sarebbe eradicato e si tratta di stimare l'aumento della speranza di vita. Partendo dallo stesso numero di nascite P_0, chiamiamo $P^*(x)$ il numero di persone di età x quando il vaiolo è scomparso. Allora

$$\frac{dP^*}{dx} = -m(x)\,P^*. \tag{4.6}$$

[2] Il fatto che p e q siano uguali è solo una coincidenza.

Bernoulli poteva dimostrare che

$$P^*(x) = \frac{P(x)}{1 - p + p\,\mathrm{e}^{-qx}}\,, \tag{4.7}$$

dove $P(x)$ è come sopra la popolazione di età x quando il vaiolo è presente.

Infatti, eliminando come prima $m(x)$ tra le equazioni (4.6) e (4.3), Bernoulli ottenne dopo il riordinamento

$$\frac{1}{P^*}\frac{dP}{dx} - \frac{P}{P^{*2}}\frac{dP^*}{dx} = -pq\frac{S}{P}\frac{P}{P^*}\,.$$

Ha fissato $h(x) = P(x)/P^*(x)$. Usando la formula (4.4), moltiplicò numeratore e denominatore per e^{-qx} e ottenne

$$\frac{1}{h}\frac{dh}{dx} = -pq\frac{\mathrm{e}^{-qx}}{1 - p + p\,\mathrm{e}^{-qx}}\,,$$

che è equivalente a

$$\frac{d}{dx}\log h = \frac{d}{dx}\log(1 - p + p\,\mathrm{e}^{-qx}),$$

dove log sta qui per il logaritmo naturale e non il logaritmo decimale. Ma $h(0) = 1$. Quindi $h(x) = 1 - p + p\,\mathrm{e}^{-qx}$.

Notate che il rapporto $P(x)/P^*(x)$ tende a $1 - p$ quando l'età x è abbastanza alta. La sesta colonna della Tabella 4.1 mostra $P^*(x)$. Un modo per confrontare $P(x)$ e $P^*(x)$ è stimare la speranza di vita alla nascita, la cui espressione teorica con il vaiolo è

$$\frac{1}{P_0}\int_0^{+\infty} P(x)\,dx\,.$$

Un'espressione simile con $P^*(x)$ che sostituisce $P(x)$ vale senza il vaiolo. Bernoulli usò la formula approssimativa $[\frac{1}{2}P(0) + P(1) + P(2) + \cdots]/P_0$, che è quella data dal metodo dei trapezi (figura 4.2). Continuando la tabella oltre i 24 anni fino a 84 anni (vedi tabella 2.1), ottenne infine una speranza di vita E con il vaiolo pari a $[\frac{1}{2}1300 + 1000 + \cdots + 20]/1300 \approx 26{,}57$ anni, cioè 26 anni e 7 mesi. Senza vaiolo, ha ottenuto un'aspettativa di vita E^* pari a $[\frac{1}{2}1300 + 1015 + \cdots + 23]/1300 \approx 29{,}65$ anni, cioè 29 anni e 8 mesi. L'inoculazione alla nascita aumenterebbe l'aspettativa di vita di più di tre anni.

Possiamo notare che esiste un metodo più semplice e veloce di quello usato da Bernoulli per ottenere queste formule. Partendo dall'equazione differenziale (4.1) per $S(x)$, vediamo prima che

$$S(x) = P_0\,\mathrm{e}^{-qx}\exp\left(-\int_0^x m(y)\,dy\right) .$$

Usando questa espressione nell'equazione (4.2) per $R(x)$, troviamo che

$$R(x) = P_0\,(1-p)\,(1-\mathrm{e}^{-qx})\exp\left(-\int_0^x m(y)\,dy\right) .$$

L'equazione (4.6) per $P^*(x)$ mostra che

$$P^*(x) = P_0\exp\left(-\int_0^x m(y)\,dy\right) . \tag{4.8}$$

Le formule (4.4) e (4.7) seguono immediatamente!

Naturalmente, l'inoculazione con un ceppo meno virulento di vaiolo non è completamente sicura. Se p' è la probabilità di morire di vaiolo subito dopo l'inoculazione ($p' < p$), allora la speranza di vita sarebbe $(1-p')\,E^*$ se tutti venissero inoculati alla nascita. Questa speranza di vita rimane più alta della speranza di vita «naturale» E se $p' < 1 - E/E^*$ o circa 11 %. I dati riguardanti p' erano difficili da ottenere all'epoca. Ma Bernoulli stimò che il rischio p' era inferiore all'1 %. Per lui non c'erano dubbi: l'inoculazione doveva essere promossa dallo Stato. Concludeva:

> «Desidero semplicemente che, in una questione che riguarda così da vicino il benessere della razza umana, nessuna decisione sia presa senza tutta la conoscenza che un po' di analisi e di calcolo possono fornire.»

Il lavoro di Bernoulli fu presentato all'Accademia delle Scienze di Parigi nell'aprile 1760. In novembre, d'Alembert presentò un commento intitolato *Sull'applicazione della teoria delle probabilità all'inoculazione del vaiolo*. Il commento fu pubblicato poco dopo nel secondo volume dei suoi *Opuscules mathématiques* con calcoli più dettagliati e insieme ad un altro lavoro intitolato *Teoria matematica dell'inoculazione*. D'Alembert criticò le ipotesi di Bernoulli sulla probabilità di infezione e la probabilità di morire di vaiolo essendo indipendenti dall'età. Suggerì una soluzione diversa che non richiede queste ipotesi. Chiama $v(x)$ la mortalità dovuta al vaiolo all'età x, $m(x)$

la mortalità dovuta ad altre cause e $P(x)$ il numero di persone ancora vive. Allora

$$\frac{dP}{dx} = -v(x)\,P - m(x)\,P\,. \tag{4.9}$$

Confrontando con l'equazione (4.3), vediamo che in effetti $v(x) = pqS(x)/P(x)$. Qui otteniamo

$$P^*(x) = P(x)\,\exp\Big(\int_0^x v(y)\,dy\Big)\,, \tag{4.10}$$

dove $P^*(x)$ rappresenta il numero di persone in vita all'età x quando il vaiolo è scomparso.

Figura 4.3:
D'Alembert (1717–1783)

Infatti possiamo sostituire la funzione $m(x)$ tra le equazioni (4.6) e (4.9) o usare la formula (4.8) per $P^*(x)$ e notare che la soluzione di (4.9) è data da

$$P(x) = P_0\,\exp\left(-\int_0^x [v(y)+m(y)]\,dy\right)\,.$$

La formula (4.10) data da d'Alembert non contraddice la formula di Bernoulli (4.7). Usa solo un diverso tipo di informazione $v(x)$, che non era disponibile all'epoca perché i registri delle morti includevano la causa della morte ma non l'età della vittima. D'Alembert suggerì che non si poteva realmente concludere se l'inoculazione fosse utile prima che questo tipo di dati fossero disponibili.

D'Alembert ha anche criticato l'utilità della speranza di vita come criterio di decisione, poiché dà lo stesso peso a tutti gli anni, sia in un futuro prossimo che lontano. Notò che, dal punto di vista dell'individuo o dello Stato, non tutti gli anni hanno la stessa «utilità», le età giovani e vecchie hanno meno valore delle età medie. Nonostante tutte queste critiche, d'Alembert si dichiarò a favore dell'inoculazione.

A causa dei ritardi nella pubblicazione, il lavoro di Bernoulli fu pubblicato solo nel 1766, mentre d'Alembert riuscì a far pubblicare il proprio lavoro molto rapidamente. Bernoulli espresse la sua amarezza in una lettera a Eulero:

> «Che cosa dite delle enormi banalità del grande d'Alembert sulle probabilità: poiché mi trovo troppo spesso trattato ingiustamente nelle sue pubblicazioni, ho deciso già da tempo di non leggere più nulla che esca dalla sua penna. Ho preso questa decisione in occasione di un manoscritto sull'inoculazione che ho inviato all'Accademia di Parigi otto anni fa e che è stato molto apprezzato per la novità dell'analisi. Era, oserei dire, come incorporare una nuova provincia nel corpo della matematica. Sembra che il successo di questa nuova analisi gli abbia causato dolori di cuore. L'ha criticata in mille modi tutti ugualmente ridicoli e dopo averla ben criticata, pretende di essere il primo autore di una teoria che non ha neanche sentito nominare. Egli, tuttavia, sapeva che il mio manoscritto poteva apparire solo dopo circa sette o otto anni. Poteva esserne a conoscenza solo in qualità di membro dell'Accademia. Da questo punto di vista il mio manoscritto sarebbe dovuto rimanere sacro fino a quando non fosse stato reso pubblico. *Dolus an virtus quis in hoste requirat!*»[3]

Nonostante i lavori di Bernoulli e d'Alembert, l'inoculazione non fu eseguita su larga scala in Francia. Il re Luigi XV morì di vaiolo nel 1774. I medici di corte inocularono il resto della famiglia reale poco dopo. Il problema perse la sua importanza quando Edward Jenner scoprì che inoculare il vaiolo bovino agli esseri umani («vaccinazione») proteggeva dal vaiolo ed era sicuro. Il suo lavoro, *Un'indagine sulle cause e gli effetti della vaccina variolae*, fu pubblicato nel 1798. La vaccinazione si diffuse rapidamente in tutta Europa. Tuttavia i metodi sviluppati per il calcolo dell'aumento della speranza di vita se viene rimossa una causa di morte sono ancora in uso oggi.

Nei decenni successivi, i dati relativi all'età in cui si moriva di vaiolo divennero disponibili. Il problema fu riconsiderato soprattutto da

- Johann Heinrich Lambert, un matematico dell'Accademia di Berlino, nel 1772;
- Emmanuel-Étienne Duvillard, allora responsabile delle statistiche della popolazione presso il Ministero degli Interni a Parigi, nella sua *Analisi e tabelle dell'influenza del vaiolo sulla mortalità ad ogni età* (1806);

[3] *Astuzia o coraggio, che importanza ha contro il nemico?* Virgilio: Eneide, Libro II.

- Pierre-Simon Laplace nella sua *Teoria analitica della probabilità* (1812).

Duvillard e Laplace hanno mostrato per esempio come modificare la formula (4.7) quando i parametri p e q dipendono dall'età:

$$P^*(x) = \frac{P(x)}{1 - \int_0^x p(y)\, q(y)\, \mathrm{e}^{-\int_0^y q(z)\, dz}\, dy}\,.$$

Qui, $p(x)$ è la probabilità di morire di vaiolo se infettati all'età x e $q(x)$ è la probabilità di essere infettati dal vaiolo all'età x.

Dopo questo lavoro sul vaiolo, Daniel Bernoulli non considerò nessun altro problema nella dinamica della popolazione. Morì a Basilea nel 1782. D'Alembert morì a Parigi un anno dopo.

Bibliografia

1. D'Alembert, J.: Onzième mémoire, Sur l'application du calcul des probabilités à l'inoculation de la petite vérole. In: *Opuscules mathématiques*, Tome second, 26–95. David, Paris (1761). gallica.bnf.fr
2. Bernoulli, D.: Réflexions sur les avantages de l'inoculation. *Mercure de France*, 173–190 (juin 1760). retronews.fr
3. Bernoulli, D.: Essai d'une nouvelle analyse de la mortalité causée par la petite vérole et des avantages de l'inoculation pour la prévenir. *Hist. Acad. R. Sci. Paris*, 1–45 (1760/1766). gallica.bnf.fr
4. Dietz, K., Heesterbeek, J.A.P.: Daniel Bernoulli's epidemiological model revisited. *Math. Biosci.* 180, 1–21 (2002)
5. Duvillard, E.E.: *Analyse et tableaux de l'influence de la petite vérole sur la mortalité à chaque âge*. Imprimerie impériale, Paris (1806). archive.org
6. Lambert, J.H.: *Contributions mathématiques à l'étude de la mortalité et de la nuptialité* (1765 et 1772). INED, Paris (2006).
7. Laplace, P.S.: *Théorie analytique des probabilités*. Courcier, Paris (1812). gallica.bnf.fr
8. Straub, H.: Bernoulli, Daniel. In Gillespie, C.C. (ed.) *Dictionary of Scientific Biography*, vol. 2, 36–46. Scribner, New York (1970)
9. Tent, M.B.W.: *Leonhard Euler and the Bernoullis*. A K Peters (2009)
10. Voltaire: *Lettres philosophiques*. Lucas, Amsterdam (1734). gallica.bnf.fr

Capitolo 5

Malthus e gli ostacoli alla crescita geometrica (1798)

Nel 1798 Malthus pubblicò una monografia intitolata *Saggio sul principio della popolazione*, in cui sosteneva che la quantità di cibo non poteva seguire, per un lungo periodo di tempo, la naturale tendenza della popolazione umana a crescere esponenzialmente. Se la popolazione rimaneva pressochè costante, ciò era dovuto al fatto che gran parte dell'umanità soffriva di una carenza di cibo. Malthus considerò il «principio della popolazione» come un argomento contro gli scritti di Godwin e Condorcet, che sottolineavano il progresso delle società umane. Il saggio di Malthus influenzò la teoria dell'evoluzione di Darwin e Wallace e fu criticato da Marx, ma fu messo in pratica dal governo cinese con la politica del figlio unico.

Thomas Robert Malthus nacque nel 1766 vicino a Londra, sesto di sette figli. Suo padre, amico e ammiratore di Jean-Jacques Rousseau, fu il suo primo insegnante. Nel 1784 il giovane Malthus iniziò a studiare matematica all'Università di Cambridge. Conseguì il diploma di Laurea nel 1791, divenne membro accademico del *Jesus College* nel 1793 e sacerdote anglicano nel 1797.

Figura 5.1:
Malthus (1766–1834)

Nel 1798 Malthus pubblicò, in forma anonima, un libro intitolato *Saggio sul principio della popolazione, con riferimento al futuro miglioramento*

della società, con osservazioni sulle speculazioni di Godwin, Condorcet e altri scrittori. Il libro fu scritto per ribattere alle argomentazioni riportate a seguito dell'inchiesta di Godwin sulla giustizia politica (1793) e Condorcet *Schizzo per un quadro storico del progresso della mente umana*. (1794). Nonostante gli orrori che la Rivoluzione Francese fece in nome del progresso, i due autori sostenevano che il progresso della società era inevitabile. Malthus non condivideva lo stesso ottimismo. Sosteneva anche che le leggi inglesi sui poveri, che avrebbero dovuto aiutare le famiglie povere con molti figli, favorivano la crescita della popolazione senza incoraggiare una crescita simile nella produzione di cibo. Gli sembrava che queste leggi non alleviassero le sofferenze dei poveri; al contrario. Più in generale, con la popolazione che tende a crescere sempre più velocemente della produzione di cibo, una parte della società sembrava essere condannata alla miseria, alla fame o alle epidemie: questi sono i flagelli che rallentano la crescita della popolazione e che, secondo Malthus, sono i principali ostacoli al progresso della società. Tutte le teorie che promettono il progresso sarebbero solo utopistiche. Queste idee portarono Malthus a pubblicare il suo libro nel 1798. Ecco come riassumeva la sua tesi:

> [...] «il potere di crescita della popolazione è indefinitamente più grande del potere della terra di produrre sussistenza per l'uomo. La popolazione, quando non è controllata, aumenta in un rapporto geometrico. La sussistenza aumenta solo in un rapporto aritmetico. Una conoscenza elementare dei numeri mostrerà la maggiore grandezza della prima potenza in confronto alla seconda. Per la legge della natura che rende il cibo necessario alla vita, gli effetti di queste due potenze ineguali devono essere mantenuti uguali. Ciò implica un controllo forte e costantemente operante sulla popolazione a causa della difficoltà di produrre la sussistenza. Questa difficoltà deve manifestarsi da qualche parte; e deve, necessariamente, essere fortemente percepita da una grande parte del genere umano.»

Il libro di Malthus ebbe molto successo. Conteneva pochi dati. Malthus notò, per esempio, che la popolazione degli Stati Uniti era raddoppiata ogni venticinque anni durante il XVIII secolo. Non cercò veramente di tradurre le sue tesi in modelli matematici, ma aprì la strada al lavoro successivo di Adolphe Quetelet e Pierre-François Verhulst, che sarà l'argomento del prossimo capitolo.

Dopo la pubblicazione del suo libro, Malthus viaggiò con amici prima in Germania, Scandinavia e Russia, poi in Francia e Svizzera. Mettendo in-

sieme le informazioni raccolte durante i suoi viaggi, pubblicò, sotto il suo nome, una seconda edizione molto ampliata nel 1803, con un sottotitolo diverso: *Un saggio sul principio della popolazione, o una visione dei suoi effetti passati e presenti sulla felicità umana, con un'indagine sulle nostre prospettive riguardo alla futura rimozione o attenuazione dei mali che essa provoca*. Questa nuova edizione discuteva in dettaglio gli ostacoli alla crescita della popolazione in vari paesi: matrimonio ritardato, aborto, infanticidio, carestia, guerra, epidemie, fattori economici. Per Malthus, il matrimonio ritardato era la migliore opzione per stabilizzare la popolazione. Nel 1806, 1807, 1817 e 1826 seguirono altre quattro edizioni del libro. Nel 1805 Malthus divenne professore di storia ed economia politica in una nuova scuola istituita dalla Compagnia delle Indie Occidentali per i suoi dipendenti. Pubblicò anche *Un'indagine sulla natura e l'andamento della rendita* (1815) e *Principi di economia politica* (1820). Nel 1819 Malthus fu eletto alla *Royal Society*. Society. Nel 1834 fu uno dei membri fondatori della *Statistical Society*. Morì vicino a Bath lo stesso anno.

Il lavoro di Malthus ebbe una forte influenza sullo sviluppo della teoria dell'evoluzione. Charles Darwin, di ritorno dal suo viaggio a bordo del Beagle, lesse il libro di Malthus sulla popolazione nel 1838. Ecco cosa scrisse nell'introduzione al suo famoso libro *L'origine delle specie per mezzo della selezione naturale*, pubblicato nel 1859:

> «Il capitolo successivo tratterà della lotta per l'esistenza fra tutti gli esseri organizzati del globo, lotta che necessariamente deriva dal loro moltiplicarsi in proporzione geometrica. È questa la legge di Malthus applicata a tutto il regno animale e vegetale.»

Alfred Russel Wallace, che sviluppò la teoria dell'evoluzione nello stesso periodo di Darwin, disse anche che le sue idee vennero dopo aver letto il libro di Malthus.

Al contrario, ecco il punto di vista di Karl Marx sul successo del libro di Malthus, come si può leggere in una nota a piè di pagina del suo *Capitale*:

> «Se il lettore dovesse ricordarmi il Malthus, il cui Essay on population uscì nel 1798, io gli ricorderò che questo scritto nella sua prima forma non è che un plagio superficiale da scolaretto, declamatorio in maniera pretesca, di scritti di De Foe, Sir james Steuart, Townsend, Franklin, Wallace ecc., e non contiene nemmeno una proposizione originale. Il grande scalpore destato da quest'opuscolo fu dovuto unicamente a interessi di partito. La rivoluzione francese aveva trovato nel regno britannico degli appassionati difensori; il "principio della popolazione", elaborato

> lentamente nel secolo XVIII, annunciato poi a suon di tromba nel bel mezzo di una grande crisi sociale come antidoto infallibile contro le dottrine del Condorcet e di altri, fu salutato entusiasticamente dall'oligarchia inglese come il grande sterminatore di tutte le voglie di progresso umano. Il Malthus, altamente stupito del proprio successo, si mise poi a riempire il vecchio schema di materiale compilato superficialmente e di materiale nuovo, che si era semplicemente annesso senza averlo scoperto.»

Certamente le tesi di Malthus non erano completamente nuove. Per esempio, l'idea che la popolazione tende a crescere geometricamente è spesso attribuita a lui, anche se abbiamo visto nel capitolo 3 che questa idea era già familiare a Eulero mezzo secolo prima[1]. Tuttavia, Malthus gli diede pubblicità collegandola in modo polemico a problemi legislativi reali. Ironicamente, fu nella Cina comunista che il suggerimento di Malthus di limitare le nascite avrebbe trovato la sua applicazione più eclatante (vedi capitolo 25).

Bibliografia

1. Condorcet: *Esquisse d'un tableau historique des progrès de l'esprit humain*. Agasse, Paris (1794). gallica.bnf.fr
2. Darwin, C. (trad. Canestrini, G.): *Sull'origine delle specie per la selezione naturale, ovvero conservazione delle razze perfezionate nella lotta per l'esistenza*. A. Barion Editore, Sesto San Giovanni (1933). liberliber.it
3. Godwin, W.: *An Enquiry Concerning Political Justice*. Robinson, London (1793). archive.org
4. Malthus, T.R.: *An Essay on the Principle of Population*, 1st edn. London (1798). econlib.org
5. Marx, K. (trad. Cantimori, D.): *Il Capitale. Critica dell'Economia Politica*. Editori Riuniti, Roma (1964). rottacomunista.org
6. Simpkins, D.M.: Malthus, Thomas Robert. In: Gillespie, C.C. (ed.) *Dictionary of Scientific Biography*, vol. 9, 67–71. Scribner, New York (1974)

[1] R. A. Fisher (vedi i capitoli 14 e 20) chiamerebbe «parametro malthusiano» il tasso di crescita delle popolazioni. Malthus ha citato il trattato di Süssmilch nel suo stesso libro.

Capitolo 6

Verhulst e l'equazione logistica (1838)

Nel 1838 il matematico belga Verhulst introdusse l'equazione logistica, che è una sorta di generalizzazione dell'equazione della crescita esponenziale ma con un valore massimo per la popolazione. Usò i dati di diversi paesi, in particolare del Belgio, per stimare i parametri sconosciuti. Il lavoro di Verhulst fu riscoperto solo negli anni 1920.

Pierre-François Verhulst nacque nel 1804 a Bruxelles. Ottenne un dottorato in matematica all'Università di Gand nel 1825. Si interessò anche di politica. Mentre era in Italia per curare la sua tubercolosi, chiese senza successo in favore di una costituzione per lo Stato Pontificio. Dopo la rivoluzione del 1830 e l'indipendenza del Belgio, pubblicò un saggio storico su un patriota del XVIII secolo. Nel 1835 divenne professore di matematica alla nuova Università Libera di Bruxelles.

Figura 6.1:
Verhulst (1804–1849)

Lo stesso anno 1835, il suo compatriota Adolphe Quetelet, statistico e direttore dell'osservatorio di Bruxelles, pubblicò un trattato *Sull'uomo e lo sviluppo delle sue facoltà*. Quetelet suggerì che le popolazioni non potevano crescere geometricamente per un lungo periodo di tempo perché gli ostacoli menzionati da Malthus formavano una sorta di «resistenza», che egli pensava

(per analogia con la meccanica) fosse proporzionale al quadrato della velocità di crescita della popolazione. Questa analogia non aveva basi reali, ma ispirò Verhulst.

Infatti, Verhulst pubblicò nel 1838 una *Nota sulla legge della crescita della popolazione*. Ecco alcuni estratti:

> «Sappiamo che il famoso Malthus ha dimostrato il principio che la popolazione umana tende a crescere in una progressione geometrica in modo da raddoppiare dopo un certo periodo di tempo, per esempio ogni venticinque anni. Questa proposizione è fuori discussione se si fa astrazione della crescente difficoltà di trovare cibo [...]
>
> L'aumento virtuale della popolazione è dunque limitato dalle dimensioni e dalla fertilità del paese. Di conseguenza, la popolazione si avvicina sempre più a uno stato stazionario.»

Verhulst probabilmente si rese conto che l'analogia meccanica di Quetelet non era ragionevole e propose invece la seguente equazione differenziale (ancora un po' arbitraria) per la popolazione $P(t)$ al tempo t:

$$\frac{dP}{dt} = rP\left(1 - \frac{P}{K}\right). \tag{6.1}$$

Quando la popolazione $P(t)$ è piccola rispetto al parametro K, si ottiene l'equazione approssimativa

$$\frac{dP}{dt} \approx rP\,,$$

la cui soluzione è $P(t) \approx P(0)\,\mathrm{e}^{rt}$, cioè crescita esponenziale[1]. Il tasso di crescita diminuisce quando $P(t)$ si avvicina a K. Diventerebbe addirittura negativo se $P(t)$ potesse superare K. Per ottenere l'espressione esatta della soluzione dell'equazione (6.1), possiamo procedere come Daniel Bernoulli per l'equazione (4.5).

Dividendo l'equazione (6.1) per P^2 e definendo $p = 1/P$, otteniamo $dp/dt = -rp + r/K$. Con $q = p - 1/K$, otteniamo $dq/dt = -rq$ e

$$q(t) = q(0)\,\mathrm{e}^{-rt} = (1/P(0) - 1/K)\,\mathrm{e}^{-rt}.$$

[1] Di solito si parla di crescita geometrica nei modelli a tempo discreto e di crescita esponenziale nei modelli a tempo continuo, ma è essenzialmente la stessa cosa.

Quindi possiamo dedurre $p(t)$ e $P(t)$.

Riordinando, otteniamo infine

$$P(t) = \frac{P(0)\,\mathrm{e}^{rt}}{1 + P(0)\,(\mathrm{e}^{rt} - 1)/K}\,. \tag{6.2}$$

La popolazione totale aumenta progressivamente da $P(0)$ al tempo $t = 0$ fino al limite K, che viene raggiunto solo quando $t \to +\infty$ (figura 6.2).

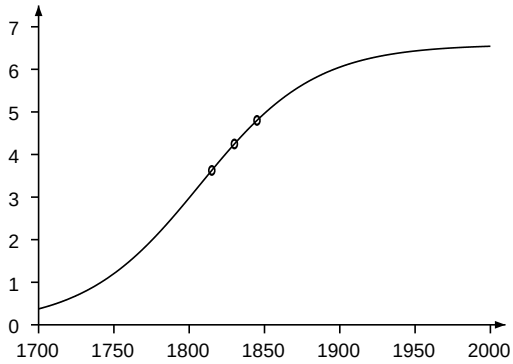

Figura 6.2: La popolazione del Belgio (in milioni) e la curva logistica. I punti dei dati corrispondono agli anni 1815, 1830 e 1845. I valori dei parametri sono quelli dell'articolo del 1845.

Senza dare i valori che ha usato per i parametri sconosciuti r e K, Verhulst ha confrontato il suo risultato con i dati riguardanti la popolazione della Francia tra il 1817 e il 1831, del Belgio tra il 1815 e il 1833, della contea di Essex in Inghilterra tra il 1811 e il 1831, e della Russia tra il 1796 e il 1827. Il risultato è stato abbastanza buono.

Nel 1840 Verhulst divenne professore alla Scuola Reale Militare di Bruxelles. L'anno seguente pubblicò un *Trattato elementare di funzioni ellittiche* e fu eletto all'Accademia Reale del Belgio. Nel 1845 continuò i suoi studi sulla popolazione con un articolo intitolato *Inchieste matematiche sulla legge della crescita della popolazione*. Per prima cosa si rifece all'osservazione di Malthus secondo cui la popolazione degli Stati Uniti era raddoppiata ogni 25 anni (tabella 6.1). Se calcoliamo il rapporto tra la popolazione dell'an-

Tabella 6.1: Censimenti ufficiali della popolazione degli USA.

anno	popolazione
1790	3 929 827
1800	5 305 925
1810	7 239 814
1820	9 638 131
1830	12 866 020
1840	17 062 566

no $n+10$ e quella dell'anno n, troviamo rispettivamente 1,350, 1,364, 1,331, 1,335 e 1,326, che è abbastanza costante. La popolazione era quindi moltiplicata in media per 1,34 ogni 10 anni e per $1{,}34^{25/10} \approx 2{,}08$ ogni 25 anni. Così aveva continuato a raddoppiare ogni 25 anni dal saggio di Malthus, quasi mezzo secolo prima. Tuttavia Verhulst aggiunse:

> «Non insisteremo sull'ipotesi della progressione geometrica, dato che può reggere solo in circostanze molto speciali; per esempio, quando un territorio fertile di dimensioni quasi illimitate capita di essere abitato da persone con una civiltà avanzata, come fu il caso delle prime colonie americane.»

Nel suo articolo Verhulst tornò anche sull'equazione (6.1), che chiamò «logistica». Notò che la curva $P(t)$ aumenta con una curvatura positiva (è convessa) finché $P(t) < K/2$ e poi continua ad aumentare verso K ma con una curvatura negativa (è concava) appena $P(t) > K/2$. Quindi la curva ha la forma di una lettera S distorta (figura 6.2).

Infatti,

$$\frac{d^2P}{dt^2} = r\,(1-2P/K)\,\frac{dP}{dt}.$$

Quindi $d^2P/dt^2 > 0$ se $P < K/2$ e $d^2P/dt^2 < 0$ se $P > K/2$.

Verhulst ha anche spiegato come i parametri r e K possono essere stimati dalla popolazione $P(t)$ in tre anni diversi ma equamente distanziati. Se P_0 è la popolazione al tempo $t=0$, P_1 quella al tempo $t=T$ e P_2 quella al tempo $t=2T$, allora partendo dall'equazione (6.2) è possibile ottenere i valori dei parametri

$$K = P_1\,\frac{P_0P_1 + P_1P_2 - 2P_0P_2}{P_1^2 - P_0P_2}\,, \quad r = \frac{1}{T}\,\log\left[\frac{1/P_0 - 1/K}{1/P_1 - 1/K}\right]\,.$$

Usando le stime per la popolazione del Belgio negli anni 1815, 1830 e 1845 (rispettivamente 3,627, 4,247 e 4,801 milioni), ottenne $K = 6{,}584$ milioni e $r = 2{,}62\,\%$ per anno. Poteva quindi utilizzare l'equazione (6.2) per prevedere che la popolazione del Belgio sarebbe stata di 4,998 milioni all'inizio dell'anno 1851 e di 6,064 milioni all'inizio dell'anno 1900 (figura 6.2). Verhulst fece uno studio simile per la Francia. Ha ottenuto $K = 39{,}685$ milioni e $r = 3{,}2\,\%$ all'anno. Poiché le popolazioni del Belgio e della Francia hanno nel frattempo ampiamente superato questi valori di K, vediamo che l'equazione logistica può essere un modello realistico solo per periodi di tempo di pochi decenni, come nell'articolo di Verhulst del 1838, ma non per periodi più lunghi.

Nel 1847 apparve una *Seconda indagine sulla legge della crescita della popolazione* in cui Verhulst rinunciò all'equazione logistica e scelse invece un'equazione differenziale che può essere scritta nella forma

$$\frac{dP}{dt} = r\left(1 - \frac{P}{K}\right) .$$

Credeva che questa equazione sarebbe stata soddisfatta quando la popolazione $P(t)$ avesse superato una certa soglia. La soluzione è

$$P(t) = K + (P(0) - K)\,\mathrm{e}^{-rt/K} .$$

Usando gli stessi dati demografici per il Belgio, Verhulst stimò nuovamente i parametri r e K. Questa volta trovò $K = 9{,}4$ milioni per la popolazione massima. Vediamo pertanto quanto il risultato può dipendere dalla scelta del modello!

Verhulst divenne presidente dell'Accademia Reale del Belgio nel 1848, ma morì l'anno successivo a Bruxelles, probabilmente di tubercolosi. Nonostante l'esitazione di Verhulst nello scegliere le equazioni per il modello, l'equazione logistica fu reintrodotta indipendentemente diversi decenni dopo da diverse persone. Robertson la usò nel 1908 per modellizzare la crescita individuale di animali, piante, esseri umani e organi del corpo. McKendrick e Kesava Pai la usarono nel 1911 per la crescita di popolazioni di microrganismi. Pearl e Reed lo usarono nel 1920 per la crescita della popolazione degli Stati Uniti, che aveva iniziato a rallentare. Nel 1922 Pearl notò finalmente il lavoro di Verhulst. Da allora in poi, l'equazione logistica ispirò molti lavori (vedi i capitoli 13, 20 e 24). La popolazione massima K alla fine divenne nota come «capacità portante».

Bibliografia

1. Lloyd, P.J.: American, German and British antecedents to Pearl and Reed's logistic curve. *Pop. Stud.* 21, 99–108 (1967)

2. McKendrick, A.G., Kesava Pai, M.: The rate of multiplication of micro-organisms: A mathematical study. *Proc. R. Soc. Edinb.* 31, 649–655 (1911)

3. Pearl, R.: *The Biology of Death*. Lippincott, Philadelphia (1922). archive.org

4. Pearl, R., Reed, L.J.: On the rate of growth of the population of the United States since 1790 and its mathematical representation. *Proc. Natl. Acad. Sci.* 6, 275–288 (1920). pnas.org

5. Quetelet, A.: *Sur l'homme et le développement de ses facultés*. Bachelier, Paris (1835). gallica.bnf.fr

6. Quetelet, A.: Pierre-François Verhulst. *Annu. Acad. R. Sci. Lett. B.-Arts Belg.* 16, 97–124 (1850). archive.org

7. Quetelet, A.: *Sciences mathématiques et physiques au commencement du* XIX*e siècle*. Mucquardt, Bruxelles (1867). gallica.bnf.fr

8. Robertson, T.B.: On the normal rate of growth of an individual and its biochemical significance. *Arch. Entwicklungsmechanik Org.* 25, 581–614 (1908)

9. Verhulst, P.-F.: Notice sur la loi que la population poursuit dans son accroissement. *Corresp. Math. Phys.* 10, 113–121 (1838). archive.org

10. Verhulst, P.-F.: Recherches mathématiques sur la loi d'accroissement de la population. *Nouv. Mém. Acad. R. Sci. B.-lett. Brux.* 18, 1–45 (1845). gdz.sub.uni-goettingen.de

11. Verhulst, P.-F.: Deuxième mémoire sur la loi d'accroissement de la population. *Mém. Acad. R. Sci. Lett. B.-Arts Belg.* 20 (1847). archive.org

Capitolo 7

Bienaymé, Cournot e l'estinzione dei nomi di famiglia (1845–1847)

Nel 1845 lo statistico francese Bienaymé comprese come calcolare la probabilità che un nome di famiglia si estingua se ogni maschio ha un numero di figli che segue una data distribuzione di probabilità. Se il numero medio di figli è inferiore o uguale a uno, il nome della famiglia si estinguerà. Se la media è maggiore di uno, la probabilità di estinzione è strettamente inferiore a uno. La dimostrazione fu pubblicata due anni dopo in un libro scritto dal suo amico Cournot. Questi lavori sono stati riscoperti solo recentemente.

Irenée Jules Bienaymé nacque nel 1796 a Parigi. Studiò alla Scuola Politecnica e fece carriera al Ministero delle Finanze, raggiungendo la massima carica di ispettore generale. Influenzato dal libro *Teoria analitica della probabilità* scritto da Laplace, Bienaymé trovò anche il tempo di pubblicare articoli su molte applicazioni della teoria della probabilità, come le statistiche demografiche e mediche (mortalità infantile, numero di nascite, aspettativa di vita), probabilità di errori nella giustizia, teoria delle assicurazioni e rappresentatività dei sistemi di voto.

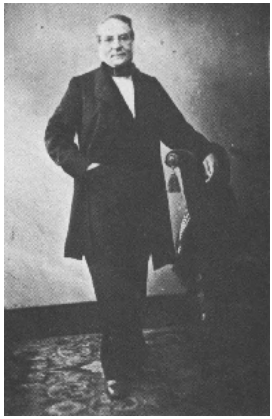

Figura 7.1:
Bienaymé (1796–1878)

Nel 1845 Bienaymé scrisse una breve nota *Sulla legge della moltiplicazione e la durata dei nomi delle famiglie*, che fu pubblicata nel bollettino della *Société Philomatique* di Parigi. Un certo numero di autori aveva già scritto su questo argomento. Nella seconda edizione di *Saggio sul principio di popolazione* (1803), Malthus incluse un capitolo sulla popolazione della Svizzera e notò che

> «nella città di Berna, dall'anno 1583 al 1654, il consiglio sovrano aveva ammesso nella borghesia 487 famiglie, di cui 379 si estinsero nello spazio di due secoli, e nel 1783 ne rimanevano solo 108.»

Nel 1842 Thomas Doubleday sostenne più in generale che le famiglie di classe superiore della nobiltà o della borghesia avevano una maggiore tendenza a scomparire rispetto alle famiglie di classe inferiore. Idee simili furono avanzate in Francia da Émile Littré nel 1844 in un testo di introduzione alla filosofia positivista di Auguste Comte e da Benoiston de Châteauneuf – un amico di Bienaymé – che pubblicò nel 1845 un saggio *Sulla durata delle famiglie nobili in Francia*.

Fu in questo contesto che Bienaymé cercò di spiegare come poteva essere che la popolazione di un paese tendesse a crescere geometricamente mentre un gran numero di famiglie scompariva. Per affrontare questo problema considerò il caso semplificato in cui tutti gli uomini avessero le stesse probabilità di avere 0, 1, 2, 3,... figli che raggiungono l'età adulta. Più precisamente, si chiese quale fosse la probabilità per un uomo di avere figli che portassero il suo nome dopo n generazioni. Se il numero medio di figli è inferiore a uno, è chiaro che questa probabilità dovrebbe tendere a zero man mano che n cresce all'infinito. Bienaymé notò che alla stessa conclusione si sarebbe arrivati [1] se il numero medio di figli fosse stato esattamente uno, ad esempio se c'è una probabilità 1/2 di non avere figli e una probabilità 1/2 di avere due figli. Ma in questo caso la probabilità di avere figli alla n-sima generazione tende a zero più lentamente: nell'esempio sarebbe ancora il 5 % dopo 35 generazioni, cioè dopo undici o dodici secoli se ci sono tre generazioni per secolo [2]. Bienaymé notò infine che se il numero medio di figli è maggiore di uno, l'estinzione del ramo familiare non è certa: la sua probabilità può essere calcolata risolvendo qualche equazione algebrica.

L'articolo di Bienaymé non conteneva ulteriori spiegazioni. Nel 1847 il suo amico Antoine-Augustin Cournot, matematico ed economista, incluse

[1] Tranne se ogni uomo ha esattamente un figlio.

[2] Come vedremo in seguito, questa probabilità è uguale a $1 - x_{35}$ con $x_{n+1} = \frac{1}{2} + \frac{1}{2}x_n^2$ e $x_0 = 0$.

alcuni dettagli in un libro intitolato *Sull'origine e sui limiti della corrispondenza tra algebra e geometria*. Presentava il problema sotto forma di gioco d'azzardo, ma riconosceva che era identico allo studio di Bienaymé sull'estinzione dei nomi di famiglia. Mantenendo l'interpretazione in termini di nomi di famiglia, Cournot considerò prima il caso speciale in cui gli uomini hanno al massimo due figli, essendo p_0, p_1 e p_2 rispettivamente la probabilità di avere 0, 1 o 2 figli. Naturalmente, $p_0 + p_1 + p_2 = 1$. Partendo da un antenato, la probabilità di estinzione dopo una sola generazione, chiamata x_1, è ovviamente uguale a p_0. La probabilità di estinzione entro due generazioni è: $x_2 = p_0 + p_1 x_1 + p_2 {x_1}^2$, o la famiglia era già estinta nella prima generazione (probabilità p_0), o c'era un solo figlio nella prima generazione che non ha avuto figli maschi (probabilità $p_1 x_1$), o c'erano due figli nella prima generazione e ognuno di loro non ha avuto figli maschi (probabilità $p_2 x_1^2$). Più in generale, la probabilità di estinzione entro n generazioni è:

$$x_n = p_0 + p_1 x_{n-1} + p_2 (x_{n-1})^2.$$

Infatti, se ci sono per esempio due figli nella prima generazione (probabilità p_2), la famiglia si estinguerà $n-1$ generazioni dopo (cioè alla n-sima generazione) con una probabilità pari a $(x_{n-1})^2$. Cournot notò che x_n è una sequenza crescente con $x_n \leqslant 1$ per tutti gli n. Quindi x_n ha un limite $x_\infty \leqslant 1$, che è una soluzione dell'equazione $x = p_0 + p_1 x + p_2 x^2$. Usando $p_1 = 1 - p_0 - p_2$, questa equazione è equivalente a: $0 = p_2(x-1)(x - p_0/p_2)$. Quindi ci sono due radici: $x = 1$ e $x = p_0/p_2$. Si possono distinguere tre casi a seconda del numero medio di figli $p_1 + 2p_2$, che è anche uguale a: $1 - p_0 + p_2$ e che chiameremo $\mathscr{R}_0$. Se $\mathscr{R}_0 < 1$, allora $p_0/p_2 > 1$. Quindi $x = 1$ è l'unico valore possibile per il limite x_∞. Sicuramente il nome della famiglia si estinguerà. Se $\mathscr{R}_0 = 1$, entrambe le radici sono uguali a 1 e la conclusione è la stessa. Se $\mathscr{R}_0 > 1$, allora Cournot sostenne che x_∞ dovrebbe essere uguale alla seconda radice p_0/p_2, poiché la probabilità di estinzione deve ovviamente essere 0 nel caso speciale in cui $p_0 = 0$.

Cournot ha menzionato brevemente il caso più generale in cui gli uomini possono avere al massimo m figli con probabilità p_0, p_1,..., p_m. La conclusione dipende allo stesso modo dal valore di

$$\mathscr{R}_0 = p_1 + 2p_2 + \cdots + mp_m,$$

il numero medio di figli, rispetto a 1. L'equazione per x_∞, che è: $x = p_0 + p_1 x + \cdots + p_m x^m$, ha sempre la radice $x = 1$. Ha solo un'altra radice positiva, che dà la probabilità di estinzione x_∞ quando $\mathscr{R}_0 > 1$.

Purtroppo l'articolo di Bienaymé e le poche pagine del libro di Cournot passarono completamente inosservati all'epoca. L'articolo fu notato solo

negli anni '70 e le pagine del libro altri venti anni dopo! Nel frattempo il problema e la sua soluzione erano stati riscoperti da altri e l'argomento si era notevolmente sviluppato. Ritorneremo su questo argomento nei capitoli 9, 17 e 18.

Bienaymé dovette lasciare il suo lavoro al Ministero delle Finanze dopo la rivoluzione del 1848. Anche la cattedra di teoria della probabilità all'Università di Parigi, per la quale era certamente il miglior candidato, fu data a qualcun altro. Tuttavia Bienaymé fu in grado di lavorare di nuovo per il Ministero delle Finanze dopo il 1850, ma si dimise nel 1852. Più tardi quell'anno, fu eletto all'Accademia delle Scienze, dove era considerato l'esperto nel campo della statistica. Nel 1853 dimostrò quella che alcuni libri di testo moderni chiamano la disuguaglianza «Bienaymé–Čebyšëv». Nel 1875 divenne presidente della neonata Société Mathématique de France. Morì a Parigi nel 1878.

Bibliografia

1. Bienaymé, I.J.: De la loi de multiplication et de la durée des familles. *Extr. p. v. séances - Soc. Philomat. Paris*, 37–39 (1845) biodiversitylibrary.org
2. Bru, B.: À la recherche de la démonstration perdue de Bienaymé. *Math. Sci. Hum.* 114, 5–17 (1991). archive.numdam.org
3. Bru, B., Jongmans, F., Seneta, E.: I.J. Bienaymé: Family information and proof of the criticality theorem. *Int. Stat. Rev.* 60, 177–183 (1992)
4. Cournot, A.-A.: *De l'origine et des limites de la correspondance entre l'algèbre et la géométrie*. Hachette, Paris (1847). archive.org
5. Doubleday, T.: *The True Law of Population* (1842). archive.org
6. Heyde, C.C., Seneta, E.: I.J. Bienaymé: *Statistical Theory Anticipated*. Springer-Verlag, New York (1977)
7. Kendall, D.G.: The genealogy of genealogy: branching processes before (and after) 1873. *Bull. Lond. Math. Soc.* 7, 225–253 (1975)
8. Littré, É.: *Conservation, révolution et positivisme* (1852). gallica.bnf.fr
9. Malthus, T.R.: *An Essay on the Principle of Population*, 2nd edn. Bensley, London (1803). archive.org
10. Martin, T.: Antoine Augustin Cournot. In: Heyde, C.C., Seneta, E. (eds.) *Statisticians of the Centuries*, 152–156. Springer, New York (2001)
11. Seneta, E.: Irenée-Jules Bienaymé. In: Heyde, C.C., Seneta, E. (eds.) *Statisticians of the Centuries*, 132–136. Springer, New York (2001)

Capitolo 8

Mendel e l'ereditarietà (1865)

Nel 1865 Mendel pubblicò i risultati dei suoi esperimenti pionieristici sull'ibridazione dei piselli. La sua analisi utilizzava aspetti elementari della teoria della probabilità. Considerò anche un modello dinamico per una popolazione di piante autofertilizzanti. Il suo lavoro, riscoperto solo nel 1900, è una pietra miliare nella storia della genetica.

Johann Mendel nacque nel 1822 in Moravia, allora parte dell'impero austriaco e oggi parte della Repubblica Ceca. Suo padre era un contadino. Con i suoi buoni risultati al liceo e la sua salute cagionevole, Mendel preferì continuare a studiare piuttosto che lavorare nella fattoria di famiglia. Ma non poteva permettersi di andare all'università. Così nel 1843 entrò nell'abbazia di San Tommaso a Bruna (oggi Brno), dove prese il nome di Gregor. Studiò teologia, ma frequentò anche alcuni corsi di agricoltura. Nel 1847 fu ordinato sacerdote. Insegnò in una scuola superiore per alcuni anni, ma non superò l'esame per diventare professore ordinario. Tra il 1851 e il 1853, grazie al sostegno dei suoi superiori religiosi, poté comunque continuare gli studi all'Università di Vienna dove frequentò corsi di fisica, matematica e scienze naturali. Dopo di che tornò a Bruna e insegnò fisica in un istituto tecnico.

Figura 8.1:
Mendel (1822–1884)

Tra il 1856 e il 1863, Mendel fece una serie di esperimenti su un gran numero di piante nel giardino della sua abbazia. Nel 1865 presentò i suoi

risultati in due riunioni della Società di Storia Naturale di Bruna, di cui era membro. Il suo lavoro, *Esperimenti sull'ibridazione delle piante*, fu pubblicato in tedesco l'anno seguente negli atti della Società. Mendel spiegò come era arrivato a studiare le variazioni dei piselli, piante che si riproducono naturalmente per autofecondazione e i cui semi possono assumere diverse forme facilmente identificabili: rotondo o rugoso, giallo o verde, ecc. Incrociando una pianta proveniente da una stirpe con semi tondi e una pianta proveniente da una stirpe con semi rugosi, notò che otteneva sempre ibridi che davano semi tondi. Chiamò il carattere «semi tondi» dominante e il carattere «semi rugosi» recessivo. Mostrò allo stesso modo che il carattere «semi gialli» era dominante e che il carattere «semi verdi» era recessivo.

Mendel notò poi che l'autofecondazione di piante cresciute da semi ibridi dava nella prima generazione nuovi semi che avevano o il carattere dominante o il carattere recessivo in proporzioni apparentemente casuali. Inoltre, notò che ripetendo l'esperimento molte volte otteneva in media circa tre volte più semi con il carattere dominante che con il carattere recessivo. Per esempio, in un primo esperimento ottenne un totale di 5 474 semi rotondi e 1 850 semi rugosi, corrispondenti a un rapporto di 2,96 a 1. Un secondo esperimento ha dato un totale di 6 022 semi gialli e 2 001 semi verdi, corrispondenti a un rapporto di 3,01 a 1 [1].

Mendel notò anche che tra le piante cresciute dai semi della prima generazione con il carattere dominante, quelle che davano per autofecondazione semi sia con il carattere dominante che recessivo erano circa il doppio di quelle che davano solo semi con il carattere dominante. Per esempio, tra le 565 piante cresciute da semi rotondi della prima generazione, 372 hanno dato sia semi rotondi che rugosi mentre 193 hanno dato solo semi rotondi; il rapporto è pari a 1,93. Allo stesso modo, tra le 519 piante cresciute da semi gialli della prima generazione, 353 hanno dato sia semi gialli che verdi mentre 166 hanno dato solo semi gialli; il rapporto è pari a 2,13.

Per spiegare questi risultati, Mendel ebbe la brillante idea di considerare il carattere apparente di un seme come il risultato dell'associazione di due fattori nascosti, ognuno di questi fattori essendo o dominante (scritto A) o recessivo (scritto a). Quindi ci sono tre combinazioni possibili: AA, Aa e aa. I semi con i fattori AA o Aa hanno lo stesso carattere dominante A. I semi con il fattore aa hanno il carattere recessivo a. Mendel assunse inoltre che

[1] Come R. A. Fisher (vedi capitolo 14) notò più tardi, la probabilità di arrivare a risultati sperimentali così vicini al valore teorico è piuttosto piccola. Mendel ha probabilmente organizzato i suoi dati. Per esempio, nel secondo esperimento riguardante $n = 6022 + 2001 = 8023$ semi, la probabilità che il rapporto differisca da 3 per meno di 0,01 è solo circa 10 %.

durante la fecondazione, i grani di polline e gli ovuli (i gameti) trasmettono solo uno dei due fattori, ciascuno con una probabilità 1/2.

Quindi l'incrocio dei lignaggi puri AA e aa dà degli ibridi che hanno tutti i fattori Aa e il carattere dominante A. I gameti dell'ibrido Aa trasmettono il fattore A con probabilità 1/2 e il fattore a con probabilità 1/2. L'autofecondazione di una pianta cresciuta da un seme ibrido Aa dà quindi AA con probabilità 1/4, Aa con probabilità $1/2$ e aa con probabilità $1/4$, come mostrato nella Tabella 8.1.

Tabella 8.1: Possibili risultati dell'autofecondazione di un ibrido Aa e loro probabilità in funzione dei fattori trasmessi dai gameti maschili (in righe) e dai gameti femminili (in colonne).

Fattore (Probabilità)	A (1/2)	a (1/2)
A (1/2)	AA (1/4)	Aa (1/4)
a (1/2)	Aa (1/4)	aa (1/4)

Mendel notò che le proporzioni $AA : Aa : aa$, che erano $1 : 2 : 1$, potevano anche essere ottenute dal calcolo formale $(A+a)^2 = AA + 2Aa + aa$. Poiché i semi AA e Aa hanno il carattere apparente A mentre solo i semi aa hanno il carattere apparente a, ci sono effettivamente tre volte più semi con il carattere A che con il carattere a. Inoltre, ci sono in media il doppio dei semi Aa che AA. L'autofecondazione delle piante cresciute dai primi dà semi con il carattere dominante (AA o Aa) o il carattere recessivo (aa). Per quanto riguarda l'autofecondazione delle piante cresciute dai semi AA, essa dà sempre semi AA con il carattere dominante. Tutte le osservazioni sono così spiegate.

Mendel considerò anche le generazioni successive. Partendo da N semi ibridi Aa e assumendo per semplicità che ogni pianta dia per autofecondazione solo quattro nuovi semi, calcolò che il numero medio di semi $(AA)_n$, $(Aa)_n$ e $(aa)_n$ nella generazione n sarebbe dato dalla Tabella 8.2, dove per chiarezza di presentazione i risultati sono stati divisi per N.

Tabella 8.2: Generazioni successive.

n	0	1	2	3	4	5
$(AA)_n$	0	1	6	28	120	496
$(Aa)_n$	1	2	4	8	16	32
$(aa)_n$	0	1	6	28	120	496
totale	1	4	16	64	256	1024

Questi numeri si ottengono semplicemente dalle formule

$$(AA)_{n+1} = (Aa)_n + 4\,(AA)_n\ , \tag{8.1}$$

$$(Aa)_{n+1} = 2\,(Aa)_n\ , \tag{8.2}$$

$$(aa)_{n+1} = (Aa)_n + 4\,(aa)_n\ , \tag{8.3}$$

che dicono che AA dà dopo l'autofecondazione quattro semi AA, che aa dà quattro semi aa e che Aa dà in media un seme AA, due semi Aa e un seme aa. Mendel notò inoltre che $(AA)_n = (aa)_n = 2^{n-1}(2^n - 1)$ e $(Aa)_n = 2^n$.

Infatti, dall'equazione (8.2) e dalla condizione iniziale $(Aa)_0 = 1$ segue che $(Aa)_n = 2^n$. Sostituendo questo nell'equazione (8.1), otteniamo che $(AA)_{n+1} = 4\,(AA)_n + 2^n$. Ci rendiamo facilmente conto che $(AA)_n = c\,2^n$ è una soluzione particolare quando $c = -1/2$. La soluzione generale dell'equazione «omogenea» $(AA)_{n+1} = 4\,(AA)_n$ è $(AA)_n = C4^n$. Infine, sommando queste due soluzioni, vediamo che $(AA)_n = C4^n - 2^{n-1}$ soddisfa la condizione iniziale $(AA)_0 = 0$ se $C = 1/2$. Per quanto riguarda la sequenza $(aa)_n$, essa soddisfa la stessa relazione di ricorrenza e la stessa condizione iniziale di $(AA)_n$. Quindi $(aa)_n = (AA)_n$.

In conclusione, la proporzione di ibridi Aa nella popolazione totale, che è $2^n/4^n = 1/2^n$, è divisa per due ad ogni generazione per autofecondazione.

Il lavoro di Mendel passò totalmente inosservato durante la sua vita. Alcuni anni dopo, Mendel tentò anche esperimenti simili con altre specie di piante, pubblicò alcuni articoli sulla meteorologia e studiò l'ereditarietà delle api. Dopo essere diventato abate nel 1868, passò la maggior parte del suo tempo a gestire problemi amministrativi. Morì nel 1884.

È solo nel 1900 che il lavoro di Mendel fu finalmente riscoperto indipendentemente e quasi contemporaneamente da Hugo De Vries ad Amsterdam, Carl Correns a Tubinga e Erich von Tschermak a Vienna. Ciò avrebbe dato inizio ad una nuova era quella che oggi chiamiamo genetica.

Bibliografia

1. Bateson, W.: *Mendel's Principles of Heredity*. Cambridge University Press (1913). archive.org
2. Mendel, J.G.: *Versuche über Pflanzenhybriden. Verh. Naturforsch. Ver. Brünn* 4, 3–47 (1866). www.esp.org
3. Fisher, R.A.: Has Mendel's work been rediscovered? *Ann. Sci.* 1, 115–137 (1936). library.adelaide.edu.au

Capitolo 9

Galton, Watson e il problema dell'estinzione (1873–1875)

Nel 1873 lo statistico britannico Galton e il suo compatriota matematico Watson considerarono il problema dell'estinzione dei nomi di famiglia senza conoscere il lavoro di Bienaymé. Watson notò che la funzione generatrice associata alla distribuzione di probabilità del numero di uomini in ogni generazione poteva essere calcolata ricorsivamente. Ma analizzò erroneamente la probabilità di estinzione.

Francis Galton nacque nel 1822, lo stesso anno di Mendel, vicino a Birmingham in Inghilterra. Era il più giovane di sette figli. Suo padre era un ricco banchiere. Per parte di madre, era il cugino di Charles Darwin. Galton ha iniziato a studiare medicina nel 1838, prima in un ospedale di Birmingham e poi a Londra. Durante l'estate del 1840, fece il suo primo lungo viaggio attraversando l'Europa fino a Istanbul. Successivamente studiò al *Trinity College* dell'Università di Cambridge per quattro anni. Ma suo padre morì nel 1844, lasciando in eredità un'ingente fortuna. Galton abbandonò l'idea di diventare medico. Viaggiò in Egitto, Sudan e Siria. Durante gli anni che seguirono mantenne uno stile di vita agiato, trascorrendo il suo tempo a cacciare, viaggiando su mongolfiere e navi o cercando di migliorare il telegrafo elettrico. Nel 1850 organizzò una spedizione esplorativa nell'Africa sud-occidentale (oggi Namibia). Al suo ritorno in Inghilterra nel 1852, fu eletto alla *Royal Geographical Society*. Lì poté avere le notizie delle spedizioni in Africa orientale alla ricerca della sorgente del Nilo. Si stabilì a Londra e scrisse una manuale per i viaggiatori che divenne un best seller. Nel 1856 fu eletto alla *Royal Society*. Si interessò allora di meteorologia e inventò la parola «anticlone». Dopo la pubblicazione nel 1859 da parte di suo cugino Darwin di *L'origine delle specie*, Galton rivolse la sua attenzione allo studio dell'ereditarietà. Pubblicò *Genio ereditario* nel 1869, in cui sosteneva che le facoltà intellettuali potevano essere trasmesse per ereditarietà.

Nel 1873 Alphonse de Candolle, un botanico svizzero, pubblicò un libro intitolato *Storia della scienza e degli scienziati negli ultimi due secoli*, che conteneva anche un saggio su *La rispettiva influenza dell'ereditarietà, della variabilità e della selezione sullo sviluppo della specie umana e sul probabile futuro di questa specie*. Lì fece le seguenti osservazioni:

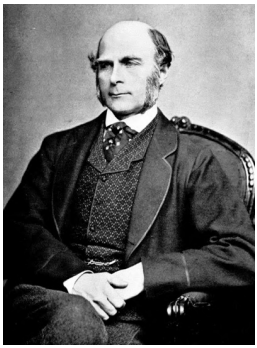

Figura 9.1: Galton (a sinistra) e Watson (a destra).

> «Tra le precise informazioni e le opinioni molto sane del signor Benoiston de Châteauneuf, di Galton e di altri statistici, non ho visto l'importante osservazione che avrebbero dovuto fare sull'inevitabile estinzione dei nomi di famiglia. Naturalmente, ogni nome deve estinguersi [...] Un matematico potrebbe calcolare in che modo possa avvenire la diminuzione dei nomi o dei titoli, conoscendo la probabilità di avere figli femmine o maschi e la probabilità di non avere figli per ogni data coppia.»

È lo stesso problema che Bienaymé aveva studiato nel 1845. Ma Candolle, che non era a conoscenza del lavoro di Bienaymé, pensava che tutte le famiglie fossero destinate all'estinzione. Galton notò il suddetto paragrafo nel libro di Candolle. Poiché anche lui non conosceva il lavoro di Bienaymé, Galton lo evidenziò come un problema aperto per i lettori di *Educational Times*:

> «Problema 4001: Una grande nazione, di cui ci occuperemo solo dei maschi adulti, in numero N, e che portano ciascuno cognomi separati, colonizza un distretto. La loro legge della popolazione, per ogni generazione, è tale che: a_0 per cento dei maschi adulti non hanno figli maschi che raggiungono la vita adulta; a_1 hanno un figlio maschio di questo tipo; a_2 ne hanno due; e così via fino a a_5 che ne hanno cinque.
>
> Trova: (1) quale proporzione dei loro cognomi si sarà estinta dopo r generazioni; (2) in quanti casi il cognome sarà posseduto

da m persone.»

Si noti che la seconda parte del problema non era stata affrontata da Bienaymé. Galton non ricevette alcuna risposta soddisfacente dai lettori della rivista e apparentemente non poteva trovare la soluzione del problema da solo. Così chiese al suo amico Henry William Watson, un matematico, di provare a risolverlo.

Watson è nato a Londra nel 1827. Suo padre era un ufficiale della marina britannica. Studiò prima al *King's College* di Londra e poi si dedicò alla matematica al *Trinity College* dell'Università di Cambridge, dal 1846 al 1850, pochi anni dopo Galton. Divenne successivamente membro del *Trinity College*, assistente alla *City of London School*, docente di matematica al *King's College* e professore di matematica alla *Harrow School* tra il 1857 e il 1865. Appassionato di alpinismo, fece parte di una spedizione che raggiunse la cima del Monte Rosa in Svizzera nel 1855. Fu ordinato diacono nel 1856 e prete anglicano due anni dopo. Dal 1865 fino al suo ritiro fu rettore di Berkswell con Barton vicino a Coventry, incarico che sottrasse abbastanza tempo per la ricerca.

Galton e Watson scrissero insieme un articolo intitolato *Sulla probabilità di estinzione delle famiglie*, che fu pubblicato nel 1875 nel *Journal of the Royal Anthropological Institute*. Galton espose il problema, Watson illustrò i suoi calcoli e le conclusioni a cui era giunto. Assunsero che gli uomini potessero avere al massimo q figli, indicando con p_k la probabilità di avere k figli ($k = 0, 1, 2, \ldots, q$). In altre parole, $p_k = a_k/100$ se usiamo le notazioni originali di Galton. Quindi

$$p_0 + p_1 + \cdots + p_q = 1.$$

Consideriamo la situazione in cui la 0-esima generazione (iniziale) consiste in un solo uomo. L'1-esima generazione consiste di s uomini con una probabilità p_s. Usando un'artificio che era ben noto ai suoi contemporanei e che era stato introdotto molto tempo prima da Abraham de Moivre, Watson considerò la funzione generatrice:

$$f(x) = p_0 + p_1 x + p_2 x^2 + \cdots + p_q x^q \tag{9.1}$$

associati alle probabilità $p_0, \ldots, p_q$. Allo stesso modo, sia $f_n(x)$ il polinomio il cui coefficiente di x^s rappresenta la probabilità di avere s maschi alla n-sima generazione, partendo da un uomo della 0-esima generazione. Allora $f_1(x) = f(x)$. Watson ha notato che:

$$f_n(x) = f_{n-1}(f(x)), \tag{9.2}$$

è una formula che permette di calcolare $f_n(x)$ ricorsivamente.

Infatti, poniamo

$$f_n(x) = p_{0,n} + p_{1,n}, x + p_{2,n}, x^2 + \cdots + p_{q^n,n}\, x^{(q^n)}.$$

Notate che ci sono al massimo q^n uomini alla n-esima generazione . Se nella $n-1$-esima generazione ci sono s uomini numerati da 1 a s, t_1,...,t_s è il numero dei loro figli maschi. In tal caso, ci saranno t uomini nella generazione n con una probabilità pari a:

$$\sum_{t_1+\cdots+t_s=t} p_{t_1} \times \cdots \times p_{t_s}\ .$$

Quando $s = 0$, bisogna capire che questa probabilità è uguale a 1 se $t = 0$ e uguale a 0 se $t \geqslant 1$. Pertanto:

$$p_{t,n} = \sum_{s\geqslant 0} p_{s,n-1} \times \sum_{t_1+\cdots+t_s=t} p_{t_1} \times \cdots \times p_{t_s}\ .$$

Ne consegue che:

$$\begin{aligned} f_n(x) &= \sum_{t\geqslant 0} p_{t,n}\, x^t \\ &= \sum_{s\geqslant 0} p_{s,n-1} \sum_{t\geqslant 0} \sum_{t_1+\cdots+t_s=t} (p_{t_1} x^{t_1}) \times \cdots \times (p_{t_s} x^{t_s}) \\ &= \sum_{s\geqslant 0} p_{s,n-1} \left[p_0 x^0 + p_1 x^1 + p_2 x^2 + \cdots\right]^s \\ &= \sum_{s\geqslant 0} p_{s,n-1} [f(x)]^s = f_{n-1}(f(x))\ . \end{aligned}$$

In particolare la probabilità x_n di estinzione del nome della famiglia entro n generazioni è uguale a $p_{0,n}$, che è uguale a $f_n(0)$. Come primo esempio, Watson considerò la seguente funzione generatrice:

$$f(x) = (1 + x + x^2)/3,$$

cioè $q = 3$ e $p_0 = p_1 = p_2 = 1/3$. Calcolò i polinomi $f_n(x)$ per $n = 1,\ldots,4$ usando l'equazione (9.2). Per esempio, ha ottenuto:

$$f_2(x) = \frac{1}{3}\left[1 + \frac{1+x+x^2}{3} + \left(\frac{1+x+x^2}{3}\right)^2\right] = \frac{13+5x+6x^2+2x^3+x^4}{27}$$

e $f_2(0) = 13/27 \approx 0{,}481$. Il calcolo di $f_n(x)$ per $n = 3$ diventa molto noioso, così noioso che Watson ha già fatto un errore per $n = 4$. Poiché $x_5 = f_5(0) = f_4(f(0))$ può evitare il calcolo di $f_5(x)$, ottenendo la seguente lista di probabilità di estinzione $x_n = f_n(0)$:

$$x_1 \approx 0{,}333, \quad x_2 \approx 0{,}481, \quad x_3 \approx 0{,}571, \quad x_4 \approx 0{,}641, \quad x_5 \approx 0{,}675\,.$$

I valori corretti sono $x_4 \approx 0{,}632$ e $x_5 \approx 0{,}677$, come si può verificare usando la semplice formula $x_n = f(x_{n-1})$ derivata da Bienaymé. Come vedremo nel capitolo 17, quest'ultima formula può anche essere derivata dall'equazione (9.2).

Watson ha notato,nel suo primo esempio, che ogni uomo ha in media $\mathscr{R}_0 = p_1 + 2\,p_2 + \cdots + q\,p_q$ figli e che $\mathscr{R}_0 = 1$. Quindi si potrebbe pensare che se il numero iniziale di membri maschili della famiglia fosse abbastanza grande, la dimensione della famiglia rimarrebbe approssimativamente costante. Tuttavia Watson sosteneva che la probabilità di estinzione x_n converge verso 1 quando $n \to +\infty$, anche se abbastanza lentamente. In altre parole l'intera famiglia raggiungerà l'estinzione come Candolle aveva suggerito. La figura 9.2a, che non è disegnata nell'articolo originale, e i risultati di Bienaymé confermano che questa conclusione per il primo esempio è corretta.

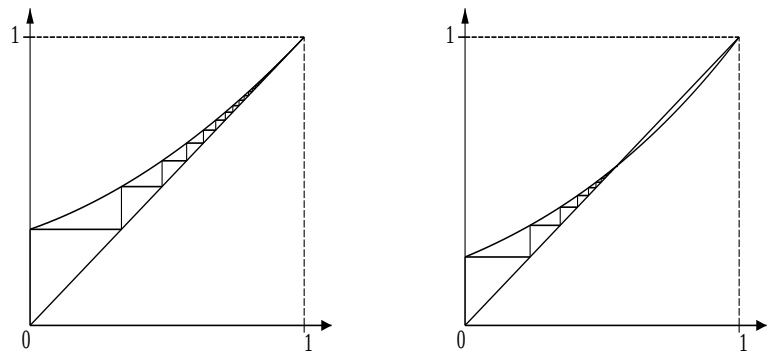

Figura 9.2: Grafico delle funzioni $y = f(x)$ e $y = x$. La probabilità di estinzione $x_n = f(x_{n-1})$ entro n generazioni è l'altezza del n-esimo «passo della scala». (a): $f(x) = (1+x+x^2)/3$. (b): $f(x) = (3+x)^5/4^5$.

Come secondo esempio, Watson ha considerato la distribuzione di probabilità binomiale

$$p_k = \binom{q}{k} \frac{a^{q-k} b^k}{(a+b)^q}, \tag{9.3}$$

per cui la funzione generatrice (9.1) è

$$f(x) = (a+bx)^q/(a+b)^q.$$

Egli calcolò $f_2(x)$ e $x_2 = f_2(0)$. A questo punto si rese conto che $x_2 = f(x_1)$ e che

$$x_n = f(x_{n-1})$$

per tutti gli n. Ma pensava che questa formula fosse vera solo per il caso speciale binomiale (9.3). Applicandola al caso in cui $q = 5$, $a = 3$ e $b = 1$, ottenne

$$x_1 \approx 0{,}237, \;\; x_2 \approx 0{,}347, \;\; x_3 \approx 0{,}410, \; \ldots \; x_9 \approx 0{,}527, \;\; x_{10} \approx 0{,}533\ldots$$

Watson si rese conto che x_n converge ad un limite x_∞ come $n \to +\infty$, che soddisfa

$$x_\infty = f(x_\infty) = (a+bx_\infty)^q/(a+b)^q.$$

Notò che $x = 1$ è una soluzione di questa equazione ma non si rese conto che potevano esserci altre soluzioni quando $\mathscr{R}_0 > 1$. Così concluse erroneamente, fuorviato da Candolle, che c'è estinzione ($x_\infty = 1$) in ogni caso, compreso l'esempio numerico che aveva considerato. La figura 9.2b mostra che non è questo il caso!

Watson notò che il numero medio di figli, in questo esempio numerico, era maggiore di 1 (si può dimostrare che $\mathscr{R}_0 = qb/(a+b) = 5/4$), il che significa che la popolazione tende ad aumentare esponenzialmente. Ma questo non lo aiutò a scoprire il suo errore. Addirittura ipotizzò che l'estinzione del cognome fosse certa per ogni distribuzione di probabilità (p_k), cioè non solo per il caso binomiale. Torneremo su questo problema nei capitoli 17 e 18.

Galton continuò il suo studio statistico delle famiglie con un libro intitolato *Uomini di scienza inglesi, la loro natura e la loro educazione*, che si concentrava sulla genealogia dei membri della *Royal Society*. Si interessò anche all'antropometria, la misura del corpo umano. Approfittò di una mostra internazionale nel 1884 a Londra per raccogliere dati su un gran numero di persone. I suoi risultati furono pubblicati nel 1889 in un libro intitolato *Eredità naturale*, la cui appendice riproduceva l'articolo scritto in collaborazione con Watson. Questo libro introdusse anche alcuni nuovi vocaboli statistici

come «percentile» e «quartile» così come la parola «eugenetica», cioè il miglioramento della specie umana dal punto di vista dei caratteri ereditari. Dopo il 1888 Galton sviluppò la tecnica di riconoscimento delle impronte digitali che sarebbe stata usata qualche anno dopo dalla polizia britannica. Continuò anche a studiare il rispettivo ruolo dell'ereditarietà (natura) e dell'ambiente (nurture) sulle caratteristiche fisiche e intellettuali dei gemelli, sulle dimensioni dei piselli cresciuti per diverse generazioni o sul colore dei topi allevati in laboratorio. Questo lo portò alla nozione di coefficiente di correlazione tra due variabili. Nel 1904 fu fondato il *Galton Laboratory* all'interno dell'*University College* di Londra. Galton fu nominato cavaliere nel 1909 e morì nel 1911.

Watson pubblicò diversi libri, in particolare un trattato sulla teoria cinetica dei gas nel 1876 e un trattato sulla teoria matematica dell'elettricità e del magnetismo in due volumi (1885 e 1889). Nel 1881 fu eletto alla *Royal Society* e nel 1903 morì a Brighton.

Nel 1924, secondo volume della sua biografia di Galton, Karl Pearson riassunse l'articolo sull'estinzione dei nomi di famiglia senza notare l'errore. Questo errore sarebbe stato finalmente notato nel 1930 (vedi capitolo 18).

Bibliografia

1. De Candolle, A.: *Histoire des sciences et des savants depuis deux siècles suivie d'autres études sur des sujets scientifiques en particulier sur la sélection dans l'espèce humaine*. Georg, Genève (1873). archive.org
2. Galton, F.: *Natural Inheritance*. Macmillan, London (1889). galton.org
3. Galton, F.: *Memories of my Life*. Methuen & Co., London (1908). galton.org
4. Kendall, D.G.: Branching processes since 1873. *J. Lond. Math. Soc.* 41, 385–406 (1966)
5. Pearson, K.: *The Life, Letters and Labours of Francis Galton*, vol. 1. Cambridge University Press (1914). galton.org
6. Pearson, K.: *The Life, Letters and Labours of Francis Galton*, vol. 2. Cambridge University Press (1924). galton.org
7. S.H.B.: Henry William Watson, 1827-1903. *Proc. R. Soc. Lond.* 75, 266–269 (1905). gallica.bnf.fr
8. Watson, H.W., Galton, F.: On the probability of the extinction of families. *J. Anthropol. Inst.* 4, 138–144 (1875). galton.org

Capitolo 10

Lotka e la teoria della popolazione stabile (1907–1911)

Nel 1907 il chimico americano Alfred Lotka iniziò a studiare la relazione tra il tasso di natalità, i tassi di mortalità specifici per età e il tasso di crescita della popolazione, usando un modello a tempo continuo. Nel 1911 in collaborazione con F.R. Sharpe pubblicò sullo stesso argomento un ulteriore articolo che includeva anche i tassi di fertilità specifici per età. L'equazione implicita che fornisce il tasso di crescita della popolazione è spesso chiamata «equazione di Lotka».

Alfred James Lotka nacque da genitori americani nel 1880 a Lemberg, che faceva parte dell'impero austro-ungarico (oggi Leopoli in Ucraina). Studiò prima in Francia e in Germania e nel 1901 si laureò in fisica e chimica all'Università di Birmingham in Inghilterra. Passò poi un anno a Lipsia, dove il ruolo della termodinamica nella chimica e nella biologia veniva evidenziato da Wilhelm Ostwald, che avrebbe ricevuto il premio Nobel per la chimica nel 1909. Lotka si stabilì a New York nel 1902 e iniziò a lavorare per la *General Chemical Company*.

Figura 10.1:
Lotka (1880–1949)

Nel 1907 e nel 1911[1], Lotka intraprese lo studio della dinamica delle popolazioni stratificate in base all'età senza conoscere il lavoro di Eulero sullo stesso argomento (vedere il capitolo 3). A differenza di Eulero, egli assume

[1] Il secondo articolo è stato scritto in collaborazione con F.R. Sharpe, un matematico della Università di Cornell.

che il tempo e l'età siano variabili continue. Sia $B(t)$ il tasso di natalità maschile (il numero di nascite maschili per unità di tempo) al tempo t, $p(x)$ la probabilità di essere ancora in vita all'età x e $h(x)$ la fertilità all'età x: $h(x)\,dx$ è la probabilità per un uomo di avere un figlio neonato tra l'età x e $x+dx$, dove dx è un infinitesimo. Allora $\int_0^{+\infty} p(x)\,dx$ è l'aspettativa di vita alla nascita. Inoltre $B(t-x)\,p(x)\,dx$ è il numero di maschi nati tra il tempo $t-x$ e $t-x+dx$, che sono ancora vivi al tempo t. Questi maschi hanno $B(t-x)\,p(x)\,h(x)\,dx$ figli per unità di tempo al tempo t. Quindi il tasso di natalità maschile totale al tempo t è

$$B(t) = \int_0^{+\infty} B(t-x)\,p(x)\,h(x)\,dx.$$

Cercando una soluzione esponenziale di questa equazione integrale con l'incognita $B(t)$ nella forma $B(t) = b\mathrm{e}^{rt}$, Lotka ottenne, dividendo membro a membro per $B(t)$, l'equazione in r

$$1 = \int_0^{+\infty} \mathrm{e}^{-rx}\,p(x)\,h(x)\,dx\,, \tag{10.1}$$

che i demografi ora chiamano «equazione di Lotka»[2]. Euler aveva ottenuto l'analoga equazione implicita (3.1) per il tasso di crescita quando il tempo e l'età sono variabili discrete. Lotka notò che il membro destro di (10.1) è una funzione decrescente di r che tende a $+\infty$ quando $r \to -\infty$ e che tende a 0 quando $r \to +\infty$. Quindi esiste un unico valore di r, che denotiamo con r^*, che soddisfa l'equazione (10.1). Inoltre, $r^* > 0$, se e solo se

$$\mathscr{R}_0 = \int_0^{+\infty} p(x)\,h(x)\,dx > 1\ . \tag{10.2}$$

Il parametro $\mathscr{R}_0$ (la notazione fu introdotta da Dublin e Lotka nel 1925) è il numero previsto di figli che un uomo può avere nel corso della sua vita. Lotka suggerì[3] che, qualunque fosse la stratificazione iniziale per età della popolazione maschile, il numero di nascite maschili per unità di tempo era effettivamente tale che $B(t) \sim b\mathrm{e}^{r^*t}$ quando $t \to +\infty$, dove b è una costante. La popolazione totale è poi data da $P(t) = \int_0^{+\infty} B(t-x)\,p(x)\,dx$. Ne consegue che anche $P(t)$ aumenta o diminuisce come e^{r^*t} quando $t \to +\infty$: il tasso di

[2] R.A. Fisher arrivò indipendentemente alla stessa equazione nel 1927 e più tardi interpretò la radice r^* come una misura della «fitness darwiniana» nella teoria dell'evoluzione per selezione naturale.

[3] Questo fu rigorosamente dimostrato nel 1941 da Willy Feller, che allora era professore di matematica alla Università Brown negli Stati Uniti. Un approccio probabilistico fu sviluppato nel 1968 da Crump, Mode e Jagers.

crescita è uguale a r^*. Inoltre, la stratificazione per età della popolazione, data da $B(t-x)\,p(x)/\,P(t)$, tende a

$$\frac{\mathrm{e}^{-r^*x}\,p(x)}{\int_0^{+\infty}\mathrm{e}^{-r^*y}\,p(y)\,dy}\,.$$

Questa è ciò che Lotka chiamava una «popolazione stabile»: la piramide dell'età mantiene la stessa forma nel tempo, ma la popolazione totale aumenta o diminuisce esponenzialmente. La conclusione è quindi la stessa del modello a tempo discreto di Eulero. Ma lo studio di Lotka tiene conto della dipendenza della fertilità dall'età. Quindi è, in un certo senso, più generale di quello di Eulero.

Lotka continuò a lavorare su questo argomento per tutta la sua vita. Nel 1908–1909 riprese i suoi studi alla Università di Cornell per ottenere un master. Lavorò per il *National Bureau of Standards* dal 1909 al 1911 e dal 1911 al 1914 come editore della rivista *Scientific American Supplement*. Nel 1912 ottenne un dottorato all'Università di Birmingham raccogliendo gli articoli che aveva pubblicato dal 1907 sulla dinamica della popolazione e sulla demografia. Durante la prima guerra mondiale, lavorò ancora per la *General Chemical Company* su come fissare l'azoto dall'atmosfera. Nel 1920 uno dei suoi articoli sulle oscillazioni biologiche (vedi capitolo 13) fece una profonda impressione su Raymond Pearl, un professore di biometria alla Università Johns Hopkins, che aveva appena «riscoperto» l'equazione logistica (vedi capitolo 6). Sperando di trovare un'occupazione al Istituto Rockefeller per la ricerca medica di New York, Lotka lavorò sui modelli matematici sviluppati da Ross per la malaria (vedi capitolo 12). Infine ottenne, dall'Università Johns Hopkins, una borsa di studio biennale che gli permise di scrivere un libro intitolato *Elementi di biologia fisica*, pubblicato nel 1925. Divenne poi capo del dipartimento di ricerca della *Metropolitan Life Insurance Company* di New York. Si concentrò sull'analisi matematica delle questioni demografiche e pubblicò diversi libri in collaborazione con un collega, lo statistico e vicepresidente della compagnia Louis Israel Dublin: *Il valore monetario di un uomo* (1930), *Durata della vita* (1936) e *Venticinque anni di progresso della salute* (1937). Fu eletto presidente della *Population Association of America* per il 1938–1939. Tra i suoi vari studi statistici, la «legge di Lotka» (risalente al 1926) afferma che il numero di autori che hanno scritto n articoli in un dato campo scientifico diminuisce più o meno come $1/n^2$ all'aumentare di n. Lotka pubblicò anche un libro in francese intitolato *Teoria analitica delle associazioni biologiche*. La prima parte, più filosofica, apparve nel 1934. La seconda parte, più tecnica, pubblicata nel 1939, riassumeva tutte le sue ricerche sulla demografia umana a partire dal 1907. Nel suo libro

Lotka presentò anche il suo contributo al problema dell'estinzione dei nomi di famiglia. Dopo la pubblicazione nel 1930 del primo articolo di Steffensen sull'argomento (vedi capitolo 18), aveva applicato la teoria ai dati contenuti nel censimento del 1920 della popolazione bianca degli Stati Uniti. Aveva notato che la distribuzione osservata $(p_k)_{k\geqslant 0}$ del numero di figli è ben approssimata da una legge geometrica decrescente in $k \geqslant 1$: $p_0 = a$, $p_k = b\,c^{k-1}$ $(k \geqslant 1)$, con $a = 0{,}4825$, $b = 0{,}2126$ e $c = 1 - b/(1-a)$. In questo modo, $\sum_{k\geqslant 0} p_k = 1$. La funzione generatrice associata è $f(x) = a + bx/(1-cx)$. Le due soluzioni dell'equazione $x = f(x)$ sono $x = 1$ e $x = a/c$. La probabilità di estinzione x_∞ è la più piccola di queste due soluzioni (vedi capitolo 7). Con i valori numerici relativi agli USA trovava $x_\infty \approx 0{,}819$, mentre il numero medio di figli era $\mathscr{R}_0 = f'(1) = (1-a)^2/b \approx 1{,}260$. Nonostante un numero medio di figli (comprendente sia figli che figlie) vicino a 2,5, la probabilità di estinzione del nome della famiglia risulta superiore all'80 %.

Lotka fu eletto presidente dell'*American Statistical Association* nel 1942. Si ritirò nel 1947 e morì nel 1949 nel New Jersey. Una nuova edizione del suo libro del 1925 apparve nel 1956 con il titolo leggermente diverso *Elementi di biologia matematica*.

Bibliografia

1. Crump, K.S., Mode, C.J.: A general age-dependent branching process. *J. Math. Anal. Appl.* 24, 494–508 (1968)
2. Dublin, L.I., Lotka, A.J.: On the true rate of natural increase. *J. Amer. Stat. Assoc.* 20, 305–339 (1925)
3. Feller, W.: On the integral equation of renewal theory. *Ann. Math. Stat.* 12, 243–267 (1941). projecteuclid.org
4. Fisher, R.A.: The actuarial treatment of official birth records. *Eugen. Rev.* 19, 103–108 (1927). digital.library.adelaide.edu.au
5. Gridgeman, N.T.: Lotka, Afred James. In Gillespie, C.C. (ed.) *Dictionary of Scientific Biography*, vol. 8, p. 512. Scribner, New York (1981)
6. Kingsland, S.E.: *Modeling Nature*, 2nd edn., University of Chicago (1995).
7. Lotka, A.J.: Relation between birth rates and death rates. *Science* 26, 21–22 (1907) → Smith & Keyfitz (1977).
8. Lotka, A.J.: *Théorie analytique des associations biologiques*, 2e partie. Hermann, Paris (1939). gallica.bnf.fr
9. Sharpe, F.R., Lotka, A.J.: A problem in age-distribution. *Philos. Mag. Ser.* 6, 21, 435–438 (1911) → Smith & Keyfitz (1977).
10. Smith, D.P., Keyfitz, N.: *Mathematical Demography*. Springer (1977)

Capitolo 11

La legge di Hardy-Weinberg (1908)

Nel 1908 il matematico britannico Hardy e il medico tedesco Weinberg, indipendentemente l'uno dall'altro scoprirono, per una popolazione infinitamente grande, cioè sufficientemente grande, per cui l'accoppiamento casuale segue le leggi di Mendel, che le frequenze dei genotipi ottenuti da due alleli rimangono costanti attraverso le generazioni. Il loro modello matematico fu uno dei punti di partenza della genetica delle popolazioni.

Godfrey Harold Hardy è nato nel 1877 nel Surrey, in Inghilterra. I suoi genitori erano insegnanti. Studiò matematica al *Trinity College*, Università di Cambridge, dal 1896, divenne membro accademico del suo college nel 1900 e docente di matematica nel 1906. Dopo un primo libro su *L'integrazione di una singola variabile* (1905), pubblicò nel 1908 *Un corso di matematica pura*, che fu ristampato molte volte e tradotto in molte lingue straniere.

Figura 11.1: Hardy (1877–1947)

All'epoca, la riscoperta del lavoro di Mendel aveva sollevato alcuni dubbi. Alcuni biologi si chiedevano perché i caratteri dominanti non diventassero più frequenti di generazione in generazione. Reginald Punnett, che aveva scritto un libro intitolato *Mendelismo* nel 1905, pose la domanda a Hardy, con il quale giocava a cricket a Cambridge. Hardy scrisse la sua soluzione

in un articolo sulle *Proporzioni mendeliane in una popolazione mista*, che fu pubblicato nel 1908. Per semplificare l'analisi, immaginò la situazione di una grande popolazione dove la scelta del partner sessuale sarebbe stata casuale. Inoltre limitò la sua attenzione a due soli fattori (o «alleli») A e a, essendo A dominante e a recessivo. Per la generazione n, sia p_n la frequenza del «genotipo» AA, $2q_n$ quella di Aa e r_n quella di aa. Naturalmente, $p_n + 2q_n + r_n = 1$. Hardy assunse anche che nessuno di questi genotipi portasse ad un eccesso di mortalità o ad una diminuzione della fertilità rispetto agli altri due genotipi. Le frequenze nella generazione $n+1$ possono essere facilmente calcolate notando che un individuo scelto a caso nella generazione n trasmette l'allele A con una probabilità $p_n + q_n$: o il genotipo è AA e l'allele A viene trasmesso con certezza o il genotipo è Aa e l'allele A viene trasmesso con il 50 % di probabilità. Allo stesso modo, l'allele a viene trasmesso con una probabilità $q_n + r_n$. Si può quindi costruire la Tabella 11.1 nello stesso modo della Tabella 8.1.

Tabella 11.1: Calcolo delle frequenze dei genotipi nella generazione $n+1$ dalle frequenze degli alleli dei genitori (le righe sono per la madre, le colonne per il padre).

Allele Frequenza	A $p_n + q_n$	a $q_n + r_n$
A $p_n + q_n$	AA $(p_n + q_n)^2$	Aa $(p_n + q_n)(q_n + r_n)$
a $q_n + r_n$	Aa $(p_n + q_n)(q_n + r_n)$	aa $(q_n + r_n)^2$

Le frequenze dei genotipi AA, Aa e aa nella $n+1$-esima generazione sono rispettivamente p_{n+1}, $2q_{n+1}$ e r_{n+1}. Così Hardy ha trovato che

$$p_{n+1} = (p_n + q_n)^2 \tag{11.1}$$

$$2q_{n+1} = 2(p_n + q_n)(q_n + r_n) \tag{11.2}$$

$$r_{n+1} = (q_n + r_n)^2 . \tag{11.3}$$

Egli studiò poi sotto quali condizioni le frequenze dei genotipi potessero rimanere costanti attraverso le generazioni essendo uguali a p, $2q$ e r. Poiché per definizione $p + 2q + r = 1$, vediamo che le equazioni (11.1)-(11.3) danno tutte la stessa condizione $q^2 = pr$.

Per esempio, la prima equazione dà

$$p = (p+q)^2 = p^2 + 2pq + q^2,$$

che è equivalente a $p(1-p-2q) = q^2$ e infine a $pr = q^2$.

Partendo da condizioni iniziali arbitrarie $(p_0, 2q_0, r_0)$ con $p_0 + 2q_0 + r_0 = 1$, Hardy ha notato che

$$q_1{}^2 = (p_0+q_0)^2(q_0+r_0)^2 = p_1 r_1.$$

Lo stato $(p_1, 2q_1, r_1)$ è quindi già un equilibrio. Quindi $(p_n, 2q_n, r_n)$ rimane uguale a $(p_1, 2q_1, r_1)$ per tutti gli $n \geqslant 1$. Se fissiamo $x = p_0 + q_0$ per la frequenza dell'allele A nella 0-esima generazione , allora $1 - x = q_0 + r_0$ è la frequenza dell'allele a. Usando il sistema (11.1)–(11.3) ancora una volta, otteniamo

$$p_n = x^2, \quad 2q_n = 2x(1-x), \quad r_n = (1-x)^2$$

per tutti gli $n \geqslant 1$ (Figura 11.2).

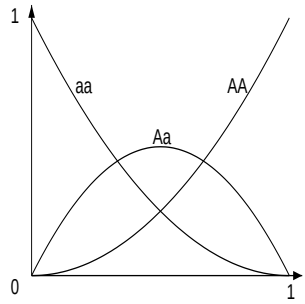

Figura 11.2: Grafici delle funzioni x^2, $2x(1-x)$ e $(1-x)^2$ corrispondenti alle frequenze di equilibrio dei genotipi AA, Aa e aa.

In conclusione, le ipotesi di cui sopra portano alla legge secondo la quale le frequenze dei genotipi AA, Aa e aa rimangono invariate attraverso le generazioni. La teoria di Mendel non porta ad un aumento progressivo della frequenza del carattere dominante come si era pensato inizialmente.

Alcuni anni dopo, Fisher avrebbe insistito su un importante corollario di questa legge: in prima approssimazione (cioè supponendo che le ipotesi del modello siano realistiche), una popolazione mantiene una varianza genetica costante. Questa osservazione risolve uno dei problemi sollevati dalla teoria dell'evoluzione per selezione naturale di Darwin. Infatti, Darwin pensava,

come i suoi contemporanei, che ad ogni generazione le caratteristiche fisiologiche dei figli fossero una sorta di media delle caratteristiche dei due genitori, ogni genitore contribuendo per metà. Questa idea era stata in seguito studiata a fondo con la statistica da Francis Galton e dal suo successore al laboratorio di biometria, Karl Pearson. Se fosse vero, la varianza di queste caratteristiche in una popolazione dovrebbe essere divisa per due ad ogni generazione e ci sarebbe presto una tale omogeneità che la selezione naturale, supposta per spiegare l'evoluzione, sarebbe impossibile. Dovettero trascorrere diversi anni perché questo meccanismo di mediazione fosse respinto, tuttavia, i biometristi difendono il punto di vista di Darwin e sono riluttanti ad ammettere che le leggi di Mendel sono inevitabili per comprendere l'evoluzione.

Dopo questo lavoro, nel 1908, Hardy tornò alla matematica pura. Nella sua autobiografia, *Apologia di un matematico*, affermò addirittura con orgoglio di aver evitato scoperte di qualsiasi utilità pratica. Nel 1910 fu eletto alla *Royal Society*. Nel 1913 scoprì il prodigio indiano Ramanujan e lo invitò a lavorare a Cambridge. Dopo la prima guerra mondiale, divenne professore all'Università di Oxford e continuò una fruttuosa collaborazione con il suo compatriota Littlewood. Tra il 1931 e il 1942 fu di nuovo professore a Cambridge. Pubblicò molti libri, spesso in collaborazione: *Ordini dell'infinito* (1910), *La teoria generale delle serie di Dirichlet* con Marcel Riesz (1915), *Disuguaglianze* con Littlewood e Pólya (1934), *Un'introduzione alla teoria dei numeri* con E. M. Wright (1938), *Ramanujan* (1940), *Serie di Fourier* con Rogosinski (1944) e *Serie divergenti* (1949). Morì a Cambridge nel 1947.

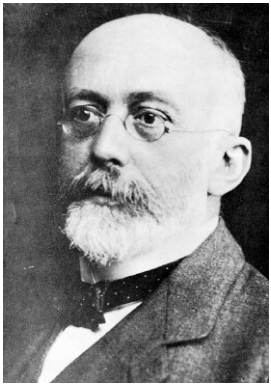

Figura 11.3:
Weinberg (1862–1937)

Diversi decenni dopo, si osservò che la legge di Hardy per le frequenze dei geni era stata anche scoperta quello stesso anno 1908 da un medico tedesco, Wilhelm Weinberg. Weinberg era nato a Stoccarda nel 1862. Dopo aver

studiato a Tübingen e Monaco di Baviera fino al suo dottorato in medicina, aveva lavorato diversi anni negli ospedali di Berlino, Vienna e Francoforte. Si era stabilito nel 1889 a Stoccarda come medico generico e ostetrico. Nonostante fosse molto occupato con il suo lavoro, aveva trovato il tempo di scrivere molti articoli in riviste scientifiche tedesche. Nel 1901 aveva studiato da un punto di vista statistico la frequenza dei gemelli dello stesso sesso. L'articolo del 1908, in cui spiegava la stessa legge trovata da Hardy, era stato pubblicato in una rivista scientifica locale e non era stato notato. Ma a differenza di Hardy, aveva continuato questo studio negli anni seguenti, scoprendo per esempio la generalizzazione al caso in cui ci sono più di due alleli. Aveva anche contribuito al settore della statistica medica. Weinberg morì nel 1937. Dopo la riscoperta del suo articolo del 1908, i genetisti chiamarono la legge di stabilità delle frequenze dei genotipi «legge di Hardy-Weinberg».

Oggi questa legge è spesso usata come segue. Se un raro allele recessivo a non ha influenza sulla sopravvivenza o sulla fertilità e se conosciamo la frequenza x^2 del genotipo aa perché aa produce un particolare fenotipo, allora possiamo calcolare x e stimare la frequenza $2x(1-x) \approx 2x$ del genotipo Aa. Per esempio, se la frequenza di Aa è $1/20000$, allora otteniamo $x \approx 1/140$. Quindi $2x \approx 1/70$ è la frequenza del genotipo Aa. L'allele recessivo a, che potrebbe sembrare molto raro dall'ispezione dei fenotipi, in realtà non è così raro.

Bibliografia

1. Hardy, G.H.: Mendelian proportions in a mixed population. *Science* 28, 49–50 (1908). esp.org
2. Hardy, G.H.: *A Mathematician's Apology*. Cambridge University Press (1940). archive.org
3. Punnett, R.C.: *Mendelism*, 2nd edn. Cambridge University Press (1907). archive.org
4. Stern, C.: The Hardy–Weinberg law. *Science* 97, 137–138 (1943)
5. Stern, C.: Wilhelm Weinberg 1862–1937. *Genetics* 47, 1–5 (1962)
6. Titchmarsh, E.C.: Godfrey Harold Hardy, 1877–1947. *Obit. Not. Fellows R. Soc.* 6, 446–461 (1949)
7. Weinberg, W.: Über den Nachweis der Vererbung beim Menschen. *Jahresh. Wuertt. Ver. vaterl. Natkd.* 64, 369–382 (1908). biodiversitylibrary.org

Capitolo 12

Ross e la malaria (1911)

Nel 1911 il medico britannico Ronald Ross, che aveva già ricevuto il premio Nobel nel 1902 per il suo lavoro sulla malaria, studiò un sistema di equazioni differenziali che modellava la diffusione di questa malattia. Mostrò che la malaria può persistere solo se il numero di zanzare è al di sopra di una certa soglia. Quindi non è necessario uccidere tutte le zanzare per sradicare la malaria: è sufficiente ucciderne solo una certa frazione. Modelli epidemici simili furono sviluppati in seguito da Kermack e McKendrick.

Ronald Ross nacque nel 1857 nel nord dell'India, dove suo padre era un ufficiale dell'esercito britannico. Studiò medicina a Londra ma preferì scrivere poesie e drammi. Dopo aver lavorato per un anno su una nave come chirurgo, riuscì ad entrare nel servizio medico indiano nel 1881. Il suo lavoro medico in India gli lasciò molto tempo libero, durante il quale scrisse opere letterarie e insegnò a se stesso un po' di matematica. In licenza in Inghilterra nel 1888, ottenne un diploma in salute pubblica e studiò batteriologia, una nuova scienza creata pochi anni prima da Pasteur e Koch. Tornato in India, Ross iniziò a studiare la malaria. Durante il suo secondo permesso nel 1894, incontrò a Londra Patrick Manson, uno specialista in medicina tropicale che gli mostrò al microscopio ciò che il medico militare francese Alphonse Laveran aveva notato nel 1880: il sangue dei pazienti con la malaria contiene parassiti. Manson suggerì che i parassiti potevano provenire dalle zanzare perché aveva scoperto lui stesso in Cina il parassita di un'altra malattia tropicale (filariasi) in questi insetti. Tuttavia, credeva che gli esseri umani fossero infettati dal parassita quando bevevano acqua contaminata dalle zanzare. Dal 1895 al 1898, Ross continuò le sue ricerche in India e testò l'idea di Manson. Nel 1897 scoprì nello stomaco di una certa specie di zanzara che non aveva studiato prima (anopheles) alcuni parassiti simili a quelli osservati da Laveran. I suoi superiori lo mandarono a Calcutta durante una stagione in cui i casi di malaria erano rari, decise di studiare la malaria negli uccelli in gabbia. Trovò il parassita nelle ghiandole salivari delle zanzare anofele e riuscì ad infettare sperimentalmente degli uccelli sani facendoli mordere dalle zanzare: questo dimostrò che la malaria si trasmette attraverso le punture delle zanzare e non attraverso l'ingestione di acqua contaminata. Nel 1899 Ross lasciò

il servizio medico indiano per insegnare alla Scuola di medicina tropicale di Liverpool, che era stata creata un anno prima. Fu eletto alla *Royal Society* nel 1901 e ricevette nel 1902 il premio Nobel in fisiologia o medicina per il suo lavoro sulla malaria. Viaggiò in Africa, alle Mauritius e nell'area mediterranea per diffondere la lotta contro le zanzare. Il metodo ebbe successo in Egitto lungo il canale di Suez, lungo il canale di Panama in costruzione, a Cuba e in Malesia. Ebbe meno successo in alcune altre aree. Ross pubblicò un *Relazione sulla prevenzione della malaria a Mauritius* nel 1908 e *La prevenzione della malaria* nel 1910.

Figura 12.1: Ross (1857–1932)

Nonostante avesse provato il ruolo di certe zanzare nella trasmissione della malaria, Ross incontrò scetticismo quando affermò che la malaria poteva essere sradicata semplicemente riducendo il numero delle zanzare. Nella seconda edizione del suo libro *La prevenzione della malaria* pubblicato nel 1911, cercò di costruire modelli matematici della trasmissione della malaria per sostenere la sua affermazione. Uno dei suoi modelli consisteva in un sistema di due equazioni differenziali. Introduciamo le seguenti notazioni:

- N: popolazione umana totale in una data area;
- $I(t)$: numero di esseri umani infettati dalla malaria al tempo t;
- n: popolazione totale di zanzare (supposta costante);
- $i(t)$: numero di zanzare infettate dalla malaria;
- b: frequenza di puntura delle zanzare;

- p (rispettivamente p'): probabilità di trasmissione della malaria da uomo a zanzara (rispettivamente da zanzara a uomo) durante una puntura;
- a: tasso di guarigione degli esseri umani dalla malaria;
- m: mortalità delle zanzare.

Durante un piccolo intervallo di tempo dt, ogni zanzara infetta punge b,dt umani, tra i quali una frazione pari a $\frac{N-I}{N}$ non è ancora infetta. Tenendo conto della probabilità di trasmissione p', ci sono $b\,p'\,i\,\frac{N-I}{N}\,dt$ nuovi esseri umani infetti. Durante lo stesso intervallo di tempo, il numero di esseri umani che guariscono è $a\,I\,dt$. Quindi,

$$\frac{dI}{dt} = b\,p'\,i\,\frac{N-I}{N} - a\,I.$$

Allo stesso modo ogni zanzara non infetta punge $b\,dt$ umani, tra i quali una frazione pari a I/N è già infetta. Tenendo conto della probabilità di trasmissione p, ci sono $b\,p\,(n-i)\,\frac{I}{N},dt$ nuove zanzare infette. Nel frattempo, assumendo che l'infezione non influenzi la mortalità, il numero di zanzare che muoiono è $m\,i\,dt$. Quindi

$$\frac{di}{dt} = b\,p\,(n-i)\,\frac{I}{N} - m\,i\,.$$

Poiché la malaria esiste permanentemente nella maggior parte dei paesi infetti, Ross ha considerato solo gli stati stazionari del suo sistema di due equazioni: il numero di esseri umani infetti $I(t)$ e il numero di zanzare infette $i(t)$ rimangono costanti nel tempo ($dI/dt = 0$ e $di/dt = 0$). In primo luogo c'è sempre lo stato stazionario con $I = 0$ e $i = 0$, che corrisponde all'assenza di malaria. In secondo luogo, Ross ha cercato uno stato stazionario tale che $I > 0$ e $i > 0$ e ha trovato che

$$I = N\,\frac{1 - a\,m\,N/(b^2\,p\,p'\,n)}{1 + a\,N/(b\,p'\,n)}\,, \quad i = n\,\frac{1 - a\,m\,N/(b^2\,p\,p'\,n)}{1 + m/(b\,p)}\,. \tag{12.1}$$

Dividendo le equazioni dello stato stazionario per il prodotto $I \times i$, il problema diventa un sistema lineare di due equazioni con due incognite $1/I$ e $1/i$:

$$\frac{b\,p'}{I} - \frac{a}{i} = \frac{b\,p'}{N}\,, \quad -\frac{m}{I} + \frac{b\,p\,n}{N\,i} = \frac{b\,p}{N}\,,$$

la cui soluzione si ottiene facilmente.

Si può notare che $I > 0$ e $i > 0$ se il numero di zanzare è sopra una soglia critica:

$$n > n^* = \frac{amN}{b^2 p p'}.$$

In questo caso lo stato stazionario corrisponde alla situazione in cui la malattia è endemica, cioè permanentemente presente. Ross ha concluso che se il numero di zanzare n si riduce al di sotto della soglia critica n^*, allora l'unico stato stazionario rimanente è $I = 0$ e $i = 0$, quindi la malaria dovrebbe scomparire. In particolare non è necessario sterminare tutte le zanzare per sradicare la malaria. Questo è precisamente il punto che Ross voleva sottolineare con il suo modello.

Per illustrare la sua teoria, Ross cercò dei valori numerici ragionevoli per i parametri del suo modello. Assunse che

- la mortalità delle zanzare è tale che solo un terzo di esse è ancora vivo dopo dieci giorni; quindi $\mathrm{e}^{-10m} = \frac{1}{3}$ e $m = (\log 3)/10$ al giorno;
- metà delle persone sono ancora infette dopo tre mesi; quindi $\mathrm{e}^{-90a} = 1/2$ e $a = (\log 2)/90$ al giorno;
- una zanzara su otto punge ogni giorno; quindi $\mathrm{e}^{-b} = 1 - 1/8$ e $b = \log(8/7)$ al giorno;
- Le zanzare infette di solito non sono contagiose durante i primi dieci giorni dopo la loro infezione perché i parassiti devono passare attraverso diversi stadi di trasformazione. Poiché un terzo delle zanzare può sopravvivere dieci giorni, Ross ha assunto che ci sia anche circa un terzo di tutte le zanzare infette che sono infettive: $p' = 1/3$;
- $p = 1/4$.

Ross poteva quindi calcolare con la formula (12.1) la frazione infetta I/N nella popolazione umana in funzione del rapporto n/N tra la zanzara e la popolazione umana. Ha mostrato i suoi risultati in una tabella che è equivalente alla figura 12.2.

La forma della curva mostra che la frazione di esseri umani infetti è più alta del 50 % già se il rapporto n/N è appena sopra il valore critico n^*/N. Ma questa frazione non cambia molto quando il rapporto n/N aumenta ulteriormente. Questo spiega perché la correlazione tra il numero di zanzare e la presenza della malaria non era mai stata notata prima. Ross notò, tuttavia, che il valore numerico della soglia n^*/N era molto sensibile a piccoli

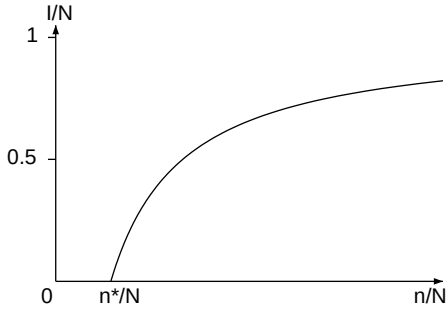

Figura 12.2: Frazione I/N di esseri umani infetti in funzione del rapporto n/N tra la zanzara e la popolazione umana.

cambiamenti nel tasso di puntura b, ma che questo non cambiava la forma generale della curva in figura 12.2. La sua spiegazione qualitativa è più importante dei risultati quantitativi, che soffrono dell'incertezza che circonda i valori numerici dei parametri.

Per interpretare la soglia critica n^* scoperta da Ross[1], consideriamo un uomo infetto introdotto in una popolazione umana ed una di zanzare entrambe libere dalla malaria. Questo essere umano rimane infetto in media per un periodo di tempo pari a $1/a$. Lui o lei riceve bn/N morsi per unità di tempo quindi in media riceve in totale $bn/(aN)$ morsi mentre è infetto. Quindi lui o lei infetta in media $bpn/(aN)$ zanzare. Ognuna di queste zanzare infette vive in media per un periodo di tempo pari a $1/m$, punge b/m umani e infetta bp'/m umani. In totale, dopo la trasmissione dal primo umano infetto alle zanzare e da queste zanzare ad altri umani, il numero medio di nuovi umani infetti è il prodotto dei due risultati precedenti, cioè

$$\mathscr{R}_0 = \frac{b^2\, p\, p'\, n}{a\, m\, N}\ . \tag{12.2}$$

Questo $\mathscr{R}_0$ è il numero di casi umani secondari dovuti a un caso umano primario. Quindi il processo di infezione che avviene continuamente nel tempo può essere considerato anche attraverso generazioni successive. La malaria può «invadere» la popolazione solo se $\mathscr{R}_0 > 1$. Questa condizione è precisamente equivalente a $n > n^*$. In conclusione, Ross si è dichiarato più in generale a favore della modellazione matematica in epidemiologia:

[1] Questa interpretazione è stata enfatizzata solo molto tempo dopo il lavoro di Ross.

> «In effetti tutta l'epidemiologia, interessata com'è alla variazione della malattia da tempo a tempo o da luogo a luogo, se dev'essere considerata scientificamente, dev'essere quantificata matematicamente, per quante variabili siano implicate. Dire che una malattia dipende da certi fattori non significa dire molto, fino a quando non possiamo anche formulare una stima di quanto ampiamente ogni fattore influenzi l'intero risultato. E il metodo di modellazione matematica non è altro che l'applicazione di un attento ragionamento ai problemi in questione.»

Ross fu nominato cavaliere nel 1911. Si trasferì a Londra e divenne consulente dell'esercito britannico durante la prima guerra mondiale. Nel 1923 pubblicò la sua autobiografia, *Memorie con un resoconto completo del grande problema della malaria e della sua soluzione*. Nel 1926 fu inaugurato il Istituto Ross di malattie tropicali (ora parte della Scuola di Igiene e Medicina Tropicale di Londra), di cui divenne direttore. Ross morì a Londra nel 1932.

Bibliografia

1. G.H.F.N.: Sir Ronald Ross, 1857-1932. *Obit. Not. Fellows Roy. Soc.* 1, 108–115 (1933)
2. Ross, R.: *The Prevention of Malaria*, 2nd edn. John Murray, London (1911) archive.org
3. Ross, R.: *Memoirs with a Full Account of the Great Malaria Problem and its Solution*. John Murray, London (1923) archive.org
4. Rowland, J.: *The Mosquito Man, The Story of Sir Ronald Ross*. Roy Publishers, New York (1958)

Capitolo 13

Lotka, Volterra e il sistema preda-predatore (1920–1926)

Nel 1920 Alfred Lotka studiò un modello preda-predatore e dimostrò che le popolazioni potevano oscillare in modo permanente. Sviluppò questo studio nel suo libro del 1925 *Elementi di biologia fisica*. Nel 1926 il matematico italiano Vito Volterra si interessò allo stesso modello per rispondere a una domanda posta dal biologo Umberto d'Ancona: perché c'erano più pesci predatori catturati dai pescatori nel mare Adriatico durante la prima guerra mondiale, quando l'attivitá di pesca si era ridotta?

Nel 1920 Lotka pubblicò un articolo intitolato *Nota analitica su certe relazioni ritmiche nei sistemi organici*. Già da alcuni anni si era interessato ad alcune reazioni chimiche che mostravano strane oscillazioni transitorie negli esperimenti di laboratorio. Lo scopo del suo articolo era quello di suggerire che anche un sistema di due specie biologiche poteva oscillare in modo permanente. L'esempio che consideró fu quello di una popolazione di erbivori che si nutre di piante. In analogia con le equazioni usate in cinetica chimica, indicando con $x(t)$ la massa totale delle piante e con $y(t)$ la massa totale degli erbivori al tempo t, Lotka usó come modello il seguente sistema di equazioni differenziali

$$\frac{dx}{dt} = ax - bxy, \tag{13.1}$$

$$\frac{dy}{dt} = -cy + dxy, \tag{13.2}$$

dove i parametri a, b, c e d sono tutti positivi. Il parametro a è il tasso di crescita delle piante quando non ci sono erbivori, mentre c è il tasso di mortalitá della popolazione di erbivori quando non ci sono piante. I termini $-bxy$ e dxy esprimono che più animali e piante ci sono, maggiore è il trasferimento di massa dalle piante verso gli animali (il trasferimento include una certa perdita di massa quindi $d \leqslant b$). Imponendo che $dx/dt = 0$ e $dy/dt = 0$, Lotka notò che ci sono due stati stazionari:

- $(x = 0, y = 0)$, la popolazione di erbivori è estinta e non ci sono più piante;

- $(x = c/d, y = a/b)$, erbivori e piante coesistono.

Scrisse, ma senza provarlo, che se al tempo $t = 0$, $(x(0), y(0))$ non è uno di questi due stati stazionari, allora le funzioni $x(t)$ e $y(t)$ oscillano periodicamente: esiste un numero $T > 0$ tale che $x(t+T) = x(t)$ e $y(t+T) = y(t)$ per tutti $t > 0$ (figura 13.1)[1]. Se per esempio le piante sono molto abbondanti, allora la popolazione di erbivori aumenterà, causando una diminuzione della massa totale delle piante. Quando questa massa diventa insufficiente per nutrire gli erbivori, alcuni animali moriranno di fame e la massa totale delle piante ricomincerà a crescere fino a raggiungere il suo valore iniziale, e, contemporaneamente, anche gli erbivori ritorneranno al loro valore iniziale. Il fenomeno si ripeterà poi periodicamente.

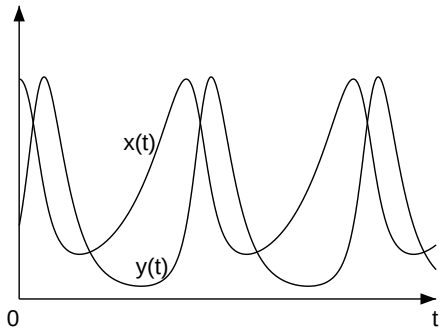

Figura 13.1: Oscillazioni della massa totale delle piante $x(t)$ e della massa totale degli erbivori $y(t)$ in funzione del tempo.

Lotka studiò ulteriormente il modello in un secondo articolo pubblicato nel 1920 intitolato *Oscillazioni smorzate derivate dalla legge dell'azione di massa*, lavoro in cui perché il sistema poteva oscillare in modo periodico. Ciò deriva dal fatto che il punto $(x(t), y(t))$ deve stare su una traiettoria chiusa nel piano con x sull'asse orizzontale e y sull'asse verticale; più precisamente, nel quadrante dove $x \geqslant 0$ e $y \geqslant 0$ (figura 13.2).

Infatti, dividendo l'equazione (13.1) per l'equazione (13.2), otteniamo

[1] Il periodo T dipende dalle condizioni iniziali, ma Lotka si é accorto di ció solo nel 1925.

dopo qualche riordinamento dei termini

$$\left(-\frac{c}{x}+d\right)\frac{dx}{dt}=\left(\frac{a}{y}-b\right)\frac{dy}{dt}\,.$$

La sua integrazione dà $d\,x(t)-c\,\log x(t)=a\,\log y(t)-b\,y(t)+K$, dove K è una costante che dipende solo dalla condizione iniziale. Quindi, il punto $(x(t),y(t))$ rimane sulla curva $d\,x-c\,\log x=a\,\log y-b\,y+K$, che è una curva chiusa (figura 13.2).

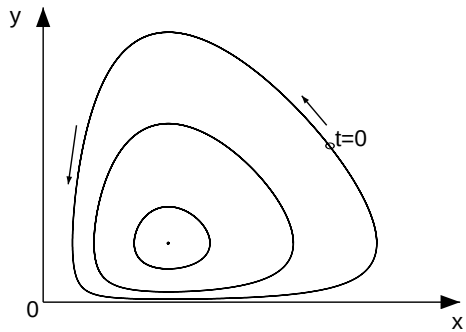

Figura 13.2: Diagramma con la massa totale delle piante $x(t)$ sull'asse orizzontale e la massa totale degli erbivori $y(t)$ sull'asse verticale. Le tre curve chiuse intorno allo stato stazionario corrispondono a diverse condizioni iniziali.

La traiettoria di $(x(t),y(t))$ gira intorno allo stato stazionario $(c/d,a/b)$ in senso antiorario come si può facilmente vedere studiando il segno di dx/dt e di dy/dt. Vicino allo stato stazionario, il sistema mostra piccole oscillazioni con un periodo uguale a $2\pi/\sqrt{a\,c}$.

Infatti, poniamo $x=\frac{c}{d}+x^*$ e $y=\frac{a}{b}+y^*$ dove $x^*\ll\frac{c}{d}$ e $y^*\ll\frac{a}{b}$. Allora

$$\frac{dx^*}{dt}=-b\,y^*\left(\frac{c}{d}+x^*\right)\approx-\frac{b\,c}{d}\,y^*\,,$$
$$\frac{dy^*}{dt}=d\,x^*\left(\frac{a}{b}+y^*\right)\approx\frac{a\,d}{b}\,x^*\,.$$

Da queste due equazioni, otteniamo

$$\frac{d^2x^*}{dt^2} \approx -acx^*, \quad \frac{d^2y^*}{dt^2} \approx -acy^*.$$

Queste equazioni sono le stesse delle oscillazioni del pendolo semplice in fisica il cui periodo è $2\pi/\sqrt{ac}$.

Raymond Pearl, che aveva comunicato il primo articolo del 1920 ai *Proceedings of the National Academy of Sciences*, aiutò Lotka ad ottenere una borsa di studio di due anni dalla Johns Hopkins University per scrivere un libro intitolato *Elementi di biologia fisica*. Il libro fu pubblicato nel 1925. La sezione che riassumeva il lavoro del 1920 menzionava anche che i sistemi di due specie, una specie ospite e una parassita o una preda e una specie predatrice, potevano essere descritti dallo stesso modello (13.1)–(13.2). Purtroppo il libro di Lotka non attirò molta attenzione quando fu pubblicato. Tuttavia, il famoso matematico Volterra riscoprì indipendentemente lo stesso modello poco dopo, mentre studiava un problema di pesca.

Vito Volterra nacque nel ghetto ebraico di Ancona nel 1860, poco prima dell'unità d'Italia, quando la città apparteneva ancora allo Stato Pontificio. Era figlio unico, e suo padre, un commerciante di stoffe, morì quando Vito aveva due anni e lasciò la famiglia senza soldi. Ottimo studente al liceo, Volterra riuscì a continuare gli studi nonostante la povertà, prima all'Università di Firenze e poi alla Scuola Normale Superiore di Pisa. Nel 1882 ottenne il dottorato in fisica e l'anno successivo divenne professore di meccanica all'Università di Pisa. Entrò all'Università di Torino nel 1892 e si trasferì alla cattedra di fisica matematica all'Università La Sapienza di Roma nel 1900. Divenne senatore nel 1905. Molte delle lezioni che tenne a Roma o in università straniere furono pubblicate in forma di libro: *Tre lezioni su alcuni recenti progressi nella fisica matematica* (Università Clark, 1909), *Lezioni sulle equazioni integrali e sulle equazioni integro-differenziali* (Roma, 1910), *Lezioni sulle funzioni di linea* (Parigi, 1912), *La teoria delle funzioni permutabili* (Princeton, 1912). Servì come ufficiale nell'esercito italiano durante la prima guerra mondiale e guidò l'Ufficio Invenzioni presso il Sottosegretariato per le Armi e Munizioni del Ministero della Guerra. Dopo la guerra, partecipò attivamente alla fondazione dell'Unione Matematica Italiana (1922) e del Consiglio Nazionale delle Ricerche (1923), diventandone il primo presidente. Divenne anche presidente della Commissione internazionale per lo studio scientifico del Mar Mediterraneo (1923) e presidente dell'Accademia dei Lincei (1924). Un'altra monografia, scritta in collaborazione con J. Pérès,

Lezioni sulla composizione e le funzioni permutabili, fu pubblicata nel 1924.

Figura 13.3:
Volterra (1860–1940) ricevette un dottorato honoris causa dall'Università di Cambridge nel 1900.

Nel 1925, all'età di 65 anni, Volterra si interessò ad uno studio dello zoologo Umberto D'Ancona, che sarebbe poi diventato suo genero, sulla proporzione di pesci cartilaginei (come squali e razze) nel totale del pescato durante gli anni 1905–1923 in tre porti del mare Adriatico: Trieste, Fiume[2] e Venezia. D'Ancona aveva notato che la proporzione di questi pesci era aumentata durante la prima guerra mondiale, quando l'attivitá di pesca era stata ridotta (Tabella 13.1).

Essendo i pesci cartilaginei predatori di pesci più piccoli, sembrava che una diminuzione dell'attivitá di pesca favorisse le specie predatrici. Volterra, che non conosceva il lavoro di Lotka, spiegò questa osservazione utilizzando lo stesso modello

$$\frac{dx}{dt} = ax - bxy\,, \quad \frac{dy}{dt} = -cy + dxy\,,$$

dove $x(t)$ rappresenta il numero di prede e $y(t)$ il numero di predatori. Notò, come Lotka, che questo sistema può oscillare in modo periodico con un periodo T che dipende dalla condizione iniziale (x_0, y_0). Notò anche che

$$\frac{d}{dt}\log x = a - by, \quad \frac{d}{dt}\log y = -c + dx.$$

Integrando su un periodo T (in modo che $x(0) = x(T)$ e $y(0) = y(T)$), ha ottenuto

$$\frac{1}{T}\int_0^T y(t)\,dt = \frac{a}{b}\,, \quad \frac{1}{T}\int_0^T x(t)\,dt = \frac{c}{d}\,.$$

[2] Ora Rijeka in Croazia.

Tabella 13.1: Percentuale di pesci cartilaginei nella pesca di Trieste, Fiume e Venezia prima, durante e dopo la prima guerra mondiale.

anno	1910	1911	1912	1913	1914	1915	1916
Trieste	5,7	8,8	9,5	15,7	14,6	7,6	16,2
Fiume	-	-	-	-	11,9	21,4	22,1
Venezia	21,8	-	-	-	-	-	-

anno	1917	1918	1919	1920	1921	1922	1923
Trieste	15,4	-	19,9	15,8	13,3	10,7	10,2
Fiume	21,2	36,4	27,3	16,0	15,9	14,8	10,7
Venezia	-	-	30,9	25,3	25,9	25,8	26,6

Quindi la media in un periodo sia del numero di prede che del numero di predatori è indipendente dalle condizioni iniziali. Inoltre, se l'attivitá di pesca diminuisce, il tasso di crescita a delle prede aumenta mentre il tasso di mortalitá c dei predatori diminuisce. Quindi la media di $x(t)$ diminuisce e la media di $y(t)$ aumenta: la proporzione di predatori aumenta. Questo è precisamente ciò che è stato osservato nelle statistiche di pesca del mare Adriatico.

Volterra pubblicò il suo articolo prima in italiano nel 1926. Un riassunto in inglese apparve qualche mese dopo su *Nature*. Lotka informò Volterra e altri scienziati della priorità del suo studio dei sistemi preda-predatore, ma il suo articolo del 1920 e il suo libro del 1925 non saranno sempre menzionati. Lotka lavorava già allora per una compagnia di assicurazioni, quindi il suo lavoro si concentrò sulla demografia umana. Volterra continuò a lavorare sulle varianti del sistema preda-predatore per un decennio. Diede una serie di lezioni nel 1928-1929 al neonato Istituto Henri Poincaré di Parigi. Le note di queste lezioni furono pubblicate nel 1931 con il titolo *Lezioni sulla teoria matematica della lotta per la vita*. Nel 1935 Volterra pubblicò in collaborazione con Umberto D'Ancona un altro libro su *Associazioni biologiche da un punto di vista matematico*.

Anche se il modello preda-predatore sembra spiegare correttamente i dati della pesca, il dibattito sul realismo dei modelli semplificati in ecologia era appena iniziato ed è ancora oggetto di dispute scientifiche. Al giorno d'oggi, il modello preda-predatore è conosciuto anche come modello Lotka-Volterra ed è uno dei più comunemente citati in ecologia.

Nel 1931 Volterra rifiutò di giurare fedeltà a Mussolini. Perse la sua cattedra all'università di Roma e fu escluso dalle accademie scientifiche italiane, di cui era uno dei membri più famosi. Da allora in poi rimase principalmente fuori dall'Italia, viaggiando per l'Europa e tenendo conferenze. Pubblicò con J. Pérès il primo volume di una *Teoria generale dei funzionali* (1936) e un libro con B. Hostinský sulle *Operazioni lineari infinitesimali* (1938). Morì a Roma nel 1940.

Bibliografia

1. Goodstein, J.R.: *The Volterra Chronicles, The Life and Times of an Extraordinary Mathematician 1860-1940*. American Mathematical Society (2007)
2. Guerraggio, A., Nastasi, P.: *Italian Mathematics between the Two World Wars*. Birkhäuser, Basel (2005)
3. Israel, G., Gasca, A.M.: *The Biology of Numbers – The Correspondence of Vito Volterra on Mathematical Biology*. Birkhäuser, Basel (2002)
4. Kingsland, S.E.: *Modeling Nature, Episodes in the History of Population Ecology*, 2nd edn. University of Chicago Press (1995)
5. Lotka, A.J.: Analytical note on certain rhythmic relations in organic systems. *Proc. Natl. Acad. Sci.* 6, 410–415 (1920) pnas.org
6. Lotka, A.J.: Undamped oscillations derived from the law of mass action. *J. Amer. Chem. Soc.* 42, 1595–1599 (1920) archive.org
7. Lotka, A.J.: *Elements of Physical Biology*. Williams & Wilkins, Baltimore (1925) archive.org
8. Volterra, V.: Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Mem. Accad. Lincei* 6, 31–113 (1926) → *Opere matematiche*, vol. 5, Accademia nazionale dei Lincei, Roma (1962) liberliber.it
9. Volterra, V.: Fluctuations in the abundance of a species considered mathematically. *Nature* 118, 558–560 (1926). → L.A. Real, J.H. Brown (eds.) *Foundations of Ecology*, 283–285. University of Chicago Press (1991)
10. Volterra, V.: *Leçons sur la théorie mathématique de la lutte pour la Vie*. Gauthier-Villars, Paris (1931)
11. Volterra, V., D'Ancona, U.: *Les Associations biologiques au point de vue mathématique*. Hermann, Paris (1935)
12. Whittaker, E.T.: Vito Volterra 1860–1940. *Obit. Not. Fellows R. Soc.* 3, 690–729 (1941)

Capitolo 14

Fisher e la selezione naturale (1922)

Nel 1922 il biologo matematico britannico Ronald Fisher pubblicò un articolo molto influente sulla genetica delle popolazioni. Questo capitolo considera solo una parte dell'articolo, che si concentra su una variante del modello Hardy-Weinberg comprendente la selezione naturale. Fisher ha dimostrato che se l'eterozigote è prevalente, allora entrambi gli alleli possono coesistere. Se uno dei due omozigoti è prevalente, allora l'altro allele scompare. Il problema di fondo è quello di spiegare perché alcuni geni possono avere diversi alleli.

Ronald Aylmer Fisher nacque a Londra nel 1890, ultimo di sei figli. Suo padre era un banditore d'asta, ma in seguito dichiarò bancarotta. Fisher studiò matematica e fisica al *Gonville and Caius College* dell'Università di Cambridge tra il 1909 e il 1913. La genetica si stava sviluppando rapidamente in quel periodo. A partire dal 1911, Fisher partecipò alle riunioni della *Eugenics Society* iniziata da Galton. Cominciò a concentrarsi su problemi statistici legati al lavoro di Galton e Mendel. Dopo aver terminato gli studi universitari, trascorse un'estate lavorando in una fattoria in Canada e poi lavorò per la *Mercantile and General Investment Company* nella Città di Londra. A causa della sua estrema miopia, non poté partecipare alla prima guerra mondiale nonostante si fosse offerto volontario. Trascorse questi anni insegnando nelle scuole superiori. Durante il suo tempo libero, si prese cura di una fattoria e continuò le sue ricerche. Ottenne nuovi importanti risultati che collegavano i coefficienti di correlazione con la genetica mendeliana. Nel 1919 iniziò a lavorare come statistico alla Stazione sperimentale Rothamsted, che si occupava di ricerche in agricoltura.

Nel 1922 Fisher pubblicò un articolo intitolato *Sul rapporto di dominanza*. Tra molte altre importanti nuove idee, questo articolo considerava un modello matematico che combinava le leggi di Mendel e l'idea della selezione naturale enfatizzata da Darwin per la teoria dell'evoluzione. Fisher considerò la stessa situazione di Hardy con due alleli, A e a, insieme all'ipotesi di accoppiamento casuale, ma ipotizzò che gli individui con i genotipi AA, Aa e aa avessero una diversa mortalità prima di raggiungere l'età adulta, in accordo con la selezione naturale. Fissando p_n, $2q_n$ e r_n per le frequenze dei tre

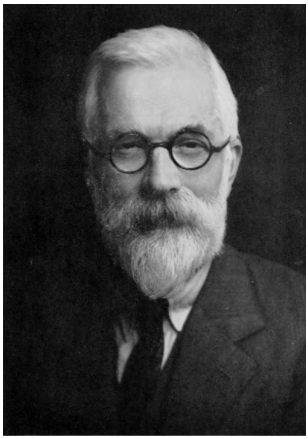

Figura 14.1:
Fisher (1890–1962)

genotipi tra gli individui adulti nella generazione n, ci sono rispettivamente

$$(p_n+q_n)^2, \quad 2(p_n+q_n)(q_n+r_n) \quad \text{e} \quad (q_n+r_n)^2$$

nuovi nati nella $n+1$ - sima generazione con questi genotipi. Siano u, v e w le rispettive probabilità di sopravvivenza dalla nascita all'età adulta. Allora le frequenze dei genotipi tra gli individui adulti della $n+1$ - esima generazione sono p_{n+1}, $2q_{n+1}$ e r_{n+1} con:

$$p_{n+1} = \frac{u(p_n+q_n)^2}{d_n} \tag{14.1}$$

$$q_{n+1} = \frac{v(p_n+q_n)(q_n+r_n)}{d_n} \tag{14.2}$$

$$r_{n+1} = \frac{w(q_n+r_n)^2}{d_n}, \tag{14.3}$$

dove abbiamo impostato per comodità:

$$d_n = u(p_n+q_n)^2 + 2v(p_n+q_n)(q_n+r_n) + w(q_n+r_n)^2.$$

Ricordando che $p_n+2q_n+r_n=1$, vediamo che quando $u=v=w$ (cioè quando non c'è selezione naturale), il sistema (14.1)-(14.3) si riduce al sistema (11.1)-(11.3) considerato da Hardy.

Sia

$$x_n = p_n+q_n$$

la frequenza dell'allele A tra gli individui adulti della n - sima generazione. Allora $q_n + r_n = 1 - x_n$ è la frequenza dell'allele a. Aggiungendo (14.1) e (14.2), otteniamo:

$$x_{n+1} = \frac{u{x_n}^2 + v x_n(1-x_n)}{u{x_n}^2 + 2v x_n(1-x_n) + w(1-x_n)^2}\,.$$

Questa equazione può essere riscritta nella forma:

$$x_{n+1} - x_n = x_n\,(1-x_n)\,\frac{(v-w)(1-x_n) + (u-v)x_n}{u{x_n}^2 + 2v x_n(1-x_n) + w(1-x_n)^2}\,. \tag{14.4}$$

Ci sono sempre almeno due stati stazionari in cui la frequenza x_n rimane costante attraverso le generazioni: $x = 0$ (la popolazione consiste interamente di omozigoti aa) e $x = 1$ (la popolazione consiste interamente di omozigoti AA).

Usando l'equazione (14.4), si può mostrare che se l'omozigote AA ha una migliore possibilità di sopravvivenza rispetto agli altri due genotipi ($u > v$ e $u > w$), allora l'allele a scomparirà progressivamente dalla popolazione. Questo caso non dovrebbe essere molto comune in natura se sappiamo che entrambi gli alleli coesistono. Se invece l'eterozigote Aa ha un vantaggio selettivo sugli omozigoti AA e aa ($v > u$ e $v > w$), allora i tre genotipi possono coesistere nella popolazione. Questo è il caso più comune e può spiegare la «vigoria» degli ibridi notato dagli agricoltori.

Infatti, lo stato stazionario $x = 1$ è stabile quando $u > v$ perché $x_{n+1} - x_n \approx (1-x_n)\,(u-v)/u$ quando x_n è vicino a 1. La popolazione tende a questo stato stazionario. Lo stato stazionario $x = 1$ è instabile quando $u < v$, nel qual caso c'è un terzo stato stazionario

$$x^* = \frac{v-w}{2v-u-w}$$

con $0 < x^* < 1$. Inoltre possiamo verificare che questo è stabile. Lo stato stazionario x^* corrisponde a una miscela tra i tre genotipi.

Quindi, combinando semplicemente le leggi di Mendel e un'ipotesi di selezione naturale (qui, differenti probabilità di sopravvivenza per i tre genotipi), possiamo spiegare le due situazioni di coesistenza o scomparsa dei genotipi. Dopo Fisher, questo modello è stato sviluppato anche da J.B.S. Haldane (vedi capitolo 17) e da Sewall Wright (vedi capitolo 19).

Anticipando il capitolo 20, si nota che se A è completamente dominante e l'omozigote aa è svantaggiato rispetto agli altri due genotipi, essendo i numeri $u : v : w$ in un rapporto $1 : 1 : 1-\varepsilon$, allora l'equazione (14.4) diventa:

$$x_{n+1} - x_n = \frac{\varepsilon x_n (1-x_n)^2}{1-\varepsilon(1-x_n)^2} \approx \varepsilon x_n (1-x_n)^2 \tag{14.5}$$

per $\varepsilon \ll 1$. Se la sopravvivenza dell'eterozigote Aa si trova a metà strada tra quella dei due omozigoti, allora i numeri $u : v : w$ sono in un rapporto $1 : 1-\varepsilon/2 : 1-\varepsilon$, si ha:

$$x_{n+1} - x_n = \frac{\varepsilon}{2} \frac{x_n(1-x_n)}{1-\varepsilon(1-x_n)} \approx \frac{\varepsilon}{2} x_n(1-x_n) \tag{14.6}$$

quando $\varepsilon \ll 1$.

A Rothamsted Fisher ha analizzato dati a lungo termine riguardanti i raccolti e la meteorologia. Ma diede anche grandi contributi alla metodologia statistica. Nel 1925 pubblicò un libro intitolato *Metodi statistici per i ricercatori*, che ebbe grande successo e fu ristampato molte volte. Divenne membro della *Royal Society* nel 1929. Nel 1930 Fisher pubblicò un libro su *La teoria genetica della selezione naturale*, una pietra miliare nella storia della genetica delle popolazioni. Divenne professore di eugenetica all'*University College* di Londra nel 1933, succedendo a Karl Pearson al Laboratorio Galton. Nel 1943 si trasferì all'Università di Cambridge ricoprendo la cattedra di genetica, questa volta succedendo a R.C. Punnett (vedi capitolo 11). Ha anche pubblicato diversi libri: *Il disegno degli esperimenti* (1935), *La teoria della consanguineità* (1949) e *Metodi statistici e inferenza scientifica* (1956). Nominato cavaliere nel 1952, si stabilì in Australia dopo essere andato in pensione nel 1959 e morì ad Adelaide nel 1962. Nel capitolo 20 torneremo su un'altra parte del suo lavoro .

Bibliografia

1. Fisher Box, J.: R.A. Fisher, *The Life of a Scientist*. John Wiley & Sons, New York (1978)
2. Fisher, R.A.: On the dominance ratio. *Proc. R. Soc. Edinb.* 42, 321–341 (1922) library.adelaide.edu.au
3. Fisher, R.A.: *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford (1930) archive.org
4. Yates, F., Mather, K.: *Ronald Aylmer Fisher 1890–1962. Biog. Mem. Fellows R. Soc.* 9, 91–120 (1963)

Capitolo 15

Yule e l'evoluzione (1924)

Nel 1924 lo statistico britannico Yule studiò un modello di evoluzione in cui le specie possono produrre nuove specie con piccole mutazioni e i generi possono produrre nuovi generi con grandi mutazioni. Il suo scopo era quello di spiegare la distribuzione del numero di specie all'interno dei generi, la maggior parte dei generi contenenti solo una specie e pochi generi contenenti un gran numero di specie. Il «processo di nascita» stocastico introdotto da Yule nel suo modello è ancora uno strumento fondamentale nello studio degli alberi filogenetici e in molte altre aree.

George Udny Yule nacque in Scozia nel 1871, suo padre ricopriva una posizione di alto livello nell'amministrazione britannica in India. All'età di 16 anni Yule iniziò a studiare all'*University College* di Londra per diventare ingegnere. Nel 1892 cambiò il suo orientamento e trascorse un anno facendo ricerca a Bonn sotto la supervisione del fisico Heinrich Hertz, che aveva dimostrato l'esistenza delle onde elettromagnetiche pochi anni prima. Quando Yule tornò in Inghilterra, Karl Pearson gli offrì un posto di assistente professore di matematica applicata all'*University College*. Yule, seguendo Pearson, iniziò a concentrarsi sulla statistica. Nel 1911 pubblicò *Un'introduzione alla teoria della statistica*, che fu ristampato 14 volte. L'anno seguente si trasferì all'Università di Cambridge. Il suo lavoro di ricerca riguardava gli aspetti teorici della statistica ma anche le applicazioni all'agricoltura e all'epidemiologia. Divenne un membro della *Royal Society* nel 1922.

Nel 1924 Yule pubblicò un articolo intitolato *Una teoria matematica dell'evoluzione basata sulle conclusioni del dottor J. C. Willis*. Willis era un collega della *Royal Society* che aveva pubblicato nel 1922 un libro intitolato *Età e area, uno studio sulla distribuzione geografica e l'origine delle specie*. Aveva studiato la distribuzione delle specie tra i diversi generi nella classificazione delle piante e degli animali. I dati che aveva compilato mostravano che la maggior parte dei generi conteneva solo una specie, che sempre meno generi contenevano un numero maggiore di specie e che c'erano ancora pochi generi che contenevano un gran numero di specie. La tabella 15.1 mostra i dati riguardanti serpenti, lucertole e due famiglie di coleotteri (i Chrysomelidae e i Cerambycinae).

Figura 15.1:
Yule (1871-1951)

Tabella 15.1: Dati compilati da Willis.

Numero di specie	Numero di generi			
	Chrysomelidae	Cerambycinae	Serpenti	Lucertole
1	215	469	131	105
2	90	152	35	44
3	38	82	28	23
4	35	61	17	14
5	21	33	16	12
6	16	36	9	7
7	15	18	8	6
8	14	17	8	4
9	5	14	9	5
10	15	11	4	5
11-20	58	74	10	17
21-30	32	21	12	9
31-40	13	15	3	3
41-50	14	8	1	2
51-60	5	4	0	0
61-70	8	3	0	1
71-80	7	0	1	0
81-90	7	1	0	0
91-100	3	1	1	0
101-	16	4	0	2
totale	627	1 024	293	259

Le 1 580 specie di lucertole conosciute all'epoca erano state classificate in 259 generi, 105 generi contenenti solo una specie, 44 solo due specie, 23 solo tre specie, ecc. e due generi contenenti più di cento specie. Per altre famiglie di animali e piante, la distribuzione dei generi secondo il numero di specie che contengono aveva una forma molto simile.

Yule suggerì a Willis di provare a tracciare i suoi dati in un grafico con scale logaritmiche. Questo ha dato un risultato sorprendente (Figura 15.2): il logaritmo del numero Q_n di generi contenenti n specie diminuisce più o meno linearmente con $\log(n)$. In altre parole, ci sono costanti $\alpha > 0$ e $\beta > 0$ tali che $Q_n \approx \alpha n^{-\beta}$. La distribuzione segue una «legge del potere». Nel suo articolo del 1924, Yule cercò un modello matematico dell'evoluzione che potesse spiegare tale distribuzione statistica.

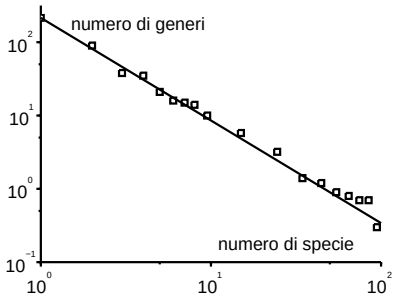

Figura 15.2: Il numero di generi in funzione del numero di specie che contengono, con scale logaritmiche decimali. Dati per i Chrysomelidae. Per smussare le fluttuazioni quando n (il numero di specie) è grande, i generi sono stati contati per intervalli di valori di n come nella Tabella 15.1. Il numero medio di generi per un singolo valore di n può quindi essere inferiore a 1.

A tal fine immaginò prima un modello stocastico[1] a tempo continuo per la crescita del numero di specie all'interno di un genere (Figura 15.3a). Partendo da una sola specie al tempo $t = 0$, assunse che la probabilità per una specie di dare vita per mutazione a una nuova specie dello stesso genere durante un «piccolo» intervallo di tempo dt (sulla scala temporale dell'evoluzione) fosse uguale a $r\,dt$ con $r > 0$.

Sia $p_n(t)$ la probabilità che ci siano n specie al tempo t (n è un numero intero ma t è un numero reale). Per calcolare $p_n(t + dt)$, Yule ha considerato diversi casi:

- se ci sono $n - 1$ specie al tempo t, ogni specie ha una probabilità $r\,dt$ di generare una nuova specie tra t e $t + dt$; nel limite $dt \to 0$, ci saranno n specie al tempo $t + dt$ con una probabilità $(n - 1)\,r\,dt$;

[1] McKendrick (vedi capitolo 16) aveva già iniziato a studiare tali modelli nella dinamica delle popolazioni in un articolo pubblicato nel 1914.

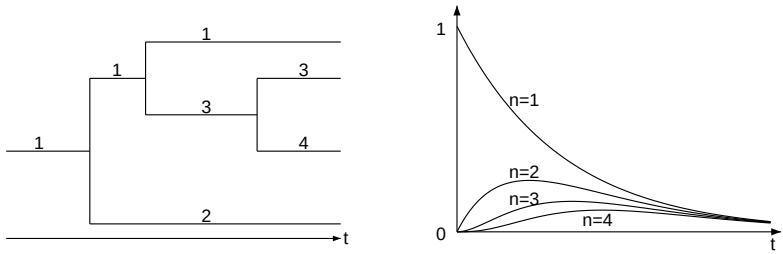

Figura 15.3: (a) Una simulazione dell'evoluzione del numero di specie all'interno di un genere. La specie 1 genera le specie 2 e 3. La specie 3 genera la specie 4. (b) La probabilità $p_n(t)$ che ci siano n specie dello stesso genere al tempo t, per $1 \leqslant n \leqslant 4$.

- se ci sono n specie al tempo t, ci saranno $n+1$ specie al tempo $t+dt$ con una probabilità $n\,r\,dt$.

Così $p_n(t)$ è dato dal seguente sistema di equazioni differenziali

$$\frac{dp_1}{dt} = -r\,p_1\,, \tag{15.1}$$

$$\frac{dp_n}{dt} = (n-1)\,r\,p_{n-1} - n\,r\,p_n \tag{15.2}$$

per tutti gli $n \geqslant 2$. Dalla prima equazione, otteniamo $p_1(t) = \mathrm{e}^{-rt}$ perché $p_1(0) = 1$. E' possibile mostrare che la soluzione della seconda equazione che soddisfa la condizione iniziale $p_n(0) = 0$ è

$$p_n(t) = \mathrm{e}^{-rt}\,\left(1-\mathrm{e}^{-rt}\right)^{n-1} \tag{15.3}$$

per tutti gli $n \geqslant 2$ (figura 15.3b). Quindi, in un tempo fisso t, la distribuzione delle probabilità $(p_n(t))_{n\geqslant 1}$ è «geometrica» con un rapporto tra due termini consecutivi uguale a $1-\mathrm{e}^{-rt}$.

Infatti, notiamo innanzitutto che l'equazione (15.2) è equivalente a:

$$\frac{d}{dt}\left[p_n\,\mathrm{e}^{nrt}\right] = (n-1)\,r\,p_{n-1}\,\mathrm{e}^{nrt}\,, \tag{15.4}$$

da cui possiamo calcolare successivamente $p_2(t)$, $p_3(t)$.... Si ottiene $p_2(t) = \mathrm{e}^{-rt}\,(1-\mathrm{e}^{-rt})$, poi $p_3(t) = \mathrm{e}^{-rt}\,(1-\mathrm{e}^{-rt})^2$, il che suggerisce

la formula (15.3) per la soluzione generale. Si può infine verificare che questa formula è una soluzione dell'equazione (15.4).

Yule ha anche dedotto dalla formula (15.3) che il numero previsto di specie aumenta esponenzialmente con il tempo: $\sum_{n=1}^{+\infty} n\, p_n(t) = \mathrm{e}^{rt}$.

Infatti, notiamo prima che per $|x| < 1$,

$$\sum_{n=1}^{+\infty} n x^{n-1} = \frac{d}{dx} \sum_{n=0}^{+\infty} x^n = \frac{d}{dx}\Big(\frac{1}{1-x}\Big) = \frac{1}{(1-x)^2}\,.$$

Allora:

$$\sum_{n=1}^{+\infty} n\, p_n(t) = \mathrm{e}^{-rt} \sum_{n=1}^{+\infty} n(1-\mathrm{e}^{-rt})^{n-1} = \mathrm{e}^{rt}\,.$$

In particolare, se T è il tempo di raddoppio definito da $\mathrm{e}^{rT} = 2$, allora la distribuzione di probabilità $(p_n(t))_{n\geqslant 1}$ del numero di specie al tempo $t = T$ è geometrica con un rapporto $1/2$: $1/2$, $1/4$, $1/8$, $1/16\ldots$ Al tempo $t = kT$, è geometrica con un rapporto $1 - 1/2^k$ e $p_1(kT) = 1/2^k$.

Yule ha poi considerato, parallelamente alla crescita del numero di specie appartenenti allo stesso genere, un processo simile dovuto a maggiori mutazioni che portano alla creazione di nuovi generi. Sia $s\,dt$ la probabilità per un genere esistente di generare un nuovo genere durante un piccolo intervallo di tempo dt. Come prima, assumendo che ci sia un solo genere al tempo $t = 0$, il numero atteso di generi al tempo t è e^{st}. Il numero medio di generi creati per unità di tempo al tempo t è la derivata $s\,\mathrm{e}^{st}$. Nel limite[2] in cui $t \to +\infty$, il numero medio di generi che all'istante t sono esistiti tra x e $x + dx$ unità di tempo è quindi $s\,\mathrm{e}^{s(t-x)}\,dx$. La probabilità all'istante t che un genere scelto a caso sia esistito tra x e $x + dx$ unità di tempo è $s\,\mathrm{e}^{-sx}\,dx$.

Se un genere scelto a caso all'istante t è esistito tra x e $x + dx$ unità di tempo, la probabilità che questo genere contenga n specie è, secondo la formula (15.3), uguale a $\mathrm{e}^{-rx}(1-\mathrm{e}^{-rx})^{n-1}$ per tutti $n \geqslant 1$. Quindi la probabilità q_n per un genere scelto a caso all'istante t di contenere n specie è: $q_n = \int_0^{+\infty} s\,\mathrm{e}^{-sx}\,\mathrm{e}^{-rx}\,(1-\mathrm{e}^{-rx})^{n-1}\,dx$. Fissiamo $u = r/s$. Un facile calcolo mostra che $q_1 = 1/(1+u)$ e che:

$$q_n = \frac{1}{1+u}\,\frac{u}{1+2u}\,\frac{2u}{1+3u}\,\cdots\,\frac{(n-1)u}{1+nu} \tag{15.5}$$

[2] Yule ha considerato anche il caso in cui t non può essere assunto molto grande rispetto al tempo di raddoppio di e^{st}. I calcoli sono un po' più complicati ma i risultati finali non sono molto diversi.

per tutti gli $n \geqslant 2$.

Infatti, abbiamo $(1-\mathrm{e}^{-rx})^{n-1} = (1-\mathrm{e}^{-rx})^{n-2}(1-\mathrm{e}^{-rx})$. Quindi:

$$q_n = q_{n-1} - s\int_0^{+\infty} \mathrm{e}^{-(r+s)x}\left(1-\mathrm{e}^{-rx}\right)^{n-2}\mathrm{e}^{-rx}\,dx.$$

Integrando per parti, otteniamo:

$$q_n = q_{n-1} - \frac{r+s}{(n-1)\,r}\,q_n, \quad q_n = \frac{(n-1)\,r/s}{1+n\,r/s}\,q_{n-1}\,.$$

La formula (15.5) mostra che la sequenza di probabilità $(q_n)_{n\geqslant 1}$ è decrescente. Quindi il massimo viene raggiunto per $n=1$: la maggior parte dei generi contiene una sola specie. Questo è precisamente ciò che i dati avevano mostrato. Inoltre, la diminuzione di q_n verso 0 quando n tende all'infinito è relativamente lenta perché $q_n/q_{n-1} \to 1$. Questo può spiegare perché alcuni generi contengono un gran numero di specie. Più precisamente, Yule ha mostrato che il $\log q_n$ diminuisce linearmente con il $\log(n)$.

Introduciamo la funzione Gamma di Eulero $\Gamma(z) = \int_0^{+\infty} t^{z-1}\,\mathrm{e}^{-t}\,dt$. Allora $\Gamma(n+1) = n! = n\times(n-1)\times\cdots\times 2\times 1$ quando n è intero e $\Gamma(z+1) = z\Gamma(z)$. Quindi (15.5) assume la forma

$$q_n = \frac{(n-1)!}{u(1+\frac{1}{u})(2+\frac{1}{u})\cdots(n+\frac{1}{u})} = \frac{\Gamma(n)\,\Gamma(1+\frac{1}{u})}{u\,\Gamma(n+1+\frac{1}{u})}\,.$$

Ma l'approssimazione di Stirling dà: $\log\Gamma(n) \approx n\log n - n - \frac{1}{2}\log n +$ cost.te. Allo stesso modo, $\log\Gamma(n+1+1/u) \approx n\log n - n + \left(\frac{1}{u}+\frac{1}{2}\right)\log n +$ cost.te. Infine: $\log q_n \approx -\left(1+\frac{1}{u}\right)\log n +$ cost.te.

Consideriamo per esempio il caso delle lucertole. Il parametro u può essere stimato dalla proporzione $q_1 = 1/(1+u)$ dei generi che contengono una sola specie. Secondo la tabella 15.1, abbiamo $q_1 = 105/259$ quindi $u \approx 1{,}467$. Possiamo quindi calcolare la probabilità teorica q_n e il numero atteso Q_n di generi contenenti n specie moltiplicando q_n per il numero totale di specie, che è 259 (Tabella 15.2). Yule ha notato che l'accordo tra le osservazioni e i calcoli è relativamente buono data la semplicità del modello[3], che non tiene

[3] Per il numero di generi che contengono più di 100 specie, Yule ha ottenuto un adattamento migliore rispetto alla tabella 15.2 considerando che t non era grande rispetto al tempo di raddoppio di e^{st}.

conto per esempio dei cataclismi che le specie hanno subito durante milioni di anni di evoluzione.

Tabella 15.2: Confronto tra dati sperimentali e teoria nel caso delle lucertole (1 580 specie classificate in 259 generi).

Numero di specie per genere	Numero osservato di generi	Numero calcolato di generi
1	105	105,0
2	44	39,2
3	23	21,3
4	14	13,6
5	12	9,6
6	7	7,2
7	6	5,6
8	4	4,5
9	5	3,7
10	5	3,1
11-20	17	16,6
21-30	9	6,9
31-40	3	3,9
41-50	2	2,6
51-60	0	1,9
61-70	1	1,4
71-80	0	1,1
81-90	0	0,9
91-100	0	0,7
101-	2	10,1
totale	259	259

Dopo il 1931 Yule si ritirò progressivamente dall'Università di Cambridge. Si interessò alla distribuzione statistica della lunghezza delle frasi per identificare gli autori dei libri. Lo applicò in particolare al libro pubblicato da John Graunt (si veda il capitolo 2) ma forse ispirato da William Petty. Nel 1944 pubblicò un libro su *Lo studio statistico del vocabolario letterario*. Morì nel 1951.

Oggi il modello di Yule è ancora utilizzato per analizzare gli «alberi filogenetici» (gli alberi genealogici delle specie). Questi alberi, simili a quello della figura 15.3, sono meglio conosciuti grazie ai nuovi dati provenienti dalla biologia molecolare. Ma le applicazioni del processo stocastico defi-

nito dalle equazioni (15.1)-(15.2) non si limitano alla teoria dell'evoluzione. Questo processo è un elemento costitutivo di molti modelli di dinamica di popolazione, dal livello microscopico (per modellizzare ad esempio colonie di batteri) al livello macroscopico (per modellizzare l'inizio di un'epidemia). È chiamato «processo di nascita puro» o «processo di Yule». Una semplice variante include una probabilità $m\,dt$ di morire durante ogni piccolo intervallo di tempo dt: la dimensione attesa della popolazione al tempo t per questo «processo di nascita e morte» è allora $\mathrm{e}^{(r-m)t}$. Per quanto riguarda la distribuzione di probabilità (15.5), essa è talvolta chiamata distribuzione Yule. Le distribuzioni con code che soddisfano le leggi di potenza hanno attirato molta attenzione in varie aree della scienza. Lo studio delle epidemie in reti casuali con una distribuzione di grado con legge di potenza è solo un esempio.

Bibliografia

1. Aldous, D.J.: Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today. *Stat. Sci.* 16, 23–34 (2001) projecteuclid.org
2. Edwards, A.W.F.: George Udny Yule. In: Heyde, C.C., Seneta, E. (eds.) *Statisticians of the Centuries*, 292–294. Springer (2001)
3. McKendrick, A.G.: Studies on the theory of continuous probabilities with special reference to its bearing on natural phenomena of a progressive nature. *Proc. Lond. Math. Soc.* 13, 401–416 (1914)
4. Simon, H.A.: On a class of skew distribution functions. *Biometrika* 42, 425–440 (1955)
5. Willis, J.C.: *Age and Area, A Study in Geographical Distribution and Origin of Species*. Cambridge University Press (1922) archive.org
6. Yates, F.: George Udny Yule. *Obit. Not. Fellows R. Soc.* 8, 308–323 (1952)
7. Yule, G.U.: A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Phil. Trans. Roy. Soc. Lond. B* 213, 21–87 (1925) gallica.bnf.fr

Capitolo 16

McKendrick e Kermack sulla modellizzazione delle epidemie (1926–1927)

Nel 1926 McKendrick studiò un modello stocastico di epidemia e trovò un metodo per calcolare la probabilità che un'epidemia raggiunga una certa dimensione finale. Propose anche l'equazione alle derivate parziali che governa le popolazioni strutturate per età in un contesto a tempo continuo. Nel 1927 Kermack e McKendrick studiarono un modello epidemico deterministico e ottennero un'equazione per la dimensione finale dell'epidemia, che mette in luce una certa soglia per la densità della popolazione. Grandi epidemie possono verificarsi sopra ma non sotto questa soglia. Questi lavori sono ancora molto usati nell'epidemiologia contemporanea.

Anderson Gray McKendrick nacque nel 1876 a Edimburgo, ultimo di cinque figli. Studiò medicina all'Università di Glasgow, dove suo padre era professore di fisiologia. Nel 1900, entrò nel servizio medico indiano. Prima di recarsi in India, accompagnò Ronald Ross in una missione per combattere la malaria in Sierra Leone. Prestò poi servizio nell'esercito per 18 mesi in Sudan. Al suo arrivo in India, fu nominato medico in una prigione del Bengala dove cercò di controllare la dissenteria. Nel 1905 entrò a far parte del nuovo Istituto centrale di ricerca medica di Kasauli (nel nord dell'India). Lavorò sulla rabbia ma studiò anche matematica. Nel 1920, avendo contratto una malattia tropicale, tornò a Edimburgo e fu nominato sovrintendente del laboratorio del *Royal College of Physicians*.

Nel 1926 McKendrick pubblicò un articolo sulle *Applicazioni della matematica ai problemi medici*, che conteneva diverse idee nuove. Introdusse in particolare un modello matematico a tempo continuo per le epidemie, che teneva conto dell'aspetto stocastico dell'infezione e della guarigione.

Consideriamo una popolazione di dimensione N con inizialmente una sola persona infetta. Le persone possono passare successivamente attraverso tre stati: lo stato di suscettibile S, lo stato di infetto I e lo stato di guarito R (Figura 16.2)[1].

[1] Il modello di Daniel Bernoulli (vedi il capitolo 4) includeva gli stati S e R ma non I, essendo la durata dell'infezione molto più breve della vita media.

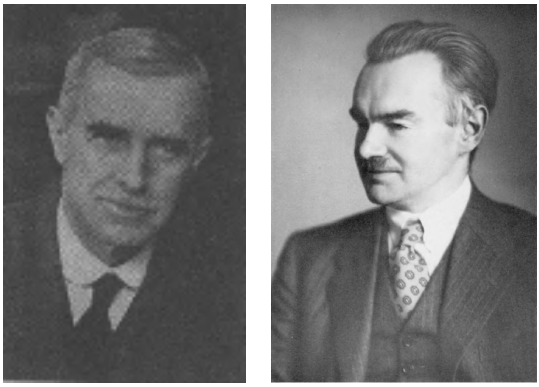

Figura 16.1: McKendrick (1876–1943) e Kermack (1898-1970)

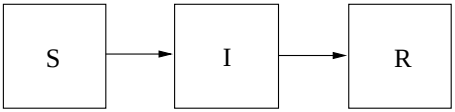

Figura 16.2: Stati possibili: suscettibile (S), infetto (I), guarito (R).

Sia $p_{i,r}(t)$ la probabilità che la popolazione contenga al tempo t esattamente i persone nello stato I e r persone nello stato R, dove i e r sono numeri interi tali che $1 \leqslant i+r \leqslant N$. In questo caso si dice che la popolazione è nello stato (i,r). Il numero di persone suscettibili è $s = N-i-r$. Seguendo il lavoro di Ross sulla malaria (vedi capitolo 12), McKendrick assunse che, durante un piccolo intervallo di tempo dt, la probabilità che si verifichi una nuova infezione è uguale a $a\,s\,i\,dt$ (cioè proporzionale sia al numero di persone suscettibili che al numero di persone infette). La probabilità di una nuova guarigione è uguale a $b\,i\,dt$. Sia a che b sono parametri positivi. Per calcolare $p_{i,r}(t+dt)$, bisogna distinguere diversi casi:

- la popolazione è nello stato $(i-1,r)$ al tempo t e una nuova infezione sposta la popolazione nello stato (i,r) tra t e $t+dt$; la probabilità di questo evento è $a\,s\,(i-1)\,dt$ con $s = N-(i-1)-r$;

- la popolazione è nello stato (i,r) al tempo t e una nuova infezione sposta la popolazione nello stato $(i+1,r)$ tra t e $t+dt$; la probabilità di questo evento è $a\,s\,i\,dt$ con $s = N-i-r$;

- la popolazione è nello stato $(i+1,r-1)$ al tempo t e una nuova guarigione sposta la popolazione nello stato (i,r) tra t e $t+dt$; la probabilità di questo evento è $b\,(i+1)\,dt$;

- la popolazione è nello stato (i,r) al tempo t e una nuova guarigione sposta la popolazione nello stato $(i-1,r+1)$ tra t e $t+dt$; la probabilità di questo evento è $b\,i\,dt$.

Quindi, McKendrick ottenne le equazioni

$$\begin{aligned}\frac{dp_{i,r}}{dt} &= a\,(N-i-r+1)\,(i-1)\,p_{i-1,r} - a\,(N-i-r)\,i\,p_{i,r}\\ &\quad + b\,(i+1)\,p_{i+1,r-1} - b\,i\,p_{i,r}\end{aligned} \tag{16.1}$$

per $1 \leqslant i+r \leqslant N$. Il primo termine sul lato destro manca quando $i=0$, mentre il terzo termine manca quando $r=0$. Le condizioni iniziali sono $p_{i,r}(0)=0$ per ogni (i,r) tranne $p_{1,0}(0)=1$.

Con questo modello McKendrick riuscì a calcolare la probabilità che l'epidemia finisca con n persone infettate, che è il limite di $p_{0,n}(t)$ quando $t \to +\infty$. In realtà, non c'è bisogno di risolvere il sistema (16.1). Basta notare che finché ci sono i persone infette e r persone guarite, la probabilità di una nuova infezione durante un piccolo intervallo di tempo dt è $a\,(N-i-r)\,i\,dt$ e la probabilità di una nuova guarigione è $b\,i\,dt$. Quindi le probabilità di transizione (come sono solitamente chiamate nella teoria delle catene di Markov) dallo stato (i,r) allo stato $(i+1,r)$ o allo stato $(i-1,r+1)$ sono rispettivamente

$$\mathscr{P}_{(i,r)\to(i+1,r)} = \frac{a\,(N-i-r)}{a\,(N-i-r)+b}\,, \qquad \mathscr{P}_{(i,r)\to(i-1,r+1)} = \frac{b}{a\,(N-i-r)+b}\,,$$

per ogni gli $i \geqslant 1$ (Figura 16.3).

Sia $q_{i,r}$ la probabilità che la popolazione passi per lo stato (i,r) durante l'epidemia. Poiché $i=1$ e $r=0$ quando $t=0$, abbiamo $q_{1,0}=1$. Gli altri stati sono raggiunti o dopo un'infezione o dopo una guarigione:

$$q_{i,r} = q_{i-1,r}\;\mathscr{P}_{(i-1,r)\to(i,r)} + q_{i+1,r-1}\;\mathscr{P}_{(i+1,r-1)\to(i,r)}\,.$$

Il primo termine del lato destro manca quando $i=0$ o $i=1$. Il secondo termine manca quando $r=0$. Da questa formula, possiamo prima calcolare $(q_{i,0})_{2\leqslant i\leqslant N}$, poi $(q_{i,1})_{0\leqslant i\leqslant N-1}$, poi $(q_{i,2})_{0\leqslant i\leqslant N-2}$ etc. La probabilità che l'epidemia infetti infine n persone è $q_{0,n}$. Nel 1926 questi calcoli erano abbastanza noiosi. Così McKendrick si limitò ad esempi riguardanti popolazioni molto piccole, per esempio una famiglia. Con $N=5$ persone e $b/a=2$, ottenne la tabella 16.1. Le maggiori probabilità corrispondono al caso in cui solo una persona della famiglia è infettata e al caso in cui l'intera famiglia è infettata.

Lo stesso articolo del 1926 contiene anche una nuova formulazione dei problemi demografici quando il tempo è considerato una variabile continua.

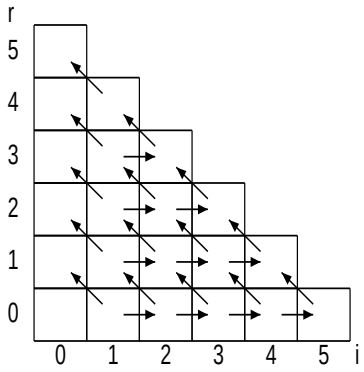

Figura 16.3: Diagramma che mostra i possibili stati di una popolazione con $N = 5$ (i sull'asse orizzontale, r sull'asse verticale) e le possibili transizioni dovute all'infezione (frecce orizzontali) o alla guarigione (altre frecce).

Tabella 16.1: Probabilità che durante un'epidemia in una famiglia di cinque persone se ne infettino n quando $b/a = 2$.

n	1	2	3	4	5
$q_{0,n}$	0,33	0,11	0,09	0,13	0,34

Per dx infinitamente piccolo, sia $P(x,t)\,dx$ la popolazione con un'età compresa tra x e $x+dx$ al tempo t. Sia $m(x)$ la mortalità all'età x. Allora

$$P(x+h,t+h) \approx P(x,t) - m(x)\,P(x,t)\,h$$

per h infinitamente piccolo. Introducendo le derivate parziali della funzione $P(x,t)$:

$$\frac{\partial P}{\partial x}(x,t) = \lim_{h\to 0} \frac{P(x+h,t)-P(x,t)}{h}, \quad \frac{\partial P}{\partial t}(x,t) = \lim_{h\to 0} \frac{P(x,t+h)-P(x,t)}{h},$$

e usando l'approssimazione

$$P(x+h,t+h) \approx P(t,x) + h\,\frac{\partial P}{\partial x}(x,t) + h\,\frac{\partial P}{\partial t}(x,t)\ ,$$

McKendrick ottenne la seguente equazione alle derivate parziali:

$$\frac{\partial P}{\partial t}(x,t) + \frac{\partial P}{\partial x}(x,t) + m(x)\,P(x,t) = 0\ .$$

Un'equazione di questo tipo si presenta in modo naturale in problemi di popolazione strutturati da una variabile continua, come l'età in demografia (vedi il capitolo 25) o il tempo dall'infezione in epidemiologia.

William Ogilvy Kermack era nato nel 1898 in una piccola città della Scozia. Aveva studiato all'Università di Aberdeen e iniziato a fare ricerca nel campo della chimica organica in un laboratorio industriale di Oxford. Nel 1921, Kermack fu nominato responsabile della sezione chimica del *Royal College of Physicians Laboratory* di Edimburgo. Sebbene divenuto completamente cieco nel 1924, dopo un'esplosione nel suo laboratorio di Edimburgo, continuò il suo lavoro di chimico con l'aiuto di colleghi e studenti. Kermack iniziò anche a collaborare con McKendrick sulla modellizzazione matematica delle epidemie. A partire dal 1927, pubblicarono insieme una serie di *Contributi alla teoria matematica delle epidemie* in cui studiavano modelli epidemici deterministici. Sia N la dimensione della popolazione con N abbastanza grande. Assumiamo, come nell'articolo del 1926, che le persone possano essere suscettibili, infette o guarite. Se la malattia è fatale, allora il terzo stato è in realtà la morte. Siano $S(t)$, $I(t)$ e $R(t)$ il numero di persone in ciascuno dei tre stati. Il modello è (in forma semplificata) un sistema di tre equazioni differenziali:

$$\frac{dS}{dt} = -aSI, \tag{16.2}$$

$$\frac{dI}{dt} = aSI - bI, \tag{16.3}$$

$$\frac{dR}{dt} = bI. \tag{16.4}$$

Quindi, il numero di nuove infezioni per unità di tempo è, come nel modello stocastico del 1926, proporzionale sia al numero di persone suscettibili che al numero di persone infette. All'inizio dell'epidemia, al tempo $t = 0$, un certo numero di persone è infetto: $S(0) = N - I_0$, $I(0) = I_0$ e $R(0) = 0$, assumendo $0 < I_0 < N$.

Anche se il sistema (16.2)-(16.4) non ha una soluzione esplicita, diverse sue proprietà possono essere dimostrate:

- la popolazione totale $S(t) + I(t) + R(t)$ rimane costante e uguale a N;
- $S(t)$, $I(t)$ e $R(t)$ rimangono non negativi (come dovrebbe essere, visto che si tratta popolazioni);
- quando $t \to +\infty$, $S(t)$ tende decrescendo ad un limite $S_\infty > 0$, $I(t)$ tende a 0 e $R(t)$ tende crescendo ad un limite $R_\infty < N$;

- inoltre la formula

$$-\log \frac{S_\infty}{S(0)} = \frac{a}{b}(N - S_\infty)\,, \tag{16.5}$$

fornisce implicitamente S_∞ e quindi anche la dimensione finale dell'epidemia $R_\infty = N - S_\infty$.

Infatti, vediamo anzitutto che $\frac{d}{dt}(S+I+R) = 0$. Quindi $S(t)+I(t)+R(t) = S(0)+I(0)+R(0) = N$. Le equazioni (16.2) e (16.3) possono essere riscritte come

$$\frac{d}{dt}\left[S(t)\,\mathrm{e}^{a\int_0^t I(\tau)\,d\tau}\right] = 0\,, \quad \frac{d}{dt}\left[I(t)\,\mathrm{e}^{bt-a\int_0^t S(\tau)\,d\tau}\right] = 0\,.$$

Ne segue che

$$S(t) = S(0)\,\mathrm{e}^{-a\int_0^t I(\tau)\,d\tau} > 0, \quad I(t) = I(0)\,\mathrm{e}^{a\int_0^t S(\tau)\,d\tau - bt} > 0.$$

Le equazioni (16.2) e (16.4) mostrano quindi che la funzione $S(t)$ è decrescente e che la funzione $R(t)$ è crescente (in particolare, $R(t) \geqslant 0$). Poiché $S(t) \geqslant 0$ e $R(t) \leqslant N$, le funzioni $S(t)$ e $R(t)$ hanno limiti quando $t \to +\infty$. Dato che $I(t) = N - S(t) - R(t)$, anche $I(t)$ ha un limite quando $t \to +\infty$, che può essere solo zero come si può vedere integrando (16.4). L'equazione (16.2) mostra anche che

$$-\frac{d}{dt}[\log S] = aI.$$

Integrando tra $t = 0$ e $t = +\infty$, troviamo

$$\log S(0) - \log S_\infty = a\int_0^{+\infty} I(t)\,dt.$$

L'equazione (16.3) può essere riscritta come

$$\frac{dI}{dt} = -\frac{dS}{dt} - b,I.$$

Integrando tra $t = 0$ e $t = +\infty$, otteniamo

$$-I(0) = S(0) - S_\infty - b\int_0^{+\infty} I(t)\,dt.$$

Combinando i due risultati, otteniamo la formula (16.5), che mostra che $S_\infty > 0$.

Quando il numero iniziale di persone infette I_0 è piccolo rispetto alla dimensione della popolazione N, come si verifica spesso all'inizio di un'epidemia in una città, la formula (16.5) può essere riscritta usando $S_\infty = N - R_\infty$ come

$$-\log\left(1 - \frac{R_\infty}{N}\right) \approx \mathscr{R}_0 \frac{R_\infty}{N}, \tag{16.6}$$

dove per definizione

$$\mathscr{R}_0 = \frac{aN}{b}.$$

L'equazione (16.6) ha una soluzione positiva solo se $\mathscr{R}_0 > 1$. Così Kermack e McKendrick arrivano alla seguente conclusione: l'epidemia infetta una frazione non trascurabile della popolazione solo se $\mathscr{R}_0 > 1$. Esiste una soglia per la densità della popolazione $N^* = b/a$ al di sotto della quale non possono verificarsi epidemie.

Quando la dimensione della popolazione N è appena sopra questa soglia ($N = N^* + \varepsilon$), si verifica un'epidemia di piccola ampiezza. Segue da (16.6) che $R_\infty \approx 2\varepsilon$. Quindi $S_\infty \approx N^* - \varepsilon$: l'epidemia porta la popolazione suscettibile sotto la soglia N^* tanto quanto era inizialmente sopra.

In effetti, usando l'approssimazione $-\log(1-x) \approx x + \frac{x^2}{2}$, l'equazione (16.6) diventa

$$\frac{R_\infty}{N} + \frac{1}{2}\left(\frac{R_\infty}{N}\right)^2 \approx \mathscr{R}_0 \frac{R_\infty}{N}.$$

Quindi

$$R_\infty \approx 2(\mathscr{R}_0 - 1)N = 2\frac{\varepsilon}{N^*}(N^* + \varepsilon) \approx 2\varepsilon.$$

Come nel modello di Ross per la malaria (Capitolo 12), la condizione $\mathscr{R}_0 > 1$ ha una semplice interpretazione. Poiché aN è il numero di persone che una persona infetta infetta per unità di tempo all'inizio dell'epidemia e poiché $1/b$ è la durata media dell'infezione, $\mathscr{R}_0 = aN/b$ è il numero medio di casi secondari dovuti a una persona infetta all'inizio dell'epidemia.

Per le malattie mortali, $R(t)$ è il numero cumulativo di morti dall'inizio dell'epidemia e dR/dt è il numero di morti per unità di tempo. Kermack e McKendrick hanno notato che il grafico della funzione dR/dt nel loro modello matematico ha la forma a campana che ci si aspetta da una curva epidemica (Figura 16.4).

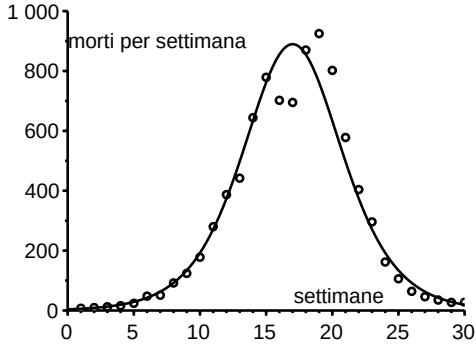

Figura 16.4: La curva dR/dt in funzione del tempo e i dati del numero di morti per settimana durante un'epidemia di peste a Bombay nel 1905-1906.

Per ricavare dR/dt, hanno diviso (16.2) per (16.4) per ottenere $dS/dR = -aS/b$. Quindi

$$S(t) = S(0)\exp(-aR(t)/b).$$

Sostituendo questo nell'equazione (16.4) e usando $S(t)+I(t)+R(t) = N$, si ottiene l'equazione

$$\frac{dR}{dt} = b\left[N - R - S(0)\exp\left(-\frac{a}{b}R\right)\right], \tag{16.7}$$

che ancora non può essere risolta esplicitamente. Tuttavia, se $\frac{a}{b}R(t)$ rimane piccolo durante l'intera epidemia, l'approssimazione $\exp(-u) \approx 1 - u + u^2/2$ dà

$$\frac{dR}{dt} \approx b\left[N - R - S(0) + S(0)\frac{a}{b}R - S(0)\frac{a^2}{2b^2}R^2\right]. \tag{16.8}$$

Questa è una cosiddetta equazione di Riccati con due soluzioni costanti, una positiva R_+ e una negativa R_-, date dalle radici del polinomio del secondo ordine in R che compare a destra in (16.8). Sia $\tilde{R}(t)$ la soluzione esatta di (16.8) e si ponga $Q(t) = \tilde{R}(t) - R_+$. Allora $Q(t)$ soddisfa un'equazione differenziale di Bernoulli simile a quelle incontrate da Daniel Bernoulli e Verhulst (vedi (4.5) e (6.1). Si può quindi adattare direttamente la formula (6.2) per ottenere $Q(t)$. Un facile ma tedioso calcolo

mostra che dQ/dt è della forma

$$\frac{\alpha}{\cosh^2(\beta t - \gamma)},$$

dove α, β e γ sono costanti che dipendono in modo complicato dai parametri del modello. Dato che $dR/dt \approx d\tilde{R}/dt = dQ/dt$, Kermack e McKendrick potevano scegliere (α, β, γ) per adattare i loro dati. Naturalmente i computer e i software moderni possono facilmente risolvere numericamente l'equazione differenziale (16.7) senza passare attraverso queste approssimazioni.

La curva dR/dt così ottenuta si adattava bene ai dati del numero di morti per settimana durante l'epidemia di peste a Bombay tra dicembre 1905 e luglio 1906 (Figura 16.4).

Kermack e McKendrick hanno anche considerato il modello più generale in cui l'infettività $a(x)$ dipende dal tempo x dall'infezione e dove anche il tasso di guarigione $b(x)$ dipende da x. L'equazione che dà la dimensione finale dell'epidemia (quando il numero iniziale di casi infetti è piccolo) è ancora (16.6) ma con

$$\mathscr{R}_0 = N \int_0^{+\infty} a(x) \exp\left(-\int_0^x b(y)\,dy\right) dx. \tag{16.9}$$

Il parametro $\mathscr{R}_0$ ha la stessa interpretazione del caso precedente: è il numero medio di casi secondari dovuti a una persona infetta all'inizio dell'epidemia. Si noti la somiglianza tra (16.9) e la formula di Lotka (10.2) per $\mathscr{R}_0$ in demografia: l'età è sostituita dal tempo dall'infezione, la sopravvivenza dalla probabilità

$$\exp\left(-\int_0^x b(y)\,dy\right)$$

di essere ancora infetti, la fertilità dal tasso di contatto $N\,a(x)$.

Kermack e McKendrick svilupparono diversi altri modelli matematici di epidemie durante gli anni '30. Questi sono ancora gli elementi costitutivi della maggior parte dei modelli più complessi usati oggi in epidemiologia. Il parametro $\mathscr{R}_0$ gioca ancora un ruolo centrale nell'analisi del modello.

McKendrick si ritirò nel 1941 e morì nel 1943. Tra il 1930 e il 1933, Kermack fu coautore di alcuni articoli di fisica matematica con William McCrea e Edmund Whittaker, entrambi del dipartimento di matematica dell'Università di Edimburgo. Durante gli anni '30 e '40, il team di chimici di Kermack cercò di sintetizzare nuove molecole con attività antimalarica, ma con poco successo. Nel 1938 Kermack fu coautore con Philip Eggleton di un libro popolare

sulla biochimica elementare, *La materia di cui siamo fatti*. Fu eletto membro della Royal Society nel 1944 e ottenne la cattedra di biochimica all'Università di Aberdeen nel 1949. In seguito fu preside della Facoltà di Scienze. Si ritirò nel 1968 e morì nel 1970.

Bibliografia

1. Advisory Committee appointed by the Secretary of State for India, the Royal Society and the Lister Institute: Reports on plague investigations in India, XXII. *J. Hyg.* 7, 724–798 (1907) ncbi.nlm.nih.gov
2. Davidson, J.N., Yates, F., McCrea, W.H.: William Ogilvy Kermack 1898–1970. *Biog. Mem. Fellows R. Soc.* 17, 399–429 (1971)
3. Gani, J.: A.G. McKendrick. In: Heyde, C.C., Seneta, E. (eds.) *Statisticians of the Centuries*, 323–327. Springer, New York (2001)
4. Harvey, W.F.: A.G. McKendrick 1876–1943. *Edinb. Med. J.* 50, 500–506 (1943)
5. McKendrick, A.G.: Applications of mathematics to medical problems. *Proc. Edinb. Math. Soc.* 13, 98–130 (1926)
6. Kermack, W.O., McKendrick, A.G.: A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A* 115, 700–721 (1927) gallica.bnf.fr

Capitolo 17

Haldane e le mutazioni (1927)

In un'altro paragrafo del suo articolo *Sul rapporto di dominanza* pubblicato nel 1922, Fisher considerò il problema di un gene mutante che può essere trasmesso ad un numero casuale di figli con una data distribuzione di probabilità. Il problema era formalmente lo stesso dell'estinzione dei nomi di famiglia, ma in un contesto genetico. Fisher dimostrò che se la distribuzione di probabilità era una distribuzione di Poisson e se il gene mutante non aveva alcun vantaggio selettivo, allora il gene mutante poteva scomparire dalla popolazione molto lentamente. Nel 1927 il biologo britannico Haldane spinse ulteriormente lo studio di questo modello e mostrò che la probabilità che un gene mutante dominante non scomparisse era il doppio del suo vantaggio selettivo. Ha anche dato un trattamento più rigoroso del problema dell'estinzione.

John Burdon Sanderson Haldane è nato nel 1892 a Oxford, dove suo padre era professore di fisiologia all'università. Haldane studiò all'*Eton College* e dopo il 1911 al *New College* dell'Università di Oxford. Dopo essersi concentrato sulla matematica nel suo primo anno, si dedicò alle materie umanistiche. Chiamato alle armi, i suoi studi furono interrotti dalla prima guerra mondiale, durante la quale fu inviato a combattere in Francia e in Iraq. Essendo stato ferito, fu inviato come istruttore militare in India. Nel 1915 pubblicò un primo articolo in cui si riportavano i risultati degli esperimenti genetici sui topi che aveva iniziato prima della guerra. Nel 1919 divenne membro accademico del *New College*, insegnando fisiologia e studiando la fisiologia della respirazione come suo padre. Nel 1923 si unì al laboratorio di biochimica di F. G. Hopkins[1] all'Università di Cambridge, dove si concentrò sulla cinetica degli enzimi. Pubblicò anche un romanzo di fantascienza, *Dedalo o Scienza e futuro* (1923), e un saggio intitolato *Callinicus, Una difesa della guerra chimica* (1925). Tra il 1924 e il 1934, scrisse una serie di dieci articoli intitolati *Una teoria matematica della selezione naturale e artificiale*.

Nel quinto articolo della serie, pubblicato nel 1927, Haldane riconsiderò un altro modello genetico che Fisher aveva studiato nel 1922, un modello incentrato sulle mutazioni. Fisher aveva studiato la probabilità per un gene

[1] Frederick Gowland Hopkins, che ha ricevuto il premio Nobel per la medicina e fisiologia nel 1929 per il suo lavoro sulle vitamine.

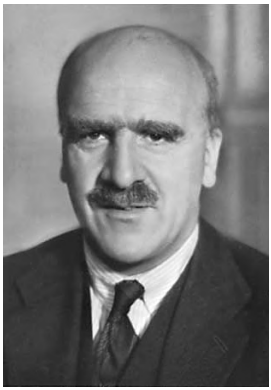

Figura 17.1:
Haldane (1892–1964)

mutante di prevalere in una popolazione o di scomparire. Questo problema è formalmente lo stesso di quello di Bienaymé, Galton e Watson sull'estinzione dei nomi di famiglia. Ma Fisher non ha fatto alcun riferimento a questi lavori, anche se potrebbe aver letto l'articolo di Galton e Watson riprodotto nell'appendice del libro di Galton del 1889 *L'eredità naturale*. Come nel capitolo 9, chiamò p_k la probabilità che un gene sia trasmesso a k figli nella prima generazione ($k \geqslant 0$). Fisher considerò anche la funzione generatrice

$$f(x) = p_0 + p_1 x + p_2 x^2 + \cdots + p_k x^k + \cdots,$$

senza fissare alcun limite superiore per k: la somma può includere un numero infinito di termini. Si rese conto che, partendo da un individuo con il gene mutante nella 0-esima generazione , la probabilità che questo gene sia in k individui è data dal coefficiente di x^k in $f_1(x) = f(x)$ per l'1-esima generazione, in $f_2(x) = f(f(x))$ per la 2-esima generazione , in $f_3(x) = f(f(f(x)))$ per la 3 -esima generazione ecc. In questo modo, diventa chiaro che è valida la seguente equazione:

$$f_n(x) = f(f_{n-1}(x)) \tag{17.1}$$

Questa equazione è molto più pratica dell'equazione $f_n(x) = f_{n-1}(f(x))$ derivata da Watson. In particolare, segue da (17.1) che la probabilità di estinzione entro n generazioni $x_n = f_n(0)$ soddisfa la formula di iterazione $x_n = f(x_{n-1})$, come Bienaymé aveva già notato.

Come esempio, Fisher ha considerato il caso di una pianta con un gene mutante che può produrre N semi, ogni seme ha una probabilità q di sopravvivere per produrre una nuova pianta. La probabilità p_k di ottenere k figli con

il gene mutante è binomiale:

$$p_k = \binom{N}{k} q^k (1-q)^{N-k}$$

per $0 \leqslant k \leqslant N$ e con $p_k = 0$ per $k > N$. La funzione generatrice è quindi $f(x) = (1-q+qx)^N$. Sia $\mathscr{R}_0 = Nq$ il numero medio di semi che sopravvivono per produrre una nuova pianta. Quando N è grande e q è piccolo, allora:

$$f(x) = \left(1 + \frac{\mathscr{R}_0}{N}(x-1)\right)^N \approx \mathrm{e}^{\mathscr{R}_0(x-1)} = \mathrm{e}^{-\mathscr{R}_0} \sum_{k=0}^{+\infty} \frac{(\mathscr{R}_0 x)^k}{k!} .$$

La distribuzione di probabilità (p_k) tende a $\mathrm{e}^{-\mathscr{R}_0}(\mathscr{R}_0)^k/k!$, nota come distribuzione di Poisson. Fisher ha poi calcolato la probabilità di estinzione entro n generazioni, usando $x_0 = 0$, $x_n \approx \exp(\mathscr{R}_0[x_{n-1} - 1])$ e i valori numerici $N = 80$ e $q = 1/80$. In questo caso, $\mathscr{R}_0 = Nq = 1$. Un noioso calcolo mostra che $x_{100} \approx 0{,}98$: un gene mutante senza vantaggio selettivo ($\mathscr{R}_0 = 1$) scompare molto lentamente. C'è ancora un 2 % di possibilità che il gene sia presente nella popolazione dopo 100 generazioni. Nel 1922 Fisher non approfondì ulteriormente lo studio di questo modello.

Continuando il lavoro di Fisher, Haldane notò per la prima volta nel suo articolo del 1927 che, per qualsiasi distribuzione di probabilità (p_k) tale che $p_0 > 0$, l'equazione $x = f(x)$ ha esattamente due radici nell'intervallo $(0,1]$ quando il numero medio di figli che portano il gene mutante $\mathscr{R}_0$ è strettamente maggiore di 1, cioè quando il gene mutante ha un vantaggio selettivo. Inoltre, la probabilità di estinzione x_∞, che è il limite di x_n come $n \to +\infty$, è la più piccola delle due radici di $x = f(x)$: il gene ha una probabilità non nulla di stabilirsi nella popolazione. A differenza di Bienaym'e e Cournot, Haldane fornì una prova per questa conclusione.

Infatti, $f'(x) \geqslant 0$ e $f''(x) \geqslant 0$ sull'intervallo $[0,1]$. In altre parole, la funzione $f(x)$ è non decrescente e convessa. Le ipotesi $f(0) = p_0 > 0$ e $f'(1) = \mathscr{R}_0 = p_1 + 2p_2 + 3p_3 + \cdots > 1$ implicano che l'equazione $f(x) = x$ ha esattamente due soluzioni nell'intervallo $(0,1]$: $x = 1$ e x^* tali che $0 < x^* < 1$. Haldane fece poi riferimento ad un articolo di Gabriel Koenigs del 1883, che mostrava che se $x_n = f(x_{n-1})$ e $x_n \to x_\infty$, allora $x_\infty = f(x_\infty)$ e $|f'(x_\infty)| \leqslant 1$. Quando $f'(1) > 1$, l'unica possibilità è che $x_\infty = x^*$.

Nel caso di una distribuzione di Poisson con $f(x) = \mathrm{e}^{\mathscr{R}_0(x-1)}$ e $\mathscr{R}_0$ appena più grande di 1, la probabilità di estinzione x_∞ è molto vicina a 1. L'equazione

$f(x_\infty) = x_\infty$ è quindi equivalente a:

$$\mathscr{R}_0(x_\infty - 1) = \log x_\infty \approx (x_\infty - 1) - \frac{(x_\infty - 1)^2}{2}\,.$$

Ne segue che $1 - x_\infty \approx 2(\mathscr{R}_0 - 1)$. Haldane concluse che la probabilità che il gene mutante non si estingua è il doppio del suo vantaggio selettivo $\mathscr{R}_0 - 1$. Senza citare Haldane, Fisher prese come esempio nel suo libro del 1930 il caso in cui $\mathscr{R}_0 = 1{,}01$, che dà una probabilità del 2 % che il gene mutante non si estingua.

Haldane divenne un membro accademico della *Royal Society* nel 1932. Lasciò Cambridge per diventare professore di genetica e poi di biometria all'University College di Londra. Era allora particolarmente interessato alla genetica umana: stima dei tassi di mutazione, mappe genetiche dei cromosomi ecc. Oltre ai suoi libri scientifici (*Biologia animale* nel 1927 con Julian Huxley, *Enzimi* nel 1930 e *Le cause dell'evoluzione* nel 1932, *La biochimica della genetica* nel 1954), pubblicò un gran numero di articoli sulla scienza nella stampa (per esempio, sull'origine della vita) e alcuni saggi (*L'ineguaglianza dell'uomo* nel 1932, *La filosofia di un biologo* nel 1935, *La filosofia marxista e le scienze* nel 1938, *Heredità e politica* nel 1938 e *I progressi della scienza* nel 1947). Dopo diverse visite in Spagna durante la guerra civile, cercò di convincere il suo paese a costruire rifugi contro i bombardamenti aerei. Durante la seconda guerra mondiale, lavorò sui problemi della respirazione nei sottomarini. Membro del partito comunista dal 1942, si dimise nel 1950 a causa del rifiuto ufficiale della genetica mendeliana in URSS a causa dell'influenza di Lysenko. Nel 1957 si stabilì in India, dove continuò le sue ricerche, prima all'Indian Statistical Institute di Calcutta e poi a Bhubaneswar. Divenuto cittadino indiano, morì nel 1964.

Bibliografia

1. Clark, R.: *J.B.S., The Life and Work of J.B.S. Haldane*. Hodder and Stoughton, London (1968)
2. Haldane, J.B.S.: A mathematical theory of natural and artificial selection, Part V, Selection and mutation. *Proc. Camb. Philos. Soc.* 23, 838–844 (1927)
3. Haldane, J.B.S.: *The Causes of Evolution*. Longmans (1932) archive.org
4. Pirie, N.W.: John Burdon Sanderson Haldane 1892-1964. *Biog. Mem. Fellows R. Soc.* 12, 218–249 (1966)

Capitolo 18

Erlang e Steffensen sul problema dell'estinzione (1929–1933)

Nel 1929 l'ingegnere delle telecomunicazioni Erlang considerò ancora una volta il problema dell'estinzione dei nomi di famiglia. Il suo connazionale, lo statistico Steffensen, elaborò una soluzione completa del problema. Mostrò in particolare che l'aspettativa del numero di figli in ogni generazione cresce esponenzialmente, creando un collegamento tra i modelli di popolazione stocastici e deterministici.

Agner Krarup Erlang è nato nel 1878 a Lønborg, Danimarca. Suo padre era un maestro di scuola. Tra il 1896 e il 1901, il giovane Erlang studiò matematica, fisica e chimica all'Università di Copenhagen. Ha poi insegnato diversi anni nelle scuole superiori, pur mantenendo un interesse per la matematica, in particolare per la teoria della probabilità. Incontrò Jensen, ingegnere capo della Compagnia Telefonica di Copenaghen e matematico dilettante, che lo convinse nel 1908 ad unirsi al nuovo laboratorio di ricerca della compagnia. Erlang iniziò a pubblicare articoli sulle applicazioni della teoria della probabilità alla gestione delle chiamate telefoniche. Nel 1917 scoprì una formula per i tempi di attesa, che fu rapidamente utilizzata dalle compagnie telefoniche di tutto il mondo. I suoi articoli, pubblicati inizialmente in danese, furono poi tradotti in diverse altre lingue.

Figura 18.1:
Erlang (1878–1929)

Nel 1929 Erlang si interessò allo stesso problema dell'estinzione che Bienaymé, Galton e Watson avevano studiato prima di lui per i nomi di famiglia e che Fisher e Haldane avevano studiato per i geni mutanti. Come i suoi predecessori, non era a conoscenza di tutti i lavori che erano stati pubblicati. Chiamando ancora p_k la probabilità per un individuo di avere k figli, notò che la probabilità x_n di estinzione entro n generazioni soddisfa l'equazione: $x_n = p_0 + p_1 x_{n-1} + p_2 (x_{n-1})^2 + \cdots = f(x_{n-1})$ con $x_0 = 0$. Notò anche che la probabilità complessiva di estinzione x_∞, che è il limite di x_n come $n \to +\infty$, è una soluzione dell'equazione $x_\infty = f(x_\infty)$. Si rese conto che $x = 1$ era sempre una soluzione e che un'altra soluzione esisteva tra 0 e 1 quando il numero medio di figli $\mathscr{R}_0 = f'(1)$ è maggiore di 1. Ma sembra che non riuscisse a capire quale di queste due soluzioni fosse quella giusta. Come Galton, sottopose il problema nel 1929 a una rivista matematica danese, *Matematisk Tidsskrift*:

> «Domanda 15. Nel caso in cui la probabilità che un individuo abbia k figli è p_k, dove $p_0 + p_1 + p_2 + \cdots = 1$, trovare la probabilità che la sua famiglia si estingua.»

Purtroppo, Erlang morì quello stesso anno 1929 all'età di 51 anni. Di fatto, morì senza figli. [1].

Un professore di matematica attuariale all'Università di Copenaghen, Johan Frederik Steffensen, riprese il quesito di Erlang. Pubblicò nel 1930 la sua soluzione nella stessa rivista danese: la probabilità di estinzione x_∞ è sempre la più piccola radice dell'equazione $x = f(x)$ nell'intervallo chiuso $[0,1]$, come avevano già notato Bienaymé e Haldane. La dimostrazione di Steffensen è quella che si trova nei moderni libri di testo.

> Infatti, abbiamo visto che la probabilità di estinzione x_∞ è una soluzione di $x = f(x)$ nell'intervallo chiuso $[0,1]$. Sia x^* la più piccola di queste soluzioni. Per definizione, $x^* \leqslant x_\infty$. Steffensen notò per primo che $x^* = f(x^*) \geqslant p_0 = x_1$. Assumiamo per induzione che $x^* \geqslant x_n$. Allora $x^* = f(x^*) \geqslant f(x_n) = x_{n+1}$ poiché la funzione $f(x)$ è crescente. Quindi $x^* \geqslant x_n$ per tutti gli n. Prendendo il limite, $x^* \geqslant x_\infty$. Quindi $x_\infty = x^*$.

Steffensen ha dato anche una spiegazione più formale del perché $x = 1$ è l'unica radice di $x = f(x)$ quando il numero medio di figli $\mathscr{R}_0 = f'(1)$ è minore o uguale a 1 (figura 18.2a) e perché c'è solo un'altra radice diversa

[1] In sua memoria, il Comitato Consultivo Telefonico Internazionale decise nel 1946 di chiamare «erlang» l'unità di misura dell'intensità del traffico telefonico. «Erlang» è anche il nome dato a un linguaggio di programmazione della società Ericsson.

da $x = 1$ nel caso in cui $\mathscr{R}_0 > 1$ (figura 18.2b). Si noti che $\mathscr{R}_0 = f'(1)$ è la pendenza della funzione $f(x)$ a $x = 1$.

Ha notato che per qualsiasi radice di $x = f(x)$,

$$1 - x = 1 - f(x) = 1 - p_0 - \sum_{k=1}^{+\infty} p_k x^k = \sum_{k=1}^{+\infty} p_k (1 - x^k) \, .$$

Assumendo $x \neq 1$ e dividendo per $1 - x$, otteniamo $1 = p_1 + p_2(1+x) + p_3(1+x+x^2) + \cdots$. Quando x aumenta da 0 a 1, il lato destro dell'equazione aumenta da $1 - p_0$ a $\mathscr{R}_0 = f'(1)$. Se $\mathscr{R}_0 < 1$, allora l'equazione non ha soluzione. Se $\mathscr{R}_0 \geqslant 1$ e se escludiamo il caso banale in cui $p_1 = 1$, allora il lato destro dell'equazione è una funzione strettamente crescente di x. Altrimenti non ci sarebbe nessun $k \geqslant 2$ tale che $p_k \neq 0$ e $\mathscr{R}_0$ sarebbe uguale a $p_1 < 1$. In conclusione, l'equazione ha una e una sola soluzione nell'intervallo $[0, 1]$ quando $\mathscr{R}_0 \geqslant 1$.

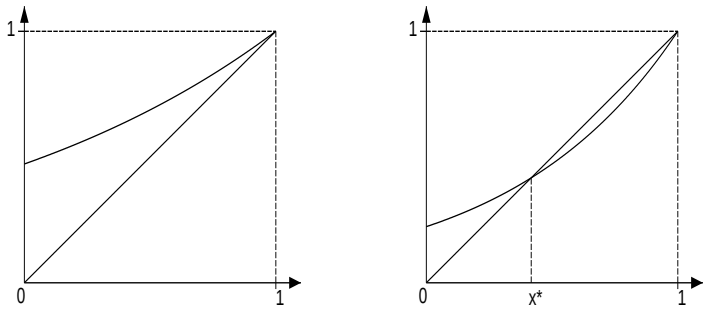

Figura 18.2: Grafico delle funzioni $y = x$ e $y = f(x)$ nell'esempio del capitolo 17, $f(x) = \exp(\mathscr{R}_0(x-1))$, con $\mathscr{R}_0 = 0{,}75 < 1$ [(a)] o $\mathscr{R}_0 = 1{,}5 > 1$ [(b)].

Steffensen, che era anche presidente della Società attuariale danese e della Società matematica danese, fu invitato all'Università di Londra nel 1930. Il suo collega britannico W. P. Elderton gli parlò del lavoro di Galton e Watson. Nel 1933 Steffensen pubblicò un nuovo articolo negli annali dell'Istituto Henri Poincaré, dove aveva tenuto una conferenza nel 1931. Riassunse i risultati del suo articolo in danese e li confrontò con quelli di Watson. Mostrò anche

che l'aspettativa matematica del numero di figli nella generazione n è uguale a $(\mathscr{R}_0)^n$.

Infatti, sia $p_{k,n}$ la probabilità che ci siano k figli nella generazione n, partendo da un individuo nella generazione 0. Nel suo articolo del 1930, Steffensen aveva notato come i suoi predecessori che la funzione generatrice $f_n(x) = \sum_{k=0}^{+\infty} p_{k,n} x^k$ relativa alla generazione n soddisfa $f_1(x) = f(x)$ e $f_n(x) = f(f_{n-1}(x))$. Sia M_n l'aspettativa del numero di figli nella generazione n. Allora $M_n = \sum_{k=1}^{+\infty} k\, p_{k,n} = f_n'(1)$. Otteniamo $f_n'(x) = f'(f_{n-1}(x)) \times f_{n-1}'(x)$. Quindi $M_n = f_n'(1) = f'(f_{n-1}(1)) \times f_{n-1}'(1) = f'(1) \times M_{n-1} = \mathscr{R}_0 \times M_{n-1}$. Poiché $M_1 = f_1'(1) = f'(1) = \mathscr{R}_0$, segue che $M_n = (\mathscr{R}_0)^n$ per tutti gli n.

Quindi il numero atteso di figli aumenta o diminuisce geometricamente a seconda che $\mathscr{R}_0$ sia maggiore o minore di 1. Il numero atteso di figli si comporta come nei modelli deterministici di crescita della popolazione considerati da Eulero, Malthus ecc. Tuttavia, anche quando $\mathscr{R}_0 > 1$, c'è una probabilità non nulla x_∞ che la famiglia si estingua. Questa possibilità non si verifica nei modelli deterministici.

Il processo stocastico studiato da Steffensen e dai suoi predecessori è ancora l'elemento base di molti modelli più realistici di dinamica di popolazione. Citeremo un'ultima volta questo problema nel capitolo 20. Per quanto riguarda Steffensen, rimase professore all'Università di Copenhagen fino al 1943 e morì nel 1961.

Bibliografia

1. Brockmeyer, E., Halstrøm, H.L., Jensen, A.: The life and works of A.K. Erlang. *Trans. Dan. Acad. Techn. Sci.* 2 (1948)
2. Erlang, A.K.: Opgave Nr. 15. *Mat. Tidsskr. B*, 36 (1929) → Guttorp (1995)
3. Guttorp, P.: Three papers on the history of branching processes. *Int. Stat. Rev.* 63, 233–245 (1995) www.stat.washington.edu/research/reports/1992/tr242.pdf
4. Heyde, C.C.: Agner Krarup Erlang. In: Heyde, C.C., Seneta, E. (eds.) *Statisticians of the Centuries*, 328–330. Springer, New York (2001)
5. Ogborn, M.E.: Johan Frederik Steffensen, 1873–1961. *J. R. Stat. Soc. Ser. A* 125, 672–673 (1962)
6. Steffensen, J.F.: Om Sandssynligheden for at Afkommet uddør. *Mat. Tidsskr. B*, 19–23 (1930) → Guttorp (1995)
7. Steffensen, J.F.: Deux problèmes du calcul des probabilités. *Ann. Inst. Henri Poincaré* 3, 319–344 (1933) archive.numdam.org

Capitolo 19

Wright e la deriva genetica casuale (1931)

Nel 1931 il biologo americano Sewall Wright sviluppò lo studio di un modello stocastico nella genetica delle popolazioni, che si basa sugli stessi presupposti della legge di Hardy-Weinberg, tranne che la popolazione non è assunta infinitamente grande. Le frequenze dei genotipi non sono più costanti. Uno dei due alleli scomparirà infatti, ma forse dopo molto tempo. L'interpretazione di questo modello è rimasta oggetto di controversia tra Wright e Fisher, quest'ultimo avendo considerato che la selezione naturale giocasse un ruolo più importante nell'evoluzione rispetto alla stocasticità.

Sewall Wright è nato nel Massachusetts nel 1889. Fece i suoi studi universitari in un piccolo college dell'Illinois dove suo padre insegnava economia. Dopo un master in biologia all'Università di Urbana nell'Illinois e una scuola estiva al *Cold Spring Harbor Laboratory*, Wright conseguì un dottorato all'Università di Harvard con una tesi sull'eredità del colore del mantello della cavia. Tra il 1915 e il 1925, continuò a lavorare su esperimenti di consanguineità con le cavie presso la divisione zootecnica dello Dipartimento dell'Agricoltura degli Stati Uniti a Washington. Sviluppò il «metodo dei coefficienti di percorso» per analizzare questi esperimenti. Poi si unì al dipartimento di zoologia dell'Università di Chicago.

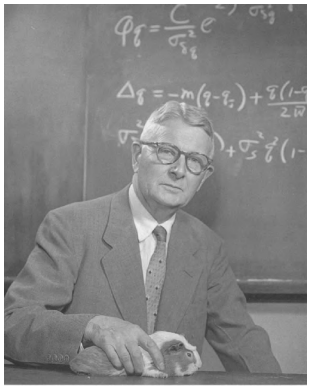

Figura 19.1: Wright (1889–1988)

Influenzato dall'articolo di Fisher del 1922 sulla genetica delle popolazioni (si veda il capitolo 14), Wright scrisse nel 1925 un lungo articolo intitolato *Evoluzione nelle popolazioni mendeliane*, che fu infine pubblicato nel 1931. Egli studiò in particolare un modello matematico che appariva anche implicitamente nel libro di Fisher del 1930 su *La teoria genetica della selezione naturale*. Come nella legge di Hardy–Weinberg, questo modello considera il caso in cui ci sono solo due possibili alleli A e a per un locus, ma non si assume che la popolazione sia infinitamente grande. Il punto è vedere se la rimozione di questa assunzione ha qualche influenza sulla composizione genetica della popolazione. Sia N il numero totale di individui, che si suppone sia lo stesso in tutte le generazioni. Ogni individuo ha due alleli. Quindi c'è un totale di $2N$ alleli nella popolazione in ogni generazione. Nel modello si assume anche che l'accoppiamento avvenga in modo casuale. Se ci sono i alleli A e $2N-i$ alleli a nella generazione n, allora un allele scelto a caso tra gli individui della $n+1$ - esima generazione sarà A con una probabilità $\frac{i}{2N}$ e a con una probabilità $1-\frac{i}{2N}$. Il numero di alleli A nella $n+1$ - esima generazione sarà quindi uguale a j con una probabilità:

$$p_{i,j} = \binom{2N}{j}\left(\frac{i}{2N}\right)^j\left(1-\frac{i}{2N}\right)^{2N-j}, \tag{19.1}$$

dove $\binom{2N}{j}$ è il coefficiente binomiale uguale a $(2N)!/[j!\,(2N-j)!]$. Sia X_n il numero di alleli A nella n-esima generazione, questa è una variabile casuale (figura 19.2). Si può mostrare che l'aspettativa di X_{n+1}, sapendo che $X_n = i$,

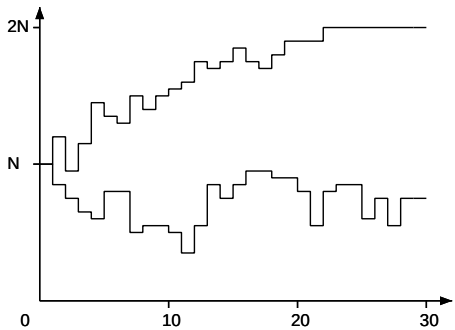

Figura 19.2: Due simulazioni che mostrano le variazioni del numero X_n di alleli A durante 30 generazioni se $N = 20$ e $X_0 = 10$.

è uguale a i: questo ricorda la legge di Hardy–Weinberg, dove la frequenza dell'allele A è rimasta costante attraverso le generazioni.

Infatti, si consideri la funzione generatrice

$$f(x) = \sum_{j=0}^{2N} p_{i,j} x^j = \left(1 - \frac{i}{2N} + \frac{ix}{2N}\right)^{2N},$$

L'aspettativa di X_{n+1} sapendo che $X_n = i$ è allora

$$\sum_{j=0}^{2N} j\, p_{i,j} = f'(1) = i. \tag{19.2}$$

Tuttavia è possibile in questo modello che, partendo da una condizione iniziale $X_0 = i$ con $0 < i < 2N$, l'evento $X_n = 0$ si verifichi casualmente dopo un certo numero di generazioni. In tal caso, tutti gli alleli sarebbero di tipo a e X_n rimarrebbe uguale a 0 in tutte le generazioni future. La stessa situazione si avrebbe con l'allele A se $X_n = 2N$ dopo un certo numero di generazioni. In sintesi, quando la popolazione è assunta infinitamente grande come nel modello di Hardy–Weinberg, i due alleli non possono scomparire perché le loro frequenze rimangono costanti. Quando si tiene conto della dimensione finita delle popolazioni, come nel modello Fisher–Wright, le frequenze dei due alleli fluttuano e uno degli alleli può (e potrà) scomparire.

Partendo da $X_0 = i$, si può facilmente calcolare la probabilità Q_i che la popolazione sia costante nello stato $X = 0$. Infatti, Q_i deve soddisfare le «condizioni limite»

$$Q_0 = 1, \quad Q_{2N} = 0. \tag{19.3}$$

Inoltre,

$$Q_i = \sum_{j=0}^{2N} p_{i,j}\, Q_j\,, \tag{19.4}$$

perché $p_{i,j}\, Q_j$ è la probabilità di essere costante nello stato $X = 0$ partendo da $X_0 = i$ e passando per $X_1 = j$. Dato che

$$\sum_{j=0}^{2N} p_{i,j} = 1,$$

vediamo, usando (19.2), che

$$Q_i = 1 - \frac{i}{2N}$$

è la soluzione del sistema (19.3)-(19.4). Quindi la probabilità che, partendo da i alleli di tipo A in una popolazione di dimensione N, il sistema evolva verso una popolazione contenente solo l'allele a è uguale a $1-\frac{i}{2N}$. Allo stesso modo, la probabilità che evolva verso una popolazione contenente solo l'allele A è uguale a $\frac{i}{2N}$.

Wright è riuscito a dimostrare che il numero di generazioni che trascorrono prima del raggiungimento della stazionarietà in uno dei due stati estremi è dell'ordine di $2N$ generazioni (figura 19.3). Per popolazioni di diversi milioni di individui, questo tempo sarebbe così lungo che le frequenze degli alleli potrebbero essere considerate quasi costanti, come nella legge di Hardy–Weinberg.

Infatti, supponiamo che ci siano i_0 alleli di tipo A nella popolazione alla 0 - esima generazione. Sia $u_i^{(n)}$ la probabilità che ci siano i alleli di tipo A nella popolazione alla n -esima generazione. Allora:

$$u_j^{(n+1)} = \sum_{i=0}^{2N} u_i^{(n)} p_{i,j}$$

per tutti $j = 0,\dots,2N$. Abbiamo già visto che, quando $n \to +\infty$,

$$u_0^{(n)} \to 1-\frac{i_0}{2N}, \quad u_{2N}^{(n)} \to \frac{i_0}{2N}, \quad u_i^{(n)} \to 0$$

per tutti $0 < i < 2N$. Wright ha notato che se $u_i^{(n)} = v$ per tutti $i = 1,\dots,2N-1$, allora:

$$u_j^{(n+1)} = v \binom{2N}{j} \sum_{i=1}^{2N-1} \left(\frac{i}{2N}\right)^j \left(1-\frac{i}{2N}\right)^{2N-j} \tag{19.5}$$

per tutti $1 < j < 2N$ perché $p_{0,j} = p_{2N,j} = 0$. Quando N è abbastanza grande,

$$\begin{aligned} \frac{1}{2N} \sum_{i=1}^{2N-1} \left(\frac{i}{2N}\right)^j \left(1-\frac{i}{2N}\right)^{2N-j} &\approx \int_0^1 x^j (1-x)^{2N-j}\, dx \\ &= \frac{j!\,(2N-j)!}{(2N+1)!}, \end{aligned} \tag{19.6}$$

il valore dell'integrale è ottenuto da successive integrazioni per parti. Combinando (19.5) e (19.6), si arriva infine per $0 < j < 2N$ a:

$$u_j^{(n+1)} \approx \frac{2N}{2N+1} v = \left(1 - \frac{1}{2N+1}\right) u_j^{(n)} .$$

Quindi le probabilità $u_j^{(n)}$ per tutti $0 < j < 2N$ diminuiscono ad un tasso di circa $1/2N$ per generazione. Questo tasso è molto lento se N è grande. Non c'è quasi nessuna diminuzione se, per esempio, N è dell'ordine di grandezza di milioni.

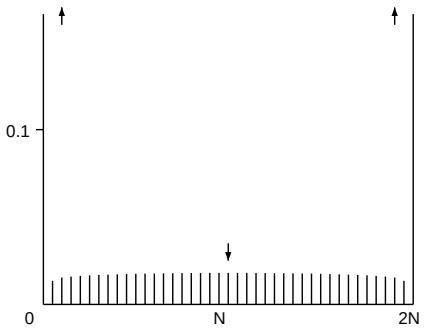

Figura 19.3: Probabilità che ci siano i alleli A nella popolazione ($i = 0, \ldots, 2N$ sull'asse orizzontale) dopo 30 generazioni se $N = 20$ e $X_0 = 10$.

Nel 1922 Fisher aveva già cercato di stimare questo tasso costante ($1/2N$) ma aveva omesso un fattore 2. In ogni caso, i due scienziati non erano d'accordo sulla dimensione tipica N delle popolazioni riproduttrici. Per la teoria dell'evoluzione, il lavoro di Wright suggeriva che la deriva genetica casuale in una piccola popolazione poteva essere un meccanismo per l'origine delle specie. I biologi che lavoravano alla classificazione delle specie avevano infatti notato che le differenze tra specie o sottospecie spesso non avevano alcuna spiegazione apparente in termini di selezione naturale. Questa idea fu fortemente contrastata negli anni '40 e '50 da Fisher e dal suo collega E. B. Ford, ritenendo che la deriva genetica casuale fosse trascurabile rispetto alla selezione naturale. Si riferirono in particolare al loro studio sulle fluttuazioni delle frequenze geniche in una piccola popolazione isolata di falene (*Panaxia dominula*) vicino a Oxford, dove i tre genotipi per un certo gene

(omozigote comune, eterozigote e omozigote raro) potevano essere distinti a vista. Un'altra famosa controversia sulla rispettiva influenza della selezione naturale e della deriva casuale si è concentrata sulle lumache del genere *Cepaea*. Modelli più realistici di evoluzione ora combinano la deriva casuale, selezione, mutazione, migrazione, accoppiamento non casuale, ecc. Il ruolo della deriva casuale è stato successivamente ribadito dallo scienziato giapponese Motoo Kimura con la sua «teoria neutrale dell'evoluzione molecolare». Un'altra conseguenza fu lo sviluppo della teoria della coalescenza (introdotta da John Kingman nel 1982), che traccia l'ascendenza dei geni a ritroso nel tempo fino al punto in cui hanno un unico antenato comune.

Wright divenne membro dell'Accademia Nazionale delle Scienze nel 1934. Ha lavorato per molti anni con Theodosius Dobzhansky sulla genetica delle popolazioni naturali di mosche (*Drosophila pseudoobscura*) nella regione della Valle della Morte. Si ritirò dall'Università di Chicago nel 1955 ma continuò altri cinque anni come professore all'Università del Wisconsin-Madison. Tra il 1968 e il 1978 pubblicò un trattato in quattro volumi che riassumeva il suo lavoro *Evoluzione e genetica delle popolazioni*. Ha ricevuto il Premio Balzan nel 1984 ed è morto nel 1988 all'età di 98 anni.

Bibliografia

1. Fisher, R.A.: *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford (1930) archive.org
2. Hill, W.G.: Sewall Wright, 21 December 1889–3 March 1988. *Biog. Mem. Fellows R. Soc.* 36, 568–579 (1990)
3. Kimura, M.: *The Neutral Theory of Molecular Evolution*. Cambridge University Press (1983)
4. Provine, W.B.: *Sewall Wright and Evolutionary Biology*. University of Chicago Press (1989)
5. Wright, S.: Evolution in Mendelian populations. *Genetics* 16, 97–159 (1931) www.esp.org
6. Wright, S.: *Evolution and the Genetics of Populations, Vol. 2, Theory of Gene Frequencies*. University of Chicago Press (1969)

Capitolo 20

La diffusione dei geni (1937)

Nel 1937 Ronald Fisher e tre matematici russi, Kolmogorov, Petrovsky e Piskunov, studiarono indipendentemente un'equazione alle derivate parziali per la diffusione geografica di un gene vantaggioso. Mostrarono che la frequenza del gene si comportava come un'onda che viaggiava ad una velocità ben definita che dipendeva dal vantaggio del gene e da un coefficiente di diffusione. I loro lavori furono il punto di partenza per la teoria delle equazioni di reazione-diffusione.

Nel 1937 furono pubblicati due articoli che introducevano un nuovo approccio allo studio dell'eterogeneità spaziale nella dinamica della popolazione. Fisher fu l'autore del primo articolo, intitolato *L'onda di avanzamento dei geni vantaggiosi*, che apparve negli *Annals of Eugenics*. Studiò la propagazione spaziale di un gene favorevole in una popolazione. Come semplificazione, considerò uno spazio ridotto ad una sola dimensione e chiamò $u(x,t)$ la frazione della popolazione situata nel punto x al tempo t che possiede il gene favorevole. Perciò $0 \leqslant u(x,t) \leqslant 1$. Per includere la selezione naturale, ha usato l'equazione (14.6) con una variabile temporale continua

$$\frac{\partial u}{\partial t} = a\,u\,(1-u)\ ,$$

dove a è un parametro positivo. Per un dato valore di x, riconosciamo l'equazione logistica di Verhulst (vedi capitolo 6) con una soluzione $u(x,t)$ che tende a 1 quando $t \to +\infty$. Inoltre, Fisher suppose che la prole di un individuo situato nel punto x con il gene favorevole non rimane nello stesso punto ma si disperde casualmente nelle vicinanze di x. Per analogia con la fisica, ha sostenuto che si deve aggiungere un termine di diffusione all'equazione per $u(x,t)$, ottenendo l'equazione alle derivate parziali

$$\frac{\partial u}{\partial t} = a\,u\,\big(1-u\big) + D\,\frac{\partial^2 u}{\partial x^2}\ . \tag{20.1}$$

Quando il coefficiente di selezione a è zero, questa si riduce all'equazione di diffusione introdotta da Fourier nella sua teoria del calore e poi usata da Fick per la diffusione delle particelle fisiche. Nel 1904, Ronald Ross aveva iniziato a considerare la dispersione casuale nella dinamica delle popolazioni. Si

chiedeva allora come la densità delle zanzare diminuisse all'aumentare della distanza da un luogo di riproduzione. Il problema era giunto all'attenzione di Karl Pearson e Lord Rayleigh. Nel 1937 il numero delle pubblicazioni scientifiche sulle equazioni di diffusione era cresciuto considerevolmente, in particolare dopo il lavoro di Einstein sul moto browniano.

Fisher dimostrò che esistono soluzioni dell'equazione (20.1) della forma $u(x,t) = U(x+vt)$ che soddisfano le tre condizioni

$$0 \leqslant u(x,t) \leqslant 1, \quad u(x,t) \underset{x\to-\infty}{\longrightarrow} 0, \quad u(x,t) \underset{x\to+\infty}{\longrightarrow} 1,$$

a condizione che $v \geqslant v^*$ dove

$$v^* = 2\sqrt{aD}\,.$$

Queste soluzioni collegano lo stato stazionario $u = 1$ con il gene favorevole allo stato stazionario $u = 0$ senza tale gene. Rappresentano onde che si propagano a velocità v nella direzione di valori decrescenti di x. Infatti, $u(x-vT,t+T) = u(x,t)$: la parte dell'onda che era nella posizione x al tempo t si sposta nella posizione $x - vT$ al tempo $t+T$.

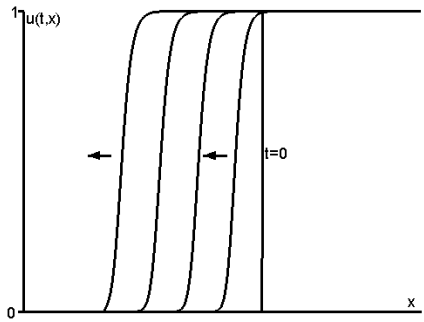

Figura 20.1: Propagazione da destra a sinistra di un gene favorevole alla velocità v^*. La frequenza del gene $u(t,x)$ a $t = 0$ è una funzione gradino.

Infatti, ponendo $z = x+v,t$, Fisher notò che se $u(x,t) = U(z)$, allora

$$\frac{\partial u}{\partial t} = vU'(z), \quad \frac{\partial u}{\partial x} = U'(z), \quad \frac{\partial^2 u}{\partial x^2} = U''(z).$$

Se u è una soluzione dell'equazione (20.1), allora

$$vU'(z) = aU(z)(1-U(z)) + DU''(z) \,. \tag{20.2}$$

Quando u è vicino a 0, cioè quando $z \to -\infty$, Fisher si aspetta che $U(z) \to 0$ e $U'(z) \to 0$. Chiamando k il limite di $U'(z)/U(z)$ quando $z \to -\infty$, sappiamo dalla regola di L'Hôpital che anche $U''(z)/U'(z)$ tende a k. Pertanto, $U''(z)/U(z) = [U''(z)/U'(z)] \times [U'(z)/U(z)]$ tende a k^2. Dividendo l'equazione (20.2) per $U(z)$ e lasciando che z tenda a $-\infty$, si arriva a un'equazione del secondo ordine $Dk^2 - vk + a = 0$. Ma k deve essere un numero reale. Quindi il discriminante di questa equazione deve essere non negativo: $v^2 - 4aD \geqslant 0$, o $v \geqslant 2\sqrt{aD} = v^*$. Quindi, $v \geqslant v^*$ è una condizione necessaria per l'esistenza di un'onda che si propaga alla velocità v. È anche una condizione sufficiente, come spiegato di seguito.

Fisher ha notato che solo l'onda che si propaga esattamente alla velocità v^* è selezionata per una grande classe di condizioni iniziali, ad esempio, per la funzione gradino: $u(x,0) = 0$ per $x < 0$, $u(x,0) = 1$ per $x > 0$. La figura 20.1 mostra come questa condizione iniziale discontinua diventi progressivamente un'onda liscia che si propaga nella direzione della diminuzione di x alla velocità v^*.

Nello stesso anno 1937, e indipendentemente dal lavoro di Fisher, Andrej Nikolaevič Kolmogorov, Ivan Georgievič Petrovsky e Nikolaj Semenovič Piskunov studiarono lo stesso problema della propagazione di un gene dominante.

Kolmogorov è nato nel 1903 a Tambov, in Russia. Durante i suoi studi di matematica all'Università Statale di Mosca, fece un lavoro importante sulle serie trigonometriche. Divenne ricercatore all'Istituto di Matematica e Meccanica nel 1929 e professore universitario nel 1931. Lavorò sui processi stocastici e sul loro legame con le equazioni differenziali ordinarie ed alle derivate parziali. Nel 1933 pubblicò un trattato che poneva le basi moderne della teoria della probabilità. I suoi interessi di ricerca includevano la topologia, la teoria dell'approssimazione, le catene di Markov, il moto browniano e anche le applicazioni ai problemi biologici. Nel 1935 pubblicò un articolo sulla genetica discutendo i risultati di Hardy, Fisher e Wright. Nel 1936 pubblicò un articolo su una generalizzazione del sistema Lotka-Volterra.

Petrovsky è nato nel 1901 a Sevsk. Studiò matematica all'Università Statale di Mosca, dove divenne professore nel 1933. Lavorò principalmente sulla teoria delle equazioni differenziali parziali e sulla topologia delle curve algebriche reali, ma scrisse anche alcuni articoli sulle equazioni differenziali

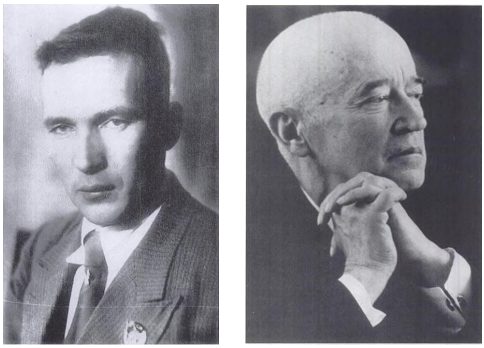

Figura 20.2: Kolmogorov (1903–1987) e Petrovsky (1901–1973)

ordinarie e sulla teoria della probabilità. Piskunov, nato nel 1908, era un altro ex studente di matematica all'Università Statale di Mosca.

Durante gli anni '30 Kolmogorov ebbe contatti con A. S. Serebrovsky, un pioniere della genetica delle popolazioni a Mosca. Stava diventando sempre più pericoloso difendere la genetica mendeliana in URSS a causa dell'ascesa di Lysenko, un agronomo che era riuscito a convincere Stalin che la genetica mendeliana era solo «pseudoscienza borghese». Il Settimo Congresso Internazionale di Genetica, originariamente previsto per il 1937 a Mosca, fu cancellato. Molti genetisti sovietici furono giustiziati o mandati nei campi di lavoro.

Nel loro articolo del 1937 intitolato *Uno studio dell'equazione di diffusione con aumento della quantità di materia e la sua applicazione a un problema biologico*, pubblicato nel Bollettino dell'Università Statale di Mosca, Kolmogorov, Petrovsky e Piskunov utilizzarono tuttavia un modello matematico basato sulla genetica mendeliana. Il loro modello si basava su di un'equazione alle derivate parziali della forma

$$\frac{\partial u}{\partial t} = f(u) + D\,\frac{\partial^2 u}{\partial x^2} \tag{20.3}$$

dove $u(x,t)$ è di nuovo la frequenza del gene favorevole nel punto x e nel tempo t. Si suppone che la funzione $f(u)$ soddisfi diverse condizioni: $f(0) = f(1) = 0$, $f(u) > 0$ se $0 < u < 1$, $f'(0) > 0$ e $f'(u) < f'(0)$ se $0 < u \leqslant 1$. Gli autori hanno ottenuto un risultato che è analogo a quello di Fisher ma con una dimostrazione più rigorosa: se la condizione iniziale è tale che $0 \leqslant u(x,0) \leqslant 1$, $u(x,0) = 0$ per tutti $x < x_1$ e $u(x,0) = 1$ per tutti $x > x_2 \geqslant x_1$, allora il gene si propaga alla velocità $v^* = 2\sqrt{f'(0)\,D}$.

Cercare una soluzione $u(x,t) = U(z)$ dove $z = x + vt$ porta all'ovvia generalizzazione dell'equazione (20.2) $vU'(z) = f(U(z)) + DU''(z)$. Questa equazione differenziale del secondo ordine può essere riscritta come un sistema di equazioni differenziali del primo ordine

$$\frac{dU}{dz} = p, \quad \frac{dp}{dz} = \frac{vp - f(U)}{D}. \tag{20.4}$$

Ricordiamo che $U(z)$ dovrebbe essere tale che $U(z) \to 0$ per $z \to -\infty$ e $U(z) \to 1$ per $z \to +\infty$. Vicino allo stato stazionario $(U = 0, p = 0)$ del sistema (20.4), abbiamo $f(U) \approx f'(0)U$. Quindi (20.4) può essere approssimato dal sistema lineare

$$\frac{dU}{dz} = p, \quad \frac{dp}{dz} = \frac{vp - f'(0)U}{D}. \tag{20.5}$$

Cercando soluzioni esponenziali della forma $U(z) = U_0 \mathrm{e}^{kz}$ e $p(z) = p_0 \mathrm{e}^{kz}$ si ottiene l'equazione caratteristica $Dk^2 - vk + f'(0) = 0$, come nell'articolo di Fisher. Ancora una volta k deve essere reale (altrimenti u oscillerebbe e assumerebbe valori negativi). Perciò $v \geqslant 2\sqrt{f'(0)D} = v^*$. Le due radici per k sono quindi reali e positive. Se $v > v^*$, le due radici sono diverse e lo stato stazionario $(U = 0, p = 0)$ è un nodo instabile. Se $v = v^*$, le due radici sono identiche e $(U = 0, p = 0)$ è un nodo degenere instabile come mostrato nella figura 20.3. Allo stesso modo, il sistema (20.4) vicino allo stato stazionario $(U = 1, p = 0)$ porta al sistema lineare

$$\frac{d(U-1)}{dz} = p, \quad \frac{dp}{dz} = \frac{vp - f'(1)(U-1)}{D}$$

e all'equazione caratteristica $Dk^2 - vk + f'(1) = 0$. Il discriminante è $v^2 - 4Df'(1) \geqslant 0$ poiché $f'(1) \leqslant 0$. Se $f'(1) < 0$, ci sono due radici reali di segno opposto e $(U = 1, p = 0)$ è un punto di sella. Se $f'(1) = 0$, una radice è zero e l'altra è positiva (vedi figura 20.3). Un'analisi dettagliata mostra infatti che per tutti i $v \geqslant 2\sqrt{f'(0)D}$ esiste un'unica curva integrale che unisce i due stati stazionari $(U = 0, p = 0)$ e $(U = 1, p = 0)$, come nel caso speciale della figura 20.3.

Kolmogorov, Petrovsky e Piskunov hanno poi dimostrato rigorosamente che l'equazione alle derivate parziali (20.3) ha una soluzione unica $u(x,t)$ che soddisfa la condizione iniziale, che questa soluzione è tale che $0 < u(x,t) \leqslant 1$ per tutti gli x e $t > 0$, che $u(x,t)$ rimane una funzione

crescente di x se lo è a $t = 0$ e infine che $u(x,t)$ converge verso un profilo d'onda che si propaga alla velocità v^*. Le dimostrazioni sono troppo lunghe per essere riassunte qui.

Si noti che la funzione $f(u) = a\,u\,(1-u)$ usata da Fisher soddisfa tutte queste condizioni con $f'(0) = a$. Ispirandosi all'equazione (14.5), Kolmogorov, Petrovsky e Piskunov hanno considerato la funzione $f(u) = a\,u\,(1-u)^2$, che soddisfa le stesse condizioni e dà la stessa velocità di propagazione.

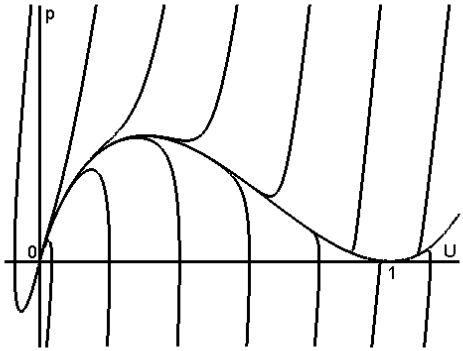

Figura 20.3: Diagramma (U,p) che mostra alcune curve integrali del sistema (20.5) e in particolare la curva unica che unisce $(U = 1, p = 0)$ a $(U = 0, p = 0)$, che è quella che dà la forma dell'onda di propagazione. Qui, $f(u) = au(1-u)^2$, $a = 1$, $D = 1$ e $v = v^* = 2$.

Gli articoli di Fisher e di Kolmogorov, Petrovsky e Piskunov furono il punto di partenza per la costruzione di molti modelli matematici a diffusione geografica in genetica, ecologia ed epidemiologia. Questi modelli sono conosciuti come «sistemi di reazione-diffusione».

Per quanto riguarda Kolmogorov, a partire dal 1938 studiò anche il problema dell'estinzione dei cognomi considerato da Bienaymé, Galton, Watson, Fisher, Haldane, Erlang e Steffensen: chiamò il processo stocastico che è comune a tutti questi lavori il «processo di ramificazione». Nel 1939 divenne membro dell'Accademia delle Scienze dell'URSS. In seguito diede importanti contributi al problema della turbolenza nella meccanica dei fluidi (1941), alla teoria dei sistemi dinamici legati alla meccanica celeste (1953) e alla teoria dell'informazione (a partire dal 1956). Ha anche contribuito alla stesura di

un'enciclopedia e di libri di testo per scuole superiori e università, a fondare una scuola superiore sperimentale e ha curato una rivista di scienze popolari. Ha ricevuto molti premi internazionali (tra cui il premio Balzan nel 1963 e il premio Wolf nel 1980) ed è morto a Mosca nel 1987.

Petrovskij divenne preside della Facoltà di Meccanica e Matematica dell'Università Statale di Mosca nel 1940. Fu rettore dell'università dal 1951 fino alla sua morte nel 1973. Fu membro a pieno titolo dell'Accademia delle Scienze dell'URSS dal 1946 e presidente del Congresso Internazionale dei Matematici che si tenne a Mosca nel 1966. Scrisse anche libri di testo sulle equazioni differenziali ordinarie, sulle equazioni alle derivate parziali e sulle equazioni integrali. Piskunov divenne professore in un'accademia militare. Il suo libro di testo sul calcolo differenziale e integrale fu usato in molti politecnici. Morì nel 1977.

Bibliografia

1. Fisher, R.A.: The wave of advance of advantageous genes. *Ann. Eugen.* 7, 355–369 (1937) digital.library.adelaide.edu.au
2. Kolmogoroff, A.N., Petrovsky, I.G., Piscounoff, N.S.: Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique. *Bull. Univ. État Moscou, Ser. Int., Sect. A, Math. et Mecan.* 1(6):1–25 (1937). → V.M. Tikhomirov (ed.) *Selected Works of A. N. Kolmogorov*, vol. 1, 242–270. Kluwer, Dordrecht (1991). → *I. G. Petrowsky Selected Works*, Part II, 106–132. Gordon and Breach, Amsterdam (1996)
3. Oleinik, O.A.: I.G. Petrowsky and modern mathematics. In: *I. G. Petrowsky Selected Works*, Part I, 4–30. Gordon and Breach, Amsterdam (1996)
4. Pearson, K.: *Mathematical Contributions to the Theory of Evolution, XV, A Mathematical Theory of Random Migration*. Dulau, London (1906) archive.org
5. Rosenfeld, B.A.: Reminiscences of Soviet Mathematicians. In: Zdravkovska, S., Duren, P.L. (eds.) *Golden Years of Moscow Mathematics*, 2nd edn., 75–100. American Mathematical Society (2007)
6. Shiryaev, A.N. (ed.): *Selected Works of A. N. Kolmogorov*, vol. 2. Kluwer, Dordrecht (1992)
7. Shiryaev, A.N.: Andrei Nikolaevich Kolmogorov (April 25, 1903 to October 20, 1987). In: *Kolmogorov in Perspective*, 1–88. American Mathematical Society (2000)

Capitolo 21

La matrice di Leslie (1945)

Nel 1945 l'ecologista britannico P.H. Leslie analizzò un modello a matrice per una popolazione di roditori strutturata per età, adattando così il lavoro di Lotka al caso di tempo discreto. Sottolineò che il tasso di crescita corrisponde a un autovalore e la struttura di età stabile a un autovettore. Inoltre stimò numericamente il tasso netto di riproduzione R_0 per il ratto marrone.

Patrick Holt Leslie nacque nel 1900 nei pressi di Edimburgo, in Scozia, e studiò al *Christ Church College* dell'Università di Oxford, ottenendo nel 1921 una laurea in fisiologia. Non poté finire i suoi studi di medicina a causa di problemi di salute. Dopo alcuni anni di lavoro come assistente di batteriologia nel dipartimento di patologia, si dedicò alla statistica ed entrò nel 1935 nel *Bureau of Animal Population*, un nuovo centro di ricerca creato da Charles Elton. Lo scopo di questo centro era quello di studiare le fluttuazioni delle popolazioni animali attraverso studi sul campo ed esperimenti di laboratorio. La maggior parte delle ricerche furono fatte sui roditori. Si occupò dell'analisi dei cicli della lepre e del suo predatore, la lince, utilizzando gli archivi della Compagnia della Baia di Hudson in Canada, del controllo dell'espansione territoriale dello scoiattolo grigio a spese dello scoiattolo rosso in Inghilterra, della raccolta di dati sulle arvicole nei dintorni di Oxford e così via. Leslie applicò ai dati sulle arvicole i metodi sviluppati da Lotka per la demografia umana. Durante la seconda guerra mondiale, la ricerca del centro si concentrò sui metodi di controllo dei ratti e dei topi nei silos.

Nel 1945 Leslie pubblicò il suo articolo più famoso su *Biometrika*, una rivista che era stata fondata da Galton, Pearson e Weldon nel 1901. L'articolo era intitolato *Sull'uso delle matrici in alcune matematiche di popolazione*. In esso Leslie considerò un modello per la crescita del numero di femmine in una popolazione animale, per esempio una popolazione di ratti (ma potrebbe essere anche umana). La popolazione è divisa in $K+1$ gruppi di età. Il numero di femmine di età k al tempo n ($k = 0,1,\ldots,K$; $n = 0,1,\ldots$) è rappresentato da $P_{k,n}$. Chiamiamo f_k la fertilità all'età k o più precisamente il numero di figlie nate per femmina tra il tempo n e il tempo $n+1$. Quindi K è l'età massima con fertilità non nulla ($f_K > 0$). Chiamiamo s_k la probabilità per un animale di età k di sopravvivere almeno fino all'età $k+1$. Allora la

Figura 21.1:
P. H. Leslie (1900–1972)

struttura per età della popolazione è data dalla seguente serie di equazioni:

$$\begin{cases} P_{0,n+1} &= f_0 P_{0,n} + f_1 P_{1,n} + \cdots + f_K P_{K,n} \\ P_{1,n+1} &= s_0 P_{0,n} \\ P_{2,n+1} &= s_1 P_{1,n} \\ \vdots & \quad \vdots \\ P_{K,n+1} &= s_{K-1} P_{K-1,n} \,. \end{cases}$$

Tutti i numeri f_k sono non negativi, mentre s_k soddisfa $0 < s_k < 1$. Tra la fine del diciannovesimo e l'inizio del ventesimo secolo, i matematici avevano preso l'abitudine di scrivere tali sistemi di equazioni nella forma abbreviata[1]

$$P_{n+1} = M P_n \,, \tag{21.1}$$

dove P_n è il vettore di colonne $(P_{0,n}, \ldots, P_{K,n})$ e M è la matrice quadrata (cioè la tabella dei numeri con $K+1$ righe e $K+1$ colonne)

$$M = \begin{pmatrix} f_0 & f_1 & f_2 & \cdots & f_K \\ s_0 & 0 & 0 & \cdots & 0 \\ 0 & s_1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & s_{K-1} & 0 \end{pmatrix} .$$

Per capire il comportamento del sistema (21.1) in funzione del tempo, Leslie cercò una soluzione geometricamente crescente o decrescente $P_n = r^n V$. Il numero r e il vettore V devono soddisfare

$$M V = r V \,. \tag{21.2}$$

[1] Significa che $P_{k,n+1} = M_{k,0} P_{0,n} + M_{k,1} P_{1,n} + \cdots + M_{k,K} P_{K,n}$ for all k.

In questo caso, r è chiamato un autovalore e V un autovettore della matrice M. In altre parole, il problema è trovare la distribuzione di età V che ad ogni passo temporale è moltiplicata per una costante r. Seguendo la terminologia di Lotka, tali distribuzioni sono chiamate stabili. Tornando a notazioni più usuali, l'equazione (21.2) può essere riscritta come

$$\begin{cases} f_0 V_0 + f_1 V_1 + \cdots + f_K V_K = r V_0, \\ s_0 V_0 = r V_1, \quad s_1 V_1 = r V_2, \quad \ldots, \quad s_{K-1} V_{K-1} = r V_K \,. \end{cases}$$

Dalle ultime K equazioni segue che

$$V_1 = \frac{s_0 V_0}{r}, \quad V_2 = \frac{s_0 s_1 V_0}{r^2}, \quad \ldots \; V_K = \frac{s_0 s_1 \cdots s_{K-1} V_0}{r^K} \,.$$

Sostituendo queste quantità nella prima equazione, semplificando per V_0 e moltiplicando per r^K, Leslie ottenne l'equazione caratteristica

$$r^{K+1} = f_0 r^K + s_0 f_1 r^{K-1} + s_0 s_1 f_2 r^{K-2} + \cdots + s_0 s_1 \cdots s_{K-1} f_K \,. \tag{21.3}$$

Questa è un'equazione polinomiale in r di grado $K+1$. Quindi ci sono $K+1$ radici reali o complesse $r_1, \ldots, r_{K+1}$. Inoltre Leslie notò (usando la regola del segno di Cartesio per i polinomi) che c'è solo una radice reale positiva (che chiameremo r_1).

Leslie suggerì inoltre che, nella maggior parte delle condizioni biologicamente realistiche (che possono essere precisate usando la teoria di Perron e Frobenius per le matrici non negative), l'autovalore r_1 è strettamente più grande del modulo di tutti gli altri autovalori reali o complessi (che chiameremo $r_2, \ldots, r_{K+1}$). Inoltre, tutte le radici di (21.3) sono solitamente diverse. Per ogni autovalore r_i si può trovare un autovettore associato. Sia Q la matrice quadrata di dimensione $K+1$ le cui colonne $K+1$ contengono rispettivamente gli autovettori associati a $r_1, \ldots, r_{K+1}$, allora $MQ = QD$, dove D è la matrice diagonale $[r_1, \ldots, r_{K+1}]$. Quindi $M = QDQ^{-1}$ e

$$P_n = M^n P_0 = Q D^n Q^{-1} P_0 \,.$$

Si noti che D^n è la matrice diagonale $[(r_1)^n, \ldots, (r_{K+1})^n]$ e che

$$D^n/(r_1)^n \longrightarrow \mathscr{D} = [1, 0, \ldots, 0]$$

quando $n \to +\infty$ perché $r_1 > |r_i|$ per $i \neq 1$. Pertanto, $P_n/(r_1)^n$ converge verso $Q \mathscr{D} Q^{-1} P_0$.

Ogni componente del vettore età-struttura P_n aumenta o diminuisce come $(r_1)^n$. Se $r_1 > 1$, allora la popolazione aumenta esponenzialmente. Se $r_1 < 1$, allora diminuisce esponenzialmente. Dall'equazione (21.3), si può facilmente dimostrare che la condizione $r_1 > 1$ è vera se e solo se il parametro $\mathscr{R}_0$, definito da

$$\mathscr{R}_0 = f_0 + s_0 f_1 + s_0 s_1 f_2 + \cdots + s_0 s_1 \cdots s_{K-1} f_K \ ,$$

è strettamente più grande di 1. Si noti che $s_0 s_1 \cdots s_{k-1}$ è la probabilità di sopravvivere almeno fino all'età k. Quindi il parametro $\mathscr{R}_0$ è il numero medio di figlie nate da una femmina durante la sua vita ed è analogo alle formule (10.2), (12.2) e (16.9). Il presente modello è una sorta di analogo in tempo discreto del lavoro di Lotka (vedi capitolo 10) e una generalizzazione che include la fertilità dipendente dall'età del lavoro di Eulero (vedi capitolo 3).

Leslie illustrò il suo metodo usando i dati pubblicati da un collega americano sui coefficienti di fertilità e sopravvivenza f_k e s_k per il ratto marrone. Dopo alcune operazioni statistiche per completare i dati in modo ragionevole, ottenne $\mathscr{R}_0 \approx 26$.

La formulazione dei problemi in dinamica delle popolazioni attraverso la matrice di Leslie è ora usata da molti biologi. I calcoli sono notevolmente semplificati dai moderni computer e dai software scientifici che possono calcolare autovalori e autovettori di qualsiasi matrice. Si può facilmente calcolare sia il parametro $\mathscr{R}_0$ che il tasso di crescita r_1.

Dopo la seconda guerra mondiale, Leslie usò il suo metodo per calcolare il tasso di crescita di altre specie animali quali uccelli, coleotteri, ecc. Lavorò anche su modelli stocastici, su modelli di competizione tra specie e sull'analisi dei dati di cattura. Si ritirò nel 1967. Lo stesso anno, essendosi ritirato anche Charles Elton, il *Bureau of Animal Population* cessò di esistere come centro di ricerca indipendente e divenne parte del Dipartimento di Zoologia dell'Università di Oxford. Leslie morì nel 1972.

Bibliografia

1. Anonimo: Dr P. H. Leslie. *Nature* 239, 477–478 (1972)
2. Crowcroft, P.: *Elton's Ecologists, A History of the Bureau of Animal Population*. University of Chicago Press (1991)
3. Leslie, P.H.: On the use of matrices in certain population mathematics. *Biometrika* 33, 213–245 (1945)

Capitolo 22

Percolazione ed epidemie (1957)

Nel 1957 Hammersley e Broadbent considerarono la propagazione di un «fluido» in un reticolo regolare infinito, dove due nodi vicini sono collegati con una data probabilità. Tra gli esempi possibili, citarono la propagazione di un'epidemia in un frutteto. Mostrarono che c'è una probabilità critica al di sotto della quale non si verificano grosse epidemie e al di sopra della quale si manifestano grandi epidemie con una probabilità positiva. Il loro articolo fu il punto di partenza della teoria della percolazione.

John Michael Hammersley è nato nel 1920 in Scozia, dove suo padre lavorava per una società americana che esportava acciaio. Iniziò a studiare all'*Emmanuel College* dell'Università di Cambridge, ma dovette arruolarsi nell'esercito nel 1940. Lavorò al miglioramento dei calcoli per l'artiglieria. Dopo aver terminato i suoi studi nel 1948, divenne assistente all'Università di Oxford nel gruppo che lavorava sulla progettazione e l'analisi degli esperimenti. Nel 1955, si unì all'Istituto di ricerca sull'energia atomica di Harwell, vicino a Oxford.

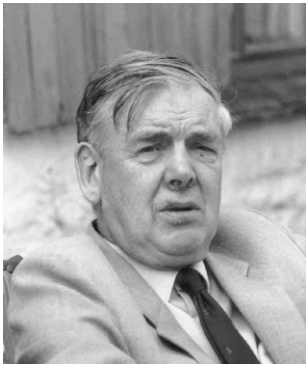

Figura 22.1:
Hammersley (1920–2004)

Simon Ralph Broadbent è nato nel 1928. Ha studiato ingegneria a Cambridge, matematica al Magdalen College di Oxford (dove ha anche scritto poesie) e ha iniziato un dottorato in statistica all'Imperial College di Londra su *Alcuni test di allontanamento dalla dispersione uniforme*. Durante il suo

dottorato ottenne un finanziamento dall'Associazione britannica per la ricerca sull'utilizzo del carbone, che gli permise di studiare problemi statistici legati alla produzione di carbone.

Nel 1954 un simposio sui metodi Monte Carlo, sponsorizzato dall'Istituto di ricerca sull'energia atomica, si tenne alla *Royal Statistical Society* di Londra. Questi metodi, iniziati negli anni 40 da John von Neumann, Stanisław Ulam e Nicholas Metropolis al Laboratorio di Los Alamos, impiegano simulazioni stocastiche al computer per stimare quantità matematiche sconosciute. Hammersley presentò al simposio di Londra un documento che aveva preparato in collaborazione con Morton, un collega di Harwell. L'articolo fu anche pubblicato nel *Journal of the Royal Statistical Society*. Durante la discussione che seguì la presentazione al simposio, Broadbent menzionò un problema interessante che avrebbe potuto essere studiato usando un metodo Monte Carlo: data una rete regolare di pori in due o tre dimensioni tale che due pori vicini sono collegati con una probabilità p, quale frazione di pori sarebbe stata riempita da un gas se fosse stato introdotto attraverso uno di questi pori? Broadbent stava infatti pensando alla progettazione delle maschere antigas per i minatori di carbone e in particolare alla dimensione dei pori che era necessaria per il loro funzionamento.

Hammersley iniziò così a lavorare con Broadbent su questo problema. Si resero conto che era solo un caso particolare di una famiglia di problemi che non era ancora stata studiata: la propagazione deterministica di un «fluido» (il significato dipende dal contesto) in un mezzo casuale. Hammersley lo chiamò «percolazione», per analogia con quello che succede in una caffettiera. All'Istituto di ricerca sull'energia atomica, Hammersley aveva anche accesso ad alcuni dei più potenti computer del suo tempo, questo gli permise di testare i metodi Monte Carlo sui problemi di percolazione.

Nel 1957 Broadbent e Hammersley pubblicarono finalmente il primo articolo sulla teoria matematica della percolazione. Tra gli esempi considerati, uno era un modello di dinamica di popolazione, cioè la propagazione di un'epidemia in un frutteto. Si suppone che gli alberi di un frutteto molto grande siano posti ai nodi di un reticolo. Ognuno dei quattro alberi più vicini ad un dato albero infetto ha una probabilità p di essere anch'esso infetto. La domanda è se un gran numero di alberi sarà infettato o se l'epidemia rimarrà localizzata. Questo dipende naturalmente dalla probabilità p, che a sua volta è legata alla distanza che separa gli alberi, cioè la larghezza della maglia del reticolo.

Broadbent e Hammersley esaminarono il caso limite in cui il frutteto è infinito e copre l'intero piano, con un solo albero infetto all'inizio. Sia $f(p)$ la probabilità che un numero infinito di alberi venga infettato, ci si aspetta che

$f(p)$ sia una funzione crescente di p con $f(0) = 0$ e $f(1) = 1$. Il loro risultato principale è che esiste una probabilità critica p^*, $0 < p^* < 1$, tale che:

- se $p < p^*$, allora $f(p) = 0$ quindi solo un numero finito di alberi è infetto;
- se $p > p^*$, allora $f(p) > 0$ e un numero infinito di alberi può essere infettato.

La dimostrazione implica un confronto con il numero di diversi «cammini auto-evitanti» nel piano a partire dalla sorgente di infezione. Questi cammini passano attraverso un certo numero di alberi vicini (ricordiamo che ogni albero ne ha quattro) senza passare per lo stesso albero più di una volta. Un cammino auto-evitante di n passi ha una probabilità p^n di essere un percorso di infezione, poiché l'infezione può essere trasmessa da ogni albero visitato al successivo con una probabilità p. Ora, sia $q(j,n)$ la probabilità che, tra tutti i cammini auto-evitanti di n passi, ce ne siano esattamente j tra di essi che sono percorsi di infezione. Se esiste un numero infinito di alberi infetti, allora per tutti gli interi n esiste almeno un cammino auto-evitante di n passi che è un percorso di infezione. Quindi

$$0 \leqslant f(p) \leqslant \sum_{j=1}^{+\infty} q(j,n) \leqslant \sum_{j=1}^{+\infty} j\, q(j,n)$$

per tutti gli n. Ma

$$\sum_{j=1}^{+\infty} j\, q(j,n)$$

è il numero atteso di cammini auto-evitanti di n passi che sono percorsi di infezione. Questo numero è uguale a $p^n s(n)$, dove $s(n)$ è il numero totale di cammini auto-evitanti a n passi. Hammersley riuscì a dimostrare in un articolo di accompagnamento che $s(n)$ cresce come $e^{\kappa n}$ per $n \to +\infty$, dove κ è chiamata la costante connettiva. Se $p < e^{-\kappa}$, allora $p^n s(n)$ tende a 0 per $n \to +\infty$ e $f(p) = 0$. Quindi $p^* \geqslant e^{-\kappa} > 0$.

In pratica è quindi meglio che gli alberi non siano troppo vicini per mantenere p sotto p^* in caso di epidemia. Ma più vicini sono gli alberi, più alta è la produzione per ettaro. Si quindi deve trovare un compromesso.

Come hanno notato Broadbent e Hammersley, c'è una certa somiglianza tra l'esistenza di una probabilità critica nei processi di percolazione e l'esistenza di una soglia nei processi di ramificazione (vedi il capitolo 7).

Si può provare a stimare numericamente la probabilità critica p^*. A questo scopo, fissiamo un valore per p e approssimiamo la rete infinita con un reticolo finito di dimensione $N \times N$, con N sufficientemente grande. Supponiamo, per esempio, che l'albero al centro della rete sia infetto. Con un computer, si possono estrarre a sorte gli alberi che infettano e gli alberi che vengono infettati. La figura 22.2a e la figura 22.2b mostrano percorsi possibili di infezione estratti a sorte, usando gli archi che collegano i nodi, come in un grafo. Nella figura 22.2a, p è più piccolo di p^*, mentre nella figura 22.2b p è più grande. Si può facilmente determinare quali alberi possono essere infettati, cioè quelli che possono essere raggiunti da un percorso di archi che parte dall'albero infetto al centro. Essi sono contrassegnati da piccoli quadrati neri nelle figure.

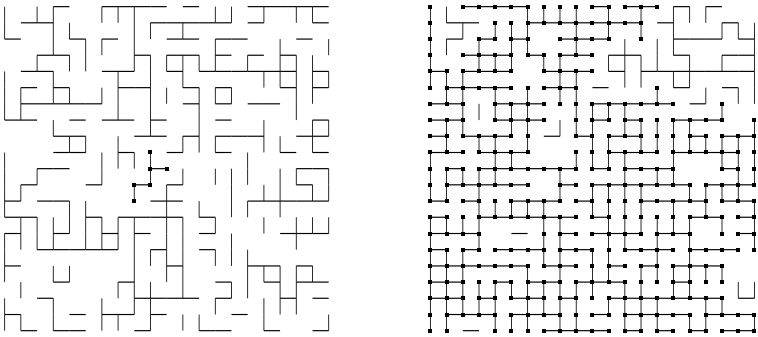

Figura 22.2: Percolazione con $p = 0{,}4$ e $p = 0{,}55$.

Si può quindi controllare se l'epidemia ha raggiunto almeno il bordo della rete $N \times N$. Se è così e se N è abbastanza grande, si può considerare che il numero di alberi infetti è «quasi infinito». Ripetendo questo tipo di simulazione molte volte, si può trovare un valore approssimativo della probabilità $f(p)$ che il numero di alberi infetti sia infinito (questo è il metodo Monte Carlo). Infine, lasciando variare p tra 0 e 1, si può ottenere un'approssimazione della soglia p^*, che è il più piccolo valore tale che $f(p) > 0$ se $p > p^*$.

L'articolo di Broadbent e Hammersley conteneva solo la prova dell'esistenza della soglia p^*. Negli anni seguenti Hammersley continuò a sviluppare la teoria matematica della percolazione, mentre Broadbent si dedicò ad altri

argomenti. Con lo sviluppo dei computer negli anni '70, divenne più facile eseguire le simulazioni descritte sopra (figura 22.3). È stato quindi congetturato che $p^* = 1/2$. Questo risultato fu finalmente dimostrato nel 1980 da Harry Kesten dell'Università di Cornell.

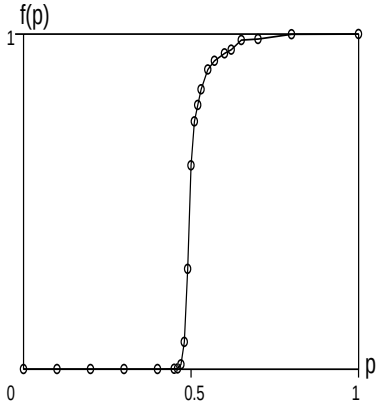

Figura 22.3: Probabilità $f(p)$ che infiniti alberi vengano infettati in funzione di p. La curva è ottenuta eseguendo 1 000 simulazioni su una rete 200 per 200.

Tra il 1959 e il 1969 Hammersley lavorò per l'Istituto di Economia e Statistica dell'Università di Oxford. Divenne fellow del *Trinity College*. Nel 1964 pubblicò in collaborazione con David Handscomb un libro intitolato *Metodi Monte Carlo*. Fu eletto alla *Royal Society* nel 1976. Si ritirò nel 1987 ma continuò a frequentare il Centro per la matematica industriale e applicata di Oxford. Morì nel 2004.

Broadbent ottenne il suo dottorato all'*Imperial College* nel 1957. Trovò un lavoro in un'azienda a carattere industriale, la *United Glass Bottle Manufacturers*. Dopo dieci anni nell'industria cominciò a lavorare in un'agenzia di stampa, la *London Press Exchange*, che faceva studi scientifici sui lettori. L'agenzia fu acquistata nel 1969 da Leo Burnett, una compagnia pubblicitaria americana. Broadbent ha lavorato a come misurare l'efficacia della pubblicità e ha pubblicato diversi libri su questo argomento: *Spendere soldi per la pubblicità* (1975), *Bilancio pubblicitario* (1989), *Pubblicità responsabile* (1997) e *Quando fare pubblicità* (1999). Nel 1980 ha contribuito ad avviare i Premi per l'efficacia della pubblicità. Ha trascorso diversi anni presso la sede centrale di Leo Burnett a Chicago come direttore economico del marchio. Ha

anche diretto una sua società di consulenza, *BrandCon Limited*. È morto nel 2002.

Bibliografia

1. Grimmett, G., Welsh, D.: John Michael Hammersley. *Biogr. Mem. Fellows R. Soc.* 53, 163–183 (2007)
2. Broadbent, S.R.: Discussion on symposium on Monte Carlo methods. *J. R. Stat. Soc. B* 16, 68 (1954)
3. Broadbent, S.R., Hammersley, J.M.: Percolation processes I: Crystals and mazes. *Proc. Camb. Philos. Soc.* 53, 629–641 (1957)
4. Broadbent, T.: Simon Broadbent – The man with a sense of fun who gave advertising a value. *Campaign*, 26 April 2002. http://www.campaignlive.co.uk/news/143366/
5. Hammersley, J.M.: Percolation processes II: The connective constant. *Proc. Camb. Philos. Soc.* 53, 642–645 (1957)
6. Hammersley, J.M.: Percolation processes: lower bounds for the critical probability. *Ann. Math. Stat.* 28, 790–795 (1957)
7. Hammersley, J.M.: Origins of percolation theory. In: Deutscher, G. Zallen, R., Adler, J. (eds.) *Percolation Structures and Processes*, 47–57. Israel Physical Society (1983)
8. Hammersley, J.M., Morton, K.W.: Poor man's Monte Carlo. *J. R. Stat. Soc. B* 16, 23–38 (1954)
9. Hammersley, J.M., Handscomb, D.C.: *Monte Carlo Methods*. Fletcher & Son, Norwich (1964)
10. Kesten, H.: The critical probability of bond percolation on the square lattice equals 1/2. *Comm. Math. Phys.* 74, 41–59 (1980)
11. Metropolis, N., Ulam, S.: The Monte Carlo method. *J. Amer. Stat. Assoc.* 44, 335–341 (1949)

Capitolo 23

Teoria dei giochi ed evoluzione (1973)

Nel 1973 Maynard Smith e Price pubblicarono un articolo che analizzava perché gli animali evitano di usare le loro armi più pericolose nel caso di conflitti intraspecifici, cioè relativi ad animali della stessa specie. Il loro modello utilizzava la teoria dei giochi ed è stato tra quelli che ha dato un notevole contributo applicativo di questa teoria matematica ai problemi evolutivi.

John Maynard Smith è nato a Londra nel 1920. Suo padre, che era un chirurgo, morì quando aveva otto anni. Maynard Smith studiò all'Eton College e si dedicò agli studi di ingegneria al *Trinity College* dell'Università di Cambridge. Fu poi membro del Partito Comunista britannico. Nel 1939, quando scoppiò la guerra, cercò di offrirsi volontario per l'esercito ma fu respinto a causa della sua scarsa vista. Finì i suoi studi di ingegneria e lavorò per alcuni anni alla progettazione di aerei militari. Infine decise di dedicarsi alla biologia, studiando genetica all'University College di Londra con Haldane come correlatore. Divenne docente di zoologia nel 1952. Lasciò il partito comunista dopo gli eventi del 1956 in Ungheria. Il suo primo libro, intitolato *La teoria dell'evoluzione*, fu pubblicato nel 1958. Nel 1965 divenne professore di biologia alla neonata Università del Sussex. In seguito pubblicò altri due libri: *Idee matematiche in biologia* (1968) e *Sull'evoluzione* (1972).

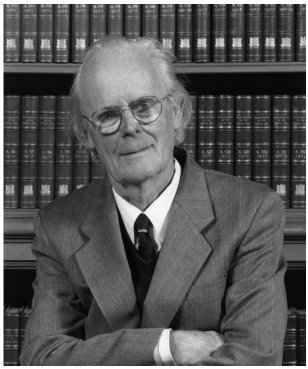

Figura 23.1: Maynard Smith (1920–2004)

George R. Price è nato nel 1922 negli Stati Uniti. Ha studiato chimica all'Università di Chicago, ottenendo un dottorato nel 1946 dopo aver lavorato al Progetto Manhattan, per la costruzione della bomba atomica. Nel 1950 divenne ricercatore associato in medicina all'Università del Minnesota. In seguito lavorò come giornalista indipendente per varie riviste prima di ritornare come ricercatore all'IBM. Nel 1967, dopo essere stato curato per un cancro alla tiroide, si stabilì in Inghilterra e si dedicò allo studio di una materia completamente diversa: la biologia evolutiva. Lavorò a Londra al Laboratorio Galton dell'*University College* dal 1968. Il suo primo articolo in questa nuova area, *Selezione e covarianza*, fu pubblicato con l'aiuto di W. D. Hamilton in un numero del 1970 di *Nature* e conteneva quella che ora è chiamata equazione di Price.

Price ha anche presentato un altro articolo alla rivista *Nature*, questa volta sui conflitti tra animali. Ma non aveva il formato giusto per questa rivista. Così Maynard Smith, che era il revisore, suggerì di preparare una versione più breve. Price iniziò a lavorare su qualcos'altro mentre Maynard Smith iniziò a sviluppare l'idea di Price per conto suo. Alla fine Maynard Smith e Price pubblicarono un articolo congiunto intitolato *La logica del conflitto animale*, che *Nature* pubblicò nel 1973. L'articolo diede un interessante contributo all'uso della teoria dei giochi nella biologia evolutiva. Prima di allora, la teoria dei giochi era stata sviluppata principalmente per l'economia e la politica, soprattutto dopo il libro del 1944 di John von Neumann e Oskar Morgenstern intitolato *Teoria dei giochi e del comportamento economico*. Il punto di partenza di Maynard Smith e Price era la formulazione della seguente domanda: come mai nei conflitti tra animali della stessa specie, le «armi» a loro disposizione (corna, artigli, veleno ecc.) sono raramente usate per uccidere? Seguendo le idee di Darwin sulla lotta per la vita, gli animali più aggressivi dovrebbero vincere più combattimenti e avere un maggior numero di figli, portando a un'escalation nell'uso delle «armi». Da notare che questo era il periodo della guerra fredda, quindi l'argomento aveva anche un sapore politico.

Maynard Smith e Price hanno immaginato una sequenza di giochi in cui due animali possono entrare in competizione per una risorsa, per esempio un territorio in un habitat favorevole. Nella presentazione semplificata che Maynard Smith userà nel suo libro del 1982 *Evoluzione e teoria dei giochi*, ogni animale adotta la «strategia del falco» o la «strategia della colomba». Nel seguito parleremo semplicemente di falchi e colombe, ma intendiamo strategie adottate da animali della stessa specie. Sia $V > 0$ il valore della risorsa, il che significa che se $\mathscr{R}_0$ è il normale numero medio di figli di un animale, il vincitore della competizione ha in media $\mathscr{R}_0 + V$ figli.

Se un falco incontra un altro falco, combattono per la risorsa: il vincitore ottiene la risorsa di valore V, il perdente subisce un «costo» $C > 0$. Ognuno dei due falchi ha una probabilità pari a 1/2 di vincere la competizione e la stessa probabilità di perdere. La vincita attesa da un combattimento tra due falchi è quindi $\frac{1}{2}(V-C)$ per i due concorrenti. Se, invece, un falco incontra una colomba, allora il falco ottiene la risorsa V, la colomba fugge senza combattere e il costo è 0. Infine, se due colombe si incontrano, una di loro ottiene la risorsa V, l'altra fugge senza combattere e a costo zero. Avendo ciascuna delle due colombe la stessa probabilità 1/2 di vincere, la vincita attesa quando due colombe si incontrano è quindi $V/2$. Le vincite possono essere riassunte come nella tabella 23.1.

Tabella 23.1: Vincite attese del gioco falco-colomba.

	un falco	una colomba
la vincita di un falco contro...	$\frac{1}{2}(V-C)$	V
la vincita di una colomba contro...	0	$V/2$

Più in generale, si possono immaginare combattimenti tra individui che possono adottare una delle due strategie, chiamate 1 e 2, con una matrice di vincita attesa $(G_{i,j})_{1\leqslant i,j\leqslant 2}$. Nell'esempio precedente, i falchi seguono la strategia 1, le colombe la strategia 2, $G_{1,1} = \frac{1}{2}(V-C)$, $G_{1,2} = V$, $G_{2,1} = 0$ e $G_{2,2} = V/2$. Nell'articolo originale del 1973, Maynard Smith e Price avevano in effetti già usato simulazioni al computer per testare più di due possibili strategie (queste erano chiamate «falco», «topolino», «prepotente», «vendicatore» e «esploratore-vendicatore»).

Immaginate ora una grande popolazione di animali della stessa specie con una proporzione x_n di falchi e una proporzione $1-x_n$ di colombe alla n-esima generazione. I falchi nella n-esima generazione hanno un numero medio di figli pari a

$$R_1(n) = \mathscr{R}_0 + x_n G_{1,1} + (1-x_n) G_{1,2}\,. \tag{23.1}$$

Allo stesso modo, le colombe hanno un numero medio di figli pari a

$$R_2(n) = \mathscr{R}_0 + x_n G_{2,1} + (1-x_n) G_{2,2}\,. \tag{23.2}$$

Il numero medio di figli nell'intera popolazione è quindi:

$$R(n) = x_n R_1(n) + (1-x_n) R_2(n)\,.$$

Dimenticando le possibili sottigliezze dovute alla riproduzione sessuale, vediamo che la proporzione di falchi nella generazione successiva è:

$$x_{n+1} = x_n R_1(n)/R(n)\,. \tag{23.3}$$

Quindi, $x_{n+1} > x_n$ se $R_1(n) > R(n)$ e $x_{n+1} < x_n$ se $R_1(n) < R(n)$. Ci sono tre possibili stati stazionari: $x = 0$, $x = 1$ e

$$x^* = \frac{G_{1,2} - G_{2,2}}{G_{2,1} - G_{1,1} + G_{1,2} - G_{2,2}}$$

purché $0 < x^* < 1$. Nel gioco del falco e della colomba, $x^* = V/C < 1$ purché $V < C$.

Infatti, $x = 0$ è un ovvio stato stazionario di (23.3). Se $x \neq 0$ è un altro stato stazionario, allora $R_1 = R = xR_1 + (1-x)R_2$. Quindi o $x = 1$ o $R_1 = R_2$. Quest'ultima possibilità è equivalente a

$$xG_{1,1} + (1-x)\,G_{1,2} = xG_{2,1} + (1-x)\,G_{2,2},$$

che dà lo stato stazionario x^*.

Lo stato stazionario $x = 1$ corrisponde ad una popolazione con il 100 % degli individui che seguono la strategia 1. Questo stato stazionario è stabile se non può essere invaso da pochi individui che seguono la strategia 2. Da (23.3), vediamo che questa condizione è equivalente ad avere $R_1(n) > R(n)$ per tutti gli x_n sufficientemente vicini a 1. Poiché $R(n) = x_n R_1(n) + (1-x_n)R_2(n)$, la condizione diventa $R_1(n) > R_2(n)$ per tutti x_n sufficientemente vicini a 1. Osservando le espressioni (23.1)-(23.2) di R_1 e R_2, arriviamo alla conclusione che $x = 1$ è stabile se e solo se una delle due seguenti condizioni è soddisfatta:

- $G_{1,1} > G_{2,1}$;
- $G_{1,1} = G_{2,1}$ e $G_{1,2} > G_{2,2}$.

Se è così, si dice che la strategia 1 è una strategia evolutivamente stabile. Nel gioco falco-colomba, la condizione $G_{1,2} > G_{2,2}$ è sempre vera. Quindi la strategia del falco è evolutivamente stabile se e solo se $G_{1,1} \geqslant G_{2,1}$, cioè $V \geqslant C$.

Lo stato stazionario $x = 0$ corrisponde a una popolazione con tutti gli individui che seguono la strategia 2. Questa situazione è simmetrica alla precedente se si scambiano gli indici 1 e 2. Nel gioco falco-colomba, abbiamo $G_{1,2} = V > G_{2,2} = V/2$ quindi lo stato stazionario $x = 0$ è sempre instabile. Introdurre un piccolo numero di falchi in una popolazione di colombe porterebbe ad una progressiva invasione da parte dei falchi.

Allo stesso modo, si può dimostrare che il terzo stato stazionario x^*, purché $0 < x^* < 1$, è sempre stabile. Nel gioco falco-colomba,

$$x^* = V/C$$

corrisponde a una popolazione mista con falchi e colombe.

In conclusione, ci sono due casi nel gioco falco-colomba. Se $V = C$, cioè se il valore della risorsa è più grande del costo possibile, allora la popolazione tende ad uno stato stazionario con falchi ma senza colombe, qualunque sia la condizione iniziale $x(0)$ con $0 < x(0) < 1$. La strategia del falco è quindi una strategia evolutivamente stabile. Se, al contrario, $V < C$, allora la popolazione tende ad uno stato stazionario misto con una proporzione x^* di falchi e una proporzione $1 - x^*$ di colombe. Quindi il modello fornisce una spiegazione del perché gli individui con comportamenti meno aggressivi possono sopravvivere quando $V < C$. La formula $x^* = V/C$ mostra inoltre che maggiore è il costo C per i perdenti, minore è la proporzione x^* di falchi nella popolazione. Quindi le specie con le «armi» più pericolose raramente le usano per combattimenti intraspecifici: preferiscono combattimenti rituali inoffensivi, in cui gli animali in competizione cercano di impressionarsi a vicenda ma evitano combattimenti reali che potrebbero causare lesioni.

L'articolo originale del 1973 di Maynard Smith e Price discuteva il concetto di strategia evolutivamente stabile e utilizzava principalmente simulazioni al computer di competizioni tra animali, registrando le vincite delle diverse strategie. L'approccio che utilizza equazioni dinamiche come (23.3) è stato sviluppato un po' più tardi, in particolare da Taylor e Jonker. Da allora molti autori hanno applicato le idee della teoria dei giochi a questioni di biologia evolutiva o viceversa hanno applicato approcci dinamici evolutivi a problemi più classici della teoria dei giochi. Oltre alle questioni riguardanti i conflitti tra animali, si possono citare per esempio i problemi dell'investimento parentale o del rapporto tra i sessi (il rapporto tra il numero di maschi e femmine alla nascita), quest'ultimo studiato già da Carl Döringer nel 1884 e da Ronald Fisher nel suo libro del 1930 *La teoria genetica della selezione naturale*. Alcuni altri modelli si concentrano sugli aspetti dinamici del «dilemma del prigioniero» o del gioco «Sasso-Carta-Forbice». Si è anche capito che il concetto di strategia evolutivamente stabile è strettamente legato al concetto di equilibrio di Nash nella teoria dei giochi.

Price, che fu un ateo convinto, ebbe un'esperienza mistica nel 1970 e si convertì alla fede cristiana. Nel 1974 ha abbandonato la sua ricerca perché sentiva che «il tipo di genetica matematica teorica che stava facendo non

era molto rilevante per i problemi umani». Ha donato tutti i suoi averi ai senzatetto e si è suicidato pochi mesi dopo.

Maynard Smith, al contrario, continuò questa linea di pensiero e fu eletto alla *Royal Society* nel 1977. Ha pubblicato molti libri: *Modelli in ecologia* (1974), *L'evoluzione del sesso* (1978), *L'evoluzione e la teoria dei giochi* (1982), *I problemi della biologia* (1986), *Darwin ha avuto ragione?* (1988) e *Genetica evolutiva* (1989). Ha anche pubblicato in collaborazione con E. Szathmáry *Le principali transizioni nell'evoluzione* (1995) e *Le origini della vita: Dalla nascita della vita all'origine del linguaggio* (1999). Si è ritirato nel 1985. Nel 1999 ha ricevuto il premio Crafoord in bioscienze dall'Accademia Reale Svedese delle Scienze per i suoi «fondamentali contributi allo sviluppo concettuale della biologia evolutiva». Nel 2003 ha pubblicato in collaborazione con D. Harper *Segnali animali*. È morto nel Sussex nel 2004.

Bibliografia

1. Charlesworth, B., Harvey, P.: John Maynard Smith, 6 January 1920–19 April 2004. *Biog. Mem. Fellows R. Soc.* 51, 253–265 (2005)
2. Edwards, A.W.F.: Carl Düsing (1884) on the regulation of the sex-ratio. *Theor. Pop. Biol.* 58, 255–257 (2000)
3. Frank, S.A.: George Price's contributions to evolutionary genetics. *J. Theor. Biol.* 175, 373–388 (1995)
4. Maynard Smith, J., Price, G.R.: The logic of animal conflict. *Nature* 246, 15–18 (1973)
5. Maynard Smith, J.: *Evolution and the Theory of Games*. Cambridge University Press (1982)
6. Schwartz, J.: Death of an altruist: Was the man who found the selfless gene too good for this world? *Lingua Franca* 10, 51–61 (2000) bio.kuleuven.be/ento/-pdfs/schwartz2000.pdf
7. Sigmund, K.: John Maynard Smith and evolutionary game theory. *Theor. Pop. Biol.* 68, 7–10 (2005)
8. Taylor, P.D., Jonker, L.B.: Evolutionary stable strategies and game dynamics. *Math. Biosci.* 40, 145–156 (1978)
9. Von Neumann, J., Morgenstern, O.: *Theory of Games and Economic Behavior*. Princeton University Press (1944) archive.org

Capitolo 24

Popolazioni caotiche (1974)

Nel 1974 Robert May, un fisico australiano divenuto ecologista, studiò l'equazione logistica a tempo discreto come modello per la dinamica delle popolazioni. Notò che si verificavano biforcazioni inaspettate e che il comportamento asintotico poteva essere addirittura caotico. In questo modo, le previsioni a lungo termine possono essere impossibili anche con un semplice modello deterministico. L'articolo di May fu tra quelli che lanciarono la «teoria del caos».

Robert McCredie May nacque nel 1936 in Australia. Dopo aver studiato fisica teorica e aver ricevuto un dottorato all'Università di Sydney nel 1959, trascorse due anni nel dipartimento di matematica applicata dell'Università di Harvard. Tornato in Australia, divenne professore di fisica teorica. Nel 1971, mentre visitava l'*Institute for Advanced Study* di Princeton, cambiò il suo argomento di ricerca e iniziò a concentrarsi sulla dinamica delle popolazioni animali. Nel 1973 divenne professore di zoologia a Princeton. Lo stesso anno pubblicò un libro intitolato *Stabilità e complessità in ecosistemi modello*.

Figura 24.1: Robert M. May

Nel 1974 May pubblicò su *Science* un articolo intitolato *Popolazioni biologiche con generazioni non sovrapposte: punti stabili, cicli stabili e caos*, in cui ha dimostrato che modelli matematici molto semplici nella dinamica delle popolazioni possono comportarsi in modo caotico.

Per capire l'origine di questo problema, bisogna tornare indietro nel tempo di circa dieci anni. Nel 1963 Edward Lorenz, un meteorologo americano che lavorava all'Istituto di tecnologia del Massachusetts (M.I.T.), aveva notato, facendo simulazioni numeriche al computer, che un modello semplificato

dell'atmosfera con solo tre equazioni differenziali poteva comportarsi in modo molto sorprendente: un minuscolo cambiamento delle condizioni iniziali poteva cambiare completamente il risultato finale di una simulazione e quindi anche le previsioni meteorologiche. Il matematico Henri Poincaré, dopo aver lavorato sul moto dei pianeti nel sistema solare, aveva infatti già pensato a questa possibilità all'inizio del XX secolo, molto prima dell'era dei computer. Ma all'inizio degli anni '70, solo pochi ricercatori cominciavano ad osservare più da vicino questa strana proprietà. All'Università del Maryland, James Yorke stava pensando al lavoro di Lorenz e ha introdotto il termine «caos» in questo contesto. L'articolo[1] che scrisse con il suo studente Tien-Yien Li, intitolato *Il periodo tre implica il caos*, apparve nel 1975.

Da parte sua, May si stava concentrando sul modello $p_{n+1} = p_n + a\,p_n(1 - p_n/K)$, dove a e K sono parametri positivi e p_n rappresenta la dimensione di una popolazione animale nell'anno n. Quando p_n è piccolo rispetto alla capacità portante dell'ambiente K, la dinamica è vicina ad una progressione geometrica $p_{n+1} \approx (1+a)\,p_n$. L'equazione per intero è una sorta di analogo in tempo discreto dell'equazione logistica introdotta da Verhulst (vedi capitolo 6). Ma a differenza di quest'ultima, May ha mostrato che l'equazione a tempo discreto può avere un comportamento molto più sorprendente, che è facile da osservare con una semplice calcolatrice tascabile facendo addizioni e moltiplicazioni (figura 24.2). Maynard Smith aveva già considerato l'equazione nel suo libro del 1968 *Idee matematiche in biologia*, ma nonostante avesse provato alcuni valori numerici per a, non si era reso conto che c'era qualcosa di speciale.

La figura 24.2, che è simile a quella dell'articolo di May del 1974, mostra che la popolazione p_n converge verso uno stato stazionario quando $0 < a < 2$. Quando $2 < a \leqslant 2{,}449$ (il limite superiore 2,449 è un'approssimazione), la popolazione p_n tende a un ciclo di periodo 2. Quando $2{,}450 \leqslant a \leqslant 2{,}544$, la popolazione p_n tende ad un ciclo di periodo 4. Quando $2{,}545 \leqslant a \leqslant 2{,}564$, p_n tende ad un ciclo di periodo 8, ecc. Gli intervalli del parametro a per cui p_n tende ad un ciclo di periodo 2^n diventano più piccoli all'aumentare di n e non superano mai 2,570. Quando $a = 2{,}570$, p_n può comportarsi in modo «caotico».

Nel 1976 May scrisse un articolo di revisione del problema, pubblicato su *Nature*, intitolato *Modelli matematici semplici con dinamiche molto complicate*. Lì raccolse non solo i suoi risultati ma anche quelli di altri ricercatori. Innanzitutto, ponendo $x_n = a\,p_n/[K\,(1+a)]$ e $r = 1+a$ (in modo che $r > 1$), vediamo che l'equazione $p_{n+1} = p_n + a\,p_n(1 - p_n/K)$ può essere

[1] Un risultato più generale è stato dimostrato da O. M. Sharkovsky nel 1964, ma il suo articolo pubblicato in una rivista matematica ucraina non era molto conosciuto.

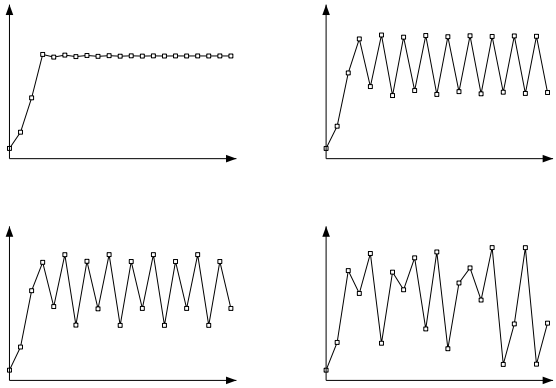

Figura 24.2: In tutte le figure: n è sull'asse orizzontale, p_n sull'asse verticale e $p_0 = K/10$. Le linee sono ottenute unendo i punti con coordinate (n, p_n). In alto a sinistra: $0 < a < 2$ (stato stazionario). In alto a destra: $2 < a \leqslant 2{,}449$ (ciclo di periodo 2). In basso a sinistra: $2{,}450\, a \leqslant 2{,}544$ (ciclo di periodo 4). In basso a destra: $2{,}570 \leqslant a \leqslant 3$ (possibilità di caos).

riscritta nella forma più semplice

$$x_{n+1} = r x_n \left(1 - x_n\right) . \tag{24.1}$$

Affinché questa equazione abbia un significato nella dinamica della popolazione, x_n dovrebbe essere non negativa per tutti gli n. Quindi assumiamo che la condizione iniziale x_0 soddisfi $0 \leqslant x_0 \leqslant 1$ e che $r \leqslant 4$. Quest'ultima condizione assicura che il lato destro di (24.1) rimanga tra 0 e 1. È da notare che il caso caotico $r = 4$ era già stato usato come generatore di numeri casuali da Stanisław Ulam e John von Neumann nel 1947. Se introduciamo la funzione $f(x) = rx(1-x)$, allora l'equazione (24.1) può essere riscritta come $x_{n+1} = f(x_n)$ e gli stati stazionari sono le soluzioni di $x = f(x)$. Graficamente, queste sono le intersezioni delle curve $y = f(x)$ e $y = x$ (figura 24.3). Notate che $x = 0$ è sempre uno stato stazionario. Poiché $r > 1$, c'è anche un altro stato stazionario $x^* > 0$ tale che $x^* = rx^*\left(1 - x^*\right)$, cioè, $x^* = 1 - 1/r$.

Poiché $r > 1$, lo stato stazionario $x = 0$ è instabile. Infatti, quando x_n è vicino a 0, abbiamo $x_{n+1} \approx r x_n$. Quindi x_n tende ad allontanarsi da 0. Per quanto riguarda lo stato stazionario x^*, è localmente stabile solo per $1 < r < 3$.

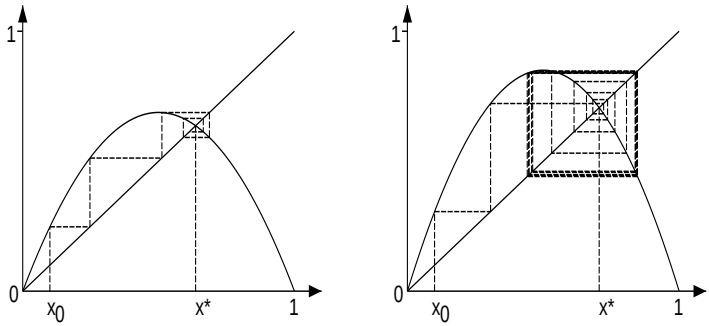

Figura 24.3: La funzione $y = f(x)$, la retta $y = x$, lo stato stazionario x^* e la sequenza definita da $x_{n+1} = f(x_n)$. (a): $r = 2{,}75$, la sequenza tende a x^*. (b): $r = 3{,}4$, lo stato stazionario x^* è instabile e la sequenza tende ad un ciclo di periodo 2.

Infatti, poniamo $y_n = x_n - x^*$. Allora (24.1) è equivalente a $y_{n+1} = (2 - r - r y_n) y_n$. Se x_n è vicino a x^*, allora y_n è vicino a 0 e $y_{n+1} \approx (2-r) y_n$. Ma se $y_{n+1} = k y_n$, allora $y_n = k^n y_0$ così che $y_n \to 0$ quando $n \to +\infty$ se e solo se $-1 < k < 1$. Qui lo stato stazionario x^* è localmente stabile se e solo se $-1 < 2 - r < 1$, cioè $1 < r < 3$.

Quando $1 < r < 3$, si può mostrare che per tutte le condizioni iniziali $0 < x_0 < 1$, la sequenza x_n tende davvero a x^* (figura 24.3a). Ma cosa succede quando $3 < r \leqslant 4$? Per rispondere a questa domanda, notiamo prima che $x_{n+2} = f(x_{n+1}) = f(f(x_n))$. Introduciamo la funzione $f_2(x) = f(f(x)) = r^2 x(1-x)\big(1 - rx(1-x)\big)$ e consideriamo le soluzioni dell'equazione $x = f_2(x)$, che si chiamano punti fissi della funzione $f_2(x)$. Graficamente, questi sono le intersezioni delle curve $y = f_2(x)$ e $y = x$ (figura 24.4).

Se $x = f(x)$, allora $x = f(f(x)) = f_2(x)$. Quindi $x = 0$ e $x = x^*$ sono punti fissi anche per la funzione $f_2(x)$. Ma quando $r > 3$, la funzione $f_2(x)$ ha altri due punti fissi, x_- e x_+, tali che $f(x_-) = x_+$ e $f(x_+) = x_-$.

Infatti notiamo che $f_2'(x) = f'(f(x)) f'(x)$ così che $f_2'(x^*) = [f'(x^*)]^2$. Ma $f'(x) = r(1-2x)$ e $x^* = 1 - 1/r$. Quindi $f'(x^*) = 2 - r$ e $f_2'(x^*) = (2-r)^2$. Quindi la pendenza della funzione $f_2(x)$ in $x = x^*$ è tale che $f_2'(x^*) > 1$ se $r > 3$. Ma poiché $f_2(0) = 0$, $f_2'(0) = r^2 > 1$ e $f_2(1) = 0$,

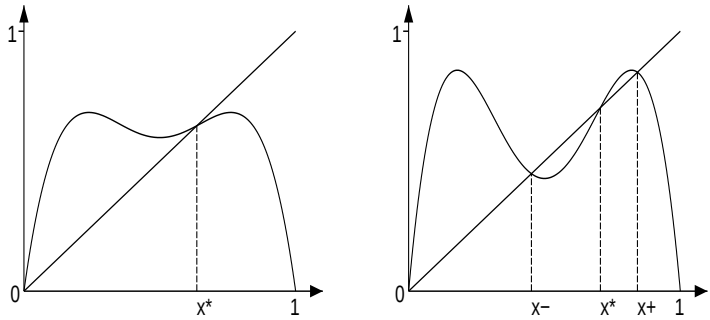

Figura 24.4: Le curve $y = f_2(x) = f(f(x))$ e $y = x$ e lo stato stazionario x^*. (a): $r = 2{,}75$. (b): $r = 3{,}4$ e le altre due soluzioni x_- e x_+ dell'equazione $x = f_2(x)$.

vediamo nella figura 24.4b che ci sono necessariamente altre due soluzioni x_- e x_+ dell'equazione $x = f_2(x)$, con $0 < x_- < x^*$ e $x^* < x_+ < 1$. Un altro modo di arrivare alla stessa conclusione consiste nel risolvere l'equazione $x = f_2(x)$, che è un'equazione polinomiale di grado 4 con due radici note: $x = 0$ e $x = x^*$. Le altre due soluzioni x_- e x_+ sono le radici del polinomio

$$x^2 - \frac{1+r}{r}x + \frac{1+r}{r^2} = 0\,. \tag{24.2}$$

Sono reali se il discriminante è positivo, cioè se $r > 3$. Poiché $f_2(f(x_-)) = f(f(f(x_-))) = f(f_2(x_-)) = f(x_-)$, il punto $f(x_-)$ è anche un punto fisso di $f_2(x)$. Ma $f(x_-) \neq x_-$ perché x_- non è un punto fisso di $f(x)$. E $f(x_-) \neq x^*$, altrimenti avremmo $x_- = f(f(x_-)) = f(x^*) = x^*$. Poiché $f(x_-) \neq 0$, concludiamo che $f(x_-) = x_+$. Allo stesso modo, $f(x_+) = x_-$.

Quindi per $r > 3$, vediamo che se per esempio $x_0 = x_-$, allora $x_1 = x_+$, $x_2 = x_-$, $x_3 = x_+$, ecc. Si può anche mostrare che per quasi tutti i dati iniziali $0 < x_0 < 1$, la sequenza x_n tende come $n \to +\infty$ verso il ciclo di periodo 2 : x_-, x_+, x_-, x_+, ecc. (figura 24.3b e 24.4b). Questo ciclo rimane stabile finché r è sotto il valore critico $r_1 = 1 + \sqrt{6} \approx 3{,}449$, dove $f_2'(x_-) = -1$.

Vediamo usando (24.2) che $f_2'(x_-) = f'(f(x_-))\,f'(x_-) = f'(x_+)\,f'(x_-) = r^2\,(1-2x_+)(1-2x_-) = r^2\,(1-2(x_+ + x_-) + 4x_+x_-) = -r^2 + 2r + 4$. Quindi $f_2'(x_-) = -1$ se $-r^2 + 2r + 5 = 0$ e in particolare se $r = 1 + \sqrt{6}$.

Per $r_1 < r < r_2$, un ciclo di periodo 4 diventa stabile: quattro nuovi punti fissi della funzione $f_4(x) = f_2(f_2(x)) = f(f(f(f(x))))$ appaiono (figura 24.5a). Per $r_2 < r < r_3$, è un ciclo di lunghezza 8, ecc. I numeri r_n tendono al limite $r_\infty \approx 3{,}570$ quando $n \to +\infty$. Quando $r_\infty < r \leqslant 4$, il sistema può anche essere caotico! La figura 24.5b mostra il diagramma di biforcazione, che dà un'idea della complessità della dinamica [2].

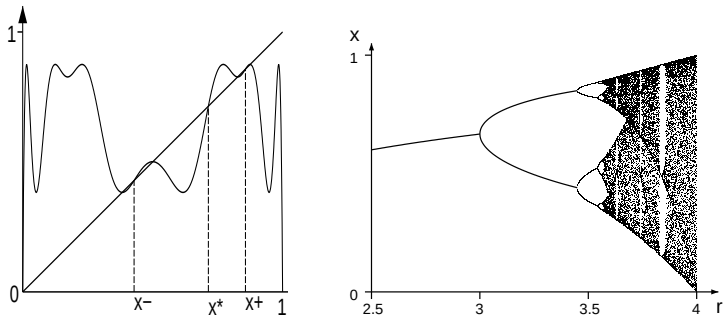

Figura 24.5: (a) La curva $y = f_4(x)$ quando $r = 3{,}5$ e la linea $y = x$. Oltre a x^*, x_+ e x_-, ci sono altri quattro punti fissi, che non sono facili da distinguere. (b) Diagramma di biforcazione dell'equazione (24.1).

R. M. May ha concluso sottolineando che anche sistemi dinamici molto semplici possono avere un comportamento molto complicato. Questo non è specifico dell'equazione $x_{n+1} = r\,x_n\,(1-x_n)$. La stessa «cascata di raddoppio dei periodi» che porta al caos appare per altre equazioni con una funzione $f(x)$ che ha la forma di un «dosso». Questo è il caso per esempio di un'altra equazione usata nella biologia delle popolazioni: $x_{n+1} = x_n \exp(r(1-x_n))$.

Questo studio suggerisce che non ci si dovrebbe sorprendere se molte serie di dati riguardanti la dinamica della popolazione sono difficili da analiz-

[2] Questo diagramma è stato ottenuto tracciando per ogni dato valore di r i punti con coordinate (r, x_{200}), (r, x_{201}),...,(r, x_{220}), dove $x_{n+1} = f(x_n)$ e $x_0 = 0{,}1$. Se x_n tende ad uno stato stazionario, vediamo solo un punto nel diagramma. Se x_n tende ad un ciclo di periodo 2, vediamo due punti, ecc.

zare. Il modello mostra anche che la distinzione tra modelli deterministici e stocastici non è così chiara come si pensava: anche con un semplice modello deterministico, può essere impossibile fare previsioni a lungo termine se i parametri sono in regime caotico.

Nel 1979 May è stato eletto membro della *Royal Society*. Dal 1988 al 1995 è stato professore all'Università di Oxford e all'*Imperial College* di Londra. Dal 1995 al 2000 è stato consigliere scientifico capo del governo britannico. Nel 1996 ha ricevuto il premio Crafoord «per la sua ricerca ecologica pionieristica sull'analisi teorica della dinamica delle popolazioni, delle comunità e degli ecosistemi». Dall'ecologia si è rivolto verso l'epidemiologia e l'immunologia, pubblicando due libri: *Malattie infettive dell'uomo* (1991, con Roy Anderson) e *Dinamica dei virus, i fondamenti matematici dell'immunologia e della virologia* (2000, con Martin Nowak). Quest'ultimo libro analizza l'interazione tra le cellule del sistema immunitario e l'HIV (il virus che causa l'AIDS) come una specie di sistema predatore-preda (vedi capitolo 13). Dal 2000 al 2005, May è stato presidente della *Royal Society*. È stato nominato cavaliere nel 1996 ed è diventato un pari a vita nel 2001.

Bibliografia

1. Gleick, J.: *Chaos, Making a New Science*. Viking Penguin, New York (1987)
2. Levin, S.A.: Robert May receives Crafoord prize. *Not. Amer. Math. Soc.* 43, 977–978 (1996) ams.org
3. Li, T.Y., Yorke, J.A.: Period three implies chaos. *Amer. Math. Monthly* 82, 985–992 (1975)
4. Lorenz, E.N.: Deterministic nonperiodic flow. *J. Atmosph. Sci.* 20, 130–141 (1963) journals.ametsoc.org
5. May, R.M.: Biological populations with nonoverlapping generations: stable points, stable cycles and chaos. *Science* 186, 645–647 (1974)
6. May, R.M.: Simple mathematical models with very complicated dynamics. *Nature* 261, 459–467 (1976)
7. May, R.M., Oster, G.F.: Bifurcations and dynamic complexity in simple ecological models. *Amer. Natur.* 110, 573–599 (1976)
8. Maynard Smith, J.: *Mathematical Ideas in Biology*. Cambridge (1968)
9. Poincaré, H.: *Science et Méthode*. Flammarion, Paris (1908) gallica.bnf.fr
10. Sharkovsky, O.M.: Co-existence of cycles of a continuous mapping of a line onto itself. *Ukr. Math. J.* 16, 61–71 (1964)
11. Ulam, S.M., von Neumann, J.: On combination of stochastic and deterministic processes. *Bull. Amer. Math. Soc.* 53, 1120 (1947) ams.org

Capitolo 25

La politica cinese del figlio unico (1980)

Nel 1980 Song Jian e i suoi collaboratori, che erano stati specialisti della teoria del controllo applicata all'ingegneria aerospaziale, calcolarono che se il tasso di natalità in Cina fosse rimasto al suo livello attuale, la popolazione avrebbe raggiunto più di due miliardi di individui durante il ventunesimo secolo. I loro risultati, basati su un modello matematico con struttura di età, hanno contribuito alla decisione del governo di passare alla politica del figlio unico.

Song Jian[1] è nato nel 1931 a Rongcheng nella provincia cinese di Shandong. Durante gli anni '50 ha studiato in Unione Sovietica all'Università Tecnica Statale Bauman di Mosca e al Dipartimento di Matematica e Meccanica dell'Università Statale di Mosca. Poi tornò in Cina e divenne il capo dell'Ufficio di Ricerca Cibernetica nell'Istituto di Matematica dell'Accademia Cinese delle Scienze. Era uno specialista dell'applicazione della teoria del controllo alla guida dei missili. Lavorò anche per il Settimo Ministero della Costruzione di Macchine, che fu poi rinominato Ministero dell'Aerospazio. Nel 1978 ha iniziato a concentrarsi sui legami tra la teoria del controllo e la demografia.

Figura 25.1: Song Jian

Per capire il contesto del lavoro di Song Jian sulla dinamica delle popolazioni, bisogna prima dare un'idea di cosa sia la «teoria del controllo». È lo

[1] Song è il nome di famiglia. È sempre scritto per primo in cinese.

studio dei sistemi dinamici il cui comportamento dipende da alcuni parametri che possono essere modificati nel tempo per ottimizzare un determinato criterio. Questa teoria è stata particolarmente sviluppata in relazione ai programmi spaziali negli Stati Uniti e nell'URSS. Infatti, gli ingegneri dovevano «controllare» la traiettoria delle navette spaziali per portare i satelliti alla loro orbita intorno alla Terra. Ma le applicazioni non erano limitate a problemi fisici o ingegneristici. Anche le politiche di controllo delle nascite potrebbero essere considerate come una sorta di problema di controllo ottimale in senso matematico.

Si dovrebbe anche menzionare il saggio intitolato *I limiti dello sviluppo, Rapporto per il progetto del Club di Roma sui dilemmi dell'umanità*, pubblicato nel 1972 e scritto da un gruppo dell'Istituto di tecnologia del Massachusetts (M.I.T.). Questo studio si basava su un modello matematico della crescita economica mondiale che teneva conto delle risorse naturali, della dimensione della popolazione e dell'inquinamento. Il rapporto suggeriva che l'economia mondiale stava andando verso una catastrofe per esaurimento delle risorse non rinnovabili, per mancanza di cibo per la popolazione o per un eccesso di inquinamento. La limitazione volontaria delle nascite era una delle soluzioni proposte. In sintesi era una sorta di versione moderna delle tesi di Malthus. Il rapporto ha ricevuto una grande eco in Occidente durante gli anni '70.

Dalla fondazione della Repubblica Popolare nel 1949, il tasso di natalità cinese era stato molto alto, tranne durante il catastrofico «Grande balzo in avanti». A metà degli anni '70 la Cina si stava lentamente riprendendo dalla rivoluzione culturale. La pianificazione familiare esortava le donne a ritardare le nascite, ad aumentare il tempo tra due nascite consecutive e ad avere meno figli. Deng Xiaoping, che emerse come nuovo leader dopo la morte di Mao Zedong nel 1976, iniziò la politica delle «Quattro modernizzazioni» nel 1978: agricoltura, industria, scienza e tecnologia e difesa nazionale. Le dimensioni e la crescita della popolazione cinese erano allora percepite come importanti ostacoli a queste modernizzazioni. Gli scienziati che avevano lavorato fino ad allora su applicazioni militari furono incoraggiati a trovare soluzioni per questo difficile problema.

Con questo sfondo, Song Jian andò nel 1978 a Helsinki per un congresso della Federazione Internazionale del Controllo Automatico. Lì notò che alcuni ricercatori in Europa avevano cercato di applicare la teoria del controllo ai problemi della popolazione, con l'idea che un rigoroso controllo delle nascite avrebbe potuto evitare le catastrofi annunciate dal rapporto *I limiti dello sviluppo*. Tornato in Cina, mise in piedi una piccola squadra, tra cui il suo collega Yu Jingyuan e l'esperto di computer Li Guangyuan, per applica-

re questo tipo di modellazione matematica ai dati riguardanti la popolazione cinese. A quel tempo la comunicazione scientifica tra la Cina e il resto del mondo era scarsa. Il team sviluppò nuovamente le equazioni che descrivono l'evoluzione della struttura dell'età di una popolazione, nello stesso modo in cui avevano fatto Lotka e McKendrick (vedi i capitoli 10 e 16). Usando un modello a tempo continuo, chiamato

- $P(x,t)$ la popolazione di età x al tempo t;
- $m(x)$ la mortalità all'età x;
- $P_0(x)$ la struttura per età della popolazione al tempo $t = 0$;
- $b(t)$ la fertilità totale delle donne al tempo t, cioè il numero medio di figli che una donna avrebbe durante la sua vita se la fertilità specifica per età rimanesse quella del tempo t;
- f la proporzione di nascite femminili;
- $h(x)$ la distribuzione di probabilità dell'età della madre quando nasce un bambino :
$$\int_0^{+\infty} h(x)\,dx = 1.$$

Con queste notazioni e ipotesi, l'evoluzione della struttura dell'età può essere modellata dall'equazione alle derivate parziali

$$\frac{\partial P}{\partial t}(x,t) + \frac{\partial P}{\partial x}(x,t) = -m(x)\,P(x,t)\,,$$

soggetta alla condizione iniziale $P(x,0) = P_0(x)$ e la condizione al bordo

$$P(0,t) = b(t)\,f \int_0^{+\infty} h(x)\,P(x,t)\,dx\,,$$

dove $b(t)$ è il parametro da controllare. Dalla analisi matematica del sistema suddetto, emerge che, se la fertilità totale delle donne è costante e sopra la soglia critica

$$b^* = 1/\left[f \int_0^{+\infty} h(x)\,\mathrm{e}^{-\int_0^x m(y)\,dy}\,dx\right],$$

allora la popolazione aumenta esponenzialmente. Questo criterio è simile a quello ottenuto da Lotka con la formula (10.2). Il team di Song Jian ha considerato anche la versione a tempo discreto del modello, che è simile al

modello di Leslie (vedi capitolo 21). Chiamiamo $P_{k,n}$ la popolazione di età k nell'anno n e introduciamo analogamente m_k, b_n e h_k. Allora

$$P_{k+1,n+1} = (1 - m_k)\, P_{k,n} \ , \quad P_{0,n+1} = b_n\, f \sum_{k \geqslant 0} h_k\, P_{k,n} \, .$$

Conoscendo dalle indagini campionarie la mortalità m_k (figura 25.2a), la proporzione di nascite femminili $f \approx 0{,}487$, la distribuzione per età delle madri h_k (Figura 25.2b), la condizione iniziale $P_{k,0}$ che è la struttura per età della popolazione nel 1978 (Figura 25.3a) e variando la fertilità totale b (assunta costante in ogni simulazione), il team di Song Jian avrebbe potuto fare proiezioni demografiche per il loro paese con un orizzonte temporale di cento anni, dal 1980 al 2080 (Figura 25.3b). Date le migliaia di addizioni e moltiplicazioni richieste (l'anno n varia tra 0 e 100 anni, l'età k tra 0 e 90 anni), era necessario un computer. All'epoca in Cina poche persone avevano accesso a tali apparecchiature, tranne quelle che lavoravano per i militari. Song Jian, uno dei maggiori esperti nella guida dei missili, era uno di questi.

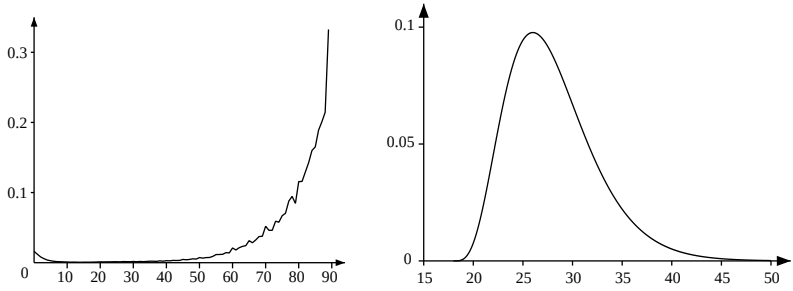

Figura 25.2: (a) Mortalità (per anno) in funzione dell'età nel 1978. (b) Forma smussata della fertilità (per anno) in funzione dell'età nel 1978.

Le proiezioni suggerivano che anche se la Cina avesse mantenuto la sua fertilità del 1978 di $b = 2{,}3$ bambini per donna, che è appena sopra la soglia critica stimata a $b^* = 2{,}19$, la popolazione sarebbe aumentata da 980 milioni nel 1980 a 2,12 miliardi nel 2080. Ma la Cina stava già utilizzando quasi tutta la terra che poteva servire per l'agricoltura. Aveva persino la tendenza a perdere parte di questa terra a causa della desertificazione e dell'urbanizzazione. Come nutrire una tale popolazione se il progresso delle rese agricole non è sufficiente? È la stessa domanda che Malthus aveva considerato due

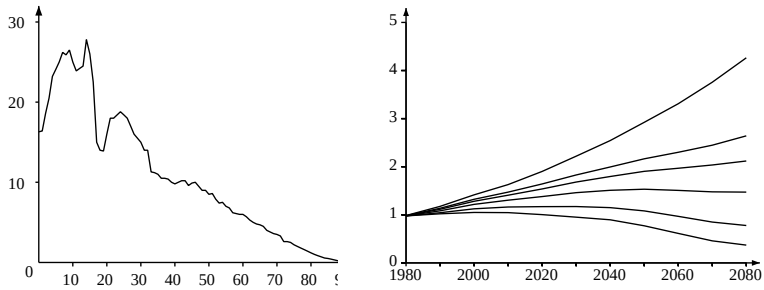

Figura 25.3: (a) Piramide dell'età nel 1978. Asse orizzontale: età. Asse verticale: popolazione (in milioni). (b) Proiezioni demografiche (in miliardi) seguendo diverse ipotesi sul numero medio di figli per donna. Dal basso verso l'alto: $b = 1{,}0$; 1,5 ; 2,0 ; 2,3 ; 2,5; 3,0

secoli prima. Con la fertilità del 1975 di $b = 3{,}0$, la popolazione potrebbe addirittura raggiungere i 4,26 miliardi nel 2080. Con $b = 2{,}0$, la popolazione raggiungerebbe un massimo di 1,53 miliardi intorno all'anno 2050 prima di cominciare a diminuire leggermente. Con $b = 1{,}5$, un massimo di 1,17 miliardi sarebbe raggiunto intorno al 2030. Con $b = 1{,}0$, il massimo sarebbe solo 1,05 miliardi e sarebbe raggiunto intorno al 2000. Sotto questa ipotesi, la popolazione tornerebbe al suo livello del 1978 solo nel 2025.

La parte più sorprendente di questo lavoro furono le sue conseguenze pratiche, in effetti di importanza ineguagliabile nella storia della dinamica matematica delle popolazioni. Infatti Li Guangyuan mostrò i risultati delle simulazioni del team nel dicembre 1979 durante un simposio sulla popolazione a Chengdu, nella provincia del Sichuan. Nel gennaio 1980, Song Jian, Yu Jingyuan e Li Guangyuan pubblicarono questi risultati in una rivista economica cinese, sostenendo tra l'altro una politica del figlio unico. Inviarono anche il loro articolo – *Una relazione sulla ricerca quantitativa sulla questione dello sviluppo demografico della Cina* – al massimo scienziato cinese Qian Xuesen, che lo inoltrò con una raccomandazione al capo dell'amministrazione per la pianificazione delle nascite. I risultati del team di Song Jian fecero una profonda impressione sulla maggior parte dei leader politici. Questi erano già convinti della necessità di un maggiore controllo delle nascite, nonostante ciò che Marx aveva scritto (vedi capitolo 5) ma erano ancora indecisi sul livello

di controllo. Nel febbraio 1980, il Consiglio di Stato e il Comitato Centrale del Partito fissarono un obiettivo per la popolazione cinese di 1,2 miliardi per l'orizzonte 2000. Nel marzo 1980, i risultati del team di Song Jian furono pubblicati sul Quotidiano del Popolo. In aprile, una commissione di leader politici e specialisti della popolazione esaminò le conseguenze ambientali ed economiche della crescita della popolazione e concluse che una politica del figlio unico era necessaria per raggiungere l'obiettivo fissato da Deng Xiaoping per il reddito pro capite nell'anno 2000. La politica è diventata ufficiale nel settembre dello stesso anno e una lettera aperta che la spiegava alla popolazione è stata pubblicata sulla prima pagina del Quotidiano del Popolo.

Nel 1983, ci furono ancora molte nascite non autorizzate. Fu deciso che un membro di ogni coppia con già due figli sarebbe stato sterilizzato e che ogni gravidanza proibita sarebbe stata interrotta. Tuttavia, a partire dal 1984, alle coppie rurali con una sola figlia fu permesso di avere un secondo figlio. Alcuni adattamenti sono stati introdotti negli ultimi anni: se in una coppia sia l'uomo che la donna erano figli unici, allora possono avere due figli. Le misure repressive contro le coppie che hanno più di un figlio sono dure: gli impiegati statali possono perdere il lavoro, si deve pagare una multa costosa per ottenere i documenti amministrativi per la scolarizzazione di un secondo figlio ecc. In sintesi, è difficile trovare nella storia della modellazione matematica un altro esempio con un impatto sociale così forte. Naturalmente il lavoro di Song Jian e dei suoi collaboratori è stato solo uno degli elementi che hanno portato alla scelta della politica del figlio unico. Ma sembra aver giocato un ruolo importante.

Come nei capitoli precedenti, il ruolo della modellazione matematica può essere oggetto di preoccupazione. Partendo da una situazione di vita reale, si costruisce un modello. Può essere analizzato matematicamente o simulato con un computer. Si può quindi capire come il modello si comporta al variare di alcuni parametri. Tuttavia, la matematica non dice se il modello è un'immagine fedele della vita reale. Alcuni aspetti molto importanti possono essere stati trascurati. Alcuni modelli contengono anche una funzione obiettivo, per esempio mantenere la popolazione cinese sotto 1,2 miliardi. La matematica non dice se questo obiettivo è appropriato[2].

Nel 1980 Song Jian fu anche coautore della nuova edizione del libro intitolato *Ingegneria cibernetica* di Qian Xuesen, il «padre» del programma spa-

[2]La popolazione nell'anno 2000 è stata stimata a 1,264 miliardi. Il reddito pro capite è cresciuto approssimativamente da 200 a 1000 dollari tra il 1980 e il 2000. Allo stesso tempo, il rapporto tra i sessi è diventato estremamente sbilanciato verso i ragazzi, principalmente a causa dell'aborto selettivo del sesso.

ziale cinese. Ha poi ricoperto varie posizioni politiche di alto livello: vice-ministro e capo scienziato-ingegnere del Ministero dell'Aerospazio (1981–1984), membro del Comitato Centrale del Partito Comunista Cinese (1982–2002), presidente della Commissione di Stato per la Scienza e la Tecnologia (1985–1998), Consigliere di Stato (1986–1998) ecc. Ha pubblicato anche altri due libri che sono stati tradotti in inglese: Il controllo della popolazione in Cina (1985, con Tuan Chi-Hsien e Yu Jingyuan) e *Controllo del sistema di popolazione* (1988, con Yu Jingyuan). Questi libri sviluppano la teoria del controllo ottimale applicata alla dinamica delle popolazioni. Song Jian è stato eletto nel 1991 all'Accademia cinese delle scienze e nel 1994 all'Accademia degli ingegneri, di cui è stato presidente dal 1998 al 2002.

Bibliografia

1. Greenhalgh, S.: Missile science, population science: The origins of China's one-child policy. *China Q.* 182, 253–276 (2005)
2. Greenhalgh, S.: *Just One Child, Science and Policy in Deng's China*. University of California Press (2008)
3. Meadows, D.H., Meadows, D.L., Randers, J., Behrens, W.W.: *The Limits to Growth, A Report for the Club of Rome's Project on the Predicament of Mankind*, 2nd edn. Universe Books, New York (1974)
4. Song, J.: Selected Works of J. Song. *Science Press*, Beijing (1999)
5. Song, J.: Some developments in mathematical demography and their application to the People's Republic of China. *Theor. Popul. Biol.* 22, 382–391 (1982)
6. Song, J., Yu, J.: *Population System Control*. Springer, Berlin (1988)

Capitolo 26

Alcuni problemi contemporanei

Questo capitolo fornisce una breve panoramica di alcuni problemi contemporanei di matematica della dinamica delle popolazioni: l'invecchiamento della popolazione in demografia; le malattie emergenti (AIDS, SARS, malattie trasmesse da vettori...) e la politica di vaccinazione in epidemiologia; le politiche di pesca in ecologia; la dispersione di organismi geneticamente modificati in genetica delle popolazioni. In questo capitolo, sono menzionate le istituzioni specializzate che lavorano in Francia sulla modellizzazione di questi problemi in demografia. Vengono anche sottolineati vari aspetti del lavoro di ricerca.

Questo capitolo fornisce una breve panoramica della ricerca contemporanea sulla modellazione matematica della dinamica delle popolazioni. Essendo l'argomento piuttosto vasto, sono qui riportati solo alcuni esempi di studi sviluppati da ricercatori in Francia.

In demografia un problema relativamente nuovo è apparso negli ultimi decenni: l'invecchiamento della popolazione. Questo problema è oggetto di preoccupazione non solo in Francia (figura 26.1) ma anche in molti altri paesi europei e in Giappone. Ha importanti conseguenze economiche e sociali su: sistemi pensionistici, politiche di immigrazione, ecc. In Francia, modelli matematici che cercano di analizzare il fenomeno dell'invecchiamento sono sviluppati dall'Istituto Nazionale di Studi Demografici (INED) e dall'Istituto Nazionale di Statistica e Studi Economici (INSEE). Una delle difficoltà delle proiezioni demografiche risiede nel fatto che i tassi di natalità possono variare considerevolmente nel tempo senza essere prevedibili nemmeno con un decennio di anticipo. Questo è particolarmente sorprendente se si guarda alle proiezioni fatte nel 1968 per la popolazione francese nel 1985: queste proiezioni non hanno anticipato la diminuzione del tasso di natalità che si è verificata durante gli anni '70. Sarebbe interessante rivedere tutte le previsioni basate su modelli matematici che si sono rivelate errate, specialmente quelle che hanno trovato un'eco nei media. Questo controbilancerebbe l'impressione di «progresso» data dal presente libro, un'impressione che può essere già apparsa sospetta al lettore dopo aver letto il capitolo sulla politica cinese del figlio unico (nota: questa politica è spesso criticata in Occidente ma sembra essere relativamente ben accettata da molti cinesi). Riguardo a quest'ultimo

argomento, un nuovo problema è ora di attualità: come ammorbidire la politica per evitare il rapido fenomeno dell'invecchiamento previsto nei prossimi decenni. Ancora una volta i modelli matematici contribuiscono al dibattito.

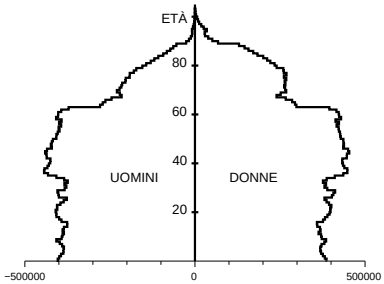

Figura 26.1: Piramide delle età della popolazione francese al 1° gennaio 2010 (www.insee.fr).

In epidemiologia, tra i nuovi problemi emersi a livello globale negli ultimi due decenni, lo sviluppo dell'epidemia di AIDS è particolarmente sorprendente. Alcuni modelli cercano di indovinare il futuro dell'epidemia nei paesi infettati più recentemente come la Russia, l'India o la Cina. È difficile prevedere se l'epidemia rallenterà come in Europa occidentale e Nord America o se raggiungerà una percentuale importante della popolazione come in alcuni paesi subsahariani. Altre malattie emergenti come Ebola in Africa, la febbre del Nilo occidentale in Nord America, la SARS (sindrome respiratoria acuta severa), l'influenza aviaria, il chikungunya o l'influenza H1N1 sono state tutte esaminate con modelli matematici, anche se con poco successo.

Per la SARS, una difficoltà di modellizzazione era che l'epidemia rimaneva relativamente limitata all'interno di ogni paese, ma poteva diffondersi molto rapidamente da paese a paese (Hong Kong e Cina, Singapore, Canada...). Il carattere casuale delle curve dell'epidemia in ogni nuovo focolaio non poteva essere trascurato. Come abbiamo visto nei capitoli 16 e 22, i modelli stocastici sono di solito più difficili da gestire.

Per l'epidemia di chikungunya che si è verificata tra il 2005 e il 2006 sull'isola della Reunion (un territorio francese d'oltremare nell'Oceano Indiano), alcuni modelli sono stati ispirati da quello di Ross per la malaria (vedi capitolo 12), essendo le due malattie trasmesse dalle zanzare. Un aspetto importante da prendere in considerazione era l'influenza della stagionalità. Infatti la popolazione di zanzare diminuisce durante l'inverno meridionale, riducendo così la trasmissione della malattia. Questo può essere visto nella figura 26.2, che mostra il numero di nuovi casi segnalati ogni settimana da una piccola rete di circa trenta medici di base che copre solo una frazione della

popolazione dell'isola. La rete non ha rilevato alcun nuovo caso durante diverse settimane di settembre e ottobre 2005, ma la trasmissione della malattia continuava. Modelli matematici dell'epidemia sono stati sviluppati all'Istituto Nazionale della Salute e della Ricerca Medica (INSERM) e all'Istituto di Ricerca Tropicale (IRD). Nonostante questi modelli, nessuno fu in grado di prevedere che l'epidemia non si sarebbe estinta prima della fine dell'inverno meridionale del 2005, quando aveva infettato solo poche migliaia di persone. Alla fine, quasi un terzo della popolazione dell'isola è stato infettato, cioè circa 266 000 persone.

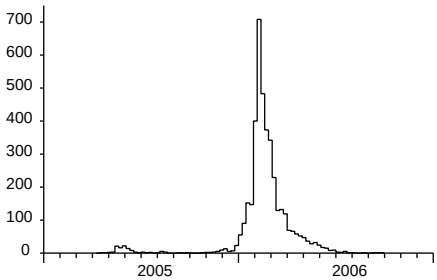

Figura 26.2: L'epidemia di chikungunya nell'isola della Riunione nel 2005–2006. Numero di nuovi casi segnalati a settimana da una piccola rete di medici in funzione del tempo. Il primo piccolo picco è stato raggiunto nel maggio 2005, il secondo grande picco nel febbraio 2006. I numeri in questa figura devono essere moltiplicati per circa 67 per ottenere la dimensione reale dell'epidemia. Fonte: www.invs.sante.fr.

Questo dimostra se ancora necessario che prevedere il futuro delle epidemie può essere abbastanza difficile e che non è così facile distinguere nei primi giorni di un'epidemia se sarà un'epidemia minore o maggiore. Un'analogia può essere fatta con le previsioni del tempo. Questo tipo di previsione si basa oggi su intense simulazioni al computer di complicati modelli matematici dell'oceano e dell'atmosfera. Tuttavia, le previsioni oltre qualche giorno non sono affidabili.

Da un punto di vista più teorico, l'epidemia di chikungunya ha sollevato la questione di come adattare la nozione di numero di riproduzione di base $\mathscr{R}_0$ al caso di modelli mediante i quali si assume che l'ambiente abbia fluttuazioni stagionali (per esempio periodiche). L'adattamento non è così semplice e questo solleva qualche preoccupazione su come il parametro $\mathscr{R}_0$ è stato

utilizzato per altre epidemie dipendenti dalla stagionalità come la pandemia dell' influenza H1N1 del 2009.

Un altro problema di crescente preoccupazione che i modellisti hanno cercato di analizzare è quello della resistenza ai farmaci (antibiotici, antimalarici). Sempre in epidemiologia, la questione ricorrente, dai tempi di Daniel Bernoulli e d'Alembert, di come bilanciare costi e benefici quando l'iniezione di un vaccino comporta un rischio potenziale è ancora oggetto di controversie e può rimanere tale man mano che cambia la sensibilità al rischio. Così, in seguito ad alcune indicazioni che il vaccino contro l'epatite B avrebbe causato gravi complicazioni in un numero molto piccolo di casi, il ministero della Sanità francese nel 1998 sospese la sua campagna di vaccinazione nelle scuole, nonostante il rischio risultasse trascurabile rispetto a quello di morire dopo l'infezione con il virus dell'epatite B.

In ecologia, lo studio della dinamica delle popolazioni dei pesci pone ancora molti problemi. Tuttavia si suppone che serva come base scientifica per la scelta delle quote di pesca e di altre restrizioni. La pesca eccessiva dell'acciuga nel Golfo di Biscaglia e del tonno rosso nel Mediterraneo sono solo due esempi recenti. Essendo la stima della riserva ittica spesso inaffidabile, i modelli che utilizzano tali dati devono essere considerati con cautela. In Francia questo tipo di studio è intrapreso principalmente dall'Istituto di ricerca per lo sfruttamento del mare (IFREMER). Alcuni modelli matematici hanno anche giocato un ruolo nelle passate decisioni della Commissione baleniera internazionale.

In genetica delle popolazioni, la dispersione degli organismi geneticamente modificati è anche un argomento di controversia che alcuni ricercatori hanno cercato di studiare utilizzando modelli di «reazione-diffusione» ispirati a quello di Fisher (vedi capitolo 20). Questo è il settore dell'Istituto Nazionale di Ricerca in Agronomia (INRA).

Sul lato più teorico della ricerca, si possono citare:

- i lavori sulle equazioni differenziali parziali come le equazioni di diffusione (si veda il capitolo 20) o le equazioni strutturate per età (si veda il capitolo 16);

- i lavori sui modelli stocastici con o senza la dimensione spaziale (vedi i capitoli 16 e 22), compresi quelli sulle reti casuali che modellano la diffusione delle epidemie e quelli che cercano approssimazioni deterministiche.

Questo tipo di ricerca è condotto principalmente dai matematici delle scienze applicate. Negli ultimi anni, diversi master in biologia matematica sono stati

introdotti nelle università francesi e in altri istituti di istruzione superiore.

Come in altri campi scientifici, lo studio matematico della dinamica delle popolazioni è organizzato principalmente attraverso:

- «società scientifiche»: *Società di biologia matematica* (dal 1973), *Società francofona di biologia teorica* (1985), *Società giapponese di biologia matematica* (1989), *Società europea di biologia matematica e teorica* (1991) ecc.
- riviste specializzate: *Bullettin of Mathematical Biology* (dal 1939), *Mathematical Biosciences* (1967), *Journal of Mathematical Biology* (1974), *Mathematical Medicine and Biology* (1984), *Mathematical Population Studies* (1988), *Mathematical Biosciences and Engineering* (2004) ecc.
- conferenze (*Riunione annuale della Società di Biologia Matematica*, *Dinamica delle popolazioni matematica e computazionale*, *Conferenza europea sulla biologia matematica e teorica* ecc.)

Si è fatto riferimento solo agli elementi di connessione tra la matematica e le applicazioni alla dinamica delle popolazioni. Ma per ogni area particolare (demografia, ecologia, genetica delle popolazioni, epidemiologia e così via), si possono trovare elementi simili che possono variare in relazione al tipo di modellazione matematica.

In conclusione, il lettore interessato è invitato a consultare gli articoli originali che sono disponibili sul *World Wide Web*. I riferimenti sono riportati nella sezione bibliografica alla fine di ogni capitolo. Come Ronald Fisher scrisse una volta su Mendel:

> «La Storia della Scienza ha sofferto molto per l'uso da parte degli insegnanti di materiale di seconda mano e la conseguente omissione delle circostanze e dell'atmosfera intellettuale in cui sono state fatte le grandi scoperte del passato. Uno studio di prima mano è sempre istruttivo e spesso pieno di sorprese.»

Bibliografia

1. Bacaër, N.: Approssimazione di $\mathscr{R}_0$ per malattie trasmesse da vettori con una popolazione periodica di vettori. *Bull. Math. Biol.* 69, 1067–1091 (2007) hal.archives-ouvertes.fr
2. Levin, S.A.: Mathematics and biology, the interface. www.bio.vu.nl/nvtb

Figure

- p. 5. Opera di Thomas Murray (ca. 1687) conservato dalla *Royal Society* di Londra. Chapman, S.: Edmond Halley, F.R.S. 1656–1742. *Notes Rec. R. Soc. Lond.* 12, 168–174 (1957) © The Royal Society.
- p. 11. Opera di Emanuel Handmann (1753) conservato dal *Kunstmuseum* di Basilea. *Leonhard Euler 1707–1783, Beiträge zu Leben und Werk*. Birkhäuser, Basel (1983)
- p. 16. Opera un tempo tenuto dalla *Petri-Kirche*, probabilmente distrutto durante la battaglia di Berlino nel 1945. Reimer, K.F.: Johann Peter Süssmilch, seine Abstammung und Biographie. *Arch. soz. Hyg. Demogr.* 7, 20–28 (1932)
- p. 22. Opera di Johann Niclaus Grooth (ca. 1750–1755) conservato dal *Naturhistorisches Museum* di Basilea. Speiser, D.: *Die Werke von Daniel Bernoulli*, Band 2. Birkhäuser, Basel (1982)
- p. 29. Opera di Quentin de La Tour (1753) conservato al *Musée du Louvre* di Parigi.
- p. 32. Opera di John Linnell (1833) conservato dall'*Haileybury College*, Inghilterra. Habakkuk, H.J.: Robert Malthus, F.R.S. (1766-1834). *Notes Rec. R. Soc. Lond.* 14, 99–108 (1959)
- p. 36. Incisione di Flameng (1850). Quetelet, A.: Pierre-François Verhulst. *Annu. Acad. R. Sci. Lett. B.-Arts Belg.* 16, 97–124 (1850)
- p. 42. Heyde, C.C., Seneta, E.: *I. J. Bienaymé, Statistical Theory Anticipated*. Springer-Verlag, New York (1977) © Académie des sciences, Institut de France.
- p. 46. Bateson, W.: *Mendel's Principles of Heredity*. Cambridge University Press (1913)
- p. 51. Pearson, K.: *The Life, Letters, and Labors of Francis Galton*, vol. 1. Cambridge University Press (1914)
- p. 51. Ritratto di Watson nella Biblioteca del *Trinity College*, Università di Cambridge. Kendall, D.G.: Branching processes since 1873. *J. Lond. Math. Soc.* 41, 385–406 (1966)
- p. 57. *Alfred J. Lotka Papers*. Documenti di politica pubblica. Dipartimento di libri rari e collezioni speciali. © Princeton University Library.
- p. 61. Titchmarsh, E. C.: Godfrey Harold Hardy 1877–1947. *Obit. Not. Fellows R. Soc.* 6, 446–461 (1949)

- p. 64. Stern, C.: Wilhelm Weinberg. *Genetics* 47, 1–5 (1962)
- p. 67. G.H.F.N.: Sir Ronald Ross 1857–1932. *Obit. Not. Fellows R. Soc.* 1, 108–115 (1933) © The Royal Society.
- p. 76. Whittaker, E.T.: Vito Volterra 1860–1940. *Obit. Not. Fellows R. Soc.* 3, 690–729 (1941)
- p. 80. Yates, F., Mather, K.: Ronald Aylmer Fisher, 1890–1962. *Biog. Mem. Fellows R. Soc.* 9, 91–120 (1963) © The Royal Society/Godfrey Argent Studio.
- p. 84. Yates, F.: George Udny Yule. *Obit. Not. Fellows R. Soc.* 8,308–323 (1952)
- p. 92. Heyde, C.C., Seneta, E. (eds.): *Statisticians of the Centuries*. Springer, New York (2001)
- p. 102. britannica.com/EBchecked/topic/252257/J-B-S-Haldane © Bassano and Vandyk Studios.
- p. 109. Hill, W.G.: Sewall Wright, 21 December 1889-3 March 1988. *Biog. Mem. Fellows R. Soc.* 36, 568–579 (1990) © Llewellyn Studios, Chicago.
- p. 105. Nybølle, H.C.: Agner Krarup Erlang f. 1. Januar 1878 - d. 3. Februar 1929. *Mat. Tidsskr. B*, 32–36 (1929)
- p. 118. Tikhomirov, V.M.: A.N. Kolmogorov. In: Zdravkovska, S., Duren, P.L. (eds.) *Golden Years of Moscow Mathematics*, 2nd edn., 101–128. American Mathematical Society (2007)
- p. 118. *I. G. Petrowsky Selected Works Part I*. Gordon and Breach, Amsterdam (1996) © Taylor and Francis Books UK.
- p. 123. P.H. Leslie fotografato da Denys Kempson. Crowcroft, P.: *Elton's Ecologists, a History of the Bureau of Animal Population*. University of Chicago Press (1991)
- p. 126. © Geoffrey Grimmett.
- p. 132. Charlesworth, B., Harvey, P.: John Maynard Smith, 6 January 1920–19 April 2004. *Biog. Mem. Fellows R. Soc.* 51, 253–265 (2005) © The Royal Society.
- p. 138. © Samuel Schlaefli / ETH Zürich.
- p. 145. Selected works of J. Song. *Science Press*, Beijing (1999) © Song Jian.

Indice

1	La successione di Fibonacci (1202)	1
2	La tavola di mortalità di Halley (1693)	4
3	Eulero e la crescita geometrica (1748–1761)	10
4	Daniel Bernoulli e l'inoculazione del vaiolo (1760)	21
5	Malthus e gli ostacoli alla crescita geometrica (1798)	32
6	Verhulst e l'equazione logistica (1838)	36
7	Bienaymé e l'estinzione dei nomi di famiglia (1845–1847)	42
8	Mendel e l'ereditarietà (1865)	46
9	Galton, Watson e il problema dell'estinzione (1873–1875)	50
10	Lotka e la teoria della popolazione stabile (1907–1911)	57
11	La legge di Hardy-Weinberg (1908)	61
12	Ross e la malaria (1911)	66
13	Lotka, Volterra e il sistema preda-predatore (1920–1926)	72
14	Fisher e la selezione naturale (1922)	79
15	Yule e l'evoluzione (1924)	83
16	McKendrick sulla modellizzazione delle epidemie (1926)	91
17	Haldane e le mutazioni (1927)	101
18	Erlang sul problema dell'estinzione (1929–1933)	105
19	Wright e la deriva genetica casuale (1931)	109
20	La diffusione dei geni (1937)	115
21	La matrice di Leslie (1945)	122
22	Percolazione ed epidemie (1957)	126
23	Teoria dei giochi ed evoluzione (1973)	132
24	Popolazioni caotiche (1974)	138
25	La politica cinese del figlio unico (1980)	145
26	Alcuni problemi contemporanei	152

Come ha sottolineato Eugene Wigner, la matematica si è dimostrata efficace, oltre ogni ragionevole aspettativa, nelle scienze fisiche e nelle loro applicazioni tecnologiche. Il ruolo della matematica nelle scienze biologiche, mediche e sociali è stato molto più modesto ma è recentemente cresciuto grazie alla capacità di simulazione offerta dai moderni computer.

Questo libro traccia la storia della dinamica delle popolazioni - un argomento teorico strettamente connesso alla genetica, all'ecologia, all'epidemiologia e alla demografia - dove la matematica ha portato intuizioni significative. Presenta una panoramica della genesi di diversi temi importanti: la crescita esponenziale, da Eulero e Malthus alla politica cinese del figlio unico; lo sviluppo di modelli stocastici, dalle leggi di Mendel e la questione dell'estinzione dei cognomi alla teoria della percolazione per la diffusione delle epidemie, e le popolazioni caotiche, dove determinismo e casualità si intrecciano.

Il lettore di questo libro vedrà, da una prospettiva diversa, i problemi che gli scienziati affrontano quando i governi chiedono previsioni affidabili per aiutare a controllare le epidemie (COVID, influenza, AIDS), gestire risorse rinnovabili (quote di pesca, diffusione di organismi geneticamente modificati) o anticipare evoluzioni demografiche come l'invecchiamento.

Con i recenti progressi nella traduzione automatica, il monopolio virtuale di una sola lingua nella letteratura scientifica non è più giustificato. La crescente alienazione linguistica nelle università può essere invertita. Con questa traduzione italiana accuratamente rivista, incoraggiamo questo nuovo percorso.

9 791034 351398

15€