

Nicolas Bacaër

amb

Xavier Bardina i Sílvia Cuadrado

**Breu història dels
models matemàtics en
dinàmica de poblacions**



**Breu història dels models matemàtics en
dinàmica de poblacions**

Nicolas Bacaër

amb

Xavier Bardina i Sílvia Cuadrado

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

Xavier Bardina / Sílvia Cuadrado
Universitat Autònoma de Barcelona
bardina@mat.uab.cat / silvia@mat.uab.cat

Qui vulgui comprar la versió en paper d'aquest llibre pot enviar un missatge a nicolas.bacaer@ird.fr.

Foto de la portada: Esteve Terrús (1857–1922), *El Racó*. ©Museu d'Art Jacint Rigau, Perpinyà.

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

Pour l'édition catalane :
© Nicolas Bacaër, Paris, 2022
ISBN : 979-10-343-9682-5
Dépôt légal : juin 2022

Introducció

La dinàmica de poblacions és l'àrea de la ciència que intenta explicar de manera mecànica i senzilla les variacions temporals de la mida i la composició de les poblacions biològiques, com les d'éssers humans, animals, plantes o microorganismes. Tot i que està relacionada amb l'àrea més descriptiva de l'estadística de poblacions és força diferent. Un punt en comú és que ambdues fan un ús ampli del llenguatge matemàtic.

La dinàmica de poblacions es troba en la intersecció de diversos camps: les matemàtiques, les ciències socials (demografia), la biologia (genètica i ecologia de poblacions) i la medicina (epidemiologia). Per això, no sol presentar-se com un tot, malgrat les similituds entre els problemes que s'aborden en les diferents aplicacions. Una notable excepció en francès és el llibre *Teories matemàtiques de la població* d'Alain Hillion¹. Però presenta el tema des del punt de vista d'un matemàtic, distingint diversos tipus de models: models a temps discret ($t = 0, 1, 2, \dots$) i models a temps continu (t és un nombre real), models deterministes (els estats futurs es coneixen amb exactitud si l'estat present es coneix amb exactitud) i models estocàstics (on intervenen les probabilitats). Per tant, el llibre considera lògicament models deterministes discrets, models deterministes continus, models estocàstics discrets i models estocàstics continus.

En aquest llibre s'intenta tractar el mateix tema però des d'un punt de vista històric. La recerca s'explica en el seu context. S'inclouen breus biografies de científics. Això hauria de facilitar la lectura del llibre als menys familiaritzats amb les matemàtiques i, en general, pot ajudar a comprendre l'origen dels problemes estudiats. Però aquest llibre no es limita a la història. També pot servir com a introducció a la modelització matemàtica. Em va semblar important incloure els detalls de la majoria dels càlculs perquè el lector pugui veure realment les limitacions dels models. Les parts tècniques es destaquen en requadres i poden saltar-se en una primera lectura. L'últim capítol se centra en els nombrosos problemes contemporanis de la dinàmica de poblacions que es poden intentar analitzar des d'un punt de vista matemàtic. Per qui vulgui aprofundir més, les llistes de referències al final de cada capítol inclouen també pàgines web d'on es poden descarregar els articles originals.

No és possible en un llibre d'aquesta extensió donar una visió completa de tots els treballs desenvolupats fins ara ni parlar de tots els científics que han

¹ *Presses Universitaires de France*, París, 1986.

contribuït al tema. L'elecció realitzada conté necessàriament una component d'arbitrarietat, sobretot pel que fa a les darreres dècades. Espero, malgrat tot, que la mostra escollida sigui prou representativa i que em disculpin les persones actives dins de l'àrea els treballs de les quals no s'esmenten.

Aquest llibre està especialment indicat per a

- Estudiants de secundària i universitaris que es pregunten sobre els vincles que existeixen entre les assignatures de matemàtiques que cursen i el món que els envolta, o estudiants que preparen algun treball sobre un tema relacionat amb la dinàmica de poblacions.
- Professorat de matemàtiques que intenta fer el seu curs més atractiu. El coneixement de les quatre operacions elementals és suficient per a entendre gran part dels capítols 1, 2 i 5. El capítol 3 pot servir d'introducció a les aplicacions dels logaritmes. Aquest llibre també cobreix: les equacions en diferències als capítols 1, 3, 8, 11, 14, 21, 23, 24; les equacions diferencials als capítols 4, 6, 12, 13, 16; les equacions en derivades parcials als capítols 20, 25; una equació integral al capítol 10; i aplicacions de la teoria de la probabilitat als capítols 2, 7, 8, 9, 15, 16, 17, 18, 19 i 22.
- Persones ja familiaritzades amb la demografia, l'epidemiologia, la genètica o l'ecologia i disposades a comparar la seva àrea favorita amb unes altres que puguin implicar models matemàtics similars.
- Lectors interessats en la història de la ciència.

Per a la traducció al català, estic molt agraït a Xavier Bardina i Sílvia Cuadrado, que van revisar i van corregir la traducció automàtica de www.soft-catala.org.

Capítol 1

La successió de Fibonacci (1202)

L'any 1202, Leonardo de Pisa, també conegut com a Fibonacci, va publicar un llibre que va popularitzar a Europa el sistema numèric decimal indi que també havien adoptat els matemàtics àrabs. Entre els molts exemples que es donen al llibre, un es refereix al creixement d'una població de conills. És probablement un dels exemples més antics d'un model matemàtic per a la dinàmica de poblacions.

Leonardo de Pisa, anomenat Fibonacci molt temps després de la seva mort, va néixer al voltant del 1170 a la República de Pisa, quan aquesta es trobava en l'apogeu del seu poder comercial i militar al món mediterrani. Al voltant del 1192, el pare de Fibonacci va ser enviat per la República al port de Bugia, actualment a Algèria, per a dirigir un establiment comercial. Poc després, el seu fill va anar a viure amb ell per a preparar-se com a comerciant. Leonardo va començar a aprendre el sistema numèric decimal que els àrabs havien portat de l'Índia i que continua utilitzant-se a dia d'avui gairebé en la mateixa forma: 0, 1, 2, 3, 4, 5, 6, 7, 8 i 9. Mentre viatjava per negocis pel Mediterrani, va comparar els diferents sistemes numèrics i va estudiar les matemàtiques àrabs. De tornada a Pisa el 1202, va acabar d'escriure un llibre en llatí titulat *Liber abaci* («El llibre del càlcul») en el qual explicava el nou sistema numèric i mostrava com utilitzar-lo per a la comptabilitat, les conversions de pes i moneda, els tipus d'interès i moltes altres aplicacions. També va compilar la major part dels resultats en àlgebra i aritmètica coneguts pels àrabs.

Fibonacci va considerar al seu llibre el que avui es diria un problema de dinàmica de poblacions. Però va aparèixer només com un exercici de càlcul enmig d'altres temes no relacionats: la secció anterior del llibre és sobre nombres perfectes, que són la suma dels seus factors propis, com $28 = 14 + 7 + 4 + 2 + 1$, i la secció següent és un problema sobre el repartiment de diners entre quatre persones que equival a un sistema lineal de quatre equacions. Aquesta és una traducció del llatí del problema poblacional:

«Vet aquí un home que té una parella de conills junts en un lloc tancat. Es vol saber quantes parelles seran creades a partir d'aquesta parella inicial en un any, si sabem que, per la seva natu-

ralesa, cada parella fèrtil en un mes pareix una nova parella i a partir del segon mes de néixer ja donen a llum.»

Si hi ha una parella de conills nounats al principi del primer mes, aquesta parella encara no serà fèrtil fins al cap d'un mes i continuarà havent-hi una sola parella de conills al principi del segon mes. Aquesta parella de conills donarà a llum a una altra parella al principi del tercer mes, per la qual cosa hi haurà dues parelles en total. La parella inicial de conills tornarà a parir una altra parella al principi del quart mes. Però la segona parella de conills encara no serà fèrtil. Només hi haurà tres parelles de conills.

Utilitzant notació moderna, sigui P_n el nombre de parelles de conills al principi del mes n . El nombre de parelles de conills P_{n+1} al mes $n+1$ és la suma del nombre P_n de parelles al mes n i del nombre de parelles nounades al mes $n+1$. Però només les parelles de conills que tenen almenys dos mesos d'edat donen lloc a noves parelles de conills al mes $n+1$. Són les parelles que ja existien al mes $n-1$ i eren P_{n-1} . Així que

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

Es tracta d'una relació de recurrència: dona la població del mes $n+1$ en funció de la població dels mesos anteriors. Per tant, Fibonacci podia construir fàcilment la següent taula, on $1+1=2$, $1+2=3$, $2+3=5$, $3+5=8$, etc.

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

De fet, Fibonacci va considerar com a condició inicial la situació al mes $n=2$. Com $P_{14} = 144 + 233 = 377$, finalment va obtenir 377 parelles de conills dotze mesos després del seu punt de partida. Va observar que aquesta successió de nombres podia continuar indefinidament.

Després de l'any 1202, Fibonacci va escriure diversos llibres més, com *Practica geometriae* el 1220 i *Liber quadratorum* («El llibre dels quadrats») el 1225. La seva fama el va portar a entrevistar-se amb l'emperador Frederic II, que apreciava la ciència. El 1240, la República de Pisa va concedir a Fibonacci una pensió anual. Es desconeix l'any de la seva mort.

Durant els segles posteriors, el problema dels conills de Fibonacci va caure en l'oblit i no va tenir cap influència en el desenvolupament dels models matemàtics de la dinàmica de poblacions. Diversos científics es van trobar amb la mateixa successió de nombres als seus estudis, però no es van referir a Fibonacci ni a cap població. Diversos llibres de Kepler contenen l'observació

que la relació P_{n+1}/P_n convergeix, quan n tendeix a l'infinit, al nombre auri $\phi = (1 + \sqrt{5})/2$. Aquest és un cas particular d'una propietat comuna a la majoria dels models de poblacions: la tendència a augmentar geomètricament (vegeu els capítols 3 i 21). El 1728 Daniel Bernoulli va obtenir la fórmula exacta

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

en estudiar sèries recurrents en general. Les obres completes de Fibonacci es van publicar al segle XIX. A partir de llavors, la successió (P_n) es troba als llibres de matemàtiques recreatives amb el nom de successió de Fibonacci.

És evident que, per a modelitzar una població de conills, les hipòtesis que condueixen a la successió de Fibonacci estan lluny de ser realistes: no hi ha mortalitat, no hi ha separació de sexes, etc. L'interès per aquesta successió en les últimes dècades en biologia prové del fet que diverses plantes contenen estructures que impliquen alguns dels nombres P_n , per exemple, 8 i 13 en les pinyes o 34 i 55 als gira-sols. Fins i tot existeix una revista científica, *The Fibonacci Quarterly*, dedicada per complet a les propietats i aplicacions de la successió de Fibonacci.

Més informació

1. Bernoulli, D.: Observaciones de seriebus... *Comment. Acad. Sci. Imp. Petropolitanae* 3, 85–100 (1728/1732) → *Die Werke von Daniel Bernoulli*, Band 2, Birkhäuser, Basel, 1982, 49–64.
2. Sigler, L.E.: *Fibonacci's Liber Abaci*. Springer (2002).
3. Vogel, K.: Leonardo Fibonacci. In: Gillespie, C.C. (ed.) *Dictionary of Scientific Biography*, vol. 4, 604–613. Scribner, New York (1971)

Capítol 2

Taula de vida d'Halley (1693)

El 1693, el famós astrònom anglès Edmond Halley va estudiar els registres de naixements i defuncions de la ciutat de Breslau, que havien estat transmesos a la *Royal Society* per Caspar Neumann. Va elaborar una taula de vida que mostrava el nombre de persones que sobreviuen fins a cada edat d'una cohort nascuda el mateix any. També va utilitzar la seva taula per a calcular el preu de les rendes vitalícies. Aquest capítol recorda aquest treball i el situa en el context de la vida d'Halley i dels primers desenvolupaments de l'«aritmètica política» i la teoria de la probabilitat, que van interessar a persones com Graunt, Petty, De Witt, Hudde, Huygens, Leibniz i de Moivre.

Edmond Halley va néixer prop de Londres el 1656. El seu pare era un ric fabricant de sabó. Edmond es va interessar per l'astronomia des de molt jove. Va començar a estudiar al *Queen's College* de la Universitat d'Oxford. Quan es va inaugurar l'Observatori de Greenwich el 1675, Halley ja va poder visitar a Flamsteed l'Astrònom Real. Va interrompre els seus estudis del 1676 al 1678 per anar a l'illa de Santa Elena i establir un catàleg de les estrelles que es poden veure des de l'hemisferi sud. Al seu retorn a Anglaterra va esdevenir membre de la *Royal Society*. Va publicar també les observacions que havia realitzat sobre la circulació dels vents durant el seu viatge a Santa Elena. El 1684 va visitar Newton a Cambridge per a discutir la relació entre les lleis del moviment planetari de Kepler i la força d'atracció exercida pel Sol. Va animar Newton a escriure els famosos *Principis matemàtics de la filosofia natural*, llibre que finalment va publicar amb els seus propis recursos. En aquell temps treballava com a secretari de la *Royal Society*. El 1689 va dissenyar una campana per a bussejar sota l'aigua, que ell mateix va provar.

Més o menys a la mateixa època, Caspar Neumann, un teòleg que vivia a Breslau, recopilava dades sobre el nombre de naixements i morts a la seva ciutat. Breslau pertanyia a l'imperi dels Habsburg (ara forma part de Polònia i es diu Wrocław). Les dades incloïen l'edat a la que havia mort la gent. D'aquesta manera, es podia construir una taula de vida que mostrava la probabilitat de sobreviure fins a una edat determinada.

La primera taula de vida es va publicar a Londres l'any 1662 en un llibre titulat *Observacions naturals i polítiques sobre les estadístiques de mortali-*

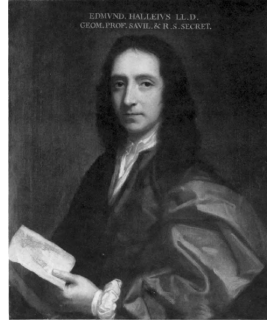


Figura 2.1:
Halley (1656–1742)

tat. Aquest llibre sol considerar-se com el text fundador tant de l'estadística com de la demografia i té una estranya particularitat: la gent es pregunta encara avui si va ser escrit per John Graunt, comerciant londinenc i autor indicat a la portada del llibre, o pel seu amic William Petty, un dels fundadors de la *Royal Society*. En qualsevol cas, la taula de vida continguda al llibre tractava d'aprofitar el contingut dels butlletins que informaven regularment dels enterraments i batejos a Londres des de principis del segle XVII. Aquests butlletins s'utilitzaven principalment per a informar la població sobre les recurrents epidèmies de pesta. Aquest és el motiu pel qual indicaven la causa de la mort i no l'edat a la que es moria. Per a obtenir una taula de vida que donés la probabilitat de sobreviure en funció de l'edat, Graunt o Petty van haver de descobrir com es relacionaven les diferents causes de mort amb els grups d'edat. A causa d'això, la seva taula de vida podia estar subjecta a grans errors. Malgrat tot, el llibre va tenir molt d'èxit, amb cinc edicions entre el 1662 i el 1676. Diverses ciutats d'Europa havien començat a publicar butlletins similars al de Londres.

Així fou com gairebé trenta anys després d'aquesta primera taula de vida, seguint el suggeriment de Leibniz, Neumann va enviar a Henry Justel, secretari de la *Royal Society*, les seves dades demogràfiques de la ciutat de Breslau pels anys 1687–1691. Justel va morir poc després, i les dades van passar a mans de Halley que va analitzar-les i el 1693 va publicar les seves conclusions a *Philosophical Transactions of the Royal Society*. El seu article es titula *Una estimació dels graus de mortalitat de la humanitat, extreta de curioses taules dels naixements i funerals a la ciutat de Breslau, amb un intent de determinar el preu de les rendes vitalícies*.

Durant el període de cinc anys estudiat, Halley va observar que el nombre de naixements a Breslau era més o menys igual al nombre de morts, de manera que la població total era gairebé constant. Per a simplificar l'anàlisi, va

suposar que la població es trobava exactament en estat estacionari: el nombre anual de naixements (denotat P_0), la població total, la població d'edat k (P_k) i el nombre anual de morts a l'edat k (D_k) són tots constants a mesura que passa el temps. Això posa en relleu una propietat addicional interessant de les dades de Breslau, ja que una simplificació com aquesta no hauria estat possible per a una ciutat de creixement ràpid com Londres, on les estadístiques també estaven esbiaixades pel flux de població procedent del camp.

Taula 2.1: La taula de vida d'Halley que mostra la població P_k d'edat k .

k	P_k	k	P_k	k	P_k	k	P_k	k	P_k	k	P_k
1	1.000	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

Les dades de Breslau tenien una mitjana de 1 238 naixements per any: aquest és el valor que Halley va prendre per a P_0 . En principi, també podia calcular a partir de les dades la mitjana anual D_k del nombre de morts entre les persones d'edat k per a tot $k \geq 0$. Utilitzant la fórmula

$$P_{k+1} = P_k - D_k, \quad (2.1)$$

va poder construir la taula 2.1 que dona P_k . De manera inversa, es poden trobar els valors de D_k que va utilitzar a partir de la fórmula $D_k = P_k - P_{k+1}$: $D_0 = 238$, $D_1 = 145$, $D_2 = 57$, $D_3 = 38$ i així successivament. De fet, Halley va reordenar una mica els seus resultats, alguns cops per obtenir nombres rodons (és el cas de D_1 , que s'ha modificat lleugerament perquè $P_1 = 1.000$) i d'altres per suavitzar algunes irregularitats causades pel nombre petit de morts a edats avançades en un estudi de cinc anys. Calculant la suma de tots els nombres P_k de la taula, Halley va obtenir una estimació de la població

total de Breslau pròxima als 34.000 habitants¹.

En resum, aquest mètode tenia el gran avantatge de no requerir un cens general, sinó només el coneixement del nombre de naixements i defuncions i de l'edat a la qual morien les persones durant uns anys.

La taula de vida d'Halley va servir de referència per a diversos treballs durant el segle XVIII (vegeu el capítol 4). En efecte, encara que els valors de P_k eren específics de la ciutat de Breslau, es podia considerar que la relació P_{k+1}/P_k era la probabilitat de sobreviure fins a l'edat $k+1$ sabent que ja s'havia assolit l'edat k . Aquesta probabilitat podria utilitzar-se raonablement per a les poblacions d'altres ciutats europees de l'època. Per exemple, caldria esperar que un nen d'un any tingués una probabilitat de 661 sobre 1.000 d'arribar als 10 anys o de 598 sobre 1.000 d'arribar als 20.

Halley també va utilitzar la seva taula de vida per a calcular el preu de les rendes vitalícies. Durant els segles XVI i XVII, diverses ciutats i estats havien venut aquest tipus de rendes vitalícies als seus ciutadans per a recaptar diners. Els compradors rebien cada any, fins a la seva mort, una quantitat fixa de diners, que equivalia a un determinat percentatge de la suma pagada inicialment, sovint al doble del tipus d'interès de l'època, però independentment de l'edat del comprador. Per descomptat, la institució s'arriscava a la fallida si massa persones amb una esperança de vida molt llarga compraven aquestes rendes vitalícies. El problema no podia abordar-se correctament sense una taula de vida fiable.

L'any 1671 Johan De Witt, primer ministre d'Holanda, i Johannes Hudde, un dels alcaldes de la ciutat d'Amsterdam, ja havien pensat en el problema de calcular el preu de les rendes vitalícies. Tement una invasió de les tropes franceses, volien recaptar diners per a reforçar l'exèrcit. Disposaven de dades relatives a les persones que havien comprat rendes vitalícies diverses dècades abans, en particular l'edat en la que havien comprat les rendes i l'edat a la que havien mort. Havien aconseguit calcular el preu de les rendes vitalícies de manera més o menys correcta, però el seu mètode va caure en l'oblit posteriorment. Holanda va ser envaïda l'any següent i De Witt va ser linxat per la multitud.

Halley va tornar a considerar el problema el 1693 amb la taula de vida de Breslau i assumint un tipus d'interès del 6%. El mètode de càlcul és senzill. Sigui i el tipus d'interès. Sigui R_k el preu al qual una persona d'edat k pot comprar una renda vitalícia de, per exemple, una lliura a l'any. Aquesta persona té una probabilitat P_{k+n}/P_k de continuar viva a l'edat $k+n$. La lliura que l'Estat es compromet a pagar si arriba a aquesta edat pot obtenir-se col·locant $1/(1+i)^n$ de la suma inicial al tipus d'interès i . Per tant, si es fa la suposició

¹Per als majors de 84 anys, Halley simplement esmenta que són 107.

simplificadora que la suma inicial s'utilitza únicament per a pagar les rendes vitalícies, el preu hauria de ser

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

Halley va obtenir així la taula 2.2, que mostra el factor R_k pel qual cal multiplicar l'anualitat desitjada per a obtenir la suma inicial necessària. Un home de 20 anys obtindria cada any $1/12,78 \approx 7,8\%$ de la suma inicial. En canvi, un home de 50 anys obtindria $1/9,21 \approx 10,9\%$, perquè li quedarien menys anys de vida. Observeu que el doble del tipus d'interès del 6% correspondria a una anualitat igual al 12% de la suma inicial, o el que és el mateix, a un preu igual al $1/12\% = 8,33$ vegades l'anualitat.

Taula 2.2: Factor multiplicador que dona el preu de les rendes vitalícies.

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

Els càlculs són, per descomptat, bastant tediosos. No obstant això, Halley va poder utilitzar taules de logaritmes per a obtenir el terme general $P_{k+n}/(1+i)^n$ més ràpidament. Com no va mostrar valors per a P_k superiors a 84 anys, no és possible comprovar els seus càlculs amb exactitud. Finalment, el treball d'Halley no va tenir cap repercussió immediata: durant diverses dècades, les rendes vitalícies a Anglaterra i en altres llocs van continuar venent-se a un preu independent de l'edat del comprador i a un preu molt més baix del que podria ser, per exemple 7 vegades la renda vitalícia.

Les qüestions derivades de les taules de vida van interessar a molts científics de l'època d'Halley. L'holandès Christiaan Huygens, autor el 1657 del primer fullet dedicat a la teoria de les probabilitats, va discutir el 1669 en la seva correspondència amb el seu germà Graunt la taula de vida i el càlcul de l'esperança de vida². Uns anys abans de posar a Neumann en contacte amb la *Royal Society*, Leibniz també va escriure sobre el càlcul de l'esperança de vida en un assaig que va romandre inèdit. El 1709 va ser el torn de Nikolaus I Bernoulli. El 1725, Abraham de Moivre va publicar tot un *Tractat sobre les anualitats*. Es va adonar, en particular, que el preu R_k podia calcular-se fàcilment per a les edats antigues, ja que la fórmula (2.2) només contenia uns

²L'esperança de vida a l'edat k ve donada per la fórmula (2.2) amb $i = 0$.

pocs termes. Es podia llavors utilitzar la fórmula de recurrència cap enrere

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

que es demostra fàcilment a partir de (2.2). Utilitzant el valor que Halley dona per al preu als 70 anys, es poden comprovar els altres valors³ de la taula 2.2.

Després d'aquesta pausa centrada en la demografia, Halley va tornar als seus principals temes de recerca. Entre el 1698 i el 1700 va navegar per l'oceà Atlàntic per a dibuixar un mapa del camp magnètic de la Terra. El 1704 va ser nomenat catedràtic de la Universitat d'Oxford. L'any següent va publicar un llibre sobre cometes i va predir que el cometa del 1682, que Kepler havia observat el 1607, tornaria el 1758: va passar a ser conegut com «el cometa Halley». També va publicar una traducció del llibre d'Apolonio de Perga sobre les còniques. El 1720 va substituir a Flamsteed com a astrònom real. Va intentar resoldre el problema de la determinació precisa de la longitud al mar a partir de l'observació de la Lluna, un problema de gran importància pràctica per a la navegació. Va morir a Greenwich el 1742 a l'edat de 86 anys.

Més informació

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*, 3rd edn. London (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; with an attempt to ascertain the price of annuities upon lives. *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 (2001)
6. Koch, P.: Caspar Neumann. In: *Ibid.*, 29–32.
7. Le Bras, H.: *Naissance de la mortalité*. Gallimard, Paris (2000)

³Sembla que hi ha alguns errors en la taula, en particular per a les edats de 5 i 15 anys.

Capítol 3

Euler i el creixement geomètric de poblacions (1748–1761)

Euler va escriure en diverses ocasions sobre la dinàmica de poblacions. En el seu tractat del 1748, *Introducció a l'anàlisi de l'infinit*, el capítol dedicat a la funció exponencial contenia quatre exemples sobre el creixement exponencial d'una població. L'any 1760 va publicar un article que combinava aquest creixement exponencial amb una estructura d'edats per a la població. Aquest treball és un precursor de la teoria de poblacions estables, que es va desenvolupar durant el segle XX i juga un paper important en demografia. El 1761, Euler també va ajudar Süßmilch amb la segona edició del seu tractat de demografia. Va elaborar un interessant model, que és una espècie de variant de la successió de Fibonacci, però no va publicar la seva anàlisi detallada.

Leonhard Euler va néixer el 1707 a Basilea, Suïssa. El seu pare era ministre protestant. El 1720, Euler va començar a estudiar a la universitat. També va rebre classes particulars de matemàtiques de Johann Bernoulli, un dels matemàtics més famosos de la generació posterior a Leibniz i Newton. Es va fer amic de dos dels fills de Johann Bernoulli: Nikolaus II i Daniel. El 1727, Euler es va incorporar, juntament amb Daniel, a la recentment creada Acadèmia de Ciències de Sant Petersburg. A més a més de les matemàtiques, es va interessar per la física i per molts altres temes científics i tècnics. L'any 1741 el rei Frederic II de Prússia el va nomenar director de la secció de matemàtiques de l'Acadèmia de Ciències de Berlín. Euler va publicar un nombre considerable d'articles i llibres sobre tots els aspectes de la mecànica (astronomia, elasticitat, fluids, sòlids) i de les matemàtiques (teoria de nombres, àlgebra, sèries infinites, funcions elementals, nombres complexos, càlcul diferencial i integral, equacions diferencials ordinàries i en derivades parcials, optimització, geometria), però també sobre demografia. Va ser el matemàtic més prolífic de la seva època.

L'any 1748, Euler va publicar un tractat en llatí titulat *Introducció a l'anàlisi de l'infinit*. Al capítol sobre exponencials i logaritmes va considerar sis exemples: un sobre la teoria matemàtica de les escales musicals, un altre sobre la devolució d'un préstec amb interessos i quatre sobre la dinàmica de



Figura 3.1:
Euler (1707–1783)

poblacions. En aquests últims, Euler va suposar que la població P_n l'any n satisfà

$$P_{n+1} = (1+x)P_n$$

per a tot nombre enter n . La taxa de creixement x és un nombre real positiu. Partint d'una condició inicial P_0 , la població l'any n ve donada per

$$P_n = (1+x)^n P_0 .$$

Això s'anomena creixement geomètric o exponencial. Al primer exemple es pregunta:

«Si la població d'una determinada regió augmenta anualment en una trentena part i en un moment donat hi havia 100.000 habitants, ens agradaria saber la població al cap de 100 anys.»

La resposta és

$$P_{100} = (1 + 1/30)^{100} \times 100.000 \approx 2.654.874.$$

Per a aquest exemple, Euler es va inspirar en el cens de Berlín que es va realitzar l'any 1747 i que va projectar una estimació de 107.224 habitants. El seu càlcul demostra que una població pot multiplicar-se per més de deu en un segle. Això és precisament el que es va observar en el seu moment per a la ciutat de Londres.

Cal tenir en compte que calcular $(1 + 1/30)^{100}$ és molt fàcil amb una calculadora de butxaca moderna. Però a l'època d'Euler calia utilitzar logaritmes per a evitar nombroses multiplicacions a mà i obtenir el resultat ràpidament.

Es calcula primer el logaritme decimal (en base 10) de P_{100} . La propietat fonamental del logaritme $\log(ab) = \log a + \log b$ mostra que

$$\begin{aligned}\log P_{100} &= 100 \log(31/30) + \log(100.000) \\ &= 100(\log 31 - \log 30) + 5.\end{aligned}$$

Els logaritmes havien estat introduïts el 1614 per l'escocès John Napier. El seu amic Henry Briggs havia publicat la primera taula de logaritmes decimals l'any 1617. El 1628 l'holandès Adriaan Vlacq havia completat el treball de Briggs publicant una taula que contenia els logaritmes decimals dels nombres enters de l'1 al 100.000 amb una precisió de deu dígit. Aquest és el tipus de taula que Euler va utilitzar per a obtenir

$$\log 30 \approx 1,477121255, \quad \log 31 \approx 1,491361694,$$

i finalment $\log P_{100} \approx 6,4240439$. Queda per trobar el nombre P_{100} que té per logaritme aquest valor. Com que els logaritmes decimals dels enters de l'1 al 100.000 van de 0 a 5, es busca el logaritme de $P_{100}/100$, que és 4,4240439 i, per tant, dins del rang de valors de la taula. Es pot comprovar a la taula de logaritmes que $\log 26.548 \approx 4,424031809$ i $\log 26.549 \approx 4,424048168$. Substituint la funció logarítmica per una recta entre 26.548 i 26.549, Euler va obtenir que

$$\frac{P_{100}}{100} \approx 26.548 + \frac{4,4240439 - 4,424031809}{4,424048168 - 4,424031809} \approx 26.548,74.$$

Així que $P_{100} \approx 2.654.874$.

El segon exemple relatiu a la dinàmica de poblacions al llibre d'Euler és el següent:

«Atès que després del Diluvi tots els homes descendeixen d'una població de sis, si suposem que la població al cap de dos-cents anys era de 1.000.000, ens agradaria trobar la taxa de creixement anual.»

Atès que

$$10^6 = (1+x)^{200} \times 6,$$

obtenim amb una calculadora de butxaca

$$x = (10^6/6)^{1/200} - 1 \approx 0,061963.$$

Amb les taules de logaritmes cal passar per

$$\log(10^6) = 200 \log(1+x) + \log 6$$

per a obtenir

$$\log(1+x) = (6 - \log 6)/200 \approx 0,0261092$$

i $1+x \approx 1,061963$. Així, Euler va poder concloure que la població augmentaria en $x \approx 1/16$ per any. Per a entendre l'origen d'aquest exemple, cal recordar que els filòsofs contemporanis començaven a negar la veracitat dels relats bíblics. Una lectura literal fixaria el moment del Diluvi entorn del 2350 a.C. amb els següents supervivents: Noè, els seus tres fills i les seves esposes. El llibre del Gènesi diu:

«A partir d'aquests tres fills de Noè, es va poblar tota la terra.»

Una taxa de creixement de la població de $1/16$ (o $6,25\%$) per any després del Diluvi no li semblava a Euler massa irreal. Sent fill d'un ministre protestant i havent estat religiós tota la vida, va concloure:

«Per aquesta raó és bastant ridícul per als incrèduls objectar que en un espai de temps tan curt no es va poder poblar tota la terra començant per un sol home.»¹

Euler també es va adonar que si el creixement hagués continuat al mateix ritme fins a 400 anys després del Diluvi, la població hauria estat de $(1+x)^{400} \times 6 = (10^6/6)^2 \times 6 \approx 166$ mil milions:

«No obstant això, tota la terra mai podria sostenir aquesta població.»

Aquesta idea seria força desenvolupada per Malthus mig segle després (vegeu el capítol 5).

El tercer exemple d'Euler pregunta:

«Si cada segle la població humana es duplica, quina és la taxa de creixement anual?»

Com que

$$(1+x)^{100} = 2,$$

¹Al llibre publicat per Graunt el 1662 (vegeu el capítol 2), es troba una observació similar:

«Una parella, és a dir, Adam i Eva, duplicant-se cada 64 anys dels 5.160 anys, que és l'edat del món segons les Escripures, produiria molta més gent de la que hi ha ara. Per tant, el món no té més de 100 mil anys, com alguns imaginin vanament, ni més del que diuen les Escripures.»

obtenim amb una calculadora de butxaca $x = 2^{1/100} - 1 \approx 0,00695$. Amb les taules de logaritmes, $100 \log(1+x) = \log 2$. Així que $\log(1+x) \approx 0,0030103$ i $1+x \approx 1,00695$. Per tant, la població creix $x \approx 1/144$ cada any. El quart i últim exemple es planteja de la mateixa manera:

«Si la població humana augmenta anualment en $1/100$, ens agradaria saber quant temps trigarà la població en multiplicar-se per deu.»

Amb

$$(1 + 1/100)^n = 10,$$

trobem $n \log(101/100) = 1$. Així que $n = 1/(\log 101 - 2) \approx 231$ anys. Això és tot el que es pot trobar a la *Introducció a l'anàlisi de l'infinit* del 1748 sobre dinàmica de poblacions. Euler tornaria a tractar aquest tema amb més profunditat alguns anys després.

L'any 1760 va publicar a les actes de l'Acadèmia de Ciències de Berlín un treball titulat *Una recerca general sobre la mortalitat i la multiplicació de l'espècie humana*. Aquest treball era una espècie de síntesi entre la seva anàlisi prèvia del creixement geomètric de les poblacions i els estudis anteriors sobre les taules de vida (vegeu el capítol 2). Euler va considerar, per exemple, el problema

«Conegut el nombre de naixements i morts que es produeixen en el transcurs d'un any, trobeu el nombre total d'individus vius i el seu increment anual, per a una hipòtesi de mortalitat donada.»

Euler va assumir aquí que es coneixien els següents nombres:

- el nombre de naixements B_n durant l'any n ;
- el nombre de morts D_n durant l'any n ;
- la proporció q_k de nounats que arriben a l'edat $k \geq 1$ ².

Sigui P_n la població a l'any n . Euler va fer dues suposicions implícites addicionals:

- la població augmenta geomètricament: $P_{n+1} = rP_n$ (fixem $r = 1+x$);
- la relació entre els naixements i la població és constant: $B_n/P_n = m$.

Aquests dos supòsits impliquen que el nombre de naixements augmenta geomètricament i al mateix ritme: $B_{n+1} = rB_n$. Euler va considerar llavors l'estat de la població en un interval de cent anys, diguem entre els anys $n = 0$

²Més concretament, que continuïn vius al principi de l'any del seu aniversari k -èsim.

i $n = 100$, suposant que ningú sobreviu més enllà dels cent anys. Per fer més entenedora la presentació, diguem $P_{k,n}$ ($k \geq 1$) a la població viva al principi de l'any n , que va néixer l'any $n - k$. Diguem $P_{0,n} = B_n$ al nombre de naixements durant l'any n . A partir de la definició del coeficient de supervivència q_k , tenim que $P_{k,n} = q_k P_{0,n-k} = q_k B_{n-k}$. Per tant,

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

Dividint aquesta equació entre $r^{100} P_0$, obtenim

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

Aquesta equació de vegades s'anomena «equació d'Euler» en demografia. Comptant els naixements i les morts per separat, obtenim

$$r P_n = P_{n+1} = P_n - D_n + B_{n+1} = P_n - D_n + r B_n. \quad (3.2)$$

Per tant, el nombre de morts també augmenta geomètricament: $D_{n+1} = r D_n$. A més a més,

$$\frac{1}{m} = \frac{P_n}{B_n} = \frac{D_n/B_n - r}{1 - r}. \quad (3.3)$$

Substituint-ho a l'equació (3.1), arribem finalment a l'equació

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

on només queda una incògnita: r . Això és el que se sol anomenar una equació implícita perquè no podem extreure r com a funció dels altres paràmetres. Però podem calcular els costats esquerre i dret de l'equació (3.4) per a un valor concret de r i fer variar la r fins que els dos costats siguin iguals. El valor de r obtingut d'aquesta manera dona la taxa de creixement $x = r - 1$ de la població. Observeu que a partir de les equacions (3.1) i (3.3), obtenim per a la població P_n la següent expressió:

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

Quan la població és estacionària ($r = 1$), aquesta expressió és la mateixa que la utilitzada per Halley per a estimar la població de la ciutat de Breslau (vegeu el capítol 2).

Euler també va considerar la següent qüestió:

«Donades les hipòtesis de mortalitat i fecunditat, si es coneix el nombre total d'individus vius, trobeu quants n'hi haurà de cada edat.»

Atès que es coneixen els coeficients de supervivència q_k i el coeficient de fertilitat m , la taxa de creixement r pot calcular-se a partir de l'equació (3.1). Durant l'any n , el nombre de persones nascudes l'any $n - k$ és $q_k B_{n-k} = q_k B_n / r^k$ (amb $q_0 = 1$). Per tant, la proporció de la població total que té edat k és

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

Aquesta proporció és constant. Utilitzant la terminologia de Lotka (vegeu el capítol 10), es diu que la població és «estable»: la piràmide d'edats manté la mateixa forma al llarg del temps.

A continuació, Euler va reexaminar el problema de la construcció d'una taula de vida quan la població no és estacionària, sinó que augmenta geomètricament:

«Conegut el nombre total d'individus vius, així com el nombre de naixements i el nombre de morts de cada edat en el transcurs d'un any, trobeu la llei de mortalitat.»

Per llei de mortalitat, Euler entenia el conjunt de coeficients de supervivència q_k . La població total se suposa ara coneguda a través d'un cens, la qual cosa no era el cas d'Halley (vegeu el capítol 2). L'equació (3.2) mostra que la taxa de creixement és $r = (P_n - D_n) / (P_n - B_n)$. Sigui $D_{k,n}$ el nombre de persones que moren a l'edat k durant l'any n : aquestes persones van néixer l'any $n - k$. Per tant, $D_{k,n} = (q_k - q_{k+1}) B_{n-k}$. Però $B_{n-k} = B_n / r^k$. Per tant, els coeficients de supervivència q_k poden calcular-se amb la fórmula de recurrència $q_{k+1} = q_k - r^k D_{k,n} / B_n$ per a tot $k \geq 0$, amb $q_0 = 1$. Aquesta fórmula multiplicada per B_n retorna la fórmula (2.1) utilitzada per Halley per al cas estacionari $r = 1$. Euler va insistir, no obstant, en el fet que el seu mètode per a calcular els coeficients de supervivència q_k suposa que la població augmenta regularment, excloent accidents com epidèmies de pesta, guerres, fams, etc. Si els censos de l'època d'Euler haguessin registrat l'edat de les persones (com a Suècia), aquesta suposició hauria estat innecessària i els coeficients q_k podrien haver-se calculat més fàcilment.

Donats els coeficients de supervivència q_k , Euler també va mostrar com calcular el preu de les rendes vitalícies. No va esmentar els treballs d'Halley o de Moivre sobre aquest tema. Euler va utilitzar un tipus d'interès del 5% i la taula de vida publicada el 1742 per l'holandès Willem Kersseboom.

Euler no era l'únic científic interessat en la demografia a l'Acadèmia de Berlín. El seu col·lega Johann Peter Süssmilch havia publicat l'any 1741 un tractat en alemany titulat *L'ordre diví en els canvis de la generació humana, a través del naixement, les morts i la procreació d'aquests establerts*, que es considera avui en dia com el primer tractat dedicat completament a la demografia. Süssmilch també havia escrit el llibre *Sobre el ràpid creixement de la ciutat de Berlín* el 1752.

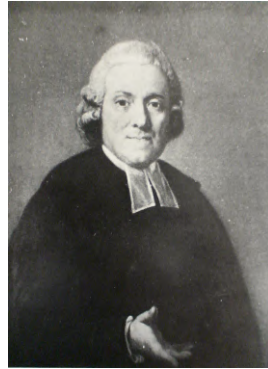


Figura 3.2:
Süssmilch (1707–1767)

L'any 1761, Süssmilch va publicar una segona edició del seu tractat. Al capítol titulat *Sobre la taxa de creixement i el temps de duplicació de les poblacions*, va incloure un interessant model matemàtic que Euler havia elaborat per a ell. El model era similar al de Fibonacci (vegeu el capítol 1) però per a una població humana. Partint d'una parella (un home i una dona) de 20 anys l'any 0, Euler va suposar que les persones moren als 40 anys i es casen als 20 anys, mentre que cada parella té sis fills: dos fills (un nen i una nena) als 22 anys, uns altres dos als 24 anys i els dos últims als 26 anys. Comptant els anys de dos en dos, de manera que B_i és el nombre de naixements durant l'any $2i$, Euler conclou que

$$B_i = B_{i-11} + B_{i-12} + B_{i-13} \quad (3.5)$$

per a tot $i \geq 1$. Les condicions inicials corresponen a $B_{-12} = 0$, $B_{-11} = 0$, $B_{-10} = 2$ i $B_i = 0$ per a $-9 \leq i \leq 0$. Així, Euler va poder calcular el nombre de naixements, com es mostra a la segona columna de la taula 3.1. El nombre de morts D_i l'any $2i$ és llavors igual al nombre de naixements l'any $2i - 40$: $D_i = B_{i-20}$ per a $i \geq 10$ mentre que $D_i = 0$ per a $i \leq 9$. Quant al nombre P_i de persones vives l'any $2i$, és igual al nombre de persones vives l'any $2i - 2$, més el nombre de naixements l'any $2i$, menys el nombre de morts l'any $2i$: $P_i = P_{i-1} + B_i - D_i$.

Taula 3.1: La taula de Euler.

i	naixements	morts	vius
0	0	0	2
1	2	0	4
2	2	0	6
3	2	0	8
4	0	0	8
5	0	0	8
6	0	0	8
7	0	0	8
8	0	0	8
9	0	0	8
10	0	2	6
11	0	0	6
12	2	0	8
13	4	0	12
14	6	0	18
15	4	0	22
16	2	0	24
17	0	0	24
18	0	0	24
19	0	0	24
20	0	0	24
21	0	2	22
22	0	2	20
23	2	2	20
24	6	0	26
25	12	0	38
26	14	0	52
27	12	0	64
28	6	0	70
29	2	0	72
30	0	0	72
31	0	0	72
32	0	2	70
33	0	4	66
34	2	6	62
35	8	4	66
36	20	2	84
37	32	0	116
38	38	0	154
39	32	0	186

i	naixements	morts	vius
40	20	0	206
41	8	0	214
42	2	0	216
43	0	2	214
44	0	6	208
45	2	12	198
46	10	14	194
47	30	12	212
48	60	6	266
49	90	2	354
50	102	0	456
51	90	0	546
52	60	0	606
53	30	0	636
54	10	2	644
55	2	8	638
56	2	20	620
57	12	32	600
58	42	38	604
59	100	32	672
60	180	20	832
61	252	8	1.076
62	282	2	1.356
63	252	0	1.608
64	180	0	1.788
65	100	2	1.886
66	42	10	1.918
67	14	30	1.902
68	16	60	1.858
69	56	90	1.824
70	154	102	1.876
71	322	90	2.108
72	532	60	2.580
73	714	30	3.264
74	786	10	4.040
75	714	2	4.752
76	532	2	5.282
77	322	12	5.592
78	156	42	5.706
79	72	100	5.678

i	naixements	morts	vius
80	86	180	5.584
81	226	252	5.558
82	532	282	5.808
83	1.008	252	6.564
84	1.568	180	7.952
85	2.032	100	9.884
86	2.214	42	12.056
87	2.032	14	14.074
88	1.568	16	15.626
89	1.010	56	16.580
90	550	154	16.976
91	314	322	16.968
92	384	532	16.820
93	844	714	16.950
94	1.766	786	17.930
95	3.108	714	20.324
96	4.608	532	24.400
97	5.814	322	29.892
98	6.278	156	36.014
99	5.814	72	41.756
100	4.610	86	46.280
101	3.128	226	49.182
102	1.874	532	50.524
103	1.248	1.008	50.764
104	1.542	1.568	50.738
105	2.994	2.032	51.700
106	5.718	2.214	55.204
107	9.482	2.032	62.654
108	13.530	1.568	74.616
109	16.700	1.010	90.306
110	17.906	550	107.662
111	16.702	314	124.050
112	13.552	384	137.218
113	9.612	844	145.986
114	6.250	1.766	150.470
115	4.664	3.108	152.026
116	5.784	4.608	153.202
117	10.254	5.814	157.642
118	18.194	6.278	169.558
119	28.730	5.814	192.474

Aquest capítol del llibre de Süßmilch acaba amb una observació que podria haver-se fet ja sobre la successió de Fibonacci:

«El gran desordre que sembla prevaler a la taula d'Euler no impedeix que el nombre de naixements segueixi una espècie de progressió que es diu sèrie recurrent [...] Independentment del desordre inicial d'aquestes progressions, esdevenen una progressió geomètrica si no s'interrompen i els desordres del principi s'esvaeixen a poc a poc i desapareixen gairebé per complet.»

El llibre no diu res més sobre les matemàtiques d'aquest model de població. Malgrat tot, Euler va portar l'estudi molt més enllà en un manuscrit titulat *Sobre la multiplicació de la raça humana*, que va romandre inèdit en vida. Buscant una solució de l'equació (3.5) de la forma $B_i = cr^i$, és a dir, de la forma d'una progressió geomètrica, va obtenir després d'una simplificació una equació polinòmica de grau 13:

$$r^{13} = r^2 + r + 1. \quad (3.6)$$

Va buscar una solució pròxima a $r = 1$ i va observar, utilitzant una taula de logaritmes per al càlcul de r^{13} , que

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

De manera que l'equació (3.6) té una arrel entre 1,09 i 1,10. Aproximant la funció $1 + r + r^2 - r^{13}$ per un segment de recta en aquest interval, Euler va obtenir

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

Com els anys es compten de dos en dos, el nombre de naixements tendeix a multiplicar-se per \sqrt{r} cada any. Aquest nombre es duplica cada n anys si $(\sqrt{r})^n = 2$, és a dir, cada $n = 2 \log 2 / \log r \approx 15$ anys. Com que asimptòticament $B_i \approx cr^i$ i com que el nombre D_i de morts l'any $2i$ és igual a B_{i-20} , obtenim $D_i \approx B_i / r^{20}$ amb $r^{20} \approx 6,25$. El nombre de naixements és aproximadament sis vegades superior al nombre de morts. Essent el nombre P_i de persones vives l'any $2i$ igual a $B_i + B_{i-1} + \dots + B_{i-19}$, obtenim també que

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

La població total és aproximadament deu vegades superior al nombre de naixements.

La prova que la successió (B_i) que es mostra a la taula 3.1 efectivament creix asimptòticament com r^i és més complicada. Se sabia des dels treballs d'Abraham de Moivre sobre sèries recurrents que, introduint la funció generatriu $f(x) = \sum_{i=0}^{+\infty} B_i x^i$, es podria expressar $f(x)$ com una funció racional. Euler havia explicat el mètode a la seva *Introducció a l'anàlisi de l'infinit* el 1748: la relació recurrent (3.5) dona efectivament

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

De manera que

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

Euler sabia que aquesta funció racional podia descompondre's de la forma

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

essent els nombres x_1, \dots, x_{13} les arrels reals o complexes de l'equació $1 - x^{11} - x^{12} - x^{13} = 0$. Així que

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

Com que B_i és el coeficient de x^i a $f(x)$, Euler va obtenir que

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

quan $i \rightarrow +\infty$, on x_k és l'arrel amb el mòdul més petit. En altres paraules, B_i tendeix a créixer geomètricament com $(1/x_k)^i$. Quedava per observar que x_k és una arrel de l'equació $1 - x^{11} - x^{12} - x^{13} = 0$ si i només si $r = 1/x_k$ és una arrel de l'equació (3.6). Alguns detalls de la prova van ser finalment aclarits per Gumbel el 1916.

Süssmilch va publicar una tercera edició del seu tractat l'any 1765 i va morir a Berlín el 1767. Estant en mala relació amb el rei de Prússia, Euler va tornar a Sant Petersburg l'any 1766. Malgrat perdre la vista, va continu-

ar publicant un gran nombre d'obres amb l'ajuda dels seus fills i col·legues, especialment sobre àlgebra, càlcul integral, òptica i construcció naval. Les seves *Cartes sobre diversos temes de la filosofia natural dirigits a una princesa alemanya*, escrites a Berlín entre el 1760 i el 1762, es van publicar entre el 1768 i el 1772 i es van convertir en un èxit de vendes a tota Europa. Euler va morir a Sant Petersburg l'any 1783. La seva contribució a la demografia matemàtica, especialment la seva anàlisi de la piràmide d'edat estable en una població de creixement exponencial, no es redescobriria fins al segle XX (vegeu els capítols 10 i 21).

Més informació

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, Leipzig (1922). gallica.bnf.fr
4. Fellmann, E.A.: *Leonhard Euler*. Birkhäuser, Basel (2007)
5. Gumbel, E.J.: Eine Darstellung statistischer Reihen durch Euler. *Jahresber. dtsh. 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 (2001)
8. Süssmilch, J.P.: *Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts aus der Geburt, dem Tode und der Fortpflanzung desselben*. Berlin (1761). echo.mpiwg-berlin.mpg.de
9. Warusfel, A.: *Euler, les mathématiques et la vie*. Vuibert, Paris (2009)

Capítol 4

Daniel Bernoulli, d'Alembert i la inoculació de la verola (1760)

L'any 1760, Daniel Bernoulli va escriure un article sobre la verola. A la seva època hi havia molta controvèrsia entorn de la inoculació, una pràctica que podia protegir la gent però que també podia ser mortal. Va utilitzar la taula de vida d'Halley i algunes dades relatives a la verola per a demostrar que la inoculació era avantatjosa si el risc de morir associat era inferior a l'11 %. La inoculació podia augmentar fins a tres anys l'esperança de vida al néixer. D'Alembert va criticar el treball de Bernoulli, que fou el primer model matemàtic en epidemiologia.

Daniel Bernoulli va néixer l'any 1700 a Groningen (Països Baixos). A la seva família ja hi havia dos matemàtics famosos: el seu pare Johann Bernoulli i el seu oncle Jakob Bernoulli. El 1705, Johann es va traslladar a Basilea, a Suïssa, on va ocupar la càtedra que havia quedat vacant per la mort de Jakob. Johann no volia que el seu fill estudiés matemàtiques. Per això, Daniel es va dedicar a la medicina i es va doctorar l'any 1721 amb una tesi sobre la respiració. Es va traslladar a Venècia i va començar a centrar-se en les matemàtiques, publicant un llibre el 1724. Aquest mateix any va guanyar un premi de l'Acadèmia de Ciències de París per un assaig *Sobre la perfecció del rellogge de sorra en un vaixell a la mar*, i va obtenir una càtedra a la nova Acadèmia de Sant Petersburg. Durant aquests anys, va treballar especialment en les sèries recurrents o en la «paradoxa de Sant Petersburg» en teoria de la probabilitat. El 1733 Daniel Bernoulli va tornar a la Universitat de Basilea, on va ensenyar successivament botànica, fisiologia i física. L'any 1738 va publicar un llibre sobre la dinàmica de fluids que continua sent famós en la història de la física. Cap al 1753 es va interessar, al mateix temps que Euler i d'Alembert, en el problema de les cordes vibrants, que va provocar una important controvèrsia matemàtica.

L'any 1760 va presentar a l'Acadèmia de Ciències de París un treball titulat *Un intent de nova anàlisi de la mortalitat causada per la verola i dels avantatges de la inoculació per a prevenir-la*. La qüestió era si la inoculació (la introducció voluntària d'una petita quantitat de verola menys virulenta en el cos per a protegir-lo contra infeccions posteriors) hauria de fomentar-se



Figura 4.1:
Daniel Bernoulli
(1700–1782)

encara que a vegades fos una operació mortal. Aquesta tècnica era coneguda des de feia temps a l'Àsia i havia estat introduïda el 1718 a Anglaterra per Lady Montagu, esposa de l'ambaixador britànic a l'Imperi Otomà. A França, malgrat la mort del fill major de Lluís XIV a causa de la verola el 1711, va persistir la reticència a la inoculació. Voltaire, que havia sobreviscut a la verola el 1723 i que havia viscut diversos anys a l'exili a Anglaterra observant les últimes innovacions, va advocar per la inoculació en les seves *Cartes filosòfiques* del 1734. El científic francès La Condamine, que també havia sobreviscut a la verola, va advocar per la inoculació a l'Acadèmia de Ciències de París el 1754.

Abans de morir a Basilea el 1759, Maupertuis va animar Daniel Bernoulli a estudiar el problema de la inoculació des d'un punt de vista matemàtic. Més concretament, el repte consistia a trobar una manera de comparar el benefici a llarg termini de la inoculació amb el risc immediat de morir. Per a això, Bernoulli va fer les següents suposicions simplificadores:

- les persones infectades amb verola per primera vegada moren amb una probabilitat p (independent de l'edat) i sobreviuen amb una probabilitat $1 - p$;
- tothom té una probabilitat q d'infectar-se cada any; més concretament, la probabilitat que un individu s'infecti entre l'edat x i l'edat $x + dx$ és $q dx$, on dx és un període de temps infinitesimal;
- les persones que sobreviuen a la verola estan protegides contra noves infeccions durant la resta de la seva vida (han estat immunitzades).

Sigui $m(x)$ la mortalitat a l'edat x deguda a causes diferents de la verola: la probabilitat que un individu mori en un període de temps infinitesimal dx ,

entre l'edat x i l'edat $x + dx$, és $m(x)dx$. Considerant un grup de P_0 persones nascudes el mateix any, anomenarem

- $S(x)$ el nombre de persones «susceptibles» que continuen vives a l'edat x sense haver-se infectat mai de verola¹;
- $R(x)$ el nombre de persones que estan vives a l'edat x i que han sobreviscut a la verola;
- $P(x) = S(x) + R(x)$ el nombre total de persones vives a l'edat x .

El naixement correspon a l'edat $x = 0$. Per tant, $S(0) = P(0) = P_0$ i $R(0) = 0$. Aplicant els mètodes de càlcul que havien estat desenvolupats a la fi del segle XVII per Newton, Leibniz i posteriorment pel seu pare, Daniel Bernoulli va observar que, entre l'edat x i l'edat $x + dx$ (amb dx infinitament petit), cada individu susceptible té una probabilitat qdx d'infectar-se de verola i una probabilitat $m(x)dx$ de morir per altres causes. Per tant, la variació del nombre de persones susceptibles és $dS = -Sqdx - Sm(x)dx$, que porta a l'equació diferencial

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

En aquesta equació, dS/dx s'anomena la derivada de la funció $S(x)$. Durant el mateix interval petit de temps, el nombre de persones que moren de verola és $pSqdx$ i el nombre de persones que sobreviuen a la verola és $(1-p)Sqdx$. A més a més, també hi ha $Rm(x)dx$ persones que moren per causes diferents a la verola. Això porta a una segona equació diferencial:

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

Sumant les dues equacions, obtenim

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

A partir de les equacions (4.1) i (4.3), Bernoulli va poder demostrar que la fracció de persones que encara són susceptibles a l'edat x és

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

¹Més exactament, és l'esperança d'aquest nombre, que pot variar contínuament i no sols en els naturals.

Per a obtenir la fórmula (4.4), Bernoulli va eliminar $m(x)$ de les equacions (4.1) i (4.3):

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

Després de reorganitzar-ho, es dedueix que

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

Observem que el primer membre és la derivada de $f(x) = S(x)/P(x)$, que és la fracció de persones susceptibles en la població d'edat x . Per tant

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

La solució d'aquest tipus d'equació es coneixia des de feia diverses dècades gràcies als treballs de Jakob Bernoulli, oncle del Daniel. Dividint l'equació per f^2 i fixant $g(x) = 1/f(x)$, veiem que $dg/dx = qg - pq$ i que $g(0) = 1/f(0) = 1$. Definint $h(x) = g(x) - p$, obtenim $dh/dx = qh$. Per tant, $h(x) = h(0)e^{qx} = (1-p)e^{qx}$. Finalment $g(x) = (1-p)e^{qx} + p$ y $f(x) = 1/g(x)$.

Per aplicar la seva teoria, Bernoulli va utilitzar la taula de vida d'Halley (vegeu el capítol 2). Aquesta taula dona el nombre de persones que continuen vives al principi de l'any x (amb $x = 1, 2, \dots$) entre una cohort de 1.238 nascuts durant l'any 0. Però en el marc del seu model, Bernoulli necessitava el nombre de persones $P(x)$ que realment assoleixen l'edat x , que és lleugerament diferent. Com Bernoulli, igual que la majoria dels seus contemporanis, no es va adonar de la diferència (l'article d'Halley no és molt explícit, per cert), va mantenir els nombres de la taula d'Halley, excepte el primer nombre 1.238, que va substituir per 1.300 per a obtenir una mortalitat realista durant el primer any de vida. Aquestes xifres apareixen a la segona columna de la taula 4.1.

Bernoulli va triar per a la probabilitat de morir de verola $p = 1/8 = 12,5\%$, que concorda amb les observacions de la seva època. La probabilitat anual de contreure la verola q no es podia estimar directament. Així que Bernoulli probablement va provar diversos valors per a q i finalment va triar aquell pel qual el nombre de morts degudes a la verola després de tots els càlculs, presentats a continuació, era aproximadament $1/13$ del nombre total de morts, una proporció que s'havia observat en diverses ciutats europe-

Taula 4.1: La taula de vida d'Halley i els càlculs de Bernoulli.

Edat x	Vius $P(x)$	Susceptibles $S(x)$	Immunes $R(x)$	Morts per verola	Sense verola $P^*(x)$
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
⋮	⋮	⋮	⋮	⋮	⋮

es. L'elecció de $q = 1/8$ per any va resultar ser un bon ajust, com veurem a continuació².

Amb la fórmula (4.4) i els valors de $P(x)$ de la segona columna de la taula, podem calcular el nombre $S(x)$ de persones susceptibles d'edat x : es tracta de la tercera columna de la taula arrodonida a l'enter més pròxim. La quarta columna mostra el nombre $R(x) = P(x) - S(x)$ de persones d'edat x que han sobreviscut a la verola. La cinquena columna mostra en la fila corresponent a l'edat x el nombre de morts per verola entre l'edat x i l'edat $x + 1$. En teoria aquest nombre hauria de ser la integral $pq \int_x^{x+1} S(t) dt$ però la fórmula $pq [S(x) + S(x + 1)]/2$ dona una bona aproximació, com s'observa a la figura 4.2: l'àrea del trapezi està prop de l'àrea sota la corba, és a dir, de la integral de la funció.

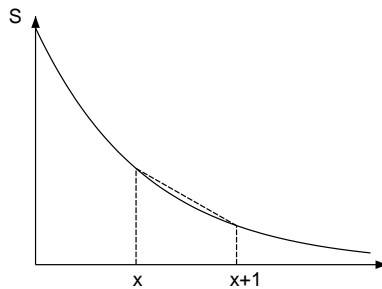


Figura 4.2: L'àrea del trapezi, en traç discontinu, aproxima la integral de la funció S entre x i $x + 1$.

Bernoulli va observar que la suma de tots els nombres de la cinquena columna dona 98 morts per verola abans dels 24 anys. Si continuéssim la taula per a edats més avançades, trobaríem només tres morts més per verola entre les 32 persones que encara són susceptibles als 24 anys. En resum, a partir de 1.300 naixements, el destí de 101 persones és morir de verola. Això és gairebé exactament la fracció esperada $1/13$.

Bernoulli va considerar llavors la situació en la qual la verola fos inculcada a tothom en néixer i, per tant, no causaria cap mort. La verola seria erradicada i la qüestió és estimar l'augment de l'esperança de vida. Partint del mateix nombre de naixements P_0 , diguem $P^*(x)$ al nombre de persones d'edat x quan la verola hagi desaparegut. Aleshores

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

²El fet que p i q siguin iguals és només una coincidència.

Bernoulli va poder demostrar que

$$P^*(x) = \frac{P(x)}{1 - p + pe^{-qx}}, \quad (4.7)$$

on $P(x)$ és com abans la població d'edat x quan la verola és present.

En efecte, eliminant com abans $m(x)$ entre les equacions (4.6) i (4.3), Bernoulli va obtenir després de reordenar

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

Va definir $h(x) = P(x)/P^*(x)$. Utilitzant la fórmula (4.4), va multiplicar el numerador i el denominador per e^{-qx} i va obtenir

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

que és equivalent a

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

on \log significa aquí el logaritme natural i no el logaritme decimal. Però $h(0) = 1$. Així que $h(x) = 1 - p + pe^{-qx}$.

Observeu que la relació $P(x)/P^*(x)$ tendeix a $1 - p$ quan l'edat x és prou gran. La sisena columna de la taula 4.1 mostra $P^*(x)$. Una manera de comparar $P(x)$ i $P^*(x)$ és estimar l'esperança de vida en néixer, l'expressió teòrica de la qual quan la verola és present és

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

Una expressió similar amb $P^*(x)$ substituint $P(x)$ és vàlida sense verola. Bernoulli va utilitzar la fórmula aproximada $[\frac{1}{2} \times P(0) + P(1) + P(2) + \dots] / P_0$, que és la donada pel mètode dels trapezis (figura 4.2). Continuant la taula més enllà dels 24 anys fins als 84 anys (veure taula 2.1), va obtenir finalment una esperança de vida E amb verola igual a $[\frac{1}{2} \times 1.300 + 1.000 + \dots + 20] / 1.300 \approx 26,57$ anys, és a dir, 26 anys i 7 mesos. Sense verola, va obtenir una esperança de vida E^* igual a $[\frac{1}{2} \times 1.300 + 1.015 + \dots + 23] / 1.300 \approx 29,65$ anys, és a dir, 29 anys i 8 mesos. La inoculació en néixer augmentaria l'esperança de vida en més de tres anys.

Podem observar que existeix un mètode més senzill i ràpid que l'utilitzat per Bernoulli per a obtenir aquestes fórmules. Partint de l'equació diferencial (4.1) per a $S(x)$, veiem primer que

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

Usant aquesta expressió en l'equació (4.2) per a $R(x)$, trobem que

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

L'equació (4.6) per a $P^*(x)$ mostra que

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

Les fórmules (4.4) i (4.7) es dedueixen de manera immediata!

Per descomptat, la inoculació amb una soca menys virulenta de verola no és completament segura. Si p' és la probabilitat de morir de verola just després de la inoculació ($p' < p$), aleshores l'esperança de vida seria $(1 - p')E^*$ si tothom passés per la inoculació al néixer. Aquesta esperança de vida continua essent més gran que l'esperança de vida «natural» E si $p' < 1 - E/E^*$ (amb les dades anteriors aquest valor límit seria al voltant d'un 11 %). Les dades relatives a p' eren difícils d'obtenir en aquella època. Però Bernoulli va estimar que el risc p' era inferior a l'1 %. Per a ell no hi havia dubte: la inoculació havia de ser promoguda per l'Estat. Va concloure:

«Simplement desitjo que, en un assumpte que concerneix tan de prop al benestar de la raça humana, no es prengui cap decisió sense tot el coneixement que una mica d'anàlisi i càlcul poden proporcionar.»

El treball de Bernoulli va ser presentat a l'Acadèmia de Ciències de París a l'abril del 1760. Al novembre, d'Alembert va presentar un comentari titulat *Sobre l'aplicació de la teoria de la probabilitat a la inoculació de la verola*. El comentari es va publicar poc després al segon volum dels seus *Opuscles mathématiques* amb càlculs més detallats i juntament amb un altre treball titulat *Teoria matemàtica de la inoculació*. D'Alembert va criticar les hipòtesis que va fer Bernoulli de que la probabilitat d'infecció i la probabilitat de morir de verola fossin independents de l'edat. Va suggerir una solució diferent que

no requereix aquestes suposicions. Anomenem $v(x)$ a la mortalitat deguda a la verola a l'edat x , $m(x)$ a la mortalitat deguda a altres causes i $P(x)$ al nombre de persones que continuen vives. Llavors

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

Comparant amb l'equació (4.3), veiem que en realitat $v(x) = pqS(x)/P(x)$. Aquí obtenim

$$P^*(x) = P(x) \exp\left(\int_0^x v(y) dy\right), \quad (4.10)$$

on $P^*(x)$ representa el nombre de persones vives a l'edat x quan la verola ha desaparegut.



Figura 4.3:
D'Alembert (1717–1783)

En efecte, podem o bé eliminar la funció $m(x)$ entre les equacions (4.6) i (4.9) o bé utilitzar la fórmula (4.8) per a $P^*(x)$ i observar que la solució de (4.9) ve donada per

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

La fórmula (4.10) donada per d'Alembert no contradiu la fórmula de Bernoulli (4.7). Simplement utilitza un tipus diferent d'informació $v(x)$, que no estava disponible en aquell moment perquè els registres de defunció inclouen la causa de la mort però no l'edat de la víctima. D'Alembert va suggerir que no es podia realment concloure si la inoculació era útil abans de disposar d'aquest tipus de dades.

D'Alembert també va criticar la utilitat de l'esperança de vida com a criteri de decisió, ja que dona el mateix pes a tots els anys, ja sigui en un futur pròxim o llunyà. Va observar que, des del punt de vista de l'individu o de l'Estat, no tots els anys tenen la mateixa «utilitat», sent menys valuoses les edats joves o velles que les mitjanes. Malgrat totes aquestes crítiques, d'Alembert es va declarar a favor de la inoculació.

A causa de retards en el procés de publicació, l'obra de Bernoulli no es va publicar fins al 1766, mentre que d'Alembert va aconseguir que la seva pròpia obra es publicqués molt ràpidament. Bernoulli va expressar la seva amargor en una carta a Euler:

«Què dir de les enormes banalitats del gran d'Alembert sobre les probabilitats: com que em trobo amb massa freqüència injustament tractat en les seves publicacions, he decidit ja fa temps no llegir res més que surti de la seva ploma. He pres aquesta decisió amb motiu d'un manuscrit sobre la inoculació que vaig enviar a l'Acadèmia de París fa vuit anys i que va ser molt benvolgut per la novetat de l'anàlisi. Va ser, m'atreveixo a dir, com incorporar una nova província al cos de les matemàtiques. Sembla que l'èxit d'aquesta nova anàlisi li va causar mal de cor. Ho ha criticat de mil maneres, totes igualment ridícules, i després de tenir-ho ben criticat, pretén ser el primer autor d'una teoria que ni tan sols va sentir esmentar. Ell, no obstant això, sabia que el meu manuscrit només podia aparèixer després d'uns set o vuit anys. Només podia tenir coneixement d'ell en la seva qualitat de membre de l'Acadèmia. En aquest sentit, el meu manuscrit s'hauria d'haver mantingut sagrat fins que es fes públic. *Dolus an virtus quis in hoste requirat?*»³

Malgrat els treballs de Bernoulli i d'Alembert, la inoculació no es va realitzar a gran escala a França. El rei Lluís XV va morir de verola l'any 1774. Els metges de la cort sí que van inocular a la resta de la família reial poc després. El problema va perdre importància quan Edward Jenner va descobrir que la inoculació de la verola de les vaques als humans («vacunació») protegia contra la verola i era segura. La seva obra, *Una recerca sobre les causes i els efectes de la variolae vaccina*, es va publicar el 1798. La vacunació es va estendre ràpidament per tota Europa. No obstant això, els mètodes desenvolupats per al càlcul de l'augment de l'esperança de vida si s'elimina una causa de mort es continuen utilitzant avui dia.

³«Què importa si per valor o per estratagema vencem a l'enemic?» Virgili: *Eneida*, Llibre II.

A les dècades següents, es va disposar de dades relatives a l'edat en què la gent moria de verola. El problema va ser reconsiderat especialment per

- Johann Heinrich Lambert, matemàtic de l'Acadèmia de Berlín, el 1772;
- Emmanuel-Étienne Duillard, llavors encarregat de les estadístiques de la població al Ministeri de l'Interior a París, en el seu *Anàlisi i taules de la influència de la verola en la mortalitat de cada edat* (1806);
- Pierre-Simon Laplace en el seu *Teoria analítica de la probabilitat* (1812).

Duillard i Laplace van mostrar, per exemple, com modificar la fórmula (4.7) quan els paràmetres p i q depenen de l'edat:

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

Aquí, $p(x)$ és la probabilitat de morir de verola si s'infecta a l'edat x i $q(x)$ és la probabilitat d'infectar-se de verola a l'edat x .

Després d'aquest treball sobre la verola, Daniel Bernoulli no va considerar cap altre problema de dinàmica de poblacions. Va morir a Basilea el 1782. D'Alembert va morir a París un any després.

Més informació

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, págs. 26–95. David, Paris (1761). gallica.bnf.fr
2. Bernoulli, D.: Réflexions sur les avantages de l'inoculation. *Mercur 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. Duillard, 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, págs. 36–46. Scribner, New York (1970)
9. Tent, M.B.W.: *Leonhard Euler and the Bernoullis*. A K Peters, Natick (2009)
10. Voltaire: *Lettres philosophiques*. Lucas, Amsterdam (1734). gallica.bnf.fr

Capítol 5

Malthus i els obstacles al creixement geomètric (1798)

L'any 1798 Malthus va publicar *Un assaig sobre el principi de la població*, en el que afirmava que el subministrament d'aliments no podia sostenir durant un llarg període de temps la tendència natural de les poblacions humanes a créixer exponencialment. Si la població es mantenia relativament constant, era perquè gran part de la humanitat sofria escassetat d'aliments. Malthus veia el «principi de la població» com un argument contra els escrits de Godwin i Condorcet, que posaven l'accent en el progrés de les societats humanes. L'assaig de Malthus va influir en la teoria de l'evolució de Darwin i Wallace i va ser criticat per Marx, però es va posar en pràctica amb la política xinesa del fill únic.

Thomas Robert Malthus va néixer el 1766 prop de Londres, sent el sisè de set fills. El seu pare, amic i admirador de Jean-Jacques Rousseau, va ser el seu primer mestre. L'any 1784, el jove Malthus va començar a estudiar matemàtiques a la Universitat de Cambridge. Va obtenir el seu diploma el 1791, va esdevenir membre del *Jesus College* el 1793 i sacerdot anglicà el 1797.



Figura 5.1:
Malthus (1766–1834)

L'any 1798, Malthus va publicar de manera anònima un llibre titulat *Un assaig sobre el principi de població, com afecta la millora futura de la societat, amb observacions sobre les especulacions del Sr. Godwin, el Sr. Condorcet*.

cet i altres escriptors. Va ser una reacció contra el llibre de Godwin *Recerca sobre la justícia política* (1793) i l'*Esbós d'un quadre històric del progrés de la ment humana* de Condorcet (1794). Malgrat els horrors que la Revolució Francesa va fer en nom del progrés, els dos autors afirmaven que el progrés de la societat era inevitable. Malthus no compartia el mateix optimisme. Sostenia que les lleis angleses de pobresa, que ajudaven les famílies pobres amb molts fills, afavorien el creixement de la població sense fomentar un creixement similar en la producció d'aliments. Li semblava que aquestes lleis no alleujaven realment als pobres, sinó tot el contrari. En general, en tendir la població a créixer sempre més ràpid que la producció d'aliments, una part de la societat semblava estar condemnada a la misèria, a la fam o a les epidèmies: aquestes són les xacres que frenen el creixement demogràfic i que, en opinió de Malthus, són els principals obstacles per al progrés de la societat. Totes les teories que prometen el progrés serien simplement utòpiques. Aquestes idees van portar Malthus a publicar el seu llibre el 1798. Així és com va resumir la seva tesi:

[...] «el poder de la població és indefinidament major que el poder de la terra per a produir la subsistència de l'home. La població, quan no es controla, augmenta en una proporció geomètrica. La subsistència només augmenta en una proporció aritmètica. Un lleuger coneixement dels nombres mostrarà la immensitat de la primera potència en comparació amb la segona. Per aquesta llei de la nostra naturalesa que fa que l'aliment sigui necessari per a la vida de l'home, els efectes d'aquests dos poders desiguals han de mantenir-se iguals. Això implica un fort i constant control de la població per la dificultat de subsistència. Aquesta dificultat ha de recaure en algun lloc; i necessàriament ha de ser sentida severament per una gran part de la humanitat.»

El llibre de Malthus va tenir molt d'èxit. Contenia, no obstant, poques dades. Malthus va observar, per exemple, que la població dels Estats Units s'havia duplicat cada vint-i-cinc anys durant el segle XVIII. En realitat, no va intentar traduir les seves tesis en models matemàtics, però va preparar el camí per als treballs posteriors d'Adolphe Quetelet i Pierre-François Verhulst, que seran objecte del pròxim capítol.

Després de la publicació del seu llibre, Malthus va viatjar amb uns amics, primer a Alemanya, Escandinàvia i Rússia, i després a França i Suïssa. Recopilant la informació recollida durant els seus viatges, va publicar el 1803, aquest cop signant amb el seu nom, una segona edició molt ampliada, amb un subtítol diferent: *Un assaig sobre el principi de població, o una visió dels*

seus efectes passats i presents sobre la felicitat humana, amb una recerca sobre les nostres perspectives respecte a la futura eliminació o mitigació dels mals que ocasiona. En aquesta nova edició s'analitzen detalladament els obstacles al creixement demogràfic en diversos països: el retard del matrimoni, l'avortament, l'infanticidi, la fam, la guerra, les epidèmies, els factors econòmics... Per a Malthus, el retard del matrimoni era la millor opció per a estabilitzar la població. Es van imprimir quatre edicions més del llibre els anys 1806, 1807, 1817 i 1826. El 1805 Malthus va esdevenir professor d'història i economia política en una nova escola creada per la Companyia de les Índies Occidentals per als seus treballadors. També va publicar *Una recerca sobre la naturalesa i el progrés de la renda* (1815) i *Principis d'economia política* (1820). El 1819 Malthus va ser escollit membre de la *Royal Society*. El 1834 va ser un dels membres fundadors de la Societat d'Estadística i va morir prop de Bath aquell mateix any.

L'obra de Malthus va tenir una forta influència en el desenvolupament de la teoria de l'evolució. Charles Darwin, de tornada del seu viatge a bord del *Beagle*, va llegir el llibre de Malthus sobre la població el 1838. Això és el que va escriure en la introducció del seu famós llibre *L'origen de les espècies*, publicat el 1859:

«En el capítol següent s'examinarà la lluita per l'existència entre tots els éssers vius arreu del món, la qual cosa és una conseqüència necessària i inevitable de la forta raó geomètrica que regeix el seu creixement. Es tracta de la doctrina de Malthus aplicada al conjunt dels regnes animal i vegetal.»

Alfred Russel Wallace, que va desenvolupar la teoria de l'evolució al mateix temps que Darwin, també va dir que les seves idees van sorgir després de llegir el llibre de Malthus.

En contrast, aquí està el punt de vista de Karl Marx sobre l'èxit del llibre de Malthus, com es pot llegir en una nota a peu de pàgina del seu *Capital*:

«Si el lector em recorda Malthus, l'*Assaig sobre població* del qual aparegué el 1798, li recordaria que aquesta obra en la seva primera forma no és més que un plagi supèrfluament escolar i lamentable de Defoe, sir James Steuart, Townsend, Franklin, Wallace, etc. i no conté ni una sola frase pensada per ell mateix. La gran atenció que aixecà aquest pamflet es correspon exclusivament a interessos de partit. La revolució francesa trobà en el Regne Unit defensors apassionats: el "principi de la població", lentament elaborat al segle XVIII, en mig d'una gran crisi

social, i proclamat amb tambors i trompetes com a antídoto infal·libre contra les doctrines de Condorcet, etc., fou saludat amb joia per l'oligarquia anglesa com el gran destructor de tots els vicis subsegüents al desenvolupament humà. Malthus, altament sorprès pel seu èxit, es dedicà a farcir el seu llibre de material superficialment compilat, i n'afegí de nou, però que ell no havia descobert, sinó que tan sols el va annexonar.»

Certament, les tesis de Malthus no eren completament noves. Per exemple, sovint se li atribueix la idea que la població tendeix a créixer geomètricament, encara que ja vam veure al capítol 3 que aquesta idea ja era coneguda per Euler mig segle abans¹. No obstant això, Malthus li va donar publicitat vinculant-ho de manera polèmica a problemes legislatius reals. Irònicament, va ser a la Xina comunista on el suggeriment de Malthus de limitar els naixements trobaria la seva aplicació més destacada (vegeu el capítol 25).

Més informació

1. Condorcet: *Esquisse d'un tableau historique des progrès de l'esprit humain*. Agasse, Paris (1794). gallica.bnf.fr
2. Darwin, C. (trad. Zulueta, A. de): *El origen de las especies por medio de la selección natural*. Calpe, Madrid (1921). cervantesvirtual.com
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.: *El Capital*, Primera part. marxists.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)

¹R. A. Fisher (vegeu els capítols 14 i 20) anomenaria «paràmetre malthusià» a la taxa de creixement de les poblacions. Malthus va citar el tractat de Süßmilch en el seu propi llibre.

Capítol 6

Verhulst i l'equació logística (1838)

L'any 1838, el matemàtic belga Verhulst va introduir l'equació logística, que és una espècie de generalització de l'equació del creixement exponencial però amb un valor màxim per a la població. Va utilitzar dades de diversos països, en particular de Bèlgica, per a estimar els paràmetres desconeguts. El treball de Verhulst no es va redescobrir fins als anys vint del segle XX.

Pierre-François Verhulst va néixer el 1804 a Brussel·les. Es va doctorar en matemàtiques a la Universitat de Gant el 1825. També es va interessar per la política. Durant la seva estada a Itàlia per a contenir la seva tuberculosi, va advocar sense èxit a favor d'una constitució per als Estats Pontificis. Després de la revolució del 1830 i la independència de Bèlgica, va publicar un assaig històric sobre un patriota del segle XVIII. El 1835 va ser nomenat catedràtic de matemàtiques a l'aleshores recentment creada Universitat Lliure de Brussel·les.



Figura 6.1:
Verhulst (1804–1849)

Aquell mateix any 1835, el seu compatriota Adolphe Quetelet, estadístic i director de l'observatori de Brussel·les, va publicar *Un tractat sobre l'home i el desenvolupament de les seves facultats*. Quetelet suggeria que les poblacions no podien créixer geomètricament durant un llarg període de temps perquè els obstacles esmentats per Malthus formaven una espècie de «resistència», que ell pensava (per analogia amb la mecànica) que era proporcional

al quadrat de la velocitat de creixement de la població. Aquesta analogia no tenia cap base real, però va inspirar a Verhulst.

En efecte, Verhulst va publicar el 1838 una *Nota sobre la llei del creixement demogràfic*. Aquí hi ha alguns extractes:

«Sabem que el cèlebre Malthus va demostrar el principi que afirma que la població humana tendeix a créixer en una progressió geomètrica fins a duplicar-se al cap d'un cert temps, per exemple cada vint-i-cinc anys. Aquesta proposició és indiscutible si es fa abstracció de la creixent dificultat per a trobar aliments. [...]

»L'augment virtual de la població està, doncs, limitat per la mida i la fertilitat del país. En conseqüència, la població s'acosta cada vegada més a un estat estacionari.»

Verhulst probablement es va adonar que l'analogia mecànica de Quetelet no era raonable i va proposar en lloc d'això la següent equació diferencial (encara una mica arbitrària) per a la població $P(t)$ a temps t :

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

Quan la població $P(t)$ és petita comparada amb el paràmetre K , obtenim l'equació aproximada $dP/dt \approx rP$, la solució de la qual és $P(t) \approx P(0)e^{rt}$, és a dir, un creixement exponencial¹. La taxa de creixement disminueix a mesura que $P(t)$ s'acosta a K . Fins i tot es tornaria negativa si $P(t)$ pogués superar K . Per a obtenir l'expressió exacta de la solució de l'equació (6.1), podem procedir com Daniel Bernoulli per a l'equació (4.5).

Dividint l'equació (6.1) entre P^2 i usant la nova variable $p = 1/P$, obtenim $dp/dt = -rp + r/K$. Prenent $q = p - 1/K$, obtenim $dq/dt = -rq$ i $q(t) = q(0)e^{-rt} = (1/P(0) - 1/K)e^{-rt}$. Així que podem deduir $p(t)$ i $P(t)$.

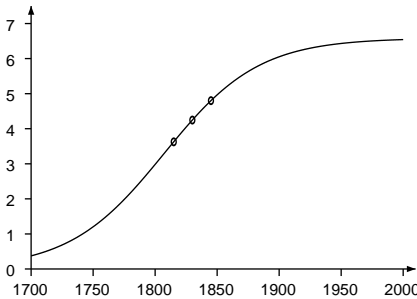
Finalment obtenim després de reordenar

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

La població total augmenta progressivament des de $P(0)$ a temps $t = 0$ fins al límit K , que s'aconsegueix només quan $t \rightarrow +\infty$ (figura 6.2).

¹Se sol parlar de creixement geomètric en els models de temps discret i de creixement exponencial en els models de temps continu, però és essencialment el mateix.

Figura 6.2: La població de Bèlgica (en milions) i la corba logística. Els punts corresponen als anys 1815, 1830 i 1845. Els valors dels paràmetres són els de l'article del 1845.



Sense donar els valors que va utilitzar per als paràmetres desconeguts r i K , Verhulst va comparar el seu resultat amb dades relatives a la població de França entre el 1817 i el 1831, de Bèlgica entre el 1815 i el 1833, del comtat d'Essex a Anglaterra entre el 1811 i el 1831, i de Rússia entre el 1796 i el 1827. L'ajust va resultar ser bastant bo.

El 1840, Verhulst va esdevenir professor de la Real Escola Militar de Brussel·les. L'any següent va publicar el *Tractat elemental de les funcions el·líptiques* i va ser escollit membre de la Reial Acadèmia de Bèlgica. El 1845 va continuar els seus estudis sobre la població amb un article titulat *Recerques matemàtiques sobre la llei del creixement demogràfic*. En primer lloc, es va referir a l'observació de Malthus, segons la qual la població dels Estats Units es duplicava cada 25 anys (taula 6.1).

Taula 6.1: Censos oficials de la població dels Estats Units.

any	població	any	població
1790	3.929.827	1820	9.638.131
1800	5.305.925	1830	12.866.020
1810	7.239.814	1840	17.062.566

Si calculem la proporció entre la població de l'any $n + 10$ i la de l'any n , trobem respectivament 1,350, 1,364, 1,331, 1,335 i 1,326, que és bastant constant. Per tant, la població s'ha multiplicat de mitjana per 1,34 cada 10 anys i per $1,34^{25/10} \approx 2,08$ cada 25 anys. Per tant, havia continuat duplicant-se cada 25 anys des de l'assaig de Malthus, gairebé mig segle abans. No obstant això, Verhulst va afegir:

«No insistirem en la hipòtesi de la progressió geomètrica, atès

que només pot sostenir-se en circumstàncies molt especials; per exemple, quan un territori fèrtil d'extensió gairebé il·limitada resulta estar habitat per persones amb una civilització avançada, com va ser el cas de les primeres colònies americanes.»

Al seu article, Verhulst també va tornar a l'equació (6.1), que va anomenar «logística». Va observar que la corba $P(t)$ augmenta amb una curvatura positiva (és convexa) mentre $P(t) < K/2$ i després continua augmentant cap a K però amb una curvatura negativa (és còncava) quan $P(t) > K/2$. Així, la corba té la forma d'una lletra S distorsionada (figura 6.2).

En efecte, $\frac{d^2P}{dt^2} = r(1 - 2P/K) \frac{dP}{dt}$. Així que $\frac{d^2P}{dt^2} > 0$ si $P < K/2$ y $\frac{d^2P}{dt^2} < 0$ si $P > K/2$.

Verhulst també va explicar com es poden estimar els paràmetres r i K a partir de la població $P(t)$ en tres anys diferents però igualment espaiats. Si P_0 és la població a l'instant $t = 0$, P_1 la de l'instant $t = T$ i P_2 la de l'instant $t = 2T$, llavors un càlcul tediós a partir de l'equació (6.2) mostra que

$$K = P_1 \frac{P_0 P_1 + P_1 P_2 - 2 P_0 P_2}{P_1^2 - P_0 P_2}, \quad r = \frac{1}{T} \log \left[\frac{1/P_0 - 1/K}{1/P_1 - 1/K} \right].$$

Utilitzant les estimacions de la població de Bèlgica en els anys 1815, 1830 i 1845 (respectivament 3,627, 4,247 i 4,801 milions), va obtenir $K = 6,584$ milions i $r = 2,6\%$ per any. Llavors va poder utilitzar l'equació (6.2) per a predir que la població de Bèlgica seria de 4,998 milions a principis de l'any 1851 i de 6,064 milions a principis de l'any 1900 (figura 6.2). Verhulst va realitzar un estudi similar per a França. Va obtenir $K = 39,685$ milions i $r = 3,2\%$ per any. Com que les poblacions de Bèlgica i França han superat àmpliament aquests valors de K , veiem que l'equació logística pot ser un model realista només per a períodes de temps d'unes poques dècades, com a l'article de Verhulst del 1838, però no per a períodes més llargs.

El 1847 va aparèixer una *Segona recerca sobre la llei del creixement demogràfic* en la qual Verhulst renunciava a l'equació logística i triava en el seu lloc una equació diferencial que pot escriure's de la forma $dP/dt = r(1 - P/K)$. Va pensar que aquesta equació es compliria quan la població $P(t)$ estigués per sobre d'un determinat llinar. La solució és $P(t) = K + (P(0) - K)e^{-rt/K}$. Utilitzant les mateixes dades demogràfiques de Bèlgica, Verhulst va estimar de nou els paràmetres r i K . Aquesta vegada va trobar $K = 9,4$ milions per a la població màxima. Veiem com pot arribar a dependre el resultat de l'elecció del model!

Verhulst va esdevenir president de la Reial Acadèmia de Bèlgica el 1848, però va morir l'any següent a Brussel·les, probablement de tuberculosi. Malgrat els dubtes de Verhulst entre els diferents models d'equacions, l'equació logística va ser reintroduïda de manera independent diverses dècades després per diferents persones. Robertson la va utilitzar el 1908 per a modelar el creixement individual d'animals, plantes, éssers humans i òrgans del cos. McKendrick i Kesava Pai la van utilitzar el 1911 per al creixement de poblacions de microorganismes. Pearl i Reed la van utilitzar el 1920 per al creixement de la població dels EUA, que havia començat a disminuir. L'any 1922 Pearl es va assabentar finalment del treball de Verhulst. A partir de llavors, l'equació logística va inspirar molts treballs (vegeu els capítols 13, 20 i 24). La població màxima K va acabar coneixent-se com la «capacitat de càrrega».

Més informació

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

Capítol 7

Bienaymé, Cournot i l'extinció dels cognoms (1845–1847)

L'estadístic francès Bienaymé el 1845 va trobar com calcular la probabilitat que un cognom s'extingeixi si cada home té un nombre de fills que segueix una determinada distribució de probabilitat. Si la mitjana de fills és inferior o igual a u , el cognom s'extingirà. Si la mitjana és major que u , la probabilitat d'extinció és estrictament menor que u . La prova del seu resultat es va publicar dos anys després en un llibre escrit pel seu amic Cournot. Aquests treballs han estat redescoberts recentment.

Irenée Jules Bienaymé va néixer l'any 1796 a París. Va estudiar a l'Escola Politècnica i va desenvolupar la seva carrera al Ministeri de Finances, aconseguint l'alt nivell d'inspector general. Inluït pel llibre *Teoria analítica de la probabilitat* escrit per Laplace, Bienaymé també va trobar temps per a publicar articles sobre moltes aplicacions de la teoria de la probabilitat, com les estadístiques demogràfiques i mèdiques (mortalitat infantil, nombre de naixements, esperança de vida), la probabilitat d'errors en la justícia, la teoria de les assegurances i la representativitat dels sistemes de votació.



Figura 7.1: Bienaymé (1796–1878) i Cournot (1801–1877)

El 1845 Bienaymé va escriure una breu nota *Sobre la llei de la multipli-*

cació i la durada de les famílies, que es va publicar al butlletí de la *Société Philomatique de Paris*. Diversos autors ja havien escrit sobre aquest tema. A la segona edició de *Un assaig sobre el principi de la població* (1803), Malthus va incloure un capítol sobre la població de Suïssa i va observar que

«a la ciutat de Berna, des de l'any 1583 fins al 1654, el consell sobirà havia admès a la burgesia 487 famílies, de les quals 379 es van extingir en l'espai de dos segles, i el 1783 només en quedaven 108.»

L'any 1842 Thomas Doubleday afirmava de forma més general que les famílies de classe alta de la noblesa o de la burgesia tenien una major tendència a desaparèixer que les de classe baixa. Idees similars van ser exposades a França per Émile Littré el 1844 en un text d'introducció a la filosofia positivista d'Auguste Comte i per Benoiston de Châteauneuf –amic de Bienaymé– que va publicar el 1845 l'assaig *Sobre la durada de les famílies nobles a França*.

En aquest context, Bienaymé va tractar d'explicar com podia ser que la població d'un país tendís a créixer geomètricament mentre desapareixia un gran nombre de famílies. Per a atacar aquest problema va considerar el cas simplificat que tots els homes tinguessin les mateixes probabilitats de tenir 0, 1, 2, 3,... fills que arribessin a l'edat adulta. Més concretament, es va preguntar quina era la probabilitat que un home tingués una descendència que portés el seu cognom després de n generacions. Si el nombre mitjà de fills és inferior a u , és evident que aquesta probabilitat ha de tendir a zero a mesura que n creix fins a l'infinit. Bienaymé va observar que la mateixa conclusió continuaria sent certa si el nombre mitjà de fills fos exactament u ¹, per exemple, si hi ha una probabilitat $1/2$ de no tenir cap fill i una probabilitat $1/2$ de tenir dos fills. Però en aquest cas la probabilitat de tenir descendència a la generació n tendeix a zero més lentament: en l'exemple continuaria sent del 5% després de 35 generacions, és a dir, després d'onze o dotze segles si hi ha tres generacions per segle². Bienaymé va observar finalment que si el nombre mitjà de fills és major que u , l'extinció de la línia familiar no és segura: la seva probabilitat pot calcular-se resolent una equació algebraica.

L'article de Bienaymé no contenia més explicacions. El 1847, el seu amic Antoine-Augustin Cournot, matemàtic i economista, va incloure alguns detalls en un llibre titulat *Sobre l'origen i els límits de la correspondència entre l'àlgebra i la geometria*. Va presentar el problema en forma de joc d'atzar, però va reconèixer que era idèntic a l'estudi de Bienaymé sobre l'extinció

¹Excepte si cada home té exactament un fill.

²Com es veurà més endavant, aquesta probabilitat és igual a $1 - x_{35}$ amb $x_{n+1} = \frac{1}{2} + \frac{1}{2}x_n^2$ i $x_0 = 0$.

dels cognoms. Si mantenim la interpretació en termes de cognoms, Cournot va considerar primer el cas especial en el qual els homes tenen com a màxim dos fills, sent p_0 , p_1 i p_2 respectivament la probabilitat de tenir 0, 1 o 2 fills. Per descomptat, $p_0 + p_1 + p_2 = 1$. Partint d'un ancestre, la probabilitat d'extinció després d'una sola generació, anomenada x_1 , és òbviament igual a p_0 . La probabilitat d'extinció en dues generacions és $x_2 = p_0 + p_1 x_1 + p_2 x_1^2$: o bé la família ja s'havia extingit en la primera generació (probabilitat p_0), o bé només hi havia un fill en la primera generació que no tenia descendència masculina (probabilitat $p_1 x_1$), o bé hi havia dos fills en la primera generació i cadascun d'ells no tenia descendència masculina (probabilitat $p_2 x_1^2$). De forma més general, la probabilitat d'extinció en n generacions és

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

En efecte, si hi ha, per exemple, dos fills en la primera generació (probabilitat p_2), la família s'extingirà $n - 1$ generacions després (és a dir, en la generació n) amb una probabilitat igual a $(x_{n-1})^2$. Cournot va observar que x_n és una successió creixent amb $x_n \leq 1$ per a tot n . Així que x_n té un límit $x_\infty \leq 1$, que és una solució de l'equació $x = p_0 + p_1 x + p_2 x^2$. Utilitzant $p_1 = 1 - p_0 - p_2$, aquesta equació és equivalent a $0 = p_2(x - 1)(x - p_0/p_2)$. Per tant, hi ha dues arrels: $x = 1$ i $x = p_0/p_2$. Es distingeixen tres casos en funció del nombre mitjà de fills $p_1 + 2p_2$, que també és igual a $1 - p_0 + p_2$ i que anomenarem \mathcal{R}_0 . Si $\mathcal{R}_0 < 1$, llavors $p_0/p_2 > 1$. Per tant, $x = 1$ és l'únic valor possible per al límit x_∞ . És segur que el cognom s'extingirà. Si $\mathcal{R}_0 = 1$, totes dues arrels són iguals a 1 i la conclusió és la mateixa. Si $\mathcal{R}_0 > 1$, aleshores Cournot va argumentar que x_∞ hauria de ser igual a la segona arrel p_0/p_2 , ja que la probabilitat d'extinció òbviament ha de ser 0 en el cas especial en què $p_0 = 0$.

Cournot va esmentar breument el cas més general en el qual els homes poden tenir com a màxim m fills amb probabilitats p_0, p_1, \dots, p_m . La conclusió depèn igualment del valor de

$$\mathcal{R}_0 = p_1 + 2p_2 + \dots + mp_m,$$

el nombre mitjà de fills que té un individu. L'equació de x_∞ , que és

$$x = p_0 + p_1 x + \dots + p_m x^m,$$

té sempre l'arrel $x = 1$. Només té una altra arrel positiva, que dona la probabilitat d'extinció x_∞ , quan $\mathcal{R}_0 > 1$.

Per desgràcia, l'article de Bienaymé i les poques pàgines del llibre de Cournot van passar completament desapercebuts en el seu moment. L'article no es va conèixer fins a la dècada del 1970 i les pàgines del llibre fins a vint

anys més tard. Mentrestant, el problema i la seva solució havien estat redescoberts per d'altres i el tema es va desenvolupar considerablement. Tornarem a això en els capítols 9, 17 i 18.

Bienaymé va haver de deixar la seva feina al Ministeri de Finances després de la revolució del 1848. La càtedra de teoria de les probabilitats de la Universitat de París, per a la qual sens dubte era el millor candidat, també va ser concedida a una altra persona. No obstant això, Bienaymé va poder tornar a treballar al Ministeri de Finances després del 1850, però va dimitir el 1852. Aquell mateix any va ser escollit membre de l'Acadèmia de Ciències, on era l'especialista en el camp de l'estadística. L'any 1853 va demostrar el que alguns llibres de text moderns denominen la desigualtat Bienaymé–Tchebychev. El 1875 va esdevenir president de l'aleshores recentment creada Societat Matemàtica de França. Va morir a París el 1878.

Més informació

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. Brun, J., Robinet, A. (éd.): A. Cournot, *études pour le centenaire de sa mort*. Economica / Vrin, Paris (1978)
5. Cournot, A.-A.: *De l'origine et des limites de la correspondance entre l'algèbre et la géométrie*. Hachette, Paris (1847). archive.org
6. Doubleday, T.: *The True Law of Population*. Simpkin, Marshall, & Co., London (1842). archive.org
7. Heyde, C.C., Seneta, E.: *I.J. Bienaymé: Statistical Theory Anticipated*. Springer (1977)
8. Kendall, D.G.: The genealogy of genealogy: branching processes before (and after) 1873. *Bull. Lond. Math. Soc.* 7, 225–253 (1975)
9. Littré, É.: *Conservation, révolution et positivisme*. Ladrance, Paris (1852). gallica.bnf.fr
10. Malthus, T.R.: *An Essay on the Principle of Population*, 2nd edn. Bensley, London (1803). archive.org
11. Martin, T.: Antoine Augustin Cournot. In: Heyde, C.C., Seneta, E. (eds.) *Statisticians of the Centuries*, 152–156. Springer (2001)
12. Seneta, E.: Irenée-Jules Bienaymé. In: *Ibid.*, 132–136.

Capítol 8

Mendel i l'herència (1865)

El 1865 Mendel va publicar els resultats dels seus experiments pioners sobre la hibridació dels pèsols. La seva anàlisi utilitzava aspectes elementals de la teoria de la probabilitat. També va considerar un model dinàmic per a una població de plantes autofecundades. El seu treball, que no va ser redescobert fins l'any 1900, és una fita en la història de la genètica.

Johann Mendel va néixer el 1822 a Moràvia, llavors part de l'imperi austríac i ara part de la República Txeca. El seu pare era pagès. Amb els seus bons resultats a l'institut i la seva mala salut, Mendel va preferir continuar estudiant en lloc de treballar a la granja familiar. Però no podia permetre's anar a la universitat. Així que el 1843 va ingressar a l'abadia de Sant Tomàs de Brünn (actual Brno), on va prendre el nom de Gregor. Va estudiar teologia, però també va assistir a alguns cursos d'agricultura. El 1847 va ser ordenat sacerdot. Va fer classes en un institut durant uns anys, però no va superar l'examen per a convertir-se en professor ordinari. Entre el 1851 i el 1853, gràcies al suport de la seva jerarquia, va poder no obstant continuar els seus estudis a la Universitat de Viena, on va assistir a cursos de física, matemàtiques i ciències naturals. Després va tornar a Brünn i va ensenyar física en una escola tècnica.



Figura 8.1:
Mendel (1822–1884)

Entre el 1856 i el 1863, Mendel va realitzar una sèrie d'experiments amb un gran nombre de plantes al jardí de la seva abadia. L'any 1865, va presentar els seus resultats en dues reunions de la Societat d'Història Natural de Brünn, de la qual era membre. El seu treball, *Experiments sobre la hibridació de plantes*, es va publicar en alemany l'any següent en les actes de la Societat. Mendel va explicar com havia arribat a estudiar les variacions dels pèsols, plantes que es reproduïen naturalment per autofecundació i les llavors de les quals poden adoptar diferents formes fàcilment identificables: rodones o arrugades, grogues o verdes, etc. En creuar una planta procedent d'un llinatge amb llavors rodones i una altra procedent d'un llinatge amb llavors arrugades, va observar que sempre obtenia híbrids que donaven llavors rodones. Va anomenar al caràcter «llavors rodones» dominant i al caràcter «llavors arrugades» recessiu. També va demostrar que el caràcter «llavors grogues» era dominant i que el caràcter «llavors verdes» era recessiu.

Mendel va observar aleshores que l'autofecundació de plantes conreades a partir de llavors híbrides donava en la primera generació noves llavors que tenien el caràcter dominant o el recessiu en proporcions aparentment aleatòries. A més a més, es va adonar que repetint l'experiment moltes vegades obtenia, de mitjana, unes tres vegades més llavors amb el caràcter dominant que amb el recessiu. Per exemple, en un primer experiment va obtenir un total de 5.474 llavors rodones i 1.850 llavors arrugades, la qual cosa correspon a una proporció de 2,96 a 1. Un segon experiment va donar un total de 6.022 llavors grogues i 2.001 verdes, la qual cosa correspon a una proporció¹ de 3,01 a 1.

Mendel també va observar que entre les plantes conreades a partir de les llavors de la primera generació amb el caràcter dominant, les que donaven per autofecundació llavors amb el caràcter dominant o recessiu eren aproximadament el doble de les que donaven llavors amb el caràcter dominant solament. Per exemple, entre les 565 plantes conreades a partir de llavors rodones de la primera generació, 372 van donar tant llavors rodones com arrugades, mentre que 193 van donar només llavors rodones; la proporció és igual a 1,93. De la mateixa manera, entre les 519 plantes conreades a partir de llavors grogues de la primera generació, 353 van donar llavors tant grogues com verdes, mentre que 166 només van donar llavors grogues; la proporció és igual a 2,13.

Per explicar aquests resultats, Mendel va tenir la brillant idea de considerar el caràcter aparent d'una llavor com el resultat de l'associació de dos

¹Com R. A. Fisher (vegeu el capítol 14) va observar més tard, la probabilitat d'arribar a resultats experimentals tan pròxims al valor teòric és bastant petita. És probable que Mendel arreglés les seves dades. Per exemple, en el segon experiment sobre $n = 6.022 + 2.001 = 8.023$ llavors, la probabilitat que la proporció difereixi de 3 en menys de 0,01 és només d'un 10 %.

factors ocults, cadascun dels quals és o bé dominant (escrit A) o bé recessiu (escrit a). Així, hi ha tres combinacions possibles: AA , Aa i aa . Les llavors amb els factors AA o Aa tenen el mateix caràcter dominant A . Les llavors amb els factors aa tenen el caràcter recessiu a . Mendel va suposar a més a més que durant la fecundació, els grans de pol·len i els òvuls (els gàmetes) transmeten només un dels dos factors, cadascun amb una probabilitat $1/2$.

Per tant, l'encreuament dels llinatges purs AA i aa dona lloc a híbrids que tenen tots els factors Aa i el caràcter dominant A . Els gàmetes de l'híbrid Aa transmeten el factor A amb probabilitat $1/2$ i el factor a amb probabilitat $1/2$. L'autofecundació d'una planta conreada a partir d'una llavor híbrida Aa dona, per tant, AA amb probabilitat $1/4$, Aa amb probabilitat $1/2$ i aa amb probabilitat $1/4$, com es mostra a la taula 8.1.

Taula 8.1: Possibles resultats de l'autofecundació d'un híbrid Aa i les seves probabilitats en funció dels factors transmesos pels gàmetes masculins (en files) i pels gàmetes femenins (en columnes).

Factor (Probabilitat)	$A (1/2)$	$a (1/2)$
$A (1/2)$	$AA (1/4)$	$Aa (1/4)$
$a (1/2)$	$Aa (1/4)$	$aa (1/4)$

Mendel es va adonar que les proporcions $AA : Aa : aa$, que eren $1 : 2 : 1$, també podien obtenir-se mitjançant el càlcul formal $(A + a)^2 = AA + 2Aa + aa$. Com les llavors AA i Aa tenen el caràcter aparent A mentre que només les llavors aa tenen el caràcter aparent a , hi ha efectivament tres vegades més llavors amb el caràcter A que amb el caràcter a . A més a més, hi ha una mitjana del doble de llavors Aa que AA . L'autofecundació de les plantes conreades a partir de les primeres dona llavors amb el caràcter dominant (AA o Aa) o amb el caràcter recessiu (aa). Mentre que l'autofecundació de plantes conreades a partir de llavors AA , sempre dona llavors AA amb el caràcter dominant. Totes les observacions queden així explicades.

Mendel també va examinar les generacions següents. Partint de N llavors híbrides Aa i suposant per simplicitat que cada planta dona per autofecundació només quatre llavors noves, va calcular que el nombre mitjà de llavors $(AA)_n$, $(Aa)_n$ i $(aa)_n$ en la generació n vindria donat per la taula 8.2 en la qual, per claredat de la presentació, els resultats s'han dividit per N .

Aquests nombres s'obtenen simplement a partir de les fórmules

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

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

Taula 8.2: Les generacions successives.

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
total	1	4	16	64	256	1.024

i $(aa)_{n+1} = (Aa)_n + 4(aa)_n$, que diuen que AA dona després de l'autofecundació quatre llavors AA , que aa dona quatre llavors aa i que Aa dona en mitjana una llavor AA , dues llavors Aa i una llavor aa . Mendel va observar a més a més que $(AA)_n = (aa)_n = 2^{n-1}(2^n - 1)$ i $(Aa)_n = 2^n$.

En efecte, de l'equació (8.2) i de la condició inicial $(Aa)_0 = 1$ es dedueix que $(Aa)_n = 2^n$. Substituint això en l'equació (8.1), obtenim que $(AA)_{n+1} = 4(AA)_n + 2^n$. Ens adonem fàcilment que $(AA)_n = c2^n$ és una solució particular quan $c = -1/2$. La solució general de l'equació «homogènia» $(AA)_{n+1} = 4(AA)_n$ és $(AA)_n = C4^n$. Finalment, sumant aquestes dues solucions, veiem que $(AA)_n = C4^n - 2^{n-1}$ satisfà la condició inicial $(AA)_0 = 0$ si $C = 1/2$. Quant a la successió $(aa)_n$, satisfà la mateixa relació de recurrència i la mateixa condició inicial que $(AA)_n$. Per tant, $(aa)_n = (AA)_n$.

En conclusió, la proporció d'híbrids Aa en la població total, que és $2^n/4^n = 1/2^n$, es divideix per la meitat en cada generació per autofecundació.

El treball de Mendel va passar totalment desapercbut durant la seva vida. Alguns anys més tard, Mendel també va intentar realitzar experiments similars amb altres espècies vegetals, va publicar alguns articles sobre meteorologia i va investigar l'herència de les abelles. Després de convertir-se en abat el 1868, va dedicar la major part del seu temps a gestionar problemes administratius. Va morir l'any 1884. No va ser fins al 1900 quan el treball de Mendel va ser finalment redescobert de manera independent i gairebé simultània per Hugo De Vries a Amsterdam, Carl Correns a Tubinga i Erich von Tschermak a Viena. Amb això s'inicia una nova era en el que avui en dia coneixem com a genètica.

Més informació

1. Bateson, W.: *Mendel's Principles of Heredity*. Cambridge (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

Capítol 9

Galton, Watson i el problema de l'extinció (1873–1875)

L'any 1873, l'estadístic britànic Galton i el seu compatriota, el matemàtic Watson, es van plantejar el problema de l'extinció dels cognoms sense conèixer els treballs de Bienaymé. Watson va observar que la funció generatriu associada a la distribució de probabilitat del nombre d'homes en cada generació podia calcular-se recursivament. Però va analitzar incorrectament la probabilitat d'extinció.

Francis Galton va néixer el 1822, el mateix any que Mendel, prop de Birmingham, a Anglaterra. Era el menor de set fills. El seu pare era un ric banquer. Per part de la seva mare, era cosí de Charles Darwin. Galton va començar a estudiar medicina el 1838, primer en un hospital de Birmingham i després a Londres. Durant l'estiu del 1840, va realitzar el seu primer viatge llarg per Europa fins a Istanbul. Després va estudiar al *Trinity College* de la Universitat de Cambridge durant quatre anys. Però el seu pare va morir el 1844, deixant una fortuna important. Galton va abandonar la idea de convertir-se en metge. Va viatjar a Egipte, Sudan i Síria. Durant els anys següents va mantenir un estil de vida ociós, passant el seu temps caçant, viatjant en globus i vaixell o intentant millorar el telègraf elèctric. El 1850 va organitzar una expedició d'exploració al sud-oest de l'Àfrica (actual Namíbia). Al seu retorn a Anglaterra, el 1852, va ser escollit membre de la *Royal Geographical Society*. Allà va poder seguir les notícies de les expedicions a l'Àfrica oriental a la recerca del naixement del Nil. Es va instal·lar a Londres i va escriure una guia per a viatgers que es va convertir en un best seller. El 1856 va ser escollit membre de la *Royal Society*. Es va interessar llavors per la meteorologia i va inventar la paraula «anticicló». Després de la publicació el 1859 pel seu cosí Darwin de *L'origen de les espècies*, Galton es va dedicar a l'estudi de l'herència. El 1869 va publicar *El geni hereditari*, en el qual afirmava que les facultats intel·lectuals podien transmetre's per herència.

L'any 1873, Alphonse de Candolle, un botànic suís, va publicar un llibre titulat *Història de la ciència i dels científics en els dos últims segles*, que contenia també un assaig sobre *La influència respectiva de l'herència, la variabilitat i la selecció en el desenvolupament de l'espècie humana i en el futur*

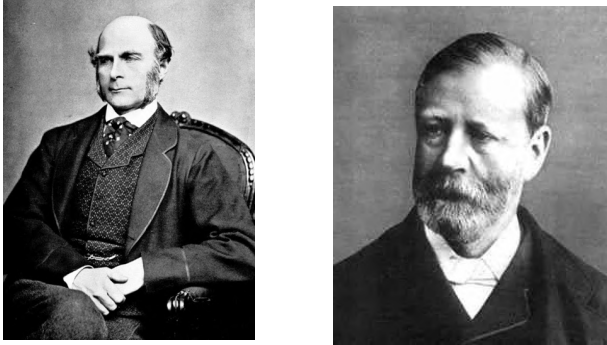


Figura 9.1: Galton (esquerra) i Watson (dreta).

probable d'aquesta espècie. Allà feia les següents observacions:

«Entre les informacions precises i les opinions molt assenyades del Sr. Benoiston de Châteauneuf, de Galton i d'altres estadístics, no he vist la important observació que haurien d'haver fet sobre la inevitable extinció dels cognoms. Per descomptat, tots els cognoms han d'extingir-se [...] Un matemàtic podria calcular com es produiria la disminució dels cognoms o dels títols, coneixent la probabilitat de tenir una filla o un fill i la probabilitat de no tenir cap descendent per part d'una parella.»

És el mateix problema que Bienaymé havia estudiat el 1845. Però Candolle, que no coneixia els treballs de Bienaymé, pensava que totes les famílies estaven destinades a l'extinció. Galton es va fixar en el paràgraf anterior del llibre de Candolle. Com que tampoc coneixia el treball de Bienaymé, Galton el va posar com un problema obert per als lectors de l'*Educational Times*:

«Problema 4.001: Una nació gran, de la qual només ens ocuparem dels homes adults, suposem que són N , i que porten cadascun cognoms diferents colonitzen un districte. La seva llei de població és tal que, en cada generació, a_0 per cent dels homes adults no tenen fills homes que arribin a la vida adulta; a_1 tenen un fill home d'aquest tipus; a_2 en tenen dos; i així successivament fins a_5 que en tenen cinc.

»Trobeu (1) quina proporció dels seus cognoms s'haurà extingit

després de r generacions; i (2) quants casos hi haurà de cognoms amb exactament m persones.»

Observem que la segona part del problema no havia estat abordada per Bienaymé. Galton no va rebre cap resposta satisfactòria dels lectors de la revista i aparentment no va poder trobar la solució del problema per si mateix. Així que va demanar al seu amic Henry William Watson, un matemàtic, que intentés resoldre'l.

Watson va néixer a Londres l'any 1827. El seu pare era oficial de la marina britànica. Va estudiar primer al *King's College* de Londres i després es va dedicar a les matemàtiques al *Trinity College* de la Universitat de Cambridge, des del 1846 fins al 1850, pocs anys després de Galton. Va ser successivament investigador del *Trinity College*, ajudant de professor a la *City of London School*, professor titular de matemàtiques al *King's College* i catedràtic de matemàtiques a la *Harrow School* entre el 1857 i el 1865. Aficionat a l'alpinisme, va formar part d'una expedició que va aconseguir fer el cim del Mont Rosa a Suïssa el 1855. Es va ordenar com a diaca el 1856 i com a sacerdot anglicà dos anys després. Des del 1865 fins a la seva jubilació va ser rector de Berkswell amb Barston, prop de Coventry, càrrec que li deixava prou temps per a la recerca.

Galton i Watson van escriure junts un article titulat *Sobre la probabilitat d'extinció de les famílies*, que es va publicar el 1875 al *Journal of the Royal Anthropological Institute*. Galton va presentar el problema i Watson va explicar els càlculs i les conclusions a les que havia arribat. Suposaven que els homes tenen com a màxim q fills, sent p_k la probabilitat de tenir k fills ($k = 0, 1, 2, \dots, q$). En altres paraules, $p_k = a_k/100$ si utilitzem la notació original de Galton. Per tant, $p_0 + p_1 + \dots + p_q = 1$. Considerem la situació en la qual la generació 0 està formada per un sol home. La generació 1 està formada per s homes amb probabilitat p_s . Utilitzant un truc ben conegut a la seva època i que havia estat introduït molt abans per Abraham de Moivre, Watson va considerar la funció generatriu

$$f(x) = p_0 + p_1x + p_2x^2 + \dots + p_qx^q \quad (9.1)$$

associada a les probabilitats p_0, \dots, p_q . De manera similar, sigui $f_n(x)$ el polinomi per al qual el coeficient de x^s és la probabilitat de tenir s homes en la generació n partint d'un sol home en la generació 0. Llavors $f_1(x) = f(x)$. Watson va observar que

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

una fórmula que permet calcular $f_n(x)$ recursivament.

En efecte, sigui

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

Observeu que hi ha com a màxim q^n homes a la generació n . Si a la generació $n - 1$ hi ha s homes numerats d'1 a s , direm t_1, \dots, t_s al nombre dels seus fills homes. En aquest cas, hi haurà t homes a la generació n amb una probabilitat igual a

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

Quan $s = 0$, s'ha d'entendre que aquesta probabilitat és igual a 1 si $t = 0$ i igual a 0 si $t \geq 1$. Per tant,

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

D'això es dedueix que

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

En particular, la probabilitat x_n d'extinció del cognom en n generacions és igual a $p_{0,n}$, que és el mateix que $f_n(0)$. Com a primer exemple, Watson va prendre

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

és a dir, $q = 3$ i $p_0 = p_1 = p_2 = 1/3$. Va calcular els polinomis $f_n(x)$ per a $n = 1, \dots, 4$ utilitzant l'equació (9.2). Va obtenir, per exemple

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

i $f_2(0) = 13/27 \approx 0,481$. El càlcul de $f_n(x)$ per a $n \geq 3$ resulta molt tediós, tant que Watson ja es va equivocar per a $n = 4$. Com que $x_5 = f_5(0) = f_4(f(0))$ es pot evitar el càlcul de $f_5(x)$. Va obtenir la següent llista de probabilitats d'extinció $x_n = f_n(0)$:

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

Els valors correctes són $x_4 \approx 0,632$ i $x_5 \approx 0,677$, com es pot comprovar utilitzant la fórmula senzilla $x_n = f(x_{n-1})$ obtinguda per Bienaymé. Com veurem al capítol 17, aquesta última fórmula també es pot obtenir a partir de l'equació (9.2).

Watson va observar que cada home té de mitjana

$$\mathcal{R}_0 = p_1 + 2p_2 + \dots + qp_q$$

fills i que $\mathcal{R}_0 = 1$ al seu primer exemple. Per tant, es podria pensar que si el nombre inicial de membres masculins de la família era prou gran, la grandària de la família es mantindria més o menys constant. No obstant això, Watson va afirmar que la probabilitat d'extinció x_n convergeix cap a 1 quan $n \rightarrow +\infty$, encara que de forma molt lenta. En altres paraules, tota família s'extingirà, com havia suggerit Candolle. La figura 9.2a, que no està dibuixada a l'article original, i els resultats de Bienaymé confirmen que aquesta conclusió per al primer exemple és correcta.

Com a segon exemple, Watson va considerar la distribució de probabilitat binomial

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

per a la qual la funció generatriu (9.1) és

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

Va calcular $f_2(x)$ i $x_2 = f_2(0)$. En aquest punt es va adonar que $x_2 = f(x_1)$ i que $x_n = f(x_{n-1})$ per a tot n . Però va pensar que aquesta fórmula era certa només per al cas especial de la binomial (9.3). Aplicant-la al cas en què $q = 5$, $a = 3$ i $b = 1$, va obtenir

$$x_1 \approx 0,237, \quad x_2 \approx 0,347, \quad x_3 \approx 0,410, \quad \dots$$

$$x_9 \approx 0,527, \quad x_{10} \approx 0,533 \dots$$

Watson es va adonar que x_n ha de convergir a un límit x_∞ quan $n \rightarrow +\infty$, que satisfà

$$x_\infty = f(x_\infty) = \frac{(a + bx_\infty)^q}{(a + b)^q}.$$

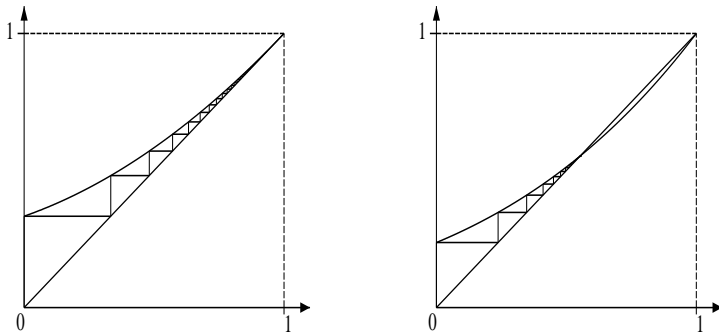


Figura 9.2: Gràfica de les funcions $y = f(x)$ i $y = x$. La probabilitat d'extinció $x_n = f(x_{n-1})$ en n generacions és l'altura del n -èsim «pas de l'escala». (a): $f(x) = (1 + x + x^2)/3$. (b): $f(x) = (3 + x)^5/4^5$.

Va observar que $x = 1$ és una solució d'aquesta equació, però no es va adonar que podia haver-hi altres solucions quan $\mathcal{R}_0 > 1$. Així que va concloure erròniament, induït a error per Candolle, que hi ha extinció ($x_\infty = 1$) en tots els casos, inclòs l'exemple numèric que acabava de considerar. La figura 9.2b mostra que no és així.

Watson es va adonar que el nombre mitjà de fills en aquest exemple numèric era major que 1 (es pot demostrar que $\mathcal{R}_0 = qb/(a+b) = 5/4$), cosa que significa que la població tendeix a augmentar exponencialment. Però això no el va ajudar a descobrir el seu error. Fins i tot va conjecturar que l'extinció del cognom era certa per a tota distribució de probabilitat (p_k), és a dir, no sols per al cas binomial. Tornarem a aquest problema als capítols 17 i 18.

Galton va prosseguir el seu estudi estadístic de les famílies amb un llibre titulat *Homes de ciència anglesos, la seva naturalesa i la seva educació*, que se centrava en la genealogia dels membres de la *Royal Society*. També es va interessar per l'antropometria, la mesura del cos humà. Va aprofitar una exposició internacional celebrada el 1884 a Londres per a recopilar dades sobre un gran nombre de persones. Els seus resultats es van publicar el 1889 en un llibre titulat *Herència natural*, l'apèndix del qual reproduïa l'article escrit en col·laboració amb Watson. Aquest llibre també va introduir nou vocabulari estadístic com «percentil» i «quartil», així com la paraula «eugenèsia», és a dir, la millora de l'espècie humana des del punt de vista dels caràcters heredi-

taris. A partir del 1888 Galton va desenvolupar la tècnica de reconeixement d'empremtes dactilars que seria utilitzada uns anys més tard per la policia britànica. També va continuar estudiant el paper respectiu de l'herència (naturalesa) i de l'entorn (criança) en les característiques físiques i intel·lectuals dels bessons, en la grandària dels pèsols conreats durant diverses generacions o en el color dels ratolins criats en un laboratori. Això el va portar a la noció de «coeficient de correlació» entre dues variables. El 1904 es va fundar el Laboratori Galton a l'*University College* de Londres. Galton va ser nomenat cavaller el 1909 i va morir el 1911.

Watson va publicar diversos llibres, en particular un tractat sobre la teoria cinètica dels gasos el 1876 i un tractat sobre la teoria matemàtica de l'electricitat i el magnetisme en dos volums (1885 i 1889). Va ser escollit membre de la *Royal Society* el 1881 i va morir a Brighton l'any 1903.

El 1924, en el segon volum de la seva biografia de Galton, Karl Pearson va resumir l'article sobre l'extinció dels cognoms sense adonar-se de l'error. Aquest error seria finalment descobert el 1930 (vegeu el capítol 18).

Més informació

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](#)

Capítol 10

Lotka i la teoria de la població estable (1907–1911)

L'any 1907, el químic estatunidenc Alfred Lotka va començar a estudiar la relació entre la taxa de natalitat, la taxa de mortalitat per edats i la taxa de creixement de la població mitjançant un model a temps continu. El 1911 va publicar un altre article sobre el mateix tema amb Sharpe, que també incloïa la taxa de fertilitat per edats. L'equació implícita que dona la taxa de creixement de la població es coneix com a «equació de Lotka».

Alfred James Lotka, fill de pares estatunidencs, va néixer el 1880 a Lemberg, que formava part de l'Imperi Austrohongarès (actualment Lviv a Ucraïna). Va estudiar primer a França i Alemanya i el 1901 es va llicenciar en física i química per la Universitat de Birmingham (Anglaterra). Després va passar un any a Leipzig, on Wilhelm Ostwald, que rebria el premi Nobel de Química el 1909, va destacar el paper de la termodinàmica en la química i la biologia. Lotka es va instal·lar a Nova York el 1902 i va començar a treballar per a la *General Chemical Company*.



Figura 10.1:
Lotka (1880–1949)

Els anys 1907 i 1911¹, Lotka va abordar l'estudi de la dinàmica de po-

¹El segon article va ser escrit en col·laboració amb Sharpe, un matemàtic de la Universitat de Cornell.

blacions estructurades per l'edat sense conèixer els treballs d'Euler sobre el mateix tema (vegeu el capítol 3). A diferència d'Euler, va assumir que el temps i l'edat són variables contínues. Sigui $B(t)$ la taxa de natalitat masculina (el nombre de naixements masculins per unitat de temps) a temps t , $p(x)$ la probabilitat de continuar viu a l'edat x i $h(x)$ la fertilitat a l'edat x : $h(x)dx$ és la probabilitat que un home tingui un fill entre l'edat x i l'edat $x+dx$, si dx és infinitèsimament petit. Aleshores $\int_0^{+\infty} p(x)dx$ és l'esperança de vida en néixer. A més a més, $B(t-x)p(x)dx$ és el nombre d'homes nascuts entre el temps $t-x$ i el temps $t-x+dx$, que continuen vius a temps t . Aquests homes tenen $B(t-x)p(x)h(x)dx$ fills per unitat de temps en l'instant t . Per tant, la taxa total de natalitat masculina a temps t ve donada per $B(t) = \int_0^{+\infty} B(t-x)p(x)h(x)dx$, que és una equació de renovació on implícitament s'ha suposat una història de naixements coneguda $B(\theta) = \phi(\theta)$, $\theta \leq 0$. Buscant una solució exponencial per a aquesta equació integral suposant $B(t)$ de la forma $B(t) = be^{rt}$, Lotka va obtenir dividint tots dos costats per $B(t)$ l'equació

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

que actualment s'anomena «equació de Lotka» en demografia². Euler havia obtingut l'equació implícita anàloga (3.1) per a la taxa de creixement quan el temps i l'edat són variables discretes. Lotka va observar que el costat dret de (10.1) és una funció decreixent de r que tendeix a $+\infty$ quan $r \rightarrow -\infty$ i que tendeix a 0 quan $r \rightarrow +\infty$. Així que hi ha un valor únic de r , que anomenarem r^* , pel qual l'equació (10.1) es compleix. A més a més, $r^* > 0$ si i només si

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

El paràmetre \mathcal{R}_0 és el famós nombre reproductiu bàsic (la notació va ser introduïda per Dublin i Lotka el 1925) i s'interpreta com el nombre esperat de fills que un home pot tenir al llarg de la seva vida.

Lotka va suggerir³ que, sigui quina sigui l'estructura d'edat inicial en la població, el nombre de naixements d'homes per unitat de temps era, en efecte, tal que $B(t) \sim be^{r^*t}$ quan $t \rightarrow +\infty$, on b és una constant. La població total ve donada llavors per $P(t) = \int_0^{+\infty} B(t-x)p(x)dx$. D'això es dedueix que $P(t)$ també augmenta o disminueix com e^{r^*t} quan $t \rightarrow +\infty$: la taxa de

²R. A. Fisher va arribar de manera independent a la mateixa equació el 1927 i més tard va interpretar l'arrel r^* com una mesura de l'«aptitud Darwiniana» a la teoria de l'evolució per selecció natural.

³Això va ser demostrat rigorosament el 1941 per Feller. El 1968, Crump, Mode i Jagers van desenvolupar un enfocament probabilístic.

creixement és igual a r^* . A més a més, l'estructura de la població en edats (perfil d'edats), donada per $B(t-x)p(x)/P(t)$, tendeix a $\frac{e^{-r^*x}p(x)}{\int_0^{+\infty} e^{-r^*y}p(y)dy}$. Això és el que Lotka anomenava una «població estable»: la piràmide d'edat manté la mateixa forma al llarg del temps, però la població total augmenta o disminueix exponencialment en el temps. La conclusió és, doncs, la mateixa que en el model a temps discret d'Euler. No obstant això l'estudi de Lotka té en compte la dependència de la fertilitat respecte de l'edat. Per tant, és en certa manera més general que el d'Euler.

Lotka va continuar treballant en aquest tema durant tota la seva vida. Els anys 1908-1909 va reprendre els seus estudis a la Universitat de Cornell per a obtenir un màster. Va treballar per al *National Bureau of Standards* des del 1909 fins al 1911 i com a editor de la revista *Scientific American Supplement* des del 1911 fins al 1914. El 1912 va obtenir un doctorat a la Universitat de Birmingham en recopilar els articles que havia publicat des del 1907 sobre dinàmica de poblacions i demografia. Durant la Primera Guerra Mundial, va tornar a treballar per a la *General Chemical Company* en com fixar el nitrogen de l'atmosfera. El 1920, un dels seus articles sobre les oscil·lacions biològiques (vegeu el capítol 13) va causar una profunda impressió a Raymond Pearl, un professor de biometria de la Universitat Johns Hopkins que acabava de «redescobrir» l'equació logística (vegeu el capítol 6). Amb l'esperança d'obtenir una feina a l'Institut Rockefeller de Recerca Mèdica de Nova York, Lotka va treballar en els models matemàtics desenvolupats per Ross per a la malària (vegeu el capítol 12). Finalment va aconseguir una beca de dos anys a la Universitat Johns Hopkins, que li va permetre escriure un llibre titulat *Elements de biologia física*, publicat el 1925 i reeditat el 1956 sota el nom d'*Elements de biologia matemàtica*. A continuació, va passar a dirigir el departament de recerca de la *Metropolitan Life Insurance Company* de Nova York. Es va centrar en l'anàlisi matemàtica de qüestions demogràfiques i va publicar diversos llibres en col·laboració amb un col·lega, l'estadístic i vicepresident de la companyia Louis Israel Dublin: *El valor monetari d'un home* (1930), *Durada de la vida* (1936) i *Vint-i-cinc anys de progrés sanitari* (1937). Va ser elegit president de la *Population Association of America* per al període 1938–1939. Entre els seus diversos estudis estadístics, la «lleï de Lotka» (que es remunta al 1926) afirma que el nombre d'autors que han escrit n articles en un camp científic determinat disminueix més o menys com $1/n^2$ a mesura que n augmenta. Lotka també va publicar un llibre en francès titulat *Teoria analítica de les associacions biològiques*. La primera part, més filosòfica, va aparèixer el 1934. La segona part, més tècnica, es va publicar el 1939 i resumia tota la seva recerca sobre demografia humana des del 1907. En el

seu llibre, Lotka també va presentar la seva contribució al problema de l'extinció dels cognoms. Després de la publicació el 1930 del primer article de Steffensen sobre el tema (vegeu el capítol 18), va aplicar la teoria a les dades contingudes en el cens del 1920 de la població blanca dels Estats Units. Va observar que la distribució observada $(p_k)_{k \geq 0}$ del nombre de fills s'aproxima bé per una llei geomètrica decreixent per a tot $k \geq 1$: $p_0 = a$, $p_k = b c^{k-1}$ ($k \geq 1$), amb $a = 0,4825$, $b = 0,2126$ i $c = 1 - b/(1 - a)$. D'aquesta manera, $\sum_{k \geq 0} p_k = 1$. La funció generatriu associada és $f(x) = a + \frac{bx}{1-cx}$. Les dues solucions de l'equació $x = f(x)$ són $x = 1$ i $x = a/c$. La probabilitat d'extinció x_∞ és la més petita d'aquestes dues solucions (vegeu el capítol 7). Amb els valors numèrics dels paràmetres per als EUA va trobar $x_\infty \approx 0,819$, mentre que el nombre mitjà de fills era $\mathcal{R}_0 = f'(1) = (1 - a)^2/b \approx 1,260 > 1$. Malgrat un nombre mitjà de fills (incloent fills i filles) pròxim a 2,5, la probabilitat d'extinció del cognom és superior al 80 %.

Lotka va ser elegit president de l'*American Statistical Association* el 1942. Es va jubilar el 1947 i va morir el 1949 a Nova Jersey. Com ja s'ha esmentat, el 1956 es va publicar una nova edició del seu llibre del 1925 amb un títol lleugerament diferent: *Elements de biologia matemàtica*.

Més informació

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, pág. 512. Scribner, New York (1981)
6. Jagers, P.: Age-dependent branching processes allowing immigration. *Theor. Probab. Appl.* 13, 225–236 (1968).
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*, 2^e 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)
11. Tanner, A.: *Von Molekülen, Parasiten und Menschen, Lotka und die Mathematisierung des Lebens*. Zürich (2014) doi:10.3929/ethz-a-010209129

Capítol 11

La llei de Hardy-Weinberg (1908)

L'any 1908, el matemàtic britànic Hardy i el metge alemany Weinberg van descobrir per separat que en una població infinitament gran que es reproduïx a l'atzar segons les lleis de Mendel, les freqüències dels genotips obtinguts de dos al·lels es mantenen constants al llarg de les generacions. El seu model matemàtic va ser un dels punts de partida de la genètica de poblacions.

Godfrey Harold Hardy va néixer el 1877 a Surrey, Anglaterra. Els seus pares eren professors. Va estudiar matemàtiques al *Trinity College*, Universitat de Cambridge, des de l'any 1896, després va esdevenir investigador d'aquesta universitat el 1900 i finalment professor permanent de matemàtiques el 1906. Després d'un primer llibre titulat *La integració de funcions d'una sola variable* (1905), va publicar el 1908 *Un curs de matemàtiques pures*, que va ser reeditat moltes vegades i traduït a moltes llengües.



Figura 11.1:
Hardy (1877–1947)

En aquella època, el redescobriments dels treballs de Mendel havia plantejat alguns dubtes. Alguns biòlegs es preguntaven per què els caràcters dominants no es feien més freqüents de generació en generació. Reginald Punnett, que havia escrit el 1905 un llibre titulat *Mendelisme*, va plantejar la qüestió a Hardy, amb qui jugava a criquet a Cambridge. Hardy va escriure la seva solució en un article titulat *Proporcions mendelianes en una població mixta*, que es va publicar el 1908. Per a simplificar l'anàlisi, va imaginar la situació

d'una població gran en la qual l'elecció de la parella sexual fos aleatòria. A més a més, va restringir la seva atenció a només dos factors (o «al·lells») A i a , essent A dominant i a recessiu. Per a la generació n , sigui p_n la freqüència del «genotip» (o tipus genètic) AA , $2q_n$ la de Aa i r_n la de aa . Naturalment, $p_n + 2q_n + r_n = 1$. Hardy va suposar també que cap d'aquests genotips provocava un excés de mortalitat o una disminució de la fertilitat en comparació amb els altres dos genotips. Les freqüències en la generació $n + 1$ poden calcular-se fàcilment observant que un individu escollit a l'atzar en la generació n transmet l'al·lel A amb una probabilitat $p_n + q_n$: en efecte, o bé el genotip és AA i l'al·lel A es transmet amb seguretat o bé el genotip és Aa i l'al·lel A es transmet amb una probabilitat del 50%. De la mateixa manera, l'al·lel a es transmet amb una probabilitat $q_n + r_n$. Així es pot construir la taula 11.1 de la mateixa manera que la taula 8.1.

Taula 11.1: Càlcul de les freqüències dels genotips en la generació $n + 1$ a partir de les freqüències dels al·lells dels pares (les files són per a la mare, les columnes per al pare).

Al·lel Freqüència	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 $(q_n + r_n)^2$

Les freqüències dels genotips AA , Aa i aa en la generació $n + 1$ són respectivament p_{n+1} , $2q_{n+1}$ i r_{n+1} . Per tant, Hardy va trobar que

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

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

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

A continuació, va investigar en quines condicions les freqüències dels genotips podien romandre constants al llarg de les generacions sent iguals a p , $2q$ i r . Ja que per definició $p + 2q + r = 1$, veiem que les equacions (11.1)-(11.3) donen totes la mateixa condició $q^2 = pr$.

Per exemple, la primera equació dona $p = (p + q)^2 = p^2 + 2pq + q^2$, la qual cosa equival a $p(1 - p - 2q) = q^2$ i finalment a $pr = q^2$.

Partint de condicions inicials arbitràries $(p_0, 2q_0, r_0)$ amb $p_0 + 2q_0 + r_0 =$

1, Hardy va observar que

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

Per tant, l'estat $(p_1, 2q_1, r_1)$ ja és un equilibri. Així doncs, $(p_n, 2q_n, r_n)$ és igual a $(p_1, 2q_1, r_1)$ per a tots els $n \geq 1$. Si fixem $x = p_0 + q_0$ per a la freqüència de l'al·lel A en la generació 0, llavors $1 - x = q_0 + r_0$ és la freqüència de l'al·lel a . Usant el sistema (11.1)–(11.3) una vegada més, obtenim

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

per a tot $n \geq 1$ (figura 11.2).

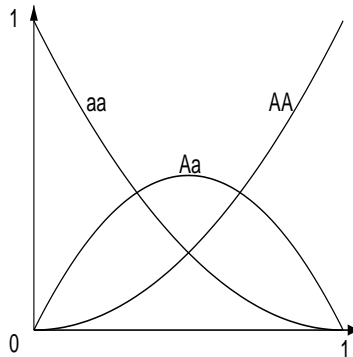


Figura 11.2: Gràfiques de les funcions x^2 , $2x(1-x)$ i $(1-x)^2$ corresponents a les freqüències d'equilibri dels genotips AA , Aa i aa .

En conclusió, les hipòtesis anteriors ens porten a la llei segons la qual les freqüències dels genotips AA , Aa i aa es mantenen invariants al llarg de les generacions. La teoria de Mendel no porta a un augment progressiu de la freqüència del caràcter dominant com s'havia pensat en un principi.

Alguns anys més tard, Fisher insistiria en un corol·lari important d'aquesta llei: en una primera aproximació (és a dir, suposant que les hipòtesis del model siguin realistes), una població manté una variància genètica constant. Aquesta observació resol un dels problemes plantejats per la teoria de l'evolució per selecció natural de Darwin. En efecte, Darwin pensava, igual que els seus contemporanis, que en cada generació les característiques fisiològiques

dels fills eren una mena de mitjana de les característiques dels dos progenitors, aportant cadascun d'ells una meitat. Aquesta idea va ser posteriorment estudiada a fons utilitzant estadística per Francis Galton i el seu successor en el laboratori de biometria, Karl Pearson. Si això fos cert, la variància d'aquestes característiques en una població hauria de dividir-se per dos en cada generació i aviat hi hauria una homogeneïtat tal que la selecció natural, suposada per a explicar l'evolució, seria impossible. No obstant, va caldre que passessin diversos anys perquè aquest mecanisme de mitjana fos rebutjat, ja que hi havia molts biomètrics que defensaven el punt de vista original de Darwin i eren reticents a admetre que les lleis de Mendel eren inevitables per a entendre l'evolució.

Després d'aquest treball, el 1908, Hardy va retornar a les matemàtiques pures. En la seva autobiografia, *La disculpa d'un matemàtic*, fins i tot va afirmar amb orgull haver evitat descobriments que tinguessin qualsevol ús pràctic. El 1910 va ser escollit membre de la *Royal Society*. El 1913 va descobrir al prodigi indi Ramanujan i el va convidar a treballar a Cambridge. Després de la Primera Guerra Mundial, va esdevenir professor de la Universitat d'Oxford i va continuar una fructífera col·laboració amb el seu compatriota Littlewood. Entre els anys 1931 i 1942 va tornar a ser professor a Cambridge. Va publicar molts llibres, sovint en col·laboració: *Ordres d'infininitat* (1910), *La teoria general de les sèries de Dirichlet* amb Marcel Riesz (1915), *Desigualtats* amb Littlewood i Pólya (1934), *Introducció a la teoria de nombres* amb E. M. Wright (1938), *Ramanujan* (1940), *Sèries de Fourier* amb Rogosinski (1944) i *Sèries divergents* (1949). Va morir a Cambridge el 1947.

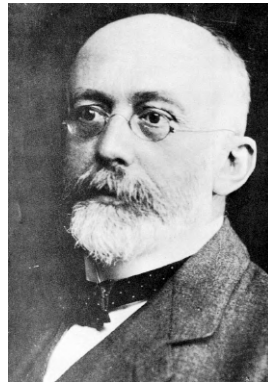


Figura 11.3:
Weinberg (1862–1937)

Diverses dècades després, es va saber que la llei de Hardy per a les freqüències genètiques també havia estat descoberta aquell mateix any, el 1908,

per un metge alemany, Wilhelm Weinberg. Weinberg va néixer a Stuttgart el 1862. Després d'estudiar a Tübingen i Munic fins a doctorar-se en medicina, va treballar diversos anys en hospitals de Berlín, Viena i Frankfurt. Va fixar la seva residència el 1889 a Stuttgart com a metge practicant general i d'obstetrícia. Malgrat estar molt ocupat amb la seva feina, havia trobat temps per a escriure molts articles en revistes científiques alemanyes. El 1901 havia estudiat des d'un punt de vista estadístic la freqüència de bessons del mateix sexe. L'article del 1908, en el qual explicava la mateixa llei que havia trobat Hardy, havia estat publicat en una revista científica local i no havia cridat l'atenció. Però, a diferència de Hardy, havia continuat aquest estudi els anys següents, descobrint, per exemple, la generalització al cas en què hi ha més de dos al·lels. També havia contribuït a l'àrea de l'estadística mèdica. Weinberg va morir el 1937. Després del redescobriment del seu article del 1908, els genetistes van anomenar a la llei d'estabilitat de les freqüències dels genotips la «llei de Hardy-Weinberg».

Avui en dia, aquesta llei s'utilitza sovint de la següent manera: si un al·lel recessiu poc freqüent a no té cap influència en la supervivència o la fertilitat i si coneixem la freqüència x^2 del genotip aa perquè aa produeix un fenotip particular, llavors podem calcular x i estimar la freqüència $2x(1-x) \approx 2x$ del genotip Aa . Com a exemple, si la freqüència de aa és $1/20\,000$, llavors tenim que $x \approx 1/140$. Així que $2x \approx 1/70$ és la freqüència del genotip Aa . L'al·lel recessiu a , que podria semblar molt rar si inspeccionem els fenotips, no és en realitat tan infreqüent.

Més informació

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

Capítol 12

Ross i la malària (1911)

L'any 1911, el metge britànic Ronald Ross, que ja havia rebut el premi Nobel el 1902 pels seus treballs sobre la malària, va estudiar un sistema d'equacions diferencials que modelitzava la propagació d'aquesta malaltia. Va demostrar que la malària només pot persistir si el nombre de mosquits supera un determinat llindar. Per tant, no és necessari matar tots els mosquits per a erradicar la malària, sinó que n'hi ha prou amb matar-ne només una determinada fracció. Kermack i McKendrick van desenvolupar posteriorment models epidèmics similars.

Ronald Ross va néixer el 1857 al nord de l'Índia, on el seu pare era oficial de l'exèrcit britànic. Va estudiar medicina a Londres, però va preferir escriure poemes i drames. Després de treballar un any en un vaixell com a cirurgià, va aconseguir entrar al Servei Mèdic de l'Índia el 1881. La seva feina com a metge a l'Índia li deixava molt temps lliure, durant el qual va escriure obres literàries i va aprendre pel seu compte una mica de matemàtiques. De permís a Anglaterra, el 1888, va obtenir un diploma en salut pública i va estudiar bacteriologia, una nova ciència creada uns anys abans per Pasteur i Koch. De tornada a l'Índia, Ross va començar a estudiar la malària. Durant el seu segon permís, el 1894, va conèixer a Londres a Patrick Manson, un especialista en medicina tropical que li va mostrar al microscopi el que el metge militar francès Alphonse Laveran havia observat el 1880: la sang dels pacients amb malària conté paràsits. Manson va suggerir que els paràsits podien procedir dels mosquits perquè ell mateix havia descobert a la Xina el paràsit d'una altra malaltia tropical (la filariasis) en aquests insectes. No obstant, creia que els humans s'infectaven amb el paràsit en beure aigua contaminada pels mosquits. Des del 1895 fins al 1898, Ross va continuar la seva recerca a l'Índia i va posar a prova la idea de Manson. El 1897 va descobrir a l'estómac d'una espècie de mosquit que no havia estudiat abans (*Anopheles*) uns paràsits similars als observats per Laveran. Com que els seus superiors el van enviar a Calcuta durant una temporada, on els casos de paludisme eren escassos, va decidir estudiar el paludisme als ocells de gàbia. Va trobar el paràsit a les glàndules salivals dels mosquits *Anopheles* i va aconseguir infectar experimentalment ocells sans deixant que els mosquits els piquessin: això va demostrar que el paludisme es transmet per picades de mosquits i no

per ingestió d'aigua contaminada. El 1899 Ross va deixar el Servei Mèdic de l'Índia per ensenyar a l'Escola de Medicina Tropical de Liverpool, que s'havia creat un any abans. Va ser escollit membre de la *Royal Society* el 1901 i va rebre el 1902 el Premi Nobel de Fisiologia o Medicina pels seus treballs sobre el paludisme. Va viatjar a l'Àfrica, a Maurici i a la zona del Mediterrani per a popularitzar la lluita contra els mosquits. El mètode va tenir èxit a Egipte al llarg del canal de Suez, al llarg del canal de Panamà en construcció en aquell moment, a Cuba i a Malàisia. Va tenir menys èxit en altres zones. Ross va publicar un *Informe sobre la prevenció de la malària a Maurici* el 1908 i *La prevenció de la malària* el 1910.



Figura 12.1:
Ross (1857–1932)

Malgrat les seves evidències sobre el paper d'un cert tipus de mosquits en la transmissió de la malària, Ross va topiar amb l'escepticisme quan va afirmar que la malària podia erradicar-se simplement reduint el nombre de mosquits. En la segona edició del seu llibre sobre *La prevenció del paludisme*, publicat el 1911, va intentar construir models matemàtics de la transmissió del paludisme per donar suport a la seva afirmació. Un dels seus models consistia en un sistema de dues equacions diferencials. Introduïm les següents notacions:

- N : població humana total en una zona determinada (se suposa constant);
- $I(t)$: nombre d'humans infectats de malària a l'instant t ;
- n : població total de mosquits (se suposa constant);
- $i(t)$: nombre de mosquits infectats de malària a l'instant t ;
- b : freqüència de picada dels mosquits;
- p (respectivament p'): probabilitat de transmissió de la malària d'humà a mosquit (respectivament de mosquit a humà) durant una picada;

- a : ritme de recuperació de la malària en els éssers humans;
- m : mortalitat dels mosquits.

Durant un interval petit de temps dt , cada mosquit infectat pica bdt humans, dels quals una fracció igual a $\frac{N-I}{N}$ encara no està infectada. Tenint en compte la probabilitat de transmissió p' , hi ha $b p' i \frac{N-I}{N} dt$ nous humans infectats. Durant el mateix interval de temps, el nombre d'humans que es recuperen és $aI dt$. Per tant, obtenim l'equació

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

De manera anàloga, cada mosquit no infectat pica a bdt humans, dels quals una fracció igual a I/N ja està infectada. Tenint en compte la probabilitat de transmissió p , hi ha $b p (n-i) \frac{I}{N} dt$ nous mosquits infectats. Mentrestant, suposant que la infecció no influeix en la mortalitat, el nombre de mosquits que moren és $mi dt$. Per tant, s'obté l'equació

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

Atès que la malària existeix de manera permanent en la majoria dels països infectats, Ross va considerar només els estats estacionaris (o d'equilibri) del seu sistema de dues equacions diferencials: el nombre d'humans infectats $I(t)$ i el nombre de mosquits infectats $i(t)$ es mantenen constants al llarg del temps ($dI/dt = 0$ i $di/dt = 0$). En primer lloc, sempre existeix l'estat estacionari amb $I = 0$ i $i = 0$, que correspon a l'absència de malària. En segon lloc, Ross va buscar un estat estacionari tal que $I > 0$ i $i > 0$ i va trobar que

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

Dividint les equacions de l'estat estacionari pel producte $I \times i$, el problema es converteix en un sistema lineal de dues equacions amb dues incògnites $1/I$ i $1/i$,

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

La seva solució es pot obtenir fàcilment.

Es pot observar que $I > 0$ i $i > 0$ si el nombre de mosquits està per sobre d'un llindar crític:

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

En aquest cas, l'estat estacionari correspon a la situació en la qual la malaltia és endèmica, és a dir, és present de manera permanent. Ross va arribar a la conclusió que si el nombre de mosquits n es redueix per sota del llindar crític n^* , l'únic estat estacionari possible és $I = 0$ i $i = 0$, per la qual cosa la malària hauria de desaparèixer. En concret, no és necessari exterminar tots els mosquits per a erradicar la malària. Aquest és precisament el punt que Ross volia destacar amb el seu model.

Per a il·lustrar la seva teoria, Ross va buscar valors numèrics raonables per als paràmetres del seu model. Va suposar que

- la mortalitat dels mosquits és tal que només un terç d'ells continua viu després de deu dies; per tant, $e^{-10m} = \frac{1}{3}$ i $m = (\log 3)/10$ per dia;
- la meitat de les persones segueixen infectades després de tres mesos; per tant, $e^{-90a} = 1/2$ i $a = (\log 2)/90$ per dia;
- un de cada vuit mosquits pica cada dia; per tant, $e^{-b} = 1 - 1/8$ i $b = \log(8/7)$ per dia.
- els mosquits infectats no solen ser infecciosos durant els primers deu dies després de la seva infecció perquè els paràsits han de passar per diverses etapes de transformació. Com que un terç dels mosquits pot sobreviure deu dies, Ross va suposar que també hi ha un terç de tots els mosquits infectats que són infecciosos: $p' = 1/3$;
- $p = 1/4$.

Ross va poder llavors calcular amb la fórmula (12.1) la fracció infectada I/N en la població humana en funció de la relació n/N entre la població de mosquits i la població humana. Va mostrar els seus resultats en una taula equivalent a la figura 12.2.

La forma de la corba mostra que la fracció d'humans infectats ja és superior al 50% si la relació n/N es troba lleugerament per sobre del valor crític n^*/N . Però aquesta fracció no canvia molt quan la relació n/N augmenta més. Això explica per què la correlació entre el nombre de mosquits i la presència de malària mai s'havia observat abans. No obstant això, Ross va observar que el valor numèric del llindar n^*/N era molt sensible a petits canvis en la taxa de picades b , però que això no canviava la forma general de la corba de la figura 12.2. La seva explicació qualitativa és més important que els resultats quantitius, que es veuen afectats per la incertesa que envolta els valors numèrics dels paràmetres.

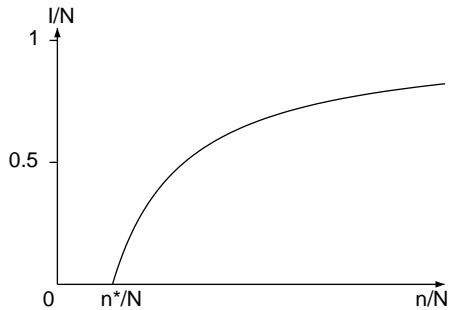


Figura 12.2: Fracció I/N d'humans infectats en funció de la relació n/N entre la població de mosquits i la població humana.

Per a interpretar el llindar crític n^* descobert per Ross¹, considerem un humà infectat introduït en una població d'humans i mosquits, tots dos lliures de malària. Aquest humà es manté infectat de mitjana durant un període de temps igual a $1/a$. Rep bn/N picades per unitat de temps, per la qual cosa de mitjana rep $bn/(aN)$ picades en total mentre està infectat. Així que infecta una mitjana de $bpn/(aN)$ mosquits. Cadascun d'aquests mosquits infectats viu de mitjana durant un període de temps igual a $1/m$, pica a b/m humans i infecta bp'/m humans. En total, després de la transmissió del primer humà infectat als mosquits i d'aquests a altres humans, el nombre mitjà de nous humans infectats és el producte dels dos resultats anteriors, és a dir

$$\mathcal{R}_0 = \frac{b^2 p p' n}{amN}. \quad (12.2)$$

Aquest \mathcal{R}_0 , el nombre reproductiu bàsic en un context epidemiològic, és interpretat com el nombre de casos humans secundaris deguts a un cas humà primari. Així, el procés d'infecció que es produeix de manera contínua en el temps també pot considerar-se a través de generacions successives. La malària pot «envair» la població només si $\mathcal{R}_0 > 1$. Aquesta condició equival precisament a $n > n^*$.

En conclusió, Ross va advocar de forma més general per la modelització matemàtica en epidemiologia:

«De fet, tota l'epidemiologia, que s'ocupa de la variació de la malaltia d'un temps a un altre o d'un lloc a un altre, ha de considerar-se matemàticament, a pesar que puguin haver-hi moltes

¹Aquesta interpretació no es va posar en relleu fins molt després del treball de Ross.

variables implicades, si es vol considerar científicament. Dir que una malaltia depèn de certs factors no és dir molt, fins que puguem també donar una estimació de la mesura en què cada factor influeix en el resultat global. I el mètode matemàtic de tractament no és més que l'aplicació d'un raonament acurat als problemes en qüestió.»

Ross va ser nomenat cavaller el 1911. Es va traslladar a Londres i es va convertir en assessor de l'exèrcit britànic durant la Primera Guerra Mundial. El 1923 va publicar la seva autobiografia, *Memòries amb un relat complet del gran problema de la malària i la seva solució*. El 1926 es va inaugurar l'Institut Ross de Malalties Tropicals (actualment part de l'Escola d'Higiene i Medicina Tropical de Londres), del qual va arribar a ser director. Ross va morir a Londres el 1932.

Més informació

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)

Capítol 13

Lotka, Volterra i el sistema depredador-presa (1920–1926)

L'any 1920 n'Alfred Lotka va estudiar un model depredador-presa i va demostrar que les poblacions podien oscil·lar permanentment. Va desenvolupar aquest estudi en el seu llibre del 1925 *Elements de biologia física*. El 1926, el matemàtic italià Vito Volterra es va interessar pel mateix model per a respondre una pregunta plantejada pel biòleg Umberto d'Ancona: per què hi havia més peixos depredadors capturats pels pescadors al mar Adriàtic durant la Primera Guerra Mundial, quan l'esforç pesquer era baix?

El 1920 Lotka va publicar un article titulat *Nota analítica sobre certes relacions rítmiques als sistemes orgànics*. Des de feia alguns anys, s'havia interessat per algunes reaccions químiques que mostraven estranyes oscil·lacions transitòries en experiments de laboratori. L'objectiu del seu article era suggerir que un sistema de dues espècies biològiques podia fins i tot oscil·lar permanentment. L'exemple que considerava era el d'una població d'herbívors alimentant-se de plantes. Per analogia amb les equacions utilitzades a la cinètica química, denotant per $x(t)$ la massa total de les plantes i per $y(t)$ la massa total dels herbívors a temps t , Lotka va utilitzar com a model el següent sistema d'equacions diferencials

$$\frac{dx}{dt} = ax - bxy, \quad (13.1)$$

$$\frac{dy}{dt} = -cy + dxy, \quad (13.2)$$

on els paràmetres a , b , c i d són tots positius. El paràmetre a és la taxa (per capita) de creixement de les plantes en absència d'herbívors, mentre que c és la taxa (per capita) de decreixement de la població d'herbívors quan no hi ha plantes disponibles. Els termes $-bxy$ i dxy expressen que com més animals i plantes hi ha, major és la transferència de massa de les plantes cap als animals (la transferència inclou alguna pèrdua de massa per la qual cosa $d \leq b$). Fixant $dx/dt = 0$ i $dy/dt = 0$, Lotka va observar que hi ha dos estats estacionaris:

- Un equilibri trivial ($x = 0, y = 0$), la població d'herbívors s'extingeix i no hi ha plantes;
- Un equilibri de coexistència ($x = c/d, y = a/b$), on conviuen els herbívors i les plantes.

També va escriure sense demostrar-ho que si a temps $t = 0$, $(x(0), y(0))$ no és un d'aquests dos estats estacionaris, llavors les funcions $x(t)$ i $y(t)$ oscil·len periòdicament: hi ha un nombre $T > 0$ tal que $x(t+T) = x(t)$ i $y(t+T) = y(t)$ per a tot $t > 0$ (figura 13.1)¹. Si, per exemple, les plantes són molt abundants, la població d'herbívors augmentarà, provocant una disminució de la massa total de plantes. Quan aquesta massa sigui insuficient per a alimentar als herbívors, alguns animals moriran de fam i la massa total de plantes començarà a créixer de nou fins a aconseguir un nivell igual al seu valor inicial. El fenomen es repetirà de manera indefinida al llarg del temps.

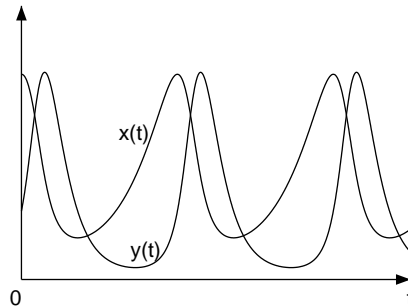


Figura 13.1: Oscil·lacions de la biomassa total de plantes $x(t)$ i de la biomassa total d'herbívors $y(t)$ en funció del temps.

Lotka va continuar estudiant el model en un segon article publicat el 1920, titulat *Oscil·lacions sense amortiment derivades de la llei d'acció de masses*. Va explicar perquè el sistema podia oscil·lar de manera periòdica. Això es dedueix del fet que el punt $(x(t), y(t))$ ha de romandre en una trajectòria tancada al pla amb x a l'eix horitzontal i y a l'eix vertical; més concretament, al quadrant on $x \geq 0$ i $y \geq 0$ (figura 13.2).

En efecte, dividint l'equació (13.1) per l'equació (13.2), obtenim, des-

¹El període T depèn de les condicions inicials, però Lotka no es va adonar d'aquest fet fins al 1925.

prés de reordenar

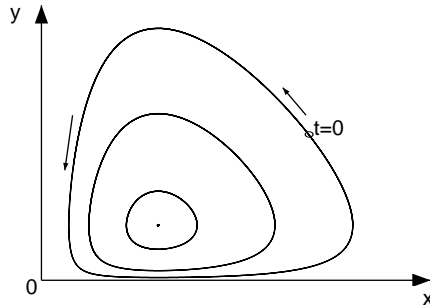
$$\left(-\frac{c}{x} + d\right) \frac{dx}{dt} = \left(\frac{a}{y} - b\right) \frac{dy}{dt}.$$

Integrant obtenim

$$dx(t) - c \log x(t) = a \log y(t) - by(t) + K,$$

on K és una constant que depèn només de la condició inicial. Per tant, el punt $(x(t), y(t))$ es manté a la corba $dx - c \log x = a \log y - by + K$, que resulta ser una corba tancada al pla (figura 13.2).

Figura 13.2: Diagrama amb la biomassa total de plantes $x(t)$ a l'eix horitzontal i la biomassa total d'herbívors $y(t)$ a l'eix vertical. Les tres corbes tancades al voltant de l'estat estacionari corresponen a diferents condicions inicials.



La trajectòria de $(x(t), y(t))$ gira entorn de l'estat estacionari $(c/d, a/b)$ en sentit contrari a les agulles del rellotge, com pot veure's fàcilment estudiant el signe de dx/dt i de dy/dt . A prop de l'estat estacionari, el sistema presenta petites oscil·lacions amb un període igual a $2\pi/\sqrt{ac}$.

En efecte, fixem $x = \frac{c}{d} + x^*$ i $y = \frac{a}{b} + y^*$ on $|x^*| \ll \frac{c}{d}$ i $|y^*| \ll \frac{a}{b}$. Llavors

$$\begin{aligned} \frac{dx^*}{dt} &= -by^* \left(\frac{c}{d} + x^*\right) \approx -\frac{bc}{d} y^*, \\ \frac{dy^*}{dt} &= dx^* \left(\frac{a}{b} + y^*\right) \approx \frac{ad}{b} x^*. \end{aligned}$$

A partir d'aquestes dues equacions, obtenim

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

Aquestes equacions són les mateixes que les que modelitzen les oscil·lacions del pèndol simple a la física. El període és $2\pi/\sqrt{ac}$.

Raymond Pearl, que havia estat l'editor del primer article del 1920 als *Proceedings of the National Academy of Sciences*, va ajudar Lotka a aconseguir una beca de dos anys a la Universitat Johns Hopkins per a escriure un llibre titulat *Elements de biologia física*. El llibre es va publicar el 1925. La secció que resumia el treball del 1920 també esmentava que sistemes de dues espècies, una espècie com a hoste i una altra com a paràsit o una espècie com a presa i una altra com a depredador, podien descriure's amb el mateix model (13.1)–(13.2). Desgraciadament, el llibre de Lotka no va tenir molta repercussió quan es va publicar. No obstant, el famós matemàtic Volterra va redescobrir de manera independent aquest mateix model poc després mentre estudiava un problema relacionat amb la pesca.

Vito Volterra va néixer al gueto jueu d'Ancona el 1860, poc abans de la unificació d'Itàlia, quan la ciutat encara pertanyia als Estats Pontificis. Era fill únic. El seu pare, un comerciant de teixits, va morir quan Vito tenia dos anys i va deixar sense diners la família. Bon estudiant a l'institut, Volterra va aconseguir continuar estudiant malgrat la pobresa, primer a la Universitat de Florència i després a la *Scuola Normale Superiore* di Pisa. El 1882 es va doctorar en física i l'any següent va esdevenir professor de mecànica a la Universitat de Pisa. El 1892 es va incorporar a la Universitat de Torí i el 1900 es va traslladar a la Universitat de la Sapienza a Roma ocupant una càtedra de física matemàtica. Va esdevenir senador el 1905. Moltes de les conferències que va pronunciar a Roma o a universitats estrangeres es van publicar en forma de llibre: *Tres lliçons sobre alguns progressos recents de la física matemàtica* (Universitat de Clark, 1909), *Lliçons sobre equacions integrals i integrodiferencials* (Roma, 1910), *Lliçons sobre funcions lineals* (París, 1912), *La teoria de les funcions permutables* (Princeton, 1912). Va ser oficial de l'exèrcit italià durant la Primera Guerra Mundial i va dirigir l'oficina d'invents de guerra. Després de la guerra, va participar activament en la fundació de la Unió Matemàtica Italiana (1922) i del Consell Nacional Italià de Recerca (1923), essent el primer president d'aquest últim. També va ser president de la Comissió Internacional per a l'Estudi Científic de la Mar Mediterrània (1923) i president de la històrica Accademia dei Lincei (1924). L'any 1924 es va publicar una altra monografia, escrita en col·laboració amb

J. Péres, *Lliçons sobre composició i funcions permutables.*

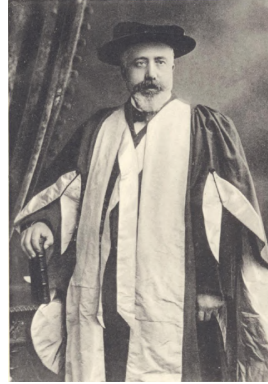


Figura 13.3: Volterra (1860–1940) va rebre el doctorat *honoris causa* per la Universitat de Cambridge el 1900.

El 1925, a l'edat de 65 anys, Volterra es va interessar per un estudi del zoòleg Umberto D'Ancona, que més tard es convertiria en el seu gendre, sobre la proporció de peixos cartilaginosa (com els taurons i les ratlles) pescats durant els anys 1905–1923 a tres ports del mar Adriàtic: Trieste, Fiume² i Venècia. D'Ancona havia observat que la proporció d'aquests peixos havia augmentat durant la Primera Guerra Mundial, quan l'esforç pesquer s'havia reduït (taula 13.1).

Taula 13.1: Percentatge de peixos cartilaginosa en les pesqueres de Trieste, Fiume i Venècia abans, durant i després de la Primera Guerra Mundial.

any	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
Venècia	21,8	-	-	-	-	-	-

any	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
Venècia	-	-	30,9	25,3	25,9	25,8	26,6

Essent els peixos cartilaginosa depredadors dels peixos més petits, semblava que una disminució de l'esforç pesquer hauria d'afavorir a les espècies

²Actualment Rijeka a Croàcia.

depredadors. Volterra, que no coneixia els treballs de Lotka, va explicar aquesta observació utilitzant el mateix model

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

on $x(t)$ representa el nombre de preses i $y(t)$ el nombre de depredadors. Va notar, igual que Lotka, que aquest sistema pot oscil·lar de manera periòdica amb un període T que depèn de la condició inicial (x_0, y_0) . També va observar que

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

Integrant sobre un període T (de manera que $x(0) = x(T)$ i $y(0) = y(T)$), va obtenir

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

Per tant, la mitjana al llarg d'un període tant del nombre de preses com del nombre de depredadors és independent de les condicions inicials. A més a més, si l'esforç pesquer disminueix, la taxa (per capita) de creixement a de les preses augmenta, mentre que la taxa (per capita) d'extinció c dels depredadors disminueix. En conclusió, la mitjana de $x(t)$ disminueix i la mitjana de $y(t)$ augmenta: la proporció de depredadors augmenta. Això és precisament el que s'havia observat a les estadístiques pesqueres del mar Adriàtic.

Volterra va publicar el seu article primer en italià el 1926. Uns mesos després va aparèixer un resum en anglès a *Nature*. Lotka va informar Volterra i altres científics de l'existència prèvia dels seus estudis dels sistemes depredador-presa. Però el seu article del 1920 i el seu llibre del 1925 no serien sempre citats. Lotka ja treballava llavors per a una companyia d'assegurances, per la qual cosa el seu treball es va centrar en demografia humana. Volterra va continuar treballant en variants del sistema depredador-presa durant una dècada. Els anys 1928-1929 va impartir una sèrie de conferències al recentment creat Institut Henri Poincaré de París. Les notes d'aquestes conferències es van publicar el 1931 amb el títol de *Lliçons sobre la teoria matemàtica de la lluita per la vida*. El 1935 Volterra va publicar en col·laboració amb Umberto D'Ancona un altre llibre titulat *Associacions biològiques des d'un punt de vista matemàtic*.

Tot i que el model depredador-presa sembla explicar correctament les dades de les pesqueres, el debat sobre el realisme dels models simplificats en ecologia acabava tan sols de començar i continua sent encara objecte de disputa científica. Avui en dia, el model depredador-presa es coneix àmpliament com el model de Lotka-Volterra i és un dels més citats en treballs d'ecologia.

El 1931, Volterra es va negar a signar el jurament de lleialtat al feixisme promogut pel govern de Mussolini. Va perdre la seva condició de catedràtic a la universitat de Roma i va ser expulsat de les acadèmies científiques italianes, de les quals era un dels membres més famosos. A partir de llavors va romanre principalment fora d'Itàlia, viatjant per Europa i donant conferències. Va publicar amb J. Péres el primer volum d'una *Teoria general dels funcionals* (1936) i un llibre amb B. Hostinský sobre *Operacions lineals infinitesimals* (1938). Va morir a Roma el 1940.

Més informació

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*, 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) gallica.bnf.fr
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)

Capítol 14

Fisher i la selecció natural (1922)

L'any 1922, el biomatemàtic britànic Ronald Fisher va publicar un article molt influent sobre genètica de poblacions. Aquest capítol considera només una part de l'article, que se centra en una variant del model de Hardy-Weinberg que inclou la selecció natural. Fisher va demostrar que si l'heterozigot es veu afavorit, tots dos al·lells poden coexistir. Si un dels dos homozigots es veu afavorit, llavors l'altre al·lel desapareix. El problema subjacent és el d'explicar per què alguns gens poden tenir diversos al·lells.

Ronald Aylmer Fisher va néixer a Londres el 1890. Era el petit de sis germans. El seu pare era subhastador, però més tard es va declarar en fallida. Fisher va estudiar matemàtiques i física al *Gonville and Caius College* de la Universitat de Cambridge entre els anys 1909 i 1913. El camp de la genètica va tenir un desenvolupament ràpid en aquesta època. A partir del 1911, Fisher va participar a les reunions de la Societat d'Eugenèsia iniciades per Galton. Va començar a centrar-se en problemes estadístics relacionats amb els treballs de Galton i Mendel. Després de finalitzar els seus estudis universitaris, va passar un estiu treballant a una granja al Canadà i després va treballar per a la *Mercantile and General Investment Company* a la *City* de Londres. A causa de la seva extrema miopia, no va poder participar en la Primera Guerra Mundial malgrat haver-se presentat com a voluntari. Va passar aquests anys fent classes en instituts de secundària. Durant el seu temps lliure, es va ocupar d'una granja i va continuar amb la seva recerca. Va obtenir nous i importants resultats que relacionaven els coeficients de correlació amb la genètica mendeliana. El 1919 va començar a treballar com a estadístic a l'Estació Experimental de Rothamsted, centrada en l'agricultura.

El 1922, Fisher va publicar un article titulat *Sobre la raó de dominància*. Entre altres importants noves idees, aquest article considerava un model matemàtic que combinava les lleis de Mendel i la idea de la selecció natural destacada per Darwin a la teoria de l'evolució. Fisher va considerar la mateixa situació que Hardy amb dos al·lells A i a i amb la hipòtesi de l'aparellament aleatori. Però va suposar que els individus amb genotips AA , Aa i aa tenen mortalitats diferents abans d'arribar a l'edat adulta, imitant així la selecció natural. Establint p_n , $2q_n$ i r_n per a les freqüències dels tres genotips

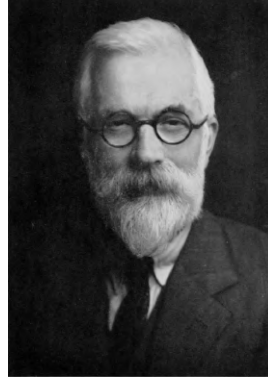


Figura 14.1:
Fisher (1890–1962)

entre els individus adults a la generació n , hi ha respectivament $(p_n + q_n)^2$, $2(p_n + q_n)(q_n + r_n)$ i $(q_n + r_n)^2$ nounats a la generació $n + 1$ que tenen aquests genotips. Siguin u , v i w les respectives probabilitats de supervivència des del naixement fins a l'edat adulta. Llavors les freqüències dels genotips entre els individus adults de la generació $n + 1$ són p_{n+1} , $2q_{n+1}$ i r_{n+1} amb

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

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

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

on fixem per conveniència

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

Recordant que $p_n + 2q_n + r_n = 1$, veiem que quan $u = v = w$ (és a dir, quan no hi ha selecció natural), el sistema (14.1)-(14.3) es redueix al sistema (11.1)-(11.3) considerat per Hardy.

Sigui $x_n = p_n + q_n$ la freqüència de l'al·lel A entre els individus adults de la generació n . Llavors $q_n + r_n = 1 - x_n$ és la freqüència de l'al·lel a . Sumant (14.1) i (14.2), obtenim

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

Aquesta equació es pot reescriure de la forma

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

Sempre hi ha com a mínim dos estats estacionaris en els quals la freqüència x_n es manté constant al llarg de les generacions: $x = 0$ (la població està formada totalment per *aa* homozigots) i $x = 1$ (la població està formada totalment per *AA* homozigots).

Utilitzant l'equació (14.4), es pot demostrar que si l'homozigot *AA* té més possibilitats de sobreviure que els altres dos genotips ($u > v$ i $u > w$), llavors l'al·lel *a* desapareixerà progressivament de la població. Aquest cas no hauria de ser molt comú a la natura si sabem que tots dos al·lells coexisteixen. No obstant, si l'heterozigot *Aa* té un avantatge selectiu sobre els homozigots *AA* i *aa* ($v > u$ i $v > w$), llavors els tres genotips poden coexistir en la població. Aquest és el cas més comú i pot explicar el «vigor» dels híbrids observat pels agricultors.

En efecte, l'estat estacionari $x = 1$ és estable quan $u > v$ perquè $x_{n+1} - x_n \approx (1 - x_n)(u - v)/u$ quan x_n és pròxim a 1. La població tendeix a aquest estat estacionari. L'estat estacionari $x = 1$ és inestable quan $u < v$, i en aquest cas hi ha un tercer estat estacionari

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

amb $0 < x^* < 1$. A més a més, podem comprovar que aquest estat és estable. L'estat estacionari x^* correspon a una mescla entre els tres genotips.

Per tant, combinant simplement les lleis de Mendel i una hipòtesi de selecció natural (aquí, diferents probabilitats de supervivència per als tres genotips), podem explicar les dues situacions de coexistència o desaparició de genotips. Després de Fisher, aquest model també va ser desenvolupat per Haldane (vegeu el capítol 17) i per Wright (vegeu el capítol 19).

Anticipant-nos al capítol 20, observeu que si *A* és completament dominant i l'homozigot *aa* està en desavantatge respecte als altres dos genotips, amb els nombres $u : v : w$ en una proporció $1 : 1 : 1 - \epsilon$, llavors l'equació (14.4) es converteix en

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

per a $\varepsilon \ll 1$. Si la supervivència de l'heterozigot Aa es troba a mig camí entre la dels dos homozigots, aleshores els nombres $u : v : w$ estan en una proporció $1 : 1 - \varepsilon/2 : 1 - \varepsilon$ i

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

quan $\varepsilon \ll 1$.

A Rothamsted, Fisher va analitzar dades a llarg termini sobre el rendiment dels cultius i la meteorologia. Però també va fer grans aportacions a la metodologia estadística. El 1925 va publicar un llibre titulat *Mètodes estadístics per als investigadors*, que va tenir un gran èxit i es va reeditar moltes vegades. El 1929 va ser nomenat membre de la *Royal Society*. El 1930 Fisher va publicar un llibre sobre *La teoria genètica de la selecció natural*, una fita en la història de la genètica de poblacions. El 1933 va esdevenir professor d'eugenesia a l'*University College* de Londres, succeint Karl Pearson al Laboratori Galton. El 1943 es va traslladar a la Universitat de Cambridge amb una càtedra de genètica, aquesta vegada succeint R.C. Punnett (vegeu el capítol 11). També va publicar diversos llibres: *El disseny d'experiments* (1935), *La teoria de l'endogàmia* (1949) i *Mètodes estadístics i inferència científica* (1956). Nomenat cavaller britànic el 1952, es va instal·lar a Austràlia després de jubilar-se el 1959 i va morir a Adelaide el 1962. Tornarem a tractar una altra part de la seva obra al capítol 20.

Més informació

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)

Capítol 15

Yule i l'evolució (1924)

L'any 1924, l'estadístic britànic Yule va estudiar un model d'evolució en el qual les espècies poden produir noves espècies mitjançant petites mutacions i els gèneres poden produir nous gèneres mitjançant grans mutacions. El seu objectiu era explicar la distribució del nombre d'espècies dins de cada gènere, la majoria dels quals contenen una sola espècie, mentre que uns pocs gèneres contenen un gran nombre d'espècies. El «procés de naixement» estocàstic que va introduir Yule en el seu model continua sent una eina bàsica en l'estudi dels arbres filogenètics i en moltes altres àrees.

George Udny Yule va néixer a Escòcia el 1871. El seu pare ocupava un alt càrrec a l'administració britànica a l'Índia. Als 16 anys, Yule va començar a estudiar enginyeria a l'*University College* de Londres. L'any 1892 va canviar d'orientació i va passar un any investigant a Bonn sota la supervisió del físic Heinrich Hertz, que havia demostrat l'existència de les ones electromagnètiques uns anys abans. Quan Yule va tornar a Anglaterra, Karl Pearson li va oferir una plaça de professor adjunt de matemàtica aplicada a l'*University College*. Yule, seguint els passos de Pearson, va començar a centrar-se en l'estadística. El 1911 va publicar *Introducció a la teoria de l'estadística*, que es va reeditar 14 vegades. L'any següent es va traslladar a la Universitat de Cambridge. La seva recerca es va centrar en els aspectes teòrics de l'estadística, però també en les aplicacions a l'agricultura i a l'epidemiologia. El 1922 va ser nomenat membre de la *Royal Society*.

El 1924 Yule va publicar un article titulat *Una teoria matemàtica de l'evolució basada en les conclusions del Dr. J. C. Willis*. Willis era un col·lega de la *Royal Society* que havia publicat el 1922 un llibre titulat *Edat i àrea, un estudi sobre la distribució geogràfica i l'origen de les espècies*. Havia estudiat la distribució de les espècies entre els diferents gèneres en la classificació de plantes i animals. Les dades que havia recopilat mostraven que la majoria dels gèneres contienien una sola espècie, que cada vegada un nombre menor de gèneres contienien un major nombre d'espècies i que encara hi havia uns pocs gèneres que contienien un gran nombre d'espècies.

La taula 15.1 mostra les dades relatives a les serps, els llangardaixos i dues famílies d'escarabats (els *Chrysomelidae* i els *Cerambycinae*). Les 1.580



Figura 15.1:
Yule (1871-1951)

espècies de llangardaixos que es coneixien en aquell moment s'havien classificat en 259 gèneres, 105 dels quals contenien només una espècie, 44 només dues espècies, 23 només tres espècies, etc., i dos gèneres que contenien més de cent espècies. Per a altres famílies d'animals i plantes, la distribució dels gèneres segons el nombre d'espècies que contenien tenia una forma molt similar.

Yule va suggerir a Willis que intentés representar les seves dades en un gràfic amb escales logarítmiques. Això va donar un resultat sorprenent (figura 15.2): el logaritme del nombre Q_n de gèneres que contenen n espècies disminueix més o menys linealment amb $\log(n)$. En altres paraules, existeixen constants $\alpha > 0$ i $\beta > 0$ tals que

$$Q_n \approx \alpha n^{-\beta}.$$

La distribució segueix una «lleï potencial». Al seu article del 1924, Yule va buscar un model matemàtic d'evolució que pogués explicar aquesta distribució estadística.

Per a formular-lo va considerar primer un model estocàstic¹ a temps continu per al creixement del nombre d'espècies dins d'un gènere (figura 15.3a). Partint d'una sola espècie a l'instant $t = 0$, va suposar que la probabilitat que una espècie donés lloc per mutació a una nova espècie del mateix gènere durant un interval de temps «petit» (en l'escala temporal de l'evolució) era $r dt$ amb $r > 0$.

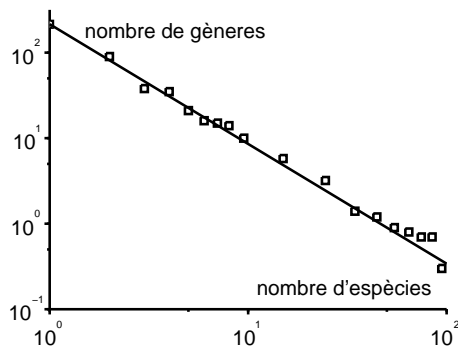
Sigui $p_n(t)$ la probabilitat que hi hagi n espècies a l'instant t (n és un nombre enter però t és un nombre real). Per a calcular $p_n(t + dt)$, Yule va

¹McKendrick (vegeu el capítol 16) ja havia començat a estudiar aquest tipus de models en dinàmica de poblacions en un treball publicat el 1914.

Taula 15.1: Dades compilades per Willis : (a) Escarabats *Chrysomelidae*, (b) Escarabats *Cerambycinae*, (c) Serps, (d) Llangardaixos.

Nombre d'espècies	Nombre de gèneres			
	(a)	(b)	(c)	(d)
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
total	627	1.024	293	259

Figura 15.2: El nombre de gèneres en funció del nombre d'espècies que contenen, en escala logarítmica decimal. Dades pels *Chrysomelidae*. Per a suavitzar les fluctuacions quan n (el nombre d'espècies) és gran, es van comptar els gèneres per a rangs de valors de n , com a la taula 15.1. El nombre mitjà de gèneres per a un únic valor de n pot ser, per tant, inferior a 1.



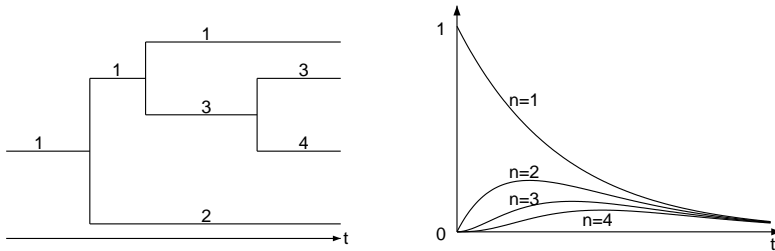


Figura 15.3: (a) Una simulació de l'evolució del nombre d'espècies dins d'un gènere. L'espècie 1 genera les espècies 2 i 3. L'espècie 3 genera l'espècie 4. (b) La probabilitat $p_n(t)$ que hi hagi n espècies dins del mateix gènere en el moment t , per a $1 \leq n \leq 4$.

considerar diversos casos:

- si hi ha $n - 1$ espècies a temps t , cada espècie té una probabilitat $r dt$ de generar una nova espècie entre t i $t + dt$; en el límit $dt \rightarrow 0$, hi haurà n espècies a temps $t + dt$ amb una probabilitat $(n - 1) r dt$;
- si hi ha n espècies a temps t , hi haurà $n + 1$ espècies a temps $t + dt$ amb una probabilitat $n r dt$.

Així, $p_n(t)$ ve donada pel següent sistema d'equacions diferencials

$$\frac{dp_1}{dt} = -r p_1, \quad \frac{dp_n}{dt} = (n - 1) r p_{n-1} - n r p_n \quad (15.1)$$

per a tot $n \geq 2$. De la primera equació obtenim $p_1(t) = e^{-rt}$ perquè $p_1(0) = 1$. Es pot demostrar que la solució de la segona equació que satisfà la condició inicial $p_n(0) = 0$ és

$$p_n(t) = e^{-rt} (1 - e^{-rt})^{n-1} \quad (15.2)$$

per a tot $n \geq 2$ (figura 15.3b). Així doncs, per a un temps fix t , la distribució de probabilitats $(p_n(t))_{n \geq 1}$ és una progressió geomètrica de raó igual a $1 - e^{-rt}$.

En efecte, observem en primer lloc que l'equació (15.1) és equivalent a

$$\frac{d}{dt} [p_n e^{nrt}] = (n - 1) r p_{n-1} e^{nrt}, \quad (15.3)$$

a partir de la qual, i usant les condicions inicials, podem calcular successivament $p_2(t)$, $p_3(t)$, Obtenim $p_2(t) = e^{-rt} (1 - e^{-rt})$, després

$p_3(t) = e^{-rt}(1 - e^{-rt})^2$, la qual cosa suggereix la fórmula (15.2) per a la solució general. Finalment es pot comprovar que aquesta fórmula és una solució de l'equació (15.3).

Yule també va deduir de la fórmula (15.2) que el nombre esperat d'espècies augmenta exponencialment amb el temps: $\sum_{n=1}^{+\infty} n p_n(t) = e^{rt}$.

En efecte, observem primer que per a $|x| < 1$,

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

Per tant,

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

En particular, si T és el temps de duplicació definit per $e^{rT} = 2$, llavors la distribució de probabilitat $(p_n(t))_{n \geq 1}$ del nombre d'espècies a temps $t = T$ és una progressió geomètrica de raó $1/2$: $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16} \dots$. A temps $t = kT$, és una progressió geomètrica de raó $1 - 1/2^k$ i $p_1(kT) = 1/2^k$.

A continuació, Yule va considerar, paral·lelament al creixement del nombre d'espècies pertanyents a un mateix gènere, un procés similar degut a grans mutacions que condueixen a la creació de nous gèneres. Sigui $s dt$ la probabilitat que un gènere existent generi un nou gènere durant un interval petit de temps dt . Com abans, suposant que només hi ha un gènere a l'instant $t = 0$, el nombre esperat de gèneres a l'instant t és e^{st} . El nombre mitjà de gèneres creats per unitat de temps a l'instant t és la derivada se^{st} . En el límit² quan $t \rightarrow +\infty$, el nombre mitjà de gèneres que a temps t han existit entre x i $x + dx$ unitats de temps és llavors $se^{s(t-x)} dx$. La probabilitat a temps t que un gènere triat a l'atzar hagi existit entre x i $x + dx$ unitats de temps és $se^{-sx} dx$.

Si un gènere triat a l'atzar a temps t ha existit entre x i $x + dx$ unitats de temps, la probabilitat que aquest gènere contingui n espècies és, segons la fórmula (15.2), igual a $e^{-rx}(1 - e^{-rx})^{n-1}$ per a tot $n \geq 1$. Així, la probabilitat q_n que un gènere triat a l'atzar a temps t contingui n espècies és $q_n = \int_0^{+\infty} se^{-sx} e^{-rx}(1 - e^{-rx})^{n-1} dx$. Fixem $u = r/s$. Un càlcul senzill dona

²Yule va considerar també el cas en el qual t no pot suposar-se molt gran en comparació amb el temps de duplicació de e^{st} . Els càlculs són una mica més complicats però els resultats finals no són molt diferents.

$$q_1 = 1/(1+u) \text{ i}$$

$$q_n = \frac{1}{1+u} \frac{u}{1+2u} \frac{2u}{1+3u} \cdots \frac{(n-1)u}{1+nu} \quad (15.4)$$

per a tot $n \geq 2$.

De fet, tenim $(1 - e^{-rx})^{n-1} = (1 - e^{-rx})^{n-2} (1 - e^{-rx})$. Per tant,

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

Integrant per parts, obtenim

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

La fórmula (15.4) mostra que la successió de probabilitats $(q_n)_{n \geq 1}$ és decreixent. Per tant, el màxim s'assoleix per a $n = 1$: la majoria dels gèneres contenen una sola espècie. Això és precisament el que mostren les dades. A més a més, el decreixement de q_n cap a 0 quan n tendeix a infinit és relativament lent perquè $q_n/q_{n-1} \rightarrow 1$. Això pot explicar perquè alguns gèneres contenen un gran nombre d'espècies. Més concretament, Yule va demostrar que $\log q_n$ disminueix linealment amb $\log(n)$.

Per definició, la funció Gamma d'Euler és $\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt$. Llavors $\Gamma(n+1) = n! = n \times (n-1) \times \cdots \times 2 \times 1$ si n és enter, i $\Gamma(z+1) = z\Gamma(z)$. Per tant (15.4) és de la forma

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

Però l'aproximació de Stirling ens dona que $\log \Gamma(n) \approx n \log n - n - \frac{1}{2} \log n + \text{constant}$. De la mateixa manera, $\log \Gamma(n+1+1/u) \approx n \log n - n + (\frac{1}{u} + \frac{1}{2}) \log n + \text{constant}$. Finalment

$$\log q_n \approx - \left(1 + \frac{1}{u} \right) \log n + \text{constant}.$$

Considerem, per exemple, el cas dels llangardaixos. El paràmetre u pot estimar-se a partir de la proporció $q_1 = 1/(1+u)$ de gèneres que contenen una

sola espècie. Segons la taula 15.1, tenim $q_1 = 105/259$, és a dir, $u \approx 1,467$. Llavors podem calcular la probabilitat teòrica q_n i el nombre esperat Q_n de gèneres que contenen n espècies multiplicant q_n pel nombre total d'espècies, que és 259 (taula 15.2).

Yule va observar que la similitud entre les observacions i els càlculs era relativament bona³ donada la simplicitat del model, que no té en compte, per exemple, els cataclismes que han patit les espècies al llarg de milions d'anys d'evolució.

Taula 15.2: Comparació entre les dades reals i les teòriques en el cas dels llangardaixos (1.580 espècies classificades en 259 gèneres).

Nombre d'espècies per gènere	Nombre observat de gèneres	Nombre calculat de gèneres
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
total	259	259

Després del 1931, Yule es va retirar progressivament de la Universitat de Cambridge. Es va interessar per la distribució estadística de la longitud de les frases per a identificar autors de llibres. Ho va aplicar, en particular, al llibre publicat per John Graunt (vegeu el capítol 2), però probablement inspirat per

³Per al nombre de gèneres que contenen més de 100 espècies, Yule va obtenir un millor ajust que a la Taula 15.2 en considerar que t no era gran en comparació amb el temps de duplicació de e^{st} .

William Petty. El 1944 va publicar un llibre titulat *L'estudi estadístic del vocabulari literari*. Va morir l'any 1951.

En l'actualitat, el model de Yule es continua utilitzant per analitzar els «arbres filogenètics» (els arbres genealògics de les espècies). Aquests arbres, similars al de la figura 15.3, són més coneguts gràcies a les noves dades procedents de la biologia molecular. Però les aplicacions del procés estocàstic definit per les equacions (15.1) no es limiten a la teoria de l'evolució. Aquest procés és la base de molts models en dinàmica de poblacions, tant a nivell microscòpic (per a modelitzar, per exemple, colònies de bacteris) com a nivell macroscòpic (per a modelitzar l'inici d'una epidèmia). S'anomena «procés de naixement pur» o «procés de Yule». Una variant senzilla inclou una probabilitat mdt de morir durant qualsevol interval petit de temps dt : la mida esperada de la població a temps t per a aquest «procés de naixement i mort» és llavors $e^{(r-m)t}$. Pel que fa a la distribució de probabilitat (15.4), s'anomena de vegades distribució de Yule. Les distribucions amb cues que satisfan les lleis de potència han interessat a diverses àrees de la ciència. L'estudi de les epidèmies sobre xarxes complexes aleatòries amb una distribució de graus (nombre de connexions de cada node) de tipus llei potencial n'és només un exemple.

Més informació

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*. 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. *Phil. Trans. Roy. Soc. Lond. B* 213, 21–87 (1925) gallica.bnf.fr

Capítol 16

McKendrick i Kermack sobre la modelització d'epidèmies (1926–1927)

L'any 1926 McKendrick va estudiar un model estocàstic d'epidèmies i va trobar un mètode per a calcular la probabilitat que una epidèmia assoleixi una determinada mida final. També va descobrir l'equació en derivades parcials que governa la densitat respecte a l'edat de les poblacions en un context a temps continu. El 1927 Kermack i McKendrick van estudiar un model epidèmic determinista i van obtenir una equació per a la mida final de l'epidèmia, que posa l'accent en un determinat llindar per a la densitat de població. Poden produir-se grans epidèmies per damunt d'aquest llindar però no per sota. Aquests treballs continuen sent molt utilitzats en l'epidemiologia contemporània.

Anderson Gray McKendrick va néixer el 1876 a Edimburg, sent el darrer de cinc fills. Va estudiar medicina a la Universitat de Glasgow, on el seu pare era professor de fisiologia. El 1900 va ingressar al Servei Mèdic de l'Índia. Abans d'anar a l'Índia, va acompanyar Ronald Ross en una missió per combatre la malària a Sierra Leone. Després va servir a l'exèrcit durant 18 mesos a Sudan. En arribar a l'Índia, va treballar com a metge a una presó de Bengala, on va intentar controlar la disenteria. El 1905 es va incorporar al nou Institut Central de Recerca Mèdica de Kasauli (al nord de l'Índia). Va treballar sobre la ràbia i també va estudiar matemàtiques. El 1920, després de contagiar-se d'una malaltia tropical, va tornar a Edimburg i va esdevenir superintendent del laboratori del *Royal College of Physicians*.

El 1926 McKendrick va publicar un article titulat *Aplicacions de les matemàtiques als problemes mèdics*, que contenia diverses idees noves. En particular, va introduir un model matemàtic a temps continu per a les epidèmies que tenia en compte el caràcter estocàstic dels processos d'infecció i recuperació.

Considerem una població de mida N amb una sola persona infectada inicialment. Les persones poden passar successivament per tres estats: l'estat susceptible (de ser infectat) S , l'estat infectat I i l'estat recuperat (amb immunitat permanent) R (figura 16.2)¹.

¹El model de Daniel Bernoulli (vegeu el capítol 4) incloïa els estats S i R però no I , essent la

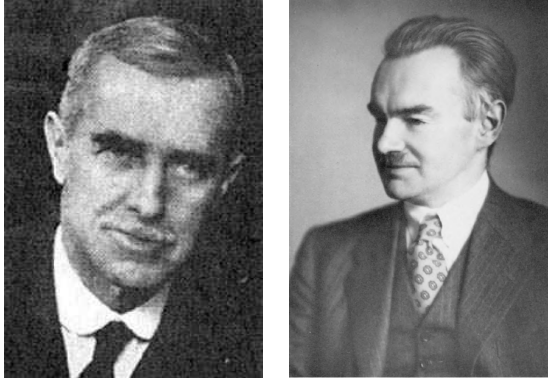


Figura 16.1: McKendrick (1876–1943) y Kermack (1898–1970)

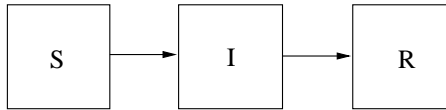


Figura 16.2: Estats possibles: susceptible (S), infectat (I), recuperat (R).

Sigui $p_{i,r}(t)$ la probabilitat que la població tingui a temps t exactament i persones a l'estat I i r persones a l'estat R , on i i r són nombres enters tals que $1 \leq i + r \leq N$. En aquest cas es diu que la població està a l'estat (i, r) . El nombre de persones susceptibles és llavors $s = N - i - r$. Seguint el treball de Ross sobre la malària (vegeu el capítol 12), McKendrick va assumir que, durant un interval petit de temps dt , la probabilitat que es produeixi una nova infecció és igual a $asi dt$ (és a dir, proporcional al nombre de persones susceptibles i al nombre de persones infectades). La probabilitat que es produeixi una nova recuperació és igual a $bidt$. Tant a com b són paràmetres positius. Per a calcular $p_{i,r}(t + dt)$, cal distingir diversos casos:

- la població està a l'estat $(i - 1, r)$ a temps t i una nova infecció fa que la població passi a l'estat (i, r) entre els temps t i $t + dt$; la probabilitat d'aquest succés és $as(i - 1) dt$ amb $s = N - (i - 1) - r$;
- la població està a l'estat (i, r) a temps t i una nova infecció fa que la població passi a l'estat $(i + 1, r)$ entre els temps t i $t + dt$; la probabilitat d'aquest succés és $asidt$ amb $s = N - i - r$;

durada de la infecció molt més curta que l'esperança de vida mitjana.

- la població està a l'estat $(i + 1, r - 1)$ a temps t i una nova recuperació fa que la població passi a l'estat (i, r) entre els temps t i $t + dt$; la probabilitat d'aquest succés és $b(i + 1)dt$;
- la població està a l'estat (i, r) a temps t i una nova recuperació fa que la població passi a l'estat $(i - 1, r + 1)$ entre els temps t i $t + dt$; la probabilitat d'aquest succés és $bi dt$.

Per tant, McKendrick va obtenir les equacions

$$\frac{dp_{i,r}}{dt} = a(N - i - r + 1)(i - 1)p_{i-1,r} - a(N - i - r)ip_{i,r} + b(i + 1)p_{i+1,r-1} - bip_{i,r} \quad (16.1)$$

per a $1 \leq i + r \leq N$. El primer terme del costat dret ha de ser nul quan $i = 0$, mentre que el tercer terme ha de ser nul quan $r = 0$. Les condicions inicials són $p_{i,r}(0) = 0$ per a tot (i, r) excepte $p_{1,0}(0) = 1$.

Amb aquest model McKendrick va aconseguir calcular la probabilitat que al final de l'epidèmia s'hagin infectat n persones, que és el límit de $p_{0,n}(t)$ quan $t \rightarrow +\infty$. En efecte, no és necessari resoldre el sistema (16.1), n'hi ha prou amb observar que, mentre hi hagi i persones infectades i r persones recuperades, la probabilitat d'una nova infecció durant un interval petit de temps dt és $a(N - i - r)idt$ i la probabilitat d'una nova recuperació és $bidt$. Així, les probabilitats de transició (com se solen anomenar en la teoria de les cadenes de Markov) de l'estat (i, r) a l'estat $(i + 1, r)$ o a l'estat $(i - 1, r + 1)$ són respectivament

$$\mathcal{P}_{(i,r) \rightarrow (i+1,r)} = \frac{a(N - i - r)}{a(N - i - r) + b}, \quad \mathcal{P}_{(i,r) \rightarrow (i-1,r+1)} = \frac{b}{a(N - i - r) + b},$$

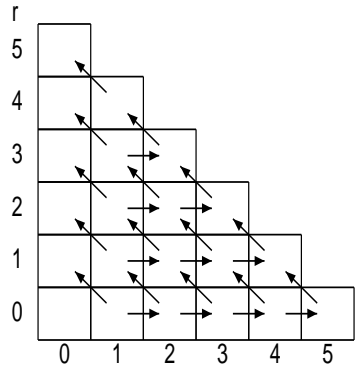
per a tot $i \geq 1$ (figura 16.3).

Sigui $q_{i,r}$ la probabilitat que la població passi per l'estat (i, r) durant l'epidèmia. Com que $i = 1$ i $r = 0$ quan $t = 0$, tenim $q_{1,0} = 1$. Els altres estats s'assoleixen després d'una infecció o després d'una recuperació:

$$q_{i,r} = q_{i-1,r} \mathcal{P}_{(i-1,r) \rightarrow (i,r)} + q_{i+1,r-1} \mathcal{P}_{(i+1,r-1) \rightarrow (i,r)}.$$

El primer terme del costat dret ha de ser nul quan $i = 0$ o $i = 1$. El segon terme ha de ser nul quan $r = 0$. A partir d'aquesta fórmula, podem calcular primer $(q_{i,0})_{2 \leq i \leq N}$, després $(q_{i,1})_{0 \leq i \leq N-1}$, després $(q_{i,2})_{0 \leq i \leq N-2}$, etc. La probabilitat que l'epidèmia infecti finalment a n persones és $q_{0,n}$. El 1926 aquests càlculs eren bastant tediosos. Per això McKendrick es va limitar a exemples amb

Figura 16.3: Diagrama que mostra els possibles estats d'una població amb $N = 5$ (i a l'eix horitzontal, r a l'eix vertical) i les possibles transicions degudes a la infecció (fletxes horitzontals) o a la recuperació (resta de fletxes).



poblacions molt petites, per exemple una família. Amb $N = 5$ persones i $b/a = 2$, va obtenir la taula 16.1. Les probabilitats més grans corresponen al cas en què només una persona de la família s'infecta i al cas en què tota la família s'infecta.

Taula 16.1: Probabilitat que una epidèmia en una família de cinc persones n'infecti n quan $b/a = 2$.

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

El mateix article del 1926 conté també una nova formulació de problemes demogràfics considerant el temps com una variable contínua. Per a dx infinitament petit, sigui $P(x, t)$ la població amb una edat entre x i $x + dx$ a temps t . Sigi $m(x)$ la taxa de mortalitat a l'edat x . Llavors

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

per a h infinitament petit. Introduïm la definició de les derivades parcials de la funció $P(x, t)$ respecte a x i t :

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

Usant que

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

McKendrick va obtenir la següent equació en derivades parcials:

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

Una equació d'aquest tipus apareix de manera natural en problemes de dinàmica de poblacions estructurades per una variable contínua, com l'edat en demografia/ecologia (vegeu el capítol 25) o el temps des de la infecció (edat d'infecció) en epidemiologia.

El 1921, William Ogilvy Kermack havia estat nomenat responsable de la secció química del Laboratori del Real Col·legi de Metges d'Edimburg. Kermack va néixer el 1898 en un poble petit d'Escòcia. Va estudiar a la Universitat d'Aberdeen i va començar a investigar en el camp de la química orgànica a un laboratori industrial d'Oxford. Malgrat quedar-se completament cec després d'una explosió al seu laboratori d'Edimburg el 1924, va continuar la seva recerca en química amb l'ajuda de col·legues i estudiants. Kermack també va començar a col·laborar amb McKendrick en la modelització matemàtica de les epidèmies. A partir del 1927, van publicar junts una sèrie de *Contribucions a la teoria matemàtica de les epidèmies* on estudiaven models epidèmics deterministes. Sigui N la mida de la població amb N prou gran. Suposem, com a l'article del 1926, que les persones poden ser susceptibles, infectades o recuperades. Si la malaltia pot ser mortal, el tercer estat inclou també els morts. Siguin $S(t)$, $I(t)$ i $R(t)$ el nombre de persones en cadascun dels tres estats a temps t . El model és (de forma simplificada) un sistema de tres equacions diferencials ordinàries:

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

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

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

Per tant, el nombre de noves infeccions per unitat de temps és, com al model estocàstic del 1926, proporcional tant al nombre de persones susceptibles com al nombre de persones infectades. A l'inici de l'epidèmia, a l'instant $t = 0$, hi ha un cert nombre de persones infectades: $S(0) = N - I_0$, $I(0) = I_0$ i $R(0) = 0$, suposant $0 < I_0 < N$.

Tot i que el sistema (16.2)-(16.4) no té una solució tancada (solució explícita), es poden demostrar algunes propietats:

- la població total $S(t) + I(t) + R(t)$ es manté constant i igual a N al llarg del temps;
- $S(t)$, $I(t)$ i $R(t)$ es mantenen no negatius (com ha de ser ja que són poblacions);
- quan $t \rightarrow +\infty$, $S(t)$ disminueix fins a un límit $S_\infty > 0$, $I(t)$ tendeix a 0 i $R(t)$ augmenta fins a un límit $R_\infty < N$;
- A més a més, l'equació

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

dona implícitament S_∞ i per tant també la mida final de l'epidèmia $R_\infty = N - S_\infty$.

En efecte, veiem primer que

$$\frac{d}{dt}(S + I + R) = 0.$$

Per tant $S(t) + I(t) + R(t) = S(0) + I(0) + R(0) = N$. Les equacions (16.2) i (16.3) poden reescriure's com

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

D'aquí es dedueix que

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

Les equacions (16.2) i (16.4) proven que la funció $S(t)$ és decreixent i que la funció $R(t)$ és creixent (en particular, $R(t) \geq 0$). Atès que $S(t) \geq 0$ i $R(t) \leq N$, les funcions $S(t)$ i $R(t)$ tenen límit quan $t \rightarrow +\infty$. Com que $I(t) = N - S(t) - R(t)$, $I(t)$ també té límit quan $t \rightarrow +\infty$, que només pot ser zero com pot comprovar-se integrant (16.4). L'equació (16.2) també mostra que

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

Integrant entre $t = 0$ i $t = +\infty$, trobem

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

L'equació (16.3) pot reescriure's com

$$\frac{dI}{dt} = -\frac{dS}{dt} - bI.$$

Integrant entre $t = 0$ i $t = +\infty$, obtenim

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

Combinant els dos resultats, obtenim la fórmula (16.5), que prova que $S_\infty > 0$.

Quan el nombre inicial de persones infectades I_0 és petit en comparació amb la mida de la població N , és a dir $S(0) \approx N$, cosa que acostuma a passar al començament d'una epidèmia en una ciutat o regió, la fórmula (16.5) pot reescriure's utilitzant $S_\infty = N - R_\infty$ com

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

on per definició

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

L'equació (16.6) té una solució positiva només si $\mathcal{R}_0 > 1$. Així, Kermack i McKendrick arriben a la següent conclusió: l'epidèmia infecta una fracció no menyspreable de la població només si $\mathcal{R}_0 > 1$. Existeix un llindar per a la mida de la població $N^* = b/a$ per sota del qual no hi ha epidèmia.

Quan la mida de la població N està just per sobre d'aquest llindar ($N = N^* + \varepsilon$), es produeix una epidèmia d'amplitud petita. Es dedueix de (16.6) que $R_\infty \approx 2\varepsilon$. Per tant, $S_\infty \approx N^* - \varepsilon$: *l'epidèmia porta la població susceptible tant per sota del llindar N^* com inicialment estava per sobre.*

En efecte, utilitzant l'aproximació $-\log(1-x) \approx x + \frac{x^2}{2}$, l'equació (16.6) esdevé

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

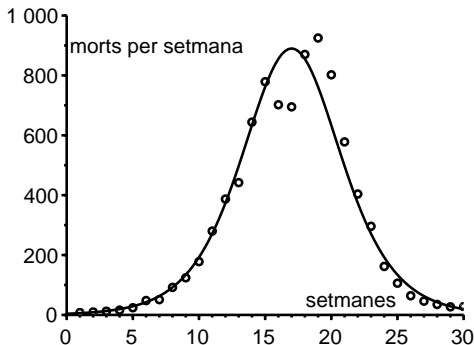
D'on,

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

Igual que al model de la malària de Ross (capítol 12), la condició $\mathcal{R}_0 > 1$ té una interpretació senzilla. Com que aN és el nombre de persones que una persona infectada infecta per unitat de temps al principi de l'epidèmia i com que $1/b$ és el període infeccios mitjà, $\mathcal{R}_0 = aN/b$ és el nombre mitjà de casos secundaris deguts a una persona infectada a l'inici de l'epidèmia (nombre reproductiu bàsic).

Per a les malalties mortals, $R(t)$ és el nombre acumulat de morts des del començament de l'epidèmia i dR/dt és el nombre de morts per unitat de temps. Kermack i McKendrick van observar que el gràfic de la funció dR/dt en el seu model matemàtic té la forma de campana que s'espera d'una corba epidèmica (figura 16.4).

Figura 16.4: La corba dR/dt en funció del temps i les dades del nombre de morts per setmana durant una epidèmia de pesta a Bombai entre el 1905 i el 1906.



Per a obtenir dR/dt , van dividir (16.2) per (16.4) obtenint $dS/dR = -aS/b$. Per tant,

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

Substituint-ho a l'equació (16.4) i utilitzant $S(t) + I(t) + R(t) = N$, van obtenir l'equació

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

que encara no pot resoldre's explícitament. No obstant, si $\frac{a}{b}R(t)$ es manté petit durant tota l'epidèmia, l'aproximació $\exp(-u) \approx 1 - u + u^2/2$ dona

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

Es tracta d'una equació de Riccati amb dues solucions constants, una positiva R_+ i una altra negativa R_- , donades per les arrels del polinomi de segon grau en R al costat dret de (16.8). Sigui $\tilde{R}(t)$ la solució exacta de (16.8) i sigui $Q(t) = \tilde{R}(t) - R_+$. Llavors $Q(t)$ satisfà una equació diferencial de Bernoulli similar a les trobades per Daniel Bernoulli i Verhulst (vegeu (4.5) i (6.1)). Per tant, es pot adaptar directament la fórmula (6.2) per a obtenir $Q(t)$. Un càlcul fàcil però tediós mostra que dQ/dt és de la forma $\alpha/\cosh^2(\beta t - \gamma)$, on α , β i γ són constants que depenen de manera complicada dels paràmetres del model. Com que $dR/dt \approx d\tilde{R}/dt = dQ/dt$, Kermack i McKendrick podien escollir (α, β, γ) per a ajustar les seves dades. Per descomptat, els ordinadors i programes moderns poden resoldre fàcilment de manera numèrica l'equació diferencial (16.7) sense passar per aquestes aproximacions.

La corba de dR/dt obtinguda d'aquesta manera s'ajustava bé a les dades del nombre de morts per setmana durant l'epidèmia de pesta a Bombai entre desembre del 1905 i juliol del 1906 (figura 16.4).

Kermack i McKendrick també van considerar el model més general on la infectivitat $a(x)$ depèn del temps x des de la infecció (l'anomenada edat d'infecció) i en el qual la taxa de recuperació $b(x)$ també depèn de x . L'equació que dona la mida final de l'epidèmia (quan el nombre inicial de casos infectats és petit) continua essent (16.6) però amb

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

El paràmetre \mathcal{R}_0 té la mateixa interpretació que en el cas anterior: és el nombre mitjà de casos secundaris provocats per una persona infectada a l'inici de l'epidèmia. Observeu la similitud entre (16.9) i la fórmula de Lotka (10.2) per a \mathcal{R}_0 en demografia: l'edat se substitueix pel temps transcorregut des de la infecció, la supervivència per la probabilitat $\exp(-\int_0^x b(y) dy)$ de seguir infectat, la fertilitat per la taxa de contacte $Na(x)$.

Kermack i McKendrick van desenvolupar altres models matemàtics d'epidèmies durant la dècada de 1930. Aquests continuen sent la base de la

majoria dels models més complexos utilitzats avui en dia en epidemiologia. El paràmetre \mathcal{R}_0 continua exercint un paper central en l'anàlisi dels models.

McKendrick es va retirar el 1941 i va morir el 1943. Entre els anys 1930 i 1933, Kermack va ser coautor d'alguns articles sobre física matemàtica amb William McCrea i Edmund Whittaker, tots dos del departament de matemàtiques de la Universitat d'Edimburg. Durant les dècades de 1930 i 1940, l'equip de químics de Kermack va intentar sintetitzar noves molècules amb activitat antipalúdica, però amb èxit limitat. El 1938, Kermack va ser coautor, juntament amb Philip Eggleton, d'un popular llibre sobre bioquímica elemental, *El material del que estem fets*. Va ser escollit membre de la *Royal Society* el 1944 i va ocupar una càtedra de bioquímica a la Universitat d'Aberdeen el 1949. Posteriorment va ser degà de la Facultat de Ciències. Es va jubilar el 1968 i va morir el 1970.

Més informació

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 (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

Capítol 17

Haldane i les mutacions (1927)

En una altra secció del seu article del 1922, Fisher va considerar el problema d'un gen mutant que pot transmetre's a un nombre aleatori de descendents amb una distribució de probabilitat determinada. El problema era formalment el mateix que el de l'extinció dels cognoms però en un context de genètica. Fisher va demostrar que si la distribució de probabilitat era una distribució de Poisson i si el gen mutant no tenia cap avantatge selectiu, llavors el gen mutant podia desaparèixer de la població molt lentament. El 1927, el biòleg britànic Haldane va aprofundir en l'estudi d'aquest model i va demostrar que la probabilitat que un gen mutant avantatjós no s'extingís era el doble del seu avantatge selectiu. També va donar un tractament més rigorós al problema de l'extinció.

John Burdon Sanderson Haldane va néixer el 1892 a Oxford, on el seu pare era professor de fisiologia a la universitat. Haldane va estudiar a l'*Eton College* i després del 1911 al *New College* de la Universitat d'Oxford. Després de centrar-se en les matemàtiques durant el seu primer any, es va dedicar a les humanitats. Els seus estudis es van veure interromputs per la Primera Guerra Mundial, durant la qual va servir a França i a l'Iraq. Després de ser ferit, va ser enviat com a instructor militar a l'Índia. El 1915 va publicar un primer article sobre els experiments genètics amb ratolins que havia iniciat abans de la guerra. El 1919 va esdevenir membre del *New College*, ensenyant fisiologia i estudiant la respiració com el seu pare. El 1923 es va incorporar al laboratori de bioquímica de F. G. Hopkins¹ a la Universitat de Cambridge, on es va centrar en la cinètica dels enzims. També va publicar una novel·la de ciència-ficció, *Daedalus o la ciència i el futur* (1923), i un assaig titulat *Callinicus, una defensa de la guerra química* (1925). Entre els anys 1924 i 1934, va escriure una sèrie de deu articles titulats *Una teoria matemàtica de la selecció natural i artificial*.

En el cinquè article de la sèrie, publicat el 1927, Haldane va considerar de nou un model genètic que Fisher havia estudiat el 1922, un model centrat en les mutacions. Fisher havia estudiat la probabilitat que un gen mutant

¹Frederick Gowland Hopkins, que va rebre el Premi Nobel de Fisiologia o Medicina el 1929 pels seus treballs sobre les vitamines.

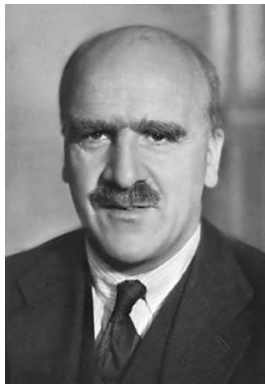


Figura 17.1:
Haldane (1892–1964)

envaís una població o desaparegués. Aquest problema és formalment el mateix que el de Bienaymé, Galton i Watson sobre l'extinció dels cognoms. Però Fisher no va fer cap referència a aquests treballs, encara que és possible que hagués llegit l'article de Galton i Watson reproduït en l'apèndix del llibre de Galton del 1889 *L'herència natural*. Com al capítol 9, denotarem per p_k la probabilitat que un gen es transmeti a k descendents a la primera generació ($k \geq 0$). Fisher va considerar també la funció generatriu $f(x) = p_0 + p_1 x + p_2 x^2 + \dots + p_k x^k + \dots$, però no va fixar cap límit superior per a k : la suma pot incloure un nombre infinit de termes. Es va adonar que, partint d'un individu amb el gen mutant a la generació 0, la probabilitat que aquest gen estigui en k individus és el coeficient de x^k en $f_1(x) = f(x)$ per a la generació 1, en $f_2(x) = f(f(x))$ per a la generació 2, en $f_3(x) = f(f(f(x)))$ per a la generació 3, etc. D'aquesta manera, queda clar que es compleix la següent equació.

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

Aquesta equació és molt més pràctica que l'equació $f_n(x) = f_{n-1}(f(x))$ obtinguda per Watson. En particular, es dedueix de (17.1) que la probabilitat d'extinció en n generacions $x_n = f_n(0)$ satisfà la fórmula d'iteració $x_n = f(x_{n-1})$, com ja havia notat Bienaymé.

Com a exemple, Fisher va considerar el cas d'una planta amb un gen mutant que pot produir N llavors, tenint cada llavor una probabilitat q de sobreviure per a produir una nova planta. La probabilitat p_k d'obtenir k descendents amb el gen mutant és binomial:

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

per a tot $0 \leq k \leq N$ i $p_k = 0$ per a $k > N$. La funció generatriu és llavors $f(x) = (1 - q + qx)^N$. Sigui $\mathcal{R}_0 = Nq$ el nombre mitjà de llavors que sobreviuen per a produir una nova planta. Quan N és gran i q és petit, es té

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

La distribució de probabilitat (p_k) tendeix a

$$e^{-\mathcal{R}_0} \frac{(\mathcal{R}_0)^k}{k!},$$

que s'anomena distribució de Poisson. Fisher va calcular aleshores la probabilitat d'extinció en n generacions, utilitzant $x_0 = 0$, $x_n \approx e^{\mathcal{R}_0(x_{n-1}-1)}$ i els valors numèrics $N = 80$ i $q = 1/80$. En aquest cas, $\mathcal{R}_0 = Nq = 1$. Un càlcul tediós dona $x_{100} \approx 0.98$: un gen mutant sense avantatge selectiu ($\mathcal{R}_0 = 1$) desapareix molt lentament. Encara hi ha una probabilitat del 2 % que el gen sigui present a la població després de 100 generacions. El 1922 Fisher no va continuar amb l'estudi d'aquest model.

Continuant amb el treball de Fisher, Haldane va observar per primera vegada en el seu article del 1927 que, per a qualsevol distribució de probabilitat (p_k) tal que $p_0 > 0$, l'equació $x = f(x)$ té exactament dues arrels a l'interval $(0, 1]$ quan el nombre mitjà de descendents que porten el gen mutant \mathcal{R}_0 és estrictament major que 1, és a dir, quan el gen mutant té un avantatge selectiu. A més a més, la probabilitat d'extinció x_∞ , que és el límit de x_n quan $n \rightarrow +\infty$, és la menor de les dues arrels de $x = f(x)$: el gen té una probabilitat no nul·la d'establir-se a la població. A diferència de Bienaymé i Cournot, Haldane va demostrar aquesta conclusió.

En efecte, $f'(x) \geq 0$ i $f''(x) \geq 0$ a l'interval $[0, 1]$. En altres paraules, la funció $f(x)$ és no decreixent i convexa. Les hipòtesis $f(0) = p_0 > 0$ i $f'(1) = \mathcal{R}_0 = p_1 + 2p_2 + 3p_3 + \dots > 1$ impliquen que l'equació $f(x) = x$ té exactament dues solucions a l'interval $(0, 1]$: $x = 1$ i x^* tal que $0 < x^* < 1$. Haldane va fer referència llavors a un article de Gabriel Koenigs del 1883, que demostrava que si $x_n = f(x_{n-1})$ i $x_n \rightarrow x_\infty$, aleshores $x_\infty = f(x_\infty)$ i $|f'(x_\infty)| \leq 1$. Quan $f'(1) > 1$, l'única possibilitat és que $x_\infty = x^*$.

Per al cas d'una distribució de Poisson amb $f(x) = e^{\mathcal{R}_0(x-1)}$ i \mathcal{R}_0 lleugerament més gran que 1, la probabilitat d'extinció x_∞ és molt pròxima a 1. L'equació $f(x_\infty) = x_\infty$ és llavors equivalent a

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

D'on es dedueix que

$$1 - x_{\infty} \approx 2(\mathcal{R}_0 - 1).$$

Haldane va concloure que la probabilitat que el gen mutant no s'extingeixi és el doble del seu avantatge selectiu $\mathcal{R}_0 - 1$. Sense citar Haldane, Fisher va prendre com a exemple en el seu llibre del 1930 el cas en que $\mathcal{R}_0 = 1,01$, la qual cosa dona una probabilitat del 2 % que el gen mutant no s'extingeixi.

Haldane va ser nomenat membre de la *Royal Society* el 1932. Va deixar Cambridge per a esdevenir professor de genètica i posteriorment de biometria a l'*University College* de Londres. En aquell moment estava particularment interessat en la genètica humana: estimació de les taxes de mutació, mapes genètics dels cromosomes, etc. A més a més dels seus llibres científics (*Biologia animal* el 1927 amb Julian Huxley, *Enzims* el 1930, *Les causes de l'evolució* el 1932 i *La bioquímica de la genètica* el 1954), va publicar un gran nombre d'articles sobre ciència a la premsa (per exemple sobre l'origen de la vida) i alguns assajos (*La desigualtat de l'home* el 1932, *La filosofia d'un biòleg* el 1935, *La filosofia marxista i les ciències* el 1938, *Herència i política* el 1938 i *Avanços de la ciència* el 1947). Després de diverses visites a Espanya durant la guerra civil, va intentar convèncer el seu propi país que construís refugis contra els bombardejos aeris. Durant la Segona Guerra Mundial, va treballar en els problemes de respiració als submarins. Membre del partit comunista des del 1942, va dimitir el 1950 a causa del rebuig oficial de la genètica mendeliana a l'URSS per influència de Lysenko. El 1957 es va instal·lar a l'Índia, on va continuar la seva recerca, primer a l'Institut Indi d'Estadística de Calcuta i després a Bhubaneswar. Després d'obtenir la nacionalitat índia, va morir el 1964.

Més informació

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)

Capítol 18

Erlang i Steffensen sobre el problema de l'extinció (1929–1933)

L'any 1929, l'enginyer danès Erlang va tornar a plantejar-se el problema de l'extinció dels cognoms. El seu compatriota, l'estadístic Steffensen, va elaborar una solució completa del problema. Va demostrar, en particular, que l'esperança del nombre de descendents a cada generació creix exponencialment, construint així un pont entre els models de població estocàstics i deterministes.

Agner Krarup Erlang va néixer el 1878 a Lønborg, Dinamarca. El seu pare era mestre d'escola. Entre els anys 1896 i 1901, el jove Erlang va estudiar matemàtiques, física i química a la Universitat de Copenhaguen. A continuació, va fer classes durant diversos anys en instituts d'ensenyament secundari, al mateix temps que s'interessava per les matemàtiques, especialment per la teoria de la probabilitat. Va conèixer Jensen, enginyer en cap de la Companyia Telefònica de Copenhaguen i matemàtic aficionat, que el va convèncer el 1908 perquè s'incorporés al nou laboratori de recerca de l'empresa. Erlang va començar a publicar articles sobre les aplicacions de la teoria de la probabilitat a la gestió de les trucades telefòniques. El 1917 va descobrir una fórmula per als temps d'espera, que va ser ràpidament utilitzada per les companyies telefòniques de tot el món. Els seus articles, publicats primer en danès, es van traduir després a altres idiomes.

El 1929 Erlang es va interessar pel mateix problema d'extinció que Bi-enaymé, Galton i Watson havien estudiat abans que ell per als cognoms i que Fisher i Haldane havien estudiat per als gens mutants. Igual que els seus predecessors, no coneixia tots els treballs que s'havien publicat. Anomenant de nou p_k a la probabilitat que un individu tingui k descendents, va observar que la probabilitat x_n d'extinció en n generacions satisfà $x_n = p_0 + p_1 x_{n-1} + p_2 (x_{n-1})^2 + \dots = f(x_{n-1})$ amb $x_0 = 0$. Va observar també que la probabilitat d'extinció global x_∞ , que és el límit de x_n quan $n \rightarrow +\infty$, és una solució de l'equació $x_\infty = f(x_\infty)$. Es va adonar que $x = 1$ era sempre una solució i que existia una altra solució entre 0 i 1 quan el nombre mitjà de descendents $\mathcal{R}_0 = f'(1)$ és major que 1. Però sembla que no va poder esbrinar quina d'aquestes dues solucions era la correcta. Igual que Galton, va presen-

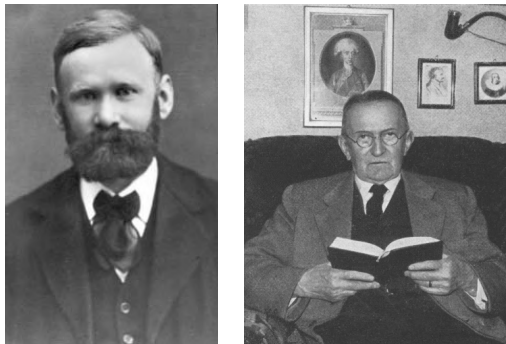


Figura 18.1: Erlang (1878–1929) i Steffensen (1873–1961)

tar el problema l'any 1929 a una revista danesa de matemàtiques, *Matematisk Tidsskrift*:

«Pregunta 15. Si la probabilitat que un individu tingui k fills és p_k , on $p_0 + p_1 + p_2 + \dots = 1$, trobi la probabilitat que la seva família s'extingeixi.»

Desafortunadament, Erlang va morir aquell mateix any 1929 a l'edat de 51 anys. De fet, va morir sense fills¹. Un professor de matemàtiques actuàries de la Universitat de Copenhaguen, Johan Frederik Steffensen, va reprendre el problema plantejat per Erlang. El 1930 va publicar la seva solució a la mateixa revista danesa: la probabilitat d'extinció x_∞ és sempre l'arrel més petita de l'equació $x = f(x)$ a l'interval tancat $[0, 1]$, com ja havien observat Bienaymé i Haldane. La prova de Steffensen és la que es troba als llibres de text moderns.

En efecte, vam veure que la probabilitat d'extinció x_∞ és una solució de $x = f(x)$ a l'interval tancat $[0, 1]$. Sigui x^* la solució més petita. Per definició, $x^* \leq x_\infty$. Steffensen va observar en primer lloc que $x^* = f(x^*) \geq p_0 = x_1$. Suposem per inducció que $x^* \geq x_n$. Llavors $x^* = f(x^*) \geq f(x_n) = x_{n+1}$ ja que la funció $f(x)$ és creixent. De manera que $x^* \geq x_n$ per a tot n . Prenent el límit, $x^* \geq x_\infty$. Aleshores $x_\infty = x^*$.

Steffensen va donar també una explicació més formal de perquè $x = 1$ és l'única arrel de $x = f(x)$ quan el nombre mitjà de descendents $\mathcal{R}_0 = f'(1)$ és

¹En la seva memòria, el Comitè Consultiu Internacional de Telefonia va decidir el 1946 anomenar «erlang» a la unitat de mesura de la intensitat del trànsit telefònic. També s'anomena «Erlang» un llenguatge de programació de l'empresa Ericsson.

menor o igual que 1 (figura 18.2a) i perquè només hi ha una altra arrel diferent de $x = 1$ en el cas en què $\mathcal{R}_0 > 1$ (figura 18.2b). Observeu que $\mathcal{R}_0 = f'(1)$ és el pendent de la funció $f(x)$ a $x = 1$.

Steffensen va observar que per a qualsevol arrel de $x = f(x)$,

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

Suposant que $x \neq 1$ i dividint per $1 - x$, obtenim

$$1 = p_1 + p_2(1+x) + p_3(1+x+x^2) + \dots \quad (18.1)$$

Quan x augmenta de 0 a 1, el costat dret de l'equació (18.1) augmenta de $1 - p_0$ a $\mathcal{R}_0 = f'(1)$. Si $\mathcal{R}_0 < 1$, llavors l'equació (18.1) no té solució. Si $\mathcal{R}_0 \geq 1$ i si exclouem el cas trivial on $p_1 = 1$, llavors el costat dret de l'equació (18.1) és una funció estrictament creixent de x . En cas contrari no existiria $k \geq 2$ tal que $p_k \neq 0$ i \mathcal{R}_0 seria igual a $p_1 < 1$. En conclusió, (18.1) té una i només una solució a l'interval $[0, 1]$ quan $\mathcal{R}_0 \geq 1$.

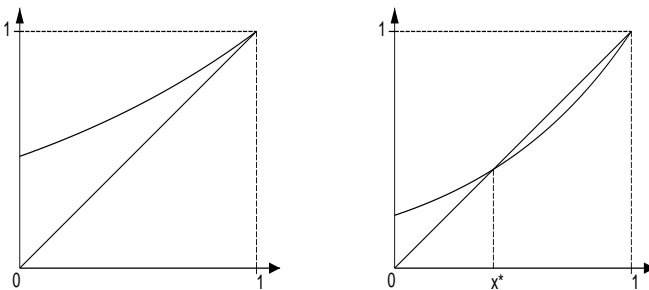


Figura 18.2: Gràfica de les funcions $y = x$ i $y = f(x)$ a l'exemple del capítol 17, $f(x) = e^{\mathcal{R}_0(x-1)}$, amb $\mathcal{R}_0 = 0,75 < 1$ (esquerra) o $\mathcal{R}_0 = 1,5 > 1$ (dreta).

Steffensen, que també era president de la Societat Actuarial Danesa i de la Societat Matemàtica Danesa, va ser convidat a la Universitat de Londres el 1930. El seu col·lega britànic W. P. Elderton li va parlar dels treballs de Galton i Watson. El 1933, Steffensen va publicar un nou article als annals de l'Institut Henri Poincaré, on havia impartit una conferència el 1931. Va resumir els resultats del seu article en danès i els va comparar amb els de Watson. També

va demostrar que l'esperança matemàtica del nombre de descendents a la generació n és igual a $(\mathcal{R}_0)^n$.

En efecte, sigui $p_{k,n}$ la probabilitat que hi hagi k descendents a la generació n , partint d'un sol individu a la generació 0. Al seu article del 1930, Steffensen havia observat, com els seus predecessors, que la funció generatriu $f_n(x) = \sum_{k=0}^{+\infty} p_{k,n} x^k$ relativa a la generació n satisfà $f_1(x) = f(x)$ i $f_n(x) = f(f_{n-1}(x))$. Sigui M_n l'esperança del nombre de descendents a la generació n . Llavors $M_n = \sum_{k=1}^{+\infty} k p_{k,n} = f'_n(1)$. Obtenim $f'_n(x) = f'(f_{n-1}(x)) \times f'_{n-1}(x)$. Així, $M_n = f'_n(1) = f'(f_{n-1}(1)) \times f'_{n-1}(1) = f'(1) \times M_{n-1} = \mathcal{R}_0 \times M_{n-1}$. Com que $M_1 = f'_1(1) = f'(1) = \mathcal{R}_0$, es dedueix que $M_n = (\mathcal{R}_0)^n$ per a tot n .

Per tant, el nombre esperat de descendents augmenta o disminueix geomètricament depenent de si \mathcal{R}_0 és major o menor que 1. El nombre esperat de descendents es comporta com als models deterministes de creixement de població considerats per Euler, Malthus, etc. No obstant, fins i tot quan $\mathcal{R}_0 > 1$, existeix una probabilitat no nul·la x_∞ que la família s'extingeixi. Aquesta possibilitat no es dona en els models deterministes.

El procés estocàstic estudiat per Steffensen i els seus predecessors continua sent l'element bàsic de molts models més realistes de dinàmica de poblacions. Farem menció a aquest problema per última vegada al capítol 20. Pel que fa a Steffensen, va continuar sent professor de la Universitat de Copenhaguen fins al 1943 i va morir l'any 1961.

Més informació

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*, págs. 328–330. Springer (2001)
5. Nørdlund, N.E.: Johan Frederik Steffensen in memoriam. *Nordisk Mat. Tidsskr.* 10, 105–107 (1962)
6. Ogborn, M.E.: Johan Frederik Steffensen, 1873–1961. *J. R. Stat. Soc. Ser. A* 125, 672–673 (1962)
7. Steffensen, J.F.: Om Sandssynligheden for at Afkommet uddør. *Mat. Tidsskr. B*, 19–23 (1930) → Guttorp (1995)
8. Steffensen, J.F.: Deux problèmes du calcul des probabilités. *Ann. Inst. Henri Poincaré* 3, 319–344 (1933) archive.numdam.org

Capítol 19

Wright i la deriva genètica aleatòria (1931)

L'any 1931, el biòleg estatunidenc Sewall Wright va desenvolupar l'estudi d'un model estocàstic en genètica de poblacions, que es basa en les mateixes hipòtesis que la llei de Hardy-Weinberg, tret que la població no se suposa infinitament gran. Les freqüències dels genotips ja no són constants, de fet, un dels dos al·lels desapareixerà, però potser després de molt de temps. La interpretació d'aquest model va continuar sent objecte de disputa entre Wright i Fisher, estimant aquest últim que la selecció natural exerceix un paper més important en l'evolució que l'estocasticitat.

Sewall Wright va néixer a Massachusetts el 1889. Va realitzar els seus estudis universitaris en una petita facultat d'Illinois on el seu pare ensenyava economia. Després d'un màster en biologia a la Universitat d'Illinois a Urbana i una escola d'estiu al Laboratori de Cold Spring Harbor, Wright va fer un doctorat a la Universitat de Harvard sobre l'herència del color del pelatge en els cobais. Entre els anys 1915 i 1925, va continuar treballant en experiments d'endogàmia amb cobais a la Divisió de Ramaderia del Departament d'Agricultura dels Estats Units a Washington. Va desenvolupar el «mètode dels coeficients de trajectòria» per a analitzar aquests experiments. Més tard es va incorporar al departament de zoologia de la Universitat de Chicago.

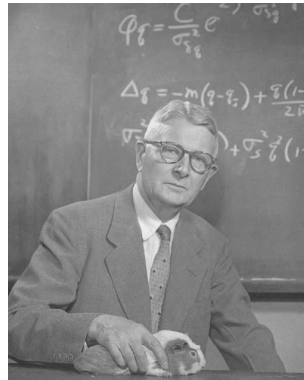


Figura 19.1:
Wright (1889–1988)

Influït per l'article de Fisher del 1922 sobre genètica de poblacions (vegeu el capítol 14), Wright va escriure el 1925 un llarg article titulat *Evolució en poblacions mendelianes*, que finalment es va publicar el 1931. Va estudiar en particular un model matemàtic que apareixia també implícitament al llibre de Fisher del 1930 titulat *La teoria genètica de la selecció natural*. Igual que a la llei de Hardy-Weinberg, aquest model considera el cas en el qual només hi ha dos al·lells possibles A i a per a un locus, però se suposa que la població no és infinitament gran. L'objectiu és veure si l'eliminació d'aquesta hipòtesi té alguna influència en la composició genètica de la població. Anomenarem N el nombre total d'individus, que se suposa que és el mateix a totes les generacions. Cada individu té dos al·lells. Per tant, hi ha un total de $2N$ al·lells en la població a cada generació. El model suposa també que l'aparellament es produeix a l'atzar. Si hi ha i al·lells A i $2N - i$ al·lells a a la generació n , llavors un al·lel triat a l'atzar entre els individus de la generació $n + 1$ serà A amb probabilitat $\frac{i}{2N}$ i a amb probabilitat $1 - \frac{i}{2N}$. El nombre d'al·lells A a la generació $n + 1$ serà per tant igual a j amb probabilitat¹

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

on el coeficient binomial $\binom{2N}{j}$ és igual a $\frac{(2N)!}{j!(2N-j)!}$. Sigui X_n el nombre d'al·lells A a la generació n : és una variable aleatòria (figura 19.2).

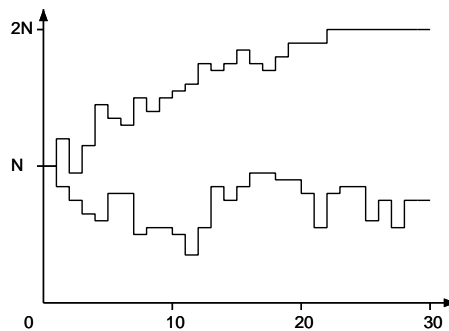


Figura 19.2: Dues simulacions que mostren les variacions del nombre X_n d'al·lells A durant 30 generacions si $N = 20$ i $X_0 = 10$.

Es pot demostrar que l'esperança de X_{n+1} sabent que $X_n = i$ és igual a i : això recorda la llei de Hardy-Weinberg, on la freqüència de l'al·lel A es mantenia constant a través de les generacions.

¹Aquesta formulació en el llenguatge de les cadenes de Markov es deu a Malécot (1944).

En efecte, considerem la funció generatriu

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

L'esperança de X_{n+1} sabent que $X_n = i$ és llavors

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

No obstant, és possible en aquest model que, partint d'una condició inicial $X_0 = i$ amb $0 < i < 2N$, l'esdeveniment $X_n = 0$ es produeixi per atzar després d'un cert nombre de generacions. En aquest cas, tots els al·lells serien del tipus a i X_n continuaria sent igual a 0 a totes les generacions futures. El mateix passaria amb l'al·lel A si $X_n = 2N$ després d'un cert nombre de generacions. En resum, quan se suposa que la població és infinitament gran com al model de Hardy-Weinberg, els dos al·lells no poden desaparèixer perquè les seves freqüències es mantenen constants. Quan es té en compte la mida finita de les poblacions, com al model de Fisher-Wright, les freqüències dels dos al·lells fluctuen i un dels al·lells pot desaparèixer (i, de fet, desapareixerà).

Partint de $X_0 = i$, es pot calcular fàcilment la probabilitat Q_i que la població quedi fixada en l'estat $X = 0$. En efecte, Q_i ha de satisfer les «condicions de frontera»

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

A més a més,

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

ja que $p_{i,j} Q_j$ és la probabilitat de quedar-se en l'estat $X = 0$ partint de $X_0 = i$ i passant per $X_1 = j$. Com que

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

veiem fent servir (19.2) que

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

és la solució del sistema (19.3)-(19.4). Per tant, la probabilitat que, partint d' i al·lells de tipus A en una població de mida N , el sistema evolucioni cap a una

població que contingui només l'al·lel a és igual a $1 - \frac{i}{2N}$. De la mateixa manera, la probabilitat que evolucioni cap a una població que contingui només l'al·lel A és $\frac{i}{2N}$.

Wright va aconseguir demostrar que el nombre de generacions que transcorren abans de la fixació en un dels dos estats extrems és de l'ordre de $2N$ generacions. Per a poblacions de diversos milions d'individus, aquest temps seria tan llarg que les freqüències dels al·lells podrien considerar-se gairebé constants, com a la llei de Hardy-Weinberg.

En efecte, suposem que hi ha i_0 al·lells de tipus A en la població a la generació 0. Sigui $u_i^{(n)}$ la probabilitat que hi hagi i al·lells de tipus A a la població a la generació n (figura 19.3). Llavors

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

per a tot $j = 0, \dots, 2N$. Ja hem vist que, quan $n \rightarrow +\infty$,

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

per a tot $0 < i < 2N$. Wright va observar que si $u_i^{(n)} = v$ per a tot $i = 1, \dots, 2N - 1$, llavors

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

per a tot $1 < j < 2N$ perquè $p_{0,j} = p_{2N,j} = 0$. Quan N és prou gran,

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

el valor de la integral s'obté mitjançant successives integracions per parts. Combinant (19.5) i (19.6), arribem finalment, per a $0 < j < 2N$, a que

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

Així que les probabilitats $u_j^{(n)}$ per a tots els $0 < j < 2N$ disminueixen a un ritme d'aproximadament $1/2N$ per generació. Aquest ritme és molt lent si N és gran. Gairebé no hi ha disminució si, per exemple, N és de l'ordre de magnitud de milions.

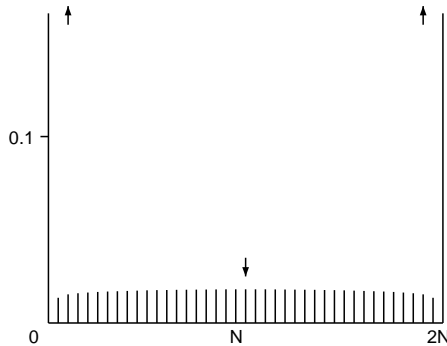


Figura 19.3: Probabilitat que hi hagi i al·lels A a la població ($i = 0, \dots, 2N$ en l'eix horitzontal) després de 30 generacions si $N = 20$ i $X_0 = 10$.

El 1922, Fisher ja havia intentat estimar aquesta taxa de fixació ($1/2N$), però se li va passar per alt un factor 2. En qualsevol cas, els dos científics discrepaven sobre la mida típica N de les poblacions reproductores. Per a la teoria de l'evolució, el treball de Wright suggeria que la deriva genètica aleatòria en una població petita podia ser un mecanisme per a l'origen de les espècies. En efecte, els biòlegs que treballaven en la classificació de les espècies havien observat que les diferències entre espècies o subespècies sovint no tenien una explicació aparent en termes de selecció natural. Aquesta idea va ser fortament criticada durant els anys 40 i 50 per Fisher i el seu col·lega E. B. Ford, que pensaven que la deriva genètica aleatòria era insignificant en comparació amb la selecció natural. Es van referir en particular al seu estudi de les fluctuacions de les freqüències gèniques en una població petita i aïllada d'arnes (*Panaxia dominula*) prop d'Oxford, on els tres genotips per a un determinat gen (homozigot comú, heterozigot i homozigot rar) podien distingir-se a simple vista. Una altra famosa controvèrsia sobre la influència respectiva de la selecció natural i de la deriva aleatòria tenia a veure amb els caragols del gènere *Cepaea*. Models més realistes d'evolució combinen ara

la deriva aleatòria, la selecció, la mutació, la migració, l'aparellament no aleatori, etc. El paper de la deriva aleatòria va ser reafirmat posteriorment pel científic japonès Motoo Kimura amb la seva «teoria neutralista de l'evolució molecular». Una altra conseqüència va ser el desenvolupament de la teoria de la coalescència (introduïda per John Kingman el 1982), que rastreja l'ascendència dels gens endarrere en el temps fins al punt en què tenen un únic ancestre comú.

Wright va esdevenir membre de l'Acadèmia Nacional de Ciències el 1934. Va treballar durant molts anys amb Theodosius Dobzhanski en la genètica de poblacions naturals de mosques (*Drosophila pseudoobscura*) a la regió de la Vall de la Mort. Es va retirar de la Universitat de Chicago el 1955, però va continuar durant cinc anys més com a professor a la Universitat de Wisconsin-Madison. Entre els anys 1968 i 1978, va publicar un tractat de quatre volums que resumia la seva recerca sobre *Evolució i genètica de poblacions*. Va rebre el Premi Balzan el 1984 i va morir el 1988 a l'edat de 98 anys.

Més informació

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. Malécot, G.: Sur un problème de probabilités en chaîne que pose la génétique. *C. R. Acad. Sci. Paris* 219, 379–381 (1944)
5. Provine, W.B.: *Sewall Wright and Evolutionary Biology*. University of Chicago Press (1989)
6. Wright, S.: Evolution in Mendelian populations. *Genetics* 16, 97–159 (1931) www.esp.org
7. Wright, S.: *Evolution and the Genetics of Populations*, Vol. 2, *Theory of Gene Frequencies*. University of Chicago Press (1969)

Capítol 20

La difusió dels gens (1937)

L'any 1937, Ronald Fisher i tres matemàtics russos, Kolmogórov, Petrovskij i Piskunov, van estudiar per separat una equació en derivades parcials per a la difusió geogràfica d'un gen avantatjós. Van demostrar que la freqüència del gen es comporta com una ona que viatja a una velocitat definida en funció de l'avantatge del gen i d'un coeficient de difusió. Els seus treballs van ser el punt de partida de la teoria de les equacions de reacció-difusió.

El 1937 es van publicar dos articles que introduïen un nou enfoc a l'estudi de l'heterogeneïtat espacial en dinàmica de poblacions. Fisher va ser l'autor del primer article, titulat *L'ona d'avanç dels gens avantatjosos*, que va aparèixer als *Annals of Eugenics*. Va estudiar la propagació espacial d'un gen favorable en una població. Com a simplificació, va considerar un espai reduït a només una dimensió i va anomenar $u(x,t)$ la proporció de la població situada en el punt x a temps t que posseeix el gen favorable. Així doncs, $0 \leq u(x,t) \leq 1$. Per a incloure la selecció natural, va utilitzar l'equació (14.6) amb una variable temporal contínua

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

on a és un paràmetre positiu. Per a un valor donat de x , reconeixem l'equació logística de Verhulst (vegeu el capítol 6) amb una solució $u(x,t)$ que tendeix a 1 quan $t \rightarrow +\infty$. A més a més, Fisher va suposar que la descendència d'un individu situat al punt x amb el gen favorable no es queda al mateix punt sinó que es dispersa aleatòriament en un entorn de x . Per analogia amb la física, va argumentar que cal afegir un terme de difusió (amb coeficient D) a l'equació de $u(x,t)$, la qual cosa porta a l'equació en derivades parcials

$$\frac{\partial u}{\partial t} = au(1-u) + D \frac{\partial^2 u}{\partial x^2}. \quad (20.1)$$

Quan el coeficient de selecció a és zero, es redueix a l'equació de difusió introduïda per Fourier en la seva teoria de la calor i posteriorment utilitzada per Fick per a la difusió de partícules físiques. El 1904, Ronald Ross havia

començat a considerar la dispersió aleatòria en dinàmica de poblacions. Es preguntava com disminueix la densitat dels mosquits a mesura que augmenta la distància d'un lloc de cria. El problema havia atret l'atenció de Karl Pearson i Lord Rayleigh. El 1937, la literatura científica sobre les equacions de difusió havia augmentat considerablement, en particular després dels treballs d'Einstein sobre el moviment brownià.

Fisher va demostrar que existeixen solucions de l'equació (20.1) de la forma $u(x, t) = U(x + vt)$ que satisfan les tres condicions

$$0 \leq u(x, t) \leq 1, \quad u(x, t) \xrightarrow{x \rightarrow -\infty} 0, \quad u(x, t) \xrightarrow{x \rightarrow +\infty} 1,$$

sempre que $v \geq v^*$ on

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

Aquestes solucions connecten l'estat estacionari $u = 1$ de la població amb el gen favorable i l'estat estacionari $u = 0$ de la població sense aquest gen. Representen ones que es propaguen a velocitat v en la direcció de valors decreixents de x . En efecte, $u(x - vT, t + T) = u(x, t)$: la part de l'ona que estava a la posició x a temps t es desplaça a la posició $x - vT$ a temps $t + T$.

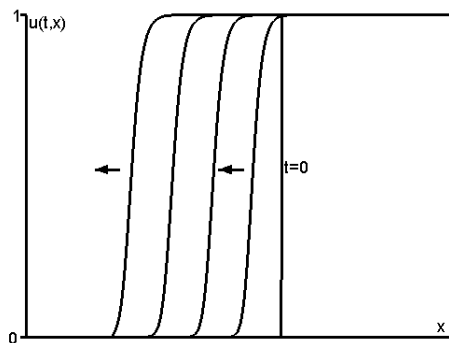


Figura 20.1: Propagació de dreta a esquerra d'un gen favorable a velocitat v^* . La freqüència del gen $u(t, x)$ a $t = 0$ és una funció esglaonada.

En efecte, fixant $z = x + vt$, Fisher va observar que si $u(x, t) = U(z)$, llavors

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

Si u és una solució de l'equació (20.1), llavors

$$vU'(z) = aU(z)(1 - U(z)) + DU''(z). \quad (20.2)$$

Quan u està prop de 0, és a dir, quan $z \rightarrow -\infty$, Fisher esperava que $U(z) \rightarrow 0$ i $U'(z) \rightarrow 0$. Anomenant k al límit de $U'(z)/U(z)$ quan $z \rightarrow -\infty$, sabem per la regla de L'Hôpital que $U''(z)/U'(z)$ també tendeix a k . Per tant,

$$U''(z)/U(z) = [U''(z)/U'(z)] \times [U'(z)/U(z)]$$

tendeix a k^2 . Dividint l'equació (20.2) per $U(z)$ i fent tendir z cap a $-\infty$, obtenim una equació de segon grau

$$Dk^2 - vk + a = 0.$$

Però k ha de ser un nombre real. Per tant el discriminant d'aquesta equació ha de ser no negatiu: $v^2 - 4aD \geq 0$, o $v \geq 2\sqrt{aD} = v^*$. Així, $v \geq v^*$ és una condició necessària per a l'existència d'una ona que es propagui a velocitat v . També és una condició suficient, com s'explica a continuació.

Fisher va observar que només se selecciona l'ona que es propaga exactament a velocitat v^* per a una classe gran de condicions inicials, per exemple, per a la funció esglaonada: $u(x,0) = 0$ per a $x < 0$, $u(x,0) = 1$ per a $x \geq 0$. La figura 20.1 mostra com aquesta condició inicial discontinua es converteix progressivament en una ona suau que es propaga en la direcció de decreixement de x a velocitat v^* .

Aquell mateix any 1937, i independentment del treball de Fisher, Andrej Nikolaevich Kolmogórov, Ivan Georgievich Petrovskij i Nikolaj Semënovich Piskunov van estudiar el mateix problema de propagació d'un gen dominant.

Kolmogórov va néixer el 1903 a Tambov, Rússia. Durant els seus estudis de matemàtiques a la Universitat Estatal de Moscou, va fer treballs importants sobre sèries trigonomètriques. Va esdevenir investigador de l'Institut de Matemàtiques i Mecànica el 1929 i professor universitari el 1931. Va treballar en processos estocàstics i la seva relació amb les equacions diferencials i les equacions en derivades parcials. El 1933 va publicar un tractat que establí les bases modernes de la teoria de la probabilitat. Els seus interessos de recerca incloïen la topologia, la teoria de l'aproximació, les cadenes de Markov, el moviment brownià i també aplicacions a problemes biològics. El 1935 va publicar un article sobre genètica en el que discutia els resultats de Hardy, Fisher i Wright. El 1936 va publicar un article sobre una generalització del

sistema de Lotka-Volterra.

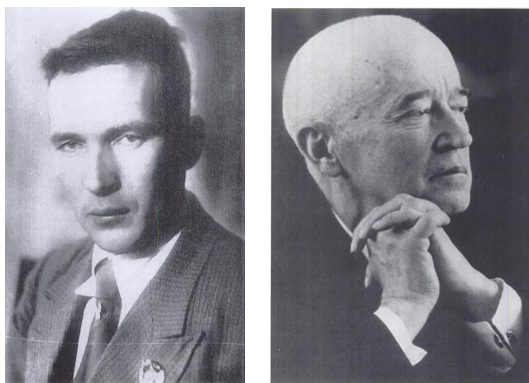


Figura 20.2: Kolmogórov (1903–1987) i Petrovskij (1901–1973)

Petrovskij va néixer el 1901 a Sevsk, Rússia. Va estudiar matemàtiques a la Universitat Estatal de Moscou, on va arribar a ser catedràtic el 1933. Va treballar principalment en la teoria de les equacions en derivades parcials i en la topologia de les corbes algebraïques reals, encara que també va escriure alguns articles sobre equacions diferencials ordinàries i sobre la teoria de la probabilitat. Piskunov, nascut el 1908, també va ser estudiant de matemàtiques a la Universitat Estatal de Moscou.

Durant els anys 30, Kolmogórov va mantenir contactes amb A. S. Serebrovskij, pioner de la genètica de poblacions a Moscou. La defensa de la genètica mendeliana a l'URSS era cada vegada més perillosa per culpa de l'ascens de Lysenko, un agrònom que havia aconseguit convèncer Stalin que la genètica mendeliana era una mera «pseudociència burgesa». El Setè Congrés Internacional de Genètica, previst inicialment pel 1937 a Moscou, va ser cancel·lat. Molts genetistes soviètics van ser executats o enviats a camps de treball.

Al seu article de l'any 1937 titulat *Un estudi de l'equació de difusió amb augment de la quantitat de substància i la seva aplicació a un problema biològic*, que va ser publicat al Butlletí de la Universitat Estatal de Moscou, Kolmogórov, Petrovskij i Piskunov van utilitzar, no obstant, un model matemàtic basat en la genètica mendeliana. El seu model era una equació en derivades parcials de la forma

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

on $u(x, t)$ és de nou la freqüència del gen favorable al punt x i a temps t . Se suposa que la funció $f(u)$ satisfà algunes condicions: $f(0) = f(1) = 0$, $f(u) > 0$ si $0 < u < 1$, $f'(0) > 0$ i $f'(u) < f'(0)$ si $0 < u \leq 1$. Els autors van obtenir un resultat anàleg al de Fisher però amb una demostració més rigorosa: si la condició inicial és tal que $0 \leq u(x, 0) \leq 1$, $u(x, 0) = 0$ per a tot $x < x_1$ i $u(x, 0) = 1$ per a tot $x > x_2 \geq x_1$, llavors el gen es propaga a velocitat $v^* = 2\sqrt{f'(0)D}$.

Buscant una solució $u(x, t) = U(z)$ del tipus $z = x + vt$ s'obté la generalització òbvia de l'equació (20.2) $vU'(z) = f(U(z)) + DU''(z)$. Aquesta equació diferencial de segon ordre pot reescriure's com un sistema d'equacions diferencials de primer ordre

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

Recordem que $U(z)$ ha de ser tal que $U(z) \rightarrow 0$ quan $z \rightarrow -\infty$ i $U(z) \rightarrow 1$ quan $z \rightarrow +\infty$. Prop de l'estat estacionari ($U = 0, p = 0$) del sistema (20.4), tenim $f(U) \approx f'(0)U$. De manera que (20.4) pot ser aproximat pel sistema lineal

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

Buscant solucions exponencials de la forma $U(z) = U_0 e^{kz}$ i $p(z) = p_0 e^{kz}$ s'obté l'equació característica

$$Dk^2 - vk + f'(0) = 0,$$

com a l'article de Fisher. De nou, k ha de ser real (sinó, u oscil·laria i prendria valors negatius). Així doncs, $v \geq 2\sqrt{f'(0)D} = v^*$. Les dues arrels per a k són llavors reals i positives. Si $v > v^*$, les dues arrels són diferents i l'estat estacionari ($U = 0, p = 0$) és un node inestable. Si $v = v^*$, les dues arrels són idèntiques i ($U = 0, p = 0$) és un node inestable degenerat com es mostra a la figura 20.3. De la mateixa manera, el sistema (20.4) prop de l'estat estacionari ($U = 1, p = 0$) porta al sistema lineal

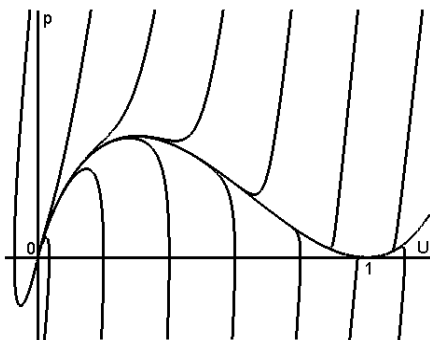
$$\frac{d(U - 1)}{dz} = p, \quad \frac{dp}{dz} = \frac{vp - f'(1)(U - 1)}{D}$$

i a l'equació característica $Dk^2 - vk + f'(1) = 0$. El discriminant és $\Delta = v^2 - 4Df'(1)$ i $\Delta \geq 0$ perquè $f'(1) \leq 0$. Si $f'(1) < 0$, hi ha dues arrels reals de signe contrari i $(U = 1, p = 0)$ és un punt de sella. Si $f'(1) = 0$, una arrel és zero i l'altra és positiva (vegeu la figura 20.3). Una anàlisi detallada mostra, de fet, que per a tot $v \geq 2\sqrt{f'(0)D}$ existeix una única corba integral que uneix els dos estats estacionaris $(U = 0, p = 0)$ i $(U = 1, p = 0)$, com al cas especial de la figura 20.3.

Kolmogórov, Petrovskij i Piskunov van continuar treballant en la demostració rigorosa del fet que l'equació en derivades parcials (20.3) té una solució única $u(x, t)$ que satisfà la condició inicial, que aquesta solució és tal que $0 < u(x, t) \leq 1$ per a tot x i $t > 0$, que $u(x, t)$ continua sent una funció creixent de x si ho és en $t = 0$ i, finalment, que $u(x, t)$ efectivament convergeix cap a un perfil d'ona que es propaga a velocitat v^* . Les demostracions són massa llargues per a resumir-les aquí.

Observeu que la funció $f(u) = au(1 - u)$ utilitzada per Fisher satisfà totes aquestes condicions amb $f'(0) = a$. Inspirant-se en l'equació (14.5), Kolmogórov, Petrovskij i Piskunov van considerar la funció $f(u) = au(1 - u)^2$, que satisfà les mateixes condicions i dona la mateixa velocitat de propagació.

Figura 20.3: Diagrama (U, p) que mostra algunes corbes integrals del sistema (20.5) i en particular la única corba que uneix $(U = 1, p = 0)$ amb $(U = 0, p = 0)$, que és la que dona la forma de l'ona que es propaga. Aquí, $f(u) = au(1 - u)^2$, $a = 1$, $D = 1$ i $v = v^* = 2$.



Els articles de Fisher i de Kolmogórov, Petrovskij i Piskunov van ser el punt de partida per a la construcció de molts models matemàtics amb difusió geogràfica en genètica, ecologia i epidemiologia. Aquests models es coneixen com «sistemes de reacció-difusió».

Pel que fa a Kolmogórov, a partir del 1938 també va estudiar el problema de l'extinció dels cognoms considerat per Bienaymé, Galton, Watson, Fisher,

Haldane, Erlang i Steffensen: va anomenar el procés estocàstic comú a tots aquests treballs «procés de ramificació». El 1939 va ser nomenat membre de l'Acadèmia de Ciències de l'URSS. Posteriorment va realitzar importants contribucions al problema de la turbulència en mecànica de fluids (1941), a la teoria dels sistemes dinàmics vinculats a la mecànica celeste (1953) i a la teoria de la informació (a partir del 1956). També va contribuir a la redacció d'una enciclopèdia i de llibres de text per a instituts de secundària i universitat, va ajudar a crear un institut experimental i va dirigir una revista de divulgació científica. Va rebre nombrosos premis internacionals (entre ells el Balzan el 1963 i el Wolf el 1980) i va morir a Moscou l'any 1987.

Petrovskij va ser nomenat degà de la Facultat de Mecànica i Matemàtiques de la Universitat Estatal de Moscou el 1940. Va ser rector de la universitat des del 1951 fins a la seva mort el 1973. Va ser membre de ple dret de l'Acadèmia de Ciències de l'URSS des del 1946 i president del Congrés Internacional de Matemàtics que es va celebrar a Moscou el 1966. També va escriure llibres de text sobre equacions diferencials ordinàries, equacions en derivades parcials i equacions integrals. Piskunov va esdevenir professor d'una acadèmia militar. El seu llibre de text sobre càlcul diferencial i integral ha estat utilitzat en moltes universitats tècniques. Va morir l'any 1977.

Més informació

1. Fisher, R.A.: The wave of advance of advantageous genes. *Ann. Eugen.* 7, 355–369 (1937) digital.library.adelaide.edu.au
2. Kolmogorov, A.N., Petrovskii, I.G., Piskunov, 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 Math. Mec.* 1:6, 1–26 (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. Shiryayev, A.N. (ed.): *Selected Works of A. N. Kolmogorov*, vol. 2. Kluwer (1992)
7. Shiryayev, A.N.: Andrei Nikolaevich Kolmogorov (April 25, 1903 to October 20, 1987). In: *Kolmogorov in Perspective*, 1–88. Am. Math. Soc. (2000)

Capítol 21

La matriu de Leslie (1945)

L'any 1945, l'ecòleg britànic P. H. Leslie va analitzar un model matricial per a una població de rosegadors amb estructura en edat, adaptant així el treball de Lotka a un marc de temps discret. Va destacar que la taxa de creixement ve donada per un valor propi de la matriu, i l'estructura estable en edats per un vector propi. També va estimar numèricament el nombre bàsic de reproducció neta \mathcal{R}_0 per a una població de rates comunes.

Patrick Holt Leslie va néixer l'any 1900 prop d'Edimburg, a Escòcia. Va estudiar al *Christ Church College* de la Universitat d'Oxford i va obtenir el 1921 una llicenciatura en fisiologia. No obstant, no va poder acabar els seus estudis de medicina per problemes de salut. Després d'uns anys treballant com a assistent de bacteriologia al departament de patologia, es va dedicar a l'estadística i es va incorporar el 1935 a l'Oficina de Població Animal, un nou centre de recerca creat per Charles Elton. Aquest centre es dedicava a l'anàlisi de les fluctuacions de les poblacions animals mitjançant estudis de camp i experiments de laboratori. La major part de la recerca es va realitzar sobre rosegadors: anàlisi dels cicles de la llebre i del seu depredador el linx utilitzant arxius de la Companyia de la Badia de Hudson al Canadà, seguiment de l'expansió territorial de l'esquirol gris a costa de l'esquirol vermell a Anglaterra, recol·lecció de dades sobre els talps a l'entorn d'Oxford, etc. Leslie va aplicar a les dades sobre els talps els mètodes desenvolupats per Lotka per a la demografia humana. Durant la Segona Guerra Mundial, la recerca del centre es va centrar en els mètodes de control de rates i ratolins a les sitges.

El 1945 Leslie va publicar el seu article més famós a *Biometrika*, una revista que havia estat fundada per Galton, Pearson i Weldon el 1901. L'article es titulava *Sobre l'ús de matrius en unes determinades matemàtiques de poblacions*. Leslie considerava un model per al creixement del nombre de femelles en una població animal, per exemple una població de rates (però també podria ser humana). La població es divideix en $K + 1$ grups d'edat: $P_{k,n}$ és el nombre de femelles d'edat k a temps n ($k = 0, 1, \dots, K; n = 0, 1, \dots$). Denotem per f_k la fertilitat a l'edat k o, més exactament, al nombre de filles nascudes per cada femella entre el temps n i el temps $n + 1$. Així, K és l'edat màxima amb fertilitat no nul·la ($f_K > 0$). Denotem per s_k la probabilitat que



Figura 21.1: P. H. Leslie (1900–1972)

un animal d'edat k sobrevisqui almenys fins l'edat $k + 1$. Llavors l'evolució en el temps de l'estructura en edat de la població (vector de població) ve donada pel següent conjunt d'equacions:

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

Tots els nombres f_k són no negatius, mentre que les probabilitats de supervivència s_k satisfan $0 < s_k < 1$. A finals del segle XIX i principis del XX, els matemàtics havien adquirit el costum d'escriure aquests sistemes d'equacions en la forma abreujada¹

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

on P_n és el vector columna $(P_{0,n}, \dots, P_{K,n})$ i M és la matriu quadrada (és a dir, la taula de nombres amb $K + 1$ files i $K + 1$ columnes)

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

Per a entendre el comportament del sistema (21.1) en funció del temps, Leslie va buscar una solució geomètricament creixent o decreixent $P_n = r^n V$. El nombre r i el vector V han de satisfer

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

¹El que significa que $P_{k,n+1} = M_{k,0} P_{0,n} + M_{k,1} P_{1,n} + \dots + M_{k,K} P_{K,n}$ per a tot k .

En aquest cas, r s'anomena valor propi i V vector propi de la matriu M . En altres paraules, el problema és trobar la distribució d'edats V que en cada pas de temps es multiplica per una constant r . Seguint la terminologia de Lotka, aquestes distribucions s'anomenen «estables». Tornant a notacions més habituals, l'equació (21.2) pot reescriure's com

$$\begin{cases} f_0 V_0 + f_1 V_1 + \dots + f_K V_K = r V_0, \\ s_0 V_0 = r V_1, \quad s_1 V_1 = r V_2, \quad \dots, \quad s_{K-1} V_{K-1} = r V_K. \end{cases}$$

De les darreres K equacions es dedueix que

$$V_1 = \frac{s_0 V_0}{r}, \quad V_2 = \frac{s_0 s_1 V_0}{r^2}, \quad \dots, \quad V_K = \frac{s_0 s_1 \dots s_{K-1} V_0}{r^K}.$$

Substituint-ho a la primera equació, dividint per V_0 i multiplicant per r^K , Leslie va obtenir l'equació característica

$$r^{K+1} = f_0 r^K + s_0 f_1 r^{K-1} + s_0 s_1 f_2 r^{K-2} + \dots + s_0 s_1 \dots s_{K-1} f_K. \quad (21.3)$$

Es tracta d'una equació polinòmica en r de grau $K + 1$, de manera que existeixen $K + 1$ arrels reals o complexes r_1, \dots, r_{K+1} . A més a més, Leslie va observar (utilitzant la regla dels signes de Descartes per a polinomis) que només hi ha una arrel real positiva. Denotem-la per r_1 .

Leslie va suggerir també que, en la majoria de condicions biològicament realistes (que poden precisar-se utilitzant la teoria de Perron i Frobenius per a matrius no negatives), el valor propi r_1 és estrictament major que el mòdul de tots els altres valors propis reals o complexos (que anomenarem r_2, \dots, r_{K+1}). A més a més, totes les arrels de (21.3) acostumen a ser diferents. Per a cada valor propi r_i , es pot trobar un vector propi associat. Sigui Q la matriu quadrada de mida $K + 1$ les $K + 1$ columnes de la qual contenen els vectors propis associats respectivament a r_1, \dots, r_{K+1} . Llavors $MQ = QD$, on D és la matriu diagonal dels valors propis $[r_1, \dots, r_{K+1}]$. Aleshores $M = QDQ^{-1}$ i $P_n = M^n P_0 = QD^n Q^{-1} P_0$. Noteu que D^n és la matriu diagonal $[(r_1)^n, \dots, (r_{K+1})^n]$ i que $D^n / r_1^n \rightarrow \mathcal{D} = [1, 0, \dots, 0]$ quan $n \rightarrow +\infty$ perquè $r_1 > |r_i|$ per a $i \neq 1$. Per tant, $P_n / (r_1)^n$ convergeix cap a $Q \mathcal{D} Q^{-1} P_0$.

A la llarga, cada component del vector d'estructura en edat P_n augmenta o disminueix com $(r_1)^n$. Si $r_1 > 1$, la població creix exponencialment. Si $r_1 < 1$, llavors decreix exponencialment cap a l'extinció. A partir de l'equació

(21.3), es pot demostrar fàcilment que la condició $r_1 > 1$ és certa si i només si el paràmetre \mathcal{R}_0 , el nombre reproductiu bàsic, definit per

$$\mathcal{R}_0 = f_0 + s_0 f_1 + s_0 s_1 f_2 + \cdots + s_0 s_1 \cdots s_{K-1} f_K,$$

és estrictament major que 1. Noteu que $s_0 s_1 \cdots s_{k-1}$ és la probabilitat de sobreviure fins almenys l'edat k . Per tant, el paràmetre \mathcal{R}_0 és el nombre mitjà de filles nascudes d'una femella al llarg de la seva vida i és anàleg a les fórmules (10.2), (12.2) i (16.9). Aquest model és una espècie d'anàleg a temps discret del treball de Lotka (vegeu el capítol 10) i una generalització del treball d'Euler que inclou fertilitats dependents de l'edat (vegeu el capítol 3).

Leslie va il·lustrar el seu mètode utilitzant dades publicades per un col·lega estatunidenc sobre els coeficients de fertilitat i supervivència f_k i s_k per a una població de rates comunes. Després d'algunes operacions estadístiques per a completar les dades de manera raonable, va obtenir $\mathcal{R}_0 \approx 26$.

La formulació matricial de Leslie per als problemes de dinàmica de poblacions és utilitzada actualment per molts biòlegs. Els càlculs se simplifiquen molt gràcies als ordinadors moderns i programes científics que poden calcular els valors i vectors propis de qualsevol matriu. Es pot calcular fàcilment tant el paràmetre \mathcal{R}_0 com la taxa de creixement r_1 .

Després de la Segona Guerra Mundial, Leslie va utilitzar el seu mètode per a calcular la taxa de creixement d'altres espècies animals: ocells, escarabats, etc. També va treballar en models estocàstics, en models de competència entre espècies i en l'anàlisi de dades de captura-recaptura. Es va jubilar el 1967. Aquell mateix any, en jubilar-se també Charles Elton, l'Oficina de Població Animal va deixar d'existir com a centre de recerca independent i va passar a formar part del Departament de Zoologia de la Universitat d'Oxford. Leslie va morir el 1972.

Més informació

1. Anònim: 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)

Capítol 22

Percolació i epidèmies (1957)

L'any 1957, Hammersley i Broadbent van considerar la propagació d'un «fluid» en una xarxa quadrada regular infinita, en la que dos nodes veïns estan connectats amb una probabilitat determinada. Entre els exemples possibles, van esmentar la propagació d'una epidèmia en una plantació d'arbres. Van demostrar que existeix una probabilitat crítica per sota de la qual no pot produir-se una gran epidèmia i per sobre de la qual es produeixen grans epidèmies amb una probabilitat positiva. El seu article va ser el punt de partida de la teoria de la percolació.

John Michael Hammersley va néixer el 1920 a Escòcia, on el seu pare treballava per a una empresa americana d'exportació d'acer. Va començar a estudiar a l'*Emmanuel College* de la Universitat de Cambridge, però va haver d'allistar-se a l'exèrcit el 1940. Va treballar en la millora dels càlculs per a l'artilleria. Després de finalitzar els seus estudis el 1948, va treballar com a professor ajudant a la Universitat d'Oxford al grup que treballava en el disseny i anàlisi d'experiments. El 1955 es va incorporar a l'Establiment de Recerca de l'Energia Atòmica a Harwell, prop d'Oxford.



Figura 22.1:
Hammersley (1920–2004)

Simon Ralph Broadbent va néixer el 1928. Va estudiar enginyeria a Cambridge, matemàtiques al *Magdalen College* d'Oxford (on també va escriure poesia) i va començar un doctorat en estadística a l'*Imperial College* de Londres titulat *Alguns tests de partida des de la dispersió uniforme*. Durant el seu doctorat, va obtenir un ajut de l'Associació Britànica de Recerca sobre

la Utilització del Carbó per a investigar problemes estadístics que poguessin estar relacionats amb la producció de carbó.

El 1954 es va celebrar a la *Royal Statistical Society* de Londres un simposi sobre els mètodes de Montecarlo patrocinat per l'Establiment de Recerca de l'Energia Atòmica. Aquests mètodes, iniciats a la dècada del 1940 per John von Neumann, Stanisław Ulam i Nicholas Metropolis al Laboratori dels Àlbers, utilitzen simulacions estocàstiques per ordinador per a estimar quantitats matemàtiques desconegudes. Hammersley va presentar al simposi de Londres un article que havia preparat en col·laboració amb Morton, un col·lega de Harwell. El treball es va publicar també al *Journal of the Royal Statistical Society*. Durant el debat que va seguir a la presentació al simposi, Broadbent va esmentar un problema interessant que podria estudiar-se utilitzant algun mètode de Montecarlo: donada una xarxa regular de porus en dues o tres dimensions tal que dos porus veïns estan connectats amb una probabilitat p , quina proporció de la xarxa s'ompliria amb un gas que s'introduís a través d'un d'aquests porus? De fet, Broadbent pensava en el disseny de les màscares de gas per als miners de carbó i, en particular, en la mida dels porus necessària per al seu funcionament.

Hammersley va començar aleshores a treballar amb Broadbent en aquest problema de la màscara de gas. Es van adonar que no era més que un prototip d'una família de problemes que encara no s'havia estudiat: la propagació determinista d'un «fluid» (el significat concret depèn del context) en un mitjà aleatori. Hammersley ho va anomenar «percolació», per analogia amb el que passa a una cafetera. A l'Institut de Recerca de l'Energia Atòmica, Hammersley també va tenir accés a alguns dels ordinadors més potents de la seva època per a testejar els mètodes de Montecarlo en problemes de percolació.

El 1957 Broadbent i Hammersley van publicar finalment el primer article sobre la teoria matemàtica de la percolació. Entre els exemples que van considerar, un era un model de dinàmica de poblacions, concretament la propagació d'una epidèmia en un camp d'arbres fruiters. Se suposa que els arbres d'un camp molt gran estan col·locats als nodes d'una xarxa quadrada. Cadascun dels quatre arbres més pròxims a un determinat arbre infectat té una probabilitat p d'infectar-se també. La pregunta és si s'infectarà un nombre gran d'arbres o si l'epidèmia es mantindrà localitzada. Això depèn, per descomptat, de la probabilitat p , que a la vegada està relacionada amb la distància que separa els arbres, és a dir, l'amplària de la xarxa.

Broadbent i Hammersley van estudiar el cas límit en que el camp és infinit i cobreix tot el pla, amb un sol arbre infectat al principi. Sigui $f(p)$ la probabilitat que un nombre infinit d'arbres s'infecti a partir d'aquesta font. És d'esperar que $f(p)$ sigui una funció creixent de p amb $f(0) = 0$ i $f(1) = 1$.

El seu principal resultat va ser demostrar que existeix una probabilitat crítica p^* , $0 < p^* < 1$, tal que:

- si $p < p^*$, llavors $f(p) = 0$ pel que només s'infectarà un nombre finit d'arbres;
- si $p > p^*$, llavors $f(p) > 0$ i es pot infectar un nombre infinit d'arbres.

La demostració involucra una comparació amb el nombre de diferents «passejos auto-evitats» al pla que comencen al focus d'infecció. Aquests passejos passen per un cert nombre d'arbres veïns (recordem que cada arbre té quatre veïns) sense visitar cap arbre més d'una vegada. Un passeig auto-evitat de n passos té una probabilitat p^n de ser un camí d'infecció, ja que la infecció pot transmetre's de cada arbre visitat al següent amb una probabilitat p . Sigui ara $q(j, n)$ la probabilitat que, entre tots els passejos auto-evitats de n passos, hi hagi exactament j passejos que siguin camins d'infecció. Si hi ha un nombre infinit d'arbres infectats, llavors per a tot nombre enter n existeix almenys un passeig auto-evitat de n passos que és un camí d'infecció. Així que

$$0 \leq f(p) \leq \sum_{j=1}^{+\infty} q(j, n) \leq \sum_{j=1}^{+\infty} j q(j, n)$$

per a tot n . Però

$$\sum_{j=1}^{+\infty} j q(j, n)$$

és el nombre esperat de passejos auto-evitats de n passos que són camins d'infecció. Aquest nombre és igual a $p^n s(n)$, on $s(n)$ és el nombre total de passejos auto-evitats de n passos. Hammersley va poder demostrar a un article complementari que $s(n)$ creix com $e^{\kappa n}$ quan $n \rightarrow +\infty$, on κ s'anomena la constant connectiva. Si $p < e^{-\kappa}$, llavors $p^n s(n)$ tendeix a 0 quan $n \rightarrow +\infty$ i $f(p) = 0$. Per tant, $p^* \geq e^{-\kappa} > 0$.

Per tant, a la pràctica és millor que els arbres no estiguin massa a prop per a mantenir p per sota de p^* en cas d'epidèmia. No obstant això, com més a prop estiguin els arbres, major serà la producció per hectàrea. Cal arribar a un equilibri.

Com van observar Broadbent i Hammersley, existeix una certa similitud entre l'existència d'una probabilitat crítica en els processos de percolació i l'existència d'un llindar en els processos de ramificació (vegeu el capítol 7).

Es pot intentar estimar numèricament la probabilitat crítica p^* . Per a això, es fixa un valor per a p i s'aproxima la xarxa infinita per una xarxa quadrada finita de mida $N \times N$ amb N prou gran. Suposem, per exemple, que l'arbre del centre de la xarxa està infectat. Amb un ordinador, es pot escollir a l'atzar quins arbres poden infectar altres arbres. Les figures 22.2a i 22.2b mostren els camins d'infecció escollits a l'atzar utilitzant arestes com en un graf. A la figura 22.2a, p és menor que p^* . A la figura 22.2b, p és major que p^* . Es pot determinar fàcilment quins arbres poden ser infectats, és a dir, aquells als quals es pot arribar per un camí d'arestes que parteixi de l'arbre infectat del centre. Estan marcats amb petits quadrats negres a les figures.

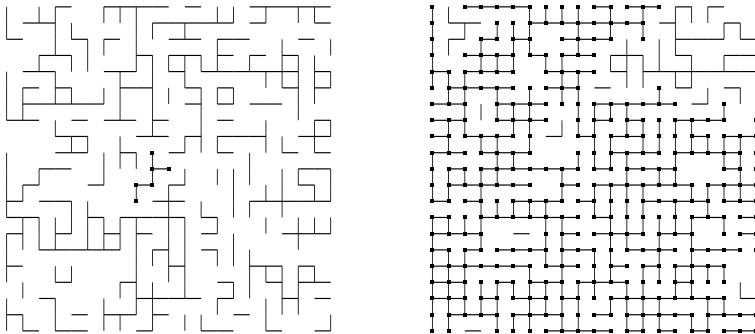


Figura 22.2: Percolació amb $p = 0,4$ o $p = 0,55$.

A continuació, es pot comprovar si l'epidèmia ha assolit almenys la frontera de la xarxa $N \times N$. Si això és així i si N és prou gran, es pot considerar que el nombre d'arbres infectats és «gairebé infinit». Repetint aquest tipus de simulació moltes vegades, es pot trobar un valor aproximat de la probabilitat $f(p)$ que el nombre d'arbres infectats sigui infinit (aquest és el mètode de Montecarlo). Finalment, fent variar p entre 0 i 1, es pot obtenir una aproximació del llindar p^* , que és el valor més petit tal que $f(p) > 0$ si $p > p^*$.

L'article de Broadbent i Hammersley només contenia la demostració de l'existència del llindar p^* . Durant els anys següents, Hammersley va continuar desenvolupant la teoria matemàtica de la percolació, mentre que Broadbent es va dedicar a altres temes. Amb el desenvolupament dels ordinadors a la dècada del 1970, va resultar més fàcil realitzar les simulacions descrites

anteriorment (figura 22.3). Llavors es va conjecturar que $p^* = 1/2$. Aquest resultat va ser finalment demostrat el 1980 per Harry Kesten, de la Universitat de Cornell.

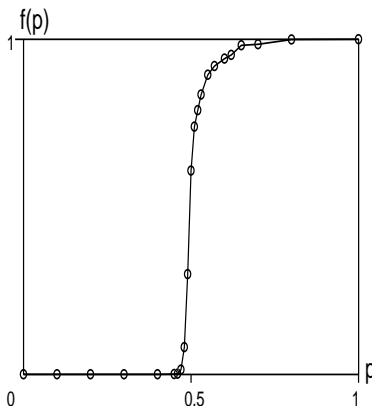


Figura 22.3: Probabilitat $f(p)$ que s'infectin infinits arbres en funció de p . La corba s'obté realitzant 1 000 simulacions en una xarxa de 200×200 .

Entre els anys 1959 i 1969, Hammersley va treballar a l'Institut d'Economia i Estadística de la Universitat d'Oxford. Va arribar a ser membre del *Trinity College*. El 1964 va publicar, en col·laboració amb David Handscomb, un llibre titulat *Mètodes de Montecarlo*. Va ser escollit membre de la *Royal Society* el 1976. Es va jubilar el 1987, però va continuar visitant el Centre de Matemàtiques Industrials i Aplicades d'Oxford. Va morir l'any 2004.

Broadbent es va doctorar a l'*Imperial College* el 1957. Va trobar feina a una empresa industrial, la *United Glass Bottle Manufacturers*. Després de deu anys a la indústria, va començar a treballar a una agència de notícies, la *London Press Exchange*, que realitzava estudis científics sobre els seus lectors. L'agència va ser comprada el 1969 per *Leo Burnett*, una empresa de publicitat estatunidenca. Broadbent va treballar en com mesurar l'eficàcia de la publicitat i va publicar diversos llibres sobre aquest tema: *Gastar diners en publicitat* (1975), *Pressupost de publicitat* (1989), *Publicitat responsable* (1997) i *Quan fer publicitat* (1999). El 1980 va ajudar a crear els Premis a l'Eficàcia Publicitària. Va passar diversos anys a la seu de *Leo Burnett* a Chicago com a director d'economia de la marca. També va dirigir la seva pròpia consultoria, *BrandCon Limited*. Va morir l'any 2002.

Més informació

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)

Capítol 23

Teoria de jocs i evolució (1973)

L'any 1973 Maynard Smith i Price van publicar un article on analitzaven per què els animals evitaven utilitzar les seves armes més perilloses en els conflictes intraespecífics. El seu model utilitzava la teoria de jocs i va ser un dels que va iniciar l'aplicació d'aquesta teoria matemàtica als problemes evolutius.

John Maynard Smith va néixer a Londres el 1920. El seu pare, que era cirurgià, va morir quan ell tenia vuit anys. Maynard Smith va estudiar a l'*Eton College* i posteriorment va estudiar enginyeria al *Trinity College* de la Universitat de Cambridge. Després es va afiliar al Partit Comunista de la Gran Bretanya. El 1939, quan va esclatar la guerra, va intentar allistar-se com a voluntari a l'exèrcit, però va ser rebutjat per la seva mala visió. Va acabar els seus estudis d'enginyeria i va treballar durant alguns anys en el disseny d'avions militars. Finalment va decidir dedicar-se a la biologia, estudiant genètica a l'*University College* de Londres amb Haldane com a supervisor. Va esdevenir professor de zoologia el 1952. Va abandonar el Partit Comunista després dels esdeveniments del 1956 a Hongria. El seu primer llibre, titulat *La teoria de l'evolució*, es va publicar el 1958. El 1965 va esdevenir professor de biologia a l'aleshores recentment fundada Universitat de Sussex. Llavors va publicar dos altres llibres: *Idees matemàtiques en biologia* (1968) i *Sobre l'evolució* (1972).

George R. Price va néixer l'any 1922 als Estats Units. Va estudiar química a la Universitat de Chicago i es va doctorar el 1946 després d'haver treballat al Projecte Manhattan en la construcció de la bomba atòmica. El 1950 va esdevenir investigador associat de medicina a la Universitat de Minnesota. Posteriorment va treballar com a periodista independent per a diverses revistes abans de tornar a la recerca a IBM. El 1967, després de ser tractat mèdicament d'un càncer de tiroides, es va instal·lar a Anglaterra i es va dedicar a l'estudi d'un tema completament diferent: la biologia evolutiva. A partir del 1968 va treballar a Londres al Laboratori Galton de l'*University College*. El seu primer article en aquesta nova àrea, *Selecció i covariància*, es va publicar amb l'ajuda de W. D. Hamilton a un número del 1970 de *Nature* i contenia el que ara es coneix com l'equació de Price.

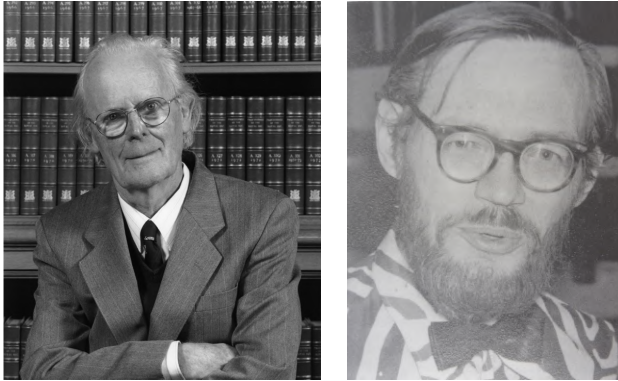


Figura 23.1: Maynard Smith (1920–2004) i Price (1922–1975)

Price també va presentar un altre treball a *Nature*, aquesta vegada sobre conflictes amb animals. En no tenir el format adequat per a aquesta revista, Maynard Smith, que era el revisor, va suggerir preparar una versió més curta. Price es va posar a treballar en una altra cosa mentre Maynard Smith va començar a desenvolupar la idea de Price pel seu compte. Finalment, Maynard Smith i Price van publicar un article conjunt titulat *La lògica del conflicte animal*, que *Nature* va publicar el 1973. L'article va suposar una interessant contribució a l'ús de la teoria de jocs en la biologia evolutiva. Abans d'això, la teoria de jocs s'havia desenvolupat principalment per a l'economia i la política, sobretot després del llibre de John von Neumann i Oskar Morgenstern del 1944 titulat *Teoria de jocs i comportament econòmic*. El punt de partida de Maynard Smith i Price va ser la següent pregunta: com és que als conflictes entre animals de la mateixa espècie, les «armes» de què disposen (banyes, arpes, verí, etc.) rares vegades s'utilitzen per a matar? Seguint les idees de Darwin sobre la lluita per la vida, els animals més agressius haurien de guanyar més combats i tenir un major nombre de cries, la qual cosa portaria a una escalada en l'ús de les «armes». Cal tenir en compte que es tractava de l'època de la Guerra Freda, per la qual cosa el tema també tenia un cert regust polític.

Maynard Smith i Price van considerar una successió de jocs en la qual dos animals poden competir per un recurs com, per exemple, un territori en un hàbitat favorable. En la presentació simplificada que Maynard Smith utilitzaria al seu llibre del 1982 *Evolució i teoria de jocs*, cada animal adopta o bé l'«estratègia del falcó» o bé l'«estratègia del colom». A partir d'ara parlarem simplement de falcons i coloms, però ens referim a estratègies adoptades

per animals de la mateixa espècie. Sigui $V > 0$ el valor del recurs, la qual cosa significa que si \mathcal{R}_0 és el nombre mitjà normal de cries d'un animal, el guanyador de la competició té de mitjana $\mathcal{R}_0 + V$ cries.

Si un falcó es troba amb un altre falcó, lluiten pel recurs: el guanyador obté el recurs de valor V , el perdedor pateix un «cost» $C > 0$. Cadascun dels dos falcons té una probabilitat igual a $1/2$ de guanyar la competició i la mateixa probabilitat de perdre. La recompensa esperada d'una baralla entre dos falcons és, per tant, $\frac{1}{2}(V - C)$ per als dos competidors. En canvi, si un falcó es troba amb un colom, el falcó obté el recurs V , el colom s'escapa sense lluitar i el cost és 0. Finalment, si dos coloms es troben, un d'ells obté el recurs V , l'altre s'escapa sense lluitar i sense cap cost. Cadascun dels dos coloms té la mateixa probabilitat $1/2$ de guanyar, per la qual cosa el resultat esperat quan dos coloms es troben és $V/2$. Les recompenses es poden resumir com a la Taula 23.1.

Taula 23.1: Recompenses esperades del joc del falcó i el colom.

	un falcó	un colom
recompensa d'un falcó contra...	$\frac{1}{2}(V - C)$	V
recompensa d'un colom contra...	0	$V/2$

De forma més general, es poden imaginar baralles entre individus que poden adoptar una de les dues estratègies, diguin-se 1 i 2, amb una matriu de recompenses esperades $(G_{i,j})_{1 \leq i,j \leq 2}$. En l'exemple anterior, els falcons segueixen l'estratègia 1, els coloms la 2, $G_{1,1} = \frac{1}{2}(V - C)$, $G_{1,2} = V$, $G_{2,1} = 0$ i $G_{2,2} = V/2$. A l'article original del 1973, Maynard Smith i Price ja havien utilitzat, de fet, simulacions per ordinador per a provar més de dues estratègies possibles (que van anomenar «falcó», «ratolí», «assetjador», «represaliador» i «sondejador-represaliador»).

Imaginem ara una població gran d'animals de la mateixa espècie amb una proporció x_n de falcons i una proporció $1 - x_n$ de coloms a la generació n . Els falcons a la generació n tenen un nombre mitjà de cries igual a

$$R_1(n) = \mathcal{R}_0 + x_n G_{1,1} + (1 - x_n) G_{1,2}. \quad (23.1)$$

De la mateixa manera, els coloms tenen un nombre mitjà de cries igual a

$$R_2(n) = \mathcal{R}_0 + x_n G_{2,1} + (1 - x_n) G_{2,2}. \quad (23.2)$$

Per tant, el nombre mitjà de cries en tota la població és

$$R(n) = x_n R_1(n) + (1 - x_n) R_2(n).$$

Oblidant les possibles subtileses degudes a la reproducció sexual, veiem que la proporció de falcons a la generació següent és

$$x_{n+1} = x_n R_1(n) / R(n). \quad (23.3)$$

Per tant, $x_{n+1} > x_n$ si $R_1(n) > R(n)$ i $x_{n+1} < x_n$ si $R_1(n) < R(n)$. Hi ha tres possibles estats estacionaris: $x = 0$, $x = 1$ i

$$x^* = \frac{G_{1,2} - G_{2,2}}{G_{2,1} - G_{1,1} + G_{1,2} - G_{2,2}}$$

sempre que $0 < x^* < 1$. Al joc del falcó i el colom, $x^* = V/C < 1$ sempre que $V < C$.

De fet, $x = 0$ és un estat estacionari obvi de (23.3). Si $x \neq 0$ és un altre estat estacionari, llavors $R_1 = R = xR_1 + (1-x)R_2$. Per tant, o bé $x = 1$ o bé $R_1 = R_2$. Aquesta última possibilitat equival a $xG_{1,1} + (1-x)G_{1,2} = xG_{2,1} + (1-x)G_{2,2}$, la qual cosa dona l'estat estacionari x^* .

L'estat estacionari $x = 1$ correspon a una població amb el 100 % dels individus que segueixen l'estratègia 1. Aquest estat estacionari és estable si no pot ser envaït per uns pocs individus que segueixen l'estratègia 2. A partir de (23.3), veiem que aquesta condició és equivalent a tenir $R_1(n) > R(n)$ per a tots els x_n prou propers a 1. Atès que $R(n) = x_n R_1(n) + (1-x_n) R_2(n)$, la condició es converteix en $R_1(n) > R_2(n)$ per a tots els x_n prou propers a 1. Observant les expressions (23.1)-(23.2) de R_1 i R_2 , arribem a la conclusió que $x = 1$ és estable si i només si es compleix una de les dues condicions següents:

- $G_{1,1} > G_{2,1}$;
- $G_{1,1} = G_{2,1}$ i $G_{1,2} > G_{2,2}$.

En aquest cas, es diu que l'estratègia 1 és una «estratègia evolutivament estable». Al joc del falcó i el colom, la condició $G_{1,2} > G_{2,2}$ és sempre certa. Per tant, l'estratègia del falcó és evolutivament estable si i només si $G_{1,1} \geq G_{2,1}$, és a dir, si $V \geq C$.

L'estat estacionari $x = 0$ correspon a una població amb tots els individus seguint l'estratègia 2. Aquesta situació és simètrica a l'anterior si intercanviem els índexs 1 i 2. Al joc de falcons i coloms, tenim $G_{1,2} = V > G_{2,2} = V/2$ pel que l'estat estacionari $x = 0$ és sempre inestable. La introducció d'un petit nombre de falcons en una població de coloms conduiria a una invasió progressiva per part dels falcons.

De la mateixa manera, es pot demostrar que el tercer estat estacionari x^* , sempre que $0 < x^* < 1$, és sempre estable. Al joc dels falcons i els coloms, $x^* = V/C$ correspon a una població mixta amb falcons i coloms.

En conclusió, hi ha dos casos possibles al joc dels falcons i els coloms. Si $V \geq C$, és a dir, si el valor del recurs és major que el possible cost, llavors la població tendeix a un estat estacionari amb falcons i sense coloms, sigui quina sigui la condició inicial $x(0)$ amb $0 < x(0) < 1$. L'estratègia del falcó és llavors una estratègia evolutivament estable. Si, per contra, $V < C$, la població tendeix a un estat estacionari mixt amb una proporció x^* de falcons i una proporció $1 - x^*$ de coloms. Per tant, el model ofereix una explicació de per què els individus amb comportaments menys agressius poden sobreviure quan $V < C$. La fórmula $x^* = V/C$ mostra, a més a més, que com més gran sigui el cost C per als perdedors, menor serà la proporció x^* de falcons a la població. D'aquí ve que les espècies amb les «armes» més perilloses rares vegades les utilitzin per a les baralles intraespecífiques: prefereixen baralles rituals inofensives, en les quals els animals que competeixen tracten d'impressionar-se mútuament, però eviten les baralles reals que podrien causar lesions.

L'article original del 1973 de Maynard Smith i Price discutia el concepte d'estratègia evolutivament estable i utilitzava principalment simulacions per ordinador de disputes d'animals, registrant les recompenses de les diferents estratègies. L'anàlisi mitjançant equacions dinàmiques com (23.3) es va desenvolupar una mica més tard, en particular per Taylor i Jonker. Des de llavors, molts autors han aplicat idees de la teoria de jocs a qüestions de biologia evolutiva o, al revés, han aplicat enfocaments dinàmics evolutius a problemes més clàssics de la teoria de jocs. A més a més de les qüestions relatives als conflictes entre animals, es poden citar, per exemple, els problemes de la inversió parental de recursos o de la proporció entre sexes (la raó entre el nombre de mascles i femelles en néixer), aquests últims ja estudiats per Carl Düsing el 1884 i per Ronald Fisher al seu llibre del 1930 sobre *La teoria genètica de la selecció natural*. Altres models se centren en els aspectes dinàmics del «dilema del presoner» o bé del joc de «pedra, paper, tisores». També s'ha vist que el concepte d'estratègia evolutivament estable està estretament relacionat amb el concepte d'equilibri de Nash a la teoria de jocs.

Price, que havia estat un ateu convençut, va tenir una experiència mística el 1970 i es va convertir a la fe cristiana. Va abandonar la seva recerca el 1974 perquè considerava que «el tipus de genètica matemàtica teòrica que feia no era molt rellevant per als problemes humans». Va regalar totes les seves pertinences a persones sense llar i es va suïcidar uns mesos després.

Maynard Smith, en canvi, va continuar aquesta línia de pensament i va ser escollit membre de la *Royal Society* el 1977. Va publicar molts llibres:

Models en ecologia (1974), *L'evolució del sexe* (1978), *L'evolució i la teoria de jocs* (1982), *Els problemes de la biologia* (1986), *Tenia raó Darwin?* (1988) i *Genètica evolutiva* (1989). També va publicar en col·laboració amb E. Szathmáry *Les principals transicions en l'evolució* (1995) i *Els orígens de la vida: del naixement de la vida a l'origen del llenguatge* (1999). Es va jubilar el 1985. El 1999 va rebre el premi Crafoord de biociències de la Reial Acadèmia Sueca de Ciències per les seves «contribucions fonamentals al desenvolupament conceptual de la biologia evolutiva». El 2003 va publicar, en col·laboració amb D. Harper, *Senyals animals*. Va morir a Sussex l'any 2004.

Més informació

1. Charlesworth, B., Harvey, P.: John Maynard Smith, 6 January 1920–19 April 2004. *Biol. 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. Harman, O.: *The Price of Altruism*. W. W. Norton, London (2010)
5. Maynard Smith, J., Price, G.R.: The logic of animal conflict. *Nature* 246, 15–18 (1973)
6. Maynard Smith, J.: *Evolution and the Theory of Games*. Cambridge University Press (1982)
7. 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/entopdfs/schwartz2000.pdf
8. Sigmund, K.: John Maynard Smith and evolutionary game theory. *Theor. Pop. Biol.* 68, 7–10 (2005)
9. Taylor, P.D., Jonker, L.B.: Evolutionary stable strategies and game dynamics. *Math. Biosci.* 40, 145–156 (1978)
10. Von Neumann, J., Morgenstern, O.: *Theory of Games and Economic Behavior*. Princeton University Press (1944) archive.org

Capítol 24

Poblacions caòtiques (1974)

L'any 1974, Robert May, un físic australià reconvertit en ecòleg, va estudiar l'equació logística a temps discret com a model de dinàmica de poblacions. Va observar que es produïen bifurcacions inesperades i que el comportament asimptòtic podia ser fins i tot caòtic. Així que les pre-diccions a llarg termini poden ser impossibles fins i tot amb un model determinista simple. L'article de May va ser un dels que va començar la «teoria del caos».

Robert McCredie May va néixer el 1936 a Austràlia. Després d'estudiar física teòrica i doctorar-se per la Universitat de Sydney el 1959, va passar dos anys al departament de matemàtica aplicada de la Universitat de Harvard. De tornada a Austràlia, va treballar de professor de física teòrica. El 1971, durant una visita a l'Institut d'Estudis Avançats de Princeton, va canviar de tema de recerca i va començar a centrar-se en dinàmica de poblacions animals. El 1973 va començar a treballar de professor de zoologia a Princeton. Aquell mateix any va publicar un llibre titulat *Estabilitat i complexitat en models d'ecosistemes*.



Figura 24.1:
Robert M. May (1936–2020)

El 1974 May va publicar a *Science* un article titulat *Poblacions biològiques amb generacions no superposades: punts estables, cicles estables i caos*, en el qual demostrava que models matemàtics molt senzills de dinàmica de poblacions poden comportar-se de manera caòtica.

Per a entendre l'origen d'aquest problema, cal remuntar-se uns deu anys enrere. El 1963, Edward Lorenz, un meteoròleg estatunidenc que treballa-

va a l'Institut Tecnològic de Massachusetts (M.I.T.), s'havia adonat, mentre realitzava simulacions numèriques amb ordinador, que un model simplificat de l'atmosfera amb només tres equacions diferencials podia comportar-se de forma molt sorprenent: un canvi minúscul de les condicions inicials podia canviar per complet el resultat final d'una simulació i, per tant, també les previsions meteorològiques. El matemàtic Henri Poincaré, després d'haver treballat sobre el moviment dels planetes al Sistema Solar, ja havia pensat en aquesta possibilitat a principis del segle XX, molt abans de l'era dels ordinadors. Però a principis dels anys setanta, només uns pocs investigadors començaven a examinar amb més detall aquesta estranya propietat. A la Universitat de Maryland, James Yorke estudiava els treballs de Lorenz i va introduir el terme «caos» en aquest context. L'article¹ que va escriure en col·laboració amb el seu alumne Tien-Yien Li, titulat *El període tres implica el caos*, va aparèixer el 1975.

Per part seva, May es va centrar en el model

$$p_{n+1} = p_n + a p_n(1 - p_n/K), \quad (24.1)$$

on a i K són paràmetres positius i p_n representa la mida d'una població animal l'any n . Quan p_n és petit en comparació amb la capacitat del medi K , la dinàmica s'aproxima a un creixement geomètric $p_{n+1} \approx (1 + a)p_n$. L'equació completa és una mena d'anàleg a temps discret de l'equació logística introduïda per Verhulst (vegeu el capítol 6). Però a diferència d'aquesta altra, May va demostrar que l'equació a temps discret pot tenir un comportament molt més sorprenent, que és fàcil d'observar amb una simple calculadora de butxaca fent sumes i multiplicacions (figura 24.2). Maynard Smith ja havia considerat l'equació (24.1) al seu llibre del 1968 *Idees matemàtiques en biologia*. Però malgrat haver provat alguns valors numèrics per a a , no s'havia adonat que hi havia quelcom d'especial.

La figura 24.2, que és similar a la de l'article de May del 1974, mostra que la població p_n convergeix a un estat estacionari quan $0 < a < 2$. Quan $2 < a \leq 2,449$ (el límit superior 2,449 és una aproximació), la població p_n tendeix a un cicle de període 2. Quan $2,450 \leq a \leq 2,544$, la població p_n tendeix a un cicle de període 4. Quan $2,545 \leq a \leq 2,564$, p_n tendeix a un cicle de període 8, etc. Els intervals del paràmetre a per als quals p_n tendeix a un cicle de període 2^n es redueixen a mesura que augmenta n i mai superen el valor 2,570. Quan $a \geq 2,570$, p_n pot comportar-se de forma «caòtica».

El 1976 May va escriure una revisió del problema, publicada a *Nature*, titulada *Models matemàtics senzills amb dinàmiques molt complicades*. Allà

¹Un resultat més general va ser demostrat per O. M. Sharkovskij el 1964, però el seu article, publicat en una revista ucraïnesa de matemàtiques, no era molt conegut.

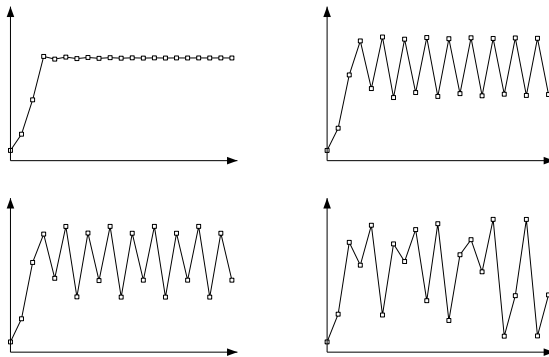


Figura 24.2: A totes les figures n està a l'eix horitzontal, p_n a l'eix vertical i $p_0 = K/10$. Les línies s'obtenen unint els punts amb coordenades (n, p_n) . A dalt a l'esquerra: $0 < a < 2$ (estat estacionari). A dalt a la dreta: $2 < a \leq 2,449$ (cicle de període 2). A baix a l'esquerra: $2,450 \leq a \leq 2,544$ (cicle de període 4). A baix a la dreta: $2,570 \leq a \leq 3$ (possiblement caos).

recollia no només els seus resultats, sinó també els d'altres investigadors. En primer lloc, denotant per $x_n = \frac{a p_n}{K(1+a)}$ i $r = 1 + a$ (de manera que $r > 1$), veiem que l'equació (24.1) pot reescriure's en la forma més senzilla

$$x_{n+1} = r x_n (1 - x_n). \quad (24.2)$$

Perquè aquesta equació tingui sentit en dinàmica de poblacions, x_n ha de ser no negativa per a tot n . Suposem que la condició inicial x_0 satisfà $0 \leq x_0 \leq 1$ i que $r \leq 4$. Aquesta última condició garanteix que el costat dret de (24.2) es manté entre 0 i 1. Cal remarcar que el cas caòtic $r = 4$ ja havia estat utilitzat com a generador de nombres aleatoris per Stanisław Ulam i John von Neumann el 1947. Si introduïm la funció $f(x) = r x (1 - x)$, llavors l'equació (24.2) pot reescriure's com $x_{n+1} = f(x_n)$ i els estats estacionaris són les solucions de $x = f(x)$. Gràficament correspon a les interseccions de les corbes $y = f(x)$ i $y = x$ (figura 24.3). Observeu que $x = 0$ és sempre un estat estacionari. Atès que $r > 1$, existeix també un altre estat estacionari $x^* > 0$ tal que $x^* = r x^* (1 - x^*)$, és a dir, $x^* = 1 - 1/r$.

Com que $r > 1$, l'estat estacionari $x = 0$ és inestable. En efecte, quan x_n s'aproxima a 0, tenim $x_{n+1} \approx r x_n$. De manera que x_n tendeix a allunyar-se de 0. Pel que fa a l'estat estacionari x^* , és localment estable només per a $1 < r < 3$.

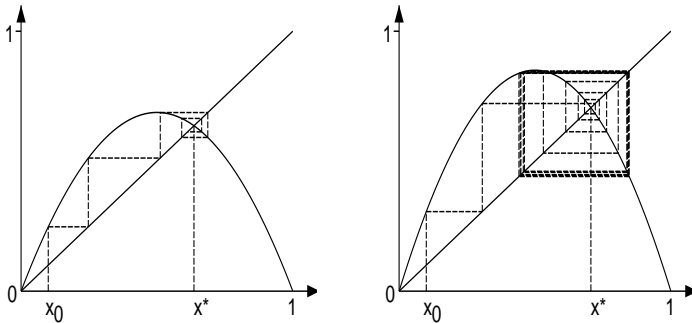


Figura 24.3: La corba $y = f(x)$, la recta $y = x$, l'estat estacionari x^* i la successió definida per $x_{n+1} = f(x_n)$. (a) : $r = 2,75$, la successió tendeix a x^* . (b) : $r = 3,4$, l'estat estacionari x^* és inestable i la successió tendeix a un cicle de període 2.

En efecte, sigui $y_n = x_n - x^*$. Llavors (24.2) és equivalent a $y_{n+1} = (2 - r - r y_n) y_n$. Si x_n està a prop de x^* , llavors y_n està a prop de 0 i $y_{n+1} \approx (2 - r) y_n$. Però si $y_{n+1} = k y_n$, llavors $y_n = k^n y_0$, de manera que $y_n \rightarrow 0$ quan $n \rightarrow +\infty$ si i només si $-1 < k < 1$. Aleshores l'estat estacionari x^* és localment estable si i només si $-1 < 2 - r < 1$, és a dir, quan $1 < r < 3$.

Quan $1 < r < 3$, es pot demostrar que per a totes les condicions inicials $0 < x_0 < 1$, la successió x_n tendeix realment a x^* (figura 24.3a). Però què succeeix quan $3 < r \leq 4$? Per a respondre a aquesta pregunta, observem primer que $x_{n+2} = f(x_{n+1}) = f(f(x_n))$. Introduïm la funció $f_2(x) = f(f(x)) = r^2 x(1-x)[1-rx(1-x)]$ i considerem les solucions de l'equació $x = f_2(x)$, que s'anomenen punts fixos de la funció $f_2(x)$. Gràficament són les interseccions de les corbes $y = f_2(x)$ i $y = x$ (figura 24.4).

Si $x = f(x)$, llavors $x = f(f(x)) = f_2(x)$. Per tant $x = 0$ i $x = x^*$ són també punts fixos de la funció $f_2(x)$. Però quan $r > 3$, la funció $f_2(x)$ té dos punts fixos més, x_- i x_+ , tals que $f(x_-) = x_+$ i $f(x_+) = x_-$.

En efecte, observem que $f_2'(x) = f'(f(x)) f'(x)$ pel que $f_2'(x^*) = [f'(x^*)]^2$. Com que $f'(x) = r(1-2x)$ i $x^* = 1 - 1/r$, tenim que $f'(x^*) = 2 - r$ i $f_2'(x^*) = (2 - r)^2$. Per tant, el pendent de la funció $f_2(x)$ a $x = x^*$ és tal que $f_2'(x^*) > 1$ si $r > 3$. A més a més, com que $f_2(0) = 0$, $f_2'(0) = r^2 > 1$

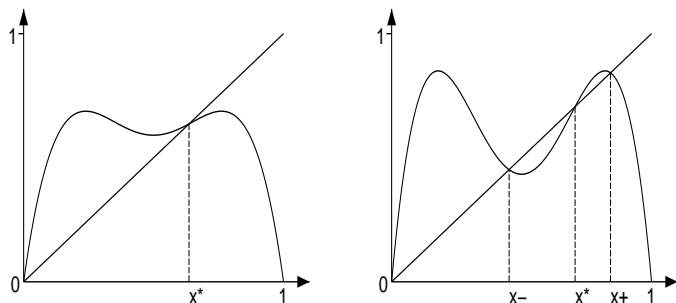


Figura 24.4: Les corbes $y = f_2(x) = f(f(x))$ i $y = x$ i l'estat estacionari x^* . (a) : $r = 2,75$. (b) : $r = 3,4$ i les altres dues solucions x_- i x_+ de l'equació $x = f_2(x)$.

i $f_2(1) = 0$, veiem a la figura 24.4b que necessàriament hi ha dues altres solucions x_- i x_+ de l'equació $x = f_2(x)$, amb $0 < x_- < x^* < x_+ < 1$. Una altra manera d'arribar a la mateixa conclusió consisteix a resoldre l'equació $x = f_2(x)$, que és una equació polinòmica de grau 4 amb dues arrels conegudes: $x = 0$ i $x = x^*$. Les altres dues solucions x_- i x_+ són les arrels del polinomi

$$x^2 - \frac{1+r}{r}x + \frac{1+r}{r^2} = 0. \quad (24.3)$$

Són reals si el discriminant és positiu, és a dir, si $r > 3$. Com que $f_2(f(x_-)) = f(f(f(x_-))) = f(f_2(x_-)) = f(x_-)$, el punt $f(x_-)$ és també un punt fix de $f_2(x)$. Però $f(x_-) \neq x_-$ perquè x_- no és un punt fix de $f(x)$. I $f(x_-) \neq x^*$, ja que en cas contrari tindríem $x_- = f(f(x_-)) = f(x^*) = x^*$. Atès que $f(x_-) \neq 0$, podem concloure que $f(x_-) = x_+$. Anàlogament, $f(x_+) = x_-$.

Per tant, per a $r > 3$, veiem que si per exemple $x_0 = x_-$, llavors $x_1 = x_+$, $x_2 = x_-$, $x_3 = x_+$, etc. També es pot demostrar que per a gairebé tota condició inicial $0 < x_0 < 1$, la successió x_n tendeix quan $n \rightarrow +\infty$ cap al cicle de període 2: x_-, x_+, x_-, x_+ , etc. (figura 24.3b i 24.4b). Aquest cicle es manté estable sempre que r estigui per sota del valor crític $r_1 = 1 + \sqrt{6} \approx 3,449$, on $f_2'(x_-) = -1$.

En efecte, veiem, utilitzant (24.3), que $f_2'(x_-) = f'(f(x_-))f'(x_-) = f'(x_+)f'(x_-) = r^2(1-2x_+)(1-2x_-) = r^2(1-2(x_+ + x_-) + 4x_+x_-) = r^2[1 - 2\frac{1+r}{r} + 4\frac{1+r}{r}] = -r^2 + 2r + 4$. Per tant $f_2'(x_-) = -1$ si $-r^2 + 2r + 5 = 0$ i, en particular, si $r = 1 + \sqrt{6}$.

En un cert interval $r_1 < r < r_2$, un cycle de període 4 esdevé estable: apareixen quatre nous punts fixos de la funció $f_4(x) = f_2(f_2(x)) = f(f(f(f(x))))$ (figura 24.5a). En un cert interval $r_2 < r < r_3$, hi ha un cycle de longitud 8, etc. Els nombres r_n tendeixen a un límit $r_\infty \approx 3,570$ quan $n \rightarrow +\infty$. Quan $r_\infty < r \leq 4$, el sistema pot ser fins i tot caòtic! La figura 24.5b mostra el diagrama² de bifurcació, que dona una idea de la complexitat de la dinàmica.

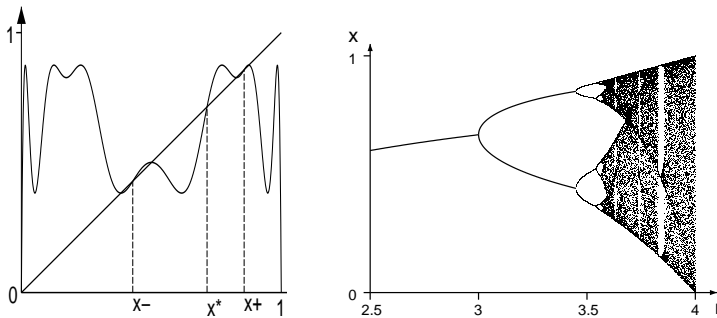


Figura 24.5: (a) La corba $y = f_4(x)$ quan $r = 3,5$ i la recta $y = x$. A més a més de x^* , x_+ i x_- , hi ha quatre punts fixos més, que no són fàcils de distingir. (b) Diagrama de bifurcació de l'equació (24.2).

R. M. May va concloure que fins i tot els sistemes dinàmics molt simples poden tenir un comportament molt complicat. Això no és específic de l'equació $x_{n+1} = rx_n(1 - x_n)$. La mateixa «cascada de duplicació de períodes» que condueix al caos apareix en altres equacions amb una funció $f(x)$ que té la forma d'un «gep». Aquest és el cas, per exemple, d'una altra equació utilitzada en dinàmica de poblacions: $x_{n+1} = x_n \exp(r(1 - x_n))$.

Aquest estudi suggereix que no cal sorprendre's si molts conjunts de dades relatives a la dinàmica de poblacions són difícils d'analitzar. El model també mostra que la distinció entre models deterministes i estocàstics no és

²Aquest diagrama s'ha obtingut dibuixant per a cada valor donat de r els punts amb coordenades $(r, x_{200}), (r, x_{201}), \dots, (r, x_{220})$, on $x_{n+1} = f(x_n)$ i $x_0 = 0.1$. Si x_n tendeix a un estat estacionari, només veiem un punt al diagrama. Si x_n tendeix a un cycle de període 2, veiem dos punts, etc.

tan clara com es pensava: fins i tot amb un model determinista simple, pot ser impossible fer previsions a llarg termini si els paràmetres estan en el règim caòtic.

El 1979 May va ser escollit membre de la *Royal Society*. Del 1988 al 1995 va ser catedràtic a la Universitat d'Oxford i a l'*Imperial College* de Londres. Del 1995 al 2000, va ser assessor científic en cap del govern britànic. El 1996 va rebre el premi Crafoord «per la seva recerca ecològica pionera en l'anàlisi teòrica de la dinàmica de poblacions, comunitats i ecosistemes». De l'ecologia va passar a l'epidemiologia i a la immunologia, i va publicar dos llibres: *Malalties infeccioses de l'ésser humà* (1991, amb Roy Anderson) i *Dinàmica dels virus, els fonaments matemàtics de la immunologia i la virologia* (2000, amb Martin Nowak). Aquest últim llibre analitza la interacció entre les cèl·lules del sistema immunitari i el VIH (el virus causant de la SIDA) com un tipus de sistema depredador-presa (vegeu el capítol 13). Del 2000 al 2005, May va ser president de la *Royal Society*. Va ser nomenat cavaller el 1996 i es va convertir en par vitalici el 2001. Va morir l'any 2020.

Més informació

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.: Coexistence of cycles of a continuous mapping of the line into itself (en rus). *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

Capítol 25

La política xinesa d'un sol fill (1980)

L'any 1980, Song Jian i els seus col·laboradors, especialistes en teoria de control aplicada a l'enginyeria de l'espai aeri, van calcular que si la taxa de natalitat de la Xina es mantenia al nivell d'aquell moment, la població assoliria més de dos mil milions d'habitants durant el segle XXI. Els seus resultats, basats en un model matemàtic estructurat per edats, van contribuir a la decisió del govern de recórrer a la política del fill únic.

Song Jian¹ va néixer el 1931 a Rongcheng, a la província xinesa de Shandong. A la dècada dels 50 va estudiar a la Unió Soviètica, a la Universitat Tècnica Estatal Bauman de Moscou i al Departament de Matemàtiques i Mecànica de la Universitat Estatal de Moscou. Després va tornar a la Xina i va treballar com a cap de l'Oficina de Recerca Cibernètica de l'Institut de Matemàtiques de l'Acadèmia Xinesa de Ciències. Va ser especialista en l'aplicació de la teoria de control al guiatge de míssils. També va treballar per al Setè Ministeri de Construcció de Maquinària, que més tard va passar a anomenar-se Ministeri Aeroespacial. El 1978 va començar a centrar-se en els vincles entre la teoria de control i la demografia.



Figura 25.1: Song Jian

Per a entendre el context de la recerca de Song Jian sobre dinàmica de poblacions, primer cal donar una idea del que és la «teoria de control». Es

¹Song és el cognom. Sempre s'escriu primer en xinès.

tracta de l'estudi de sistemes dinàmics, el comportament dels quals depèn d'uns paràmetres que poden modificar-se amb el pas del temps, per a optimitzar un criteri determinat. Aquesta teoria es va desenvolupar especialment en relació amb els programes espacials dels Estats Units i de l'URSS. En efecte, els enginyers havien de «controlar» la trajectòria de les llançadores espacials per tal de portar els satèl·lits a la seva òrbita al voltant de la Terra. Però les aplicacions no es limiten als problemes físics o d'enginyeria. Les polítiques de control de la natalitat també podrien considerar-se com una espècie de problema de control òptim en el sentit matemàtic.

També cal fer esment a l'assaig titulat *Els límits del creixement: un informe per al projecte del Club de Roma sobre la situació de la humanitat*, publicat el 1972 i redactat per un grup de l'Institut Tecnològic de Massachusetts (M.I.T.). Aquest estudi es basava en un model matemàtic del creixement econòmic mundial que tenia en compte els recursos naturals, la mida de la població i la contaminació. L'informe suggeria que l'economia mundial es dirigia cap a una catàstrofe per l'esgotament dels recursos no renovables, per la falta d'aliments per a la població o per un excés de contaminació. La limitació voluntària dels naixements era una de les solucions proposades. En resum, era una espècie de versió moderna de les tesis de Malthus. L'informe va tenir un gran ressò a Occident durant els anys 70.

Des de la fundació de la República Popular el 1949, la taxa de natalitat xinesa havia estat molt alta, excepte durant el catastròfic «Gran Salt Endavant». A mitjans de la dècada dels 70, la Xina es recuperava lentament de la Revolució Cultural. La planificació familiar instava les dones a retardar els naixements, a augmentar el temps entre dos parts consecutius i a tenir menys fills. Deng Xiaoping, que es va convertir en el nou líder després de la mort de Mao Zedong el 1976, va iniciar el 1978 la política de les «quatre modernitzacions»: agricultura, indústria, ciència i tecnologia i defensa nacional. La mida i el creixement de la població xinesa es van percebre llavors com a importants obstacles per a aquestes modernitzacions. Es va animar als científics que fins aleshores havien treballat en aplicacions militars a trobar solucions per a aquest difícil problema.

Amb aquests antecedents, Song Jian va anar el 1978 a Hèlsinki a un congrés de la Federació Internacional de Control Automàtic. Allà es va adonar que alguns investigadors d'Europa havien intentat aplicar la teoria de control als problemes de població amb la idea que un control estricte de la natalitat podria arribar a evitar les catàstrofes anunciades per l'informe *Els límits del creixement*. De tornada a la Xina, va crear un petit equip, que incloïa el seu col·lega Yu Jingyuan i l'informàtic Li Guangyuan, per a aplicar aquest tipus de models matemàtics a les dades relatives a la població xinesa. En aquella

època la comunicació científica entre la Xina i la resta del món era escassa. L'equip va tornar a obtenir les equacions que descriuen l'evolució de l'estructura en edat d'una població, de la mateixa manera que havien fet Lotka i McKendrick (vegeu els capítols 10 i 16). Utilitzant un model a temps continu, denotarem

- $P(x, t)$ la població d'edat x a l'instant t ;
- $m(x)$ la mortalitat a l'edat x ;
- $P_0(x)$ l'estructura d'edat de la població a l'instant $t = 0$;
- $b(t)$ la fertilitat total de les dones a l'instant t , és a dir, el nombre mitjà de fills que tindria una dona durant la seva vida si la fertilitat específica per edat continués sent la mateixa que la de l'instant t ;
- f la proporció de naixements de dones;
- $h(x)$ la distribució de probabilitat de l'edat de la mare quan neix un nen ($\int_0^{+\infty} h(x) dx = 1$).

Amb aquestes notacions i hipòtesis, l'evolució de l'estructura en edat pot modelitzar-se mitjançant l'equació en derivades parcials

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

amb condició inicial $P(x, 0) = P_0(x)$ i condició de contorn

$$P(0, t) = b(t) f \int_0^{+\infty} h(x) P(x, t) dx,$$

on $b(t)$ és el paràmetre que cal controlar. Si la fertilitat total de les dones és constant i està per sobre del llindar crític

$$b^* = 1 / \left[f \int_0^{+\infty} h(x) \exp \left(- \int_0^x m(y) dy \right) dx \right],$$

aleshores la població augmenta exponencialment. Aquest criteri és similar a l'obtingut per Lotka amb la fórmula (10.2). L'equip de Song Jian va considerar també la versió a temps discret del model, que és similar al model de Leslie (vegeu el capítol 21). Anomenem $P_{k,n}$ a la població d'edat k l'any n . Introduïm de manera similar m_k , b_n i h_k . Llavors

$$P_{k+1,n+1} = (1 - m_k) P_{k,n}, \quad P_{0,n+1} = b_n f \sum_{k \geq 0} h_k P_{k,n}.$$

Coneixent a partir d'enquestes per mostreig la mortalitat m_k (figura 25.2a), la proporció de naixements femenins $f \approx 0,487$, la distribució per edats de les

mares h_k (figura 25.2b), la condició inicial $P_{k,0}$ que és l'estructura per edats de la població el 1978 (figura 25.3a) i la variació de la fertilitat total b (que se suposa constant al llarg de cada simulació), l'equip de Song Jian va poder fer projeccions demogràfiques per al seu país amb un horitzó temporal de cent anys, des del 1980 fins al 2080 (figura 25.3b).

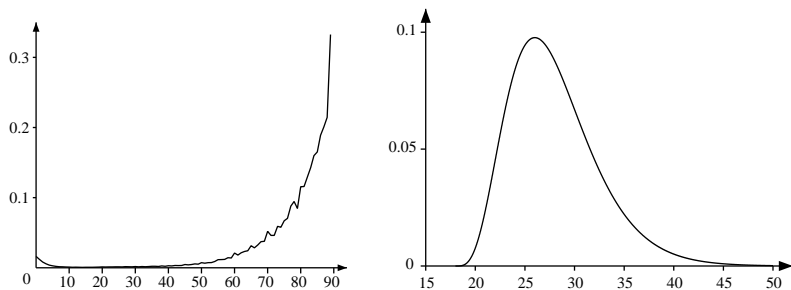


Figura 25.2: (a) Mortalitat (per any) en funció de l'edat el 1978. (b) Forma suavitzada de la fertilitat (per any) en funció de l'edat el 1978.

Donats els milers de sumes i multiplicacions necessàries (l'any n varia entre 0 i 100 anys, l'edat k entre 0 i 90 anys), era necessari l'ús d'un ordinador. En aquella època, a la Xina, poques persones tenien accés a aquesta mena d'equips, excepte els que treballaven per a l'exèrcit. Song Jian, un destacat expert en guiatge de míssils, era un d'ells.

Les projeccions suggerien que, encara que la Xina mantingués la seva fertilitat del 1978 de $b = 2,3$ fills per dona, que està just per sobre del llindar crític estimat en $b^* = 2,19$, la població passaria de 980 milions el 1980 a 2.120 milions el 2080. Però la Xina ja estava utilitzant gairebé tota la terra que podia servir per a l'agricultura. Fins i tot tenia tendència a perdre part d'aquestes terres a causa de la desertització i la urbanització. Com alimentar una població d'aquesta magnitud si el progrés dels rendiments agrícoles no era suficient? És la mateixa pregunta que es va plantejar Malthus dos segles abans. Amb una fertilitat el 1975 de $b = 3,0$, la població podria arribar fins i tot a 4.260 milions el 2080. Amb $b = 2,0$, la població assoliria un màxim de 1.530 milions cap a l'any 2050 abans de començar a disminuir lleugerament. Amb $b = 1,5$, s'assoliria un màxim de 1.170 milions cap a l'any 2030. Amb $b = 1,0$, el màxim seria només de 1.050 milions i s'assoliria als voltants de l'any 2000. Amb aquesta hipòtesi, la població no tornaria al seu nivell del

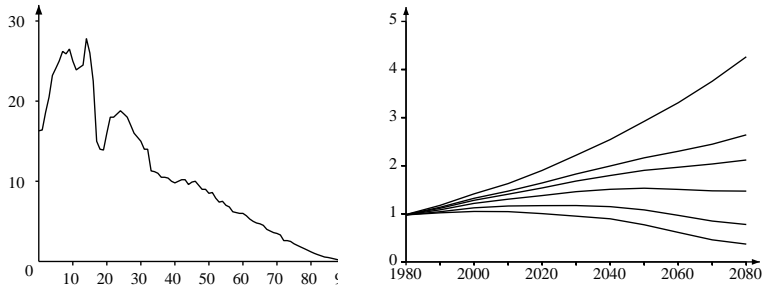


Figura 25.3: (a) Piràmide d'edat el 1978. Eix horitzontal: edat. Eix vertical: població (en milions). (b) Projeccions demogràfiques (en milers de milions) segons diferents hipòtesis sobre el nombre mitjà de fills per dona. De baix a dalt: $b = 1,0; 1,5; 2,0; 2,3; 2,5; 3,0$.

1978 fins al 2025.

El més sorprenent d'aquest treball van ser les seves conseqüències pràctiques d'una importància inigualable en la història dels models matemàtics en dinàmica de poblacions. De fet, Li Guangyuan va mostrar els resultats de les simulacions de l'equip el desembre del 1979 durant un simposi sobre població a Chengdu, província de Sichuan². Al gener del 1980, Song Jian, Yu Jingyuan i Li Guangyuan van publicar aquests resultats en una revista econòmica xinesa, advocant de pas per la política d'un sol fill. També van enviar el seu article - *Un informe sobre la recerca quantitativa de la qüestió del desenvolupament demogràfic de la Xina*- al principal científic xinès, Qian Xuesen, que el va remetre amb recomanació al cap de l'administració de planificació de la natalitat. Els resultats de l'equip de Song Jian van causar una profunda impressió en la majoria dels líders polítics. Aquests ja estaven convençuts de la necessitat d'un major control de la natalitat malgrat el que havia escrit Marx (vegeu el capítol 5), però continuaven dubtant sobre el nivell de control. Al febrer del 1980, el Consell d'Estat i el Comitè Central del Partit van fixar un objectiu per a la població xinesa de 1 200 milions per a l'horitzó de l'any 2000. Al març del 1980, els resultats de l'equip de Song Jian es van publicar al *Diari del Poble*. A l'abril, una comissió de dirigents polítics i especialistes en població va examinar les conseqüències mediambientals i econòmiques

²A continuació, resumim el relat detallat de Susan Greenhalgh [1,2].

del creixement demogràfic i va concloure que era necessària una política d'un sol fill per a aconseguir l'objectiu fixat per Deng Xiaoping de renda per capita l'any 2000. La política es va fer oficial al setembre d'aquell mateix any i es va publicar una carta oberta explicant-la a la població a la primera pàgina del *Diari del Poble*.

El 1983, encara hi hauria molts naixements no autoritzats. Es va decidir que un membre de cada parella que ja tingués dos fills seria esterilitzat i que s'interrompia cada embaràs prohibit. No obstant, a partir del 1984, es va permetre a les parelles rurals amb una sola filla tenir un segon fill. La política del fill únic va durar fins al 2015. Es van introduir progressivament algunes adaptacions: si en una parella tant l'home com la dona eren fills únics, llavors podien tenir dos fills. Les mesures repressives contra les parelles que tenien més d'un fill eren dures: els empleats públics podien perdre la feina, calia pagar una costosa multa per a aconseguir els papers administratius per a l'escolarització d'un segon fill, etc. En resum, és difícil trobar en la història de la modelització matemàtica un altre exemple amb un impacte social tan fort. Per descomptat, el treball de Song Jian i els seus col·laboradors va ser només un dels elements que van conduir a l'elecció de la política del fill únic. Però sembla haver exercit un paper important.

Tant en aquest capítol com en els anteriors, el paper de la modelització matemàtica s'ha de tractar amb cura. Partint d'una situació de la vida real, es construeix un model. Pot analitzar-se matemàticament o simular-se amb un ordinador. Així es pot entendre com es comporta el model quan varien alguns paràmetres. Malgrat tot, les matemàtiques no diuen si el model representa una imatge fidel de la vida real. És possible que no s'hagin tingut en compte alguns aspectes molt importants. Alguns models també contenen una funció objectiu, per exemple, mantenir la població xinesa per sota dels 1.200 milions. Les matemàtiques no diuen si aquest objectiu era adequat³.

El 1980, Song Jian va ser també coautor de la nova edició del llibre titulat *Enginyeria cibernètica* de Qian Xuesen, el «pare» del programa espacial xinès. Posteriorment, va ocupar diversos càrrecs polítics d'alt nivell: viceministre i científic-enginyer en cap del Ministeri d'Aeroespacial (1981-1984), membre del Comitè Central del Partit Comunista Xinès (1982-2002), president de la Comissió Estatal de Ciència i Tecnologia (1985-1998), Conseller d'Estat (1986-1998), etc. També ha publicat dos altres llibres que han estat traduïts a l'anglès: *Control de la població a la Xina* (1985, amb Tuan

³Es calcula que la població l'any 2000 era de 1.264 milions d'habitants. La renda per capita ha crescut aproximadament de 200 a 1.000 dòlars entre els anys 1980 i 2000. Al mateix temps, la proporció de sexes s'ha inclinat extremadament cap als homes, principalment a causa de l'avortament selectiu per sexe.

Chi-Hsien i Yu Jingyuan) i *Control del sistema de població* (1988, amb Yu Jingyuan). Aquests llibres desenvolupen la teoria de control òptim aplicada a la dinàmica de poblacions. Song Jian va ser escollit el 1991 membre de l'Acadèmia Xinesa de Ciències i el 1994 de l'Acadèmia d'Enginyers, de la qual va ser president del 1998 al 2002.

Més informació

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)

Capítol 26

Alguns problemes actuals

Aquest capítol ofereix una breu visió d'alguns problemes contemporanis de la modelització matemàtica en dinàmica de poblacions: l'envelliment de la població en demografia; les malalties emergents (SIDA, SARS, malalties transmeses per vectors...) i la política de vacunació en epidemiologia; les polítiques pesqueres en ecologia; la dispersió d'organismes genèticament modificats en genètica de poblacions. S'esmenten les institucions especialitzades que treballen a França en la modelització d'aquests problemes i es posa èmfasi en diversos aspectes de la recerca en aquests àmbits.

Aquest capítol ofereix una breu visió de la recerca contemporània sobre la modelització matemàtica en dinàmica de poblacions. Atès que el tema és bastant ampli, només s'ofereixen aquí alguns exemples d'estudis desenvolupats per investigadors a França.

En demografia ha aparegut un problema relativament nou en les últimes dècades: l'envelliment de la població. Aquest problema és motiu de preocupació no només a França (figura 26.1) sinó també en molts altres països europeus, així com al Japó. Té importants conseqüències econòmiques i socials: sistemes de pensions, polítiques d'immigració, etc. A França, l'Institut Nacional d'Estudis Demogràfics (INED) i l'Institut Nacional d'Estadística i Estudis Econòmics (INSEE) desenvolupen models matemàtics per a analitzar el fenomen de l'envelliment. Una de les dificultats de les projeccions demogràfiques resideix en el fet que les taxes de natalitat poden variar considerablement al llarg del temps sense que sigui possible preveure-les ni tan sols amb una dècada d'antelació. Això és especialment rellevant si s'observen les projeccions realitzades el 1968 per a la població francesa el 1985: aquestes projeccions no van anticipar el descens de la natalitat que es va produir durant els anys 70. Seria interessant repassar totes les prediccions basades en models matemàtics que van resultar ser errònies, especialment les que van trobar ressò als mitjans de comunicació. Això contrarestarà la impressió de «progrés» que dona aquest llibre, una impressió que pot haver semblat ja sospitosa al lector després de llegir el capítol sobre la política xinesa d'un sol fill. Respecte a aquest últim tema, un nou problema està ara d'actualitat: com

suavitzar aquesta política per a evitar el ràpid fenomen de l'envelliment que s'espera les pròximes dècades. De nou, els models matemàtics contribueixen al debat.

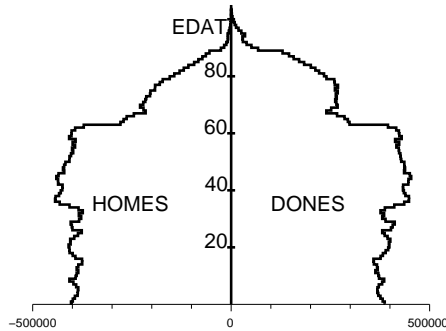


Figura 26.1: Piràmide d'edat de la població francesa a 1 de gener del 2010. Font: www.insee.fr.

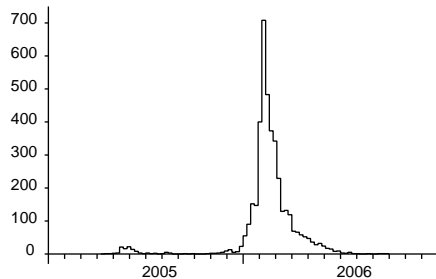
En epidemiologia, entre els nous problemes que han sorgit a nivell mundial en les dues últimes dècades, destaca el desenvolupament de l'epidèmia de la sida. Alguns models tracten de pronosticar el futur de l'epidèmia als països amb una infecció més recent, com Rússia, l'Índia o la Xina. És difícil predir si l'epidèmia s'alentirà com a Europa Occidental i Amèrica del Nord o si afectarà a un percentatge important de la població com en alguns països subsaharians. D'altres malalties emergents com el virus d'Ebola a l'Àfrica, la febre del Nil Occidental a Amèrica del Nord, la SARS (síndrome respiratòria aguda greu), la grip aviària, el chikungunya o la grip H1N1 han estat analitzades amb models matemàtics, encara que cal reconèixer que amb poc èxit.

En el cas de la SARS, una de les dificultats de la modelització era que l'epidèmia es mantenia relativament limitada dins de cada país, però podia estendre's molt ràpidament d'un país a un altre (Hong Kong i la Xina, Singapur, el Canadà...). El caràcter aleatori de les corbes de l'epidèmia en cada nou focus no podia passar-se per alt. Com vam veure als capítols 16 i 22, els models estocàstics solen ser més difícils de tractar.

Per a l'epidèmia de chikungunya que es va produir entre els anys 2005 i 2006 a l'illa de la Reunió (territori francès d'ultramar a l'oceà Índic), alguns models es van inspirar en el de Ross per a la malària (vegeu el capítol 12), ja que les dues malalties es transmeten per mitjà de mosquits. Un aspecte important a tenir en compte era la influència de l'estacionalitat. En efecte, la població de mosquits disminueix durant l'hivern austral, la qual cosa redueix

la transmissió de la malaltia. Això es pot veure a la figura 26.2, que mostra el nombre de nous casos notificats cada setmana per una petita xarxa d'uns trenta metges de capçalera que cobreixen només una fracció de la població de l'illa. La xarxa no va detectar cap cas nou durant diverses setmanes de setembre i octubre del 2005, però la transmissió de la malaltia continuava. A l'Institut Nacional de Salut i Recerca Mèdica (INSERM) i a l'Institut francès de Recerca per al Desenvolupament (IRD) es van elaborar models matemàtics de l'epidèmia. Malgrat aquests models, ningú va poder preveure que l'epidèmia no s'extingiria abans del final de l'hivern austral del 2005, quan només havia infectat a uns pocs milers de persones. Finalment, gairebé un terç de la població de l'illa es va infectar, és a dir, unes 266.000 persones. Això demostra, si encara cal, que predir el futur de les epidèmies pot ser bastant difícil i que no és tan fàcil distingir en els primers dies d'una epidèmia si serà una epidèmia menor o major. Es pot establir un paral·lelisme amb la previsió meteorològica. Aquest tipus de previsió es basa avui dia en intenses simulacions per ordinador de complicats models matemàtics de l'oceà i de l'atmosfera. Malgrat tot, les prediccions més enllà d'uns pocs dies no són fiables.

Figura 26.2: L'epidèmia de chikungunya a l'Illa de la Reunió entre 2005–2006. Nombre de nous casos notificats per setmana per una petita xarxa de metges en funció del temps. Cal multiplicar les xifres d'aquest gràfic per aproximadament 67 per a obtenir la mida real de l'epidèmia. Font: www.invs.sante.fr.



Des d'un punt de vista més teòric, l'epidèmia de chikungunya va plantejar la qüestió de com adaptar la noció de nombre bàsic de reproducció \mathcal{R}_0 en models que suposen que l'entorn té fluctuacions estacionals (per exemple, periòdiques). L'adaptació no és tan senzilla i això suscita una certa preocupació sobre com s'ha utilitzat el paràmetre \mathcal{R}_0 per a d'altres epidèmies influïdes per l'estacionalitat, com la pandèmia de grip H1N1 del 2009.

Un altre problema cada vegada més preocupant que els modelitzadors han intentat analitzar és el de la resistència als medicaments (antibiòtics, antiparlútics). Sempre en l'àmbit de l'epidemiologia, la qüestió recurrent des dels temps de Daniel Bernoulli i d'Alembert de com equilibrar els costos i els be-

nefícis quan la injecció d'una vacuna comporta un risc potencial continua sent objecte de controvèrsia i pot continuar sent-ho a mesura que canviï la sensibilitat al risc. Així, arran d'algunes indicacions que la vacuna contra l'hepatitis B podria causar complicacions greus en un nombre molt reduït de casos, el Ministeri de Sanitat francès va suspendre el 1998 la seva campanya de vacunació a les escoles, encara que el risc semblava insignificant en comparació amb el de morir després de la infecció pel virus de l'hepatitis B.

En ecologia, l'estudi de la dinàmica de les poblacions de peixos continua plantejant molts problemes. Malgrat això, se suposa que serveix de base científica per a l'elecció de quotes de pesca i altres restriccions. La sobrepesca de l'anxova al Golf de Biscaia i de la tonyina vermella al Mediterrani són només dos exemples recents. Atès que l'estimació de les poblacions de peixos sol ser poc fiable, els models que utilitzen aquestes dades han de considerar-se amb precaució. A França, aquest tipus d'estudis els realitza principalment l'Institut de Recerca per a l'Explotació de la Mar (IFREMER). Alguns models matemàtics també han jugat un paper en les decisions preses per la Comissió Ballenera Internacional.

En genètica de poblacions, la dispersió dels organismes modificats genèticament és també un tema de controvèrsia que alguns investigadors han tractat d'estudiar utilitzant models de «reacció-difusió» inspirats en el de Fisher (vegeu el capítol 20). Aquest és l'àmbit de l'Institut Nacional de Recerca Agronòmica (INRAE).

Des del punt de vista més teòric de la recerca, es poden esmentar:

- els treballs sobre equacions en derivades parcials, com les equacions de difusió (vegeu el capítol 20) o les equacions amb estructura d'edat (vegeu el capítol 16);
- els treballs sobre models estocàstics amb o sense dimensió espacial (vegeu els capítols 16 i 22), inclosos els relatius a les xarxes aleatòries que modelen la propagació d'epidèmies i els que busquen aproximacions deterministes.

Aquest tipus de recerca és duta a terme principalment per investigadors en matemàtica aplicada. En els últims anys s'han introduït diversos cursos de màster en biologia matemàtica a les universitats franceses i altres centres d'ensenyament superior.

Igual que altres camps científics, l'estudi matemàtic de la dinàmica de poblacions s'organitza principalment a través de:

- «Societats científiques»: Societat Neerlandesa de Biologia Teòrica (des del 1970), Societat de Biologia Matemàtica (1973), Societat Francesa

de Biologia Teòrica (1985), Societat Xinesa de Biologia Matemàtica (1985), Societat Japonesa de Biologia Matemàtica (1989), Societat Europea de Biologia Matemàtica i Teòrica (1991), Societat Llatinoamericana de Biologia Matemàtica (2002), etc.

- revistes especialitzades: *Bulletin of Mathematical Biology* (des del 1939), *Mathematical Biosciences* (1967), *Journal of Mathematical Biology* (1974), *Mathematical Medicine and Biology* (1984), *Mathematical Population Studies* (1988), *Mathematical Biosciences and Engineering* (2004), *International Journal of Biomathematics* (2008), *Biomath* (2012) etc.
- conferències (Reunió Anual de la Societat de Biologia Matemàtica, Dinàmica de Poblacions Matemàtica i Computacional, Conferència Europea de Biologia Matemàtica i Teòrica, etc.).

Només s'ha fet referència als elements que afirmen explícitament estar situats a la interfície entre les matemàtiques i les seves aplicacions a la dinàmica de poblacions. Però en cada àmbit concret (demografia, ecologia, genètica de poblacions, epidemiologia, etc.) es poden trobar elements similars amb una dosi variable de modelització matemàtica.

Per a concloure, es convida al lector interessat a consultar els articles originals que estan disponibles a la *World Wide Web*. Les adreces figuren a les referències al final de cada capítol. Com va escriure Ronald Fisher sobre Mendel:

«La Història de la Ciència ha patit molt per l'ús per part del professorat de material de segona mà i la consegüent obliteració de les circumstàncies i l'ambient intel·lectual en el que es van fer els grans descobriments del passat. Un estudi de primera mà sempre és instructiu i sovint ... està ple de sorpreses.»

Més informació

1. Bacaër, N.: Aproximación de \mathcal{R}_0 para enfermedades transmitidas por vectores con una población periódica de vectores. *Bull. Math. Biol.* 69, 1067–1091 (2007) hal.archives-ouvertes.fr
2. Levin, S.A.: Mathematics and biology, the interface. bio.vu.nl/nv-tb/Contents.html

Figures

- pàg. 5. Retrat de Thomas Murray (ca. 1687) conservat per la *Royal Society* de Londres. Chapman, S.: Edmond Halley, F.R.S. 1656–1742. *Notes Rec. R. Soc. Lond.* 12, 168–174 (1957) © The Royal Society.
- pàg. 11. Retrat de Emanuel Handmann (1753) conservat en el *Kunstmuseum* de Basilea. *Leonhard Euler 1707–1783, Beiträge zu Leben und Werk*. Birkhäuser, Basel (1983)
- pàg. 17. Retrat que va estar en poder de Petri-Kirche, probablement destruït durant la batalla de Berlín el 1945. Reimer, K.F.: Johann Peter Süssmilch, seine Abstammung und Biographie. *Arch. soz. Hyg. Demogr.* 7, 20–28 (1932)
- pàg. 23. Retrat de Johann Niclaus Gooth (ca. 1750–1755) conservat al Museu d'Història Natural de Basilea. Speiser, D.: *Die Werke von Daniel Bernoulli*, Band 2. Birkhäuser, Basel (1982)
- pàg. 30. Retrat de Maurice Quentin de La Tour (1753) conservat al Museu del Louvre de París.
- pàg. 33. Retrat de John Linnell (1833) en possessió del *Haileybury College*, Anglaterra. Habakkuk, H.J.: Robert Malthus, F.R.S. (1766-1834). *Notes Rec. R. Soc. Lond.* 14, 99–108 (1959)
- pàg. 37. Gravat de Flameng (1850). Quetelet, A.: Pierre-François Verhulst. *Annu. Acad. R. Sci. Lett. B.-Arts Belg.* 16, 97–124 (1850)
- pàg. 42. Heyde, C.C., Seneta, E.: *I. J. Bienaymé, Statistical Theory Anticipated*. Springer (1977) © Académie des sciences, Institut de France.
- pàg. 42. Brun, J., Robinet, A. (éd.): *A. Cournot, études pour le centenaire de sa mort*. Economica / Vrin, Paris (1978)
- pàg. 46. Bateson, W.: *Mendel's Principles of Heredity*. Cambridge University Press (1913)
- pàg. 51. Pearson, K.: *The Life, Letters, and Labors of Francis Galton*, vol. 1. Cambridge University Press (1914)
- pàg. 51. Retrat de Watson a la Biblioteca del *Trinity College*, Universitat de Cambridge. Kendall, D.G.: Branching processes since 1873. *J. Lond. Math. Soc.* 41, 385–406 (1966)
- pàg. 57. Documents d'Alfred J. Lotka. © Princeton University Library.
- pàg. 61. Titchmarsh, E. C.: Godfrey Harold Hardy 1877–1947. *Obit. Not. Fellows R. Soc.* 6, 446–461 (1949)
- pàg. 64. Stern, C.: Wilhelm Weinberg. *Genetics* 47, 1–5 (1962)
- pàg. 67. G.H.F.N.: Sir Ronald Ross 1857–1932. *Obit. Not. Fellows R. Soc.* 1, 108–115 (1933) © The Royal Society.
- pàg. 76. Whittaker, E.T.: Vito Volterra 1860–1940. *Obit. Not. Fellows R. Soc.* 3, 690–729 (1941)

- pàg. 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.
- pàg. 84. Yates, F.: George Udny Yule. *Obit. Not. Fellows R. Soc.* 8, 308–323 (1952)
- pàg. 92. Heyde, C.C., Seneta, E. (eds.): *Statisticians of the Centuries*. Springer (2001)
- pàg. 102. www.britannica.com/EBchecked/topic/252257/J-B-S-Haldane © Basano and Vandyk Studios.
- pàg. 109. Hill, W.G.: Sewall Wright, 21 December 1889–3 March 1988. *Biog. Mem. Fellows R. Soc.* 36, 568–579 (1990) © Llewellyn Studios, Chicago.
- pàg. 106. Nybølle, H.C.: Agner Krarup Erlang f. 1. Januar 1878 - d. 3. Februar 1929. *Mat. Tidsskr. B*, 32–36 (1929)
- pàg. 106. Nørðlund, N.E.: Johan Frederik Steffensen in memoriam. *Nordisk Mat. Tidsskr.* 10, 105–107 (1962)
- pàg. 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)
- pàg. 118. *I. G. Petrowsky Selected Works Part I*. Gordon and Breach, Amsterdam (1996) © Taylor and Francis Books UK.
- pàg. 123. Fotografia de Denys Kempson. Crowcroft, P.: *Elton's Ecologists, a History of the Bureau of Animal Population*. University of Chicago Press (1991)
- pàg. 126. © Geoffrey Grimmett.
- pàg. 133. 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.
- pàg. 133. Harman, O.: *The Price of Altruism*. W. W. Norton, London (2010)
- pàg. 138. © Samuel Schläefli / ETH Zürich.
- pàg. 145. Selected works of J. Song. *Science Press*, Beijing (1999) © Song Jian.

Índex

Introducció	iii
1 La successió de Fibonacci (1202)	1
2 Taula de vida d'Halley (1693)	4
3 Euler i el creixement geomètric (1748–1761)	10
4 Daniel Bernoulli i la inoculació de la verola (1760)	22
5 Malthus i els obstacles al creixement geomètric (1798)	33
6 Verhulst i l'equació logística (1838)	37
7 Bienaymé i l'extinció dels cognoms (1845)	42
8 Mendel i l'herència (1865)	46
9 Galton, Watson i el problema de l'extinció (1873–1875)	50
10 Lotka i la teoria de la població estable (1907–1911)	57
11 La llei de Hardy-Weinberg (1908)	61
12 Ross i la malària (1911)	66
13 Lotka, Volterra i el sistema depredador-presa (1920–1926)	72
14 Fisher i la selecció natural (1922)	79
15 Yule i l'evolució (1924)	83
16 McKendrick sobre la modelització d'epidèmies (1926–1927)	91
17 Haldane i les mutacions (1927)	101
18 Erlang sobre el problema de l'extinció (1929)	105
19 Wright i la deriva genètica aleatòria (1931)	109
20 La difusió dels gens (1937)	115
21 La matriu de Leslie (1945)	122
22 Percolació i epidèmies (1957)	126
23 Teoria de jocs i evolució (1973)	132
24 Poblacions caòtiques (1974)	138
25 La política xinesa d'un sol fill (1980)	145
26 Alguns problemes actuals	152

Aquest llibre recorre la història de la dinàmica de poblacions –un tema teòric estretament relacionat amb la genètica, l’ecologia, l’epidemiologia i la demografia– en la que les matemàtiques han aportat importants coneixements. Presenta una visió general de la gènesi de diversos temes de gran rellevància: el creixement exponencial, des d’Euler i Malthus fins a la política xinesa del fill únic; el desenvolupament dels models estocàstics, des de les lleis de Mendel i la qüestió de l’extinció de cognoms fins a la teoria de la percolació per a la propagació d’epidèmies, i les poblacions amb dinàmiques caòtiques, en les quals s’entremesclen el determinisme i l’atzar.

Amb els recents avanços de la traducció automàtica, ja no està justificat el virtual monopoli d’una sola llengua a la literatura científica. La creixent alienació lingüística a les universitats pot revertir-se. Amb aquesta traducció al català, acuradament revisada per dos especialistes, esperem mostrar aquest nou camí.

ISBN : 979-10-343-9682-5



15€