Nicolas Bacaër

Lenka Přibylová, Veronika Eclerová,
Zdeněk Pospíšil, Luděk Berec, Vlastimil Křivan

Stručná historie matematické populační dynamiky

Stručná historie matematické populační dynamiky

Nicolas Bacaër

Lenka Přibylová, Veronika Eclerová,
Zdeněk Pospíšil, Luděk Berec, Vlastimil Křivan

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

Veronika Eclerová, Zdeněk Pospíšil, Lenka Přibylová
Masarykova univerzita
eclerova@math.muni.cz, pospisil@math.muni.cz, pribylova@math.muni.cz

Luděk Berec, Vlastimil Křivan
Jihočeská univerzita v Českých Budějovicích
lberec@prf.jcu.cz, vlastimil.krivan@gmail.com

Čtenáři, kteří si chtějí zakoupit papírovou verzi této knihy, mohou poslat e-mail na adresu nicolas.bacaer@ird.fr.

Obrázek na obálce: Alfons Mucha (1860–1939), Fragment vlysu zdobícího pavilon Bosny a Hercegoviny na Světové výstavě v Paříži v roce 1900, © Muzeum Orsay.

Titre original : Histoires de mathématiques et de populations

Pour l'édition tchèque :

ISBN : 979-10-343-9751-8
Dépôt légal : avril 2022

Předmluva

Populační dynamika je vědní obor, který se snaží jednoduchým mechanistickým způsobem vysvětlit časové změny velikosti a složení biologických populací, jako jsou populace lidí, zvířat, rostlin nebo mikroorganismů. Je příbuzná s popisnější oblastí populační statistiky, ale přesto se od ní značně liší. Jejich společným rysem je, že hojně využívají matematický jazyk.

Populační dynamika je průnikem různých oborů: matematiky, společenských věd (demografie), biologie (populační genetiky a ekologie) a medicíny (epidemiologie). V důsledku toho není často prezentována jako celek, a to navzdory podobnostem mezi problémy, s nimiž se setkáváme v různých aplikacích. Významnou výjimkou je kniha Alaina Hilliona *Matematické teorie populace* ve francouzštině (1986). Představuje však toto téma z pohledu matematika a rozlišuje různé typy modelů: modely diskrétního času ($t = 0, 1, 2...$) a modely spojitého času (t je reálné číslo), deterministické modely (budoucí stavy jsou přesně známy, pokud je přesně znám současný stav) a stochastické modely (kde hrají roli pravděpodobnosti). V knize jsou pak uvažovány logicky diskrétní deterministické modely, spojité deterministické modely, diskrétní stochastické modely a spojité stochastické modely.

V této knize jsem se pokusil pojednat o stejném tématu, ale z historického hlediska. Výzkum je vysvětlen v jeho souvislostech. Součástí jsou krátké životopisy vědců. To by mělo usnadnit čtení knihy těm, kteří se v matematice orientují méně, a obvykle může pomoci pochopit původ zkoumaných problémů. Tato kniha však není jen o historii. Může sloužit také jako úvod do matematického modelování. Aby čtenář skutečně pochopil omezení modelů, bylo důležité zahrnout do knihy podrobnosti většiny výpočtů. Technické části jsou zdůrazněny v šedých rámečcích a při prvním čtení je lze přeskočit. Poslední kapitola se zaměřuje na četné současné problémy populační dynamiky, které se lze pokusit analyzovat z matematického hlediska. V české variantě knihy je přidáno postscriptum, které přibližuje přínos matematických modelů pro analýzu a řízení aktuální pandemie čínského koronaviru z pohledu matematiček aktivně zapojených do tohoto procesu. Pro ty, kteří by se chtěli dozvědět více, jsou v seznamu literatury na konci každé kapitoly uvedeny také webové stránky, z nichž lze stáhnout původní články.

V knížce tohoto rozsahu samozřejmě nebylo možné načrtnout úplný obraz o veškeré doposud se rozvíjející práci, ani pojednat o všech vědcích, kteří se na jejich tématech podíleli. Zejména v částech o posledních desetiletích, je výběr nutně trochu nahodilý. Přesto doufám, že zvolený vzorek je dostatečně reprezentativní a že z něho lidé působící v oboru, který není zmíněn, nebudou otráveni.

Kniha je určena ideálně těmto čtenářům:

- Studentům středních a vysokých škol, které zajímá, jaké souvislosti existují mezi matematickými předměty, které studují, a okolním světem, nebo studentům, kteří připravují práce na téma související s populační dynamikou.

- Učitelům matematiky, kteří se snaží zatraktivnit svůj předmět. K pochopení většiny kapitol 1, 2 a 5 stačí znalost čtyř základních operací. Kapitola 3 může sloužit jako úvod do aplikací logaritmů. Tato kniha se také zabývá rekurentními rovnicemi v kapitolách 1, 3, 8, 11, 14, 21, 23, 24; obyčejnými diferenciálními rovnicemi v kapitolách 4, 6, 12, 13, 16; parciálními diferenciálními rovnicemi v kapitolách 20, 25; integrálními rovnicemi v kapitole 10; a aplikacemi teorie pravděpodobnosti v kapitolách 2, 7, 8, 9, 15, 16, 17, 18, 19, 22.

- Lidem, kteří se již seznámili s demografií, epidemiologií, genetikou nebo ekologií a chtějí porovnat svou oblíbenou oblast s jinými, které mohou zahrnovat podobné matematické modely.

- Čtenářům se zájmem o historii vědy.

Tato kniha je v podstatě překladem francouzského vydání, které vyšlo v nakladatelství Cassini (Paříž) v roce 2008 pod názvem *Histoires de mathématiques et de populations*. Rád bych poděkoval Lence Přibylové, Veronice Eclerové, Zdeňku Pospíšilovi, Luďku Berecovi a Vlastimilu Křivanovi, kteří upravili překlad do češtiny vytvořený automaticky softwarem DeepL. Při své práci také nahlíželi do anglického překladu *A Short History of Mathematical Population Dynamics*, který vydalo nakladatelství Springer-Verlag (Londýn) v roce 2011.

Kapitola 1

Fibonacciho posloupnost (1202)

V roce 1202 vydal Leonardo z Pisy, zvaný Fibonacci, knihu, která v Evropě zpopularizovala indickou desítkovou číselnou soustavu, kterou před tím již přijali arabští matematici. Mezi mnoha v knize uvedenými příklady se jeden týká růstu populace králíků. Jedná se o jeden z nejstarších příkladů matematického modelu dynamiky populace.

Leonardo Pisánský, pojmenovaný Fibonacci až dlouho po své smrti, se narodil kolem roku 1170 v republice Pisa, která byla na vrcholu své obchodní a vojenské moci ve středomořském světě. Fibonacciho otce poslala republika kolem roku 1192 do přístavu Bejaia, který se dnes nachází v Alžírsku, aby zde řídil obchodní středisko. Jeho syn se k němu krátce nato připojil, aby se připravil na dráhu obchodníka. Leonardo se začal učit desítkovou číselnou soustavu, kterou Arabové přivezli z Indie a která se v téměř stejné podobě používá dodnes: 0, 1, 2, 3, 4, 5, 6, 7, 8 a 9. Při obchodních cestách kolem Středozemního moře porovnával různé číselné soustavy a studoval arabskou matematiku. Po návratu do Pisy dokončil v roce 1202 latinsky psanou knihu *Liber abaci* („Kniha o početním umění“), v níž vysvětlil novou číselnou soustavu a ukázal, jak ji používat pro účetnictví, převody váhy a měny, úrokové sazby a mnoho dalších aplikací. Shromáždil také většinu výsledků algebry a aritmetiky známých Arabům.

Fibonacci se ve své knize zabýval tím, co bychom dnes nazvali problémem populační dynamiky. Ten ale uvedl pouze jako početní cvičení mezi mnoha jinými nesouvisejícími úlohami: předchozí část knihy je o dokonalých číslech, která jsou součtem svých dělitelů, jako například $28 = 14 + 7 + 4 + 2 + 1$, a následující část je problém o rozdělení peněz mezi čtyři osoby, který je ekvivalentní lineární soustavě čtyř rovnic. Zde je překlad latinského textu (pořízený Eduardem Čechem) oné populační úlohy:

> „Kdosi umístil pár králíků na určitém místě, se všech stran ohrazeném zdí, aby poznal, kolik párů králíků se při tom zrodí během roku, jestliže u králíků je tomu tak, že pár králíků přivede na svět měsíčně jeden pár a že králíci počínají rodit ve dvou měsících svého věku.“

Pokud je na začátku prvního měsíce pár novorozených králíků, nebude tento pár po měsíci ještě plodný a na začátku druhého měsíce bude stále jen jeden pár králíků. Tento pár králíků porodí na začátku třetího měsíce další pár, takže celkem budou existovat dva páry. Počáteční pár králíků opět porodí další pár na začátku čtvrtého měsíce. Druhý pár králíků však ještě nebude plodný. Budou tedy jen tři páry králíků.

Pokud použijeme moderní zápis, bude P_n počet párů králíků na začátku měsíce n. Počet párů králíků P_{n+1} v měsíci $n+1$ je součtem počtu párů P_n v měsíci n a počtu nově narozených párů v měsíci $n+1$. Ale pouze páry králíků, které jsou staré alespoň dva měsíce, porodí v měsíci $n+1$ nové páry králíků. Jedná se o páry, které již existovaly v měsíci $n-1$ a jejich počet je P_{n-1}. Takže

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

Jedná se o rekurentní vztah: udává velikost populace v měsíci $n+1$ jako funkci velikosti populace v předchozích měsících. Fibonacci tedy mohl snadno sestavit následující tabulku, kde $1+1=2$, $1+2=3$, $2+3=5$, $3+5=8$ atd.

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

Fibonacci totiž považoval za výchozí stav situaci v měsíci $n=2$. Protože $P_{14} = 144+233 = 377$, získal nakonec 377 párů králíků dvanáct měsíců po svém výchozím bodě. Všiml si, že tato posloupnost čísel může pokračovat donekonečna.

Po roce 1202 Fibonacci napsal několik dalších knih, například *Practica geometriæ* v roce 1220 a *Liber quadratorum* („Kniha čtverců“) v roce 1225. Jeho pověst jej dovedla k setkání s císařem Fridrichem II, který si vědy vážil. V roce 1240 udělila Republika Pisa Fibonaccimu roční důchod. Rok jeho úmrtí není znám.

V následujících staletích byl Fibonacciho králičí problém zapomenut a neměl žádný vliv na vývoj matematických modelů populační dynamiky. Několik vědců se ve svých studiích setkalo se stejnou posloupností čísel, ale neodvolávali se na Fibonacciho ani na žádnou populaci. Několik Keplerových knih obsahuje poznámku, že poměr P_{n+1}/P_n konverguje pro n jdoucí do nekonečna k poměru zlatého řezu $\phi = (1+\sqrt{5})/2$. To je zvláštní případ vlastnosti společné většině populačních modelů: blíží se ke geometrickému růstu (viz kapitoly 3 a 21). V roce 1728 získal Daniel Bernoulli přesný vzorec

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

při studiu obecných rekurentních řad. Kompletní Fibonacciho dílo bylo vydáno v devatenáctém století. Od té doby je možné posloupnost (P_n) nalézt v knihách o rekreační matematice pod názvem Fibonacciho posloupnost.

Je zřejmé, že pro modelování populace králíků nejsou předpoklady vedoucí k Fibonacciho posloupnosti zdaleka reálné: žádná úmrtnost, žádné rozlišení pohlaví atd. Náš zájem o tuto posloupnost v posledních desetiletích v biologii pramení ze skutečnosti, že některé rostliny obsahují struktury, které zahrnují některé z čísel P_n, například 8 a 13 u borových šišek nebo 34 a 55 u slunečnic. Vědecký časopis *The Fibonacci Quarterly* je dokonce celý věnován vlastnostem a aplikacím Fibonacciho posloupnosti!

Další čtení

1. Bernoulli, D.: Observationes de seriebus... *Comment. Acad. Sci. Imp. Petropolitanae* 3, 85–100 (1728/1732) → *Die Werke von Daniel Bernoulli*, Band 2, Birkhäuser, Basel, 1982, 49–64.
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)

Kapitola 2

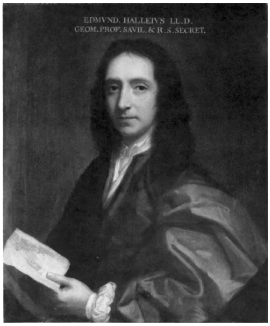
Halleyova úmrtnostní tabulka (1693)

V roce 1693 studoval slavný anglický astronom Edmond Halley záznamy o narozeních a úmrtích ve městě Vratislav, které mu poskytl člen Královské společnosti Caspar Neumann. Sestavil úmrtnostní tabulku, v níž uvedl počet lidí, rozdělených podle předpokládaného věku dožití, z kohorty narozené v jednom roce. Svou tabulku použil také k výpočtu nákladů na doživotní renty. Tato kapitola připomíná tuto práci a zasazuje ji do kontextu Halleyova života a raného vývoje „politické aritmetiky" a teorie pravděpodobnosti, o kterou se zajímali lidé jako Graunt, Petty, De Witt, Hudde, Huygens, Leibniz a de Moivre.

Edmond Halley se narodil v roce 1656 nedaleko Londýna. Jeho otec byl bohatý výrobce mýdla. Edmond se od mládí zajímal o astronomii. Začal studovat na *Queen's College* Oxfordské univerzity. Když byla v roce 1675 slavnostně otevřena Greenwichská observatoř, mohl Halley navštívit královského astronoma Flamsteeda. V letech 1676 až 1678 přerušil studium, aby se vydal na ostrov Svatá Helena a sestavil katalog hvězd, které lze pozorovat z jižní polokoule. Po návratu do Anglie se stal členem Královské společnosti. Zveřejnil také pozorování o cirkulaci větrů, která provedl během své cesty na Svatou Helenu. V roce 1684 navštívil Newtona v Cambridgi, aby s ním diskutoval o souvislosti mezi Keplerovými zákony o pohybu planet a přitažlivou silou, kterou působí Slunce. Podnítil Newtona k napsání slavných *Matematických principů přírodní filozofie*, knihy, kterou nakonec vydal vlastním nákladem. V té době pracoval jako úředník Královské společnosti. V roce 1689 navrhl zvon pro potápění, který sám vyzkoušel.

Přibližně ve stejné době shromažďoval teolog Caspar Neumann, žijící ve Vratislavi, údaje o počtu narozených a zemřelých ve svém městě. Vratislav byla tehdy součástí habsburského impéria a jmenovala se Breslau, nyní se nachází v Polsku a jmenuje se Wrocław. Údaje zahrnovaly i věk, ve kterém lidé umírali. Bylo tedy možné z nich sestavit úmrtnostní tabulku, která ukazovala pravděpodobnost dožití se určitého věku.

První úmrtnostní tabulka byla publikována v Londýně v roce 1662 v knize *Přirozená a politická pozorování založená na seznamech zemřelých*. Tato kniha je obvykle považována za zakládající text pro obory statistiky i demografie. S textem se pojí jedna nevyřešená záhada: lidé se dodnes ptají, zda



Obrázek 2.1:
Halley (1656–1742)

ji napsal londýnský obchodník a autor uvedený na obálce knihy John Graunt, nebo jeho přítel William Petty, jeden ze zakladatelů Královské společnosti.

Úmrtnostní tabulka obsažená v knize využívala zpravodaje, které od počátku 17. století pravidelně informovaly o pohřbech a křtech v Londýně. Tyto zpravodaje sloužily především k informování o opakujících se epidemiích moru. Proto se v nich uváděla příčina úmrtí, ale nikoli věk, v němž lidé zemřeli. Aby Graunt nebo Petty získali tabulku úmrtnosti udávající šanci na přežití v závislosti na věku, museli odhadnout souvislost mezi příčinami úmrtí a věkovými skupinami, proto jejich úmrtnostní tabulka mohla být zatížena velkou chybou. Přesto byla kniha velmi úspěšná a v letech 1662–1676 se dočkala pěti vydání. Několik měst v Evropě začalo vydávat zpravodaje podobné tomu londýnskému.

Téměř třicet let po vydání této první úmrtnostní tabulky Neumann na základě Leibnizova návrhu zaslal Henrymu Justelovi, tajemníkovi Královské společnosti, svá demografická data zratislavi za roky 1687–1691. Justel krátce nato zemřel a data se tak dostala k Halleymu, který je analyzoval a v roce 1693 publikoval své závěry ve *Filozofických zprávách královské společnosti*. Jeho článek se jmenuje *Odhadovaná úmrtnost lidské populace na základě zvláštních tabulek narození a pohřbů ve městě Breslau a pokus o zjištění nákladů na doživotní renty*.

Za sledované období pěti let Halley zjistil, že počet narozených dětí ve Vratislavi se víceméně rovnal počtu zemřelých, takže celkový počet obyvatel byl téměř konstantní. Pro zjednodušení analýzy předpokládal, že populace je přesně v ustáleném stavu: roční počet narozených (nazvěme jej P_0), celkový počet obyvatel, počet obyvatel ve věku k (P_k) a roční počet zemřelých ve věku k (D_k) jsou v průběhu času konstantní. Tím zdůrazňuje další zajímavou vlastnost údajů z Vratislavi, protože takové zjednodušení by nebylo možné u rychle rostoucího města, jakým byl v té době Londýn, ve kterém byly statistiky zkresleny například přílivem obyvatelstva stěhujícího se z venkova.

Tabulka 2.1: Halleyova úmrtnostní tabulka zobrazující velikost populace P_k ve věku 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

Průměrný počet obyvatel narozených za rok ve Vratislavi činil v té době 1 238, a je to také hodnota, kterou Halley použil jako P_0. Dále mohl ze sledovaných dat získat také D_k, tj. roční průměrný počet úmrtí lidí ve věku k pro všechna $k \geqslant 0$. A pomocí vzorce

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

mohl sestrojit tabulku 2.1, která udává počet obyvatel P_k věku k pro $k \geqslant 0$. Zpětně lze zjistit hodnoty D_k, které použil, pomocí vzorce $D_k = P_k - P_{k+1}$: $D_0 = 238$, $D_1 = 145$, $D_2 = 57$, $D_3 = 38$ atd. Ukazuje se, že ve skutečnosti Halley své výsledky trochu upravil, aby získal zaokrouhlená čísla (to je případ D_1, který byl mírně pozměněn tak, že $P_1 = 1\,000$), nebo aby vyhladil určité nepravidelnosti způsobené malým počtem úmrtí ve velmi pokročilém věku v rámci jeho pětileté studie. Součtem všech čísel P_k v tabulce získal Halley odhad celkového počtu obyvatel Vratislavi blížící se 34 000 osob[1]. Tato metoda měla velkou výhodu v tom, že nevyžadovala plošné sčítání lidu, ale pouze znalost počtu narozených a zemřelých a věku, ve kterém lidé během několika posledních let zemřeli.

Na Halleyovu úmrtnostní tabulku se v osmnáctém století odkazovaly různé práce (viz kapitola 4). Ačkoli hodnoty P_k byly specifické pro Vratislav, lze poměr P_{k+1}/P_k obecně považovat za pravděpodobnost dožití se věku $k+1$ za

[1]U osob starších 84 let Halley právě uvedl, že jejich počet je 107.

podmínky, že člověk již dosáhl věku k. Tuto pravděpodobnost by bylo možné rozumně použít i pro populaci jiných evropských měst té doby. Dalo by se například očekávat, že roční dítě bude mít šanci 661 ku 1 000, že dosáhne věku 10 let, nebo šanci 598 ku 1 000, že dosáhne věku 20 let.

Halley také použil svou úmrtnostní tabulku k výpočtu nákladů na doživotní renty. V průběhu šestnáctého a sedmnáctého století prodávala různá města a státy svým občanům tyto renty, aby získaly peníze. Kupující dostávali každý rok až do své smrti pevně stanovenou částku peněz, která se rovnala určitému procentu původně zaplacené částky, často dvojnásobku tehdejší úrokové míry, avšak nezávisle na věku kupujícího. Instituce samozřejmě riskovala bankrot, pokud by si tyto renty koupilo příliš mnoho lidí s velmi dlouhou očekávanou délkou dožití. Tento problém nebylo možné správně řešit bez spolehlivé úmrtnostní tabulky.

Již v roce 1671 se Johan De Witt, předseda nizozemské vlády, a Johannes Hudde, jeden ze starostů města Amsterdamu, zamýšleli nad problémem výpočtu nákladů na doživotní renty. Obávali se vpádu francouzských vojsk, a proto chtěli získat peníze na posílení armády. Měli k dispozici údaje o lidech, kteří si koupili doživotní rentu o několik desetiletí dříve, zejména o věku, v němž byla renta zakoupena, a o věku, v němž lidé zemřeli. Podařilo se jim vypočítat náklady na renty víceméně správně, ale jejich metoda byla později zapomenuta. Následujícího roku bylo Holandsko napadeno a De Witt byl davem zlynčován.

V roce 1693 se Halley zabýval tímto problémem znovu, přičemž použil úmrtnostní tabulku z Vratislavi a předpokládal úrokovou míru 6 %. Metoda výpočtu je jednoduchá. Nechť i je úroková míra. Nechť R_k je cena, za kterou si osoba ve věku k může koupit rentu, řekněme ve výši jedné libry ročně. Tato osoba má pravděpodobnost P_{k+n}/P_k, že bude ve věku $k+n$ ještě naživu. Libru, kterou mu stát slíbí vyplatit, pokud se tohoto věku dožije, lze získat uložením $1/(1+i)^n$ libry z počáteční částky s úrokovou sazbou i. Pokud tedy přijmeme zjednodušující předpoklad, že počáteční částka bude použita pouze na výplatu renty, pak by cena měla být

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

Halley tímto způsobem získal tabulku 2.2, která ukazuje faktor R_k, kterým je třeba vynásobit požadovanou rentu, abychom získali potřebný počáteční vklad. Muž ve věku 20 let by tedy každý rok dostal $1/12{,}78 \approx 7{,}8\,\%$ počáteční částky. Muž ve věku 50 let by však dostal $1/9{,}21 \approx 10{,}9\,\%$, protože by mu zbývalo méně let života. Všimněte si, že dvojnásobná úroková míra by odpovídala rentě ve výši 12 % původní částky, neboli ceně rovnající se 8,33 násobku renty.

Tabulka 2.2: Faktor R_k, kterým je potřeba vynásobit požadovanou rentu, abychom získali potřebný počáteční vklad.

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

Výpočty byly ve své době samozřejmě poměrně zdlouhavé. Halley nicméně mohl použít logaritmické tabulky, aby obecný člen $P_{k+n}/(1+i)^n$ získal rychleji. Protože neuváděl hodnoty P_k nad 84 let, není možné jeho výpočty přesně ověřit. Halleyova práce neměla žádný bezprostřední dopad: po několik desetiletí se v Anglii i jinde nadále prodávaly doživotní renty za cenu nezávislou na věku kupujícího a za cenu, která byla mnohem nižší, než by mohla být, například ve výši sedminásobku požadované renty.

Otázky vyplývající z úmrtnostních tabulek zajímaly v Halleyově době mnoho vědců. Nizozemec Christiaan Huygens, autor první knížky věnované teorii pravděpodobnosti z roku 1657, se v roce 1669 ve své korespondenci s bratrem Grauntem zabýval úmrtnostní tabulkou a výpočtem střední délky života[2].

Několik let předtím, než Neumann navázal kontakt s Královskou společností, napsal Leibniz v eseji, která zůstala nepublikována, o výpočtu průměrné délky života. V roce 1709 obhájil disertaci o teorii pravděpodobnosti Nikolaus I. Bernoulli. V roce 1725 vydal Abraham de Moivre celé *Pojednání o rentách*. Všiml si zejména, že cenu R_k lze snadno vypočítat pro vyšší věk, protože vzorec (2.2) obsahuje jen několik členů. Pak bylo možné použít vzorec zpětné rekurence

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

což lze snadno dokázat na základě (2.2). Pomocí hodnoty, kterou Halley uvádí pro náklady na rentu od 70 let věku, lze tedy ověřit další hodnoty z tabulky 2.2 [3].

Po této přestávce zaměřené na demografii se Halley vrátil ke svým hlavním výzkumným tématům. V letech 1698 až 1700 obeplul Atlantský oceán, aby nakreslil mapu magnetického pole Země. V roce 1704 se stal profesorem na Oxfordské univerzitě. V následujícím roce vydal knihu o kometách

[2]Střední délka života ve věku k je dána vzorcem (2.2) s $i = 0$.

[3]Zdá se, že v tabulce je několik chyb, zejména ve věku 5 a 15 let.

a předpověděl, že kometa z roku 1682, kterou Kepler pozoroval v roce 1607, se vrátí v roce 1758. Stala se známou jako „Halleyho kometa“. Vydal také překlad knihy Apollonia z Pergy o kuželosečkách. V roce 1720 získal funkci královského astronoma po Flamsteedovi. Pokusil se vyřešit problém přesného určení zeměpisné délky na moři na základě pozorování Měsíce, což byl problém s velkým praktickým významem pro navigaci. Zemřel v Greenwichi roku 1742 ve věku 86 let.

Další čtení

1. Fox, M.V.: *Scheduling the Heavens*. Morgan Reynolds (2007)
2. Graunt, J.: *Natural and Political Observations Mentioned in a Following Index and Made upon the Bills of Mortality* (1665). echo.mpiwg-berlin.mpg.de
3. Hald, A.: *A History of Probability and Statistics and Their Applications before 1750*. Wiley, Hoboken, New Jersey (2003).
4. Halley, E.: An estimate of the degrees of the mortality of mankind. *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)

Kapitola 3

Euler, Süssmilch a boží řád (1748–1761)

Euler napsal několik děl zabývajících se populační dynamikou. Jeho pojednání *Introductio in analysin infinitorum* z roku 1748 obsahuje kapitolu o exponenciálách a logaritmech. Čtyři problémy uvedené v této kapitole jsou řešeny prostřednictvím modelu exponenciálního růstu populace. V roce 1760 publikoval článek kombinující tento exponenciální růst s věkovou strukturou populace. Tato práce je předchůdcem teorie „stabilizovaných" populací, která byla rozvinuta ve dvacátém století a hraje důležitou roli v demografii. V roce 1761 pomohl Euler také Süssmilchovi s druhým vydáním jeho pojednání o demografii. Vypracoval zajímavý model, který je jakousi variantou Fibonacciho posloupnosti, ale svou podrobnou analýzu nepublikoval.

Leonhard Euler se narodil v roce 1707 ve švýcarské Basileji. Jeho otec byl protestantský duchovní. V roce 1720 začal Euler studovat na univerzitě. Soukromé hodiny matematiky mu dával také Johann Bernoulli, jeden z nejslavnějších matematiků generace po Leibnizovi a Newtonovi. Přátelil se se dvěma syny Johanna Bernoulliho: Nikolausem II. a Danielem. V roce 1727 se Euler připojil k Danielovi na nově vzniklé Akademii věd v Petrohradě. Kromě matematiky se zajímal také o fyziku a mnoho dalších vědeckých a technických oborů. V roce 1741 ho pruský král Fridrich II. navrhl, aby se stal ředitelem matematické sekce Akademie věd v Berlíně. Euler publikoval značné množství článků a knih o všech aspektech mechaniky (astronomie, pružnost, kapaliny, pevné látky) a matematiky (teorie čísel, algebra, nekonečné řady, elementární funkce, komplexní čísla, diferenciální a integrální počet, diferenciální a parciální diferenciální rovnice, optimalizace, geometrie), ale také o demografii. Byl nejplodnějším matematikem své doby.

V roce 1748 vydal Euler latinsky psané pojednání s názvem *Introductio in analysin infinitorum*. V kapitole o exponenciálách a logaritmech se zabýval šesti problémy: jedním o matematické teorii hudebních stupnic, dalším o splácení půjčky s úroky a čtyřmi dalšími o populační dynamice. V části věnující se populační dynamice Euler předpokládal, že populace P_n v roce n splňuje následující podmínky

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

Obrázek 3.1:
Euler (1707–1783)

pro všechna přirozená n. Míra růstu x je kladné reálné číslo. Vycházíme-li z počátečního stavu P_0, je populace v roce n dána vztahem

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

Tento typ růstu se nazývá geometrický nebo exponenciální. První příklad se ptá:

> „Víme, že počet obyvatel v určitém regionu se ročně zvyšuje o jednu třicetinu a dále víme, že v daný okamžiky zde žilo 100 000 obyvatel. Chceme určit počet obyvatel po 100 letech.“

Řešením je $P_{100} = (1 + 1/30)^{100} \times 100\,000 \approx 2\,654\,874$. Tento problém byl inspirován sčítáním lidu v Berlíně, které proběhlo v roce 1747 a jehož výsledkem byl odhad velikosti populace města na 107 224 obyvatel. Jeho výpočet ukazuje, že počet obyvatel se může během jednoho století zvýšit více než desetkrát. Přesně to se v té době dělo v Londýně.

Je třeba si uvědomit, že výpočet $(1 + 1/30)^{100}$ je s moderní kalkulačkou velmi snadný. V Eulerově době lidé k vyhodnocování obdobných výrazů užívali logaritmy, aby se vyhnuli četným ručním násobením a rychle získali výsledek. Postup k získání výsledku je následující. Nejdříve se vypočítá dekadický logaritmus (logaritmus o základu 10) z P_{100}. Užitím základního pravidla pro počítání s logaritmy $\log(ab) = \log a + \log b$ dostáváme

$$\log P_{100} = 100 \log(31/30) + \log(100\,000) = 100\,(\log 31 - \log 30) + 5 .$$

Logaritmy zavedl v roce 1614 Skot John Napier. Jeho přítel Henry Briggs zveřejnil první tabulku dekadických logaritmů v roce 1617. V roce 1628 Holanďan Adriaan Vlacq doplnil Briggsovu práci a vydal tabulku obsahující dekadické logaritmy celých čísel od 1 do 100 000 s přesností na deset desetinných míst. Právě tuto tabulku použil Euler k získání

$$\log 30 \approx 1{,}477121255, \quad \log 31 \approx 1{,}491361694.$$

Získal tak odhad $\log P_{100} \approx 6{,}4240439$. Zbývalo najít číslo P_{100}, jehož logaritmus je znám. Protože dekadické logaritmy celých čísel od 1 do 100 000 nabývají hodnot od 0 do 5, hledal Euler místo hodnoty P_{100} hodnotu čísla $P_{100}/100$, jehož logaritmus je roven 4,4240439. V tabulce logaritmů lze ověřit, že

$$\log 26\,548 \approx 4{,}424031809 \quad \text{a} \quad \log 26\,549 \approx 4{,}424048168.$$

Aproximací logaritmické funkce přímkou mezi 26 548 a 26 549 Euler získal

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

Takže $P_{100} \approx 2\,654\,874$.

Druhý příklad týkající se populační dynamiky v Eulerově pojednání je následující:

> „Po potopě světa všichni lidé pocházeli z šestičlenné populace. Předpokládáme, že počet obyvatel po dvou stech letech činil 1 000 000. Chceme zjistit roční rychlost růstu populace.“

Protože

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

dostaneme pomocí kalkulačky $x = (10^6/6)^{1/200} - 1 \approx 0{,}061963$. Pomocí tabulek logaritmů musíme z rovnice

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

vyjádřit $\log(1+x)$. Získáme $\log(1+x) = (6 - \log 6)/200 \approx 0{,}0261092$ a tedy $1+x \approx 1{,}061963$. Euler tedy došel k závěru, že počet obyvatel se bude zvyšovat o $x \approx 1/16$ ročně. Abychom znali pozadí, které stálo za volbou tohoto příkladu, musíme si uvědomit, že soudobí filozofové začali popírat pravdivost biblických příběhů. Doslovný výklad bible by určil dobu potopy světa kolem roku 2350 př. n. l. přežili ji Noé s manželkou a také jejich tři synové se svými ženami. Kniha Genesis říká:

> „Ti tři jsou synové Noé, a ti se rozprostřeli po vší zemi.“

Tempo růstu populace $1/16$ (neboli 6,25 %) ročně po potopě světa se Eulerovi nezdálo nereálné. Protože byl synem protestantského pastora a celý život zůstal věřící, dospěl k závěru:

> „Z tohoto důvodu je zcela směšné namítání skeptiků, že za tak krátkou dobu nemohla být celá Země osídlena jediným člověkem“[1].

Euler si také všiml, že kdyby růst pokračoval stejným tempem i 400 let po potopě světa, počet obyvatel by byl $(1+x)^{400} \times 6 = (10^6/6)^2 \times 6 \approx 166$ miliard:

> „Celá Země by však nikdy nebyla schopna takovou populaci uživit.“

Tuto myšlenku o půl století později významně rozvinul Malthus (viz kapitola 5).

Třetí Eulerův příklad řeší otázku:

> „Pokud se každé století lidská populace zdvojnásobí, jaká je roční míra růstu?“

Protože

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

dostaneme pomocí kalkulačky $x = 2^{1/100} - 1 \approx 0{,}00695$. Pomocí tabulek logaritmů $100\log(1+x) = \log 2$. Takže $\log(1+x) \approx 0{,}0030103$ a $1+x \approx 1{,}00695$. Populace tedy roste každý rok o $x \approx 1/144$. Čtvrtý a poslední příklad řeší obdobnou otázku:

> „Pokud se lidská populace ročně zvětšuje o 1/100, chtěli bychom vědět, za jak dlouho se populace zvětší desetkrát.“

Při

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

najdeme $n\log(101/100) = 1$. Takže $n = 1/(\log 101 - 2) \approx 231$ let. To jsou veškeré myšlenky uvedené v *Introductio in analysin infinitorum* z roku 1748,

[1] V knize vydané Grauntem v roce 1662 (viz kapitola 2) najdeme podobnou poznámku:

> „Lidský pár, například Adam a Eva, zdvojnásobí svůj počet každých 64 let z celkových 5 160 let, což je podle Písma svatého stáří našeho světa. Zplodí tedy mnohem více lidí, než kolik jich je na světě nyní. Proto svět není starší než 100 tisíc let, jak si někteří marně představují, ani starší než jej činí Písmo svaté.“

které se týkají populační dynamiky. Euler se k tomuto tématu vrátí důkladněji o několik let později.

V roce 1760 publikoval ve sborníku Akademie věd v Berlíně práci nazvanou *Obecné zkoumání úmrtnosti a množení lidského druhu*. Tato práce byla jakousi syntézou jeho předchozí analýzy geometrického růstu populací a dřívějších studií o úmrtnostních tabulkách (viz kapitola 2). Euler se zabýval například následujícím problémem:

> „Známe počet narození a pohřbů, ke kterým dojde v průběhu jednoho roku. Chceme zjistit velikost populace a její roční přírůstek.“

Euler zde předpokládal, že následující hodnoty jsou známy:

- počet narozených dětí B_n během roku n;
- počet úmrtí D_n během roku n;
- podíl q_k novorozenců, kteří dosáhnou věku $k \geqslant 1$.

Nechť P_n je velikost populace v roce n[2]. Euler učinil dva další implicitní předpoklady:

- populace roste geometricky: $P_{n+1} = r P_n$ (označíme $r = 1 + x$);
- podíl počtu narozených dětí k počtu obyvatel je konstantní: $B_n / P_n = m$.

Z těchto dvou předpokladů vyplývá, že počet narozených dětí roste geometricky se stejnou rychlostí růstu: $B_{n+1} = r B_n$. Euler dále uvažoval stav populace v intervalu sta let, řekněme mezi roky $n = 0$ a $n = 100$, za předpokladu, že žádná osoba nežije déle než sto let. Pro přehlednost nazvěme $P_{k,n}$ ($k \geqslant 1$) počet osob žijících na začátku roku n, které se narodily v roce $n - k$. Počet narozených během roku n nazvěme $P_{0,n} = B_n$. Z definice koeficientu přežití q_k vyplývá, že $P_{k,n} = q_k P_{0,n-k} = q_k B_{n-k}$. Takže

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

[2]Přesněji P_n je velikost populace na konci roku n, která se skládá z osob, které jsou po celý rok n naživu a osob, které se narodí během roku n

Vydělením této rovnice číslem $r^{100} P_0$, dostáváme

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

Tato rovnice se někdy v demografii nazývá „Eulerova rovnice“. Pokud počítáme narození a úmrtí zvlášť, dostaneme následující

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

Počet úmrtí tedy také roste geometrickou řadou: $D_{n+1} = r D_n$. Navíc,

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

Dosazením do rovnice (3.1), dostaneme nakonec rovnici

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

kde zbývá jediná neznámá: r. Taková rovnice se nazývá implicitní, protože nemůžeme vyjádřit r jako funkci ostatních parametrů. Můžeme však vyhodnotit levou a pravou stranu rovnice (3.4) pro pevnou hodnotu r a r můžeme potom postupně měnit, dokud se obě strany nerovnají. Takto získaná hodnota r udává míru růstu populace $x = r - 1$. Všimněte si, že z rovnic (3.1) a (3.3) získáme pro velikost populace P_n následující vyjádření:

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

Pokud uvažujeme konstantní velikost populace ($r = 1$), je tento výraz stejný jako ten, který použil Halley k odhadu populace Vratislavi (viz kapitola 2).

Euler se zabýval také následující otázkou:

> „Je dána míra úmrtnosti a porodnosti. Pokud známe celkový počet žijících osob, chceme pro každý věk určit celkový počet osob daného stáří.“

Protože je známá míra přežití q_k pro věk k a míra porodnosti m, lze z rovnice (3.1) vypočítat míru růstu r. V průběhu roku n je počet lidí narozených v roce $n-k$ roven $q_k B_{n-k} = q_k B_n / r^k$ (pro $q_0 = 1$). Podíl populace ve věku k ku celkové velikosti populace je tedy následující

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

Tento podíl je konstantní. S použitím Lotkovy terminologie (viz kapitola 10) se říká, že populace je „stabilizovaná“: věková pyramida si v čase zachovává stejný tvar.

Euler také zkoumal problém konstrukce úmrtnostní tabulky, když velikost populace není konstantní, ale roste geometricky:

> „Známe počet všech žijících, počet nově narozených a počet zemřelých v každém věku v průběhu jednoho roku. Cílem je určit zákon úmrtnosti.“

Zákonem úmrtnosti myslel Euler množinu koeficientů přežití q_k. Nyní se předpokládá, že celkový počet obyvatel je znám ze sčítání lidu, což nebyl Halleyův případ (viz kapitola 2). Rovnice (3.2) ukazuje, že míra růstu je

$$r = \frac{P_n - D_n}{P_n - B_n}\,.$$

Nechť $D_{k,n}$ je počet lidí, kteří zemřeli ve věku k během roku n: tito lidé se narodili v roce $n-k$. Takže $D_{k,n} = (q_k - q_{k+1})\,B_{n-k}$, kde $B_{n-k} = B_n/r^k$. Koeficienty přežití q_k lze tedy vypočítat pomocí rekurentního vzorce

$$q_{k+1} = q_k - \frac{r^k\,D_{k,n}}{B_n}$$

pro všechna $k \geqslant 0$, přičemž $q_0 = 1$. Tento vzorec vynásobený B_n dává zpět vzorec (2.1) použitý Halleyem pro stacionární případ $r = 1$. Euler nicméně trval na tom, že jeho metoda výpočtu koeficientů přežití q_k předpokládá postupný přírůstek obyvatelstva a zanedbává nenadálé situace jako jsou morové epidemie, války, hladomory apod. Kdyby se při sčítání lidu v Eulerově době zaznamenával věk obyvatel (jako to bylo například ve Švédsku), nebyl by tento předpoklad nutný a koeficienty q_k by se počítaly snadněji.

Díky znalosti koeficientům přežití q_k, dokázal Euler vypočítat náklady na doživotní renty. Práce Halleyho nebo de Moivreho na toto téma však nezmínil. Euler použil úrokovou míru 5 % a úmrtnostní tabulku, kterou v roce 1742 publikoval Holanďan Willem Kersseboom.

Euler nebyl jediným vědcem, který se na berlínské akademii zabýval demografií. Jeho kolega Johann Peter Süssmilch vydal v roce 1741 německy psané pojednání s názvem *Božský řád v proměnách lidského pokolení skrze zrození, úmrtí a rozmnožování*, které je dnes považováno za první pojednání zcela věnované demografii. V roce 1752 napsal Süssmilch také knihu *O rychlém růstu města Berlína*.

Obrázek 3.2:
Süssmilch (1707–1767)

V roce 1761 vydal Süssmilch druhé vydání svého pojednání. V kapitole nazvané „O rychlosti růstu a zdvojnásobení populace“ uvedl zajímavý matematický model, který pro něj vypracoval Euler. Model byl podobný Fibonacciho modelu (viz kapitola 1), ale pro lidskou populaci. Euler vycházel z páru (jeden muž a jedna žena), kterému je v roce 0 20 let, a předpokládal, že lidé umírají ve věku 40 let a žení se ve věku 20 let, přičemž každý pár má šest dětí: dvě děti (chlapce a dívku) ve věku 22 let, další dvě ve věku 24 let a poslední dvě ve věku 26 let. Počítáme-li roky po dvou tak, že B_i je počet porodů během roku $2i$, dospěl Euler k závěru, že

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

pro všechna $i \geqslant 1$. Počáteční podmínky odpovídají $B_{-12} = 0$, $B_{-11} = 0$, $B_{-10} = 2$ a $B_i = 0$ pro $-9 \leqslant i \leqslant 0$. Euler tedy mohl vypočítat počet narozených dětí, jak je uvedeno ve druhém sloupci tabulky 3.1. Počet úmrtí D_i v roce $2i$ se pak rovná počtu narozených v roce $2i-40$: $D_i = B_{i-20}$ pro $i \geqslant 10$, zatímco $D_i = 0$ pro $i \leqslant 9$. Pokud jde o počet P_i lidí žijících v roce $2i$, je roven počtu lidí žijících v roce $2i-2$ plus počet narozených v roce $2i$ minus počet zemřelých v roce $2i$: $P_i = P_{i-1} + B_i - D_i$.

Tato kapitola Süssmilchovy knihy končí poznámkou, která již mohla být vyslovena o Fibonacciho posloupnosti:

> „Velká neuspořádanost, která v Eulerově tabulce zdánlivě převládá, nebrání tomu, aby se počet nově narozených postupně vyvíjel prostřednictvím rekurentních posloupností [...] Jakkoli tento vývoj působí ze začátku neuspořádaně, mění se postupně v geo-

Tabulka 3.1: Eulerova tabulka.

i	Narození	Úmrtí	Populace
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
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
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

> metrickou posloupnost, pokud nenastane nějaké nenadálá událost, neuspořádanost postupně mizí až téměř úplně vymizí.“

Kniha neuvádí více o matematice tohoto populačního modelu. Euler však tuto studii posunul mnohem dále v rukopise nazvaném *O množení lidského rodu*, který zůstal za jeho života nepublikován. Při hledání řešení rovnice (3.5) ve tvaru $B_i = c\,r^i$, tj. ve tvaru geometrické posloupnosti, získal po zjednodušení polynomiální rovnici stupně 13. V této rovnici se objevilo řešení $B_i = c\,r^i$:

$$r^{13} = r^2 + r + 1\,. \tag{3.6}$$

Hledal řešení blízké hodnotě $r = 1$ a pomocí tabulky logaritmů pro výpočet r^{13} zjistil, že

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

Rovnice (3.6) má tedy kořen někde mezi 1,09 a 1,10. Aproximací grafu funkce $1 + r + r^2 - r^{13}$ úsečkou na tomto intervalu získal Euler následující výsledek

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

Protože se roky počítají po dvou, počet narozených se každý rok násobí $\sqrt{r}$. Toto číslo se zdvojnásobí každých n let, pokud $(\sqrt{r})^n = 2$, tj. každých $n = 2\log 2/\log r \approx 15$ let. Protože asymptoticky $B_i \approx c\,r^i$ a protože počet úmrtí D_i v roce $2i$ je roven B_{i-20}, dostaneme $D_i \approx B_i/r^{20}$ s $r^{20} \approx 6{,}25$. Počet narozených je přibližně šestkrát větší než počet zemřelých. Počet P_i lidí žijících v roce $2i$ je roven $B_i + B_{i-1} + \cdots + B_{i-19}$, a proto také dostaneme, že

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

Celkový počet obyvatel je asi desetkrát vyšší než počet narozených.

Důkaz, že posloupnost (B_i) uvedená v tabulce 3.1 skutečně roste asymptoticky jako r^i, je složitější. Již od prací Abrahama de Moivra o rekurentních řadách bylo známo, že zavedením „vytvořující funkce“

$$f(x) = \sum_{i=0}^{+\infty} B_i\,x^i\,,$$

lze vyjádřit $f(x)$ jako racionální funkci. Euler tuto metodu vysvětlil ve svém *Introductio in analysin infinitorum* v roce 1748: rekurentní vztah

(3.5) dává skutečně

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

Takže

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

Euler věděl, že takovou racionální funkci lze rozložit ve tvaru

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

čísla $x_1, \ldots, x_{13}$ jsou reálné nebo komplexní kořeny rovnice

$$1 - x^{11} - x^{12} - x^{13} = 0.$$

Takže

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

Protože B_i je koeficient x^i v $f(x)$, Euler získal, že

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

pro $i \to +\infty$, kde x_k je v absolutní hodnotě nejmenší kořen. Jinými slovy, B_i má tendenci růst geometricky jako $(1/x_k)^i$. Zbývá poznamenat, že x_k je kořenem rovnice $1 - x^{11} - x^{12} - x^{13} = 0$ tehdy a jen tehdy, když $r = 1/x_k$ je kořenem rovnice (3.6). Některé detaily důkazu nakonec objasnil Gumbel v roce 1916.

Süssmilch vydal třetí vydání svého pojednání v roce 1765 a zemřel v Berlíně v roce 1767. Euler se v roce 1766 vrátil do Petrohradu, protože měl špatné vztahy s pruským králem. Přestože ztratil zrak, pokračoval s pomocí svých synů a kolegů v publikování velkého množství prací, zejména o algebře, integrálním počtu, optice a stavbě lodí. Jeho *Dopisy na různá témata přírodní filosofie psané pro německou princeznu*, napsané v Berlíně v letech 1760 až 1762, vyšly v letech 1768 až 1772 a staly se bestsellerem v celé Evropě. Euler zemřel v Petrohradě v roce 1783. Jeho přínos k matematické demografii, zejména jeho analýza „stabilizované“ věkové pyramidy v exponenciálně rostoucí populaci, byl znovuobjeven až ve dvacátém století (viz kapitoly 10 a 21).

Další čtení

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, Tom. prim.* (1748) $\rightarrow$ *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. dtsch. Math. Ver*. 25, 251–264 (1917). digizeitschriften.de
6. Reimer, K.F.: Johann Peter Süssmilch, seine Abstammung und Biographie. *Arch. soz. Hyg. Demogr*. 7, 20–28 (1932)
7. Rohrbasser, J.M.: Johann Peter Süssmilch. In: Heyde, C.C., Seneta, E. (eds.) *Statisticians of the Centuries*, 72–76. Springer (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)

Kapitola 4

Daniel Bernoulli, d'Alembert a očkování proti neštovicím (1760)

V roce 1760 napsal Daniel Bernoulli článek, ve kterém modeloval populaci ovlivněnou neštovicemi. V jeho době se vedly velké spory o očkování, které mohlo lidi ochránit, ale mohlo být i smrtelně nebezpečné. Použil Halleyovu tabulku přežívání a některé údaje týkající se neštovic, aby ukázal, že očkování je výhodné, pokud je související riziko úmrtí menší než 11 %. Očkování mohlo prodloužit očekávanou délku života při narození až o tři roky. D'Alembert Bernoulliho práci, která byla prvním matematickým modelem v epidemiologii, kritizoval.

Daniel Bernoulli se narodil v roce 1700 v nizozemském Groningenu. V jeho rodině již byli dva slavní matematici: otec Johann Bernoulli a strýc Jakob Bernoulli. V roce 1705 se Johann přestěhoval do Basileje ve Švýcarsku, kde nastoupil na místo profesora uvolněné Jakobovou smrtí. Johann si nepřál, aby jeho syn studoval matematiku. Daniel se proto zaměřil na medicínu a roku 1721 získal doktorský titul za práci o dýchání. Přestěhoval se do Benátek, kde se začal věnovat matematice a v roce 1724 vydal knihu. V témže roce získal cenu Pařížské akademie věd za esej *O dokonalosti přesýpacích hodin na lodi na moři* a získal profesuru na nové Petrohradské akademii. V těchto letech pracoval zejména na rekurentních řadách nebo na „petrohradském paradoxu" v teorii pravděpodobnosti. V roce 1733 se Daniel Bernoulli vrátil na univerzitu v Basileji, kde vyučoval postupně botaniku, fyziologii a fyziku. Roku 1738 vydal knihu o dynamice tekutin, která v dějinách fyziky zůstala slavnou. Kolem roku 1753 se současně s Eulerem a d'Alembertem začal zajímat o problém vibrujících strun, který vyvolal významnou matematickou polemiku.

V roce 1760 předložil Akademii věd v Paříži práci nazvanou *Pokus o novou analýzu úmrtnosti způsobené neštovicemi a výhod očkování pro jejich prevenci*. Šlo o to, zda by očkování (dobrovolné vpravení malého množství méně virulentních neštovic do těla za účelem ochrany před pozdějšími infekcemi) mělo být podporováno, i když je to někdy smrtelně nebezpečný zákrok. Tato technika byla v Asii známá již velmi dlouho a do Anglie ji přinesla v roce 1718 lady Montaguová, manželka britského velvyslance v Osmanské říši. Ve

Obrázek 4.1:
Daniel Bernoulli (1700–1782)

Francii se i přes úmrtí nejstaršího syna Ludvíka XIV. na neštovice v roce 1711 o očkování uvažovalo neochotně. Voltaire, který neštovice v roce 1723 přežil a který žil několik let v anglickém exilu a sledoval nejnovější inovace, se v roce 1734 ve svých *Filosofických listech* pro očkování vyslovil. Francouzský vědec La Condamine, který rovněž přežil neštovice, se pro očkování vyslovil v roce 1754 v pařížské Akademii věd.

Než Maupertuis v roce 1759 v Basileji zemřel, podnítil Daniela Bernoulliho, aby problém očkování studoval z matematického hlediska. Přesněji řečeno, šlo o to najít způsob, jak porovnat dlouhodobý přínos očkování s okamžitým rizikem úmrtí. Za tímto účelem Bernoulli přijal následující zjednodušující předpoklady:

- lidé nakažení neštovicemi poprvé umírají s pravděpodobností p (nezávisle na věku) a přežívají s pravděpodobností $1-p$;
- každý má pravděpodobnost q, že se každý rok nakazí; přesněji řečeno, pravděpodobnost, že se daný jedinec nakazí mezi věkem x a věkem $x+dx$, je $q\,dx$, kde dx je velmi krátké časové období;
- lidé, kteří přežili neštovice, jsou po zbytek života chráněni proti novým infekcím (jsou chráněni získanou imunitou).

Nechť $m(x)$ je úmrtnost ve věku x z jiných příčin než neštovic: pravděpodobnost, že daný jedinec zemře ve velmi krátkém časovém období dx mezi věkem x a věkem $x+dx$, je $m(x)\,dx$. Uvažujeme-li skupinu P_0 lidí narozených ve stejném roce, označme

- $S(x)$ počet „vnímavých“ lidí[1], kteří jsou ve věku x stále naživu, aniž by se kdy nakazili neštovicemi;

- $R(x)$ počet lidí, kteří stále žijí ve věku x a kteří přežili neštovice;

- $P(x) = S(x) + R(x)$ celkový počet lidí žijících ve věku x.

Narození odpovídá věku $x = 0$. Takže $S(0) = P(0) = P_0$ a $R(0) = 0$. Při použití analytických metod, které na konci 17. století vyvinuli Newton, Leibniz a později i jeho otec Johann, si Daniel Bernoulli všiml, že mezi věkem x a věkem $x + dx$ (přičemž dx je nekonečně malé) má každý vnímavý jedinec pravděpodobnost $q\,dx$, že se nakazí neštovicemi, a pravděpodobnost $m(x)\,dx$, že zemře z jiných příčin. Variace počtu vnímavých osob je tedy $dS = -S\,q\,dx - S\,m(x)\,dx$, což vede k diferenciální rovnici

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

V této rovnici je dS/dx derivací funkce $S(x)$. Během stejného malého časového intervalu je počet lidí, kteří zemřou na neštovice, $p\,S\,q\,dx$ a počet lidí, kteří neštovice přežijí, je $(1-p)\,S\,q\,dx$. Navíc existuje $R\,m(x)\,dx$ lidí, kteří během dx zemřou z jiných příčin než na neštovice. To vede k druhé diferenciální rovnici:

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

Sečtením těchto dvou rovnic získáme

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

Řešením rovnic (4.1) a (4.3) mohl Bernoulli ukázat, že podíl lidí, kteří jsou ve věku x stále vnímaví, činí

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

Pro získání vzorce (4.4) odstranil Bernoulli $m(x)$ z rovnic (4.1) a (4.3):

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

[1] Přesněji řečeno, je to střední hodnota tohoto počtu, která se může měnit průběžně a ne jen o jednotky.

Po přeuspořádání této rovnice dostáváme

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

Všimněte si, že levá strana je derivací funkce $f(x) = S(x)/P(x)$, což je podíl vnímavých osob v populaci ve věku x. Takže

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

Řešení tohoto typu diferenciální rovnice bylo známo již několik desetiletí díky práci Jakoba Bernoulliho, Danielova strýce. Když tuto rovnici vydělíme f^2 a položíme $g(x) = 1/f(x)$, zjistíme, že

$$\frac{dg}{dx} = qg - pq$$

a že $g(0) = 1/f(0) = 1$. Pokud označíme $h(x) = g(x) - p$, máme $dh/dx = qh$. Takže $h(x) = h(0)\,e^{qx} = (1-p)\,e^{qx}$. Konečně $g(x) = (1-p)\,e^{qx} + p$ a $f(x) = 1/g(x)$.

Pro aplikaci své teorie Bernoulli použil Halleyovu tabulku přežívání (viz kapitola 2). Tato tabulka udává počet lidí, kteří ještě žijí na začátku roku x (s $x = 1, 2...$) z kohorty 1238 narozených během roku 0. V rámci svého modelu však Bernoulli potřeboval počet lidí $P(x)$, kteří skutečně dosáhnou věku x, což je poněkud odlišné. Protože si Bernoulli – stejně jako většina jeho současníků – nebyli tohoto rozdílu vědomi (Halleyův článek v tomto skutečně není příliš určitý), ponechal čísla v Halleyově tabulce kromě prvního čísla 1238, které nahradil číslem 1300, aby získal realistickou úmrtnost během prvního roku života. Tato čísla jsou uvedena ve druhém sloupci tabulky 4.1.

Bernoulli zvolil pro pravděpodobnost úmrtí na neštovice hodnotu $p = 1/8 = 12{,}5\,\%$, která je v souladu s tehdejšími pozorováními. Roční pravděpodobnost onemocnění neštovicemi q nebylo možné odhadnout přímo. Bernoulli tedy pravděpodobně vyzkoušel několik hodnot pro q a nakonec zvolil takovou, aby počet úmrtí na neštovice po všech níže uvedených výpočtech činil přibližně 1/13 celkového počtu úmrtí, což byl podíl, který byl tehdy pozorován v několika evropských městech. Ukázalo se, že dobře odpovídá volba $q = 1/8$ za rok, jak nyní uvidíme[2].

[2] To, že p a q jsou stejné, je jen náhoda.

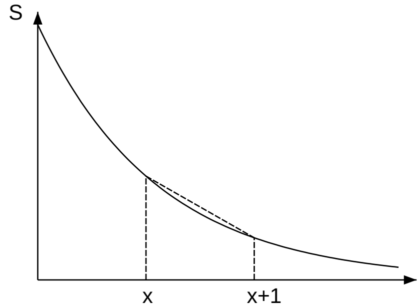
Tabulka 4.1: Halleyova tabulka přežívání a Bernoulliho výpočty.

Věk x	Živí $P(x)$	Vnímaví $S(x)$	Imunní $R(x)$	Úmrtí na neštovice	Bez neštovic $P^*(x)$
0	1300	1300	0	17,2	1300
1	1000	896	104	12,3	1015
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
⋮	⋮	⋮	⋮	⋮	⋮

Pomocí vzorce (4.4) a hodnot $P(x)$ ve druhém sloupci tabulky můžeme vypočítat počet $S(x)$ náchylných osob ve věku x: jedná se o třetí sloupec tabulky zaokrouhlený na nejbližší celé číslo. Čtvrtý sloupec udává počet $R(x) = P(x) - S(x)$ lidí ve věku x, kteří neštovice přežili. V pátém sloupci je v řádku odpovídajícím věku x uveden počet úmrtí na neštovice mezi věkem x a věkem $x+1$. Teoreticky by tento počet měl být integrálem

$$pq \int_x^{x+1} S(t)\, dt\,,$$

ale vzorec $pq[S(x)+S(x+1)]/2$ poskytuje dobrou aproximaci, jak je naznačeno na obrázku 4.2: plocha lichoběžníku se blíží ploše pod křivkou, tj. integrálu funkce.



Obrázek 4.2: Plocha čárkovaného lichoběžníku aproximuje integrál funkce S mezi x a $x+1$.

Bernoulli si všiml, že součet všech čísel v pátém sloupci dává 98 úmrtí na neštovice do věku 24 let. Kdybychom v tabulce pokračovali pro starší věk, našli bychom mezi 32 lidmi, kteří jsou ve věku 24 let ještě vnímaví, jen tři další úmrtí na neštovice. Souhrnně lze říci, že z 1300 narozených je osudem 101 lidí zemřít na neštovice. To je téměř přesně očekávaný zlomek 1/13.

Bernoulli pak uvažoval o situaci, kdy by neštovice byly naočkovány všem při narození a nezpůsobily by žádné úmrtí. Neštovice by byly vymýceny a otázkou je, jak by se zvýšila očekávaná délka života. Vycházíme-li ze stejného počtu narozených P_0, nazvěme $P^*(x)$ počet lidí ve věku x, když neštovice vymizely. Pak

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

Bernoulli mohl ukázat, že

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

kde $P(x)$ je, jak je uvedeno výše, populace ve věku x při výskytu neštovic.

Pokud totiž mezi rovnicemi (4.6) a (4.3) vyloučíme stejně jako dříve $m(x)$, dostaneme po přeuspořádání

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

Bernoulli označil $h(x) = P(x)/P^*(x)$. Použil vztah (4.4) a vynásobil čitatele a jmenovatele číslem e^{-qx}, čímž získal výsledek

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

což je ekvivalentní rovnici

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

kde ln zde znamená přirozený logaritmus a nikoliv desítkový logaritmus. Ale $h(0) = 1$. Takže $h(x) = 1-p+pe^{-qx}$.

Všimněte si, že poměr $P(x)/P^*(x)$ směřuje při dostatečně vysokém věku x k hodnotě $1-p$. Šestý sloupec tabulky 4.1 ukazuje $P^*(x)$. Způsob, jak porovnat $P(x)$ a $P^*(x)$, je odhadnout očekávanou délku života při narození, jejíž teoretické vyjádření v případě přítomnosti neštovic je

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

Analogický výraz s $P^*(x)$ nahrazujícím $P(x)$ platí při absenci neštovic v populaci. Bernoulli použil přibližný vzorec

$$\left[\tfrac{1}{2}P(0)+P(1)+P(2)+\cdots\right]/P_0,$$

který je dán lichoběžníkovou metodou (obr 4.2). Pokračováním v tabulce za hodnotu 24 let až do 84 let (viz tabulka 2.1) získal nakonec očekávanou délku života E s neštovicemi rovnou $\left[\tfrac{1}{2}1300+1000+\cdots+20\right]/1300 \approx 26{,}57$ let, tj. 26 let a 7 měsíců. Bez neštovic pak získal očekávanou délku života E^* rovnající se $\left[\tfrac{1}{2}1300+1015+\cdots+23\right]/1300 \approx 29{,}65$ let, tj. 29 let a 8 měsíců. Očkování při narození by prodloužilo očekávanou délku života o více než tři roky.

Můžeme si všimnout, že pro získání odvozených vztahů existuje jednodušší a rychlejší metoda než ta, kterou použil Bernoulli. Vyjdeme-li z diferenciální rovnice (4.1) pro $S(x)$, vidíme nejprve, že

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

Použitím tohoto výrazu v rovnici (4.2) pro $R(x)$ zjistíme, že

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

Rovnice (4.6) pro $P^*(x)$ ukazuje, že

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

Vztahy (4.4) a (4.7) tak plynou okamžitě!

Očkování méně virulentním kmenem neštovic samozřejmě není zcela bezpečné. Je-li p' pravděpodobnost úmrtí na neštovice těsně po očkování ($p' < p$), pak by očekávaná délka života byla $(1-p')\,E^*$, pokud by každý prošel očkováním při narození. Tato očekávaná délka života zůstává vyšší než „přirozená“ očekávaná délka života E, pokud $p' < 1 - E/E^*$ nebo přibližně 11 %. Údaje o p' bylo v té době obtížné získat. Bernoulli však odhadoval, že riziko p' je menší než 1 %. Pro něj nebylo pochyb: očkování by mělo být podpořeno stát. Došel k závěru:

> „Přeji si jen, aby v záležitosti, která se tak úzce dotýká blaha lidského rodu, nebylo přijato žádné rozhodnutí bez všech znalostí, které může poskytnout trocha analýzy a výpočtů.“

Bernoulliho práce byla přednesena na zasedání Akademie věd v Paříži v dubnu 1760. V listopadu pak přednesl d'Alembert komentář s názvem *O aplikaci teorie pravděpodobnosti na očkování proti neštovicím*. Tento komentář byl krátce poté publikován ve druhém svazku jeho *Opuscules mathématiques* s podrobnějšími výpočty a spolu s další prací nazvanou *Matematická teorie očkování*. D'Alembert kritizoval Bernoulliho předpoklady, že pravděpodobnost nákazy a pravděpodobnost úmrtí na neštovice nezávisejí na věku. Navrhl jiné řešení, které tyto předpoklady nevyžaduje. Nazvěme $v(x)$ úmrtnost na neštovice ve věku x, $m(x)$ úmrtnost z jiných příčin a $P(x)$ počet lidí, kteří ještě žijí. Pak

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

Porovnáním s rovnicí (4.3) zjistíme, že $v(x) = p\,q\,S(x)/P(x)$. Zde dostáváme

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

kde $P^*(x)$ reprezentuje počet lidí žijících ve věku x, pokud neštovice vymizely.

Obrázek 4.3:
D'Alembert (1717–1783)

Můžeme totiž buď dosadit funkci $m(x)$ z rovnice (4.6) do rovnice (4.9), nebo použít vzorec (4.8) pro $P^*(x)$ a uvědomit si, že řešení rovnice (4.9) je dáno vztahem

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

Vzorec (4.10) uvedený d'Alembertem není v rozporu s Bernoulliho vzorcem (4.7). Používá pouze jiný typ informace $v(x)$, která v té době nebyla k dispozici, protože úmrtní matriky obsahovaly příčinu smrti, ale ne věk oběti. D'Alembert tvrdil, že nelze skutečně dospět k závěru, zda je očkování užitečné, dokud nebude tento typ údajů k dispozici.

D'Alembert také kritizoval užitečnost očekávané délky života jako kritéria pro rozhodování, protože dává stejnou váhu všem létům, ať už v blízké nebo vzdálené budoucnosti. Všiml si, že z hlediska jednotlivce nebo státu nemají všechny roky stejnou „užitečnost“, přičemž mladý a starý věk jsou méně hodnotné než věk střední. Přes všechny tyto výtky se d'Alembert vyslovil pro očkování.

Kvůli prodlevám v publikování vyšla Bernoulliho práce až v roce 1766, zatímco d'Alembertovi se podařilo svou práci vydat velmi rychle. Bernoulli vyjádřil svou hořkost v dopise Eulerovi:

> Co říkáte na obrovské banality velkého d'Alemberta o pravděpodobnostech: protože se mi zdá, že se mi v jeho publikacích příliš často křivdí, rozhodl jsem se už před časem, že už nebudu číst nic, co pochází z jeho pera. Toto rozhodnutí jsem učinil u příležitosti rukopisu o očkování, který jsem před osmi lety poslal Akademii v Paříži a který byl velmi ceněn pro novost použité analýzy. Troufám si říci, že to bylo jako začlenění nové provincie do korpusu matematiky. Zdá se, že úspěch této nové analýzy mu způsobil bolest srdce. Kritizoval ji na tisíc způsobů, všechny stejně směšné, a poté, co ji dobře zkritizoval, se vydává za prvního autora teorie, o níž neslyšel ani zmínku. Věděl však, že můj rukopis může vyjít až po nějakých sedmi či osmi letech. Mohl o něm vědět pouze jako člen Akademie. V tomto ohledu měl můj rukopis zůstat nedotknutelný, dokud nebyl zveřejněn. *Dolus an virtus quis in hoste requirat?*[3]

Navzdory pracím Bernoulliho a d'Alemberta se očkování ve Francii neprovádělo ve velkém měřítku. Král Ludvík XV. zemřel na neštovice v roce 1774. Dvorní lékaři však krátce naočkovali zbytek královské rodiny. Problém ztratil na významu, když Edward Jenner objevil, že očkování lidí kravskými neštovicemi (tzv. „vakcinace“) chrání před neštovicemi a je bezpečné. Jeho práce *Vyšetřování příčin a účinků vakcíny variolae* byla publikována v roce 1798. Vakcinace se rychle rozšířila po celé Evropě. Nicméně metody vyvinuté pro počítání zvýšení očekávané délky života, pokud se odstraní jedna příčina smrti, se používají dodnes.

V následujících desetiletích byly konečně k dispozici údaje o věku, ve kterém lidé na neštovice umírali. Problém pak přehodnotila řada vědců, zejména

- Johann Heinrich Lambert, matematik z berlínské akademie, v roce 1772;
- Emmanuel-Étienne Duvillard, v té době pověřený vedením statistiky obyvatelstva na Ministerstvu vnitra v Paříži, ve své *Analýze a tabulkách vlivu neštovic na úmrtnost v jednotlivých věkových kategoriích* (1806);

[3] „Zdali to chrabrost či lest – zda ptá se kdo ve válce na to?“ Vergilius, *Aeneis*, kniha II.

- Pierre-Simon Laplace ve své *Analytické teorii pravděpodobnosti* (1812).

Duvillard a Laplace například ukázali, jak upravit vzorec (4.7), když parametry p a q závisí na věku:

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

Zde $p(x)$ je pravděpodobnost úmrtí na neštovice, pokud se nakazíte ve věku x, a $q(x)$ je pravděpodobnost nakažení se neštovicemi ve věku x.

Po této práci o neštovicích se Daniel Bernoulli nezabýval žádným dalším problémem populační dynamiky. Zemřel v Basileji v roce 1782. D'Alembert zemřel v Paříži o rok později.

Další čtení

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

Kapitola 5

Malthus a potíže s geometrickým růstem (1798)

V roce 1798 Malthus publikoval *Esej o principu populace*, ve kterém tvrdil, že zásobování potravinami nemůže dlouhodobě sledovat exponenciální růst lidské populace. Pokud populace zůstávala relativně konstantní, bylo to proto, že velká část lidstva trpěla nedostatkem potravin. Malthus považoval „princip populace“ za argument proti spisům Godwina a Condorceta, kteří zdůrazňovali pokrok v lidských společnostech. Malthusova esej ovlivnila Darwinovu a Wallaceovu evoluční teorii a byla kritizována Marxem, do praxe však byla uvedena čínskou politikou jednoho dítěte.

Thomas Robert Malthus se narodil v roce 1766 nedaleko Londýna jako šesté ze sedmi dětí. Jeho otec, přítel a obdivovatel Jeana-Jacquese Rousseaua, byl jeho prvním učitelem. V roce 1784 začal mladý Malthus studovat matematiku na univerzitě v Cambridgi. Diplom získal v roce 1791, v roce 1793 se stal členem *Jesus College* a v roce 1797 anglikánským knězem.

Obrázek 5.1:
Malthus (1766–1834)

V roce 1798 vydal Malthus anonymně knihu s názvem *Esej o principu populace, jak ovlivňuje budoucí zlepšení společnosti, s poznámkami ke spekulacím pana Godwina, pana Condorceta a dalších autorů*. Byla to reakce na Godwinovo *Zkoumání společenské spravedlnosti* (1793) a Condorcetovu *Náčrt historického obrazu pokroku lidského ducha* (1794). Navzdory hrůzám, které ve jménu pokroku napáchala Francouzská revoluce, oba autoři tvrdili, že pokrok společnosti je nevyhnutelný. Malthus stejný optimismus nesdílel. Tvrdil také, že anglické chudinské zákony, které pomáhaly chudým rodinám s mnoha dětmi, podporovaly růst populace, aniž by podporovaly obdobně rychlý růst produkce potravin. Zdálo se mu, že tyto zákony chudým život ve skutečnosti neulehčují, spíše naopak. Obecněji řečeno, protože populace měla tendenci růst vždy rychleji než produkce potravin, zdálo se, že část společnosti je odsouzena k bídě, hladu nebo epidemiím: to jsou pohromy, které zpomalují růst populace a které jsou podle Malthusova názoru hlavními překážkami pokroku společnosti. Všechny teorie slibující pokrok by byly jen utopické. Tyto myšlenky vedly Malthuse k vydání jeho knihy v roce 1798. Na následujícím úryvku je vidět, jak shrnul své teze:

> [...] „síla populace je nekonečně větší než síla Země, která pro člověka produkuje potravu. Popukace, je-li nekontrolovaná, narůstá geometrickou řadou. Potraviny rostou jen aritmetickou řadou. Už povrchní pohled na čísla ukáže neúměrnost první síly v poměru ke druhé. Zákon přírody, jenž produkuje jídlo nutné k obživě člověka, ale rovněž určuje, že účinky těchto dvou nestejných sil – růst populace i růst potravinových zdrojů – musí být udržovány v rovnováze. To v sobě zahrnuje účinnou a trvalou kontrolu množství populace v souladu s obtížemi, jež nastávají při získávání potravy. Tato potíž musí dopadnout kamkoliv a musí být vážně pociťována velkou částí lidstva.“

Malthusova kniha byla velmi úspěšná. Obsahovala jen málo dat. Malthus si například všiml, že počet obyvatel USA se během 18. století každých pětadvacet let zdvojnásobil. Sám se nepokusil převést své teze do matematických modelů, ale připravil půdu pro pozdější práce Adolpha Queteleta a Pierra-Françoise Verhulsta, kterým bude věnována následující kapitola.

Po vydání své knihy cestoval Malthus s přáteli nejprve do Německa, Skandinávie a Ruska, poté do Francie a Švýcarska. Když dal dohromady informace získané během svých cest, vydal pod svým jménem v roce 1803 velmi rozšířené druhé vydání s jiným podtitulem: *Esej o principu populace neboli pohled na její minulý a současný vliv na lidské štěstí, se zkoumáním našich vyhlídek na budoucí odstranění nebo zmírnění zla, které způsobuje*. Toto nové vydání se podrobně zabývalo překážkami růstu populace v různých zemích: opožděné sňatky, potraty, infanticida (záměrné usmrcení dítěte), hladomor, válka, epidemie, ekonomické faktory... Pro Malthuse bylo odložení sňatků nejlepší možností, jak stabilizovat populaci. Následovala čtyři další vydání knihy v letech 1806, 1807, 1817 a 1826. V roce 1805 se Malthus stal profesorem historie a politické ekonomie na nové škole, kterou pro své zaměstnance zřídila Západoindická společnost. Vydal také knihu *Zkoumání povahy a pokroku renty* (1815) a *Principy politické ekonomie* (1820). V roce 1819 byl Malthus zvolen členem Královské společnosti. V roce 1834 byl jedním ze zakládajících členů Statistické společnosti. Téhož roku zemřel poblíž Bathu.

Malthusova práce měla silný vliv na vývoj evoluční teorie. Charles Darwin, který se vrátil ze své cesty na lodi *Beagle*, si v roce 1838 přečetl Malthusovu knihu o růstu populace. Toto napsal v úvodu své slavné knihy *O vzniku druhů přírodním výběrem*, vydané v roce 1859:

> „V další kapitole pojednáme o boji o přežití mezi všemi organismy celého světa, který je nevyhnutelný kvůli jejich množení geometrickou řadou. Jde o Malthusovo učení, uplatněné na celou živočišnou a rostlinou říši.“

Alfred Russel Wallace, který vytvořil evoluční teorii ve stejné době jako Darwin, také uvedl, že jeho myšlenky vznikly po přečtení Malthusovy knihy.

Opačný názor na význam Malthusovy knihy měl Karl Marx, jak se lze dočíst v poznámce pod čarou v jeho knize *Kapitál*:

> „Pokud si čtenář pamatuje na Malthuse, jehož spis *Essay on Population* se objevil v roce 1798, tak si já pamatuji, že tento spis ve své první formě nebyl ničím jiným než školácky povrchním a otrocky opsaným plagiátem ze spisů, které napsal Defoe, Sir James Steuart, Townsend, Franklin, Wallace, a další, a neobsahoval ani jednu původní myšlenku. Ten velký rozruch, který tento pamflet vzbudil, vznikl pouze kvůli stranickým zájmům. Francouzská revoluce si v britském království našla horlivé zastánce; ten ‚populační princip‘, který byl pomalu vypracován v 18. století, byl poté uprostřed velké společenské krize roztrubován jako neomylný protijed proti učení Condorceta, a dalších, anglická oligarchie ho nadšeně vítala jako prostředek pro potlačení všech choutek na nějaký další lidský rozvoj. Malthus, svým úspěchem nanejvýš zaskočen, se pak začal věnovat tomu, že ten povrchně sebraný materiál podle starého schematu vycpával a přidával nový materiál – nikoli však vytvořený Malthusem, nýbrž přebraný “

Malthusovy teze jistě nebyly zcela nové. Například myšlenka, že populace má tendenci růst geometrickou řadou, je často připisována právě jemu[1],

[1] R. A. Fisher (viz kapitoly 14 a 20) by nazval „malthusiánským parametrem“ rychlost růstu populace. Malthus se ve své knize zmiňuje o Süssmilchově pojednání.

i když jsme v kapitole 3 viděli, že tuto myšlenku znal již Euler o půl století dříve. Malthus ji však zpopularizoval tím, že ji polemicky spojil se skutečnými legislativními problémy. Paradoxně právě v komunistické Číně našel Malthusův návrh na omezení porodnosti své nejvýraznější uplatnění (viz kapitola 25).

Další čtení

1. Condorcet: *Esquisse d'un tableau historique des progrѷs de l'esprit humain.* Agasse, Paris (1794) gallica.bnf.fr; česky: *Náčrt historického obrazu pokroků lidského ducha*. Praha (1968).
2. Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.* John Murray, London (1859) darwin-online.org.uk; česky: *O vzniku druhů přírodním výběrem.* Academia, Praha (2007)
3. Godwin, W.: *An Enquiry Concerning Political Justice* (1793). archive.org
4. Malthus, T.R.: *An Essay on the Principle of Population.* London (1789). econlib.org; česky: *Esej o principu populace.* „Zvláštní vydání...“, Brno (2002)
5. Marx, K.: *Das Kapital, Erster Band*. In: Marx, K., Engels, F., *Werke, Band 23*, Dietz Verlag, Berlin (1962); česky: *Kapitál*, Praha (1978) joza.pytel.sweb.cz
6. Simpkins, D.M.: Malthus, Thomas Robert. In: Gillespie, C.C. (ed.) *Dictionary of Scientific Biography*, vol. 9, 67–71. Scribner, New York (1974)

Kapitola 6

Verhulst a logistická rovnice (1838)

V roce 1838 zavedl belgický matematik Verhulst logistickou rovnici, která je zobecněním rovnice pro exponenciální růst, ale s maximální hodnotou velikosti populace. K odhadu neznámých parametrů použil údaje z několika zemí, zejména z Belgie. Verhulstova práce byla znovuobjevena až ve 20. letech 20. století.

Pierre-François Verhulst se narodil v roce 1804 v Bruselu. V roce 1825 získal doktorát z matematiky na univerzitě v Gentu. Zajímal se také o politiku. Během svého pobytu v Itálii, kde se léčil s tuberkulózou, se neúspěšně zasazoval o přijetí ústavy pro papežské státy. Po revoluci v roce 1830, kdy Belgie získala nezávislost, vydal historický esej o vlastenci z osmnáctého století. Roku 1835 se stal profesorem matematiky na nově vzniklé Svobodné univerzitě v Bruselu.

Obrázek 6.1:
Verhulst (1804–1849)

V témže roce 1835 vydal jeho krajan Adolphe Quetelet, statistik a ředitel bruselské observatoře, *Traktát o člověku a vývoji jeho schopností*. Quetelet se domníval, že populace nemůže dlouhodobě růst geometrickou řadou, protože překážky, o nichž se zmiňoval Malthus, tvoří jakýsi „odpor“, který je podle něj (analogicky k mechanice) úměrný čtverci rychlosti růstu populace. Tato analogie neměla reálný základ, ale inspirovala Verhulsta.

V roce 1838 Verhulst publikoval *Poznámku k zákonu populačního růstu*. Zde je několik výňatků:

> „Víme, že slavný Malthus prokázal zásadu, že lidská populace má tendenci růst geometrickou řadou tak, že se po určité době, například každých pětadvacet let, zdvojnásobí. Toto tvrzení je nezpochybnitelné, pokud se abstrahuje od rostoucích obtíží při hledání potravy [...].
>
> Virtuální nárůst počtu obyvatel je tedy omezen velikostí a porodností země. V důsledku toho se populace stále více přibližuje k ustálenému stavu.“

Verhulst si pravděpodobně uvědomil, že Queteletova mechanická analogie není rozumná, a místo toho navrhl následující (stále poněkud heuristickou) diferenciální rovnici pro velikost populace $P(t)$ v čase t:

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

Pokud je populace $P(t)$ malá v porovnání s parametrem K, dostaneme přibližnou rovnici

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

jejímž řešením je $P(t) \approx P(0)\,e^{rt}$, tj. exponenciální růst[1]. Rychlost růstu se snižuje s tím, jak se $P(t)$ blíží K. Dokonce by se stala zápornou, kdyby $P(t)$ mohlo překročit K. Abychom získali řešení rovnice (6.1), můžeme postupovat jako Daniel Bernoulli u rovnice (4.5).

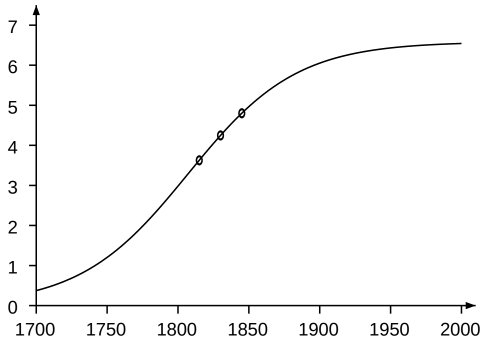
Když rovnici (6.1) vydělíme P^2 a položíme $p = 1/P$, dostaneme lineární diferenicální rovnici $dp/dt = -rp + r/K$. Jestliže dále položíme $q = p - 1/K$ dostaneme $dq/dt = -rq$ a tedy $q(t) = q(0)\,e^{-rt} = (1/P(0) - 1/K)\,e^{-rt}$. Odtud můžeme odvodit vzorce pro $p(t)$ a $P(t)$.

Nakonec dostaneme

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

Celková populace roste z počáteční velikosti $P(0) < K$ v čase $t = 0$ až po mezní hodnotu K, které je dosaženo při $t \to +\infty$ (obrázek 6.2).

[1] Obvykle se hovoří o geometrickém růstu v modelech s diskrétním časem a o exponenciálním růstu v modelech se spojitým časem, ale v podstatě jde o totéž.



Obrázek 6.2: Počet obyvatel Belgie (v milionech) a logistická křivka. Datové body odpovídají rokům 1815, 1830 a 1845. Hodnoty parametrů odpovídají hodnotám z článku z roku 1845.

Aniž by Verhulst uvedl hodnoty neznámých parametrů r a K, porovnal svůj výsledek s údaji o počtu obyvatel Francie v letech 1817–1831, Belgie v letech 1815–1833, anglického hrabství Essex v letech 1811–1831 a Ruska v letech 1796–1827. Ukázalo se, že shoda je docela dobrá.

V roce 1840 se Verhulst stal profesorem na Královské vojenské škole v Bruselu. V následujícím roce vydal *Elementární traktát o eliptických funkcích* a byl zvolen členem belgické Královské akademie. V roce 1845 pokračoval ve svých populačních studiích článkem nazvaným *Matematické zkoumání zákona růstu populace*. Nejprve se vrátil k Malthusově poznámce, podle níž se počet obyvatel USA zdvojnásobil každých 25 let (tabulka 6.1).

Tabulka 6.1: Oficiální počty obyvatel USA.

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

Vypočítáme-li poměr mezi počtem obyvatel v roce $n+10$ a počtem obyvatel v roce n, zjistíme, že tento poměr je 1,350, 1,364, 1,331, 1,335 a 1,326, což je poměrně konstantní hodnota. Počet obyvatel se tedy každých 10 let v průměru znásobil o 1,34 a každých 25 let o $1{,}34^{25/10} \approx 2{,}08$. Od vydání Malthusovy eseje o téměř půl století dříve se tedy populace každých 25 let zdvojnásobovala. Verhulst však dodal:

> „Nebudeme trvat na hypotéze geometrické progrese, protože ta může platit jen za velmi zvláštních okolností, například když

> úrodné území téměř neomezené velikosti obývají lidé s vyspělou civilizací, jako tomu bylo v případě prvních amerických kolonií.“

Verhulst se ve svém článku také vrátil k rovnici (6.1), kterou nazval „logistickou“. Všiml si, že křivka $P(t)$ roste s kladnou křivostí (je konvexní), dokud $P(t) < K/2$, a pak pokračuje v růstu směrem ke K, ale se zápornou křivostí (je konkávní), jakmile $P(t) > K/2$. Křivka má tedy tvar zkresleného písmene S (obrázek 6.2).

Skutečně, $\frac{d^2P}{dt^2} = r(1-2P/K)\,\frac{dP}{dt}$. Takže $\frac{d^2P}{dt^2} > 0$, pokud $P < K/2$, a $\frac{d^2P}{dt^2} < 0$, pokud $P > K/2$.

Verhulst také vysvětlil, jak lze parametry r a K odhadnout z populace $P(t)$ ve třech různých, ale stejně vzdálených letech. Jestliže P_0 je populace v čase $t = 0$, P_1 v čase $t = T$ a P_2 v čase $t = 2T$, pak výpočet, vycházející z rovnice (6.2) ukazuje, že

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

Na základě odhadů počtu obyvatel Belgie v letech 1815, 1830 a 1845 (3,627, 4,247 a 4,801 milionu) získal $K = 6{,}584$ milionu a $r = 2{,}6\,\%$ ročně. Pomocí rovnice (6.2) pak mohl předpovědět, že počet obyvatel Belgie bude na začátku roku 1851 činit 4,998 milionu a na začátku roku 1900 6,064 milionu (obrázek 6.2). Verhulst provedl podobnou studii pro Francii. Získal $K = 39{,}685$ milionů a $r = 3{,}2\,\%$ ročně. Jelikož populace Belgie a Francie v mezidobí tyto hodnoty K značně překročily, vidíme, že logistická rovnice může být realistickým modelem pouze pro časové úseky několika desetiletí, jako ve Verhulstově článku z roku 1838, nikoli však pro delší období.

V roce 1847 vyšla práce *Druhé zkoumání zákona populačního růstu*, v níž Verhulst upustil od logistické rovnice a místo ní zvolil diferenciální rovnici

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

Domníval se, že tato rovnice bude platit, pokud populace $P(t)$ překročí určitou mez. Řešení je

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

Na základě stejných demografických údajů pro Belgii Verhulst nově odhadl parametry r a K. Tentokrát zjistil, že maximální počet obyvatel je $K = 9{,}4$ milionu. Vidíme, jak dalece může výsledek záviset na volbě modelu!

V roce 1848 se Verhulst stal prezidentem Belgické královské akademie, ale následujícího roku v Bruselu zemřel, pravděpodobně na tuberkulózu. Navzdory Verhulstově váhání mezi modelovými rovnicemi byla logistická rovnice o několik desetiletí později nezávisle na sobě znovu objevena různými lidmi. Robertson ji v roce 1908 použil k modelování individuálního růstu zvířat, rostlin, lidí a tělesných orgánů, McKendrick a Kesava Pai ji v roce 1911 použili pro růst populací mikroorganismů a Pearl s Reedem ji v roce 1920 použili pro růst populace USA, který se začal zpomalovat. V roce 1922 si Pearl konečně všiml Verhulstovy práce. Od té doby logistická rovnice inspirovala mnoho prací (viz kapitoly 13, 20 a 24). Maximální velikost populace K se nakonec začala označovat jako „nosná kapacita prostředí“.

Další čtení

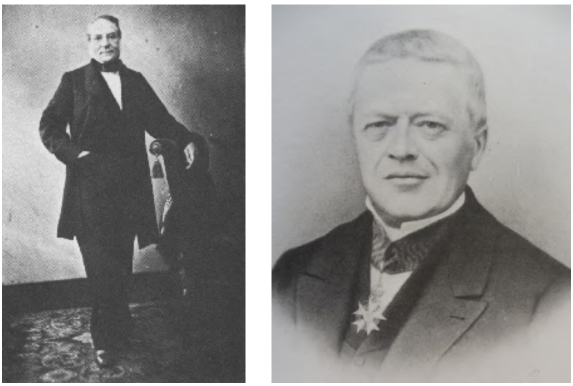
1. Lloyd, P.J.: American, German and British antecedents to Pearl and Reed's logistic curve. *Pop. Stud.* 21, 99–108 (1967)
2. McKendrick, A.G., Kesava Pai, M.: The rate of multiplication of micro-organisms: A mathematical study. *Proc. R. Soc. Edinb.* 31, 649–655 (1911)
3. Pearl, R.: *The Biology of 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). 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

Kapitola 7

Bienaymé, Cournot a zánik příjmení (1845–1847)

Francouzský statistik Bienaymé v roce 1845 pochopil, jak počítat pravděpodobnost zániku příjmení, pokud má každý muž určitý počet synů daný rozdělením pravděpodobnosti. Pokud je průměrný počet synů menší nebo roven jedné, příjmení zanikne jistě. Je-li průměr větší než jedna, je pravděpodobnost vymření ostře menší než jedna. Důkaz svého výsledku zveřejnil o dva roky později v knize, kterou napsal jeho přítel Cournot. Tyto práce byly znovu objeveny teprve nedávno.

Irenée Jules Bienaymé se narodil v roce 1796 v Paříži a studoval na *École Polytechnique*. Pracoval na ministerstvu financí, kde dosáhl vysoké pozice generálního inspektora. Pod vlivem Laplaceovy knihy *Analytická teorie pravděpodobnosti* si ale Bienaymé našel také čas na publikování článků o mnoha aplikacích teorie pravděpodobnosti, jako je demografická a lékařská statistika (dětská úmrtnost, počet narozených dětí, střední délka života), pravděpodobnost chyb v soudnictví, teorie pojištění a reprezentativnost volebních systémů.



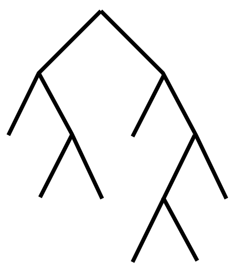
Obrázek 7.1: Bienaymé (1796–1878) a Cournot (1801–1877)

V roce 1845 napsal Bienaymé krátkou poznámku *O zákonu množení a trvání rodin*, která byla zveřejněna v bulletinu *Société Philomatique* v Paříži. Na toto téma již psala řada autorů. Například i ve druhém vydání knihy *Esej* *o principu populace* (1803) zařadil Malthus kapitolu o obyvatelstvu Švýcarska a poznamenal, že

> „v Bernu přijala panovnická rada v letech 1583 až 1654 do buržoazie 487 rodin, z nichž 379 během dvou století vymřelo a v roce 1783 jich zůstalo jen 108.“

V roce 1842 Thomas Doubleday dokonce vyslovil obecnější tvrzení, že rodiny z vyšších šlechtických nebo měšťanských vrstev mají větší tendenci vymírat než rodiny z nižších vrstev. Podobné myšlenky se objevily i v úvodním textu od Émila Littrého k pozitivistické filozofii Augusta Comta uveřejněné ve Francii v roce 1844. Také Bienaymého přítel Benoiston de Châteauneuf publikoval v roce 1845 esej *O trvání šlechtických rodů ve Francii*.

Bienaymé se snažil v této souvislosti vysvětlit, jak je možné, že počet obyvatel země geometricky roste, zatímco velké množství rodin mizí. Aby dokázal tento paradox pochopit, uvažoval o zjednodušeném případu, kdy by všichni muži měli stejnou pravděpodobnost, že budou mít 0, 1, 2, 3... synů, kteří dosáhnou dospělosti. Přesněji řečeno, položil si otázku, jaká je pravděpodobnost, že muž bude mít po n generacích potomky nesoucí jeho jméno. Je-li průměrný počet synů menší než jedna, je jasné, že tato pravděpodobnost by měla s růstem n do nekonečna směřovat k nule. Bienaymé si všiml, že stejný závěr by zůstal pravdivý[1], i kdyby byl průměrný počet synů přesně 1, tedy např. kdyby pravděpodobnost, že nebude mít žádného syna, byla 1/2, stejně jako, že bude mít dva syny (obr. 7.2). V takovém případě však pravděpodobnost mít potomka v n-té generaci směřuje k nule pomaleji: v tomto případě by byla 5 % i po 35 generacích, tj. po jedenácti nebo dvanácti stoletích, pokud uvažujeme na století tři generace[2]. Bienaymé si zároveň všiml, že je-li průměrný počet synů větší než jedna, není vymření rodové linie jisté a tuto pravděpodobnost lze vypočítat řešením jisté algebraické rovnice.



Obrázek 7.2: Umělý příklad rodokmenu. Předek je na vrcholu stromu. V každé generaci mají muži pravděpodobnost 1/2, že nebudou mít žádného syna, a pravděpodobnost 1/2, že budou mít právě dva syny.

[1] vyjma případu, kdy má každý muž přesně jednoho syna

[2] Jak uvidíme dále, tato pravděpodobnost je rovna $1 - x_{35}$ s $x_{n+1} = \frac{1}{2} + \frac{1}{2} x_n^2$ a $x_0 = 0$.

Bienayméův článek ale neobsahoval bližší vysvětlení. V roce 1847 jeho přítel, matematik a ekonom, Antoine-Augustin Cournot, uvedl některé podrobnosti v knize s názvem *O původu a hranicích korespondence mezi algebrou a geometrií*. K problému přistoupil s pomocí náhodné hry, přičemž uvedl, že odpovídá Bienayméo studii o vymírání rodových jmen. My zde ponecháme výklad v termínech rodových jmen. Cournot se nejprve zabýval zvláštním případem, kdy muži mají nejvýše dva syny, přičemž p_0, p_1 a p_2 znamenají pravděpodobnost, že budou mít 0, 1 nebo 2 syny. Samozřejmě, že $p_0 + p_1 + p_2 = 1$. Začneme-li s prvním předkem, je pravděpodobnost x_1 vymření po pouhé jedné generaci zřejmě rovna p_0. Pravděpodobnost vymření během dvou generací je $x_2 = p_0 + p_1 x_1 + p_2 x_1{}^2$, protože buď rod vymřel již v první generaci (pravděpodobnost p_0), nebo byl v první generaci jen jeden syn, který neměl mužského potomka (pravděpodobnost $p_1 x_1$), nebo byli v první generaci dva synové a každý z nich neměl mužského potomka (pravděpodobnost $p_2 x_1^2$). Obecněji řečeno, pravděpodobnost vymření během n generací je

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

Pokud se totiž v první generaci narodí například dva synové (pravděpodobnost p_2), rodina vymře o $n-1$ generací později (tj. v generaci n) s pravděpodobností rovnou $(x_{n-1})^2$. Cournot si všiml, že x_n je rostoucí posloupnost s $x_n \leqslant 1$ pro všechna n. Takže x_n má limitu $x_\infty \leqslant 1$, což je řešení rovnice

$$x = p_0 + p_1 x + p_2 x^2.$$

Substitucí $p_1 = 1 - p_0 - p_2$ dostaneme ekvivalentní rovnici $0 = p_2(x-1)(x-p_0/p_2)$, takže existují dva kořeny: $x = 1$ a $x = p_0/p_2$. Můžeme pak rozlišit tři různé případy v závislosti na průměrném počtu synů $\mathscr{R}_0 = p_1 + 2p_2$, který je rovněž roven $1 - p_0 + p_2$. Pokud platí $\mathscr{R}_0 < 1$, pak $p_0/p_2 > 1$ a jedinou možnou hodnotou limity x_∞ je proto $x = 1$. V takovém případě tedy rodové jméno určitě zanikne. Pokud $\mathscr{R}_0 = 1$, jsou oba kořeny rovny 1 a závěr je stejný. Pokud $\mathscr{R}_0 > 1$, pak Cournot vyvodil, že x_∞ by se mělo rovnat druhému kořeni p_0/p_2, protože ve zvláštním případě $p_0 = 0$ musí být pravděpodobnost vymření zřejmě rovna 0.

Cournot se krátce zmínil o obecnějším případu, kdy muži mohou mít nejvýše m synů s pravděpodobnostmi p_0, p_1,…, p_m. Závěr závisí stejným způsobem na hodnotě

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

průměrného počtu synů, vzhledem k 1. Rovnice pro x_∞, která je

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

má vždy kořen $x = 1$. Má pouze jeden další kladný kořen, který udává pravděpodobnost zániku x_∞, když $\mathscr{R}_0 > 1$.

Bohužel Bienayméův článek a několik stránek v Cournotově knize zůstaly v té době zcela bez povšimnutí. Článek byl zaznamenán až v sedmdesátých letech a stránky knihy o dalších dvacet let později! Mezitím problém a jeho řešení znovu objevili jiní a téma se značně rozvinulo. K tomu se vrátíme v kapitolách 9, 17 a 18.

Bienaymé musel po revoluci v roce 1848 opustit své místo na ministerstvu financí. Stejně tak vedení katedry teorie pravděpodobnosti na pařížské univerzitě, na které byl jistě nejlepším kandidátem, získal někdo jiný. Přesto mohl Bienaymé po roce 1850 opět pracovat na ministerstvu financí, ale v roce 1852 podal výpověď. Ještě téhož roku byl zvolen do Akademie věd, kde byl odborníkem v oblasti statistiky. V roce 1853 dokázal to, co některé moderní učebnice nazývají Bienaymého-Čebyševovou nerovností. V roce 1875 se stal předsedou nově vytvořené *Société Mathématique de France*. Zemřel v Paříži v roce 1878.

Další čtení

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* (1842). archive.org
7. Heyde, C.C., Seneta, E.: *I.J. Bienaymé*. 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* (1852). gallica.bnf.fr
10. Malthus, T.R.: *An Essay on the Principle of Population* (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.

Kapitola 8

Mendel a dědičnost (1865)

V roce 1865 Mendel zveřejnil výsledky svých průkopnických pokusů s hybridizací hrachu. Jeho analýza využívala základní aspekty teorie pravděpodobnosti. Uvažoval také o dynamickém modelu populace samoplodných rostlin. Jeho práce, která byla znovu objevena až v roce 1900, je milníkem v dějinách genetiky.

Johann Mendel se narodil v roce 1822 na Moravě, která byla tehdy součástí Rakouského císařství a dnes je součástí České republiky. Jeho otec byl rolník. Vzhledem k dobrým výsledkům na střední škole a špatnému zdraví dal Mendel přednost dalšímu studiu před prací na rodinném statku. Studium na univerzitě si však nemohl dovolit. Proto v roce 1843 vstoupil do Augustiniánského opatství svatého Tomáše v Brně, kde přijal jméno Gregor. Studoval teologii, ale navštěvoval také několik zemědělských kurzů. V roce 1847 byl vysvěcen na kněze. Několik let vyučoval na gymnáziu, ale neuspěl u zkoušky na řádného profesora. Přesto mohl v letech 1851–1853 díky podpoře svého představeného opata pokračovat ve studiu na vídeňské univerzitě, kde navštěvoval kurzy fyziky, matematiky a přírodních věd. Poté se vrátil do Brna a vyučoval fyziku na technické škole.

Obrázek 8.1:
Mendel (1822–1884)

V letech 1856 až 1863 provedl Mendel v zahradě svého kláštera řadu pokusů s velkým počtem rostlin. V roce 1865 prezentoval své výsledky na dvou schůzích Přírodovědné společnosti v Brně, jejímž byl členem. Jeho práce *Pokusy s rostlinnými hybridy* byla následujícího roku publikována v němčině ve sborníku této společnosti. Mendel vysvětlil, jak se dostal ke studiu variací hrachu, rostlin, které se přirozeně rozmnožují samooplozením a jejichž semena mohou nabývat různých snadno rozpoznatelných forem: kulatá nebo hranatá, žlutá nebo zelená atd. Křížením rostliny pocházející z linie s kulatými semeny a rostliny pocházející z linie s hranatými semeny zjistil, že vždy získá křížence (hybridy), které dávají kulatá semena. Znak „kulatá semena" nazval dominantním a znak „hranatá semena" recesivním. Stejným způsobem ukázal, že znak „žlutá semena" je dominantní a znak „zelená semena" je recesivní.

Mendel si pak všiml, že rostliny vypěstované z hybridních semen plodí v první generaci nová semena, která mají buď dominantní, nebo recesivní znak ve zdánlivě náhodném poměru. Kromě toho si všiml, že opakováním pokusu získal v průměru asi třikrát více semen s dominantním znakem než s recesivním znakem. Například při prvním pokusu získal celkem 5 474 kulatých semen a 1 850 hranatých semen, což odpovídá poměru 2,96:1. Ve druhém pokusu získal celkem 6 022 žlutých semen a 2 001 zelených semen, což odpovídá[1] poměru 3,01 ku 1.

Mendel si také všiml, že mezi rostlinami vypěstovanými ze semen první generace s dominantním znakem bylo těch, které plodily jak semena s dominantním, tak s recesivním znakem, přibližně dvakrát více než těch, které plodily semena pouze s dominantním znakem. Například mezi 565 rostlinami vypěstovanými z kulatých semen první generace jich 372 plodilo jak kulatá, tak hranatá semena, zatímco 193 plodilo pouze kulatá semena; tento poměr je roven 1,93. Podobně mezi 519 rostlinami vypěstovanými ze žlutých semen první generace plodilo 353 rostlin žlutá i zelená semena, zatímco 166 rostlin plodilo pouze žlutá semena; poměr je roven 2,13.

Pro vysvětlení těchto výsledků přišel Mendel s geniálním nápadem považovat pozorovatelný znak semene za výsledek spojení dvou skrytých faktorů, přičemž každý z těchto faktorů je buď dominantní (psáno A), nebo recesivní (psáno a). Existují tedy tři možné kombinace: AA, Aa a aa. Semena s faktory AA nebo Aa mají stejný dominantní znak A. Semena s faktory aa mají recesivní znak a. Mendel navíc předpokládal, že při oplození přenášejí pylová zrna a vajíčka (gamety) pouze jeden z obou faktorů, každý s pravděpodobností 1/2.

---

[1] Jak si později všiml R. A. Fisher (viz kapitola 14), pravděpodobnost, že experimentální výsledky se budou tak blížit teoretické hodnotě, je poměrně malá. Mendel pravděpodobně uspořádal svá data. Například ve druhém pokusu týkajícím se $n = 6022 + 2001 = 8023$ semen je pravděpodobnost, že se poměr bude lišit od 3 o méně než 0,01, jen asi 10 %.

Křížením čistých linií AA a aa tedy vzniknou kříženci, kteří mají všechny faktory Aa a dominantní znak A. Gamety křížence Aa přenášejí faktor A s pravděpodobností 1/2 a faktor a s pravděpodobností 1/2. Při samooplození rostliny vypěstované z hybridního osiva Aa tedy vzniká AA s pravděpodobností 1/4, Aa s pravděpodobností $1/2$ a aa s pravděpodobností $1/4$, jak ukazuje tabulka 8.1.

Tabulka 8.1: Možné výsledky samooplození hybridu Aa a jejich pravděpodobnosti v závislosti na faktorech přenášených samčími gametami (v řádcích) a samičími gametami (ve sloupcích).

Faktor Pravděpodobnost	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 si všiml, že poměry AA : Aa : aa, které jsou 1 : 2 : 1, lze získat také formálním výpočtem $(A+a)^2 = AA + 2Aa + aa$. Protože semena AA a Aa mají zdánlivý znak A, zatímco pouze semena aa mají zdánlivý znak a, je skutečně třikrát více semen se znakem A než se znakem a. Navíc je v průměru dvakrát více semen Aa než AA. Při samooplození rostlin vypěstovaných z prvních z nich vznikají semena buď s dominantním znakem (AA nebo Aa), nebo s recesivním znakem (aa). Pokud jde o samooplození rostlin vypěstovaných ze semen AA, dává vždy semena AA s dominantním znakem. Všechna pozorování jsou tedy vysvětlena.

Mendel se zabýval i následujícími generacemi. Vycházel z N hybridních semen Aa a pro zjednodušení předpokládal, že každá rostlina dává samooplozením pouze čtyři nová semena, a vypočítal, že průměrný počet semen $(AA)_n$, $(Aa)_n$ a $(aa)_n$ v generaci n by byl dán tabulkou 8.2, kde byly výsledky pro přehlednost vyděleny N.

Tabulka 8.2: Následné generace.

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
celkem	1	4	16	64	256	1 024

Tato čísla se jednoduše získají ze vzorců

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

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

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

které říkají, že AA dává po samooplození čtyři semena AA, že aa dává čtyři semena aa a že Aa dává v průměru jedno semeno AA, dvě semena Aa a jedno semeno aa. Mendel si dále všiml, že

$$(AA)_n = (aa)_n = 2^{n-1}(2^n - 1), \qquad (Aa)_n = 2^n.$$

Z rovnice (8.2) a z počáteční podmínky $(Aa)_0 = 1$ přímo vyplývá, že $(Aa)_n = 2^n$. Substitucí tohoto výrazu v rovnici (8.1) dostaneme rovnici $(AA)_{n+1} = 4\,(AA)_n + 2^n$. Snadno vidíme, že jejím partikulárním řešením je $(AA)_n = 2^n c$ pro $c = -1/2$. Obecné řešení přidružené homogenní rovnice $(AA)_{n+1} = 4\,(AA)_n$ je $(AA)_n = 4^n C$. Sečtením těchto dvou řešení dostáváme $(AA)_n = C\,4^n - 2^{n-1}$, přičemž počáteční podmínka $(AA)_0 = 0$ implikuje $C = 1/2$. Posloupnost $(aa)_n$ splňuje stejnou relaci rekurence a stejnou počáteční podmínku jako $(AA)_n$, takže $(aa)_n = (AA)_n$.

Závěrem lze říci, že podíl hybridů Aa v celkové populaci, který je $2^n/4^n = 1/2^n$, se v každé generaci vzniklé samooplozením dělí dvěma.

Mendelova práce zůstala za jeho života zcela bez povšimnutí. O několik let později se Mendel pokusil o podobné experimenty i s jinými druhy rostlin, publikoval několik článků o meteorologii a zkoumal dědičnost včel. Poté, co se v roce 1868 stal opatem, trávil většinu času řešením administrativních problémů. Zemřel v roce 1884.

Teprve v roce 1900 Mendelovu práci konečně nezávisle na sobě a téměř současně znovuobjevili Hugo De Vries v Amsterdamu, Carl Correns v Tübingenu a Erich von Tschermak ve Vídni. Tím začala nová éra v oblasti, kterou dnes nazýváme genetikou.

Další čtení

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

Kapitola 9

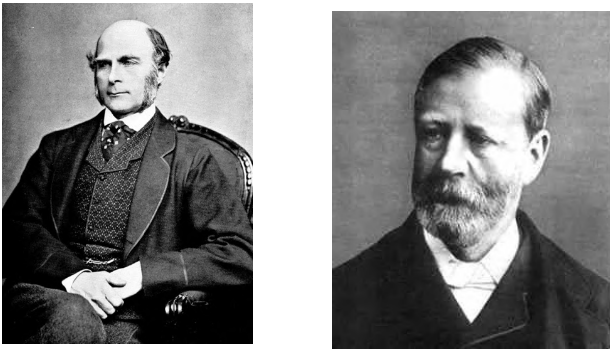
Galton, Watson a problém vymírání (1873–1875)

V roce 1873 se britský statistik Galton a jeho krajan matematik Watson zabývali problémem zániku příjmení (resp. vymření rodů), aniž by znali Bienaymého práci. Watson si všiml, že vytvořující funkci spojenou s rozdělením pravděpodobnosti počtu mužů v každé generaci lze vypočítat rekurentním vztahem. Došel však k nesprávným výsledků, týkajících se pravděpodobnosti zániku příjmení.

Francis Galton se narodil v roce 1822, tedy ve stejném roce jako Mendel, nedaleko Birminghamu v Anglii. Byl nejmladší ze sedmi dětí. Jeho otec byl bohatý bankéř. Jeho bratrancem z matčiny strany byl Charles Darwin. Galton začal v roce 1838 studovat medicínu, nejprve v nemocnici v Birminghamu a později v Londýně. V létě roku 1840 podnikl svou první dlouhou cestu po Evropě, dojel až do Istanbulu. Poté čtyři roky studoval na *Trinity College* na univerzitě v Cambridge. Jeho otec však v roce 1844 zemřel a zanechal po sobě značné jmění. Galton se vzdal myšlenky stát se lékařem. Procestoval Egypt, Súdán a Sýrii. Během několika následujících let si udržoval svobodný způsob života, trávil čas lovem, cestováním v balonech a na lodích nebo se snažil zdokonalit elektrický telegraf. V roce 1850 se vydal na průzkumnou expedici do jihozápadní Afriky (dnešní Namibie). Po návratu do Anglie v roce 1852 byl zvolen do Královské geografické společnosti. Tam mohl sledovat zprávy z výprav do východní Afriky, které hledaly pramen Nilu. Usadil se v Londýně a napsal průvodce pro cestovatele, který se stal bestsellerem. V roce 1856 byl zvolen do Královské společnosti. V té době se zajímal o meteorologii a zavedl pojem „anticyklóna“. Poté, co jeho bratranec Darwin v roce 1859 vydal knihu *O původ druhů*, se Galton začal věnovat studiu dědičnosti. V roce 1869 vydal knihu *Dědičnost geniality*, v níž tvrdil, že intelektuální schopnosti se mohou přenášet dědičností.

V roce 1873 vydal švýcarský botanik Alphonse de Candolle knihu s názvem *Dějiny vědy a učenců v posledních dvou stoletích*, která obsahovala také esej „O vlivu dědičnosti, variability a výběru na vývoj lidského druhu a o jeho pravděpodobné budoucnosti“. Zde uvedl následující poznámky:

> „Mezi přesnými informacemi a velmi rozumnými názory pana Benoistona de Châteauneuf, Galtona a dalších statistiků jsem ne-



Obrázek 9.1: Galton (vlevo) a Watson (vpravo).

> zaznamenal důležitou poznámku, kterou měli učinit k nevyhnutelnému zániku příjmení. Každé příjmení samozřejmě musí zaniknout [...] Matematik by mohl vypočítat, jak k úbytku příjmení nebo šlechtických titulů dojde, kdyby znal pravděpodobnost narození dítěte ženského či mužského pohlaví a pravděpodobnost, že se danému páru žádné dítě nenarodí.“

Jedná se o stejný problém, který Bienaymé studoval v roce 1845. Candolle, který o Bienaymého práci nevěděl, se však domníval, že všechny rody jsou odsouzeny k zániku (příjmení vymizí). Galton si všiml výše uvedeného odstavce v Candolleově knize. Protože ani on o Bienaymého práci nevěděl, položil následující otevřený problém pro čtenáře *Educational Times*:

> Problém 4 001: Velký národ, z něhož se budeme zabývat pouze dospělými muži v počtu N, z nichž každý nese samostatné příjmení, kolonizuje určitý okres. Jejich populační zákon je takový, že v každé generaci a_0 procent dospělých mužů nemá žádného mužského potomka, který by se dožil dospělosti; a_1 má jednoho takového mužského potomka; a_2 má dva a tak dále až k a_5, která jich mají pět.
>
> Zjistěte, (1) podíl příjmení, která zaniknou po r generacích a (2) kolik příjmení ponese právě m mužů.

Všimněte si, že druhou částí problému se Bienaymé nezabýval. Galton nedostal od čtenářů časopisu žádnou uspokojivou odpověď a zřejmě ani sám nedokázal najít řešení problému. Požádal tedy svého přítele Henryho Williama Watsona, matematika, aby se pokusil o jeho vyřešení.

Watson se narodil v Londýně v roce 1827. Jeho otec byl důstojníkem britského námořnictva. Nejprve studoval na *King's College* v Londýně a poté se věnoval matematice na *Trinity College* na Cambridgeské univerzitě, a to od roku 1846 do roku 1850, tedy jen několik let po Galtonovi. Postupně se stal členem *Trinity College*, asistentem na *City of London School*, lektorem matematiky na *King's College* a v letech 1857–1865 profesorem matematiky na *Harrow School*. Měl rád horolezectví a v roce 1855 se zúčastnil expedice, která dosáhla vrcholu masivu Monte Rosa ve Švýcarsku. V roce 1856 byl vysvěcen na jáhna[1] a o dva roky později na anglikánského kněze. Od roku 1865 až do svého odchodu do důchodu byl farářem v Berkswellu a Bartonu nedaleko Coventry, což mu ponechávalo dostatek času na výzkum.

Galton a Watson společně napsali článek s názvem *O pravděpodobnosti vymírání rodin*, který byl publikován v roce 1875 v časopise *Journal of the Royal Anthropological Institute*. Galton představil problém a Watson vysvětlil své výpočty a závěry, ke kterým dospěl. Předpokládali, že muži mají nejvýše q synů, přičemž p_k je pravděpodobnost, že budou mít k synů ($k = 0$, 1, 2, …, q). Jinými slovy, $p_k = a_k/100$, pokud použijeme původní Galtonův zápis. Takže

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

Uvažujme situaci, kdy generaci 0 tvoří jediný člověk. Generaci 1 tvoří s mužů s pravděpodobností p_s. Pomocí triku, který byl v jeho době dobře znám a který zavedl již dávno před ním Abraham de Moivre, Watson uvažoval vytvořující funkci

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

obsahující s pravděpodobnosti p_0, …, p_q. Podobně nechť $f_n(x)$ je polynom, jehož koeficient u x^s je pravděpodobnost, že v generaci n bude s mužů, počínaje jedním mužem v generaci 0. Pak $f_1(x) = f(x)$. Watson si všiml, že

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

vzorec, který umožňuje rekurentní výpočet $f_n(x)$.

Vskutku, uvažujme

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

[1] Jáhen či diákon je nejnižší anglikánský duchovní.

Všimněte si, že v generaci n je nejvýše q^n mužů. Jestliže v generaci $n-1$ existuje s mužů, které označíme čísly 1 až s, nazvěme $t_1,\dots,t_s$ počet jejich mužských potomků. V takovém případě bude v generaci n t mužů s pravděpodobností rovnou

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

Pokud $s=0$, je tato pravděpodobnost je rovna 1 pro $t=0$, a je rovna 0 pro $t \geqslant 1$. Proto

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

Z toho vyplývá, že

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

Zejména pravděpodobnost zániku příjmení x_n během n generací je rovna $p_{0,n}$, což je stejné jako $f_n(0)$. Jako první příklad Watson uvažoval

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

tj. $q=3$ a $p_0=p_1=p_2=1/3$. Vypočítal polynomy $f_n(x)$ pro $n=1,\dots,4$ pomocí rovnice (9.2). Získal např.

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

a $f_2(0)=13/27 \approx 0{,}481$. Výpočet $f_n(x)$ pro $n \geqslant 3$ je velmi náročný. Tak náročný, že Watson udělal chybu už pro $n=4$. Protože $x_5 = f_5(0) = f_4(f(0))$, výpočet $f_5(x)$ již není potřeba. Watson získal následující seznam pravděpodobností zániku příjmení $x_n = f_n(0)$:

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

Správné hodnoty jsou $x_4 \approx 0{,}632$ a $x_5 \approx 0{,}677$, což lze ověřit pomocí jednoduchého vzorce

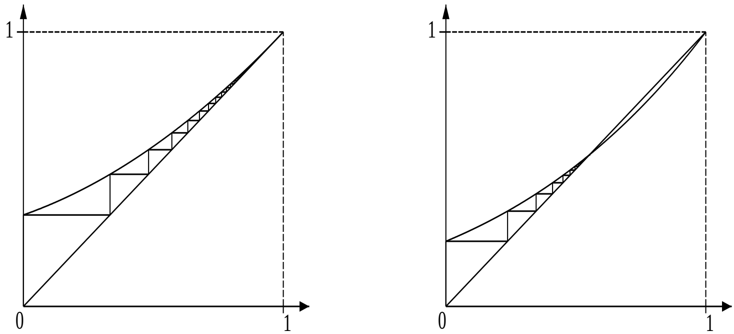
$$x_n = f(x_{n-1}),$$

který odvodil Bienaymé. Jak uvidíme v kapitole 17, tento vzorec lze mimo jiné odvodit z rovnice (9.2).

Watson si všiml, že každý člověk má v průměru

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

synů a že $\mathscr{R}_0 = 1$ v jeho prvním příkladu. Je možné se tedy domnívat, že pokud bude počáteční počet mužských členů rodu dostatečně velký, zůstane velikost rodu zhruba konstantní. Nicméně Watson tvrdil, že pravděpodobnost zániku x_n konverguje k 1, když $n \to +\infty$, i když poměrně pomalu. Jinými slovy, celý rod vymře a příjmení zanikne, jak předpokládal Candolle. Obrázek 9.2a, který není v původním článku nakreslen, a Bienayméovy výsledky potvrzují, že tento závěr pro tento první příklad je správný.



Obrázek 9.2: Graf funkcí $y = f(x)$ a $y = x$. Pravděpodobnost zániku příjmení $x_n = f(x_{n-1})$ během n generací je výškou „schodu" n. Vlevo: $f(x) = (1 + x + x^2)/3$. Vpravo: $f(x) = (3 + x)^5/4^5$.

Jako druhý příklad Watson uvedl binomické rozdělení pravděpodobnosti.

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

pro které je vytvořující funkce (9.1)

$$f(x) = \frac{(a+bx)^q}{(a+b)^q}.$$

Vypočítal $f_2(x)$ a $x_2 = f_2(0)$. V tomto případě si uvědomil, že $x_2 = f(x_1)$ a že $x_n = f(x_{n-1})$ pro všechna n. Domníval se však, že tento vzorec platí pouze pro speciální binomický případ (9.3). Když jej aplikoval na případ, kdy $q = 5$, $a = 3$ a $b = 1$, dostal následující výsledky

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

Watson si uvědomil, že x_n musí konvergovat k limitě x_∞ pro $n \to +\infty$, která splňuje

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

Všiml si, že $x = 1$ je řešením této rovnice, ale neuvědomil si, že mohou existovat i jiná řešení pro $\mathscr{R}_0 > 1$. Tak chybně odvodil, naveden Candollem, že k zániku ($x_\infty = 1$) dochází vždy, i v tomto druhém uvažovaném případě. Obrázek 9.2b ukazuje, že to neplatí!

Watson si všiml, že průměrný počet synů v každé generaci je v tomto číselném příkladu je větší než 1 (lze ukázat, že $\mathscr{R}_0 = qb/(a+b) = 5/4$), což znamená, že populace má tendenci exponenciálně růst. To mu však nepomohlo odhalit vlastní chybu. Dokonce vyslovil domněnku, že zánik příjmení je jistý pro každé rozdělení pravděpodobnosti (p_k), tedy nejen pro binomický případ. K tomuto problému se vrátíme v kapitolách 17 a 18.

Galton pokračoval ve statistickém studiu rodů knihou *Anglické vědecké osobnosti, jejich původ a výchova*, která se zaměřovala na genealogii členů Královské společnosti. Začal se také zajímat o antropometrii, měření lidského těla. Využil mezinárodní výstavy v Londýně v roce 1884 a shromáždil údaje o velkém počtu lidí. Své výsledky publikoval v roce 1889 v knize *Přírodní dědičnost*, jejíž příloha reprodukovala článek napsaný ve spolupráci s Watsonem. Tato kniha také zavedla některé nové statistické pojmy, jako je „percentil“ a „kvartil“, a také slovo „eugenika“, tj. zdokonalování lidského druhu z hlediska dědičných znaků. Po roce 1888 Galton vyvinul techniku rozpoznávání otisků prstů, kterou o několik let později začala používat britská policie. Pokračoval také ve studiu role dědičnosti (přírody) a prostředí (výchovy) na fyzické a intelektuální vlastnosti dvojčat, na velikost hrachu pěstovaného po několik generací nebo na barvu myší chovaných v laboratoři. To ho přivedlo k pojmu „korelační koeficient“ mezi dvěma proměnnými. V roce 1904 byla v rámci *University College* v Londýně založena Galtonova laboratoř. Galton byl v roce 1909 povýšen do rytířského stavu a zemřel v roce 1911.

Watson vydal několik knih, zejména pojednání o kinetické teorii plynů v roce 1876 a dvousvazkové pojednání o matematické teorii elektřiny a magnetismu (1885 a 1889). V roce 1881 byl zvolen do Královské společnosti a zemřel v Brightonu v roce 1903.

V roce 1924 shrnul Karl Pearson ve druhém díle Galtonova životopisu článek o zániku příjmení, aniž by si všiml chyby, které se Galton dopustil. Tato chyba byla nakonec objevena až v roce 1930 (viz kapitola 18).

Další čtení

1. De Candolle, A.: *Histoire des sciences et des savants depuis deux siècles*. 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. 1/2. Cambridge University Press (1914/1924). galton.org
6. S.H.B.: Henry William Watson, 1827-1903. *Proc. R. Soc. Lond.* 75, 266–269 (1905). gallica.bnf.fr
7. Watson, H.W., Galton, F.: On the probability of the extinction of families. *J. Anthropol. Inst.* 4, 138–144 (1875). galton.org

Kapitola 10

Lotka a teorie stabilizované populace (1907–1911)

V roce 1907 začal americký chemik Alfred Lotka studovat vztah mezi porodností, věkově specifickou úmrtností a tempem růstu populace pomocí modelu se spojitým časem. V roce 1911 publikoval spolu s F. R. Sharpem další článek na stejné téma, který zahrnoval i věkově specifické míry porodnosti. Implicitní rovnice udávající míru růstu populace se často nazývá „Lotkova rovnice".

Alfred James Lotka se narodil americkým rodičům v roce 1880 ve Lvově, který byl součástí Rakouska-Uherska (dnes Lviv na Ukrajině). Studoval nejprve ve Francii a v Německu a v roce 1901 získal bakalářský titul z fyziky a chemie na univerzitě v anglickém Birminghamu. Poté strávil jeden rok v Lipsku, kde roli termodynamiky v chemii a biologii zdůrazňoval Wilhelm Ostwald, který měl v roce 1909 obdržet Nobelovu cenu za chemii. V roce 1902 se Lotka usadil v New Yorku a začal pracovat pro *General Chemical Company*.

Obrázek 10.1:
Lotka (1880–1949)

V letech 1907 a 1911[1] se Lotka zabýval dynamikou věkově strukturovaných populací, aniž by věděl o Eulerově práci na stejné téma (viz kapitola 3). Na rozdíl od Eulera předpokládal, že čas a věk jsou spojité proměnné. Nechť

[1] Druhý článek byl napsán ve spolupráci s F. R. Sharpem, matematikem z Cornellovy univerzity.

$B(t)$ je porodnost mužů (počet narozených mužů za jednotku času) v čase t, $p(x)$ pravděpodobnost, že budou ve věku x ještě naživu, a $h(x)$ porodnost ve věku x: $h(x)\,dx$ je pravděpodobnost, že se muži narodí jeden syn mezi věkem x a $x+dx$, je-li dx nekonečně malé. Pak

$$\int_0^{+\infty} p(x)\,dx$$

je střední délka života při narození. Dále $B(t-x)\,p(x)\,dx$ je počet mužů narozených mezi časem $t-x$ a $t-x+dx$, kteří ještě žijí v čase t. Tito muži mají $B(t-x)\,p(x)\,h(x)\,dx$ synů za jednotku času v čase t. Celková porodnost mužů v čase t je tedy následující

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

Při hledání řešení této integrální rovnice v neznámé $B(t)$ v exponenciálním tvaru $B(t) = b\,e^{rt}$ získal Lotka dělením obou stran rovnicí $B(t)$ rovnici.

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

která je dnes demografy nazývána „Lotkova rovnice“[2]. Euler získal analogickou implicitní rovnici (3.1) pro rychlost růstu, když čas a věk jsou diskrétní proměnné. Lotka si všiml, že pravá strana (10.1) je klesající funkcí r, která pro $r \to -\infty$ konverguje k $+\infty$ a pro $r \to +\infty$ konverguje k 0. Existuje tedy jediná hodnota r, řekněme r^*, pro kterou rovnice (10.1) platí. Kromě toho platí, že $r^* > 0$ tehdy a jen tehdy, když

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

Parametr $\mathscr{R}_0$ (zápis zavedli Dublin a Lotka v roce 1925) je očekávaný počet synů, které může mít jeden člověk během svého života.

Lotkův předpoklad[3], že bez ohledu na počáteční věkovou strukturu populace je počet narozených mužů za jednotku času skutečně takový, že $B(t) \sim b\,e^{r^*t}$, když $t \to +\infty$, kde b je konstanta. Celková populace je pak dána vztahem $P(t) = \int_0^{+\infty} B(t-x)\,p(x)\,dx$. Z toho vyplývá, že $P(t)$ také roste nebo

[2] R. A. Fisher dospěl nezávisle na něm ke stejné rovnici v roce 1927 a později interpretoval kořen r^* jako míru „Darwinovy zdatnosti“ v evoluční teorii přírodního výběru.

[3] Ten byl rigorózně odvozen v roce 1941 Willym Fellerem, který byl tehdy profesorem matematiky na Brownově univerzitě v USA. Pravděpodobnostní přístup vypracovali v roce 1968 Crump, Mode a Jagers.

klesá jako $e^{r^* t}$, když $t \to +\infty$: míra růstu je rovna r^*. Navíc věková struktura populace, která je daná vztahem $B(t-x)\,p(x)/P(t)$, konverguje k hodnotě

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

Lotka ji nazval „stabilizovaná populací“: věková pyramida si v čase zachovává stejný tvar, ale celkový počet obyvatel exponenciálně roste nebo klesá. Závěr je tedy stejný jako v Eulerově modelu diskrétního času. Lotkova studie však bere v úvahu závislost porodnosti na věku. Je tedy v jistém smyslu obecnější než Eulerova.

Lotka se tímto tématem zabýval po celý svůj život. V letech 1908–1909 pokračoval ve studiu na Cornellově univerzitě, aby získal magisterský titul. V letech 1909–1911 pracoval pro Národní úřad pro standardizaci a v letech 1911–1914 jako redaktor časopisu *Scientific American Supplement*. V roce 1912 získal doktorát na Birminghamské univerzitě za články, které publikoval od roku 1907 o populační dynamice a demografii. Během první světové války pracoval opět pro společnost *General Chemical Company* na způsobu získávání dusíku z atmosféry. V roce 1920 jeden z jeho článků o biologických oscilacích (viz kapitola 13) udělal hluboký dojem na Raymonda Pearla, profesora biometrie na Univerzitě Johnse Hopkinse, který právě „znovuobjevil“ logistickou rovnici (viz kapitola 6). Lotka doufal, že najde práci v Rockefellerově institutu lékařského výzkumu v New Yorku, a pracoval na matematických modelech, které Ross vyvinul pro malárii (viz kapitola 12). Nakonec získal dvouleté stipendium na Univerzitě Johna Hopkinse, které mu umožnilo napsat knihu *Základy fyzikální biologie*, vydanou v roce 1925. Poté se stal vedoucím výzkumného oddělení Metropolitní životní pojišťovny v New Yorku. Zaměřil se na matematickou analýzu demografických otázek a ve spolupráci s kolegou, statistikem a viceprezidentem společnosti Louisem Israelem Dublinem, vydal několik knih: *Peněžní hodnota člověka*. (1930), *Délka života* (1936) a *Dvacet pět let pokroku v oblasti zdraví* (1937). V letech 1938–1939 byl zvolen prezidentem Americké populační asociace. Mezi jeho různé statistické studie patří „Lotkův zákon“ (sahající až do roku 1926), který říká, že počet autorů, kteří napsali n článků v daném vědním oboru, klesá s rostoucím n víceméně jako $1/n^2$.

Lotka vydal také knihu ve francouzštině s názvem *Analytická teorie biologických asociací*. První část, která byla více filozofická, vyšla v roce 1934. Druhá, odbornější část, vydaná v roce 1939, shrnovala všechny jeho výzkumy v oblasti lidské demografie od roku 1907. V knize Lotka také představil svůj příspěvek k problematice vymírání rodových jmen. Poté, co v roce 1930 vyšel Steffensenův první článek na toto téma (viz kapitola 18), aplikoval teorii na údaje obsažené ve sčítání bělošské populace v USA z roku 1920. Všiml si, že pozorované rozdělení $(p_k)_{k\geqslant 0}$ počtu synů je dobře aproximováno klesajícím geometrickým zákonem pro všechna $k \geqslant 1$: $p_0 = a$, $p_k = b\,c^{k-1}$ $(k \geqslant 1)$, pro $a = 0{,}4825$, $b = 0{,}2126$ a $c = 1 - b/(1-a)$. Odtud vyplývá, že $\sum_{k\geqslant 0} p_k = 1$. Příslušná generující funkce je

$$f(x) = a + b \sum_{k=1}^{+\infty} c^{k-1}\, x^k = a + \frac{b\,x}{1 - c\,x}.$$

Dvě řešení rovnice $x = f(x)$ jsou $x = 1$ a $x = a/c$. Pravděpodobnost zániku x_∞ je nejmenší z těchto dvou řešení (viz kapitola 7). S číselnými hodnotami pro USA zjistil $x_\infty \approx 0{,}819$, zatímco průměrný počet synů byl $\mathscr{R}_0 = f'(1) = (1-a)^2/b \approx 1{,}260$. Navzdory průměrnému počtu dětí (včetně synů a dcer) blízkému 2,5 je pravděpodobnost zániku příjmení vyšší než 80 %.

V roce 1942 byl Lotka zvolen prezidentem Americké statistické asociace. V roce 1947 odešel do důchodu a zemřel v roce 1949 v New Jersey. Nové vydání jeho knihy z roku 1925 vyšlo v roce 1956 s mírně odlišným názvem *Základy matematické biologie*.

Další čtení

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, 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*, 2e partie. Hermann, Paris (1939) gallica.bnf.fr
9. Sharpe, F.R., Lotka, A.J.: A problem in age-distribution. *Philos. Mag. Ser.* 6, 21, 435–438 (1911) → Smith & Keyfitz (1977).
10. Smith, D.P., Keyfitz, N.: *Mathematical Demography*. Springer (1977)
11. Tanner, A.: *Von Molekülen, Parasiten und Menschen – Lotka und die Mathematisierung des Lebens*. Zürich (2014) doi:10.3929/ethz-a-010209129

Kapitola 11

Hardyho-Weinbergův zákon (1908)

V roce 1908 britský matematik Hardy a německý lékař Weinberg nezávisle na sobě odvodili, že v nekonečně velké populaci, jejíž jedinci se páří náhodně a předávají své geny podle Mendelových zákonů, frekvence genotypů získaných ze dvou alel zůstávají v průběhu generací konstantní. Jejich matematický model se stal jedním z východisek populační genetiky.

Godfrey Harold Hardy se narodil roku 1877 v anglickém hrabství Surrey. Oba jeho rodiče byli učitelé. Od roku 1896 studoval matematiku na *Trinity College* na univerzitě v Cambridge, v roce 1900 se stal vědeckým pracovníkem této koleje a v roce 1906 lektorem matematiky. Po první knize o *Integracích funkcí jedné proměnné* (1905) vydal roku 1908 učebnici *Kurz čisté matematiky*, která se dočkala mnoha vydání nejen anglických, ale i v překladech.

Obrázek 11.1:
Hardy (1877–1947)

V té době byly znovuobjeveny Mendelovy práce, které vyvolaly také určité otázky. Někteří biologové se ptali, proč frekvence dominantních znaků z generace na generaci neroste. Reginald Punnett, autor knihy *Mendelismus* z roku 1905, se na to zeptal Hardyho, s nímž hrával v Cambridgi kriket. Hardy popsal své řešení v článku *Mendelovské poměry ve smíšené populaci* publikovaném v roce 1908. Pro zjednodušení analýzy si představil situaci ve veliké populaci, kde se sexuální partneři vybírají náhodně. Navíc se omezil na pouhé dva faktory (varianty genu neboli „alely“) A a a, přičemž A označuje dominantní a a recesivní alelu. Pro n-tou generaci nechť p_n označuje frekvenci „genotypu“ AA, $2q_n$ frekvenci Aa a r_n frekvenci aa. Při tomto označení řejmě platí $p_n + 2q_n + r_n = 1$. Hardy dále předpokládal, že žádný z těchto genotypů nezvyšuje úmrtnost ani nesnižuje plodnost ve srovnání s ostatními dvěma genotypy. Frekvence alel v generaci $n+1$ lze snadno vypočítat, pokud si všimneme, že náhodně vybraný jedinec v generaci n přenáší alelu A s pravděpodobností $p_n + q_n$: buď je genotyp AA a alela A se přenáší s jistotou, nebo je genotyp Aa a alela A se přenáší s 50 % pravděpodobností. Podobně alela a se přenáší s pravděpodobností $q_n + r_n$. Tabulku 11.1 lze tedy sestavit stejným způsobem jako tabulku 8.1.

Tabulka 11.1: Výpočet četností genotypů v $(n+1)$-ní generaci z četností alel rodičů (řádky jsou pro matku, sloupce pro otce).

Alela Frekvence	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 $(q_n + r_n)^2$

Četnosti genotypů AA, Aa a aa v generaci $n+1$ jsou p_{n+1}, $2q_{n+1}$ a r_{n+1}. Hardy tedy odvodil, že

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

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

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

Poté zkoumal, za jakých podmínek zůstávají frekvence genotypů v průběhu generací konstantní a rovny p, $2q$ a r. Protože podle definice je $p + 2q + r = 1$, snadno nahlédneme, že ze všech rovnic (11.1)–(11.3) plyne stejná podmínka $q^2 = pr$.

> Například první rovnice dává $p = (p+q)^2 = p^2 + 2pq + q^2$, což je ekvivalentní s $p(1 - p - 2q) = q^2$ a dále s $pr = q^2$.

Hardy tak ukázal, že při každých počátečních podmínkách $(p_0, 2q_0, r_0)$ s vlastností $p_0 + 2q_0 + r_0 = 1$ platí

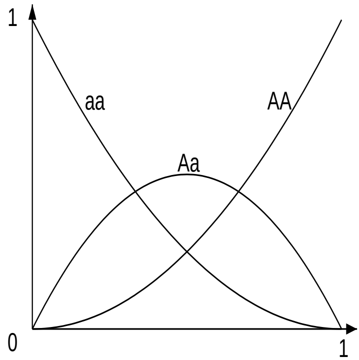
$${q_1}^2 = (p_0 + q_0)^2 (q_0 + r_0)^2 = p_1 r_1 .$$

Hned první stav $(p_1, 2q_1, r_1)$ je tedy rovnovážný. V důsledku toho je libovolný následující stav $(p_n, 2q_n, r_n)$ stejný jako $(p_1, 2q_1, r_1)$ pro všechna $n \geqslant 1$.

Pokud položíme $x = p_0 + q_0$ pro frekvenci alely A v počáteční generaci, pak frekvence alely a je rovna $1 - x = q_0 + r_0$. Použijeme-li opět systém (11.1)–(11.3), dostaneme výsledek

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

pro všechna $n \geqslant 1$ (obr. 11.2).

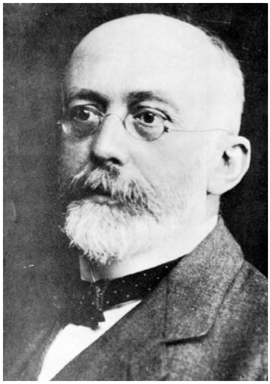


Obrázek 11.2: Grafy funkcí x^2, $2x(1-x)$ a $(1-x)^2$ odpovídající rovnovážným frekvencím genotypů AA, Aa a aa.

Provedené úvahy lze shrnout: výše uvedené předpoklady (nekonečná populace, náhodné páření a Mendelovy zákony) implikují, že *frekvence genotypů AA, Aa a aa zůstávají v průběhu generací neměnné*. Z Mendelovy teorie neplyne postupné zvyšování frekvence dominantního znaku, jak se původně myslelo.

O několik let později Fisher zdůraznil jeden důležitý důsledek tohoto zákona: v prvním přiblížení (tj. pokud jsou předpoklady modelu splněny) si populace udržuje konstantní genetickou variabilitu. Toto pozorování odstraňuje jeden z problémů Darwinovy teorie vývoje přírodním výběrem. Darwin se totiž spolu se svými současníky domníval, že v každé generaci jsou fyziologické vlastnosti potomků jakýmsi průměrem vlastností obou rodičů, přičemž každý z rodičů přispívá jednou polovinou. Tuto myšlenku, známou jako regrese k průměru, později důkladně statisticky zkoumali Francis Galton s Karlem Pearsonem, svým nástupcem v biometrické laboratoři. Pokud by byla pravdivá, rozptyl charakteristik v populaci by měl být v každé generaci poloviční a brzy by populace měla být tak homogenní, že by neumožnila přírodní výběr, který by mohl vysvětlit evoluci. Nicméně trvalo několik let, než byl mechanismus průměrování odmítnut. Biometrici hájící Darwinův názor jen zdráhavě připouštěli, že Mendelovy zákony jsou pro pochopení evoluce nezbytné.

Po této práci se Hardy v roce 1908 vrátil k čisté matematice. Ve své autobiografii *Obrana matematikova* dokonce hrdě prohlašoval, že nikdy neudělal nic praktického. V roce 1910 byl zvolen do Královské učené společnosti (*Royal society*). Roku 1913 rozpoznal v Indovi Ramanujanovi zázračné dítě a pozval ho pracovat do Cambridge. Po první světové válce se stal profesorem na Oxfordské univerzitě a rozvíjel plodnou spolupráci se svým krajanem Littlewoodem. V letech 1931 až 1942 byl opět profesorem v Cambridge. Vydal mnoho knih, často se spoluautory: *Řády nekonečna* (1910), *Obecná teorie Dirichletových řad* s Marcelem Rieszem (1915), *Nerovnosti* s Littlewoodem a Pólyou (1934), *Úvod do teorie čísel* s E. M. Wrightem (1938), *Ramanujan* (1940), *Fourierovy řady* s Rogosinskim (1944) a *Divergentní řady* (1949). Zemřel v Cambridgi roku 1947.



Obrázek 11.3:
Weinberg (1862–1937)

O několik desetiletí později vyšlo najevo, že Hardyho zákon pro genové frekvence objevil v témže roce 1908 německý lékař Wilhelm Weinberg. Ten se narodil roku 1862 ve Stuttgartu. Vytudoval medicínu v Tübingenu a Mnichově a několik let po získání doktorátu pracoval v nemocnicích v Berlíně, Vídni a Frankfurtu. V roce 1889 se usadil ve Stuttgartu jako praktický lékař a porodník. Přestože ho tato práce velice zaměstnávala, našel si i čas k napsání mnoha článků do německých odborných časopisů. V roce 1901 sledoval statistické záznamy četnosti dvojčat stejného pohlaví. Článek z roku 1908, v němž vysvětloval stejný zákon, jaký objevil Hardy, byl však publikován jen v místním lékařském časopise a v té době si nikdo jeho význam neuvědomil. Na rozdíl od Hardyho však Weinberg ve studiu zákona o frekvencích alel pokračoval i v následujících letech a zobecnil ho, například, na případ více než dvou alel. Svou prací přispěl také k rozvoji lékařské statistiky. Weinberg zemřel v roce 1937. Po znovuobjevení jeho článku z roku 1908 genetici začali pro zákon stability genotypových frekvencí používat název „Hardyho-Weinbergův zákon“.

V současnosti je tento zákon často používán následujícím způsobem. Jestliže vzácná recesivní alela a nijak neovlivňuje přežití ani plodnost a známe frekvenci x^2 genotypu aa, který se projevuje určitým fenotypem, pak můžeme vypočítat x a odhadnout frekvenci $2x(1-x) \approx 2x$ genotypu Aa. Pokud je například frekvence aa 1/20 000, dostaneme $x \approx 1/140$, takže $2x \approx 1/70$ je frekvence genotypu Aa. Recesivní alela a, která se ve fenotypu objevuje velice vzácně, ve skutečnosti tak neobvyklá není.

Další čtení

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; česky: *Obrana matematikova.* Prostor, Praha (1999)
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

Kapitola 12

Ross a malárie (1911)

V roce 1911 studoval britský lékař Ronald Ross, který již v roce 1902 obdržel Nobelovu cenu za práci o malárii, systém diferenciálních rovnic modelujících šíření této nemoci. Ukázal, že malárie může přetrvávat pouze tehdy, pokud je počet komárů vyšší než určitá kritická hodnota. K vymýcení malárie tak není nutné vyhubit všechny komáry – postačí vyhubit jen určitou část populace. Podobné epidemické modely později vypracovali Kermack s McKendrickem.

Ronald Ross se narodil v roce 1857 na severu Indie, kde jeho otec působil jako důstojník britské armády. Studoval medicínu v Londýně, ale raději psal básně a dramata. Poté, co rok pracoval na lodi jako chirurg, vstoupil v roce 1881 do Indické lékařské služby. Lékařská práce v Indii mu ponechávala dostatek volného času pro psaní literárních děl a učení se matematice. Během propustky strávené v Anglii v roce 1888 získal diplom z veřejného zdravotnictví a studoval bakteriologii, novou vědu, kterou o několik let dříve vytvořili Pasteur s Kochem. Po návratu do Indie začal Ross studovat malárii. Během své druhé propustky v roce 1894 se v Londýně setkal s Patrickem Mansonem, specialistou na tropickou medicínu, který mu pod mikroskopem ukázal to, čeho si v roce 1880 všiml francouzský vojenský lékař Alphonse Laveran: krev pacientů s malárií obsahuje parazity. Manson vyslovil domněnku, že parazité mohou pocházet z komárů, protože sám v Číně objevil u tohoto hmyzu parazita jiné tropické nemoci (filariózy). Domníval se však, že lidé se tímto parazitem nakazili pitím vody kontaminované komáry. V letech 1895 až 1898 Ross pokračoval ve výzkumu v Indii a Mansonovu myšlenku ověřoval. V roce 1897 objevil v žaludku jistého druhu komára, kterého předtím nestudoval (anofeles), určité parazity podobné těm, které pozoroval Laveran. Protože ho jeho nadřízení poslali do Kalkaty v období, kdy byly případy malárie vzácné, rozhodl se studovat malárii u ptáků chovaných v klecích. Našel parazita ve slinných žlázách komárů rodu Anofeles a experimentálně úspěšně nakazil zdravé ptáky tím, že nechal komáry je štípnout: tím dokázal, že malárie se přenáší komářím štípnutím, a nikoli požitím kontaminované vody. Roku 1899 Ross opustil Indickou lékařskou službu a začal vyučovat na Liverpoolské škole tropické medicíny, která byla založena o rok dříve. V roce 1901 byl zvolen do Královské společnosti a v roce 1902 obdržel Nobelovu cenu za fyziologii nebo lékařství za svou práci o malárii. Cestoval do Afriky, na Mauricius a do oblasti Středomoří, aby popularizoval boj proti komárům. Metoda byla úspěšná v Egyptě podél Suezského průplavu, podél budovaného Panamského průplavu, na Kubě a v Malajsii. V některých jiných oblastech však byla méně úspěšná. V roce 1908 Ross vydal práci *Zpráva o prevenci malárie na Mauriciu* a v roce 1910 knihu *Prevence malárie*.

Obrázek 12.1:
Ronald Ross (1857–1932)

Navzdory důkazům o úloze některých komárů při přenosu malárie se Ross setkal se skepticismem, když tvrdil, že malárii lze vymýtit pouhým snížením počtu komárů. Aby své tvrzení podpořil, pokusil se ve druhém vydání své knihy *Prevence malárie*, publikovaném v roce 1911, vytvořit matematické modely přenosu malárie. Jeden z jeho modelů sestával ze soustavy dvou diferenciálních rovnic. Zaveďme si následující značení:

- N: celková lidská populace v dané oblasti;
- $I(t)$: počet lidí nakažených malárií v čase t;
- n: celková populace komárů (předpokládaná konstantní);
- $i(t)$: počet komárů nakažených malárií;
- b: četnost štípání komárů;
- p (resp. p'): pravděpodobnost přenosu malárie z člověka na komára (resp. z komára na člověka) během jednoho štípnutí;
- a: rychlost, s jakou se lidé zotavují z malárie;
- m: úmrtnost komárů.

Během krátkého časového intervalu dt štípne každý infikovaný komár $b\,dt$ lidí, z nichž část $\frac{N-I}{N}$ ještě není infikovaná. Uvážíme-li pravděpodobnost přenosu p', dostáváme $b\,p'\,i\,\frac{N-I}{N}\,dt$ nově nakažených lidí. Během stejného časového intervalu je počet lidí, kteří se uzdraví, $a\,I\,dt$. Z toho vyplývá, že rychlost, s jakou se během časového intervalu dt změní počet nakažených lidí, je

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

Podobně každý neinfikovaný komár štípne $b\,dt$ lidí, z nichž část I/N je již infikovaná. Uvážíme-li pravděpodobnost přenosu p, dostáváme $b\,p\,(n-i)\,\frac{I}{N}\,dt$ nově infikovaných komárů. Za předpokladu, že infekce nemá vliv na úmrtnost, je počet uhynulých komárů $m\,i\,dt$. Rychlost, s jakou se během časového intervalu dt změní počet infikovaných komárů, tedy je

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

Protože malárie existuje ve většině postižených zemí trvale, uvažoval Ross pouze ustálené stavy své soustavy rovnic: počet infikovaných lidí $I(t)$ a počet infikovaných komárů $i(t)$ zůstávají v čase konstantní ($dI/dt = 0$ a $di/dt = 0$). Vždy existuje ustálený stav s $I = 0$ a $i = 0$, který odpovídá nepřítomnosti malárie. Ross však hledal také takový ustálený stav, kdy $I > 0$ a $i > 0$, a odvodil, že pro něj platí

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

> Vydělením rovnic ustáleného stavu součinem $I \times i$ se z původní soustavy stává soustava dvou lineárních rovnic se dvěma neznámými $1/I$ a $1/i$,
>
> $$\frac{b\,p'}{I} - \frac{a}{i} = \frac{b\,p'}{N}\ , \quad -\frac{m}{I} + \frac{b\,p\,n}{N\,i} = \frac{b\,p}{N}\,.$$
>
> Vypočítat její řešení je snadné.

Lze si všimnout, že platí $I > 0$ a $i > 0$ je-li počet komárů nad kritickou hranicí:

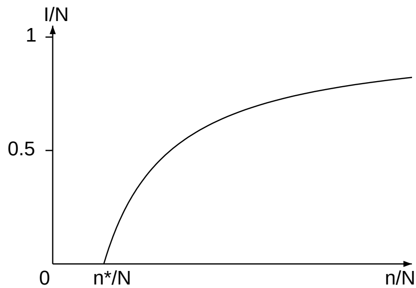
$$n > n^* = \frac{a\,m\,N}{b^2\,p\,p'}\,.$$

V takovém případě odpovídá ustálený stav situaci, kdy je nemoc endemická, tedy trvale přítomná. Ross dospěl k závěru, že pokud se počet komárů n sníží pod kritickou hranici n^*, zůstane jediný ustálený stav $I = 0$ a $i = 0$ a malárie by proto měla vymizet. Zejména není pro vymýcení malárie nutné, aby byli vyhubeni všichni komáři. A to je přesně to, co chtěl Ross svým modelem zdůraznit.

Pro ilustraci své teorie Ross vyhledal smysluplné číselné hodnoty pro parametry svého modelu. Předpokládal, že

- úmrtnost komárů je taková, že po deseti dnech je naživu pouze třetina z nich; takže $e^{-10m} = \frac{1}{3}$ odkud $m = (\log 3)/10$ za den;
- polovina nakažených lidí je infekční i po třech měsících, takže $e^{-90a} = 1/2$ odkud $a = (\log 2)/90$ denně;
- každý den někoho štípne osmina komárů, takže $e^{-b} = 1 - 1/8$ odkud $b = \log(8/7)$ za den;
- infikovaní komáři obvykle nejsou během prvních deseti dnů po nakažení infekční, neboť paraziti musí projít několika fázemi přeměny. Vzhledem k tomu, že třetina komárů může přežít deset dní, Ross předpokládal, že je také asi třetina všech nakažených komárů infekční: $p' = 1/3$;
- $p = 1/4$.

S těmito parametry mohl Ross podle vzorce (12.1) vypočítat podíl infikovaných lidí I/N jako funkci poměru n/N mezi velikostmi populací komárů a lidí. Své výsledky shrnul v tabulce, která je ekvivalentní obrázku 12.2.



Obrázek 12.2: Podíl I/N nakažených lidí v závislosti na poměru n/N mezi velikostmi populací komárů a lidí.

Tvar křivky ukazuje, že podíl nakažených lidí je vyšší než 50 % již tehdy, když je poměr n/N jen o málo vyšší než kritická hodnota n^*/N. Tento podíl se však příliš nemění, pokud se poměr n/N dále zvyšuje. To vysvětluje, proč nebyla korelace mezi počtem komárů a výskytem malárie nikdy dříve zaznamenána. Ross si však všiml, že číselná hodnota prahové hodnoty n^*/N je velmi citlivá na malé změny ve frekvenci štípání b, což ale nemění celkový tvar křivky na obr. 12.2. Jeho kvalitativní vysvětlení je důležitější než kvantitativní výsledky zahalené nejistotou danou nepřesnými číselnými hodnotami parametrů.

Pro interpretaci kritické hodnoty n^*, kterou Ross objevil[1], uvažujme jednoho nakaženého člověka zavlečeného do lidské a komáří populace, které jsou obě prosté malárie. Tento člověk zůstane nakažený v průměru po dobu rovnou $1/a$. Za jednotku času tento člověk obdrží bn/N štípnutí, takže v průměru $bn/(aN)$ štípnutí během celkové doby, kdy je nakažen. V průměru proto nakazí $bpn/(aN)$ komárů. Každý z těchto nakažených komárů žije průměrně po dobu rovnou $1/m$, kousne b/m lidí a z nich jich nakazí bp'/m. Celkově je proto po přenosu z prvního nakaženého člověka na komáry a z těchto komárů na další lidi průměrný počet nově nakažených lidí součinem předchozích dvou výsledků, tedy

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

Toto číslo $\mathscr{R}_0$ je počet sekundárních případů u lidí v důsledku jednoho primárního případu u lidí. Proces infekce, který probíhá kontinuálně v čase, lze proto uvažovat také pomocí po sobě jdoucích generací. Malárie se tak může šířit pouze tehdy, pokud $\mathscr{R}_0 > 1$. Tato podmínka je ekvivalentní podmínce $n > n^*$.

Za zmínku jistě stojí, že se Ross vyslovil pro obecnější využítí matematického modelování v epidemiologii:

> „Ve skutečnosti musí být celá epidemiologie, která se zabývá změnami výskytu nemoci v čase či prostoru, posuzována matematicky, ať už je do ní zapojeno jakkoli mnoho proměnných, má-li být vůbec posuzována vědecky. Říci, že nemoc závisí na určitých faktorech neznamená mnoho, dokud nedokážeme také odhadnout, do jaké míry jednotlivé faktory ovlivňují celkový výsledek. A matematický přístup není ve skutečnosti nic jiného než aplikace pečlivé úvahy na dané problémy.“

Ross byl v roce 1911 povýšen do rytířského stavu. Během první světové války se přestěhoval do Londýna a stal se poradcem britské armády. V roce 1923 vydal autobiografii *Paměti s vyčerpávajícím popisem velkého problému*

[1] Tato interpretace byla zdůrazněna až dlouho po Rossově práci.

malárie a jeho řešení. V roce 1926 byl slavnostně otevřen Rossův institut tropických nemocí (nyní součást Londýnské školy hygieny a tropické medicíny), jehož se stal ředitelem. Ross zemřel v Londýně v roce 1932.

Další čtení

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*. Roy Publishers, New York (1958)

Kapitola 13

Lotka, Volterra a matematická teorie boje o život (1920–1926)

V roce 1920 Alfred Lotka studoval model dravce a kořisti a ukázal, že velikosti populací mohou trvale oscilovat. Tuto studii rozpracoval ve své knize *Základy fyzikální biologie* z roku 1925. V roce 1926 se o stejný model náhodou začal zajímat italský matematik Vito Volterra, aby odpověděl na otázku, kterou si položil biolog Umberto d'Ancona: proč bylo v Jaderském moři během první světové války, kdy byl rybolov na velmi nízké úrovni, uloveno rybáři více dravých ryb?

V roce 1920 Lotka publikoval článek s názvem *Analytická poznámka o jistých rytmických vztazích v organických systémech*. Již několik let se zajímal o některé chemické reakce, které při laboratorních pokusech vykazovaly zvláštní přechodné oscilace. Cílem jeho článku bylo naznačit, že systém dvou biologických druhů může oscilovat i trvale. Jako příklad uvedl populaci býložravců živících se rostlinami. Analogicky k rovnicím používaným v chemické kinetice nechť $x(t)$ je celková hmotnost rostlin a $y(t)$ celková hmotnost býložravců v čase t. Lotka použil jako model následující soustavu diferenciálních rovnic

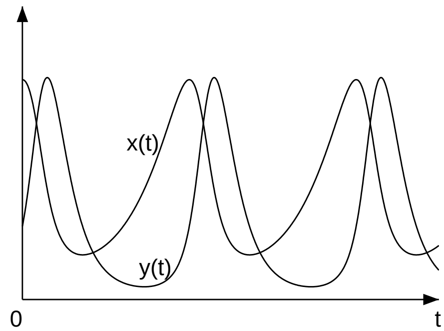
$$\frac{dx}{dt} = ax - bxy, \tag{13.1}$$

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

kde parametry a, b, c a d jsou kladné. Parametr a je rychlost růstu rostlin, když nejsou přítomni býložravci, zatímco c je rychlost poklesu populace býložravců, když nejsou přítomny rostliny. Výrazy $-bxy$ a dxy vyjadřují, že čím více je živočichů a rostlin, tím větší je přenos hmoty z rostlin směrem k živočichům (přenos zahrnuje určitou ztrátu hmoty, takže $d \leqslant b$). Při nastavení $dx/dt = 0$ a $dy/dt = 0$ si Lotka všiml, že existují dva ustálené stavy:

- $(x = 0, y = 0)$, populace býložravců vymřela a rostliny došly;
- $(x = c/d, y = a/b)$, koexistují býložravci a rostliny.

Bez důkazu také napsal, že pokud v čase $t = 0$ není $(x(0), y(0))$ jedním z těchto dvou ustálených stavů, pak funkce $x(t)$ a $y(t)$ periodicky oscilují: Existuje číslo $T > 0$ takové, že $x(t+T) = x(t)$ a $y(t+T) = y(t)$ pro všechny $t > 0$ (obrázek 13.1)[1]. Pokud jsou například rostliny velmi hojné, pak se populace býložravců zvýší, což způsobí pokles celkové hmotnosti rostlin. Když tato hmota přestane býložravcům stačit k obživě, někteří živočichové zemřou hlady a celková hmota rostlin začne opět růst, až dosáhne úrovně rovnající se její původní hodnotě. Tento jev se bude opakovat.



Obrázek 13.1: Oscilace celkové hmotnosti rostlin $x(t)$ a celkové hmotnosti býložravců $y(t)$ v závislosti na čase.

Lotka se tímto modelem zabýval o něco hlouběji v druhém článku publikovaném v roce 1920 s názvem *Netlumené oscilace odvozené ze zákona působení hmoty*. Vysvětlil, proč může systém periodicky kmitat. Vyplývá to ze skutečnosti, že bod $(x(t), y(t))$ musí zůstat na uzavřené trajektorii v rovině s x na vodorovné ose a y na svislé ose; přesněji řečeno v kvadrantu, kde $x \geqslant 0$ a $y \geqslant 0$ (obrázek 13.2).

Pokud totiž rovnici (13.1) vydělíme rovnicí (13.2), získáme po přeposkládání

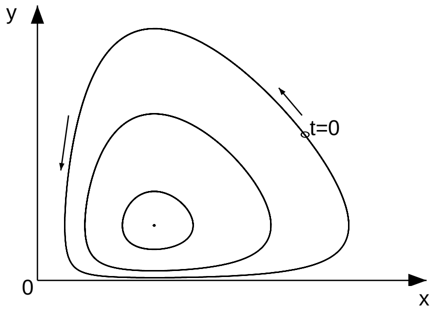
$$\left(-\frac{c}{x} + d\right)\frac{dx}{dt} = \left(\frac{a}{y} - b\right)\frac{dy}{dt}.$$

Integrací dostaneme

$$d\,x(t) - c\log x(t) = a\log y(t) - b\,y(t) + K,$$

kde K je konstanta, která závisí pouze na počátečních podmínkách. Proto bod $(x(t), y(t))$ zůstává na křivce $d\,x - c\log x = a\log y - b\,y + K$, která je uzavřenou křivkou (obrázek 13.2).

[1]Perioda T závisí na počátečních podmínkách, ale Lotka si tuto skutečnost uvědomil až v roce 1925.



Obrázek 13.2: Diagram s celkovou hmotností rostlin $x(t)$ na vodorovné ose a celkovou hmotností býložravců $y(t)$ na svislé ose. Tři uzavřené křivky kolem ustáleného stavu odpovídají různým počátečním podmínkám.

Trajektorie $(x(t), y(t))$ se otáčí kolem ustáleného stavu $(c/d, a/b)$ proti směru hodinových ručiček, jak lze snadno zjistit studiem znaménka dx/dt a dy/dt. V blízkosti ustáleného stavu systém vykazuje malé oscilace s periodou rovnou $2\pi/\sqrt{ac}$.

Nastavte $x = \frac{c}{d} + x^*$ a $y = \frac{a}{b} + y^*$, kde $|x^*| \ll \frac{c}{d}$ a $|y^*| \ll \frac{a}{b}$. Pak

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

Z těchto dvou rovnic získáme

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

Tyto rovnice jsou stejné jako pro kmitání jednoduchého kyvadla ve fyzice. Perioda je $2\pi/\sqrt{ac}$.

Raymond Pearl, který první článek z roku 1920 poslal do *Proceedings of the National Academy of Sciences*, pomohl Lotkovi získat dvouleté stipendium na Univerzitě Johnse Hopkinse, aby mohl napsat knihu s názvem *Základy fyzikální biologie*. Kniha vyšla v roce 1925. V části shrnující práci z roku 1920 je také zmíněno, že systémy dvou druhů, jednoho druhu hostitele a jednoho druhu parazita nebo jednoho druhu kořisti a jednoho druhu predátora, lze popsat stejným modelem (13.1)–(13.2). Bohužel Lotkova kniha v době svého vydání nevzbudila velkou pozornost. Slavný matematik Volterra však stejný model brzy poté nezávisle znovu objevil při studiu problému rybolovu.

Vito Volterra se narodil v židovském ghettu v Anconě v roce 1860, krátce před sjednocením Itálie, kdy město ještě patřilo papežskému státu. Byl jedináčkem. Jeho otec, obchodník s látkami, zemřel, když byly Vitovi dva roky, a rodina zůstala bez peněz. Volterra byl dobrým středoškolákem a navzdory chudobě se mu podařilo pokračovat ve studiu, nejprve na univerzitě ve Florencii a později na *Scuola Normale Superiore* v Pise. V roce 1882 získal doktorát z fyziky a následujícího roku se stal profesorem mechaniky na univerzitě v Pise. V roce 1892 nastoupil na univerzitu v Turíně a v roce 1900 přešel na katedru matematické fyziky na univerzitu *La Sapienza* v Římě. Roku 1905 se stal senátorem. Řada jeho přednášek, které přednesl v Římě nebo na zahraničních univerzitách, vyšla knižně: *Tři lekce o některých nejnovějších pokrocích v matematické fyzice* (Clark univerzita, 1909), *Přednášky o integrálních a integro-diferenciálních rovnicích* (Řím, 1910), *Přednášky o funkcích na křivkách* (Paříž, 1912), *Teorie permutabilních funkcí* (Princeton, 1912). Za první světové války sloužil jako důstojník italské armády a vedl úřad pro válečné vynálezy. Po válce se aktivně podílel na založení Italské matematické unie (1922) a Italské národní rady pro výzkum (1923) a stal se jejím prvním předsedou. Stal se také předsedou Mezinárodní komise pro vědecké studium Středozemního moře (1923) a předsedou Accademia dei Lincei (1924). Další monografie, napsaná ve spolupráci s J. Pérèsem, *Přednášky o složených a permutabilních funkcích*, vyšla v roce 1924.

Obrázek 13.3: Volterra (1860–1940) získal v roce 1900 doktorát *honoris causa* na univerzitě v Cambridge.

V roce 1925, ve svých 65 letech, se Volterra zajímal o studii zoologa Umberta D'Ancony, který se později stal jeho zetěm, o podílu paryb (jako jsou žraloci a rejnoci) vyložených při rybolovu v letech 1905–1923 ve třech přístavech Jaderského moře: Terst, Fiume[2] a Benátky. D'Ancona si všiml, že podíl těchto ryb se zvýšil během první světové války, kdy se snížilo intenzita rybolovu (tabulka 13.1).

Tabulka 13.1: Podíl chrupavčitých ryb v rybářství Terstu, Fiume a Benátek před první světovou válkou, během ní a po ní.

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

rok	1917	1918	1919	1920	1921	1922	1923
Terst	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
Benátky	-	-	30,9	25,3	25,9	25,8	26,6

Paryby jsou predátory menších ryb, a proto se zdá, že snížení intenzity rybolovu zvýhodňuje dravé druhy. Volterra, který o Lotkově práci nevěděl, vysvětlil toto pozorování pomocí stejného modelu.

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

kde $x(t)$ znamená počet kořisti a $y(t)$ počet predátorů. Stejně jako Lotka si všiml, že tento systém může periodicky oscilovat s periodou T, která závisí na počáteční podmínce (x_0, y_0). Všiml si také, že

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

Integrací přes jednu periodu T (takže $x(0) = x(T)$ a $y(0) = y(T)$) získal následující výsledky

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

Průměr počtu kořisti i predátorů za jedno období je tedy nezávislý na počátečních podmínkách. Navíc pokud se intenzita rybolovu sníží, zvýší se míra růstu kořisti a, zatímco míra vymírání predátorů c se sníží. Proto se průměr

[2] Nyní Rijeka v Chorvatsku.

$x(t)$ snižuje a průměr $y(t)$ zvyšuje: podíl predátorů se zvyšuje. Přesně to bylo pozorováno u statistik rybolovu v Jaderském moři.

Volterra publikoval svůj článek nejprve v italštině v roce 1926. Anglické shrnutí vyšlo o několik měsíců později v časopise *Nature*. Lotka upozornil Volterru a další vědce na důležitost studia systémů predátor-kořist. Jeho článek z roku 1920 a kniha z roku 1925 však nebudou vždy zmiňovány. Lotka tehdy již pracoval pro pojišťovací společnost, takže se jeho práce zaměřovala na demografii lidí. Volterra pokračoval v práci na variantách systému predátor-kořist ještě deset let. V letech 1928–1929 přednesl sérii přednášek na nově vytvořeném Institutu Henriho Poincarého v Paříži. Zápisky z těchto přednášek byly publikovány v roce 1931 pod názvem *Přednášky matematické teorie boje o život*. V roce 1935 vydal Volterra ve spolupráci s Umbertem D'Anconou další knihu *Biologické asociace z matematického hlediska*.

Ačkoli se zdá, že model predátor-kořist vysvětluje údaje o rybolovu správně, debata o realističnosti zjednodušených modelů v ekologii teprve začínala a stále je předmětem vědeckých sporů. V současné době je model predátor-kořist známý také jako Lotka-Volterra model a je jedním z nejčastěji citovaných modelů v ekologii.

V roce 1931 Volterra odmítl vyjádřit věrnost Mussolinimu. Přišel o profesuru na univerzitě v Římě a byl vyloučen z italských vědeckých akademií, jejichž byl jedním z nejznámějších členů. Od té doby se zdržoval převážně mimo Itálii, cestoval po Evropě a přednášel. Společně s J. Pérèsem vydal první díl obecné teorie funkcionálů. (1936) a spolu s B. Hostinským knihu *Infinitezimální lineární operace* (1938). Zemřel v Římě v roce 1940.

Další čtení

1. Goodstein, J.R.: *The Volterra Chronicles, The Life and Times of an Extraordinary Mathematician 1860–1940*. American Mathematical Society (2007)
2. Guerraggio, A., Nastasi, P.: *Italian Mathematics between the Two World Wars*. Birkhäuser, Basel (2005)
3. Israel, G., Gasca, A.M.: *The Biology of Numbers – The Correspondence of Vito Volterra on Mathematical Biology*. Birkhäuser, Basel (2002)
4. Kingsland, S.E.: *Modeling Nature, Episodes in the History of Population Ecology*, 2nd edn. University of Chicago Press (1995)
5. Lotka, A.J.: Analytical note on certain rhythmic relations in organic systems. *Proc. Natl. Acad. Sci.* 6, 410–415 (1920) pnas.org
6. Lotka, A.J.: Undamped oscillations derived from the law of mass action. *J. Amer. Chem. Soc.* 42, 1595–1599 (1920) archive.org
7. Lotka, A.J.: *Elements of Physical Biology*. Williams & Wilkins, Baltimore (1925) archive.org

8. Volterra, V.: Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Mem. Accad. Lincei* 6, 31–113 (1926) → *Opere matematiche*, vol. 5, Accademia nazionale dei Lincei, Roma (1962) liberliber.it
9. Volterra, V.: Fluctuations in the abundance of a species considered mathematically. *Nature* 118, 558–560 (1926). → L.A. Real, J.H. Brown (eds.) *Foundations of Ecology*, 283–285. University of Chicago Press (1991)
10. Volterra, V.: *Leçons sur la théorie mathématique de la lutte pour la vie*. Gauthier-Villars, Paris (1931) 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)

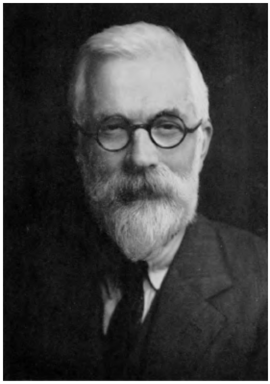
Kapitola 14

Fisher a přírodní výběr (1922)

V roce 1922 publikoval britský matematický biolog Ronald Fisher velmi vlivný článek o populační genetice. Tato kapitola se zabývá pouze jednou jeho částí zaměřenou na variantu Hardyho-Weinbergova modelu, která která navíc obsahuje přírodní výběr. Fisher ukázal, že pokud má heterozygot největší zdatnost, pak mohou obě alely existovat současně. Pokud ji má menší než některý z obou homozygotů, pak jedna z alel zanikne. Tím odpověděl na základní otázku, proč geny mají více alel.

Ronald Aylmer Fisher se narodil v Londýně v roce 1890 jako poslední ze šesti dětí. Jeho otec byl aukcionářem, který později zbankrotoval. Fisher studoval v letech 1909–1913 matematiku a fyziku na *Gonville and Caius College* Cambridgeské univerzity. V té době došlo k rychlému rozvoji genetiky. Od roku 1911 se Fisher účastnil schůzí Eugenické společnosti založené Galtonem. Začal se věnovat statistickým problémům souvisejícím s Galtonovou a Mendelovou prací. Po ukončení univerzitních studií strávil jedno léto prací na farmě v Kanadě a poté pracoval pro *Mercantile and General Investment Company* v londýnské City. Kvůli své silné krátkozrakosti nebojoval v první světové válce, přestože se hlásil jako dobrovolník. Válečná léta strávil vyučováním na středních školách. Ve volném čase se staral o farmu a pokračoval ve výzkumu. Získal nové důležité výsledky spojující korelační koeficienty s mendelovskou genetikou. V roce 1919 začal pracovat jako statistik na Rothamstedově experimentální stanici zaměřené na zemědělství.

V roce 1922 publikoval Fisher článek s názvem *O poměru dominance*. Vedle několika důležitých nových myšlenek se tento článek zabýval matematickým modelem kombinujícím Mendelovy zákony a přírodní výběr, který tvoří jeden ze základů Darwinovy evoluční teorie. Fisher uvažoval stejnou situaci jako Hardy se dvěma alelami A a a a s náhodným křížením. Navíc však předpokládal, že jedinci s genotypy AA, Aa a aa mají různou úmrtnost před dosažením dospělosti, čímž modeloval přírodní výběr. Označíme-li p_n, $2q_n$ a r_n četnosti těchto tří genotypů mezi dospělými jedinci v generaci n, pak v generaci $n+1$ jsou novorozenci s těmito genotypy $(p_n+q_n)^2$, $2(p_n+q_n)(q_n+r_n)$ a $(q_n+r_n)^2$. Nechť u, v a w jsou příslušné pravděpodobnosti přežití od narození do dospělosti. Potom četnosti genotypů mezi dospělými jedinci v gene-



Obrázek 14.1:
Fisher (1890–1962)

raci $n+1$ jsou p_{n+1}, $2q_{n+1}$ a r_{n+1}, přičemž

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

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

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

kde jsme pro zjednodušení vzorců položili

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

Připomeňme si, že $p_n + 2q_n + r_n = 1$. Vidíme, že v případě $u = v = w$ (tj. když neexistuje přírodní výběr), se systém (14.1)–(14.3) redukuje na systém (11.1)–(11.3) uvažovaný Hardym.

Nechť $x_n = p_n + q_n$ je frekvence alely A mezi dospělými jedinci v generaci n. Potom $q_n + r_n = 1 - x_n$ je frekvence alely a. Sečteme-li (14.1) a (14.2), získáme výsledek

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

Tuto rovnici lze přepsat ve tvaru

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

Tato diferenční rovnice má vždy alespoň dva rovnovážné body, tj. stavy, kdy frekvence x_n zůstává v průběhu generací konstantní: $x = 0$ (populace se skládá výhradně z homozygotů aa) a $x = 1$ (populace se skládá výhradně z homozygotů AA).

Pomocí rovnice (14.4) můžeme ukázat, že pokud má homozygot AA větší šanci na přežití než ostatní dva genotypy ($u > v$ a $u > w$), pak alela a bude z populace postupně mizet. Tento případ by neměl být v přírodě příliš častý, pokud víme, že obě alely existují současně. Pokud však heterozygot Aa má selektivní výhodu oproti homozygotům AA a aa ($v > u$ a $v > w$), pak mohou v populaci koexistovat tyto tři genotypy. Tento případ je v reálných populacích nejčastější. Lze jím vysvětlit „životnost“ kříženců pozorovanou zemědělci.

Rovnovážný stav $x = 1$ je totiž stabilní, pokud $u > v$, neboť $x_{n+1} - x_n \approx (1 - x_n)(u - v)/u$, když x_n je blízko 1. Populace směřuje k tomuto rovnovážnému stavu. Podobně, pokud $w > v$, pak je rovnovážný stav $x = 0$ stabilní. Pokud ale $u < v$ a $w < v$, pak existuje třetí rovnovážný stav

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

splňující nerovnosti $0 < x^* < 1$ a který je stabilní, jak si můžeme ověřit. Rovnovážný stav x^* odpovídá směsi tří genotypů.

Jednoduchou kombinací Mendelových zákonů a hypotézy přírozeného výběru (zde vyjádřenému rozdílnými pravděpodobnostmi přežití do dospělosti pro tři možné genotypy) lze tedy vysvětlit dvě různé situace – koexistenci nebo zánik genotypů. Po Fisherovi tento model rozpracovali také John Haldane (viz kapitola 17) a Sewall Wright (viz kapitola 19).

V předstihu před kapitolou 20 si ještě všimněme situace, kdy alela A je dominantní a fenotyp s A je zdatnější než recesivní homozygot aa. Jestliže je v takovém případě číselný poměr $u : v : w$ roven $1 : 1 : 1 - \varepsilon$, pak rovnice (14.4) nabývá tvaru

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

pro $\varepsilon \ll 1$. Jestliže pravděpodobnost přežití heterozygota Aa leží v polovině mezi pravděpodobnostmi přežitídvou homozygotů, pak poměr $u : v : w$ se rovná poměru $1 : 1 - \varepsilon/2 : 1 - \varepsilon$ a

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

pro $\varepsilon \ll 1$.

Ve Rothamstedově stanici Fisher analyzoval dlouhodobé údaje týkající se výnosů plodin a meteorologie. Významně však také přispěl k metodologii statistiky. V roce 1925 vydal velice úspěšnou knihu *Statistické metody pro výzkumné pracovníky*, která vyšla v mnoha vydáních. Roku 1929 se Fisher stal členem Královské společnosti. V roce 1930 vydal knihu *Genetická teorie přírodního výběru*, která se stala milníkem v historii populační genetiky. V roce 1933 se stal profesorem eugeniky na *University College* v Londýně a vystřídal Karla Pearsona v Galtonově laboratoři. V roce 1943 přešel na katedru genetiky na Cambridgeské univerzitě, tentokrát jako nástupce R. C. Punnetta (viz kapitola 11). Vydal také několik dalších knih: *Plánování experimentů* (1935), *Teorie příbuzenské plemenitby* (1949) a *Statistické metody a vědecká inference* (1956). V roce 1952 byl pasován na rytíře a po odchodu do důchodu v roce 1959 se usadil v Austrálii. Zemřel v Adelaide roku 1962. K další části jeho díla se vrátíme v kapitole 20.

Další čtení

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)

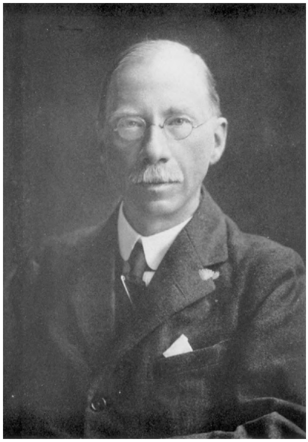
Kapitola 15

Yule a evoluce (1924)

V roce 1924 britský statistik Yule studoval model evoluce, ve kterém nové druhy mohou vznikat mutacemi malého rozsahu a nové rody mutacemi velkého rozsahu. Jeho cílem bylo vysvětlit rozložení počtu druhů v rámci rodů, kdy většina rodů obsahuje pouze jeden druh a jen několik málo rodů obsahuje velký počet druhů. Stochastický „proces zrodu“, který Yule ve svém modelu zavedl, je dodnes základním nástrojem při studiu fylogenetických stromů a v mnoha dalších oblastech.

George Udny Yule se narodil v roce 1871 ve Skotsku, jeho otec zastával vysokou funkci v britské správě v Indii. V 16 letech začal Yule studovat na *University College* v Londýně, aby se stal inženýrem. V roce 1892 změnil své zaměření a strávil jeden rok výzkumem v Bonnu pod vedením fyzika Heinricha Hertze, který o několik let dříve prokázal existenci elektromagnetických vln. Když se Yule vrátil do Anglie, nabídl mu Karl Pearson místo docenta aplikované matematiky na *University College* v Londýně. Yule se po Pearsonově vzoru začal věnovat statistice. V roce 1911 vydal knihu *Úvod do teorie statistiky*, která se dočkala čtrnácti dotisků. V následujícím roce se přestěhoval na univerzitu v Cambridge. Ve své výzkumné práci se zabýval teoretickými aspekty statistiky, ale také aplikacemi v zemědělství a epidemiologii. V roce 1922 se stal členem Královské společnosti.

V roce 1924 Yule publikoval článek s názvem *Matematická teorie evoluce založená na závěrech Dr. J. C. Willise*. Willis byl jeho kolega z Královské společnosti, který v roce 1922 vydal knihu s názvem *Věk a oblast, studie o prostorovém rozšíření a původu druhů*. Studoval v ní rozšíření druhů mezi jednotlivými rody v rámci klasifikace rostlin a živočichů. Údaje, které shromáždil, ukázaly, že většina rodů obsahuje pouze jeden druh, že stále méně rodů obsahuje vyšší počet druhů a že stále existuje několik rodů obsahujících velké množství druhů. V tabulce 15.1 jsou uvedeny údaje týkající se hadů, ještěrů a dvou čeledí brouků (*Chrysomelidae* a *Cerambycinae*). Tehdy známých 1580 druhů ještěrů bylo zařazeno do 259 rodů, přičemž 105 rodů obsahovalo pouze jeden druh, 44 pouze dva druhy, 23 pouze tři druhy atd. Dva rody obsahovaly více než sto druhů. U ostatních čeledí živočichů a rostlin mělo rozdělení rodů podle počtu druhů, které obsahují, velmi podobný tvar.

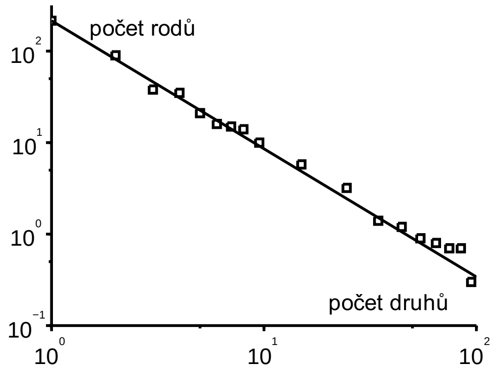


Obrázek 15.1:
Yule (1871–1951)

Tabulka 15.1: Údaje sestavil Willis.

Počet druhů	Počty rodů *Chrysomelidae*	*Cerambycinae*	Hadi	Ještěři
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
celkem	627	1 024	293	259

Yule navrhl, aby se Willis pokusil zakreslit svá data do grafu se stupnicemi v logarimickém měřítku. To přineslo pozoruhodný výsledek (obrázek 15.2): logaritmus počtu rodů $\log(Q_n)$ obsahujících n druhů klesá téměř lineárně v závislosti na $\log(n)$. Jinými slovy, existují konstanty $\alpha > 0$ a $\beta > 0$ takové, že $Q_n \approx \alpha n^{-\beta}$: rozdělení se řídí „mocninným zákonem“. Ve svém článku z roku 1924 hledal Yule matematický model evoluce, který by takové statistické rozdělení vysvětlil.



Obrázek 15.2: Počet rodů v závislosti na počtu druhů, měřítkem na osách je dekadický logaritmus. Údaje pro čeleď Chrysomelidae. Pro vyhlazení výkyvů při velkém počtu druhů n byl počet rodů počítán jako průměr přes rozsahy hodnot obsahující n, viz tabulka 15.1. Proto může být průměrný počet rodů pro jednu hodnotu n menší než 1.

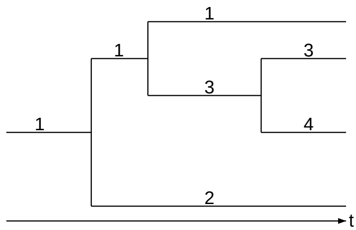
Za tímto účelem si nejprve představil stochastický model se spojitým časem[1] pro růst počtu druhů v rámci jednoho rodu (obrázek 15.3). Vycházel z předpokladu, že v čase $t = 0$ existuje pouze jeden druh a že pravděpodobnost, že se z jednoho druhu zrodí mutací nový druh téhož rodu během „malého“ (na časové škále evoluce) časového intervalu dt, je rovna $r\,dt$, přičemž $r > 0$.

Nechť $p_n(t)$ je pravděpodobnost, že v čase t existuje n druhů (n je celé číslo, t je reálné číslo). Pro výpočet $p_n(t+dt)$ uvažoval Yule několik případů:

- jestliže v čase t existuje $n-1$ druhů, má každý druh pravděpodobnost $r\,dt$ vzniku jednoho nového druhu mezi t a $t+dt$; v limitě pro $dt \to 0$ bude v čase $t+dt$ existovat n druhů s pravděpodobností $(n-1)\,r\,dt$;

- jestliže v čase t existuje n druhů, bude v čase $t+dt$ existovat $n+1$ druhů s pravděpodobností $n\,r\,dt$.

[1] McKendrick (viz kapitola 16) začal studovat tyto modely v populační dynamice již v článku publikovaném v roce 1914.



Obrázek 15.3: Simulace vývoje počtu druhů v rámci jednoho rodu. Druh 1 vytváří druhy 2 a 3. Druh 3 vytváří druh 4.

$p_n(t)$ je tedy dáno následující soustavou diferenciálních rovnic

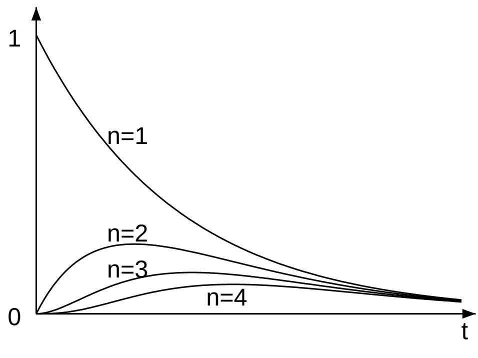
$$\frac{dp_1}{dt} = -r\,p_1\,, \tag{15.1}$$

$$\frac{dp_n}{dt} = (n-1)\,r\,p_{n-1} - n\,r\,p_n \quad \text{pro všechna } n \geqslant 2. \tag{15.2}$$

Z první rovnice dostaneme $p_1(t) = e^{-rt}$, protože $p_1(0) = 1$. Lze ukázat, že řešení druhé rovnice, které splňuje počáteční podmínku $p_n(0) = 0$, je

$$p_n(t) = e^{-rt}\,(1 - e^{-rt})^{n-1} \quad \text{pro všechna } n \geqslant 2, \tag{15.3}$$

viz obrázek 15.4. V pevně zvoleném t je tedy rozdělení pravděpodobností $(p_n(t))_{n\geqslant 1}$ „geometrické“ s poměrem mezi dvěma po sobě následujícími členy rovným $1 - e^{-rt}$.



Obrázek 15.4: Pravděpodobnost $p_n(t)$, že v čase t existuje n druhů stejného rodu, pro $1 \leqslant n \leqslant 4$.

Nejprve si všimneme, že rovnice (15.2) je ekvivalentní rovnici

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

ze kterých můžeme postupně vypočítat $p_2(t)$, $p_3(t)$... Dostaneme

$$p_2(t) = e^{-rt}\,(1-e^{-rt}), \quad \text{pak} \quad p_3(t) = e^{-rt}\,(1-e^{-rt})^2,$$

což naznačuje vzorec (15.3) pro obecné řešení. Nakonec lze dosazením ověřit, že tento vzorec (15.3) je řešením rovnice (15.4).

Yule také ze vzorce (15.3) odvodil, že střední hodnota počtu druhů roste v čase exponenciálně: $\sum\limits_{n=1}^{+\infty} n\, p_n(t) = e^{rt}$.

Nejprve si všimněme, že pro $|x| < 1$,

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

Pak

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

Je-li T doba, za kterou se počet druhů zdvojnásobí, definovaná vztahem $e^{rT} = 2$, pak pravděpodobnostní rozdělení $(p_n(t))_{n\geqslant 1}$ počtu druhů v čase $t = T$ je geometrické s kvocientem $1/2$: $1/2$, $1/4$, $1/8$, $1/16$... V čase $t = kT$ je geometrické s kvocientem $1 - 1/2^k$ a $p_1(kT) = 1/2^k$.

Yule dále uvažoval, že souběžně s růstem počtu druhů patřících do stejného rodu probíhá podobný proces v důsledku mutací většího rozsahu vedoucích ke vzniku nových rodů. Nechť $s\,dt$ je pravděpodobnost, že existující rod vytvoří nový rod během malého časového intervalu dt. Stejně jako v předchozím případě, za předpokladu, že v čase $t = 0$ existuje pouze jeden rod, střední hodnota počtu rodů v čase t je e^{st}. Průměrný počet rodů vytvořených za jednotku času v čase t je derivace tohoto výrazu a je rovna $s\,e^{st}$. V limitě[2] pro $t \to +\infty$, pak průměrný počet rodů, které v čase t existovaly mezi x a $x+dx$

[2] Yule uvažoval i případ, kdy t nelze považovat za příliš velké ve srovnání s dobou e^{st}, za kterou se velikost populace zdvojnásobí. Výpočty jsou o něco složitější, ale konečné výsledky se příliš neliší.

jednotkami času, je roven $s\,e^{s\,(t-x)}\,dx$. Pravděpodobnost, že v čase t bude náhodně vybraný rod existovat mezi x a $x+dx$ jednotkami času, je $s\,e^{-s\,x}\,dx$.

Jestliže náhodně vybraný rod v čase t existoval mezi x a $x+dx$ jednotkami času, pravděpodobnost, že tento rod obsahuje n druhů, je podle vzorce (15.3) rovna $e^{-r\,x}\,(1-e^{-r\,x})^{n-1}$ pro všechna $n \geqslant 1$. Pravděpodobnost q_n, že náhodně vybraný rod bude v čase t obsahovat n druhů, je tedy:

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

Nechť $u = r/s$. Lze snadno ukázat, že $q_1 = 1/(1+u)$ a že

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

pro všechna $n \geqslant 2$.

> Skutečně máme $(1-e^{-r\,x})^{n-1} = (1-e^{-r\,x})^{n-2}\,(1-e^{-r\,x})$. Takže
>
> $$q_n = q_{n-1} - s\int_0^{+\infty} e^{-(r+s)\,x}\,(1-e^{-r\,x})^{n-2}\,e^{-r\,x}\,dx\,.$$
>
> Integrací metodou per-partes dostáváme
>
> $$q_n = q_{n-1} - \frac{r+s}{(n-1)\,r}\,q_n\,, \qquad q_n = \frac{(n-1)\,r/s}{1+n\,r/s}\,q_{n-1}\,.$$

Vzorec (15.5) ukazuje, že posloupnost pravděpodobností $(q_n)_{n\geqslant 1}$ je klesající. Maxima je tedy dosaženo pro $n=1$: většina rodů obsahuje pouze jeden druh. To přesně odpovídá získaným datům. Navíc pokles q_n směrem k 0 pro n jdoucí do nekonečna, je relativně pomalý, protože $q_n/q_{n-1} \to 1$. To může vysvětlovat, proč některé rody obsahují velký počet druhů. Přesněji řečeno, Yule ukázal, že $\log q_n$ klesá lineárně s $\log(n)$.

> Zavedeme Eulerovu gama funkci $\Gamma(z) = \int_0^{+\infty} t^{z-1}\,e^{-t}\,dt$. Pak $\Gamma(n+1) = n! = n \times (n-1) \times \cdots \times 2 \times 1$, když n je celé číslo, a $\Gamma(z+1) = z\Gamma(z)$. Takže (15.5) má tvar
>
> $$q_n = \frac{(n-1)!}{u(1+\frac{1}{u})(2+\frac{1}{u})\cdots(n+\frac{1}{u})} = \frac{\Gamma(n)\,\Gamma(1+\frac{1}{u})}{u\,\Gamma(n+1+\frac{1}{u})}\,.$$
>
> Stirlingova aproximace však dává $\log\Gamma(n) \approx n\log n - n - \frac{1}{2}\log n + \text{konst.}$

Podobně, $\log \Gamma(n+1+1/u) \approx n \log n - n + (\frac{1}{u} + \frac{1}{2}) \log n + \text{konst}$. Konečně $\log q_n \approx -(1+\frac{1}{u}) \log n + \text{konst}$.

Uvažujme například podřád ještěrů. Parametr u lze odhadnout z podílu $q_1 = 1/(1+u)$ rodů, které obsahují pouze jeden druh. Podle tabulky 15.1 máme $q_1 = 105/259$, takže $u \approx 1{,}467$. Teoretickou pravděpodobnost q_n a očekávaný počet Q_n rodů obsahujících n druhů pak můžeme vypočítat vynásobením q_n celkovým počtem druhů, který je 259 (tabulka 15.2). Yule si všiml, že vypočtené hodnoty vcelku dobře odráží skutečnost vzhledem k jednoduchosti modelu[3], který nebere v úvahu například katastrofy, jimiž druhy prošly během milionů let evoluce.

Tabulka 15.2: Srovnání reálných a vypočtených počtů druhů v rodě v případě ještěrů (1 580 druhů zařazených do 259 rodů).

Počet druhů v rodě	Zjištěný počet rodů	Vypočtený počet rodů
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
celkem	259	259

[3]Pro počet rodů obsahujících více než 100 druhů získal Yule lepší výsledky než v tabulce 15.2, když uvážil, že t není velké ve srovnání s dobou e^{st}, za kterou se populace zdvojnásobí.

Po roce 1931 Yule postupně odcházel z Cambridgeské univerzity do důchodu. Začal se zajímat o statistické rozložení délky vět, aby bylo možné identifikovat autory knih. Používal zejména knihy vydané Johnem Grauntem (viz kapitola 2), ale pravděpodobně se inspiroval Williamem Pettym. V roce 1944 vydal knihu *Statistické studium literární slovní zásoby*. Zemřel v roce 1951.

V dnešní době se Yuleho model stále používá k analýze „fylogenetických stromů" (rodokmenů druhů). Tyto stromy, podobné tomu na obrázku 15.3, jsou lépe poznány díky novým datům pocházejícím z molekulární biologie. Použití stochastického procesu definovaného rovnicemi (15.1)–(15.2) se však neomezuje pouze na teorii evoluce. Tento proces je základním stavebním kamenem mnoha modelů v populační dynamice, od mikroskopické úrovně (pro modelování např. kolonií bakterií) až po makroskopickou úroveň (pro modelování počátku epidemie). Nazývá se „čistý proces zrodu" nebo „Yuleho proces". Jednoduchá varianta zahrnuje pravděpodobnost $m\,dt$ úmrtí během libovolně malého časového intervalu dt: očekávaná velikost populace v čase t pro tento „proces zrodu a zániku" je pak $e^{(r-m)t}$. Pokud jde o rozdělení pravděpodobnosti (15.5), někdy se nazývá Yuleho rozdělení. Rozdělení s konci splňujícími mocninné zákony přitahují velkou pozornost v různých oblastech vědy. Jedním z příkladů je studium epidemií na náhodných sítích, kde se pravděpodobnostní rozložení stupňů vrcholů řídí mocninným zákonem.

Další čtení

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 (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

Kapitola 16

McKendrick a Kermack o modelování epidemií (1926–1927)

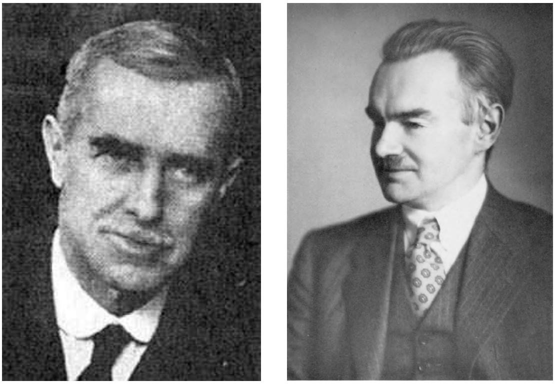
V roce 1926 McKendrick studoval stochastický model epidemie a nalezl metodu výpočtu pravděpodobnosti, že epidemie dosáhne určité dané velikosti. Odvodil také parciální diferenciální rovnici popisující vývoj věkově strukturované populace ve spojitém čase. V roce 1927 studovali Kermack a McKendrick deterministický model epidemie a určili rovnici pro konečnou velikost epidemie, obsahující určitý práh velikosti populace. Velké epidemie mohou nastat nad touto prahovou hodnotou, ale nikoliv pod ní. Tyto práce se v současné epidemiologii stále velmi využívají.

Anderson Gray McKendrick se narodil v roce 1876 v Edinburghu jako poslední z pěti dětí. Vystudoval medicínu na univerzitě v Glasgow, kde jeho otec působil jako profesor fyziologie. V roce 1900 vstoupil do Indické lékařské služby. Předtím doprovázel Ronalda Rosse v Sieře Leone, kde bojovali proti malárii. Poté sloužil 18 měsíců v armádě v Súdánu. Po příjezdu do Indie byl jmenován lékařem ve věznici v Bengálsku, kde se snažil zvládnout úplavici. V roce 1905 nastoupil do nového Ústředního ústavu pro lékařský výzkum v Kasauli (na severu Indie). Pracoval na výzkumu vztekliny, ale také studoval matematiku. V roce 1920 se po nákaze tropickou nemocí vrátil do Edinburghu a stal se superintendantem laboratoře v *Royal College of Physicians*.

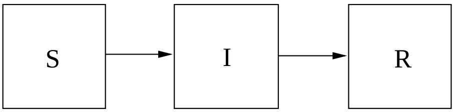
V roce 1926 McKendrick publikoval článek *Aplikace matematiky na lékařské problémy*, který obsahoval několik nových myšlenek. Představil zejména matematický model epidemií ve spojitém čase, který zohledňoval stochastický charakter nakažení se a uzdravení.

Uvažujme populaci o velikosti N s na počátku pouze jednou nakaženou osobou. Lidé mohou postupně procházet třemi stavy: náchylným či vnímavým stavem S, infekčním stavem I způsobeným nákazou a uzdraveným stavem R (obrázek 16.2)[1].

[1]Model Daniela Bernoulliho (viz kapitola 4) zahrnoval stavy S a R, ale ne I, přičemž doba trvání infekce byla mnohem kratší než průměrná délka života.



Obrázek 16.1: McKendrick (1876–1943) a Kermack (1898–1970)



Obrázek 16.2: Možné stavy: náchylný / vnímavý (S), infekční (I), uzdravený (R).

Nechť $p_{i,r}(t)$ je pravděpodobnost, že populace obsahuje v čase t přesně i lidí ve stavu I a r lidí ve stavu R, kde i a r jsou celá čísla taková, že $1 \leqslant i+r \leqslant N$. V takovém případě řekneme, že populace je ve stavu (i,r). Počet vnímavých lidí je $s = N - i - r$. V návaznosti na Rossovu práci o malárii (viz kapitola 12) McKendrick předpokládal, že během malého časového intervalu dt je pravděpodobnost výskytu jedné nové nákazy rovna $a\,s\,i\,dt$ (tj. úměrná počtu vnímavých osob i počtu infekčních osob). Pravděpodobnost jednoho nového uzdravení je rovna $b\,i\,dt$. Oba parametry a a b jsou kladné. Pro výpočet pravděpodobnosti $p_{i,r}(t+dt)$ je třeba rozlišit několik případů:

- populace je v čase t ve stavu $(i-1,r)$ a jedna nová nákaza přesune populaci do stavu (i,r) mezi časy t a $t+dt$; pravděpodobnost této události je $a\,s\,(i-1)\,dt$, kde $s = N-(i-1)-r$;

- populace je v čase t ve stavu (i,r) a jedna nová nákaza přesune populaci do stavu $(i+1,r)$ mezi časy t a $t+dt$; pravděpodobnost této události je $a\,s\,i\,dt$, kde $s = N-i-r$;

- populace je v čase t ve stavu $(i+1,r-1)$ a jedno nové uzdravení přesune populaci do stavu (i,r) mezi časy t a $t+dt$; pravděpodobnost této události je $b\,(i+1)\,dt$;

- populace je v čase t ve stavu (i,r) a jedno nové uzdravení přesune populaci do stavu $(i-1,r+1)$ mezi časy t a $t+dt$; pravděpodobnost této události je $b\,i\,dt$.

McKendrick tedy získal rovnice

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

pro $1 \leqslant i+r \leqslant N$. První člen na pravé straně chybí, je-li $i=0$, zatímco třetí člen chybí, pokud $r=0$. Počáteční podmínky jsou $p_{i,r}(0)=0$ pro všechny (i,r) kromě $p_{1,0}(0)=1$.

Pomocí tohoto modelu se McKendrickovi podařilo vypočítat pravděpodobnost, že epidemie skončí s celkovým počtem n nakažených lidí, což je limita $p_{0,n}(t)$ pro $t \to +\infty$. Ve skutečnosti není třeba řešit systém (16.1). Stačí si všimnout, že pokud existuje i infekčních lidí a r uzdravených lidí, je pravděpodobnost nové nákazy během malého časového intervalu dt rovna $a\,(N-i-r)\,i\,dt$ a pravděpodobnost nového uzdravení je $b\,i\,dt$. Pravděpodobnosti přechodu (jak se obvykle nazývají v teorii Markovových řetězců) ze stavu (i,r) do stavu $(i+1,r)$ nebo do stavu $(i-1,r+1)$ jsou tak následující

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

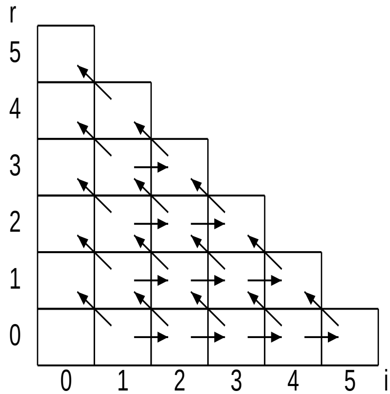
$$\mathscr{P}_{(i,r)\to(i-1,r+1)} = \frac{b}{a\,(N-i-r)+b}\,,$$

pro všechny $i \geqslant 1$ (obrázek 16.3).

Nechť $q_{i,r}$ je pravděpodobnost, že populace projde během epidemie stavem (i,r). Protože pro $t=0$ je $i=1$ a $r=0$, máme $q_{1,0}=1$. Ostatních stavů se dosáhne buď po nakažení, nebo po uzdravení:

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

První člen pravé strany chybí, je-li $i=0$ nebo $i=1$. Druhý člen chybí, když $r=0$. Z tohoto vzorce můžeme nejprve vypočítat hodnoty $(q_{i,0})_{2\leqslant i\leqslant N}$, pak $(q_{i,1})_{0\leqslant i\leqslant N-1}$, pak $(q_{i,2})_{0\leqslant i\leqslant N-2}$ atd. Pravděpodobnost, že se při epidemii nakonec nakazí n lidí, je $q_{0,n}$. V roce 1926 byly takové výpočty poměrně zdlouhavé. McKendrick se proto omezil na příklady týkající se velmi malých populací, například rodiny. Pro hodnoty $N=5$ lidí a $b/a=2$ získal tabulku



Obrázek 16.3: Diagram zobrazující možné stavy populace s $N = 5$ (i na vodorovné ose, r na svislé ose) a možné přechody v důsledku nákazy (vodorovné šipky) nebo uzdravení (ostatní šipky).

Tabulka 16.1: Pravděpodobnost, že se v pětičlenné rodině nakazí n lidí, když $b/a = 2$.

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

16.1. Největší pravděpodobnosti odpovídají situacím, kdy je v rodině nakažena pouze jedna osoba a kdy je nakažena celá rodina.

Stejný článek z roku 1926 obsahuje i novou formulaci demografických problémů, kdy je čas považován za spojitou proměnnou. Nechť pro dostatečně malé dx je $P(x,t)\,dx$ populace ve věku mezi x a $x+dx$ v čase t. Nechť $m(x)$ je úmrtnost ve věku x. Pak

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

pro h dostatečně malé. Zavedeme parciální derivace funkce $P(x,t)$:

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

Užitím vztahu

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

získal McKendrick následující parciální diferenciální rovnici:

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

Taková rovnice se přirozeně objevuje v populačních problémech strukturovaných nějakou spojitou proměnnou, jako je například věk v demografii (viz kapitola 25) nebo doba od nákazy v epidemiologii.

William Ogilvy Kermack se narodil v roce 1898 v malém městečku ve Skotsku. Vystudoval Aberdeenskou univerzitu a začal se věnovat výzkumu v oblasti organické chemie v průmyslové laboratoři v Oxfordu. V roce 1921 byl Kermack jmenován vedoucím chemické sekce laboratoře *Royal College of Physicians* v Edinburghu. Přestože po výbuchu ve své edinburské laboratoři roku 1924 zcela oslepl, pokračoval s pomocí kolegů a studentů ve své chemické práci. Kermack také začal spolupracovat s McKendrickem na matematickém modelování epidemií. Od roku 1927 společně publikovali sérii článku s názvem *Příspěvky k matematické teorii epidemií*, kde se zabývali deterministickými modely epidemií. Nechť N je velikost populace s dostatečně velkým N. Předpokládejme stejně jako v článku z roku 1926, že lidé mohou být buď vnímaví, infekční nebo uzdravení. Pokud je nemoc smrtelná, pak třetím stavem je ve skutečnosti smrt. Nechť $S(t)$, $I(t)$ a $R(t)$ jsou počty lidí v každém z těchto tří stavů. Model je (ve zjednodušené podobě) soustavou tří diferenciálních rovnic:

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

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

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

Počet nových nákaz za jednotku času je tedy stejně jako ve stochastickém modelu z roku 1926 úměrný počtu vnímavých osob i počtu infekčních osob. Na počátku epidemie, v čase $t = 0$, je určitý počet lidí infekčních: $S(0) = N - I_0$, $I(0) = I_0$ a $R(0) = 0$, za předpokladu $0 < I_0 < N$.

Přestože systém (16.2)–(16.4) nemá uzavřené řešení, lze dokázat několik jeho vlastností:

- celková populace $S(t) + I(t) + R(t)$ zůstává konstantní a rovná se N;
- $S(t)$, $I(t)$ a $R(t)$ zůstávají nezáporné (jak by také měly být, neboť se jedná o velikost populací);
- když $t \to +\infty$, $S(t)$ klesá k limitní hodnotě $S_\infty > 0$, $I(t)$ směřuje k 0 a $R(t)$ roste k limitní hodnotě $R_\infty < N$;
- vztah

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

implicitně udává hodnotu S_∞, a tedy i konečnou velikost epidemie $R_\infty = N - S_\infty$.

Nejprve si všimněme, že $\frac{d}{dt}(S+I+R) = 0$. Takže $S(t)+I(t)+R(t) = S(0)+I(0)+R(0) = N$. Rovnice (16.2) a (16.3) lze přepsat takto

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

Z toho na jedné straně vyplývá, že $S(t) = S(0)\, e^{-a\int_0^t I(\tau)\,d\tau} > 0$ a na druhé straně, že $I(t) = I(0)\, e^{a\int_0^t S(\tau)\,d\tau - bt} > 0$. Rovnice (16.2) a (16.4) pak ukazují, že funkce $S(t)$ je klesající a že funkce $R(t)$ je rostoucí (zejména $R(t) \geqslant 0$). Protože $S(t) \geqslant 0$ a $R(t) \leqslant N$, funkce $S(t)$ a $R(t)$ mají pro $t \to +\infty$ své limity. Protože $I(t) = N - S(t) - R(t)$, má $I(t)$ také limitu, když $t \to +\infty$, která však může být pouze nulová, jak je vidět z integrace rovnice (16.4). Rovnice (16.2) také ukazuje, že

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

Integrací tohoto vztahu od $t = 0$ do $t = +\infty$ zjistíme, že

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

Rovnici (16.3) lze přepsat jako $\frac{dI}{dt} = -\frac{dS}{dt} - bI$. Integrací tohoto vztahu od $t = 0$ do $t = +\infty$ dostaneme

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

Kombinací těchto dvou výsledků získáme vzorec (16.5), který ukazuje, že $S_\infty > 0$.

Pokud je počáteční množství infekčních osob I_0 malé ve srovnání s velikostí populace N, což je častý případ na začátku epidemie ve městě, lze vzorec (16.5) přepsat s pomocí formule $S_\infty = N - R_\infty$ takto

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

kde podle definice

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

Rovnice (16.6) má kladné řešení pouze tehdy, jestliže $\mathscr{R}_0 > 1$. Kermack a McKendrick tedy dospěli k následujícímu závěru: epidemie infikuje nezanedbatelnou část populace pouze tehdy, jestliže $\mathscr{R}_0 > 1$. Existuje prahová hodnota velikosti populace $N^* = b/a$, pod kterou nemůže epidemie vzniknout.

Když je velikost populace N těsně nad tímto prahem ($N = N^* + \varepsilon$), dochází k epidemii s malou amplitudou. Z (16.6) vyplývá, že $R_\infty \approx 2\,\varepsilon$. Takže $S_\infty \approx N^* - \varepsilon$: epidemie přivede vnímavou populaci o stejnou hodnotu pod práh N^*, o jakou byla původně nad ním.

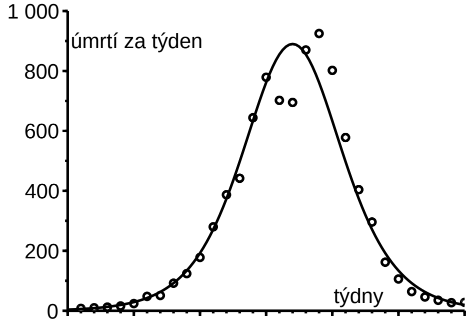
Při použití aproximace $-\log(1-x) \approx x + \frac{x^2}{2}$ lze rovnice (16.6) přepsat jako

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

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

Stejně jako v Rossově modelu malárie (kapitola 12) má podmínka $\mathscr{R}_0 > 1$ jednoduchý výklad. Protože aN je počet lidí, které jedna infekční osoba nakazí za jednotku času na počátku epidemie, a protože $1/b$ je průměrná doba infekčnosti jedince, je $\mathscr{R}_0 = aN/b$ průměrný počet sekundárních případů způsobených jednou infekční osobou na počátku epidemie.

U smrtelných nemocí je $R(t)$ kumulativní počet úmrtí od počátku epidemie a dR/dt je počet úmrtí za jednotku času. Kermack a McKendrick si všimli, že graf funkce dR/dt má v jejich matematickém modelu skutečně zvonovitý tvar, který se od epidemické křivky očekává (obrázek 16.4).



Obrázek 16.4: Křivka dR/dt jako funkce času a údaje o týdenních počtech úmrtí během morové epidemie v Bombaji v letech 1905–1906.

Pro určení dR/dt vydělili rovnici (16.2) rovnicí (16.4) a získali

$$\frac{dS}{dR} = -\frac{a}{b} S.$$

Takže

$$S(t) = S(0) \exp\left(-\frac{a}{b} R(t)\right).$$

Dosazením do rovnice (16.4) a využitím $S(t) + I(t) + R(t) = N$ získali rovnici

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

kterou stále nelze explicitně vyřešit. Nicméně pokud výraz $\frac{a}{b} R(t)$ zůstane během celé epidemie malý, aproximace $\exp(-u) \approx 1 - u + u^2/2$ dává následující výsledek

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

Toto je tzv. Riccatiho rovnice se dvěma konstantními řešeními, jedním kladným R_+ a jedním záporným R_-, danými kořeny polynomu druhého řádu v R na pravé straně (16.8). Nechť $\tilde{R}(t)$ je přesné řešení (16.8) a položme $Q(t) = \tilde{R}(t) - R_+$. Pak $Q(t)$ splňuje Bernoulliho diferenciální rovnici podobnou těm, se kterými se setkali Daniel Bernoulli a Verhulst (viz (4.5) a (6.1)). Lze tedy přímo upravit vzorec (6.2) a získat $Q(t)$. Snadný, ale zdlouhavý výpočet ukazuje, že dQ/dt má tvar

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

kde α, β a γ jsou konstanty, které složitě závisejí na parametrech modelu. Protože $dR/dt \approx d\tilde{R}/dt = dQ/dt$, mohli Kermack a McKendrick zvolit (α, β, γ) pro dobrý popis svých dat. Moderní počítače a software samozřejmě dokáží diferenciální rovnici (16.7) vyřešit snadno numericky bez těchto aproximací.

Takto získaná křivka dR/dt dobře odpovídá údajům o týdenních počtech úmrtí během morové epidemie v Bombaji mezi prosincem 1905 a červencem 1906 (obrázek 16.4).

Kermack a McKendrick se zabývali také obecnějším modelem, kde nakažlivost $a(x)$ závisí na době x od nakažení a kde rychlost uzdravení $b(x)$

závisí také na x. Rovnice udávající konečnou velikost epidemie (když je počáteční počet infekčních osob malý) je stále (16.6), ale s

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

Parametr $\mathscr{R}_0$ má stejnou interpretaci jako v předchozím případě: je to průměrný počet sekundárních případů způsobených jednou infekční osobou na počátku epidemie. Všimněte si podobnosti mezi (16.9) a Lotkovým vzorcem (10.2) pro $\mathscr{R}_0$ v demografii: věk je nahrazen dobou od nakažení, přežití pravděpodobností $e^{-\int_0^x b(y)\,dy}$, že je člověk stále infekční, plodnost mírou kontaktů $N\,a(x)$.

Kermack a McKendrick vytvořili ve 30. letech 20. století několik dalších matematických modelů epidemií. Ty jsou dodnes základním stavebním kamenem pro většinu složitějších modelů, které se dnes v epidemiologii používají. Parametr $\mathscr{R}_0$ stále hraje ústřední roli při analýze jakéhokoli epidemického modelu.

McKendrick odešel do důchodu v roce 1941 a zemřel v roce 1943. V letech 1930–1933 byl Kermack spoluautorem několika článků o matematické fyzice s Williamem McCreou a Edmundem Whittakerem z katedry matematiky na Edinburské univerzitě. Ve 30. a 40. letech 20. století se Kermackův tým chemiků pokoušel syntetizovat nové molekuly s antimalarickou aktivitou, avšak s omezeným úspěchem. V roce 1938 byl Kermack spolu s Philipem Eggletonem autorem populární knihy o základech biochemie *Z čeho jsme stvořeni*. V roce 1944 byl zvolen členem Královské společnosti a v roce 1949 převzal katedru biochemie na Aberdeenské univerzitě. Později působil jako děkan Přírodovědecké fakulty. Do důchodu odešel v roce 1968 a zemřel v roce 1970.

Další čtení

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

Kapitola 17

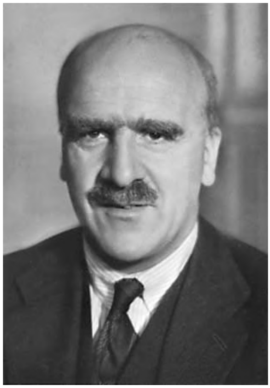
Haldane a mutace (1927)

Ve druhé části svého článku z roku 1922 se Fisher zabýval problémem mutantního genu, který může být přenesen na náhodný počet potomků s daným rozdělením pravděpodobnosti. Problém byl formálně stejný jako problém vymírání rodových jmen, ale v genetickém kontextu. Fisher ukázal, že pokud je pravděpodobnostní rozdělení Poissonovým rozdělením a pokud mutantní gen nemá žádnou selekční výhodu, pak může mizet z populace velmi pomalu. V roce 1927 britský biolog Haldane rozšířil studium tohoto modelu a ukázal, že pravděpodobnost udržení se výhodného mutantního genu v populaci je dvojnásobkem jeho selekční výhody. Důkladněji také zpracoval problém vymírání.

John Burdon Sanderson Haldane se narodil v roce 1892 v Oxfordu, kde jeho otec působil jako univerzitní profesor fyziologie. Haldane studoval na *Eton College* a po roce 1911 na *New College* Oxfordské univerzity. Poté, co se v prvním ročníku zaměřil na matematiku, se začal věnovat humanitním vědám. Jeho studia přerušila první světová válka, během níž sloužil ve Francii a Iráku. Po zranění byl poslán jako vojenský instruktor do Indie. V roce 1915 publikoval první článek pojednávající o genetických pokusech na myších, které zahájil ještě před válkou. V roce 1919 se stal členem *New College*, kde vyučoval fyziologii a studoval dýchání stejně jako jeho otec. V roce 1923 nastoupil do biochemické laboratoře F. G. Hopkinse[1] na Cambridgeské univerzitě, kde se zaměřil na kinetiku enzymů. Vydal také vědeckofantastický román *Daidalos aneb Věda a budoucnost* (1923) a esej nazvanou *Callinicus, obrana chemické války* (1925). V letech 1924–1934 napsal sérii deseti článků s názvem *Matematická teorie přírodního a umělého výběru*.

V pátém článku této série, publikovaném v roce 1927, Haldane přehodnotil další genetický model, který Fisher studoval v roce 1922, model zaměřený na mutace. Fisher studoval pravděpodobnost, s jakou mutantní gen vnikne do populace nebo zanikne. Tento problém je formálně stejný jako problém Bienayméa, Galtona a Watsona týkající se vymírání rodových jmen. Fisher se však na tyto práce neodvolával, ačkoli možná četl článek Galtona a Watsona

[1] Frederick Gowland Hopkins, který v roce 1929 obdržel Nobelovu cenu za fyziologii a medicínu za práci o vitamínech.



Obrázek 17.1:
Haldane (1892–1964)

reprodukovaný v příloze Galtonovy knihy *Přirozená dědičnost* z roku 1889. Stejně jako v kapitole 9 nazvěme p_k pravděpodobnost, že se gen přenese na k potomků v první generaci ($k \geqslant 0$). Fisher uvažoval také vytvořující funkci

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

v níž nestanovil žádnou horní hranici pro k: součet může obsahovat nekonečný počet členů. Uvědomil si, že počínaje jedním jedincem s mutantním genem v generaci 0 je pravděpodobnost výskytu tohoto genu u k jedinců rovna koeficientu x^k v $f_1(x) = f(x)$ pro generaci 1, v $f_2(x) = f(f(x))$ pro generaci 2, v $f_3(x) = f(f(f(x)))$ pro generaci 3 atd. Tímto způsobem je zřejmé, že platí následující rovnice

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

Tato rovnice je mnohem praktičtější než rovnice $f_n(x) = f_{n-1}(f(x))$ odvozená Watsonem. Zejména z (17.1) vyplývá, že pravděpodobnost vymření během n generací $x_n = f_n(0)$ splňuje rekurentní formuli $x_n = f(x_{n-1})$, jak si všiml již Bienaymé.

Jako příklad uvedl Fisher případ rostliny s mutantním genem, která může produkovat N semen, přičemž každé semeno má pravděpodobnost q, že přežije a vytvoří novou rostlinu. Pravděpodobnost p_k získání k potomků s mutantním genem je binomická:

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

pro všechny $0 \leqslant k \leqslant N$ a $p_k = 0$ pro $k > N$. Vytvořující funkce je pak

$$f(x) = (1 - q + qx)^N.$$

Nechť $\mathscr{R}_0 = N\,q$ je střední počet semen, která přežijí a vytvoří novou rostlinu. Když je N velké a q malé, pak platí, že

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

Rozdělení pravděpodobnosti (p_k) konverguje k $e^{-\mathscr{R}_0}\,(\mathscr{R}_0)^k/k!$, které se nazývá Poissonovo rozdělení. Fisher pak vypočítal pravděpodobnost vymření během n generací pomocí $x_0 = 0$,

$$x_n \approx e^{\mathscr{R}_0(x_{n-1}-1)}$$

a číselných hodnot $N = 80$ a $q = 1/80$. V tomto případě je $\mathscr{R}_0 = Nq = 1$. Zdlouhavý výpočet ukazuje, že $x_{100} \approx 0{,}98$: mutantní gen bez selekční výhody ($\mathscr{R}_0 = 1$) mizí velmi pomalu. Po 100 generacích je stále 2 % šance, že gen bude v populaci přítomen. Od roku 1922 Fisher již ve studiu tohoto modelu nepokračoval.

V návaznosti na Fisherovu práci si Haldane ve svém článku z roku 1927 poprvé všiml, že pro libovolné rozdělení pravděpodobnosti (p_k) takové, že $p_0 > 0$, má rovnice $x = f(x)$ přesně dva kořeny v intervalu $(0,1]$, když je průměrný počet potomků nesoucích mutantní gen $\mathscr{R}_0$ ostře větší než 1, tj. když má mutantní gen selekční výhodu. Navíc pravděpodobnost vymření x_∞, která je limitou x_n pro $n \to +\infty$, je menší ze dvou kořenů rovnice $x = f(x)$: gen má nenulovou pravděpodobnost prosazení se v populaci. Haldane tento závěr, na rozdíl od Bienaymého a Cournota, dokázal.

Na intervalu $[0,1]$ totiž platí $f'(x) \geqslant 0$ a $f''(x) \geqslant 0$. Jinými slovy, funkce $f(x)$ je neklesající a konvexní. Z předpokladů $f(0) = p_0 > 0$ a

$$f'(1) = \mathscr{R}_0 = p_1 + 2p_2 + 3p_3 + \cdots > 1$$

vyplývá, že rovnice $f(x) = x$ má v intervalu $(0,1]$ přesně dvě řešení: $x = 1$ a x^* takové, že $0 < x^* < 1$. Haldane pak odkázal na článek Gabriela Koenigse z roku 1883, který ukázal, že z předpokladů $x_n = f(x_{n-1})$ a $x_n \to x_\infty$ plyne $x_\infty = f(x_\infty)$ a $|f'(x_\infty)| \leqslant 1$. Když $f'(1) > 1$, jako jediná možnost zůstává $x_\infty = x^*$.

V případě Poissonova rozdělení s $f(x) = e^{\mathscr{R}_0(x-1)}$ a $\mathscr{R}_0$ jen o málo větším než 1 je pravděpodobnost vymizení x_∞ velmi blízká 1. Rovnice $f(x_\infty) = x_\infty$ je pak ekvivalentní následujícím vzorcům

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

Z toho vyplývá, že

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

Haldane došel k závěru, že pravděpodobnost, že mutantní gen nevyhyne, je dvojnásobkem jeho selekční výhody $\mathscr{R}_0 - 1$. Aniž by Haldanea citoval, Fisher ve své knize z roku 1930 uvedl jako příklad případ, kdy $\mathscr{R}_0 = 1{,}01$, což dává 2 % šanci, že mutantní gen nevyhyne.

Haldane se stal členem Královské společnosti v roce 1932. Z Cambridge odešel a stal se profesorem genetiky a později biometrie na *University College* v Londýně. Zabýval se tehdy zejména lidskou genetikou: odhadem míry mutací, genetickými mapami chromozomů atd. Vedle svých vědeckých knih (*Biologie živočichů* v roce 1927 s Julianem Huxleym, *Enzymy* v roce 1930 a *Příčiny evoluce* v roce 1932, *Biochemie genetiky* v roce 1954) publikoval velké množství článků o vědě (např. o vzniku života) a několik esejí (*Nerovnost člověka* v roce 1932, *Filosofie biologa* v roce 1935, *Marxistická filosofie a přírodní vědy* v roce 1938, *Dědičnost a politika* v roce 1938 a *Pokroky vědy* v roce 1947). Po několika návštěvách Španělska během občanské války se snažil přesvědčit svou zemi, aby vybudovala kryty proti leteckému bombardování. Za druhé světové války se zabýval problémy dýchání v ponorkách. Od roku 1942 byl členem komunistické strany, v roce 1950 z ní vystoupil kvůli oficiálnímu odmítnutí mendelovské genetiky a přijetí Lysenkovy ideologie v SSSR. V roce 1957 se usadil v Indii, kde pokračoval ve výzkumu, nejprve v Indickém ústavu statistiky v Kalkatě a později v Bhubanesváru. Získal indické občanství, zemřel v roce 1964.

Další čtení

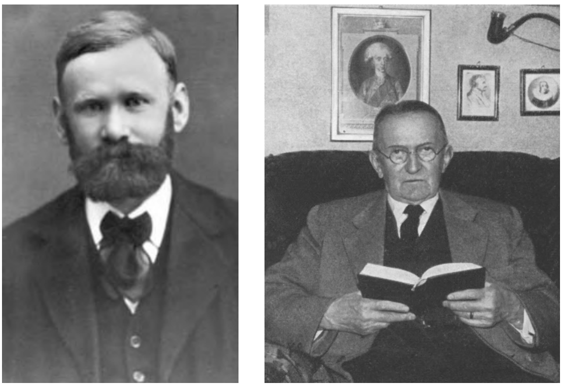
1. Clark, R.: *J. B. S., The Life and Work of J. B. S. Haldane*. London (1968)
2. Haldane, J.B.S.: A mathematical theory of natural and artificial selection, Part V, Selection and mutation. *Proc. Camb. Philos. Soc.* 23, 838–844 (1927)
3. Haldane, J.B.S.: *The Causes of Evolution*. Longmans (1932) archive.org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

Kapitola 18

Erlang a Steffensen o problému vymírání (1929–1933)

V roce 1929 se dánský telefonní inženýr Erlang vrátil k problému zániku příjmení. Jeho krajan statistik Steffensen problém úplně vyřešil. Ukázal zejména, že očekávaný počet potomků v každé generaci roste exponenciálně. Tím propojil stochastický a deterministický model populace.

Agner Krarup Erlang se narodila v roce 1878 v Lønborgu v Dánsku. Jeho otec byl učitelem. V letech 1896–1901 studoval matematiku, fyziku a chemii na Kodaňské univerzitě. Poté několik let učil na středních školách a zároveň se zajímal o matematiku, zejména o teorii pravděpodobnosti. Seznámil se s Jensenem, hlavním inženýrem Kodaňské telefonní společnosti a amatérským matematikem, který ho v roce 1908 přesvědčil, aby se připojil k nové výzkumné laboratoři společnosti. Erlang začal publikovat články o aplikacích teorie pravděpodobnosti na řízení telefonních hovorů. V roce 1917 objevil vzorec pro stanovení čekací doby, který rychle začaly používat telefonní společnosti po celém světě. Jeho články, publikované nejprve v dánštině, byly poté přeloženy do několika dalších jazyků.



Obrázek 18.1: Erlang (1878–1929) a Steffensen (1873–1961)

V roce 1929 se Erlang začal zajímat o stejný problém vymírání, který před ním zkoumali Bienaymé, Galton a Watson u příjmení a který Fisher a Haldane zkoumali u mutantních genů. Stejně jako jeho předchůdci si nebyl vědom všech prací, které byly publikovány. Opět označil p_k pravděpodobnost, že jeden jedinec bude mít k potomků, a všiml si, že pravděpodobnost x_n vymření během n generací splňuje rovnost

$$x_n = p_0 + p_1\, x_{n-1} + p_2\, (x_{n-1})^2 + \cdots = f(x_{n-1})$$

s $x_0 = 0$. Všiml si také, že celková pravděpodobnost vymření x_∞, která je limitou x_n pro $n \to +\infty$, je řešením rovnice

$$x_\infty = f(x_\infty).$$

Uvědomil si, že $x = 1$ je vždy řešením a že další řešení existuje mezi 0 a 1, když je průměrný počet potomků

$$\mathscr{R}_0 = f'(1)$$

větší než 1. Zdá se však, že nemohl přijít na to, které z těch dvou řešení je to správné. Stejně jako Galton předložil tento problém v roce 1929 dánskému matematickému časopisu *Matematisk Tidsskrift*:

> „Otázka 15. Je-li p_k pravděpodobnost, že jedinec bude mít k dětí, přičemž
>
> $$p_0 + p_1 + p_2 + \cdots = 1,$$
>
> najděte pravděpodobnost, že jeho rodina vymře.“

Erlang bohužel ještě téhož roku 1929 zemřel jako bezdětný ve věku 51 let[1].

Erlangovy otázky se ujal profesor pojistné matematiky na Kodaňské univerzitě Johan Frederik Steffensen. V roce 1930 publikoval ve stejném dánském časopise své řešení: pravděpodobnost zániku x_∞ je vždy menším kořenem rovnice $x = f(x)$ v uzavřeném intervalu $[0,1]$, jak si všimli již Bienaymé a Haldane. Steffensenův důkaz tohoto tvrzení lze nalézt v moderních učebnicích.

Viděli jsme totiž, že pravděpodobnost vymřeení x_∞ je řešením rovnice $x = f(x)$ v uzavřeném intervalu $[0,1]$. Nechť x^* je nejmenší takové řešení. Z definice vyplývá, že $x^* \leqslant x_\infty$. Steffensen si nejprve všiml, že

$$x^* = f(x^*) \geqslant p_0 = x_1.$$

[1] Na jeho památku se Mezinárodní telefonní poradní výbor v roce 1946 rozhodl nazvat „erlang“ jednotkou pro měření intenzity telefonního provozu. „Erlang“ je také název programovacího jazyka společnosti Ericsson.

Z indukčního předpokladu $x^* \geqslant x_n$ plyne, že

$$x^* = f(x^*) \geqslant f(x_n) = x_{n+1},$$

protože funkce $f(x)$ je rostoucí. Takže $x^* \geqslant x_n$ pro všechna n. Přechodem k limitě dostaneme $x^* \geqslant x_\infty$. To znamená, že $x_\infty = x^*$.

Steffensen také podal formálnější vysvětlení, proč je $x = 1$ jediným kořenem rovnice $x = f(x)$, pokud střední počet potomků $\mathscr{R}_0 = f'(1)$ nepřevýší 1 (obrázek 18.2a) a proč existuje pouze jeden další kořen různý od $x = 1$ v případě, že $\mathscr{R}_0 > 1$ (obrázek 18.2b). Všimněte si, že $\mathscr{R}_0 = f'(1)$ je sklon funkce $f(x)$ při $x = 1$.

Pro libovolný kořen rovnice $x = f(x)$ platí

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

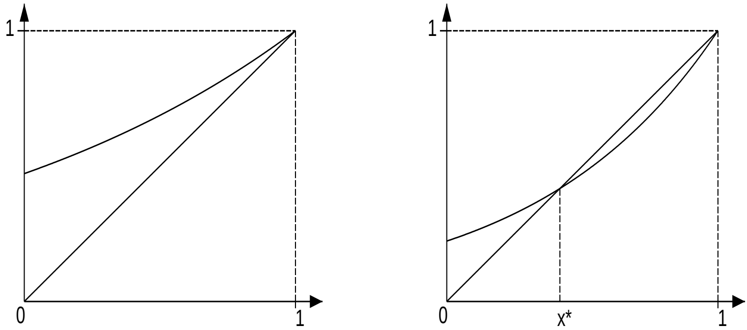
Za předpokladu, že $x \neq 1$, po vydělení výrazem $1 - x$ dostaneme rovnici

$$1 = p_1 + p_2(1 + x) + p_3(1 + x + x^2) + \cdots . \qquad (18.1)$$

Když se x zvětší z 0 na 1, pravá strana rovnice (18.1) se zvětší z $1 - p_0$ na $\mathscr{R}_0 = f'(1)$. Jestliže $\mathscr{R}_0 < 1$, pak rovnice (18.1) nemá řešení. Pokud $\mathscr{R}_0 \geqslant 1$ a pokud vyloučíme triviální případ $p_1 = 1$, pak pravá strana rovnice (18.1) je ryze rostoucí funkcí x. Jinak by neexistovalo $k \geqslant 2$ takové, aby $p_k \neq 0$ a $\mathscr{R}_0$ by bylo rovno $p_1 < 1$. Z provedených úvah lze vyvodit, že (18.1) má pro $\mathscr{R}_0 \geqslant 1$ jediné řešení v intervalu $[0, 1]$.

Steffensen, který byl také prezidentem Dánské pojistné společnosti a Dánské matematické společnosti, byl v roce 1930 pozván na Londýnskou univerzitu. Jeho britský kolega W. P. Elderton mu řekl o Galtonově a Watsonově práci. V roce 1933 Steffensen publikoval nový článek v análech Ústavu Henriho Poincarého, kde v roce 1931 přednášel na konferenci. V dánštině shrnul výsledky svého článku a porovnal je s Watsonovými výsledky. Ukázal také, že očekávaný počet potomků v n-té generaci je roven $(\mathscr{R}_0)^n$.

Nechť $p_{k,n}$ je pravděpodobnost, že v n-téeneraci bude k potomků jednoho jedince z počáteční, nulté, generace. Steffensen si ve svém článku



Obrázek 18.2: Graf funkcí $y = x$ a $y = f(x)$ v příkladu z kapitoly 17, $f(x) = e^{\mathscr{R}_0(x-1)}$, přičemž $\mathscr{R}_0 = 0,75 < 1$ (vlevo) nebo $\mathscr{R}_0 = 1,5 > 1$ (vpravo).

z roku 1930 stejně jako jeho předchůdci všiml, že vytvořující funkce

$$f_n(x) = \sum_{k=0}^{+\infty} p_{k,n} x^k$$

pro n-tou generaci splňuje rovnosti $f_1(x) = f(x)$ a

$$f_n(x) = f(f_{n-1}(x)). \tag{18.2}$$

Nechť M_n je očekávaný počet potomků v n-té generaci. Pak

$$M_n = \sum_{k=1}^{+\infty} k\, p_{k,n} = f_n'(1)\,.$$

Derivováním rovnosti (18.2) dostaneme

$$f_n'(x) = f'(f_{n-1}(x)) \cdot f_{n-1}'(x).$$

Takže

$$M_n = f_n'(1) = f'(f_{n-1}(1)) \cdot f_{n-1}'(1) = f'(1) \cdot M_{n-1} = \mathscr{R}_0 \cdot M_{n-1}.$$

Protože $M_1 = f_1'(1) = f'(1) = \mathscr{R}_0$, vyplývá z toho, že $M_n = (\mathscr{R}_0)^n$ pro všechna n.

Očekávaný počet potomků tedy geometricky roste nebo klesá podle toho, zda je $\mathscr{R}_0$ větší nebo menší než 1. Očekávaný počet potomků se chová podobně jako v deterministických modelech růstu populace, které uvažovali Euler, Malthus atd. Avšak i když $\mathscr{R}_0 > 1$, existuje nenulová pravděpodobnost x_∞, že rodina vymře. V deterministických modelech tato možnost nenastává.

Stochastický proces, který studoval Steffensen a jeho předchůdci, je stále základním prvkem mnoha realističtějších modelů populační dynamiky. Naposledy se o tomto problému zmíníme v kapitole 20. Pokud jde o Steffensena, zůstal profesorem na Kodaňské univerzitě až do roku 1943 a zemřel v roce 1961.

Další čtení

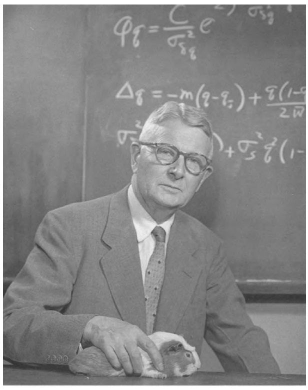
1. Brockmeyer, E., Halstrøm, H.L., Jensen, A.: The life and works of A.K. Erlang. *Trans. Dan. Acad. Techn. Sci.* 2 (1948)
2. Erlang, A.K.: Opgave Nr. 15. *Mat. Tidsskr. B*, 36 (1929) → Guttorp (1995)
3. Guttorp, P.: Three papers on the history of branching processes. *Int. Stat. Rev.* 63, 233–245 (1995) www.stat.washington.edu/research/reports/1992/tr242.pdf
4. Heyde, C.C.: Agner Krarup Erlang. In: Heyde, C.C., Seneta, E. (eds.) *Statisticians of the Centuries*, 328–330. Springer (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

Kapitola 19

Wright a náhodný genetický drift (1931)

V roce 1931 americký biolog Sewall Wright vypracoval studii stochastického modelu v populační genetice, který vychází ze stejných předpokladů jako Hardyho-Weinbergův zákon s tím rozdílem, že nepředpokládá nekonečně velkou populaci. Frekvence genotypů již nejsou konstantní. Jedna ze dvou alel skutečně vymizí, ale možná až po velmi dlouhé době. Interpretace tohoto modelu zůstala předmětem sporu mezi Wrightem a Fisherem, přičemž druhý jmenovaný předpokládal, že přírodní výběr hraje v evoluci důležitější roli než stochasticita.

Sewall Wright se narodil v roce 1889 v Massachusetts. Bakalářské studium absolvoval na malé vysoké škole v Illinois, kde jeho otec vyučoval ekonomii. Po magisterském studiu biologie na Illinoiské univerzitě v Urbane a letní škole v Cold Spring Harbor Laboratory získal Wright doktorát na Harvardově univerzitě, kde se zabýval dědičností barvy srsti u morčat. V letech 1915 až 1925 pokračoval v práci na pokusech s příbuzenskou plemenitbou (inbreedingem) morčat na ústavu pro chov zvířat amerického ministerstva zemědělství ve Washingtonu. K analýze těchto pokusů vyvinul „metodu dráhových koeficientů“. Poté nastoupil na katedru zoologie Chicagské univerzity.



Obrázek 19.1:
Wright (1889–1988)

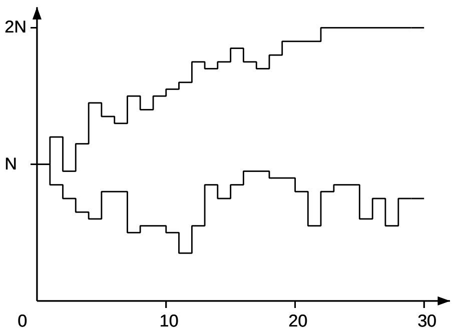
Pod vlivem Fisherova článku o populační genetice z roku 1922 (viz kapitola 14) napsal Wright v roce 1925 dlouhý článek s názvem *Evoluce v mendelovských populacích*, který byl nakonec publikován v roce 1931. Studoval zejména matematický model, který se implicitně objevil i ve Fisherově knize *Genetická teorie přírodního výběru* z roku 1930. Stejně jako v případě Hardyho-Weinbergova zákona se v tomto modelu uvažuje případ, kdy pro jeden lokus existují právě dvě možné alely A a a, ale nepředpokládá se, že populace je nekonečně velká. Jde o to zjistit, zda odstranění tohoto předpokladu má nějaký vliv na genetické složení populace. Nechť je tedy N celkový počet jedinců, o němž se předpokládá, že je ve všech generacích stejný. Každý jedinec má dvě alely. V každé generaci je tedy v populaci celkem $2N$ alel. Model také předpokládá, že k páření dochází náhodně. Jestliže v generaci n existuje i alel A a $2N-i$ alel a, pak alela náhodně vybraná mezi jedinci v generaci $n+1$ bude A s pravděpodobností $\frac{i}{2N}$ a a s pravděpodobností $1-\frac{i}{2N}$. Počet alel A v generaci $n+1$ se tedy bude rovnat j s pravděpodobností[1]

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

kde $\binom{2N}{j}$ je binomický koeficient rovný

$$\frac{(2N)!}{j!(2N-j)!}.$$

Nechť X_n je náhodná veličina udávající počet alel A v generaci n (obr. 19.2).



Obrázek 19.2: Dvě simulace ukazující změny počtu X_n alel A během 30 generací, pokud $N = 20$ a $X_0 = 10$.

[1] Autorem této formulace v jazyce Markovských řetězců je Malécot (1944).

Lze ukázat, že střední hodnota X_{n+1} za podmínky, že $X_n = i$, je rovna i. To připomíná Hardyho-Weinbergův zákon, kde frekvence alely A zůstává v rámci generací konstantní.

> Uvažujme pravděpodobnostní funkci
>
> $$f(x) = \sum_{j=0}^{2N} p_{i,j} x^j = \left(1 - \frac{i}{2N} + \frac{ix}{2N}\right)^{2N},$$
>
> Střední hodnota X_{n+1} za podmínky, že $X_n = i$, je pak následující
>
> $$\sum_{j=0}^{2N} j\, p_{i,j} = f'(1) = i. \tag{19.2}$$

V tomto modelu s počáteční podmínkou $X_0 = i$ s $0 < i < 2N$ je však možné, že po určitém počtu generací dojde k náhodné události $X_n = 0$. V takovém případě by všechny alely byly typu a a X_n by ve všech budoucích generacích zůstaly rovny 0. Stejně tak by pro alelu A platilo, že pokud by $X_n = 2N$ po určitém počtu generací, zůstalo by také zafixováno. Obecně řečeno, pokud předpokládáme nekonečně velkou populaci jako v Hardyho-Weinbergově modelu, nemůže ani jedna alela zaniknout, protože frekvence alel zůstávají konstantní. Když se vezme v úvahu konečná velikost populace, jako je tomu ve Fisherově-Wrightově modelu, frekvence obou alel kolísají a jedna z alel může vymizet, a nakonec také vymizí.

Vycházíme-li z podmínky $X_0 = i$, můžeme snadno vypočítat pravděpodobnost Q_i, že se populace ustálí ve stavu $X = 0$. Pro Q_i platí, že musí splňovat „okrajové podmínky“

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

Navíc platí, že

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

protože $p_{i,j} Q_j$ je pravděpodobnost ustálení stavu $X = 0$ počínaje $X_0 = i$ za podmínky, že $X_1 = j$. Protože

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

vidíme pomocí (19.2), že

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

je řešením systému (19.3)–(19.4). Proto pravděpodobnost, že populace o velikosti N s počátečním počtem i alel typu A, bude nakonec obsahovat pouze alely typu a, je rovna $1 - \frac{i}{2N}$. Podobně pravděpodobnost, že bude obsahovat pouze alely typu A, je rovna $\frac{i}{2N}$.

Wrightovi se podařilo ukázat, že počet generací, které uplynou, než dojde k fixaci v jednom ze dvou krajních stavů, je řádově $2N$ (obr. 19.3). Pro populace o několika milionech jedinců by tato doba byla tak dlouhá, že by frekvence alel mohly být považovány za téměř konstantní, jako v Hardyho-Weinbergově zákoně.

Předpokládejme, že v nulté generaci populace existuje i_0 alel typu A. Nechť $u_i^{(n)}$ je pravděpodobnost, že v populaci v generaci n je i alel typu A. Pak

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

pro všechna $j = 0,\dots,2N$. Jak bylo již dříve ukázáno, když $n \to +\infty$,

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

pro všechna $0 < i < 2N$. Wright si všiml, že pokud $u_i^{(n)} = v$ pro všechna $i = 1,\dots,2N-1$, pak

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

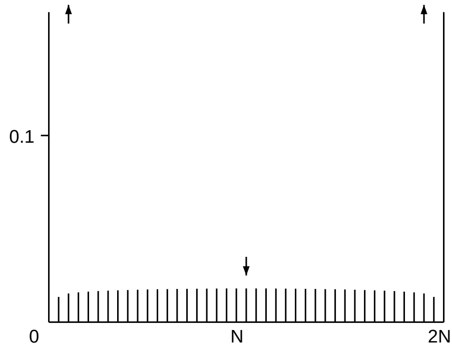
pro všechna $1 < j < 2N$, protože $p_{0,j} = p_{2N,j} = 0$. Když je N dostatečně velké,

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

hodnota integrálu se získá opakovaným použitím integrační metody per-partes. Kombinací (19.5) a (19.6) dostaneme nakonec pro $0 < j < 2N$ hodnotu

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

Tedy, že se pravděpodobnost $u_j^{(n)}$ pro všechna $0 < j < 2N$ snižuje rychlostí přibližně $1/2N$ za jednu generaci. Pro velká N je tato rychlost velmi pomalá. Pokud je například N řádově v milionech, nedochází téměř k žádnému poklesu.



Obrázek 19.3: Pravděpodobnost, že po 30 generacích bude v populaci i alel A ($i = 0,\ldots,2N$ na vodorovné ose), pokud $N = 20$ a $X_0 = 10$.

Fisher se již v roce 1922 pokusil odhadnout tuto míru fixace ($1/2N$), ale vynechal faktor 2. V každém případě se oba vědci neshodli na tom, jaký je typický počet N rozmnožujících se jedinců v populaci. Pro teorii evoluce Wrightova práce naznačovala, že náhodný genetický drift v malé populaci by mohl být mechanismem vzniku druhů. Biologové, kteří pracovali na klasifikaci druhů, si skutečně všimli, že rozdíly mezi druhy nebo poddruhy často nemají zjevné vysvětlení z hlediska přírodního výběru. Proti této myšlence se ve 40. a 50. letech 20. století ostře postavili Fisher a jeho kolega E. B. Ford, kteří se domnívali, že náhodný genetický drift je ve srovnání s přírodním výběrem zanedbatelný. Odvolávali se zejména na svou studii kolísání genových frekvencí v malé izolované populaci můr (*Panaxia dominula*) poblíž Oxfordu, kde bylo možné pohledem rozlišit tři genotypy určitého genu (běžný homozygot, heterozygot a vzácný homozygot). Další známý spor o vliv přírodního výběru a náhodného driftu se týkal plžů rodu *Cepaea*. Realističtější modely evoluce dnes kombinují náhodný drift, selekci, mutaci, migraci, nenáhodné páření atd. Úlohu náhodného driftu později znovu zdůraznil japonský vědec Motoo Kimura svou „Neutrální teorií molekulární evoluce“. Dalším výstupem byl rozvoj teorie koalescence (představené Johnem Kingmanem v roce 1982), která sleduje původ genů zpět v čase až k bodu, kde mají jediného společného předka.

V roce 1934 se Wright stal členem Národní akademie věd. Mnoho let pracoval s Theodosiem Dobzhanskym na genetice přírodních populací much (*Drosophila pseudoobscura*) v oblasti Údolí smrti. V roce 1955 odešel z Chicagské univerzity do důchodu, ale ještě pět let působil jako profesor na Wisconsinské univerzitě v Madisonu. V letech 1968–1978 vydal čtyřsvazkové pojednání shrnující jeho práci na téma *Evoluce a genetika populací*. V roce 1984 obdržel Balzanovu cenu a zemřel v roce 1988 ve věku 98 let.

Další čtení

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. University of Chicago Press (1969)

Kapitola 20

Šíření genů (1937)

V roce 1937 Ronald Fisher a tři ruští matematici, Kolmogorov, Petrovskij a Piskunov, nezávisle na sobě studovali parciální diferenciální rovnici pro prostorové šíření výhodného genu. Ukázali, že frekvence genu se chová jako vlna pohybující se přesně definovanou rychlostí v závislosti na evoluční zdatnosti genu a na koeficientu difúze. Jejich práce odstartovala studium teorie reakčně-difúzních rovnic.

V roce 1937 byly publikovány dva články, které představily nový přístup ke studiu prostorové heterogenity v populační dynamice. Autorem prvního článku s názvem *Vlna prosazení prospěšného genu*, který vyšel v časopise *Annals of Eugenics* byl Ronald Fisher. Studoval v něm prostorové šíření evolučně zdatnějšího genu v populaci. Pro zjednodušení uvažoval prostor redukovaný na jeden rozměr, přičemž označil jako $u(x,t)$ podíl populace nacházející se v bodě x v čase t s genem s vyšší zdatností. Takže $0 \leqslant u(x,t) \leqslant 1$. Aby zahrnul přirozený výběr, použil rovnici (14.4) se spojitou časovou proměnnou

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

kde a je kladný parametr. Pro danou hodnotu x jde o Verhulstovu logistickou rovnici (viz kapitola 6) s řešením $u(x,t)$, které konverguje k 1 pro $t \to +\infty$. Fisher dále předpokládal, že potomci jedince s evolučně zdatnějším genem, který se nachází v bodě x, nezůstávají ve stejném bodě, ale náhodně se rozptýlí v okolí bodu x. Využil fyzikální analogie a do rovnice pro $u(x,t)$ přidal difúzní člen, což vedlo na parciální diferenciální rovnici tvaru

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

Pokud je selekční koeficient a nulový, redukuje se na difúzní rovnici, kterou zavedl Fourier ve své teorii vedení tepla a kterou později použil Fick pro difúzi fyzikálních částic. V roce 1904 začal Ronald Ross uvažovat o náhodném rozptylu v populační dynamice. Zajímalo ho tehdy, jak klesá hustota komárů s rostoucí vzdáleností od místa rozmnožování. Problému si všimli Karl Pearson a lord Rayleigh. Do roku 1937 se množství vědecké literatury o difúzních rovnicích značně rozrostlo, zejména v návaznosti na Einsteinovu práci o Brownově pohybu.

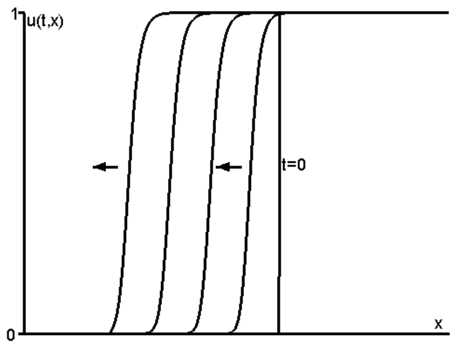
Fisher ukázal, že existují řešení rovnice (20.1) ve tvaru $u(x,t) = U(x + vt) = U(z)$ splňující tři podmínky

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

za předpokladu, že $v \geqslant v^*$, kde

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

Tato řešení spojují ustálený stav $u = 1$ s výhodným genem s ustáleným stavem $u = 0$ bez takového genu. Představují postupující vlnu šířící se rychlostí v ve směru klesajících hodnot x. Skutečně, $u(x - vT, t + T) = u(x,t)$: část vlny, která byla v čase t v poloze x, se v čase $t + T$ přesune do polohy $x - vT$.



Obrázek 20.1: Šíření evolučně zdatnějšího genu zleva doprava rychlostí v^*. Frekvence genu $u(t,x)$ při $t = 0$ je skoková funkce.

Fisher si všiml, že pokud je řešení ve tvaru postupující vlny

$$u(x,t) = U(z), \quad z = x + vt,$$

pak

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

Je-li u řešením rovnice (20.1), pak

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

Fisher očekával, že pokud se u blíží 0, tj. pro $z \to -\infty$, platí $U(z) \to 0$ a $U'(z) \to 0$. Je-li k limita $U'(z)/U(z)$ pro $z \to -\infty$, víme z l'Hospitalova pravidla, že $U''(z)/U'(z)$ také konverguje ke k. Proto se

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

blíží ke k^2. Dělíme-li rovnici (20.2) $U(z)$ a necháme-li z konvergovat k $-\infty$, dostaneme pro k kvadratickou rovnici

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

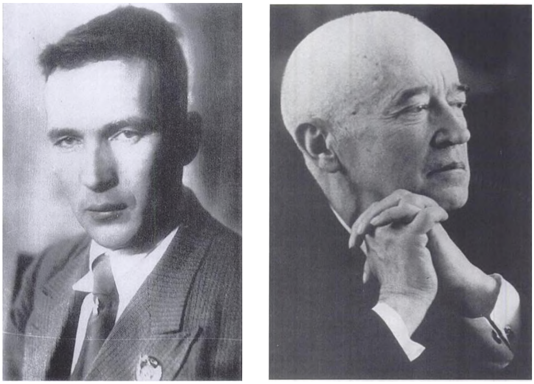
Protože k musí být reálné číslo, musí být diskriminant této rovnice nezáporný, tj. $v^2 - 4aD \geqslant 0$ neboli $v \geqslant 2\sqrt{aD} = v^*$. Proto je $v \geqslant v^*$ nutnou podmínkou existence vlny šířící se rychlostí v. Je to také postačující podmínka, jak je vysvětleno níže.

Fisher si všiml, že pro velkou třídu počátečních podmínek se vlna šíří přesně rychlostí v^*, např. pro po částech konstantní funkci $u(x,0) = 0$ pro $x < 0$, $u(x,0) = 1$ pro $x \geqslant 0$. Obrázek 20.1 ukazuje, jak se z této nespojité počáteční podmínky postupně stává hladká vlna šířící se ve směru klesajícího x rychlostí v^*.

Ve stejném roce 1937, nezávisle na Fisherově práci, studovali Andrej Nikolajevič Kolmogorov, Ivan Georgijevič Petrovskij a Nikolaj Semenovič Piskunov stejný problém šíření dominantního genu.

Kolmogorov se narodil v roce 1903 v ruském Tambově. Během studia matematiky na Moskevské státní univerzitě vytvořil několik důležitých výsledků o trigonometrických řadách. V roce 1929 se stal vědeckým pracovníkem Ústavu mechaniky a matematiky MGU a v roce 1931 univerzitním profesorem. Zabýval se stochastickými procesy a jejich souvislostí s diferenciálními a parciálními diferenciálními rovnicemi. V roce 1933 vydal pojednání, v němž položil moderní základy teorie pravděpodobnosti. Zabýval se topologií, teorií aproximace, Markovovými řetězci, Brownovým pohybem a také aplikacemi na biologické problémy. V roce 1935 publikoval článek o genetice, v němž diskutoval výsledky Hardyho, Fishera a Wrighta. V roce 1936 publikoval článek o zobecnění systému Lotka-Volterra.

Petrovskij se narodil v roce 1901 v Sevsku. Vystudoval matematiku na Moskevské státní univerzitě, kde se v roce 1933 stal profesorem. Zabýval se především teorií parciálních diferenciálních rovnic a topologií reálných algebraických křivek, ale napsal také několik článků o obyčejných diferenciálních rovnicích a o teorii pravděpodobnosti. Piskunov, který se narodil v roce 1908, byl dalším absolventem studia matematiky na Moskevské státní univerzitě.



Obrázek 20.2: Kolmogorov (1903–1987) a Petrovskij (1901–1973)

Ve 30. letech 20. století Kolmogorov udržoval kontakty s A. S. Serebrovským, průkopníkem populační genetiky v Moskvě. Obhajoba mendelovské genetiky v SSSR se tehdy stávala stále nebezpečnější kvůli nástupu agronoma Lysenka, kterému se podařilo přesvědčit Stalina, že mendelovská genetika je pouhá „buržoazní pavěda“. Sedmý mezinárodní genetický kongres, který se měl původně konat v roce 1937 v Moskvě, byl zrušen. Mnoho sovětských genetiků bylo popraveno nebo posláno do pracovních táborů.

Kolmogorov, Petrovskij a Piskunov ve svém článku z roku 1937 nazvaném *Studium difúzní rovnice s nárůstem množství látky a její aplikace na biologický problém*, který byl publikován v *Bulletinu Moskevské státní univerzity*, nicméně použili matematický model založený na Mendelově genetice. Jejich modelem byla parciální diferenciální rovnice ve tvaru

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

kde $u(x,t)$ je opět frekvence evolučně zdatnějšího genu v bodě x a čase t. Předpokládá se, že funkce $f(u)$ splňuje několik podmínek: $f(0) = f(1) = 0$, $f(u) > 0$, pokud $0 < u < 1$, $f'(0) > 0$ a $f'(u) < f'(0)$, pokud $0 < u \leqslant 1$. Autoři ukázali silnější výsledek, který je analogický Fisherovu: je-li počáteční podmínka taková, že $0 \leqslant u(x,0) \leqslant 1$, $u(x,0) = 0$ pro všechny $x < x_1$ a $u(x,0) = 1$ pro všechny $x > x_2 \geqslant x_1$, pak se gen šíří rychlostí $v^* = 2\sqrt{f'(0)\,D}$.

Hledání řešení $u(x,t) = U(z)$ kde $z = x + vt$ vede k zřejmému zobecnění rovnice (20.2)

$$vU'(z) = f(U(z)) + DU''(z).$$

Tuto diferenciální rovnici druhého řádu lze přepsat jako soustavu diferenciálních rovnic prvního řádu

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

Připomeňme, že $U(z)$ by mělo být takové, že $U(z) \to 0$, když $z \to -\infty$ a $U(z) \to 1$, když $z \to +\infty$. V blízkosti ustáleného stavu $(U = 0, p = 0)$ systému (20.4) máme $f(U) \approx f'(0)U$. Takže (20.4) lze aproximovat lineárním systémem

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

Hledáme-li exponenciální řešení ve tvaru $U(z) = U_0 e^{kz}$ a $p(z) = p_0 e^{kz}$, získáme charakteristickou rovnici

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

jako ve Fisherově článku. Opět platí, že k musí být reálné (jinak by u oscilovalo a nabývalo záporných hodnot). Tedy

$$v \geqslant 2\sqrt{f'(0)D} = v^*.$$

Oba kořeny pro k jsou pak reálné a kladné. Pokud $v > v^*$, jsou oba kořeny různé a ustálený stav $(U = 0, p = 0)$ je nestabilní uzel. Pokud $v = v^*$, jsou oba kořeny shodné a $(U = 0, p = 0)$ je nestabilní degenerovaný uzel, jak je znázorněno na obr. 20.3. Podobně systém (20.4) v blízkosti ustáleného stavu $(U = 1, p = 0)$ vede k lineárnímu systému

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

a charakteristické rovnici

$$Dk^2 - vk + f'(1) = 0.$$

Diskriminant je $v^2 - 4Df'(1) \geqslant 0$, protože $f'(1) \leqslant 0$. Jestliže $f'(1) < 0$, existují dva reálné kořeny opačného znaménka a $(U = 1, p = 0)$ je sedlový bod. Je-li $f'(1) = 0$, je jeden kořen nulový a druhý kladný (viz obr. 20.3). Podrobná analýza ve skutečnosti ukazuje, že pro všechny

$v \geqslant 2\sqrt{f'(0)D}$ existuje jedinečná integrální křivka spojující dva ustálené stavy $(U = 0, p = 0)$ a $(U = 1, p = 0)$, jako ve speciálním případě na obr. 20.3.

Kolmogorov, Petrovskij a Piskunov dále rigorózně ukázali, že parciální diferenciální rovnice (20.3) má jediné řešení $u(x,t)$ splňující počáteční podmínku, že toto řešení je takové, že $0 < u(x,t) \leqslant 1$ pro všechna x a $t > 0$, že $u(x,t)$ zůstává rostoucí funkcí x, pokud je tomu tak při $t = 0$, a konečně že $u(x,t)$ skutečně konverguje k vlnovému profilu šířícímu se rychlostí v^*. Důkazy jsou příliš dlouhé na to, abychom je zde shrnuli.

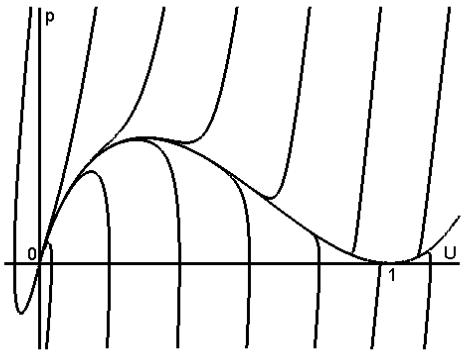
Všimněte si, že Fisherem použitá funkce

$$f(u) = au(1-u)$$

splňuje všechny tyto podmínky s $f'(0) = a$. Inspirováni rovnicí (14.5) uvažovali Kolmogorov, Petrovskij a Piskunov funkci

$$f(u) = au(1-u)^2,$$

která splňuje stejné podmínky a poskytuje stejnou rychlost šíření.



Obrázek 20.3: Schéma (U,p) zobrazující některé integrální křivky systému (20.5) a zejména jedinečnou křivku spojující $(U = 1, p = 0)$ s $(U = 0, p = 0)$, která udává tvar šířící se vlny. Zde platí $f(u) = au(1-u)^2$, $a = 1$, $D = 1$ a $v = v^* = 2$.

Články Fishera a Kolmogorova, Petrovského a Piskunova byly východiskem pro konstrukci mnoha matematických modelů s geografickou difúzí v genetice, ekologii a epidemiologii. Tyto modely jsou známy jako „reakčně-difúzní systémy“.

Pokud jde o Kolmogorova, od roku 1938 se zabýval také problémem vymírání rodových jmen, o němž uvažovali Bienaymé, Galton, Watson, Fisher, Haldane, Erlang a Steffensen: stochastický proces, který je všem těmto pracím společný, nazval „větvící proces“. V roce 1939 se stal členem Akademie věd SSSR. Později významně přispěl k problematice turbulence v mechanice tekutin (1941), k teorii dynamických systémů spojených s nebeskou mechanikou (1953) a k teorii informace (od roku 1956). Podílel se také na sepsání encyklopedie a středoškolských a vysokoškolských učebnic, pomáhal založit experimentální střední školu a redigoval populárně-vědecký časopis. Obdržel řadu mezinárodních cen (včetně Balzanovy ceny v roce 1962 a Wolfovy ceny v roce 1980) a zemřel v Moskvě v roce 1987.

Petrovskij se v roce 1940 stal děkanem Fakulty mechaniky a matematiky Moskevské státní univerzity. Od roku 1951 až do své smrti v roce 1973 byl rektorem univerzity. Od roku 1946 byl řádným členem Akademie věd SSSR a předsedou Mezinárodního kongresu matematiků, který se konal v Moskvě v roce 1966. Napsal také učebnice obyčejných diferenciálních rovnic, parciálních diferenciálních rovnic a integrálních rovnic. Piskunov se stal profesorem na vojenské akademii. Jeho učebnice diferenciálního a integrálního počtu se používala na mnoha technických univerzitách. Zemřel v roce 1977.

Další čtení

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, 1–26 (1937) $\to$ V.M. Tikhomirov (ed.) *Selected Works of A. N. Kolmogorov*, vol. 1, 242–270. Kluwer (1991).
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. Am. Math. Soc. (2007)
6. Shiryaev, A.N.: *Selected Works of A. N. Kolmogorov*, vol. 2. Kluwer (1992)
7. Shiryaev, A.N.: Andrei Nikolaevich Kolmogorov (April 25, 1903 to October 20, 1987). In: *Kolmogorov in Perspective*, 1–88. Am. Math. Soc. (2000)

Kapitola 21

Leslieho matice (1945)

V roce 1945 analyzoval britský ekolog P. H. Leslie maticový model věkově strukturované populace hlodavců, čímž adaptoval Lotkovu práci na diskrétní čas. Vyzdvihl, že rychlost růstu populace odpovídá vlastnímu číslu a stabilní věková struktura vlastnímu vektoru příslušné matice. Numericky také odhadl čistou míru reprodukce $\mathscr{R}_0$ pro populaci potkana.

Patrick Holt Leslie se narodil v roce 1900 poblíž Edinburghu ve Skotsku. Studoval na *Christ Church College* Oxfordské univerzity a v roce 1921 získal bakalářský titul z fyziologie. Studium medicíny však ze zdravotních důvodů nemohl dokončit. Potom, co několik let pracoval jako asistent bakteriologie na katedře patologie, se začal zabývat statistikou a v roce 1935 nastoupil do *Bureau of Animal Population*, nového výzkumného centra, které založil Charles Elton. Smyslem tohoto střediska bylo studovat kolísání populací zvířat prostřednictvím terénních studií a laboratorních experimentů. Většina výzkumů se týkala hlodavců: analýza populačních cyklů zajíce a jeho predátora rysa s využitím archivů kanadské společnosti *Hudson's Bay Company*, monitorování teritoriální expanze veverky popelavé na úkor veverky obecné v Anglii, sběr dat o hraboších v okolí Oxfordu a podobně. Na data o hraboších použil Leslie metody, které vyvinul Lotka pro studium lidské demografie. Během druhé světové války se výzkum střediska zaměřil na metody kontroly potkanů a myší v obilných silech.

Obrázek 21.1:
P. H. Leslie (1900–1972)

V roce 1945 Leslie publikoval svůj nejslavnější článek v časopise *Biometrika*, který v roce 1901 založili Galton, Pearson a Weldon. Článek nesl název *O užití matic v určité populační matematice*. Leslie se zde zabýval modelem růstu počtu samic v populaci zvířat, např. v populaci potkanů (ale mohlo by stejně tak jít o lidi). Populace je rozdělena do $K+1$ věkových skupin: $P_{k,n}$ je počet samic ve věku k v čase n ($k = 0,1,\dots,K$; $n = 0,1,\dots$). Nechť f_k je plodnost samic ve věku k nebo přesněji počet dcer narozených jedné samici mezi časem n a časem $n+1$. Pak K je maximální věk s nenulovou plodností ($f_K > 0$). Nazvěme dále s_k pravděpodobnost, že se zvíře ve věku k dožije alespoň věku $k+1$. Pak je věková struktura populace dána následující soustavou rovnic:

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

Všechna čísla f_k jsou kladná a s_k splňuje podmínky $0 < s_k < 1$. Na přelomu devatenáctého a dvacátého století si matematici uvykli psát takové soustavy rovnic ve zkráceném tvaru[1]

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

kde P_n je sloupcový vektor $(P_{0,n},\dots,P_{K,n})$ a M je čtvercová matice (tj. tabulka čísel s $K+1$ řádky a $K+1$ sloupci)

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

Aby pochopil chování systému (21.1) v závislosti na čase, hledal Leslie geometricky rostoucí nebo klesající řešení $P_n = r^n\,V$. Dosazením tohoto předpokládaného řešení do vztahu (21.1) zjistíme, že číslo r a vektor V musí splňovat rovnici

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

Číslo r se v tomto případě nazývá vlastní číslo matice M a vektor V vlastní vektor matice M. Problém tak spočívá v nalezení věkového rozdělení V, které

[1] To znamená, že $P_{k,n+1} = M_{k,0}\,P_{0,n} + M_{k,1}\,P_{1,n} + \cdots + M_{k,K}\,P_{K,n}$ pro všechna $k = 0,1,\dots,K$.

se pak v každém časovém kroku násobí konstantou r. Podle Lotkovy terminologie se taková rozdělení nazývají stabilní. Vrátíme-li se k obvyklejšímu zápisu, lze rovnici (21.2) rozepsat takto

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

Z posledních K rovnic plyne

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

Dosazením těchto hodnot do první rovnice, zkrácením V_0 a vynásobením r^K získal Leslie „charakteristickou rovnici“

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

Jedná se o polynomickou rovnici stupně $K+1$ v proměnné r. Existuje tedy $K+1$ reálných nebo komplexních kořenů $r_1, \ldots, r_{K+1}$ této rovnice. Leslie si navíc všiml (s použitím Descartova znaménkového pravidla pro polynomy), že existuje právě jeden reálný kladný kořen. Nazvěme jej r_1.

Leslie také uvažoval, že za většiny biologicky realistických podmínek (které lze exaktně vyjádřit s využitím Perronovy a Frobeniovy teorie pro nezáporné matice) je vlastní číslo r_1 větší než modul všech ostatních reálných či komplexních vlastních čísel matice M (nazvěme tato další vlastní čísla r_2, ... , r_{K+1}). Kromě toho jsou všechny řešení rovnice (21.3) obvykle různé. Pro každé vlastní číslo r_i lze najít přidružený vlastní vektor. Nechť Q je čtvercová matice řádu $K+1$, jejíž sloupce obsahují vlastní vektory přidružené vlastním číslům r_1, ... , r_{K+1}. Pak $MQ = QD$, kde D je diagonální matice $[r_1, \ldots, r_{K+1}]$. Odtud $M = QDQ^{-1}$ a

$$P_n = M^n P_0 = Q D^n Q^{-1} P_0 .$$

Všimněte si, že D^n je diagonální matice $[(r_1)^n, \ldots, (r_{K+1})^n]$ a že

$$D^n/(r_1)^n \longrightarrow \mathscr{D} = [1, 0, \ldots, 0]$$

pro $n \to +\infty$, neboť $r_1 > |r_i|$ pro $i \neq 1$. Proto $P_n/(r_1)^n$ konverguje k vektoru $Q \mathscr{D} Q^{-1} P_0$.

Každá složka vektoru věkové struktury P_n roste nebo klesá jako $(r_1)^n$. Pro $r_1 > 1$ tak populace roste exponenciálně. Pokud je však $r_1 < 1$, populace exponenciálně klesá. Z rovnice (21.3) lze snadno ukázat, že podmínka $r_1 > 1$ platí tehdy a jen tehdy, když parametr $\mathscr{R}_0$, definovaný vztahem

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

je větší než 1. Všimněte si, že výraz $s_0\, s_1 \cdots s_{k-1}$ určuje pravděpodobnost, že se dožijete alespoň věku k. Parametr $\mathscr{R}_0$ je tedy průměrný počet dcer narozených jedné samici během celého jejího života a je analogický vztahům (10.2), (12.2) a (16.9). Právě představený model je určitou časově diskrétní obdobou Lotkovy práce (viz kapitola 10) a zobecněním Eulerovy práce zahrnujícím plodnost závislou na věku (viz kapitola 3).

Leslie svou metodu ilustroval na datech o koeficientech plodnosti a přežití f_k a s_k u potkana, které publikoval jeho americký kolega. Po několika statistických operacích, které původní data vhodně doplnily, spočetl $\mathscr{R}_0 \approx 26$.

Leslieho maticovou formulaci problémů populační dynamiky dnes používá mnoho biologů. Výpočty jsou značně zjednodušeny moderními počítači a vědeckým softwarem, který dokáže vypočítat vlastní hodnoty a vlastní vektory libovolné matice. Snadno lze vypočítat jak parametr $\mathscr{R}_0$, tak rychlost růstu r_1.

Po druhé světové válce používal Leslie svou metodu k výpočtu rychlosti růstu populací dalších živočišných druhů: ptáků, brouků, atd. Zabýval se také stochastickými modely, modely konkurence mezi druhy a analýzou dat založených na zpětných odchytech. Do důchodu odešel v roce 1967. V témže roce odešel do důchodu i Charles Elton, *Bureau of Animal Population* přestalo existovat jako samostatné výzkumné středisko a stalo se součástí Oddělení zoologie Oxfordské univerzity. Leslie zemřel v roce 1972.

Další čtení

1. Anonymous: 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)

Kapitola 22

Perkolace a epidemie (1957)

V roce 1957 se Hammersley a Broadbent zabývali šířením „tekutiny“ v nekonečné pravidelné čtvercové mřížce, jejíž každé dva sousední uzly jsou propojeny jen s jistou pravděpodobností. Jako možný příklad takového procesu uvedli šíření epidemie v sadu. Ukázali, že existuje kritická pravděpodobnost propojení uzlů, při jejímž překročení se epidemie s kladnou pravděpodobností rozšíří na celou mřížku a při menší pravděpodobnosti nemůže vzniknout žádná velká epidemie. Jejich článek inicioval teorii perkolace.

John Michael Hammersley se narodil roku 1920 ve Skotsku, kde jeho otec pracoval pro americkou ocelářskou společnost. Zahájil studium na *Emmanuel College* Cambridgeské univerzity, ale v roce 1940 musel narukovat. V armádě pracoval na vylepšení dělostřeleckých výpočtů. Po ukončení studia v roce 1948 se stal asistentem na Oxfordské univerzitě ve skupině zabývající se navrhováním a analýzou experimentů. V roce 1955 nastoupil do Výzkumného ústavu pro atomovou energii v Harwellu u Oxfordu.

Obrázek 22.1:
Hammersley (1920–2004)

Simon Ralph Broadbent se narodil roku 1928. Vystudoval inženýrství v Cambridgi, matematiku na *Magdalen College* v Oxfordu (kde také psal básně) a zahájil doktorské studium statistiky na *Imperial College* v Londýně na téma *Některé testy odchylky od rovnoměrného rozptylu*. Během doktorátu získal podporu od Britské asociace pro výzkum využití uhlí, aby mohl zkoumat statistické problémy související s těžbou.

V roce 1954 se v Královské statistické společnosti v Londýně konalo sympozium o metodách Monte Carlo, které sponzoroval Ústav pro výzkum atomové energie. Tyto metody, jejichž vznik iniciovali ve 40. letech minulého století John von Neumann, Stanisław Ulam a Nicholas Metropolis v laboratoři v Los Alamos, využívají stochastické počítačové simulace k odhadu neznámých matematických veličin. Hammersley na londýnském sympoziu přednesl příspěvek, který připravil ve spolupráci s Mortonem, kolegou z Harwellu. Článek byl rovněž publikován v *Journal of the Royal Statistical Society*. Během diskuse, která následovala po prezentaci na sympoziu, Broadbent zmínil zajímavý problém, který by bylo možné studovat pomocí některé z metod Monte Carlo: je-li dána pravidelná síť pórů ve dvou nebo třech rozměrech tak, že dva sousední póry jsou spojeny s pravděpodobností p, jaká část sítě bude vyplněna plynem, kdyby pronikl do jednoho z těchto pórů? Broadbent přitom myslel na konstrukci plynových masek pro horníky a velikosti pórů ve filtru, která zajistí jejich funkčnost.

Hammersley pak začal s Broadbentem na problému plynových masek pracovat. Uvědomili si, že se jedná jen o jeden ze skupiny problémů, které dosud nebyly zkoumány: deterministické šíření „tekutiny“ (význam tohoto slova je určen kontextem) v náhodném prostředí. Hammersley ji nazval „perkolací“, tj. „průchodem rozpouštědla (např. vody) přes propustnou látku (např. mletou kávu) za účelem extrakce rozpustné složky“ podle tehdy populárního typu konvice na přípravu kávy. Ve Výzkumném ústavu pro atomovou energii měl Hammersley také přístup k nejvýkonnějším počítačům své doby, na kterých mohl testovat řešení perkolačních problémů metodou Monte Carlo.

V roce 1957 Broadbent a Hammersley konečně publikovali první článek o matematické teorii perkolace. Mezi příklady, které uvažovali, byl i model populační dynamiky, konkrétně šíření epidemie v ovocném sadu. Předpokládá se, že stromy velmi velkého sadu jsou umístěny v uzlech čtvercové sítě. Každý ze čtyř nejbližších sousedů nějakého nakaženého stromu má pravděpodobnost p, že bude také nakažen. Otázkou je, zda bude nakaženo velké množství stromů, nebo zda epidemie zůstane lokalizovaná. To samozřejmě závisí na pravděpodobnosti p, která zase souvisí se vzdáleností mezi jednotlivými stromy, tj. šířkou „oka sítě“.

Broadbent a Hammersley se zabývali limitním případem, kdy je sad nekonečný a pokrývá celou rovinu, přičemž na začátku je pouze jeden infikovaný strom. Nechť $f(p)$ je pravděpodobnost, že se z tohoto zdroje nakazí nekonečný počet stromů. Očekáváme, že $f(p)$ bude rostoucí funkcí pravděpodobnosti p, $f(0) = 0$ a $f(1) = 1$. Jejich hlavním výsledkem bylo, že existuje kritická pravděpodobnost p^*, $0 < p^* < 1$, taková, že:

- pokud $p < p^*$, pak $f(p) = 0$, takže je infikován pouze konečný počet stromů;
- pokud $p > p^*$, pak $f(p) > 0$ a infikován může být nekonečný počet stromů.

Důkaz spočívá v porovnání počtů různých „soběvyhýbajících se procházek“ v rovině vycházející ze zdroje infekce. Tyto procházky procházejí určitým počtem sousedních stromů (připomeňme, že každý strom má čtyři sousedy), aniž by některý strom navštívily více než jednou. Lze je interpretovat jako trasy šíření infekce. Soběvyhýbající se procházka s n kroky vznikne s pravděpodobností p^n, protože infekce se může přenést z každého navštíveného stromu na další s pravděpodobností p. Nyní nechť $q(j,n)$ je pravděpodobnost, že mezi všemi soběvyhýbajícími se procházkami s n kroky je právě j takových procházek, které jsou trasou rozšíření infekce. Jestliže existuje nekonečný počet infikovaných stromů, pak pro všechna celá čísla n existuje alespoň jedna soběvyhýbající se procházka s n kroky, která je trasou infekce. Takže

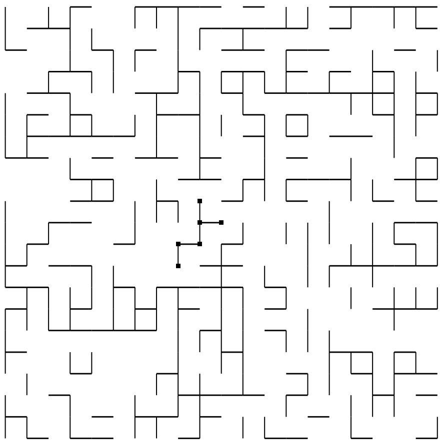
$$0 \leqslant f(p) \leqslant \sum_{j=1}^{+\infty} q(j,n) \leqslant \sum_{j=1}^{+\infty} j\, q(j,n)$$

pro všechny n. Ale $\sum_{j=1}^{+\infty} j\, q(j,n)$ je střední hodnotou počtu soběvyhýbajících se procházek s n kroky, které jsou trasami infekce. Tento počet je roven $p^n s(n)$, kde $s(n)$ je celkový počet soběvyhýbajících se procházek s n kroky. Hammersley později v navazujícím článku ukázal, že $s(n)$ roste jako $e^{\kappa n}$ pro $n \to +\infty$, kde konstanta κ se nazývá koeficient konektivity. Jestliže $p < e^{-\kappa}$, pak $p^n s(n)$ konverguje k 0 pro $n \to +\infty$ a tedy $f(p) = 0$. To znamená, že $p^* \geqslant e^{-\kappa} > 0$.

V praxi je proto lepší, když stromy nejsou příliš blízko sebe, aby se v případě nákazy pravděpodobnost p udržela pod kritickou hodnotou p^*. Čím jsou však stromy blíže, tím vyšší je produkce na hektar. Je třeba najít nějaký kompromis.

Jak si všimli již Broadbent a Hammersley, existuje jistá podobnost mezi existencí kritické pravděpodobnosti v perkolačních procesech a existencí prahové hodnoty v procesech větvení (viz kapitola 7).

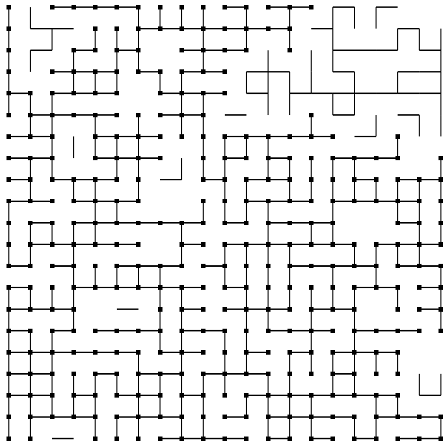
Kritickou pravděpodobnost p^* se můžeme pokusit numericky odhadnout. K tomu zvolíme nějakou hodnotu p a aproximujeme nekonečnou mřížku konečnou čtvercovou mřížkou o velikosti $N \times N$ s dostatečně velkým N. Předpokládejme například, že infikován je strom uprostřed mřížky. Pomocí počítače lze náhodně vybrat, které stromy mohou infikovat další stromy. Obrázek 22.2 a obrázek 22.3 ukazují náhodně zvolené trasy infekce jako hrany v grafu. Na obrázku 22.2 je pravděpodobnost p menší než kritická hodnota p^*. Na obrázku 22.3 je $p > p^*$. Z obrázků lze snadno určit, které stromy mohou být infikovány; jsou to ty, kterými vede cesta vycházející z infikovaného stromu uprostřed. Na obrázcích jsou označeny malými černými čtverečky.



Obrázek 22.2: Perkolace s $p = 0{,}4$.

Poté lze zkontrolovat, zda epidemie dosáhla hranice mřížky $N \times N$. Pokud tomu tak je a pokud je N dostatečně velké, můžeme považovat počet nakažených stromů za „téměř nekonečný“. Dostatečně velkým opakováním takové simulace lze zjistit přibližnou hodnotu pravděpodobnosti $f(p)$, že počet nakažených stromů je nekonečný (právě toto je metoda Monte Carlo). Nakonec necháme p probíhat hodnoty od 0 do 1 a získáme přibližnou hodnotu prahu p^*, což je nejmenší hodnota taková, že $f(p) > 0$, pokud $p > p^*$.

Broadbentův a Hammersleyho článek obsahoval pouze důkaz existence prahové hodnoty p^*. Hammersley v následujících letech dále rozvíjel matematickou teorii perkolace, Broadbent se věnoval jiným tématům. S rozvojem počítačů v sedmdesátých letech 20. století již bylo snadnější provádět výše

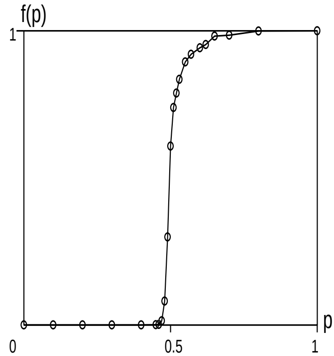


Obrázek 22.3: Perkolace s $p = 0{,}55$.

popsané simulace (obrázek 22.4). Ty vedly k odhadu $p^* = 1/2$. Tento výsledek nakonec roku 1980 dokázal Harry Kesten z Cornellovy univerzity.

V letech 1959–1969 pracoval Hammersley pro Ústav ekonomie a statistiky na Oxfordské univerzitě, stal se příslušníkem *Trinity College*. V roce 1964 vydal ve spolupráci s Davidem Handscombem knihu nazvanou *Metody Monte Carlo*. V roce 1976 byl zvolen členem Královské společnosti. V roce 1987 odešel do důchodu, ale nadále navštěvoval Oxfordské centrum pro průmyslovou a aplikovanou matematiku. Zemřel roku 2004.

Broadbent získal doktorát na *Imperial College* v roce 1957. Našel si práci ve společnosti *United Glass Bottle Manufacturers*. Po deseti letech práce v průmyslu začal pracovat ve zpravodajské agentuře *London Press Exchange*, kde prováděl vědecké studie čtenářství. Tuto agenturu koupila v roce 1969 americká reklamní společnost *Leo Burnett*. Broadbent zde pracoval na způsobu měření účinnosti reklamy a vydal na toto téma několik knih: *Výdaje na reklamu* (1975), *Rozpočet na reklamu* (1989), *Zodpovědná reklama* (1997) a *Kdy inzerovat* (1999). V roce 1980 se podílel na založení soutěže *Advertising Effectiveness Awards*. Několik let pracoval v centrále společnosti *Leo Burnett* v Chicagu jako ekonomický ředitel. Vedl také vlastní poradenskou společnost *BrandCon Limited*. Zemřel roku 2002.



Obrázek 22.4: Pravděpodobnost $f(p)$, že se nakazí nekonečně mnoho stromů, jako funkce p. Křivka je získána provedením 1 000 simulací na síti 200×200.

Další čtení

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. 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)

Kapitola 23

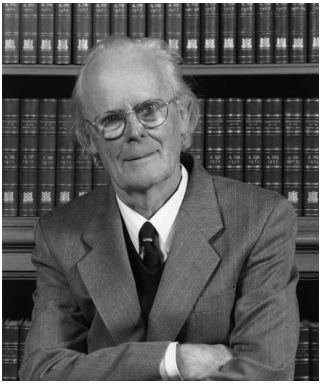
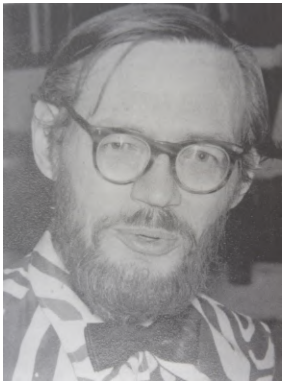
Teorie her a evoluce (1973)

V roce 1973 publikovali Maynard Smith a Price článek, ve kterém analyzovali, proč se zvířata vyhýbají použití svých nejnebezpečnějších zbraní ve vnitrodruhových konfliktech. Jejich model využívající teorii her byl jedním z těch, které zahájily aplikaci této matematické teorie na evoluční problémy.

John Maynard Smith se narodil v roce 1920 v Londýně. Jeho otec, chirurg, zemřel, když mu bylo osm let. Maynard Smith studoval na *Eton College* a poté se věnoval inženýrským studiím na univerzitě v Cambridge na *Trinity College*. V roce 1939, když vypukla válka, se hlásil jako dobrovolník do armády, ale byl odmítnut kvůli špatnému zraku. Po dokončení inženýrských studií několik let pracoval jako konstruktér vojenských letadel. Nakonec se ale rozhodl věnovat se biologii a studoval genetiku na *University College* v Londýně, kde byl jeho školitelem Haldane. V roce 1952 se stal lektorem zoologie. Po událostech v Maďarsku v roce 1956 vystoupil z komunistické strany Velké Británie, jejímž byl členem. Jeho první kniha s názvem *Teorie evoluce* vyšla v roce 1958. V roce 1965 se stal profesorem biologie na nově založené univerzitě v Sussexu. Poté vydal další dvě knihy: *Matematické myšlenky v biologii* (1968) a *O evoluci* (1972).

George R. Price se narodil v roce 1922 v USA. Vystudoval chemii na Chicagské univerzitě, kde v roce 1946 získal doktorát poté, co se podílel na projektu Manhattan konstrukce atomové bomby. V roce 1950 se stal pomocným vědeckým pracovníkem v oblasti medicíny na Minnesotské univerzitě. Později pracoval jako nezávislý novinář pro různé časopisy, načež se vrátil k výzkumu v IBM. V roce 1967, poté co se léčil s rakovinou štítné žlázy, se usadil v Anglii, kde se začal věnovat studiu evoluční biologie. Od roku 1968 pracoval v Galtonově laboratoři na *University College* v Londýně. Jeho první článek v této nové oblasti, *Selekce a kovariance*, byl publikován s pomocí W. D. Hamiltona v roce 1970 v časopise *Nature* a obsahoval to, co se dnes nazývá Priceova rovnice.

Price do časopisu *Nature* zaslal také další článek, tentokrát o konfliktech mezi zvířaty. Maynard Smith, který byl recenzentem, navrhl připravit kratší verzi. Mezitím Price začal pracovat na jiné problematice, zatímco Maynard Smith rozvíjel Priceovu myšlenku. Nakonec Maynard Smith a Price

Obrázek 23.1: Maynard Smith (1920–2004) a Price (1922–1975)

publikovali společně článek s názvem *Logika zvířecího konfliktu*, který vyšel v *Nature* roku 1973. Článek byl zajímavým příspěvkem k využití teorie her v evoluční biologii. Počátky teorie her jsou spojovány s knihou Johna von Neumanna a Oskara Morgensterna z roku 1944 s názvem *Teorie her a ekonomického chování*, kde byla tato teorie převážně aplikována na ekonomické problémy. Otázka, kterou Maynard Smith a Price řešili byla: Jak je možné, že při konfliktech mezi zvířaty stejného druhu jsou „zbraně", které mají k dispozici (rohy, drápy, jed atd.), jen zřídkakdy použity k zabíjení? V souladu s Darwinovými myšlenkami o přírodním výběru by agresivnější zvířata měla vyhrávat více soubojů a mít tedy i větší počet potomků, což by vedlo k eskalaci agresivního chování. Vzhledem k tomu, že se jednalo o období studené války, uvedená problematika měla i politický nádech.

Maynard Smith a Price si představili sled her, v nichž dva živočichové soutěží o zdroj, například o území na příznivém stanovišti. Ve zjednodušeném podání, které Maynard Smith použil ve své knize z roku 1982 *Evoluce a teorie her*, používá každé zvíře buď strategii „jestřába", nebo strategii „hrdličky". V následujícím textu budeme mluvit jednoduše o jestřábech a hrdličkách, ale budeme mít na mysli strategie používané zvířaty stejného druhu. Je-li $V > 0$ hodnota zdroje a $\mathscr{R}_0$ je normální průměrný počet potomků zvířete, má vítěz soutěže v průměru $\mathscr{R}_0 + V$ potomků.

Pokud se jestřáb setká s jiným jestřábem, dojde k souboji o zdroj: vítěz získá zdroj v hodnotě $V > 0$ a poražený nese „náklady" $C > 0$. Každý z obou jestřábů má pravděpodobnost rovnou 1/2, že v soutěži zvítězí, a stejnou pravděpodobnost, že prohraje. Průměrná výhra ze souboje dvou jestřábů je tedy $\frac{1}{2}(V - C)$ pro oba soutěžící. Pokud se však jestřáb setká s hrdličkou, pak jestřáb získá zdroj V, hrdlička unikne bez boje a nezíská nic. A konečně, potkají-li se dvě hrdličky, jedna z nich získá zdroj V, zatímco druhá unikne bez boje a s nulovými náklady. Každá z obou hrdliček má stejnou pravděpodobnost 1/2 výhry, průměrná výhra je tedy $V/2$. Výplaty lze shrnout do tabulky 23.1.

Tabulka 23.1: Očekávané výhry ve hře jestřáb-holubice.

	jestřábovi	holubici
výplata jestřába proti...	$\frac{1}{2}(V-C)$	V
výplata holubice proti...	0	$V/2$

Obecněji si lze představit souboje mezi jedinci, kteří mohou zvolit jednu ze dvou strategií, nazvěme je 1 a 2, s maticí očekávaných výher $(G_{i,j})_{1\leqslant i,j\leqslant 2}$. Ve výše uvedeném příkladu se jestřábi řídí strategií 1, hrdličky strategií 2, $G_{1,1}=\frac{1}{2}(V-C)$, $G_{1,2}=V$, $G_{2,1}=0$ a $G_{2,2}=V/2$. V původním článku z roku 1973 Maynard Smith a Price použili počítačové simulace k testování více než dvou možných strategií (ty byly nazvány „jestřáb“, „myš“, „tyran“, „odvetník“ a „pokušitel-odvetník“).

Představme si nyní velkou populaci zvířat stejného druhu s podílem x_n jestřábů a podílem $1-x_n$ hrdliček v generaci n. Jestřábi v generaci n mají průměrný počet potomků

$$R_1(n)=\mathscr{R}_0+x_n\,G_{1,1}+(1-x_n)\,G_{1,2}\,. \tag{23.1}$$

Podobně mají hrdličky průměrný počet potomků

$$R_2(n)=\mathscr{R}_0+x_n\,G_{2,1}+(1-x_n)\,G_{2,2}\,. \tag{23.2}$$

Průměrný počet potomků v celé populaci je tedy

$$R(n)=x_n\,R_1(n)+(1-x_n)\,R_2(n)\,.$$

Zapomeneme-li na případné jemnosti způsobené pohlavním rozmnožováním, vidíme, že podíl jestřábů v další generaci činí

$$x_{n+1}=x_n\,R_1(n)/R(n)\,. \tag{23.3}$$

Z toho vyplývá, že $x_{n+1}>x_n$ pokud $R_1(n)>R(n)$ a $x_{n+1}<x_n$ pokud $R_1(n)<R(n)$. Existují tři možné ustálené (stacionární) stavy: $x=0$, $x=1$ a

$$x^*=\frac{G_{1,2}-G_{2,2}}{G_{2,1}-G_{1,1}+G_{1,2}-G_{2,2}}$$

pokud $0 < x^* < 1$. Ve hře na jestřába a hrdličku platí

$$x^* = V/C < 1$$

pokud $V < C$.

Skutečně, $x = 0$ je zřejmě ustálený stav (23.3). Jestliže $x \neq 0$, tak další ustálené stavy musí splňovat

$$R_1 = R = xR_1 + (1-x)R_2.$$

Takže buď $x = 1$, nebo $R_1 = R_2$. Druhá možnost je ekvivalentní s

$$x\,G_{1,1} + (1-x)\,G_{1,2} = x\,G_{2,1} + (1-x)\,G_{2,2},$$

což dává ustálený stav x^*.

Ustálený stav $x = 1$ odpovídá populaci se 100 % jedinců, kteří se řídí strategií 1. Tento ustálený stav je stabilní, pokud nemůže být narušen několika jedinci, kteří se řídí strategií 2. Z (23.3) vidíme, že tato podmínka je ekvivalentní podmínce, že $R_1(n) > R(n)$ pro x_n dostatečně blízké 1. Protože $R(n) = x_n R_1(n) + (1 - x_n)R_2(n)$, lze tuto podmínku zapsat ve tvaru $R_1(n) > R_2(n)$ pro x_n dostatečně blízko 1. Podíváme-li se na výrazy (23.1)–(23.2) dojdeme k závěru, že $x = 1$ je stabilní tehdy a jen tehdy, je-li splněna jedna z následujících dvou podmínek:

- $G_{1,1} > G_{2,1}$;
- $G_{1,1} = G_{2,1}$ a $G_{1,2} > G_{2,2}$.

Pokud jsou tyto dvě podmínky splněny, pak se strategie 1 nazývá „evolučně stabilní strategie“[1]. Ve hře jestřáb-hrdlička platí vždy podmínka $G_{1,2} > G_{2,2}$. Jestřábí strategie je tedy evolučně stabilní tehdy a jen tehdy, když $G_{1,1} \geqslant G_{2,1}$, tj. když $V \geqslant C$.

Ustálený stav $x = 0$ odpovídá populaci, ve které se všichni jedinci řídí strategií 2. Tato situace je symetrická k předchozí, pokud vyměníme indexy 1 a 2. Ve hře jestřáb-hrdlička máme $G_{1,2} = V > G_{2,2} = V/2$, takže ustálený stav $x = 0$ je vždy nestabilní, protože strategie jestřába se bude v populaci hrdliček šířit.

[1] Přesněji řečeno, ustálený stav $x = 1$ je „evolučně stabilní stav“. Rozdíl mezi stavem a strategií je důležitý, pokud existují více než dvě strategie.

Podobně lze ukázat, že třetí ustálený stav x^* je za předpokladu, že $0 < x^* < 1$, vždy stabilní. Ve hře na jestřáby a holubice odpovídá $x^* = V/C$ smíšené populaci s jestřáby i hrdličkami.

Závěrem lze říci, že ve hře jestřáb-hrdlička existují dva případy. Je-li $V \geqslant C$, tj. je-li hodnota zdroje větší než možné náklady, pak monomorfní populace skládající se z jestřábů je stabilním stavem bez ohledu na počáteční stav $x(0)$ s $0 < x(0) < 1$. Strategie jestřába je tedy evolučně stabilní strategií. Jestliže naopak $V < C$, pak populace směřuje ke smíšenému ustálenému stavu s podílem x^* jestřábů a podílem $1 - x^*$ hrdliček. Model tedy skutečně poskytuje vysvětlení, proč jedinci s méně agresivním chováním mohou přežít, když $V < C$. Ze vzorce $x^* = V/C$ navíc vyplývá, že čím vyšší jsou náklady C pro poražené, tím menší je podíl x^* jestřábů v populaci. Proto druhy s nejnebezpečnějšími „zbraněmi“ je jen zřídka používají k vnitrodruhovým soubojům: dávají přednost neškodným rituálním soubojům při nichž se soupeřící zvířata navzájem zastrašují, ale vyhýbají se skutečným soubojům, které by mohly způsobit zranění.

Původní článek Maynarda Smithe a Price z roku 1973 pojednával o konceptu evolučně stabilní strategie a využíval především počítačové simulace zvířecích soubojů, přičemž zaznamenával výhry různých strategií. Přístup využívající dynamické rovnice, jako např. (23.3), byl rozvinut o něco později, zejména Taylorem a Jonkerem. Od té doby mnoho autorů aplikovalo myšlenky z teorie her na otázky evoluční biologie, nebo naopak použilo dynamické evoluční přístupy na klasičtější problémy teorie her. Kromě otázek týkajících se konfliktů mezi zvířaty lze uvést například problémy rodičovských investic nebo poměru pohlaví (poměr mezi počtem samců a samic při narození), které studoval již Carl Düsing v roce 1884 a Ronald Fisher ve své knize *Genetická teorie přírodního výběru* z roku 1930. Některé další modely se zaměřují na dynamické aspekty „vězňova dilematu“ nebo hry „kámen-nůžky-papír“. Pojem evolučně stabilní strategie také úzce souvisí s pojmem Nashovy rovnováhy v teorii her.

Price, který byl přesvědčeným ateistou, měl v roce 1970 mystický zážitek a konvertoval ke křesťanské víře. V roce 1974 se vzdal svého výzkumu, protože měl pocit, že „teoretická matematická genetika, kterou se zabýval, není příliš relevantní pro lidské problémy“. Všechny své věci rozdal bezdomovcům a o několik měsíců později spáchal sebevraždu.

Maynard Smith naproti tomu pokračoval v tomto myšlenkovém směru a v roce 1977 byl zvolen do Královské společnosti. Vydal řadu knih: *Modely v ekologii* (1974), *Evoluce pohlaví* (1978), *Evoluce a teorie her* (1982), *Problémy biologie* (1986), *Měl Darwin pravdu?* (1988) a *Evoluční genetika* (1989). Ve spolupráci s E. Szathmárym vydal také knihu *Hlavní přechody* *v evoluci* (1995) a *Původ života: Od zrodu života ke vzniku jazyka* (1999). V roce 1985 odešel do důchodu. V roce 1999 obdržel Crafoordovu cenu za biologické vědy od Královské švédské akademie věd za „zásadní přínos ke koncepčnímu rozvoji evoluční biologie“. V roce 2003 publikoval ve spolupráci s D. Harperem *Zvířecí signály*. Zemřel v Sussexu v roce 2004.

Další čtení

1. Charlesworth, B., Harvey, P.: John Maynard Smith, 6 January 1920 – 19 April 2004. *Biog. Mem. Fellows R. Soc.* 51, 253–265 (2005)
2. Edwards, A.W.F.: Carl Düsing (1884) on the regulation of the sex-ratio. *Theor. Pop. Biol.* 58, 255–257 (2000)
3. Frank, S.A.: George Price's contributions to evolutionary genetics. *J. Theor. Biol.* 175, 373–388 (1995)
4. 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/ento/-pdfs/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

Kapitola 24

Chaotické populace (1974)

V roce 1974 se Robert May, australský fyzik a ekolog, zabýval logistickou rovnicí s diskrétním časem jako modelem populační dynamiky. Všiml si, že dochází k neočekávaným bifurkacím a že asymptotické chování může být dokonce chaotické. Dlouhodobé předpovědi tak mohou být nemožné i u jednoduchého deterministického modelu. Mayův článek byl jedním z těch, které odstartovaly „teorii chaosu“.

Robert McCredie May se narodil v roce 1936 v Austrálii. Po studiu teoretické fyziky a získání doktorátu na univerzitě v Sydney v roce 1959 strávil dva roky na katedře aplikované matematiky na Harvardově univerzitě. Po návratu do Austrálie se stal profesorem teoretické fyziky. V roce 1971 při návštěvě Institutu pro pokročilá studia v Princetonu změnil předmět svého výzkumu a začal se věnovat populační dynamice živočichů. V roce 1973 se stal profesorem zoologie v Princetonu. V témže roce vydal knihu s názvem *Stabilita a komplexita v modelových ekosystémech*.

Obrázek 24.1:
Robert M. May (1936–2020)

V roce 1974 publikoval May v časopise *Science* článek s názvem *Biologické populace s nepřekrývajícími se generacemi: stabilní body, stabilní cykly a chaos*, v němž ukázal, že velmi jednoduché matematické modely populační dynamiky se mohou chovat chaoticky.

Abychom pochopili původ tohoto problému, musíme se vrátit asi o deset let zpět v čase. V roce 1963 si Edward Lorenz, americký meteorolog pracující na Massachusettském technologickém institutu (M.I.T.), při numerických simulacích na svém počítači všiml, že zjednodušený model atmosféry popsaný třemi diferenciálními rovnicemi se může chovat velmi překvapivým způsobem: nepatrná změna počátečních podmínek může zcela změnit konečný výsledek simulace, a tedy i meteorologické předpovědi. Matematik Henri Poincaré, který se zabýval pohybem planet ve Sluneční soustavě, ve skutečnosti o této možnosti uvažoval již na počátku dvacátého století, dávno před nástupem počítačového věku. Na počátku sedmdesátých let se však touto zvláštní vlastností začalo blíže zabývat jen několik badatelů. Na Marylandské univerzitě přemýšlel James Yorke o Lorenzově práci a zavedl v této souvislosti termín „chaos“. Článek *Perioda tři implikuje chaos*[1], který napsal se svým studentem Tien-Yienem Li, byl publikován v roce 1975.

May se zabýval modelem

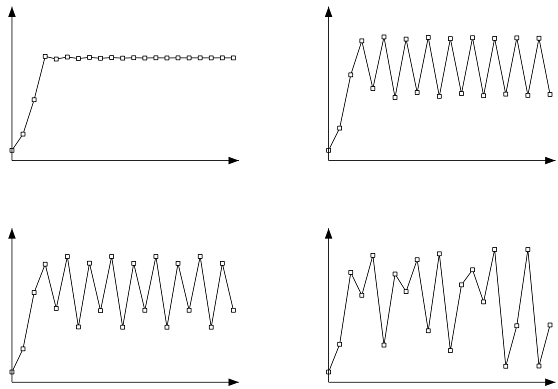
$$p_{n+1} = p_n + a\,p_n(1 - p_n/K), \tag{24.1}$$

kde a a K jsou kladné parametry a p_n je velikost populace zvířat v roce n. Je-li p_n malé ve srovnání s nosnou kapacitou K, dynamika se blíží geometrickému růstu $p_{n+1} \approx (1+a)\,p_n$. Úplná rovnice je jakousi diskrétní časovou obdobou logistické rovnice, kterou zavedl Verhulst (viz kapitola 6). Na rozdíl od ní však May ukázal, že rovnice s diskrétním časem může mít mnohem překvapivější chování, které lze snadno pozorovat pomocí jednoduché kapesní kalkulačky (obrázek 24.2). Maynard Smith se rovnicí (24.1) zabýval již ve své knize *Matematické myšlenky v biologii* z roku 1968. Ale přestože vyzkoušel několik číselných hodnot pro a, neuvědomil si, že je na ní něco mimořádného.

Obrázek 24.2, který je podobný obrázku v Mayově článku z roku 1974, ukazuje, že pro $0 < a < 2$ velikost populace p_n konverguje k ustálenému stavu. Pro $2 < a \leqslant 2{,}449$ (horní hranice 2,449 je aproximace), populace p_n konverguje k cyklu s periodou 2, pro $2{,}450 \leqslant a \leqslant 2{,}544$ k cyklu s periodou 4, pro $2{,}545 \leqslant a \leqslant 2{,}564$ k cyklu s periodou 8 atd. Intervaly parametru a, pro které p_n konverguje k cyklu s periodou 2^n, se s rostoucím n zmenšují a nikdy nepřekročí hodnotu 2,570. Pokud je $a \geqslant 2{,}570$, může se p_n chovat „chaoticky“.

V roce 1976 napsal May o tomto problému přehledový článek *Jednoduché matematické modely s velmi komplikovanou dynamikou*, který byl publikován v časopise *Nature*. Shromáždil v něm nejen své vlastní výsledky, ale

[1] Podstatně obecnější výsledek dokázal O. M. Šarkovskij v roce 1964, ale jeho článek, publikovaný v ukrajinském matematickém časopise, nebyl příliš známý.



Obrázek 24.2: Na všech obrázcích: n je na vodorovné ose, p_n na svislé ose a $p_0 = K/10$. Přímky získáme spojením bodů se souřadnicemi (n, p_n). Vlevo nahoře: $0 < a < 2$ (ustálený stav). Vpravo nahoře: $2 < a \leqslant 2{,}449$ (cyklus s periodou 2). Vlevo dole: $2{,}450 \leqslant a \leqslant 2{,}544$ (cyklus s periodou 4). Vpravo dole: $2{,}570 \leqslant a \leqslant 3$ (pravděpodobně chaos).

i výsledky dalších badatelů. Substituce

$$x_n = \frac{a\,p_n}{K\,(1+a)}$$

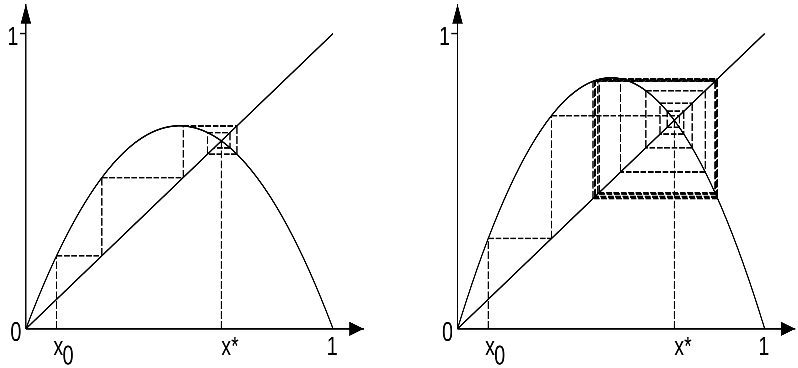
a $r = 1 + a$ (takže $r > 1$ pro $a > 0$) transformují rovnici (24.1) do jednoduššího tvaru

$$x_{n+1} = r x_n \,(1 - x_n)\ . \tag{24.2}$$

Aby tato rovnice měla v populační dynamice smysl, musí být x_n nezáporné pro všechna n. Předpokládáme tedy, že počáteční podmínka x_0 splňuje $0 \leqslant x_0 \leqslant 1$ a $r \leqslant 4$. Druhá podmínka zajišťuje, že hodnoty funkce na pravé straně (24.2) jsou mezi 0 a 1. Pozoruhodné je, že chaotický případ pro $r = 4$ použili jako generátor náhodných čísel Stanisław Ulam a John von Neumann již v roce 1947. Zavedeme-li funkci

$$f(x) = rx(1-x),$$

pak lze rovnici (24.2) přepsat jako $x_{n+1} = f(x_n)$ a ustálené stavy jsou řešení $x = f(x)$. Graficky jsou to průsečíky křivek $y = f(x)$ a $y = x$ (obrázek 24.3). Všimněte si, že $x = 0$ je vždy ustálený stav. Protože $r > 1$, existuje další ustálený stav $x^* > 0$ takový, že $x^* = rx^*\,(1 - x^*)$, tj. že $x^* = 1 - 1/r$.



Obrázek 24.3: Funkce $y = f(x)$, přímka $y = x$, ustálený stav x^* a posloupnost definovaná vztahem $x_{n+1} = f(x_n)$. Vlevo: $r = 2{,}75$, posloupnost konverguje k x^*. Vpravo: $r = 3{,}4$, ustálený stav x^* je nestabilní a posloupnost konverguje k cyklu s periodou 2.

Protože $r > 1$, je ustálený stav $x = 0$ nestabilní. Když je x_n blízko u 0, máme $x_{n+1} \approx r x_n$, takže řešení x_n má tendenci vzdalovat se od 0. Ustálený stav x^* je lokálně stabilní pro $1 < r < 3$.

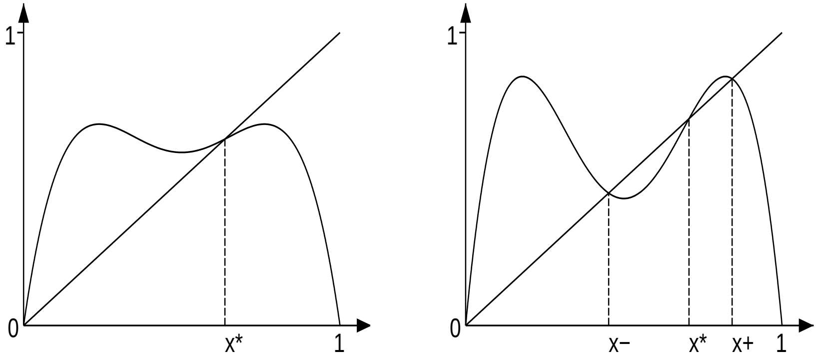
Vskutku, uvažujme $y_n = x_n - x^*$. Pak (24.2) je ekvivalentní $y_{n+1} = (2 - r - r y_n) y_n$. Je-li x_n blízko x^*, pak y_n je blízko 0 a $y_{n+1} \approx (2-r) y_n$. Je-li $y_{n+1} = k y_n$, pak $y_n = k^n y_0$, takže $y_n \to 0$ pro $n \to +\infty$ tehdy a jen tehdy, když $-1 < k < 1$. Ustálený stav x^* je tedy lokálně stabilní právě když $-1 < 2 - r < 1$, tj. $1 < r < 3$.

Pro $1 < r < 3$ lze ukázat, že pro všechny počáteční podmínky $0 < x_0 < 1$ posloupnost x_n konverguje k x^* (obrázek 24.3a). Co se však stane, když $3 < r \leqslant 4$? Abychom na tuto otázku odpověděli, všimněme si nejprve, že $x_{n+2} = f(x_{n+1}) = f(f(x_n))$. Zavedeme funkci

$$f_2(x) = f(f(x)) = r^2 x(1-x)\,(1 - r x(1-x))$$

a uvažujeme řešení rovnice $x = f_2(x)$, která se nazývají pevné body funkce $f_2(x)$. Graficky jsou to průsečíky křivek $y = f_2(x)$ a $y = x$ (obrázek 24.4).

Jestliže $x = f(x)$, pak $x = f(f(x)) = f_2(x)$. Takže $x = 0$ a $x = x^*$ jsou také pevné body funkce $f_2(x)$. Pokud však $r > 3$, má funkce $f_2(x)$ další dva pevné body, x_- a x_+, takové, že $f(x_-) = x_+$ a $f(x_+) = x_-$.



Obrázek 24.4: Křivky $y = f_2(x) = f(f(x))$ a $y = x$ a ustálený stav x^*. Vlevo: $r = 2{,}75$. Vpravo: $r = 3{,}4$ a další dvě řešení x_- a x_+ rovnice $x = f_2(x)$.

Všimneme si, že
$$f_2'(x) = f'(f(x))\,f'(x),$$
takže $f_2'(x^*) = [f'(x^*)]^2$. Protože $f'(x) = r(1-2x)$ a $x^* = 1 - 1/r$, tak $f'(x^*) = 2 - r$ a $f_2'(x^*) = (2-r)^2$. Proto je sklon funkce $f_2(x)$ v bodě $x = x^*$ takový, že pro $r > 3$ je $f_2'(x^*) > 1$. Protože však $f_2(0) = 0$, $f_2'(0) = r^2 > 1$ a $f_2(1) = 0$, vidíme na obrázku 24.4b, že nutně existují dvě další řešení x_- a x_+ rovnice $x = f_2(x)$, přičemž $0 < x_- < x^*$ a $x^* < x_+ < 1$. Jiný způsob, jak dojít ke stejnému závěru, spočívá v řešení rovnice $x = f_2(x)$, což je polynomiální rovnice stupně 4 se dvěma známými kořeny: $x = 0$ a $x = x^*$. Další dvě řešení x_- a x_+ jsou kořeny polynomu
$$x^2 - \frac{1+r}{r}x + \frac{1+r}{r^2} = 0\,. \tag{24.3}$$
Pokud je diskriminant kladný, tj. pokud $r > 3$, existují dvě reálná řešení. Protože
$$f_2(f(x_-)) = f(f(f(x_-))) = f(f_2(x_-)) = f(x_-),$$
je bod $f(x_-)$ také pevným bodem $f_2(x)$. Ale $f(x_-) \neq x_-$, protože x_-

není pevným bodem $f(x)$. A $f(x_-) \neq x^*$, jinak bychom měli

$$x_- = f(f(x_-)) = f(x^*) = x^*.$$

Protože $f(x_-) \neq 0$, docházíme k závěru, že $f(x_-) = x_+$. Podobně platí $f(x_+) = x_-$.

Proto pro $r > 3$ vidíme, že pokud například $x_0 = x_-$, pak $x_1 = x_+$, $x_2 = x_-$, $x_3 = x_+$ atd. Lze také ukázat, že téměř pro každou počáteční podmínku $0 < x_0 < 1$ posloupnost x_n konverguje pro $n \to +\infty$ k cyklu s periodou 2 x_-, x_+, x_-, x_+ atd. (obrázek 24.3b a 24.4b). Tento cyklus zůstává stabilní, dokud je r pod kritickou hodnotou $r_1 = 1 + \sqrt{6} \approx 3,449$, při které $f_2'(x_-) = -1$.

Pomocí (24.3) totiž vidíme, že

$$\begin{aligned} f_2'(x_-) &= f'(f(x_-))\, f'(x_-) = f'(x_+)\, f'(x_-) \\ &= r^2\,(1-2x_+)(1-2x_-) = r^2\,(1-2(x_+ + x_-) + 4x_+x_-) \\ &= r^2 \left(1 - 2\,\frac{1+r}{r} + 4\,\frac{1+r}{r^2}\right) = -r^2 + 2r + 4\,. \end{aligned}$$

Takže $f_2'(x_-) = -1$, pokud $-r^2 + 2r + 5 = 0$ a tedy $r = 1 + \sqrt{6}$.

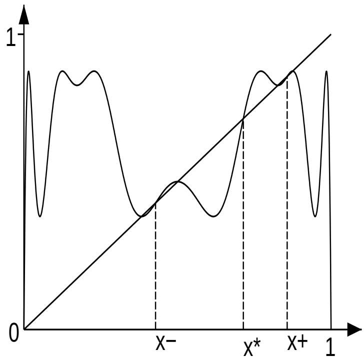
Pro $r_1 < r < r_2$ se cyklus s periodou 4 stává stabilním: objeví se čtyři nové pevné body funkce

$$f_4(x) = f_2(f_2(x)) = f(f(f(f(x))))$$

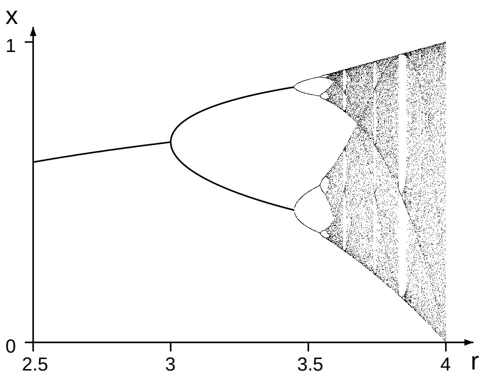
(obrázek 24.5). Pro $r_2 < r < r_3$ je to cyklus délky 8 atd. Čísla r_n konvergují k hodnotě $r_\infty \approx 3{,}570$ pro $n \to +\infty$. Když $r_\infty < r \leqslant 4$, může být systém dokonce chaotický! Obrázek 24.6 ukazuje bifurkační diagram[2], což dává představu o složitosti dynamiky.

R. M. May na závěr zdůraznil, že i velmi jednoduché dynamické systémy mohou mít velmi komplikované chování. To se netýká pouze rovnice $x_{n+1} = r\,x_n\,(1-x_n)$. Stejná kaskáda zdvojení periody vedoucí k chaosu se objevuje i u jiných rovnic s funkcí $f(x)$, která má tvar „kopce“. To je například případ jiné rovnice používané v populační biologii: $x_{n+1} = x_n\,\exp(r(1-x_n))$.

[2]Tento diagram byl získán vynesením bodů se souřadnicemi (r, x_{200}), $(r, x_{201}), \ldots, (r, x_{220})$ pro každou danou hodnotu r, kde $x_{n+1} = f(x_n)$ a $x_0 = 0,1$. Jestliže x_n směřuje k ustálenému stavu, vidíme v diagramu pouze jeden bod. Pokud x_n směřuje k cyklu s periodou 2, vidíme dva body atd.



Obrázek 24.5: Křivka $y = f_4(x)$ při $r = 3,5$ a přímka $y = x$. Kromě x^*, x_+ a x_- existují další čtyři pevné body, které není snadné rozlišit.



Obrázek 24.6: Bifurkační diagram rovnice (24.2).

Tato studie naznačuje, že bychom neměli být překvapeni, pokud je mnoho datových souborů týkajících se populační dynamiky obtížně analyzovatelných. Model také ukazuje, že rozdíl mezi deterministickými a stochastickými modely není tak jasný, jak se dříve myslelo: i u jednoduchého deterministického modelu může být nemožné vytvořit dlouhodobou předpověď, pokud jsou hodnoty parametrů v oblasti ve které dochází k chaotickému chování řešení.

V roce 1979 byl May zvolen členem Královské společnosti. V letech 1988 až 1995 působil jako profesor na Oxfordské univerzitě a na *Imperial College* v Londýně. V letech 1995 až 2000 byl hlavním vědeckým poradcem britské vlády. V roce 1996 obdržel Crafoordovu cenu „za průkopnický ekologický výzkum týkající se teoretické analýzy dynamiky populací, společenstev a ekosystémů“. Od ekologie se zaměřil na epidemiologii a imunologii a vydal dvě knihy: *Infekční nemoci lidí* (1991, spolu s Royem Andersonem) a *Dynamika virů, Matematické základy imunologie a virologie* (2000, s Martinem Nowakem). Druhá jmenovaná kniha analyzuje interakci mezi buňkami imunitního systému a virem HIV (virus způsobující AIDS) jako určitý druh systému dravec-kořist (viz kapitola 13). V letech 2000 až 2005 byl May prezidentem Královské společnosti. V roce 1996 byl povýšen do rytířského stavu a v roce 2001 se stal doživotním peerem. Zemřel v roce 2020.

Další čtení

1. Gleick, J.: *Chaos, Making a New Science*. Viking Penguin, New York (1987)
2. Levin, S.A.: Robert May receives Crafoord prize. *Not. Amer. Math. Soc.* 43, 977–978 (1996) ams.org
3. Li, T.Y., Yorke, J.A.: Period three implies chaos. *Amer. Math. Monthly* 82, 985–992 (1975)
4. Lorenz, E.N.: Deterministic nonperiodic flow. *J. Atmosph. Sci.* 20, 130–141 (1963) journals.ametsoc.org
5. May, R.M.: Biological populations with nonoverlapping generations: stable points, stable cycles and chaos. *Science* 186, 645–647 (1974)
6. May, R.M.: Simple mathematical models with very complicated dynamics. *Nature* 261, 459–467 (1976)
7. May, R.M., Oster, G.F.: Bifurcations and dynamic complexity in simple ecological models. *Amer. Natur.* 110, 573–599 (1976)
8. Maynard Smith, J.: *Mathematical Ideas in Biology*. Cambridge (1968)
9. Poincaré, H.: *Science et Méthode*. Flammarion, Paris (1908) gallica.bnf.fr
10. Sharkovsky, O.M.: Co-existence of cycles of a continuous mapping of a line onto itself. *Ukr. Math. J.* 16, 61–71 (1964)
11. Ulam, S.M., von Neumann, J.: On combination of stochastic and deterministic processes. *Bull. Amer. Math. Soc.* 53, 1120 (1947) ams.org

Kapitola 25

Čínská politika jednoho dítěte (1980)

V roce 1980 Song Jian a jeho spolupracovníci, kteří se zabývali teorií řízení aplikovanou na letecké inženýrství, spočítali, že pokud by porodnost v Číně zůstala na současné úrovni, dosáhla by populace během 21. století více než dvou miliard. Jejich výsledky založené na věkově strukturovaném matematickém modelu přispěly k rozhodnutí vlády přejít na politiku jednoho dítěte.

Song Jian[1] se narodil v roce 1931 v Rongchengu v čínské provincii Shandong. V padesátých letech studoval v Sovětském svazu na Moskevské státní technické univerzitě Bauman a také na Ústavu matematiky a mechaniky Moskevské státní univerzity. Poté se vrátil do Číny a stal se vedoucím kanceláře kybernetického výzkumu v Matematickém institutu Čínské akademie věd. Byl specialistou na aplikaci teorie řízení při navádění raket. Pracoval také pro Sedmé ministerstvo strojírenství, které bylo později přejmenováno na Ministerstvo letectví a kosmonautiky. V roce 1978 se začal věnovat souvislostem mezi teorií řízení a demografií.

Obrázek 25.1: Song Jian

Abychom pochopili kontext Song Jianovi práce o populační dynamice, musíme nejdříve pochopit, čím se zabývá „teorie řízení“. Jedná se o studium dynamických systémů, jejichž chování závisí na některých parametrech, které

[1] Song je příjmení. V čínštině se píše vždy jako první.

lze v průběhu času měnit za účelem optimalizace daného kritéria. Tato teorie byla rozvinuta zejména v souvislosti s kosmickými programy v USA a v SSSR. Inženýři totiž museli „řídit“ trajektorii kosmických dopravních prostředků tak, aby se družice dostaly na oběžnou dráhu kolem Země. Aplikace se však neomezovaly pouze na fyzikální nebo inženýrské problémy. Politiku kontroly porodnosti lze také považovat za určitý druh problému optimálního řízení v matematickém smyslu.

Za zmínku stojí také esej s názvem *Meze růstu: zpráva pro projekt řešení tíživé situace lidstva zadaný římským klubem*, publikovaná v roce 1972 a napsaná skupinou z Massachusettského technologického institutu (M.I.T.). Tato studie byla založena na matematickém modelu světového hospodářského růstu, který zohledňoval přírodní zdroje, velikost populace a znečištění. Zpráva naznačovala, že světová ekonomika směřuje ke katastrofě v důsledku vyčerpání neobnovitelných zdrojů, nedostatku potravin pro obyvatelstvo nebo nadměrného znečištění. Jedním z navrhovaných řešení bylo dobrovolné omezení porodnosti. Souhrnně řečeno šlo o jakousi moderní verzi Malthusových tezí. Zpráva měla na západě v 70. letech 20. století velký ohlas.

Od založení Čínské lidové republiky v roce 1949 byla porodnost v Číně velmi vysoká s výjimkou období katastrofálního pětiletého plánu zvaného „Velký skok vpřed“. V polovině 70. let se Čína pomalu vzpamatovávala z kulturní revoluce. V rámci plánování rodiny byly ženy nabádány, aby odkládaly porody, prodlužovaly dobu mezi dvěma po sobě jdoucími porody a měly méně dětí. Teng Siao-pching, který se po smrti Mao Ce-tunga v roce 1976 stal novým vůdcem, zahájil v roce 1978 politiku „čtyř modernizací“: zemědělství, průmyslu, vědy a techniky a národní obrany. Velikost a růst čínské populace byly tehdy vnímány jako významné překážky těchto modernizací. Vědci, kteří do té doby pracovali na vojenských aplikacích, byli vyzváni, aby našli řešení tohoto obtížného problému.

Na základě těchto skutečností se Song Jian vydal v roce 1978 do Helsinek na kongres Mezinárodní federace pro automatické řízení (IFAC). Všiml si tam, že se někteří evropští vědci pokoušeli aplikovat teorii řízení na populační problémy s myšlenkou, že přísná kontrola porodnosti by nakonec mohla zabránit katastrofám zveřejněným ve zprávě *Meze růstu*. V Číně sestavil malý tým, jehož součástí byl i jeho kolega Yu Jingyuan a počítačový expert Li Guangyuan, aby tento typ matematických modelů aplikoval na data o čínské populaci. V té době byla vědecká komunikace mezi Čínou a zbytkem světa velmi omezená. Tým znovu odvodil rovnice popisující vývoj věkové struktury populace stejným způsobem, jakým to udělali Lotka a McKendrick (viz kapitoly 10 a 16). V modelu se spojitým časem zavedeme

- $P(x,t)$ populace ve věku x v čase t;

- $m(x)$ úmrtnost ve věku x;
- $P_0(x)$ věkovou strukturu populace v čase $t = 0$;
- $b(t)$ úhrnná plodnost žen v čase t, tj. průměrný počet dětí, které by žena měla během svého života, kdyby věkově specifická plodnost zůstala stejná jako v čase t;
- f podíl narozených dětí ženského pohlaví;
- $h(x)$ rozdělení pravděpodobnosti věku matky při narození dítěte:

$$\int_0^{+\infty} h(x)\,dx = 1.$$

Pomocí tohoto označení a příslušných předpokladů lze vývoj věkové struktury populace popsat pomocí parciální diferenciální rovnice

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

s počáteční podmínkou $P(x,0) = P_0(x)$ a okrajovou podmínkou

$$P(0,t) = b(t)\,f \int_0^{+\infty} h(x)\,P(x,t)\,dx\,,$$

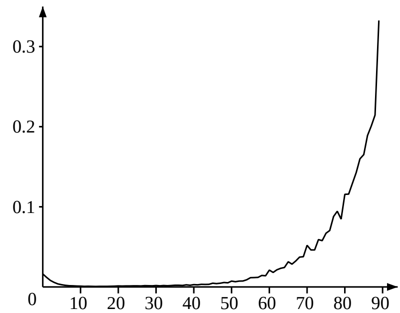
kde $b(t)$ je řídící parametr. Pokud je úhrnná plodnost žen konstantní a nad kritickou hranicí

$$b^* = 1/\left[f \int_0^{+\infty} h(x)\,e^{-\int_0^x m(y)\,dy}\,dx\right],$$

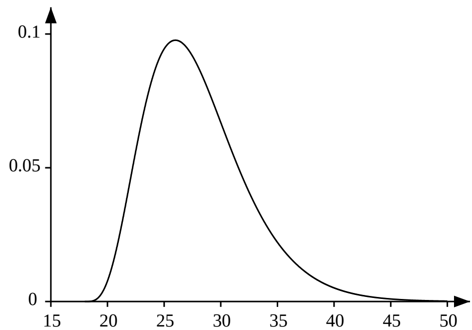
pak počet obyvatel exponenciálně roste. Toto kritérium se podobá tomu, které získal Lotka pomocí vzorce (10.2). Song Jianův tým uvažoval také časově diskrétní verzi modelu, která je podobná Leslieho modelu (viz kapitola 21). Nechť $P_{k,n}$ je populace ve věku k v roce n. Zavedeme podobně m_k, b_n a h_k. Pak

$$P_{k+1,n+1} = (1-m_k)\,P_{k,n}\,, \quad P_{0,n+1} = b_n\,f \sum_{k\geqslant 0} h_k\,P_{k,n}\,.$$

Díky výběrovým šetřením tým znal úmrtnost m_k (obrázek 25.2), podíl narozených žen $f \approx 0{,}487$, věkového rozložení matek h_k (obrázek 25.3), počáteční stavu $P_{k,0}$, což je věková struktura obyvatelstva v roce 1978 (obrázek 25.4) a měnící se úhrnnou plodnost b (předpokládané jako konstantní v průběhu každé simulace). Díky těmto znalostem mohl Song Jianův tým vytvořit demografické projekce pro svou zemi s časovým horizontem sto let, od roku 1980 do roku 2080 (obrázek 25.5). Vzhledem k množství potřebných operací sčítání a násobení k vytvoření projekce (roky jsou dány n mezi 0 a 100 lety, věk k mezi 0 a 90 lety) bylo nutné použít počítač. V té době měli v Číně k takovému zařízení přístup jen lidé pracující pro armádu. Song Jian, přední odborník na navádění raket, byl jedním z nich.



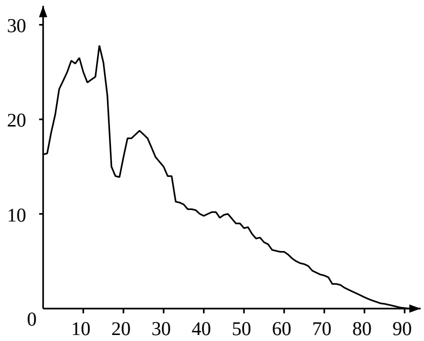
Obrázek 25.2: Úmrtnost (za rok) v závislosti na věku v roce 1978.



Obrázek 25.3: Vyhlazený tvar věkově specifické plodnosti (za rok) v roce 1978.

Projekce naznačují, že i kdyby Čína zachovala porodnost z roku 1978 na úrovni $b = 2{,}3$ dítěte na ženu, což je těsně nad kritickou hranicí, která se odhaduje na $b^* = 2{,}19$, počet obyvatel by se zvýšil z 980 milionů v roce

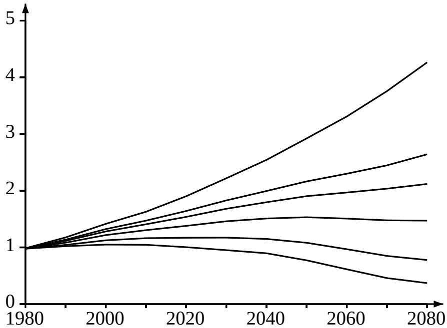
Obrázek 25.4: Věková pyramida v roce 1978. Vodorovná osa: věk. Svislá osa: počet obyvatel (v milionech).



1980 na 2,12 miliardy v roce 2080. Čína však již v té době využívala téměř veškerou půdu, která mohla sloužit pro zemědělství. Dokonce mohla o část této půdy přijít v důsledku desertifikace a urbanizace. Jak takovou populaci uživit, když pokrok ve výnosnosti zemědělských podniků není dostatečný? Je to stejná otázka, o které uvažoval Malthus o dvě století dříve. Při plodnosti $b = 3{,}0$ v roce 1975 by počet obyvatel mohl v roce 2080 dosáhnout dokonce 4,26 miliardy. Při hodnotě $b = 2{,}0$ by počet obyvatel dosáhl kolem roku 2050 svého maxima 1,53 miliardy a poté by začal mírně klesat. Při hodnotě $b = 1{,}5$ by svého maxima 1,17 miliardy dosáhl kolem roku 2030. Při hodnotě $b = 1{,}0$ by maximum činilo pouze 1,05 miliardy a bylo by dosaženo kolem roku 2000. Za tohoto předpokladu by se počet obyvatel vrátil na úroveň roku 1978 v roce 2025.

Nejpřekvapivější částí této práce byly její praktické důsledky, které v podstatě nemají v historii matematické populační dynamiky obdoby. Li Guangyuan skutečně předvedl výsledky simulací svého týmu v prosinci 1979 během sympozia o populaci v Čcheng-tu v provincii S'-čchuan[2]. V lednu 1980 Song Jian, Yu Jingyuan a Li Guangyuan publikovali tyto výsledky v čínském ekonomickém časopise a mimochodem obhajovali politiku jednoho dítěte. Svůj článek – *Zpráva o kvantitativním výzkumu v otázce populačního vývoje Číny* – zaslali také vrcholnému čínskému vědci Čchien Süe-senovi, který jej s doporučením předal vedoucímu správy plánování porodnosti. Výsledky Song Jianova týmu udělaly na většinu politických představitelů hluboký dojem. Ti již byli přesvědčeni o nutnosti zvýšené kontroly porodnosti navzdory tomu,

[2]Zde a níže shrnujeme podrobnou zprávu Susan Greenhalghové [1,2].



Obrázek 25.5: Demografické prognózy (v miliardách) podle různých hypotéz o průměrném počtu dětí na ženu. Odspodu nahoru: $b = 1{,}0$; 1,5; 2,0; 2,3; 2,5; 3,0.

co napsal Marx (viz kapitola 5), ale stále váhali nad mírou kontroly. V únoru 1980 stanovila Státní rada a ústřední výbor strany cíl pro maximální velikost čínské populace na 1,2 miliardy pro horizont roku 2000. V březnu 1980 byly výsledky Song Jianova týmu zveřejněny v *Lidovém deníku*. V dubnu komise složená z politických představitelů a odborníků na populaci zkoumala ekologické a ekonomické důsledky populačního růstu a dospěla k závěru, že k dosažení cíle stanoveného Teng Siao-pingem pro úhrný příjem na osobu v roce 2000 je nutná politika jednoho dítěte. Politika se stala oficiální v září téhož roku a otevřený dopis, který ji vysvětloval obyvatelstvu, byl zveřejněn na první straně *Lidového deníku*.

Do roku 1983 se ještě narodilo mnoho dětí bez povolení. Bylo rozhodnuto, že jeden člen každého páru, který již má dvě děti, bude sterilizován a každé zakázané těhotenství bude ukončeno. Od roku 1984 však bylo venkovským párům s jednou dcerou povoleno mít druhé dítě. Později byly zavedeny některé úpravy: pokud v páru byli muž i žena svobodní, mohli mít dvě děti. Represivní opatření proti párům, které měly více než jedno dítě, byla tvrdá: státní zaměstnanci mohli přijít o práci, bylo třeba zaplatit vysokou pokutu, aby rodiče získaly povolení od příslušných orgánů pro školní docházku druhého dítěte apod. Politika jednoho dítěte skončila v roce 2015. Souhrnně lze říci, že v historii matematického modelování až do roku 2020 lze jen těžko najít jiný příklad s tak silným společenským dopadem. Práce Song Jiana a jeho spolupracovníků byla samozřejmě jen jedním z důvodů, které vedly k rozhodnutí o politice jednoho dítěte. Zdá se však, že sehrála důležitou roli.

Stejně jako v předchozích kapitolách může být úloha matematického modelování předmětem obav. Na základě reálné situace se vytváří matematický model. Ten lze analyzovat nebo simulovat pomocí počítače. Díky tomu lze porozumět, jak se model chová při změně některých parametrů. Není však vždy možné ověřit, zda je model věrným obrazem skutečného života, protože některé jeho velmi důležité aspekty mohou být zanedbány. Některé modely obsahují také cílovou funkci, jako například udržení počtu obyvatel Číny pod 1,2 miliardami. Matematika již neříká, zda byl tento cíl vhodný[3].

V roce 1980 byl Song Jian také spoluautorem nového vydání knihy s názvem *Inženýrská kybernetika* od Čchiena Süe-sena, „otce" čínského vesmírného programu. Poté zastával různé vysoké politické funkce: náměstek ministra a hlavní vědecký inženýr ministerstva pro letectví a kosmonautiku (1981–1984), člen Ústředního výboru Komunistické strany Číny (1982–2002), předseda Státní komise pro vědu a techniku (1985–1998), státní rada (1986–1998) atd. Vydal také dvě další knihy, které byly přeloženy do angličtiny: *Řízení velikosti populace v Číně* (1985, spolu s Tuan Chi-Hsienem a Yu Jingyuanem) a *Systém řízení velikosti populace* (1988, s Yu Jingyuanem). Tyto knihy rozvíjejí teorii optimálního řízení aplikovanou na populační dynamiku. Song Jian byl v roce 1991 zvolen do Čínské akademie věd, v roce 1994 do Akademie inženýrů a v letech 1998 až 2002 byl jejím prezidentem.

Další čtení

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 (1988)

[3] Populace v roce 2000 byla odhadnuta na 1,264 miliardy. Příjem na obyvatele vzrostl v letech 1980 až 2000 přibližně z 200 na 1 000 dolarů. Současně se poměr pohlaví extrémně vychýlil ve prospěch chlapců, zejména kvůli častému umělému přerušení těhotenství u dívčích potomků.

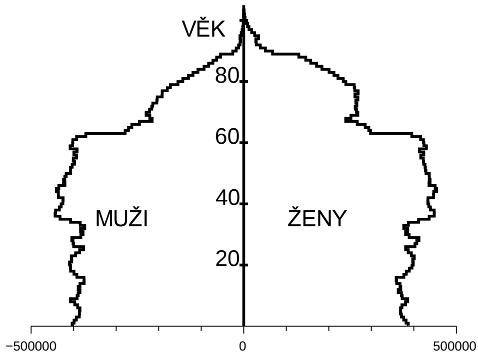
Kapitola 26

Některé současné problémy

Tato kapitola podává stručný přehled některých současných problémů matematické populační dynamiky: stárnutí populace v demografii; nově se objevující nemoci (AIDS, SARS, nemoci přenášené vektory...) a očkovací politika v epidemiologii; rybářská politika v ekologii; šíření geneticky modifikovaných organismů v populační genetice. Jsou zmíněny specializované instituce, které se ve Francii zabývají modelováním těchto problémů. Zdůrazněny jsou rovněž různé aspekty výzkumné práce.

V demografii se v posledních desetiletích objevil relativně nový problém: stárnutí populace. Tento problém je předmětem zájmu nejen ve Francii (obr. 26.1), ale i v mnoha dalších evropských zemích a také v Japonsku. Má významné ekonomické a sociální důsledky, které se dotýkají důchodových systémů, imigrační politiky apod. Matematické modely, které se snaží analyzovat stárnutí populace, vyvíjí ve Francii Národní institut demografických studií (INED) a Národní institut statistiky a ekonomických studií (INSEE). Jeden z problémů demografických prognóz spočívá v tom, že mnohé nelze předvídat ani na deset let dopředu, např. v čase se výrazně měnící porodnost. Když se podíváme na prognózy z roku 1968, které se týkaly francouzské populace v roce 1985, je to velice dobře patrné. Tyto prognózy nepředpokládaly pokles porodnosti, ke kterému došlo v 70. letech. Bylo by zajímavé přezkoumat všechny predikce založené na matematických modelech, které se ukázaly jako chybné, zejména ty medializované. To by vyvážilo dojem „pokroku“, který vyvolává tato kniha, dojem, který se čtenáři mohl zdát podezřelý již po přečtení kapitoly o čínské politice jednoho dítěte. Co se týče posledně jmenovaného tématu, je nyní aktuální nový problém: jak tuto politiku zmírnit, aby se zabránilo fenoménu rychlého stárnutí, který se očekává v příštích několika desetiletích. Do debaty opět přispívají matematické modely.

V celosvětovém měřítku (před pandemií covid-19) vyvstal v posledních třech desetiletích jako významný nový problém v epidemiologii vývoj epidemie AIDS. Některé modely se snaží odhadnout budoucí vývoj epidemie v nověji zasažených zemích, jako je Rusko, Indie nebo Čína. Je obtížné předpovědět, zda se epidemie zpomalí jako v západní Evropě a Severní Americe, nebo zda dosáhne významného procenta populace jako v některých subsaharských zemích. Pomocí matematických modelů byly zkoumány i další nově se

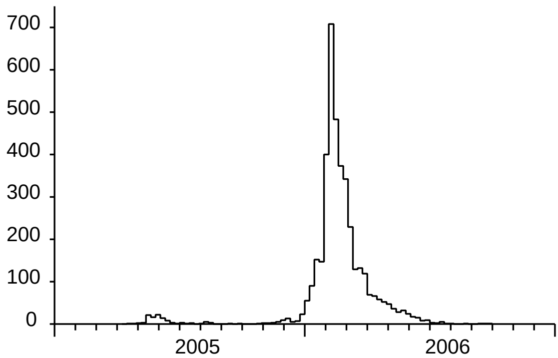


Obrázek 26.1: Věková pyramida francouzské populace k 1. lednu 2010. Zdroj: www.insee.fr.

objevující nemoci, jako je ebola v Africe, západonilská horečka v Severní Americe, SARS (těžký akutní respirační syndrom), ptačí chřipka, chikungunya nebo chřipka H1N1, ovšem popravdě s malým úspěchem.

V případě původní SARS byla jedna z potíží modelování skutečnost, že epidemie zůstávala relativně omezená v rámci jednotlivých zemí, ale mohla se velmi rychle šířit ze země do země (Hongkong a Čína, Singapur, Kanada...). Náhodný charakter epidemických křivek v každém novém ohnisku nebylo možné přehlédnout. Jak jsme viděli v kapitolách 16 a 22, stochastické modely jsou obvykle obtížněji zpracovatelné.

V případě epidemie chikungunya, která se vyskytla v letech 2005–2006 na ostrově Réunion (francouzské zámořské území v Indickém oceánu), se některé modely inspirovaly Rossovým modelem pro malárii (viz kapitola 12), protože obě nemoci přenášejí komáři. Bylo ale třeba vzít v úvahu vliv sezónnosti. Během jižní zimy se totiž populace komárů zmenšuje, čímž se přenos nemoci omezuje. To je patrné z obrázku 26.2, který ukazuje počet nových případů hlášených každý týden malou sítí asi třiceti praktických lékařů, kteří pokrývají jen zlomek populace ostrova. Během několika týdnů v září a říjnu 2005 síť nezaznamenala žádné nové případy, ale přenos nemoci stále pokračoval. Matematické modely epidemie byly vypracovány v Národním ústavu pro zdraví a lékařský výzkum (INSERM) a v Institutu tropického výzkumu (IRD). Navzdory těmto modelům nikdo nedokázal předvídat, že epidemie nevyhasne před koncem jižní zimy roku 2005, kdy se jí nakazilo jen několik tisíc lidí. Nakonec se nakazila téměř třetina obyvatel ostrova, tedy asi 266 000 lidí. To ukazuje, že předpovídat budoucnost epidemií může být poměrně obtížné a že není tak snadné rozlišit v prvních dnech epidemie, zda půjde o epidemii menšího nebo většího rozsahu. Paralelu lze vést s předpovědí počasí. Tento druh předpovědi se dnes opírá o intenzivní počítačové simulace složitých matematických modelů oceánu a atmosféry. Nicméně předpovědi delší než několik dní nejsou spolehlivé.



Obrázek 26.2: Epidemie chikungunya na ostrově Réunion v letech 2005–2006. Počet nových případů hlášených týdně malou sítí lékařů v závislosti na čase. Prvního malého vrcholu bylo dosaženo v květnu 2005, druhého velkého vrcholu v únoru 2006. Čísla na tomto obrázku je třeba vynásobit přibližně 67, abychom získali skutečný rozsah epidemie. Zdroj: www.invs.sante.fr.

Z teoretického hlediska vyvolala epidemie chikungunya otázku, jak upravit pojem základního reprodukčního čísla $\mathscr{R}_0$ v modelech, které předpokládají, že prostředí má sezónní (např. periodické) fluktuace. Přizpůsobení není tak jednoduché, což vyvolává určité obavy, jak byl parametr $\mathscr{R}_0$ použit u jiných epidemií ovlivněných sezónností, jako je pandemie chřipky H1N1 v roce 2009.

Dalším problémem, který vzbuzuje stále větší obavy a který se modeláři pokoušeli analyzovat, je rezistence vůči lékům (antibiotikům, antimalarikům...). V epidemiologii je to stále opakující se otázka od dob Daniela Bernoulliho a d'Alemberta: jak vyvážit náklady a přínosy, když injekce vakcíny s sebou nese potenciální riziko. To je stále předmětem sporů a možná jím i zůstane, protože citlivost na riziko se mění. To byl také důvod, proč náznaky, že by vakcína proti hepatitidě B mohla ve velmi malém počtu případů způsobit závažné komplikace, vedly francouzské ministerstvo zdravotnictví v roce

1998 k zastavení očkovací kampaně ve školách, přestože se riziko jevilo jako zanedbatelné ve srovnání s rizikem úmrtí po nákaze virem hepatitidy B.

Další oblasti ke studiu se objevují v ekologii při analýze dynamiky rybích populací. Ty mají sloužit jako vědecký podklad pro tvorbu rybolovných kvót a dalších omezení. Uveďme za všechny dva nedávné příklady, kterými jsou nadměrný lov sardelí v Biskajském zálivu nebo tuňáka červeného ve Středozemním moři. Vzhledem k tomu, že odhady velikosti rybí populace jsou často nespolehlivé, je třeba s modely používajícími tyto údaje posuzovat obezřetně. Ve Francii se tímto typem studií zabývá především Výzkumný ústav pro využívání moří (IFREMER). Některé matematické modely sehrály roli i v minulých rozhodnutích Mezinárodní velrybářské komise.

V populační genetice je předmětem sporů také šíření geneticky modifikovaných organismů, které se někteří výzkumníci pokusili studovat pomocí „reakčně-difúzních“ modelů inspirovaných Fisherovým modelem (viz kapitola 20). Touto oblastí se zabývá Národní institut pro výzkum v agronomii (INRAE).

Z teoretičtější stránky výzkumu lze zmínit např:

- práce o parciálních diferenciálních rovnicích, jako jsou difúzní rovnice (viz kapitola 20) nebo věkově strukturované rovnice (viz kapitola 16);

- práce o stochastických modelech s prostorovou složkou nebo bez ní (viz kapitoly 16 a 22), včetně prací o náhodných sítích modelujících šíření epidemií a prací hledajících deterministické aproximace.

Tento typ výzkumu provádějí především aplikovaní matematici. V posledních letech bylo na francouzských univerzitách a dalších vysokých školách (stejně jako na českých) zavedeno několik magisterských programů v oboru matematické biologie.

Stejně jako jiné vědní obory, i matematické studium populační dynamiky rozšiřuje poznatky především prostřednictvím:

- „učených společností“: Nizozemská společnost pro teoretickou biologii (od roku 1970), Společnost pro matematickou biologii (1973), Frankofonní společnost pro teoretickou biologii (1985), Čínská společnost pro matematickou biologii (1985), Japonská společnost pro matematickou biologii (1989), Evropská společnost pro matematickou a teoretickou biologii (1991), Latinskoamerická společnost pro matematickou biologii (2002) atd.

- specializovaných odborných časopisů: *Acta Biotheoretica* (od roku 1935), *Bulletin of Mathematical Biology* (1939), *Mathematical Bio-*

sciences (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) atd.

- konferencí: výroční zasedání Společnosti pro matematickou biologii, Výpočetní a matematická dynamika populace, Evropská konference o matematické a teoretické biologii atd.

V každé konkrétní oblasti (demografie, ekologie, populační genetika, epidemiologie atd.) lze najít další podobné společnosti, časopisy či konference s různým podílem matematického modelování. Zde jsou uvedeny pouze ty, které se výslovně hlásí k rozhraní mezi matematikou a jejími aplikacemi na populační dynamiku.

Na závěr bychom doporučili čtenáři, aby se sám podíval na původní články, které jsou k dispozici na internetu. Jejich adresy jsou uvedeny v odkazech na konci každé kapitoly. Jak kdysi napsal Ronald Fisher o Mendelovi:

> „Dějiny vědy značně utrpěly tím, že učitelé používali přepsané materiály a následně se vytratily okolnosti a intelektuální atmosféra, v níž byly učiněny velké objevy minulosti. Studium původních zdrojů je vždy poučné a často plné překvapení.“

Další čtení

1. Bacaër, N.: Approximation of the basic reproduction number $\mathscr{R}_0$ for vector-borne diseases with a periodic vector population. *Bull. Math. Biol.* 69, 1067–1091 (2007)
2. Levin, S.A.: Mathematics and biology, the interface. www.bio.vu.nl/nvtb/

Kapitola 27

Postscriptum: Matematické modelování covid-19 v Evropě[1]

Český překlad této knihy vznikl na počátku roku 2022, kdy již druhým rokem probíhala pandemie covid-19. Toto období je z hlediska rozvoje matematického modelování epidemií velmi významné, a protože se mnozí čeští vědci pomáhající s překladem této knihy do tohoto překotného vývoje zapojili, rozhodli se společně s autorem Nicolasem Bacaërem podělit o své zkušenosti z tohoto období v této kapitole.

Mnoho matematických biologů pandemie covid-19 pravděpodobně zaskočila. Snad ani ne tím, že přišla, ale nepochybně tím, jak obrovskému tlaku na tvorbu modelů, scénářů a predikcí budoucího vývoje byli najednou vystaveni. Ze svých temných kanceláří se dostali do světel ramp, kterými byli oslněni, většinou pozitivně. Výsledkem byl překotný vývoj nových metod a přístupů, využití známých modelů i postupné pochopení a učení se z nedostatků či různých chyb, které se v modelech objevují a mají vliv na to, jak realistické jejich výstupy jsou.

Matematické modely se v epidemiologii používají více než století. Základní kompartmentový epidemiologický model typu SIR je popsán v kapitole 16. Tento mechanistický přístup k modelování dynamiky epidemie se stal jádrem téměř všech pozdějších modelů, obecných i těch vyvinutých pro specifické infekce, ať už šlo šíření viru HIV či v předchozí kapitole 26 zmíněné modely nemocí SARS, eboly a chikungunya. Specifikem řady infekcí, včetně viru SARS-CoV-2 a následné nemoci covid-19, je infekčnost nakaženého jedince už před objevením se symptomů a dokonce infekčnost (byť pravděpodobně snížená) zcela asymptomatických jedinců. Právě z tohoto důvodu začali modeláři pracovat s modely odvozenými z modelu SIR (resp. SEIR s kompartmentem viru exponovaných osob E), které obsahovaly další kompartmenty osob detekovaných, asymptomatických, presymptomatických, hospitalizovaných, izolovaných apod. Podobně se pro covid-19 ukázalo jako velice podstatné model věkově strukturovat, neboť jeho klinická závažnost zcela evidentně závisela na věku [Levin, 2020]. Vznikly nové přístupy k od-

[1] Postscriptum napsala Lenka Přibylová a Veronika Eclerová.

hadům parametrů z reálných dat sbíraných v reálném čase z nemocnic i laboratoří. Právě sběr dat v reálném čase je pro kvalitní odhady budoucího vývoje epidemie klíčový. V České republice (Česku) sběr dat zajišťuje Ústav zdravotnických informací a statistiky České republiky přes Informační systém infekční nemoci ISIN a Dispečink intenzivní péče DIP.

Výkonné počítače umožnily vývoj a užití multiagentních modelů, které umožňují simulovat probíhající epidemii na skupině agentů (jedinců), jejichž stav je unikátní a určený pravděpodobnostmi přechodu potenciálně závislými na mnoha proměnných. Výhodou tohoto přístupu je možnost pracovat s kontaktní sítí, tj. vytvořit kromě množiny stavů (vnímavý S, exponovaný E, infekční detekovaný I, apod.) také spojnice mezi jednotlivými agenty a síť kontaktů tak simulovat těmito spojnicemi (hranami grafu) a jejich intenzitou. To umožnilo simulovat a analyzovat problémy, na které standardní kompartmentové modely nestačí, např. rodinné kontakty a mezigenerační přenos nemoci, izolaci části klastrů v okolí nakažené populace, detailní trasování kontaktů, varianty harmonogramu prezenční/distanční výuky ve školách v průběhu epidemie apod. Obecnou nevýhodou těchto modelů je nutnost provedení tisíců paralelních a časově náročných opakování stejných simulací a vysoká citlivost na řadu předpokladů. Skupina BISOP, jeden z důležitých hráčů na poli modelování epidemie covid-19 v Česku, takový multiagentní model vytvořila: obsahuje 30 vrstev kontaktů (rodina, škola, doprava, apod.) a přibližně 56 tisíc agentů inspirovaných jednou reginální oblastí, přičemž sama kontaktní síť je složena z kolem 2,8 milionů hran. Spuštění jedné simulace v délce 300 dní (i na rychlém počítači s 32 procesory Intel(R) Xeon(R) Gold 6226R CPU 2.90GHz) přitom trvá tisíc opakování až hodinu.

Přestože pro žádnou jinou infekci v historii nebylo v tak krátkém čase sesbíráno tolik datových sad a provedeno tolik longitudinálních studií, jsou tato data stále neúplná. Některé nedostatky jsou dány principiálně, např. v záznamech nejsou a ani nemohou být všechny nákazy proto, že se virus šíří v asymptomatické a presymptomatické době a není možné testovat a trasovat celou populaci. Neznámou proměnnou je proto i míra detekce (ascertainment rate). Chyby vznikají i při testování samotném, neboť různé testy mají různou sensitivitu a specificitu, mohou být proto falešně negativní i falešně pozitivní. Jde o efekt podmíněné pravděpodobnosti, který můžeme matematicky popsat pomocí Bayesova vzorce[2]. Jak ukazuje následující box, v případě nízké prevalence nemocných v populaci je antigenní testování nevhodné, protože vede k vysokému procentu falešně pozitivních případů. V takovém případě je vhodné provádět konfirmaci PCR testem.

[2]Koncept úplné a podmíněné pravděpodobnosti je známý již minimálně od 17. století, kdy se jím zabýval anglický duchovní Thomas Bayes.

Zkoumejme účinnost plošného antigenního testování ve dvou případech: (i) prevalence nemoci v populaci je 10 %, (ii) prevalence nemoci v populaci je 1 %. V obou případech předpokládejme, že je senzitivita antigenních testů (tj. procento nemocných, kteří budou pozitivně testováni) 95 %, a specificita těchto testů (tj. procento zdravých, kteří budou negativitě testováni) 98 %.

(i) V případě 10 % prevalence onemocnění v populaci je rozložení výsledků testů následující:

	Nemocný	Zdravý	Celkem
Pozitivní	9,50 %	1,80 %	11,30 %
Negativní	0,50 %	88,20 %	88,70 %
Celkem	10,00 %	90,00 %	100,00 %

Falešná negativita testů (tj. procento negativně testovaných, kteří jsou ve skutečnosti nemocní) je 0,56 % a falešná pozitivita testů (tj. procento pozitivně testovaných, kteří jsou ve skutečnosti zdraví) je 15,93 %.

(ii) V případě 1 % prevalence onemocnění v populaci jsou výsledky následující:

	Nemocný	Zdravý	Celkem
Pozitivní	0,95 %	1,98 %	2,93 %
Negativní	0,05 %	97,02 %	97,07 %
Celkem	1,00 %	99,00 %	100,00 %

Falešná negativita testů je 0,05 % a falešná pozitivita testů je 67,58 %.

Dokonce i testování metodou PCR (polymerázová řetězová reakce), která je tzv. zlatým standardem, může být zdrojem chyby, je-li neinfekční osoba označena za pozitivní kvůli fragmentům virové RNA ve vzorku. A k chybám pochopitelně dochází i při samotném sběru dat. Může jít o náhodné chyby jednotlivců při zápisu dat do systému, chyby při odběru, změny metodiky či systematická podhodnocení/nadhodnocení vlivem epidemie samotné. Uvedu několik příkladů. Náhodné chyby při zápisu se stanou a pravděpodobně nebudou mít žádný vychylující vliv, ale někdy může dojít k významně vychylující chybě. V létě 2021 se v celé Evropě zavedly certifikáty a dřívější prodělání covidu umožňovalo vyjet na dovolenou. Díky tomu se najednou velké množství lidí obrátilo na laboratoře, aby jim dodatečně potvrdili dříve provedené testy, které se chybou nezapsaly do systému. Tento výkyv v datech vypadal jako rozhoření epidemie. Systematická chyba se projevuje i na datech zveřejňovaných WHO pro jednotlivé země, kdy denní přírůstky hospitalizací a úmrtí se odečítají z denních kumulativních hodnot a neopravují se zpětně, přičemž samotná zatížení nemocnic epidemií ovlivňují zápis hospitalizací a úmrtí v reálném čase a zpětné zápisy a opravy se na národní úrovni běžně provádějí.

Dosud modelované epidemie probíhaly většinou přirozeně. Některé dříve vyvinuté modely sice zjednodušeně pracovaly např. s vakcinací nebo sezónností viru, ale jinak popisovaly vcelku přirozený vývoj epidemie. Epidemie covid-19 však byla výrazně ovlivněna neustálými změnami lidského chování a tedy počtu epidemiologicky významných kontaktů. Tyto změny byly vyvolány zavedením či rušením plošných nefarmaceutických opatření, omezeními pohybu, nošením roušek či respirátorů, způsobem trasování, vývojem vakcinace, ale také strachem lidí z infekce samotné. Modely na tuto situaci reagovaly zahrnutím kontaktní struktury populace, mobility, reálných dat o průběhu vakcinace, či vícevariantními verzemi zahrnujícími nástup a dynamiku nových variant. Proto je také v každé zemi zcela odlišný a unikátní průběh epidemie, liší se například hrubá míra smrtnosti CFR[3] nebo zatížení nemocniční péče. Tyto rozdíly, včetně například rozdílné věkové struktury obyvatelstva či míry a kvality sociální péče v různých státech světa, snadno vedou k dezinterpretacím. Zatímco například v Africe je dle údajů z roku 2019 podíl osob starších 65 let menší než 5 %, v Evropě tvoří tato věková skupina téměř 20 % populace.

Přes všechny tyto problémy, nedostatky a odlišnosti byla tvorba scénářů vývoje epidemie či krátkodobých predikcí založených na dostupných datech v době pandemie covid-19 nezbytná. Společenská poptávka tlačila na matematické modeláře, aby v reálném čase probíhající epidemie vytvářeli modely, analyzovali data a poskytovali nástin budoucího vývoje. Termín rychlá věda (*rapid science*) nabral zcela jinou dimenzi. Během prvních 4 měsíců pandemie bylo publikováno téměř 20 000 různých vědeckých článků o covid-19, z nichž třetina jako preprinty. K tomu statisíce mediálních výstupů snažících se reagovat na otázky politiků a veřejnosti. Často se studie a statistiky zpracovávaly po nocích a jejich výsledky byly zveřejňovány jako tiskové zprávy. Takovým případem byla i britská klinická studie Recovery testující léčbu covid-19 pomocí levného steroidu dexametazon. Ten pravděpodobně snížil počet úmrtí u kriticky nemocných pacientů na podpoře dýchání až o jednu třetinu. Když epidemiolog z Oxfordské univerzity, sir Peter William Horby, jednoho pátečního večera uviděl výsledek studie, cítil prý směs nadšení a úz-

[3]Zatímco CFR (*case fatality ratio*) je podíl úmrtí po diagnostikované nákaze, a závisí tedy na míře detekce, IFR (*infection fatality ratio*) je podíl úmrtí po nákaze, tedy i té nedetekované.

kosti. Nadšení z toho, že se mu podařilo najít život zachraňující léčbu covid-19 a úzkost z toho, zda je výsledek správný. Měl obavu zveřejnit výsledek, který by mohl být špatný nebo zavádějící. Statistici pracovali celý víkend, aby třikrát zkontrolovali všechny analýzy a hledali v datech případné díry. Teprve poté, co si Horby a jeho spoluřešitel Martin Landray byli jistí, že je výsledek spolehlivý, podělili se o něj se světem v úterý 16. června 2020 v tiskové zprávě. Ještě ten den se lék začal používat v celé Británii a celosvětově do konce roku 2021 zachránil odhadem jeden milion životů.

Oproti tomuto úspěchu stojí případy, kdy nerecenzované preprinty umožnily publikaci prací, které by v dané formě publikovány nebyly, pokud by prošly kritickou oponenturou. Takovým případem je například kauza studií o ivermektinu, což je antiparazitikum, které opravdu v mnohých studiích vykazovalo účinnost proti těžkému průběhu covid-19. Později se však ukázalo, že šlo o oblasti s vysokým výskytem parazitů, kdy nasazení léku pomohlo zlepšit imunitu velké části pacientů a účinek proti covid-19 tak byl nepřímý.

Obrovským pozitivem ve vývoji matematických modelů a vědy obecně bylo bezprecedentně rychlé sdílení informací a dat. To umožnilo výzkumným pracovníkům zlepšit a zrychlit práci, získat zpětnou vazbu od ostatních odborníků a vytvořit mezioborové kontakty. Vznikly platformy pro sdílení dat, výsledků i predikcí modelů (např. platforma Evropského střediska pro prevenci a kontrolu nemocí ECDC Forecast Hub). K tomu přispěly i sociální sítě, takže se nedá obecně říct, že během pandemie plnily jen negativní roli ve vytváření sociálních bublin, šíření dezinformací a zvýšení polarizace společnosti, ačkoli toto bylo a zůstává velkým problémem společnosti.

Evropští modeláři se na pravidelných seminářích ECDC Forecast Hubu informovali o metodách, které používají pro predikční modely. Zmiňme zde jen některé: bayesovské průměrování modelu (*Bayesian model averaging*), optimalizace na úmrtí, hospitalizace, na odhad exponovaných osob, velikost zasažených klastrů, metody citlivostní analýzy apod. Matematické modely publikované na ECDC Forecast Hubu jsou agregovány na základě optimalizace do společného modelu, který vykazuje velmi dobré predikční schopnosti. Zároveň je publikována retrospektivní evaluace všech zveřejňovaných modelů, což umožňuje jejich dlouhodobé porovnání mezi sebou i s realitou.

Jistě bude potřeba modely, výsledky, metody, i publikované články podrobit důkladné retrospektivní analýze, ale není pochyb o tom, že se matematické modelování epidemií během pár let pandemie covid-19 posunulo obrovsky kupředu.

Další čtení

1. Levin, A.T., Hanage, W.P., Owusu-Boaitey, N., Cochran, K.B., Walsh, S.P., Meyerowitz-Katz, G.: Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications. *Eur J Epidemiol.* (2020) Dec; 35(12):1123–1138.
2. Watson, C.: Rise of the preprint: how rapid data sharing during COVID-19 has changed science forever. *Nature Medicine* (2022)
3. ECDC Forecast Hub: https://covid19forecasthub.eu

Zdroje obrázků

- S. 5. Portrét Thomase Murraye (asi 1687) ve vlastnictví Královské společnosti v Londýně. Chapman, S.: Edmond Halley, F.R.S. 1656–1742. *Notes Rec. R. Soc. Lond.* 12, 168–174 (1957) © The Royal Society.
- S. 11. Portrét od Emanuela Handmanna (1753) ve vlastnictví Muzea umění v Basileji. *Leonhard Euler 1707–1783, Beiträge zu Leben und Werk*. Birkhäuser, Basel (1983)
- S. 17. Portrét z kostela sv. Petra (Petri-Kirche), který byl pravděpodobně zničen během bitvy o Berlín v roce 1945. Reimer, K.F.: Johann Peter Süßmilch, seine Abstammung und Biographie. *Arch. soz. Hyg. Demogr.* 7, 20–28 (1932)
- S. 23. Portrét Johanna Niclause Grootha (asi 1750–1755) ve vlastnictví Přírodovědeckého muzea v Basileji. Speiser, D.: *Die Werke von Daniel Bernoulli*, Band 2. Birkhäuser, Basel (1982)
- S. 30. Portrét Maurice Quentina de La Tour (1753) ve vlastnictví muzea Louvre v Paříži.
- S. 33. Portrét od Johna Linnella (1833) v majetku *Haileybury College*, Anglie. Habakkuk, H.J.: Robert Malthus, F.R.S. (1766–1834). *Notes Rec. R. Soc. Lond.* 14, 99–108 (1959)
- S. 37. Rytina od Flamenga (1850). Quetelet, A.: Pierre-François Verhulst. *Annu. Acad. R. Sci. Lett. B.-Arts Belg.* 16, 97–124 (1850)
- S. 42. Heyde, C.C., Seneta, E.: *I. J. Bienaymé, Statistical Theory Anticipated*. Springer (1977) © Académie des sciences, Institut de France.
- S. 42. Brun, J., Robinet, A. (éd.): *A. Cournot, études pour le centenaire de sa mort*. Economica / Vrin, Paris (1978)
- S. 46. Bateson, W.: *Mendel's Principles of Heredity*. Cambridge University Press (1913)
- S. 51. Pearson, K.: *The Life, Letters, and Labors of Francis Galton*, vol. 1. Cambridge University Press (1914)
- S. 51. Watsonův portrét v knihovně *Trinity College* na univerzitě v Cambridge. Kendall, D.G.: Branching processes since 1873. *J. Lond. Math. Soc.* 41, 385–406 (1966)
- S. 57. Alfred J. Lotka Papers. Public Policy Papers. Odbor vzácných knih a zvláštních sbírek. © Princeton University Library.
- S. 61. Titchmarsh, E. C.: Godfrey Harold Hardy 1877–1947. *Obit. Not. Fellows R. Soc.* 6, 446–461 (1949)
- S. 64. Stern, C.: Wilhelm Weinberg. *Genetics* 47, 1–5 (1962)
- S. 67. G.H.F.N.: Sir Ronald Ross 1857–1932. *Obit. Not. Fellows R. Soc.* 1,

108–115 (1933) © The Royal Society.

- S. 75. Whittaker, E.T.: Vito Volterra 1860–1940. *Obit. Not. Fellows R. Soc.* 3, 690–729 (1941)
- S. 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.
- S. 84. Yates, F.: George Udny Yule. *Obit. Not. Fellows R. Soc.* 8, 308–323 (1952)
- S. 92. Heyde, C.C., Seneta, E. (eds.): *Statisticians of the Centuries*. Springer (2001)
- S. 101. britannica.com/EBchecked/topic/252257/J-B-S-Haldane © Bassano and Vandyk Studios.
- S. 109. Hill, W.G.: Sewall Wright, 21 December 1889 – 3 March 1988. *Biog. Mem. Fellows R. Soc.* 36, 568–579 (1990) © Llewellyn Studios, Chicago.
- S. 104. Nybølle, H.C.: Agner Krarup Erlang f. 1. Januar 1878 – d. 3. Februar 1929. *Mat. Tidsskr. B*, 32–36 (1929)
- S. 104. Nørdlund, N.E.: Johan Frederik Steffensen in memoriam. *Nordisk Mat. Tidsskr.* 10, 105–107 (1962)
- S. 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)
- S. 118. *I. G. Petrowsky Selected Works Part I*. Gordon and Breach, Amsterdam (1996) © Taylor and Francis Books UK.
- S. 122. Foto: Denys Kempson. Crowcroft, P.: *Elton's Ecologists, a History of the Bureau of Animal Population*. University of Chicago Press (1991)
- S. 126. © Geoffrey Grimmett.
- S. 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.
- S. 133. Harman, O.: *The Price of Altruism*. W. W. Norton, London (2010)
- S. 138. © Samuel Schlaefli / ETH Zürich.
- S. 146. Selected works of J. Song. *Science Press*, Beijing (1999) © Song Jian.

Obsah

Předmluva iii

1 Fibonacciho posloupnost (1202) 1

2 Halleyova úmrtnostní tabulka (1693) 4

3 Euler, Süssmilch a boží řád (1748–1761) 10

4 Daniel Bernoulli a očkování proti neštovicím (1760) 22

5 Malthus a potíže s geometrickým růstem (1798) 33

6 Verhulst a logistická rovnice (1838) 37

7 Bienaymé, Cournot a zánik příjmení (1845–1847) 42

8 Mendel a dědičnost (1865) 46

9 Galton, Watson a problém vymírání (1873–1875) 50

10 Lotka a teorie stabilizované populace (1907–1911) 57

11 Hardyho-Weinbergův zákon (1908) 61

12 Ross a malárie (1911) 66

13 Lotka, Volterra a matematická teorie boje o život (1920–1926) 72

14 Fisher a přírodní výběr (1922) 79

15 Yule a evoluce (1924) 83

16 McKendrick a Kermack o modelování epidemií (1926–1927) 91

17 Haldane a mutace (1927) 100

18 Erlang a Steffensen o problému vymírání (1929–1933) 104

19 Wright a náhodný genetický drift (1931) 109

20 Šíření genů (1937) 115

21 Leslieho matice (1945) 122

22 Perkolace a epidemie (1957) 126

23 Teorie her a evoluce (1973) 132

24 Chaotické populace (1974) 138

25 Čínská politika jednoho dítěte (1980) 146

26 Některé současné problémy 153

27 Postscriptum: Matematické modelování covid-19 v Evropě 158

Zdroje obrázků 164

Tato kniha sleduje historii populační dynamiky – teoretického oboru úzce souvisejícího s genetikou, ekologií, epidemiologií a demografií – do kterého matematika přinesla významné vhledy. Sleduje vznik některých důležitých témat: exponenciální růst od Eulera a Malthuse až po čínskou politiku jednoho dítěte; vývoj stochastických modelů od Mendelových zákonů a otázky vymírání příjmení po šíření epidemií; prolínání nutnosti a náhody v chaotické populaci.

S nedávnými pokroky ve strojovém překladu ztrácí faktický monopol jednoho jazyka ve vědecké literatuře své opodstatnění. Rostoucí jazykové odcizení na univerzitách lze zvrátit. Tímto pečlivě revidovaným českým překladem tuto novou cestu podporujeme.

ISBN : 979-10-343-9751-8

9 791034 397518

15€

N. Bacaër, L. Přibylová, V. Eclerová a kol. Stručná historie matematické populační dynamiky