Николя Бакаэр

Редакторы перевода :

В.А. Вольперт, Д.М. Эдиев

Краткая история математической динамики населения

Краткая история математической динамики населения

Николя Бакаэр

Редакторы перевода : В.А. Вольперт, Д.М. Эдиев

Николя Бакаэр (Nicolas Bacaër)
Institut de recherche pour le développement
nicolas.bacaer@ird.fr

В.А. Вольперт
Centre national de la recherche scientifique
Université Claude-Bernard-Lyon-I
volpert@math.univ-lyon1.fr

Д.М. Эдиев
Северо-Кавказская государственная академия
ediev@ncsa.ru

Желающие приобрести бумажную версию этой книги могут отправить сообщение на nicolas.bacaer@ird.fr.

Фото на обложке: Кандинский, *Gelb-Rot-Blau*, 1925. © Centre Pompidou.

Titre original :
Histoires de mathématiques et de populations

Pour l'édition russe :

ISBN : 979-10-343-8016-9
Dépôt légal : juillet 2021

Введение

Динамика народонаселения - это область науки, которая изучает закономерности изменения численности и состава биологических популяций, таких как люди, животные, растения или микроорганизмы. Она тесно связана со статистикой народонаселения, более описательной областью, но все же довольно сильно отличается от нее. Один общий момент заключается в том, что обе эти области широко используют математический язык.

Динамика народонаселения находится на стыке различных областей: математики, социальных наук (демография), биологии (генетика и экология народонаселения) и медицины (эпидемиология). Но она не часто рассматривается в совокупности подходов, несмотря на сходство проблем, встречающихся в различных областях применения. Заметным исключением является книга на французском языке *Математические теории народонаселения*[1] Алена Хиллиона. Но в ней проблема рассматривается с точки зрения математика, выделяя различные типы моделей: модели дискретного времени ($t = 0, 1, 2...$) и континуум-временные модели (t - вещественное число), детерминистические модели (будущие состояния точно известны, если известно настоящее состояние) и стохастические модели (где играет роль вероятность). В книге рассматриваются логически дискретные детерминистические модели, непрерывные детерминистические модели, дискретные стохастические модели и непрерывные стохастические модели.

В настоящей книге я попытался представить проблему динамики народонаселения с исторической точки зрения. Исследования объясняются в их контексте и приводятся краткие биографии ученых. Это должно облегчить чтение книги для тех, кто меньше знаком с математикой, и помочь в понимании происхождения исследуемых проблем. Но эта книга не только об истории. Она также может служить введением в математическое моделирование. Представляется важным включить в нее подробности большинства вычислений, чтобы читатель мог действительно увидеть ограничения, накладываемые на модели. Технические детали выделены в рамках и могут быть пропущены при первом чтении. Последняя глава посвящена многочисленным современным проблемам в динамике

---
[1] *Presses Universitaires de France*, Париж, 1986.

народонаселения, которые можно попытаться проанализировать с математической точки зрения. Для тех, кто хочет узнать больше, списки ссылок в конце каждой главы также включают веб-сайты, с которых можно загрузить оригинальные статьи.

В книге такого объема не было возможности дать полную картину всех исследований, проведенных до настоящего времени, и рассказать обо всех ученых, которые внесли свой вклад в эту тему. Сделанный выбор с необходимостью содержит произвольную составляющую, особенно относительно работ, сделанных в последние десятилетия. Тем не менее, я надеюсь, что выбранные работы будет достаточно репрезентативными, и что исследователи, активно работающие в этой области, чьи работы не упомянуты, не будут задеты.

Идеальной аудиторией для этой книги могут быть:

- Старшеклассники и студенты вузов, задающиеся вопросом, какие связи могут существовать между курсами математики, которые они посещают, и окружающим миром, или студенты, готовящие работу по теме, связанной с динамикой народонаселения.
- Преподаватели математики, пытающиеся сделать свой курс более привлекательным. Знания четырех элементарных математических операций достаточно для понимания большинства глав 1, 2 и 5. Глава 3 может служить введением в приложения логарифмов. В этой книге также рассматриваются: уравнения повторения в главах 1, 3, 8, 11, 14, 21, 23, 24; дифференциальные уравнения в главах 4, 6, 12, 13, 16; дифференциальные уравнения в частных производных в главах 20, 25; интегральное уравнение в главе 10; и применение теории вероятности в главах 2, 7, 8, 9, 15, 16, 17, 18, 19, 22.
- Специалисты, уже знакомые с демографией, эпидемиологией, генетикой или экологией и желающие сравнить свою область науки с другими, которые могут использовать аналогичные математические модели.
- Читатели, интересующиеся историей науки.

За редактирование русского перевода я очень благодарен Виталию Вольперту и Далхату Эдиеву, которые прочитали и откорректировали автоматический перевод программы DeepL. Я также выражаю свою благодарность Павлу Шевчуку, Йордану Стоянову, Ефиму Фрисману и Геннадию Розенбергу за их замечания.

Глава 1

Последовательность Фибоначчи (1202)

В 1202 году Леонардо из Пизы, также известный как Фибоначчи, опубликовал книгу, которая популяризировала в Европе индийскую десятичную систему исчисления, которая также была принята арабскими математиками. Среди многих примеров, приведенных в книге, один относится к росту популяции кроликов. Это один из старейших примеров математической модели динамики популяции.

Леонардо из Пизы, по имени Фибоначчи (Fibonacci), родился около 1170 года в Пизанской Республике, когда она находилась на пике своего коммерческого и военного могущества в средиземноморском мире. Около 1192 года отец Фибоначчи был отправлен республикой в Беджию, которая сейчас находится в Алжире, чтобы возглавить торговый пункт. Вскоре после этого к нему присоединился сын, который готовился стать купцом. Леонардо начал изучать десятичную систему исчисления, которую арабы привезли из Индии и которая используется и по сей день почти в той же форме: 0, 1, 2, 3, 4, 5, 6, 7, 8 и 9. Путешествуя по Средиземному морю, он сравнивал различные системы исчисления и изучал арабскую математику. Вернувшись в Пизу в 1202 году, он закончил написание книги на латинском языке под названием *Liber abaci* (Книга Расчетов), в которой он объяснил новую систему исчисления и показал, как использовать ее для бухгалтерского учета, конвертации веса и валюты, процентных ставок и многих других применений. Он также собрал большую часть результатов в алгебре и арифметике, известных арабам.

Фибоначчи рассмотрел в своей книге то, что сегодня можно назвать проблемой динамики народонаселения. Но она появилась просто как вычислительное упражнение среди других не связанных между собой тем: в предыдущем разделе книги речь идет об идеальных числах, представляющих собой сумму их факторов, например, $28 = 14 + 7 + 4 + 2 + 1$, а в следующем разделе речь идет о проблеме разделения денег между четырьмя людьми, что эквивалентно линейной системе из четырех уравнений. Вот перевод с латыни проблемы народонаселения:

> «Один человек держал пару кроликов в определенном закрытом месте. Требуется посчитать, сколько кроликов появилось от из этой пары за один год, если предполагается, что они за один месяц вынашивают другую пару, а за второй месяц рождаются еще и те, кто родился для вынашивания.»

Если в начале первого месяца появляется на свет пара новорожденных кроликов, то через месяц эта пара еще не будет способна к размножению и в начале второго месяца все равно останется только одна пара кроликов. Эта пара кроликов родит другую пару в начале третьего месяца, так что всего будет две пары. Первоначальная пара кроликов снова родит другую пару в начале четвертого месяца. Но вторая пара кроликов еще не будет фертильной. Будет всего три пары кроликов.

Используя современные обозначения, пусть P_n будет числом пар кроликов в начале месяца n. Количество пар кроликов P_{n+1} в месяц $n+1$ - это сумма количества пар кроликов P_n в месяц n и количества новорожденных пар в месяц $n+1$. Но только пары кроликов, которым исполнилось не менее двух месяцев, рождают новые пары кроликов в месяце $n+1$. Это пары, которые уже были в месяце $n-1$ и их количество составляет P_{n-1}. Итак,

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

Это рекуррентное соотношение: оно дает численность населения в месяц $n+1$ в зависимости от численности населения в предыдущие месяцы. Таким образом, Фибоначчи построил следующую последовательность чисел, в которой $1+1=2$, $1+2=3$, $2+3=5$, $3+5=8$ и т.д.

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

Фактически, Фибоначчи считал начальным условием ситуацию в месяц $n=2$. Поскольку $P_{14} = 144+233 = 377$, то через двенадцать месяцев после старта он, наконец, получил 377 пар кроликов. Он заметил, что эта последовательность чисел может продолжаться бесконечно.

После 1202 году Фибоначчи написал еще несколько книг, таких как *Геометрия Практики* в 1220 году и *Книга квадратов* в 1225 году. Его репутация привела к встрече с императором Фридрихом II, который высоко ценил науку. В 1240 году Пизанская республика присудила Фибоначчи ежегодную пенсию. Год его смерти неизвестен.

На протяжении последующих столетий проблема кроликов Фибоначчи была забыта и не имела никакого влияния на развитие математических моделей динамики населения. Несколько ученых встречали в своих исследованиях одну и ту же последовательность чисел, но не ссылались ни на Фибоначчи, ни на какую-либо популяцию. В некоторых книгах Кеплера содержится замечание, что соотношение P_{n+1}/P_n сходится, когда n стремится к бесконечности, к золотому числу $\phi = (1+\sqrt{5})/2$. Это особый случай свойства, общего для большинства популяционных моделей: тенденция к увелечению в геометрической прогрессии (см. главы 3 и 21). В 1728 году Даниэль Бернулли получил точную формулу

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

во время изучения общих повторяющихся серий.

Полные произведения Фибоначчи были опубликованы в XIX веке. С тех пор последовательность (P_n) можно было найти в книгах рекреационной математики под названием последовательности Фибоначчи.

Понятно, что для моделирования популяции кроликов гипотезы, ведущие к последовательности Фибоначчи, далеки от реальности: не учитываются смертность, разделение по половому признаку и т.д. Наш интерес к этой последовательности в последние десятилетия в биологии обусловлен тем, что некоторые растения содержат структуры, которые включают в себя некоторые числа P_n, например, 8 и 13 в сосновых шишках или 34 и 55 в подсолнечнике. Научный журнал «Ежеквартальный журнал Фибоначчи» даже полностью посвящен свойствам и применению последовательности Фибоначчи!

Дополнительное чтение

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, New York (2002).
3. Vogel, K.: Leonardo Fibonacci. In: Gillespie, C.C. (ed.) *Dictionary of Scientific Biography*, vol. 4, 604–613. Scribner, New York (1971)

Глава 2

Таблица смертности Галлея (1693)

В 1693 году известный английский астроном Эдмонд Галлей изучил записи о рождениях и смертях города Бреслау, которые были переданы Королевскому обществу Каспаром Нейманом. Он составил таблицу смертности, показывающую количество людей, доживающих до любого возраста от когорты, родившейся в том же году. Он также использовал свою таблицу для расчета стоимости пожизненного аннуитета. Эта глава напоминает эту работу и ставит ее в контекст жизни Галлея и ранних разработок «политической арифметики» и теории вероятностей, которые интересовали таких людей, как Граунт, Петти, Де Витт, Хадде, Гюйгенс, Лейбниц и Муавр.

Эдмонд Галлей (Halley) родился недалеко от Лондона в 1656 году. Его отец был богатым мыловаром. Эдмонд заинтересовался астрономией в юном возрасте. Он начал учиться в Королевском колледже Оксфордского университета. Когда в 1675 году была открыта Гринвичская обсерватория, Галлей уже мог посещать Флемстида, Королевского астронома. В 1676-1678 годах он прервал учебу, чтобы отправиться на остров Святой Елены и составить каталог звезд, которые можно увидеть из южного полушария. По возвращении в Англию он стал членом Королевского общества. Он также опубликовал сделанные им наблюдения о циркуляции ветров во время своего путешествия на остров Святой Елены. В 1684 году он посетил Ньютона в Кембридже, чтобы обсудить связь между законами планетарного движения Кеплера и силой притяжения Солнца. Он призвал Ньютона написать знаменитые *Математические принципы естественной философии* - книгу, которую он в итоге опубликовал за свой счет. Затем он работал клерком Королевского общества. В 1689 году он спроектировал колокол для подводных погружений, который испытал сам.

Примерно в это же время Каспар Нейман, богослов, живущий в Бреслау, собирал данные о количестве рождений и смертей в своем городе. Бреслау принадлежал к империи Габсбургов (сейчас он находится в Польше и называется Вроцлав). Данные включали в

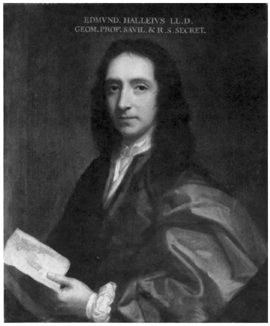


Рис. 2.1: Эдмонд Галлей (1656–1742)

себя возраст, в котором люди умирали. Таким образом, на основании этих данных можно было составить таблицу смертности, показывающую вероятность выживания до определенного возраста.

Первая таблица смертности была опубликована в Лондоне в 1662 году в книге под названием *Естественные и политические наблюдения, сделанные на основе бюллетеней о смерти*. Эта книга обычно рассматривается как основополагающий текст как статистики, так и демографии и имеет странную особенность: люди до сих пор задаются вопросом, была ли она написана Джоном Граунтом, лондонским купцом и автором, указанным на обложке книги, или его другом Уильямом Петти, одним из основателей Королевского общества. В любом случае таблица смертности, содержащаяся в книге, основывалась на бюллетенях, в которых регулярно сообщалось о погребениях и крещениях в Лондоне с начала семнадцатого века. Эти бюллетени в основном использовались для информирования людей о повторяющихся эпидемиях чумы. Именно поэтому в них указывалась причина смерти, а не возраст, в котором люди умирали. Чтобы получить таблицу смертности, дающую шанс на выживание в зависимости от возраста, Граунту или Петти приходилось соотносить различные причины смерти с возрастными группами. Такая таблица смертности может быть подвержена большим ошибкам. Тем не менее, книга оказалась очень успешной: между 1662 и 1676 годами было выпущено пять изданий. Несколько городов в Европе начали издавать бюллетени, похожие на бюллетени Лондона.

Итак, почти тридцать лет спустя после этой первой таблицы смертности, по предложению Лейбница, Нейман послал Анри Жюстелю, секретарю Королевского общества, свои демографические данные из города Бреслау за 1687-1691 годы. Вскоре после этого Жюстель умер, а Галлей получил эти данные, проанализировал их и в 1693 году опубликовал свои выводы в *Философских трудах Королевского общества.* Его статья называется *Оценка степени смертности человечества, составленная на основе любопытных таблиц рождений и похорон в городе Бреслау, с попыткой установить цену аннуитета на жизнь.*

Для исследованного пятилетнего периода, Галлей отмечал, что число рождений в Бреслау было более или менее равно числу смертей, так что общая численность населения была почти постоянной. Для упрощения анализа он предположил, что население находится в точном стационарном состоянии: годовое число рождений (назовем его P_0), общая численность населения, население в возрасте k (P_k) и годовое число умерших в возрасте k (D_k) - все это постоянные величины во времени. Это подчеркивает дополнительную интересную особенность данных из Бреслау, потому что такое упрощение было бы невозможно для быстрорастущего города, такого как Лондон, где статистика также отражала бы поток населения, приезжающего из сельской местности.

Таблица 2.1: Таблица смертности Галлея, показывающая численность населения P_k в возрасте $k > 0$.

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

Данные Бреслау имели в среднем 1 238 рождений в год: именно эту величину Галлей взял за P_0. В принципе, он также мог вычислить из этих данных среднегодовое значение D_k количества смертей среди людей в возрасте k для всех $k \geqslant 0$. Отсюда, используя формулу

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

можно построить таблицу 2.1, содержащую P_k. И, наоборот, можно найти значения D_k, которые он использовал, по формуле $D_k = P_k - P_{k+1}$: $D_0 = 238$, $D_1 = 145$, $D_2 = 57$, $D_3 = 38$ и так далее. На самом деле, Галлей немного подкорректировал свои результаты либо для получения круглых чисел (это случай D_1, который был слегка изменен так, что $P_1 = 1\,000$), либо для сглаживания определенных нарушений, связанных с небольшим количеством смертей в старости в пятилетнем исследовании. Взяв сумму всех приведенных в таблице чисел P_k, Галлей получил оценку общей численности населения Бреслау, близкую к 34 000 человек[1]. В целом, этот метод имел большое преимущество, поскольку не требовал проведения всеобщей переписи, а лишь знания о количестве рождений и смертей, а также о возрасте, в котором люди умирали в течение нескольких лет.

Таблица смертности Галлея служила справочным материалом для различных работ восемнадцатого века (см. главу 4). Действительно, хотя значения P_k были специфичны для города Бреслау, можно считать, что соотношение P_{k+1}/P_k было вероятностью дожить до возраста $k+1$, зная, что человек уже достиг возраста k. Эту вероятность можно было бы разумно использовать для населения других европейских городов того времени. Например, можно было бы ожидать, что у годовалого ребенка будет 661 шанс из 1 000 достигнуть 10-летнего возраста или 598 шансов из 1 000 достигнуть 20-летнего возраста.

Галлей также использовал таблицу смертности для расчета цены аннуитета на жизнь. В XVI и XVII веках несколько городов и государств продавали такие аннуитеты своим гражданам для сбора средств. Покупатели получали каждый год до своей смерти фиксированную сумму денег, которая равнялась определенному проценту от первоначально уплаченной суммы, часто вдвое превышающему проценты того времени, независимо от возраста покупателя. Конечно же, учреждение рисковало обанкротиться, если слишком много людей с очень длительным сроком жизни покупали эти аннуитеты. Проблема не могла быть правильно решена без надежной таблицы смертности.

[1] Для людей старше 84 лет Галлей только упомянул, что их число было 107.

В 1671 году Йохан де Витт, премьер-министр Голландии, и Йоханнес Хадде, один из мэров города Амстердам, уже задумывались над проблемой расчета стоимости пожизненного аннуитета. Опасаясь вторжения французских войск, они хотели собрать деньги на укрепление армии. У них были данные о людях, которые покупали аннуитеты за несколько десятилетий до этого, в частности, о возрасте, в котором аннуитеты были куплены, и о возрасте, в котором люди скончались. Им удалось более или менее правильно рассчитать цену аннуитетов, но позже их метод был забыт. В следующем году французская армия вторглась в Голландию. Де Витта обвинили в иностранной оккупации и его растерзала толпа на улице.

Галлей заново рассмотрел проблему в 1693 году с таблицей смертности из Бреслау и предположил, что процентная ставка составляет 6 %. Метод вычисления прост. Пусть в качестве процентной ставки будет i. Пусть R_k - это цена, по которой человек в возрасте k может купить годовой аннуитет, скажем, в один фунт. Этот человек имеет вероятность P_{k+n}/P_k быть еще живым в возрасте $k+n$. Фунт, который государство обещает заплатить, если он достигнет этого возраста, можно получить, разместив $1/(1+i)^n$ фунтов первоначальной суммы по процентной ставке i. Таким образом, если сделать упрощающее предположение, что начальная сумма используется только для выплаты аннуитетов, то цена должна быть

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

Галлей получил таким образом Таблицу 2.2, которая показывает коэффициент R_k, на который необходимо умножить желаемый аннуитет, чтобы получить необходимую начальную сумму. Таким образом, мужчина в возрасте 20 лет будет получать каждый год $1/12{,}78 \approx 7{,}8\,\%$ от начальной суммы. Но мужчина в возрасте 50 лет получал бы $1/9{,}21 \approx 10{,}9\,\%$, потому что ему оставалось бы жить меньше лет. Заметьте, что аннуитет, равный 12 %, т.е. в два раза больше процентной ставки, соответствовал бы цене, равной 8,33 аннуитета.

Вычисления, конечно, довольно утомительны. Тем не менее, Галлей мог использовать таблицы логарифмов, чтобы быстрее получить слагаемое $P_{k+n}/(1+i)^n$. Так как он не показывал значения для P_k в возрастах страше 84 лет, то проверить его вычисления точно невозможно. Работа Галлея не оказала немедленного влияния: в течение нескольких десятилетий аннуитеты, выплачиваемые

Таблица 2.2: Умножающий фактор, дающий цену аннуитетов на жизни.

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

за жизнь в Англии и других странах, продолжали продаваться по цене, не зависящей от возраста покупателя, и по цене, которая была значительно ниже, чем должна была бы быть, например, в 7 раз больше аннуитета.

Вопросы, связанные с таблицами смертности, интересовали многих ученых во времена Галлея. Голландец Кристиан Гюйгенс, автор в 1657 году первой брошюры, посвященной теории вероятностей, обсуждал в 1669 году в переписке со своим братом таблицу смертности Граунта и расчет ожидаемой продолжительности жизни[2].

За несколько лет до того, как посоветовать Нейману вступить в контакт с Королевским обществом, Лейбниц также написал о расчете продолжительности жизни в эссе, которое осталось неопубликованным. В 1709 году наступила очередь Николая I Бернулли. В 1725 году Абрахам де Муавр опубликовал целый *Трактат об аннуитетах*. Он, в частности, заметил, что цену R_k можно легко рассчитать для старости, так как формула (2.2) содержала всего несколько слагаемых. Затем можно было бы использовать рекуррентную формулу

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

которую легко вывести из (2.2). Используя значение, которое Галлей дает за цену в возрасте 70 лет, можно, таким образом, проверить остальные значения таблицы 2.2. [3]

После этого перерыва на демографию, Галлей вернулся к своим основным исследованиям. Между 1698 и 1700 годами он плавал по Атлантическому океану, чтобы нарисовать карту магнитного поля Земли. В 1704 году он стал профессором Оксфордского университета. В следующем году он опубликовал книгу о кометах и предсказал, что комета 1682 года, которую Кеплер наблюдал в 1607

[2] Ожидаемая продолжительность жизни в возрасте k дается по формуле (2.2) с $i = 0$.

[3] Похоже, что в таблице есть несколько ошибок, в частности для 5 и 15 лет.

году, вернется в 1758 году: она стала известна как «комета Галлея». Он также опубликовал перевод книги Аполлония Пергского о конических сечениях. В 1720 г. он заменил Флемстида на посту Королевского астронома. Он попытался решить проблему точного определения долготы в море по наблюдениям Луны - проблему, имевшую большое практическое значение для мореплавания. Он умер в Гринвиче в 1742 году в возрасте 86 лет.

Дополнительное чтение

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, New York (2001)
6. Koch, P.: Caspar Neumann. In: Ibid., 29–32.
7. Le Bras, H.: *Naissance de la mortalité.* Gallimard, Paris (2000)

Глава 3

Эйлер и геометрический рост популяций (1748–1761)

Эйлер несколько раз писал о динамике населения. В его трактате 1748 года *Введение в анализ бесконечно малых* глава, посвященная экспоненциальной функции, содержала четыре примера экспоненциального роста населения. В 1760 году он опубликовал статью, сочетающую этот экспоненциальный рост с возрастной структурой населения. Данная работа является предшественницей теории стабильного населения, которая была разработана в XX веке и играет важную роль в демографии. В 1761 г. Эйлер также помог Зюсмильху со вторым изданием своего трактата по демографии. Он разработал интересную модель, которая является своеобразным вариантом последовательности Фибоначчи, но не опубликовал свой подробный анализ.

Леонард Эйлер (Euler) родился в 1707 году в Базеле, Швейцария. Его отец был протестантским священником. В 1720 году Эйлер начал учиться в университете. Он также получил частные уроки математики у Иоганна Бернулли, одного из самых известных математиков поколения после Лейбница и Ньютона. Он подружился с двумя сыновьями Иоганна Бернулли: Николаем II и Даниилом. В 1727 году Эйлер поступил на работу к Даниилу во вновь созданную Академию наук в Санкт-Петербурге. Помимо математики он интересовался физикой и многими другими научными и техническими предметами. В 1741 году король Пруссии Фридрих II пригласил его стать директором математической секции Академии наук в Берлине. Эйлер опубликовал значительное количество статей и книг по всем аспектам механики (астрономии, теориям упругости, жидкостей, твёрдых тел) и математики (теории чисел, алгебре, бесконечным рядам, элементарным функциям, комплексным числам, дифференциальному и интегральному исчислению, обыкновенным дифференциальным уравнениям и уравнениям в частных производных, оптимизации, геометрии), а также по демографии. Он был самым плодовитым математиком своего времени.

Рис. 3.1:
Эйлер (1707–1783)

В 1748 году Эйлер опубликовал трактат на латинском языке под названием *Введение в анализ бесконечно малых.* В главе об экспонентах и логарифмах он рассмотрел шесть примеров: один - по математической теории музыкальных шкал, другой - о погашении кредита с процентами и четвертый - о динамике населения. В последнем Эйлер предположил, что население P_n в году n удовлетворяет уравнению

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

для любого целого n. Скорость роста x - положительное действительное число. При заданном начальном P_0, население в году n дается по формуле

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

Это называется геометрическим или экспоненциальным ростом. В первом примере спрашивается:

> «Если население того или иного региона ежегодно увеличивается на одну тридцатую и изначально насчитывает 100 000 жителей, то мы хотели бы знать численность населения через 100 лет.»

Ответ: $P_{100} = (1+1/30)^{100} \times 100\,000 \approx 2\,654\,874$. Для данного примера Эйлер был вдохновлен переписью населения Берлина, которая состоялась в 1747 году и дала оценку 107 224 человек. Его расчеты показывают, что население может увеличиться более чем в десять раз в течение одного столетия. Именно это наблюдалось в то время в Лондоне.

Следует отметить, что вычислить $(1 + 1/30)^{100}$ с помощью современного карманного калькулятора очень просто. Но во времена Эйлера приходилось использовать логарифмы, чтобы избежать многочисленных умножений вручную и быстро получить результат. Сначала вычисляется десятичный логарифм (в базе 10) P_{100}. Фундаментальное свойство логарифма $\lg(a\,b) = \lg a + \lg b$ показывает, что

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

Логарифмы были введены в 1614 году шотландцем Джоном Напье. Его друг Генри Бригс опубликовал первую таблицу десятичных логарифмов в 1617 году. В 1628 году голландец Адриан Влак закончил работу Бригса, опубликовав таблицу, содержащую десятичные логарифмы целых чисел от 1 до 100 000 с десятизначной точностью. Именно такую таблицу Эйлер использовал для получения $\lg 30 \approx 1{,}477121255$, $\lg 31 \approx 1{,}491361694$ и, наконец, $\lg P_{100} \approx 6{,}4240439$. Осталось найти число P_{100}, логарифм которого известен. Поскольку десятичные логарифмы целых чисел от 1 до 100 000 находятся в диапазоне от 0 до 5, то вместо него мы ищем логарифм $P_{100}/100$, который равен 4,4240439. В таблице логарифмов можно проверить, что $\lg 26\,548 \approx 4{,}424031809$ и $\lg 26\,549 \approx 4{,}424048168$. Заменив логарифмическую функцию на прямую между 26 548 и 26 549, Эйлер получил, что

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

Итак, $P_{100} \approx 2\,654\,874$.

Второй пример, касающийся динамики населения в книге Эйлера, выглядит следующим образом:

> «Так как после потопа все люди произошли от исходного населения в шесть человек, предположив, что население после двухсот лет составляло 1 000 000 человек, мы хотели бы найти ежегодные темпы роста.»

Из $10^6 = (1 + x)^{200} \times 6$, получаем с помщью карманного калькулятора: $x = (10^6/6)^{1/200} - 1 \approx 0{,}061963$. С помощью таблицы логарифмов нужно использовать $\lg(10^6) = 200\,\lg(1 + x) + \lg 6$, чтобы получить $\lg(1+x) = (6 - \lg 6)/200 \approx 0{,}0261092$, и $1 + x \approx 1{,}061963$. Таким образом, Эйлер мог бы сделать вывод, что население будет увеличиваться на $x \approx 1/16$ в год. Чтобы понять происхождение этого примера, нужно помнить, что современные философы начали отрицать истинность библейских историй. Буквальное чтение зафиксировало бы время потопа около 2350 г. до н.э. со следующими выжившими: Ной, его трое сыновей и их жены. В книге Бытия сказано:

> «Сии трое были сыновья Ноевы, и от них населилась вся земля.»

Темпы роста населения в 1/16 (или 6,25 %) в год после наводнения не казались Эйлеру слишком нереальными. Будучи сыном протестантского священника и оставаясь верующим всю свою жизнь, он пришел к выводу:

> «По этой причине, нелепо возражать, что за такой короткий промежуток времени вся Земля не может быть заселена, начиная с одного человека.[1]»

Эйлер также заметил, что если бы рост продолжался такими же темпами до 400 лет после потопа, численность населения составила бы $(1+x)^{400} \times 6 = (10^6/6)^2 \times 6 \approx 166$ миллиардов:

> «Однако, вся Земля никогда не сможет выдержать эту популяцию.»

Эта идея будет значительно развита Мальтусом полвека спустя (см. главу 5).

В третьем примере Эйлера спрашивается:

> «Если в каждом столетии человеческое население удваивается, каковы годовые темпы роста?»

[1] В книге, опубликованной Граунтом в 1662 году (см. главу 2), содержится аналогичное замечание:

> «Одна пара, а именно Адам и Ева, удваивая себя каждые 64 года из 5160 лет, что, согласно Писанию, является возрастом мира, должна произвести гораздо больше людей, чем сейчас в нем. А потому мир не старше 100 тысяч лет, как некоторые тщетно представляют себе, и не старше того, что дает Писание.»

Из $(1+x)^{100}=2$, мы получаем с помощью карманного калькулятора: $x=2^{1/100}-1\approx 0{,}00695$. С таблицами логарифмов, $100\lg(1+x)=\lg 2$. Так что $\lg(1+x)\approx 0{,}0030103$ и $1+x\approx 1{,}00695$. Таким образом, население растет на $x\approx 1/144$ каждый год. Наконец, в четвертом, и последнем, примере спрашивается:

> «Если человеческая популяция ежегодно увеличивается на 1/100 численности людей, то нам хотелось бы знать, сколько времени понадобится, чтобы численность населения увеличилась в десять раз.»

Из $(1+1/100)^n=10$, находим: $n\lg(101/100)=1$. Таким образом $n=1/(\lg 101-2)\approx 231$ лет. Это все, что можно найти во *Введении в анализ бесконечно малых* (1748) относительно динамики населения. Эйлер вернется к этой теме более подробно через несколько лет.

В 1760 году он опубликовал в трудах Академии наук в Берлине работу под названием *Общее исследование смертности и размножения человеческого рода.* Эта работа явилась своего рода синтезом между его предыдущим анализом геометрического роста популяций и более ранними исследованиями таблиц смертности (см. главу 2). Эйлер рассматривал, например, следующую проблему:

> «Зная количество рождений и погребений, которые происходят в течение одного года, найти количество всех живых и их ежегодный прирост, для заданной гипотезы смертности.»

Эйлер предположил, что здесь известны следующие числа:

- количество рождений B_n за год n;
- количество смертей D_n в течение года n;
- пропорция q_k новорожденных, достигающих возраста $k\geqslant 1$.

Пусть P_n равно численности населения в году n. Эйлер сделал два дополнительных неявных предположения:

- население увеличивается геометрически: $P_{n+1}=r\,P_n$ (мы обозначили $r=1+x$);
- соотношение между рождаемостью и численностью населения постоянно: $B_n/P_n=m$.

Эти два предположения подразумевают, что число рождений увеличивается геометрически и с той же самой скоростью: $B_{n+1} = r\,B_n$. Эйлер затем рассмотрел изменение населения в столетнем интервале времени, скажем, между годами $n = 0$ и $n = 100$, предполагая, что никто не выживет больше ста лет. Чтобы пояснить выкладки, обозначим $P_{k,n}$ ($k \geqslant 1$) население, которое жило в начале года n из родившихся в году $n - k$. Обозначим $P_{0,n} = B_n$ количество рождений в течение года n. Из определения коэффициента выживаемости q_k, имеем $P_{k,n} = q_k\,P_{0,n-k} = q_k\,B_{n-k}$. Итак,

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

Делением этого уравнения на $r^{100}\,P_0$, получается:

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

В демографии это уравнение иногда называют «уравнением Эйлера». Рассмотрев числа рождений и смертей по отдельности, получаем уравнение баланса:

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

Таким образом, количество смертей также увеличивается геометрически: $D_{n+1} = r\,D_n$. Более того,

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

Подставив это в уравнение (3.1), мы, наконец, приходим к уравнению.

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

где остался только один неизвестный: r. Это то, что обычно называют неявным уравнением, потому что мы не можем выразить r как функцию других параметров. Но мы можем вычислить левую и правую стороны уравнения (3.4) для фиксированного значения r и подбирать r до тех пор, пока две стороны не станут равны. Полученное таким образом значение r дает прирост населения $x = r - 1$.

Обратите внимание, что из уравнений (3.1) и (3.3) мы получаем для населения P_n следующее выражение:

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

Когда население стационарно ($r = 1$), это выражение совпадает с выражением, использованным Галлеем для оценки населения города Бреслау (см. главу 2).

Эйлер также рассмотрел следующий вопрос:

> «При заданных гипотезах смертности и плодовитости, зная общую численность населения, найти, сколько людей в каждом возрасте.»

Поскольку коэффициенты дожития q_k и коэффициент фертильности m известны, то скорость роста r можно вычислить из уравнения (3.1). В начале года n число людей, родившихся в году $n - k$, составляет $q_k\, B_{n-k} = q_k\, B_n/r^k$ (считая $q_0 = 1$). Таким образом, доля населения в возрасте k составляет

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

Эта пропорция постоянна. Используя терминологию Лотки (см. главу 10), говорят, что население "стабильно": возрастная пирамида населения сохраняет одну и ту же форму во времени.

Затем Эйлер переосмыслил проблему построения таблицы смертности в случае, когда население не стационарно, а увеличивается в геометрической прогрессии:

> «Зная число всех живых, так же, как и число рождений с числом смертей в каждом возрасте в течение одного года, найти закон смертности.»

Под законом смертности Эйлер имел в виду набор коэффициентов дожития q_k. В настоящее время предполагается само собой разумеющимся, что общая численность населения известна по результатам переписи, чего не было во время Галлея (см. главу 2). Уравнение (3.2) показывает, что скорость роста составляет

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

Пусть $D_{k,n}$ - это количество людей, которые умирают в возрасте k в течение года n: эти люди родились в году $n-k$. Так что $D_{k,n} = (q_k - q_{k+1})\, B_{n-k}$. Но $B_{n-k} = B_n/r^k$. Поэтому коэффициенты дожития q_k можно вычислить по рекуррентной формуле

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

для всех $k \geqslant 0$, при $q_0 = 1$. Эта формула, умноженная на B_n, дает формулу (2.1), использованную Галлеем для стационарного случая $r = 1$. Эйлер, тем не менее, настаивал на том, что его метод вычисления коэффициентов дожития q_k предполагает регулярный прирост населения, исключая такие несчастные случаи, как эпидемии чумы, войны, голод и т.д. Если бы переписи во времена Эйлера фиксировали возраст населения (как в Швеции), то это предположение было бы излишним, и коэффициенты q_k можно было бы вычислить легче.

Используя коэффициенты дожития q_k, Эйлер также показал, как рассчитать цену пожизненных аннуитетов. Он не упомянул работы Галлея или Муавра на эту тему. Эйлер использовал процентную ставку 5% и таблицу смертности, опубликованную в 1742 году голландцем Виллемом Керссебумом.

Эйлер был не единственным ученым, интересовавшимся демографией в Берлинской академии. Его коллега Иоганн Петер Зюсмильх (Süßmilch) опубликовал в 1741 году трактат на немецком языке под названием *Божественный порядок в изменениях рода человеческого, через рождения, смерти и размножение его*, который сегодня считается первым трактатом, полностью посвященным демографии. Зюсмильх в 1752 году также написал книгу *О стремительном росте города Берлина.*

В 1761 году Зюсмильх опубликовал второе издание своего трактата. В главу под заголовком *О темпах прироста и времени удвоения населения* он включил интересную математическую модель, которую для него разработал Эйлер. Эта модель была аналогична модели Фибоначчи (см. главу 1), но для человеческой популяции. Начиная с пары (один мужчина и одна женщина) в возрасте 20 лет в год 0, Эйлер предположил, что люди умирают в возрасте 40 лет и женятся в возрасте 20 лет, в то время как у каждой пары есть шесть детей: двое детей (мальчик и девочка) в возрасте 22 лет, двое других в возрасте 24 лет и двое последних в возрасте 26 лет. Считая годы попарно таким образом, что B_i - это число рождений в

Рис. 3.2:
Зюсмильх (1707–1767)

течение года $2i$, Эйлер пришел к выводу, что

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

для всех $i \geqslant 1$. Начальные условия соответствуют $B_{-12} = 0$, $B_{-11} = 0$, $B_{-10} = 2$ и $B_i = 0$ для $-9 \leqslant i \leqslant 0$. Таким образом, Эйлер смог вычислить число рождений, как показано во второй колонке таблицы 3.1. Число смертей D_i в году $2i$ тогда равно числу рождений в году $2i - 40$: $D_i = B_{i-20}$ для $i \geqslant 10$, в то время как $D_i = 0$ для $i \leqslant 9$. Что касается числа P_i живых людей в году $2i$, то оно равно числу живых людей в году $2i - 2$, плюс число рождений в году $2i$, минус число смертей в году $2i$: $P_i = P_{i-1} + B_i - D_i$.

Эта глава в книге Зюсмильх заканчивается замечанием, которое могло быть сделано и о последовательности Фибоначчи:

> «Несмотря на то, что в таблице Эйлера заметна большая нерегулярность, число рождений следует некой прогрессии, называемой рекуррентным рядом [...] Какими бы нерегулярными ни были эти прогрессии изначально, они обращаются в геометрическую прогрессию, если их не прерывать, а исходные нерегулярности постепенно затухают и почти полностью исчезают.»

В книге больше не сказано о математике этой модели популяции. Однако Эйлер продвинул исследование гораздо дальше в рукописи под названием *Об умножении рода человеческого*, которая оставалась неопубликованной в течение его жизни. В поисках решения уравнения (3.5) в виде $B_i = c\,r^i$, т.е. в виде геометрической прогрессии, он получил, после упрощения, уравнение полинома 13-й

Таблица 3.1: Таблица Эйлера.

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

степени:

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

Он искал решение, близкое к $r = 1$ и заметил, используя таблицу логарифмов для вычисления r^{13}, что

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

Таким образом, уравнение (3.6) имеет корень между 1,09 и 1,10. Приблизившись к функции $1 + r + r^2 - r^{13}$ отрезком прямой на этом интервале, Эйлер получил:

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

Годы, учитывая их попарный счет, имеют тенденцию к умножению числа рождений на $\sqrt{r}$ каждый год. Это число удваивается каждые n лет, если $(\sqrt{r})^n = 2$, т.е. каждые $n = 2\,\log 2/\log r \approx 15$ лет. Так как асимптотически $B_i \approx c\,r^i$, и так как число D_i смертей в году $2i$ равно B_{i-20}, мы получаем $D_i \approx B_i/r^{20}$ при том, что $r^{20} \approx 6{,}25$. Количество рождений примерно в шесть раз больше количества смертей. Поскольку количество P_i живых людей в году $2i$ равно $B_i + B_{i-1} + \cdots + B_{i-19}$, получаем, что

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

Общая численность населения примерно в десять раз превышает число рождений.

Доказательство того, что последовательность (B_i), показанная в таблице 3.1, действительно растет асимптотически как r^i, является более сложным. Со времен работы Абрахама де Муавра по рекуррентным рядам было известно, что, вводя генерирующую функцию

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

можно выразить $f(x)$ как рациональную функцию. Эйлер объяснил этот метод в своем *Введении в анализ бесконечно малых* в 1748 г.: рекуррентная связь (3.5) дает

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

Итак,

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

Эйлер знал, что такая рациональная функция может быть разложена в виде

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

где числа $x_1, \ldots, x_{13}$ - вещественные или комплексные корни уравнения $1 - x^{11} - x^{12} - x^{13} = 0$. Итак,

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

Так как B_i - это коэффициент при x^i в $f(x)$, Эйлер получил, что

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

при $i \to +\infty$, где x_k - корень с наименьшим модулем. Другими словами, B_i имеет тенденцию расти геометрически как $(1/x_k)^i$. Осталось отметить, что x_k является корнем уравнения $1 - x^{11} - x^{12} - x^{13} = 0$, если и только если $r = 1/x_k$ является корнем уравнения (3.6). Некоторые детали доказательства были окончательно уточнены Гумбелем в 1916 году.

Зюсмильх опубликовал третье издание своего трактата в 1765 году и умер в Берлине в 1767 году. Будучи в плохих отношениях с королем Пруссии, Эйлер вернулся в Петербург в 1766 году. Несмотря на то, что он потерял зрение, с помощью сыновей и коллег он продолжал издавать большое количество работ, особенно по алгебре, интегральному исчислению, оптике и кораблестроению. Его *Письма к немецкой принцессе*, написанные в Берлине между 1760 и 1762 годами, были опубликованы между 1768 и 1772 годами и стали бестселлером по всей Европе. Эйлер умер в Санкт-Петербурге в 1783 году. Его вклад в математическую демографию, особенно его анализ стабильной возрастной пирамиды при экспоненциально растущем населении, будет вновь открыт только в XX веке (см. главы 10 и 21).

Дополнительное чтение

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

Глава 4

Даниэль Бернулли, д'Аламбер и прививка от оспы (1760)

> В 1760 году Даниэль Бернулли написал статью по моделированию эпидемии оспы. В его время было много споров о прививках, использование которых могло защитить людей, но могло и привести к смертельным случаям. Он использовал таблицу жизни Галлея и некоторые данные, касающиеся оспы, чтобы показать, что инокуляция была выгодна, если связанный с ней риск смерти был меньше, чем 11%. Вакцинация может увеличить ожидаемую продолжительность жизни при рождении до трех лет. Д'Аламбер подверг критике работу Бернулли, в которой была предложена первая математическая модель в эпидемиологии.

Даниэль Бернулли (Bernoulli) родился в 1700 году в Гронингене в Нидерландах. В его семью входили уже два известных математика: его отец Иоганн Бернулли и его дядя Якоб Бернулли. В 1705 г. Иоганн переехал в Базель в Швейцарии, где после смерти Якоба занял должность профессора, которая осталась вакантной. Иоганн не хотел, чтобы его сын изучал математику. Поэтому Даниил обратился к медицине, получив докторскую степень в 1721 году с диссертацией на тему дыхания. Он переехал в Венецию и начал математические исследования, опубликовав книгу в 1724 году. Получив в том же году премию Парижской академии наук за сочинение *В совершенстве песочные часы на корабле в море*, он получил профессорскую степень в новой петербургской академии. В эти годы он работал, в частности, над периодическими последовательностями и над «Парадоксом Санкт-Петербурга» по теории вероятностей. В 1733 году Даниэль Бернулли вернулся в Базельский университет, где преподавал ботанику, физиологию и физику. В 1738 году он опубликовал книгу по гидродинамике, которая осталась известной в истории физики. Около 1753 года он заинтересовался, как Эйлер и d'Алембер, проблемой вибрирующих струн, что вызвало оживленную математическую дискуссию.

Рис. 4.1: Даниэль Бернулли (1700–1782)

В 1760 году он представил в Академию наук в Париже работу под названием *Попытка нового анализа смертности от оспы и преимуществ прививок для ее предотвращения.* Вопрос заключался в том, следует ли поощрять прививку (добровольное введение в организм небольшого количества менее вирулентной оспы для защиты его от последующих инфекций), даже если иногда это может привести к смертельному исходу. Этот метод был давно известен в Азии и был введен в 1718 году в Англии леди Монтагу, женой британского посла в Османской империи. Во Франции, несмотря на смерть старшего сына Людовика XIV от оспы в 1711 году, инокуляция считалась нежелательной. Вольтер, переживший оспу в 1723 году и проживший несколько лет в изгнании в Англии, наблюдая за последними новшествами, в своей книге *Философские письма* в 1734 году призывал делать прививки. Французский ученый Ла Кондамин, также переживший оспу, в 1754 году обратился в Академию наук в Париже с просьбой сделать ему прививку.

Перед смертью в Базеле в 1759 году Мопертуис призвал Даниэля Бернулли изучить проблему инокуляции с математической точки зрения. Точнее, задача состояла в том, чтобы найти способ сопоставить долгосрочную выгоду от инокуляции с непосредственной опасностью смерти. С этой целью Бернулли сделал следующие упрощающие предположения:

- люди, впервые инфицированные оспой, умирают с вероятностью p (независимо от возраста) и выживают с вероятностью $1 - p$;

- у каждого человека есть вероятность q быть зараженным каждый год; точнее, вероятность заражения одного человека в возрасте от x до $x + dx$ составляет $q\,dx$, где dx - это бесконечно малый временной промежуток;

- люди, выжившие от оспы, защищены от новых инфекций на всю оставшуюся жизнь (они были вакцинированы).

Пусть $m(x)$ будет смерть в возрасте x от других причин, кроме оспы: вероятность смерти одного человека в бесконечно малом промежутке времени dx между возрастом x и возрастом $x + dx$ составляет $m(x)\,dx$. Рассматривая группу P_0 людей, рожденных в том же году, обозначим через

- $S(x)$ количество восприимчивых к оспе людей, имеющих возраст x и никогда не болевших оспой;
- $R(x)$ количество людей, имеющих возраст x и выживших после оспы;
- $P(x) = S(x) + R(x)$ общее количество людей, имеющих возраст x.

Рождение соответствует возрасту $x = 0$. Таким образом, $S(0) = P(0) = P_0$ и $R(0) = 0$. Применяя методы вычисления, разработанные в конце семнадцатого века Ньютоном, Лейбницом и позже его отцом, Даниэль Бернулли заметил, что между возрастом x и $x + dx$ (с dx бесконечно малым) у каждого восприимчивого человека есть вероятность $q\,dx$ заражения оспой и вероятность $m(x)\,dx$ смерти от других причин. Таким образом, вариация числа восприимчивых людей составляет $dS = -S\,q\,dx - S\,m(x)\,dx$, что приводит к дифференциальному уравнению

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

В этом уравнении dS/dx обозначает производную функции $S(x)$. За тот же небольшой промежуток времени количество людей, умирающих от оспы, составляет $p\,S\,q\,dx$, а количество людей, выживающих от оспы, составляет $(1 - p)\,S\,q\,dx$. Кроме того, есть также $R\,m(x)\,dx$ людей, которые умирают от других причин, кроме оспы. Это приводит ко второму дифференциальному уравнению:

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

Складывая эти два уравнения, мы получаем

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

Из уравнений (4.1) и (4.3) Бернулли показал, что доля людей, которые все еще восприимчивы в возрасте x, составляет

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

Чтобы получить формулу (4.4), Бернулли выразил $m(x)$ из уравнений (4.1) и (4.3):

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

После подстановки получается, что

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

Заметим, что левая чать этого равенства является производной функции $f(x) = S(x)/P(x)$, которая представляет собой долю восприимчивых людей в популяции в возрасте x. Итак,

$$\frac{df}{dx} = -q\,f + p\,q\,f^2. \tag{4.5}$$

Решение уравнения такого типа было известно в течение нескольких десятилетий благодаря работе Якоба Бернулли, дяди Даниэля. Деля это уравнение на f^2 и полагая $g(x) = 1/f(x)$, мы видим, что $dg/dx = q\,g - p\,q$ и $g(0) = 1/f(0) = 1$. Вводя обозначение $h(x) = g(x) - p$, мы получим $dh/dx = q\,h$. Итак, $h(x) = h(0)\,e^{q\,x} = (1-p)\,e^{q\,x}$. Наконец, $g(x) = (1-p)\,e^{q\,x} + p$ и $f(x) = 1/g(x)$.

Для применения своей теории Бернулли использовал таблицу дожития Галлея (см. главу 2). В этой таблице указано количество людей, еще живущих в начале года x (при $x = 1, 2...$) среди когорты в 1,238 людей, родившихся в течение нулевого года. Но в рамках его модели Бернулли нужно было знать количество людей

$P(x)$, которые на самом деле достигают возраста x, что несколько отличается. Поскольку Бернулли, как и большинство его современников, не осознавал разницы (статья Галлея на самом деле не очень ясна), он оставил цифры в таблице Галлея, за исключением первого числа 1 238, которое он заменил на 1 300, чтобы получить реальную смертность в течение первого года жизни. Эти цифры показаны во второй колонке таблицы 4.1.

Таблица 4.1: Таблица Галлея и вычисления Бернулли.

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

Бернулли выбрал для вероятности смерти от оспы значение $p =$

$1/8 = 12{,}5\,\%$, что согласуется с наблюдениями того времени. Вероятность заболеть оспой в течение года q не может быть оценена напрямую. Поэтому Бернулли, вероятно, попробовал несколько значений для q и, в конце концов, выбрал одно из них таким образом, что число смертей от оспы после всех приведенных ниже вычислений составляет около $1/13$ от общего числа смертей - доля, которая тогда наблюдалась в нескольких европейских городах. Как мы сейчас увидим, выбор $q = 1/8$ оказался удачным[1].

С помощью формулы (4.4) и значений $P(x)$ во втором столбце таблицы можно вычислить число $S(x)$ восприимчивых людей в возрасте x: это третий столбец таблицы, округленный до ближайшего целого числа. В четвертой колонке показано число $R(x) = P(x) - S(x)$ людей в возрасте x, переживших оспу. В пятой колонке в строке, соответствующей возрасту x, показано количество смертей от оспы между возрастом x и $x+1$. Теоретически, это число должно выражаться через интеграл: $p\,q \int_x^{x+1} S(t)\,dt$, но формула $p\,q\,[S(x) + S(x+1)]/2$ дает хорошую аппроксимацию, как показано на рисунке 4.2: площадь трапеции близка к площади под кривой, т.е. к интегралу функции.

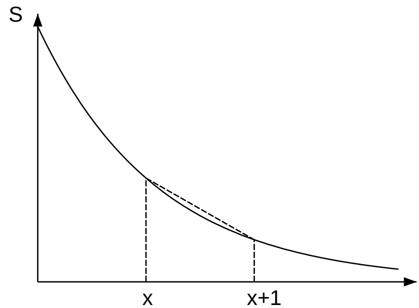


Рис. 4.2: Область трапеции, показанной пунктирной линией, аппроксимирует интеграл от функции S между x и $x+1$.

Бернулли заметил, что сумма всех чисел в пятой колонке дает 98 смертей от оспы до 24 лет. Если бы мы продолжили таблицу для более старших возрастов, то обнаружили бы только еще три смерти от оспы среди 32 человек, которые все еще подвержены этой болезни в возрасте 24 лет. Таким образом, из 1300 новорожденных 101 умрут от оспы. Это почти точно ожидаемая пропорция $1/13$.

[1] Тот факт, что p и q равны, - это просто совпадение.

Бернулли тогда рассматривал ситуацию, когда оспа будет привита всем при рождении и не вызовет никаких смертей. Оспа будет искоренена, и вопрос заключается в том, чтобы оценить увеличение продолжительности жизни. Начиная с того же числа рождений P_0, обозначим через $P^*(x)$ число людей в возрасте x при отсутствии оспы. Тогда

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

Бернулли показал, что

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

где $P(x)$, как и выше, население в возрасте x при наличии оспы.

Действительно, исключая, как и прежде, $m(x)$ из уравнений (4.6) и (4.3), Бернулли получил после преобразований

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

Полагая $h(x) = P(x)/P^*(x)$, и, умножив в (4.4) числитель и знаменатель на $e^{-q\, x}$, он вывел уравнение

$$\frac{1}{h}\, \frac{dh}{dx} = -p\, q\, \frac{e^{-q\, x}}{1 - p + p\, e^{-q\, x}}\ ,$$

что эквивалентно $\frac{d}{dx} \ln h = \frac{d}{dx} \ln(1 - p + p\, e^{-q\, x})$, где ln обозначает натуральный логарифм. Но $h(0) = 1$, так что $h(x) = 1 - p + p\, e^{-q\, x}$.

Отметим, что выражение $P(x)/P^*(x)$ стремится к $1 - p$, когда возраст x возрастает. Шестая колонка таблицы 4.1 показывает $P^*(x)$. Способом сравнения $P(x)$ и $P^*(x)$ является оценка ожидаемой продолжительности жизни при рождении, теоретическое выражение которой при оспе составляет

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

Аналогичное выражение с $P^*(x)$, заменяющим $P(x)$, имеет место при отсутствии оспы. Бернулли использовал приближенную формулу

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

которая получена методом трапеций (рисунок 4.2). Продолжая таблицу с 24 лет до 84 лет (см. таблицу 2.1), он, в результате, получил ожидаемую продолжительность жизни E с оспой, равную

$$\left[\frac{1}{2} \times 1\,300 + 1\,000 + \cdots + 20\right] / 1\,300 \approx 26{,}57,$$

т.е. 26 лет и 7 месяцев. Без оспы он получил ожидаемую продолжительность жизни E^*, равную

$$\left[\frac{1}{2} \times 1\,300 + 1\,015 + \cdots + 23\right] / 1\,300 \approx 29{,}65,$$

т.е. 29 лет и 8 месяцев. Инокуляция при рождении увеличивает ожидаемую продолжительность жизни более чем на три года.

Можно отметить, что существует более простой и быстрый метод, чем метод, используемый Бернулли для получения этих формул. Начиная с дифференциального уравнения (4.1) для $S(x)$, сначала мы видим, что

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

Используя это выражение в уравнении (4.2) для $R(x)$, мы находим, что

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

Уравнение (4.6) для $P^*(x)$ показывает, что

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

Формулы (4.4) и (4.7) вытекают отсюда.

Конечно, инокуляция с менее вирулентным штаммом оспы не совсем безопасна. Если p' - это вероятность смерти от оспы сразу после инокуляции ($p' < p$), то ожидаемая продолжительность жизни составила бы $(1 - p')\,E^*$, если бы все прошли инокуляцию при рождении. Эта ожидаемая продолжительность жизни остается выше, чем естественная ожидаемая продолжительность жизни E, если $p' < 1 - E/E^*$ или около 11 %. Данные о p' было трудно получить в то время. Но Бернулли подсчитал, что риск p' был меньше, чем 1 %. Для него не было никаких сомнений: прививку должно было проводить государство. Он сделал вывод:

> «Я просто хочу, чтобы в вопросе, который так тесно связан с благополучием человеческого рода, ни одно решение не принималось без всех знаний, которые могут быть получены в результате небольшого анализа и расчетов.»

Работы Бернулли были представлены в Академии наук в Париже в апреле 1760 года. В ноябре д'Аламбер (d'Alembert) опубликовал комментарии под названием *О применении теории вероятности к инокуляции оспы*. Эти комментарии были опубликованы вскоре после этого во втором томе его работы *Opuscules mathématiques* с более подробными расчетами и вместе с другой работой, озаглавленной *Математическая теория прививок*. Д'Аламбер критиковал предположения Бернулли о вероятности заражения и о том, что вероятность смерти от оспы не зависит от возраста. Он предложил другое решение, не требующее этих предположений. Обозначим через $v(x)$ смертность от оспы в возрасте x, через $m(x)$ смертность от других причин и через $P(x)$ количество людей, которые еще живы. Тогда

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

Сравнивая это уравнение с уравнением (4.3), мы видим, что на самом деле $v(x) = p\,q\,S(x)/P(x)$. Мы получаем

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

где $P^*(x)$ означает количество людей в возрасте x, оставшихся в живых, когда оспа исчезла.

Действительно, мы можем либо выразить функцию $m(x)$ из уравнений (4.6) и (4.9), либо использовать формулу (4.8) для

Рис. 4.3:
Д'Алембер (1717–1783)

$P^*(x)$ и заметить, что решение (4.9) дается по формуле

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

Формула (4.10), данная д'Аламбером, не противоречит формуле Бернулли (4.7). В ней просто используется другая зависимость $v(x)$, которая в то время был недоступена, так как в регистрах смерти была указана причина смерти, а не возраст жертвы. Даламбер высказал предположение, что нельзя сделать реальный вывод о том, была ли полезной инъекция до того, как этот тип данных стал доступен.

Даламбер также критиковал полезность продолжительности жизни в качестве критерия для принятия решения, поскольку она придает одинаковый вес всем возрастам, будь то в ближайшем или отдаленном будущем. Он заметил, что с точки зрения индивида или государства не все возраста имеют одинаковую «полезность», молодые и старые годы имеют меньшую ценность, чем средние. Несмотря на все эти критические замечания, Даламбер заявил о своей позиции в пользу прививок.

Из-за задержек с публикацией работа Бернулли появилась только в 1766 г., в то время как Даламберу удалось очень быстро опубликовать свой труд. Бернулли выразил свою горечь в письме к Эйлеру:

> «Что вы скажете про выдающуюся посредственность великого Даламбера в теории вероятностей: поскольку я слишком часто нахожу несправедливую критику в свой адрес в его публикациях, я уже некоторое время назад

> решил больше ничего не читать, что исходит из-под его пера. Я принял это решение в связи с рукописью о прививках, которую я отправил в Академию в Париже восемь лет назад и которая была высоко оценена из-за новизны анализа. Это было, осмелюсь сказать, как развитие новой области в математике. Кажется, что успех этого нового анализа вызвал у него боль в сердце. Он критиковал его тысячу раз одинаково нелепо, и после того, как его хорошо раскритиковали, он притворился первым автором теории, которую он не только слышал, но и упомянул. Он, однако, знал, что моя рукопись может появиться только через семь или восемь лет. Он мог знать об этом только как член Академии. В этом отношении моя рукопись должна была оставаться неприкосновенной до тех пор, пока она не была обнародована. *Dolus an virtus quis in hoste requirat!*[2]»

Несмотря на произведения Бернулли и Даламбера, во Франции крупномасштабные прививки не проводились. Король Людовик XV умер от оспы в 1774 году. Вскоре после этого придворные врачи сделали прививки остальным членам королевской семьи. Проблема потеряла свою актуальность, когда Эдуард Дженнер обнаружил, что прививка коровьей оспы людям («вакцинация») защищает от оспы и является безопасной. Его работа, *Исследование причин и последствий вакцинации против оспы*, была опубликована в 1798 году. Вакцинация быстро распространилась по всей Европе. Тем не менее, методы, разработанные для расчета увеличения продолжительности жизни в случае устранения одной из причин смерти, используются и по сей день.

В последующие десятилетия стали доступны данные о возрасте, в котором люди умирали от оспы. Проблема была пересмотрена, в частности, такими математиками как

- Иоганн Генрих Ламберт, математик из Берлинской академии, в 1772 году;
- Emmanuel-Étienne Duvillard, в то время ответственный за демографическую статистику в Министерстве внутренних дел в Париже, в своем труде *Анализ и таблицы влияния оспы на смертность в каждом возрасте* (1806);

[2] *Какая разница как победить, отвагой или хитростью, главное - победить!* Вергилий: Энейд, книга II.

- Пьер-Симон Лаплас в своей работе *Аналитическая теория вероятности* (1812).

Дювиллард и Лаплас показали, например, как модифицировать формулу (4.7), когда параметры p и q зависят от возраста:

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

Здесь $p(x)$ - вероятность умереть от оспы при заражении в возрасте x, а $q(x)$ - вероятность заражения оспой в возрасте x.

После этой работы по оспе Даниэль Бернулли не рассматривал никаких других проблем в динамике популяции. Он умер в Базеле в 1782 году. Даламбер умер в Париже год спустя.

Дополнительное чтение

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

Глава 5

Мальтус и препятствия на пути геометрического роста (1798)

В 1798 году Мальтус опубликовал *Очерк о законе народонаселения*, в котором он утверждал, что поставки продовольствия не могут в течение длительного периода времени следовать естественной тенденции человеческого населения к экспоненциальному росту. Если численность населения оставалась относительно постоянной, то это происходило потому, что большая часть человечества страдала от нехватки продовольствия. Мальтус рассматривал «принцип народонаселения» как аргумент против записок Годвина и Кондорсе, в которых подчеркивается прогресс в человеческом обществе. Эссе Мальтуса оказало влияние на теорию эволюции Дарвина и Уоллеса и было подвергнуто критике со стороны Маркса, но было реализовано на практике с помощью китайской политики в отношении одного ребенка.

Томас Роберт Мальтус (Malthus) родился в 1766 году недалеко от Лондона шестым из семи детей. Его отец, друг и поклонник Жан-Жака Руссо, был его первым учителем. В 1784 году молодой Мальтус начал изучать математику в Кембриджском университете. Он получил диплом в 1791 году, стал стипендиатом Колледжа Иисуса в 1793 году и англиканским священником в 1797 году.

В 1798 году Мальтус анонимно опубликовал книгу под названием *Очерк о законе народонаселения, как он влияет на будущее улучшение общества, с замечаниями по поводу спекуляций мистера Годвина, мистера Кондорсе и других писателей*. Он был подготовлен как реакция на работы Годвина *Запрос относительно политической справедливости* (1793) и Кондорсе *Эскиз исторической картины прогресса человеческого разума* (1794). Несмотря на ужасы, которые совершила Французская революция во имя прогресса, оба автора утверждали, что прогресс общества неизбежен. Мальтус не разделял того же оптимизма. Он также утверждал, что английские законы о бедных, которые помогали многодетным мало-

Рис. 5.1:
Мальтус (1766–1834)

имущим семьям, способствовали росту населения, не поощряя аналогичного роста производства продуктов питания. Ему казалось, что эти законы на самом деле не облегчают положение бедных, а, наоборот, усугубляют его. В более общем плане, в условиях, когда население склонно расти быстрее, чем производство продуктов питания, часть общества, казалось, была обречена на страдания, голод или эпидемии: это те бедствия, которые замедляют рост населения, и которые, по мнению Мальтуса, являются принципиальными препятствиями на пути прогресса общества. Все теории, обещающие прогресс, просто утопичны. Эти идеи побудили Мальтуса опубликовать свою книгу в 1798 году. Вот как он подытожил свою диссертацию:

> [...] «сила населения бесконечно более велика, чем способность земли дать человеку средства к существованию. Население, если его не контролировать, увеличивается в геометрической прогрессии. Средства к существованию увеличиваются только в арифметической прогрессии. Даже незначительное знакомство с числами приведет к выводу, что в сравнении со второй силой, первая сила окажется безгранична. По тому закону нашей природы, который делает пищу необходимой для жизни человека, последствия этих двух неравных сил должны постоянно быть уравнены. Это подразумевает сильное и постоянно действующее ограничение населения трудностью пропитания. Эта трудность должна где-то проявиться; и обязательно должна серьезно ощущаться значительной частью человечества.»

Книга Мальтуса была очень успешной. Она содержала мало данных. Мальтус заметил, например, что население США удваивалось каждые двадцать пять лет в течение восемнадцатого века. На самом деле он не пытался перевести свои тезисы на математические модели, но проложил путь для последующих работ Адольфа Кетеле и Пьера-Франсуа Ферхюльста, которым будет посвящена следующая глава.

После публикации книги Мальтус вместе с друзьями отправился сначала в Германию, Скандинавию и Россию, затем во Францию и Швейцарию. Собрав воедино информацию, собранную во время путешествий, он опубликовал под своим именем очень крупное второе издание в 1803 году, с другим подзаголовком: *Очерк о законе народонаселения, или взгляд на его прошлое и нынешнее воздействие на человеческое счастье, с обзором наших перспектив в отношении будущего устранения или смягчения зла, которое он совершает*. В этом новом издании подробно рассматриваются препятствия на пути роста населения в различных странах: запоздалые браки, аборты, детоубийство, голод, войны, эпидемии, экономические факторы .... Для Мальтуса отсроченные браки были лучшим вариантом стабилизации численности населения. Четыре других издания книги последовали в 1806, 1807, 1817 и 1826 годах. В 1805 году Мальтус стал профессором истории и политэкономии в новой школе, созданной Вест-индской компанией для своих сотрудников. Он также опубликовал *Исследование природы и прогресса арендной платы* (1815) и *Принципы политической экономии* (1820). В 1819 году Мальтус был избран в Королевское общество. В 1834 году он был одним из членов-учредителей Статистического общества. В том же году он умер недалеко от г. Бат.

Работа Мальтуса оказала сильное влияние на развитие теории эволюции. Чарльз Дарвин, вернувшись из путешествия на борту *Beagle*, прочитал книгу Мальтуса о населении в 1838 году. Вот что он написал во вступлении к своей знаменитой книге *О происхождении видов путем естественного отбора*, опубликованной в 1859 году:

> «Мы будем, однако, в состоянии обсудить, какие условия наиболее благоприятны для вариации. В следующей главе будет рассмотрена Борьба за Существование между всеми органическими существами во всем мире, которая неизбежно вытекает из геометрической прогрессии роста их численности. Это — доктрина Мальтуса, распро-

> страненная на оба царства — животных и растений.»

Альфред Рассел Уоллес, разработавший теорию эволюции в то же время, что и Дарвин, также сказал, что его идеи пришли после прочтения книги Мальтуса.

Для контраста, приведем точку зрения Карла Маркса на успех книги Мальтуса, которую можно прочитать в сноске его труда *Капитал*:

> «Если читатель вспомнит о Мальтусе, работа которого *Essay on Population* появилась в 1798 г., то я напомню, что эта работа в своей первоначальной форме есть не что иное, как ученически-поверхностный и поповски-напыщенный плагиат из Дефо, сэра Джемса Стюарта, Таунсенда, Франклина, Уоллеса и т.д. и не содержит ни одного самостоятельного положения. Большой шум, вызванный этим памфлетом, объясняется исключительно партийными интересами. Французская революция нашла в Британском королевстве страстных защитников: „закон народонаселения", медленно вырисовывавшийся в XVIII веке, потом с помпой возвещенный среди великого социального кризиса как несравненное противоядие против теории Кондорсе и других, был с ликованием встречен английской олигархией, которая увидела в нем великого искоренителя всех стремлений к дальнейшему человеческому развитию. Мальтус, до крайности изумлённый своим успехом, принялся тогда за то, чтобы заполнить старую схему поверхностно компилированным материалом и присоединить к нему новый, который был, однако, не открыт, а просто присвоен Мальтусом.»

Конечно, тезисы Мальтуса не были совершенно новыми. Например, ему часто приписывают идею о том[1] , что население имеет тенденцию к геометрическому росту, хотя в главе 3 мы видели, что эта идея уже была знакома Эйлеру на полвека раньше. Однако Мальтус придал ей известность, связав ее полемическим путем с реальными законодательными проблемами. По иронии судьбы именно в коммунистическом Китае предложение Мальтуса об ограничении рождаемости нашло свое самое поразительное применение (см. главу 25).

[1] Р. А. Фишер (см. главы 14 и 20) назвал «Мальтузианским параметром» темп роста популяции. Мальтус упомянул трактат Зюсмильха в своей книге.

Дополнительное чтение

1. Condorcet: *Esquisse d'un tableau historique des progrès de l'esprit humain.* Agasse, Paris (1794). gallica.bnf.fr
2. Дарвин Ч. (переводчик: А.Л. Тахтаджян): *Происхождение видов путём естественного отбора.* charles-darwin.narod.ru
3. Godwin, W.: *An Enquiry Concerning Political Justice.* Robinson, London (1793). archive.org
4. Malthus, T.R.: *An Essay on the Principle of Population*, 1st edn. London (1798). econlib.org
5. Маркс, К.: Капитал, том 1. esperanto.mv.ru/Marksismo/Kapital1
6. Simpkins, D.M.: Malthus, Thomas Robert. In: Gillespie, C.C. (ed.) *Dictionary of Scientific Biography*, vol. 9, 67–71. Scribner, New York (1974)

Глава 6

Ферхюльст и логистическое уравнение (1838)

В 1838 г. бельгийский математик Ферхюльст ввел логистическое уравнение, которое является своего рода обобщением уравнения для экспоненциального роста, но с максимальным значением для населения. Он использовал данные из нескольких стран, в частности из Бельгии, для оценки неизвестных параметров. Работа Ферхюльста была заново открыта только в 1920-х годах.

Пьер-Франсуа Ферхюльст (Verhulst) родился в 1804 году в Брюсселе. Он получил докторскую степень по математике в Гентском университете в 1825 году. Он также интересовался политикой. Находясь в Италии для сдерживания развития его туберкулеза, он безуспешно молил в пользу конституции для папских штатов. После революции 1830 года и независимости Бельгии, он опубликовал исторический очерк о патриоте восемнадцатого века. В 1835 году он стал профессором математики во вновь созданном Свободном университете в Брюсселе.

Рис. 6.1: Ферхюльст (1804–1849)

В том же 1835 году его соотечественник Адольф Кетле, статистик и директор обсерватории в Брюсселе, опубликовал *Трактат о человеке и развитии его способностей.* Кетле предполагал, что население не может расти геометрически в течение длительного периода времени, поскольку упомянутые Мальтусом препятствия образуют своего рода «сопротивление», которое, по его мнению (по аналогии с механикой), пропорционально квадрату скорости роста населения. Эта аналогия не имела под собой реальной основы, но вдохновляла Ферхюльста.

Действительно, Ферхюльст опубликовал в 1838 году *Примечание о законе роста населения.* Вот некоторые выдержки:

> «Мы знаем, что знаменитый Мальтус показал принцип, что человеческое население имеет тенденцию расти в геометрической прогрессии так, чтобы удваивать численность по прошествии определенного периода времени, например, каждые двадцать пять лет. Это предложение не подлежит сомнению, если абстрагироваться от возрастающих трудностей в поисках пищи [...]
>
> Таким образом, виртуальный прирост населения ограничивается размерами и рождаемостью страны. В результате население все ближе и ближе приближается к стабильному состоянию.»

Ферхюльст, вероятно, понял, что механическая аналогия Кетле нецелесообразна, и предложил вместо этого следующее (все еще несколько произвольное) дифференциальное уравнение для численности популяции $P(t)$ в момент времени t:

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

Когда популяция $P(t)$ мала по сравнению с параметром K, получаем приближенное уравнение $dP/dt \approx r\,P$, чьим решением является $P(t) \approx P(0)\,e^{r\,t}$, т.е. экспоненциальный рост[1]. Темп роста снижается по мере приближения $P(t)$ к K. Он даже станет отрицательным, если $P(t)$ превысит K. Для получения точного выражения решения уравнения (6.1), мы можем использовать подход Даниэля Бернулли для уравнения (4.5).

[1] Обычно говорят о геометрическом росте в дискретно-временных моделях и об экспоненциальном росте в моделях с непрерывным временем, но это, по сути, одно и то же.

Разделив уравнение (6.1) на P^2 и используя замену переменной $p = 1/P$, получим $dp/dt = -r\,p + r/K$. Заменив на $q = p - 1/K$, получим $dq/dt = -r\,q$, и $q(t) = q(0)\,e^{-r\,t} = (1/P(0) - 1/K)\,e^{-r\,t}$. Отсюда можно вывести $p(t)$ и $P(t)$.

В итоге, получаем после перестановки:

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

Общая численность населения постепенно увеличивается с $P(0)$ в момент времени $t = 0$ до предела K при $t \to +\infty$ (рисунок 6.2). Не приводя значения неизвестных параметров r и K, Ферхюльст

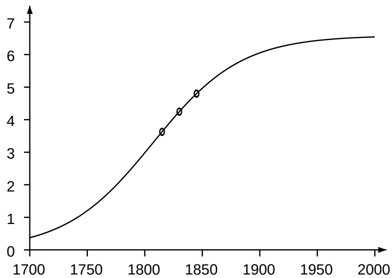


Рис. 6.2: Население Бельгии (в миллионах) и кривая логистики. Точки данных соответствуют 1815, 1830 и 1845 годам. Значения параметров соответствуют значениям статьи 1845 года.

сравнил свой результат с данными о населении Франции в 1817-1831 гг., Бельгии в 1815-1833 гг., графства Эссекс в Англии в 1811-1831 гг. и России в 1796-1827 гг. Точность подгонки оказалась довольно неплохой.

В 1840 году Ферхюльст стал профессором Королевского военного училища в Брюсселе. В следующем году он опубликовал *Элементарный трактат об эллиптических функциях* и был избран в Королевскую академию Бельгии. В 1845 году он продолжил исследования в области народонаселения, написав статью под названием *Математические исследования закона прироста населения.* Сначала он вернулся к замечанию Мальтуса, согласно которому население США удваивалось каждые 25 лет (таблица 6.1).

Если (по данным таблицы 6.1) вычислить соотношения между численностью населения за год $n+10$ и за год n, то получится 1,350, 1,364, 1,331, 1,335 и 1,326, соответственно, что является достаточно

Таблица 6.1: Официальные данные переписей населения США.

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

стабильным показателем. Таким образом, население умножалось в среднем на 1,34 каждые 10 лет и на $1{,}34^{25/10} \approx 2{,}08$ каждые 25 лет. Таким образом, с момента написания сочинения Мальтуса, почти за полвека до работы Ферхюльста, население США продолжало удваиваться каждые 25 лет. Однако Ферхюльст добавил:

> «Мы не будем настаивать на гипотезе геометрической прогрессии, учитывая, что она может протекать только в совершенно особых обстоятельствах; например, когда на плодородной территории почти неограниченного размера обитают люди с развитой цивилизацией, как это было в случае первых американских колоний.»

В своей статье Ферхюльст также вернулся к уравнению (6.1), которое он назвал «логистическим». Он заметил, что кривая $P(t)$ увеличивается с положительной кривизной (она выпуклая вниз) до тех пор, пока $P(t) < K/2$, а затем продолжает увеличиваться в направлении K, но с отрицательной кривизной (она выпуклая вверх) при $P(t) > K/2$. Таким образом, кривая имеет форму искаженной буквы S (рисунок 6.2).

Действительно,

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

Так что $d^2P/dt^2 > 0$, если $P < K/2$, и $d^2P/dt^2 < 0$, если $P > K/2$.

Ферхюльст также объяснил, как параметры r и K могут быть оценены по населению $P(t)$ за три разных, но одинаково удаленных друг от друга года. Если P_0 - это население в момент времени $t = 0$, P_1, что в момент времени $t = T$, и P_2, что в момент времени $t = 2\,T$, то утомительные вычисления, начинающиеся с уравнения

(6.2), показывают, что

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

Используя оценки численности населения Бельгии в 1815, 1830 и 1845 годах (соответственно 3,627, 4,247 и 4,801 млн. чел.), он получил $K = 6{,}584$ млн. чел. и $r = 2{,}62\,\%$ в год. Затем с помощью уравнения (6.2) он мог предсказать, что население Бельгии составит 4,998 млн. человек в начале 1851 г. и 6,064 млн. человек в начале 1900 г. (рисунок 6.2). Ферхюльст провел аналогичное исследование для Франции. Он получил $K = 39{,}685$ млн. чел. и $r = 3{,}2\,\%$ в год. Поскольку население Бельгии и Франции значительно превысило значения K, мы видим, что логистическое уравнение может быть реалистичной моделью только для периодов времени в несколько десятилетий, как в статье Ферхюльста 1838 года, но не для более длительных периодов.

В 1847 году появилось *Второе исследование о законе прироста населения*, в котором Ферхюльст отказался от логистического уравнения и выбрал вместо него дифференциальное уравнение, которое можно написать в форме $dP/dt = r\,(1 - P/K)$. Он полагал, что это уравнение будет иметь силу после того, как численность населения $P(t)$ превысит некоторое пороговое значение. Решение уравнения заключается в следующем:

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

Используя те же самые демографические данные по Бельгии, Ферхюльст заново оценил параметры r и K. На этот раз он нашел $K = 9{,}4$ миллионов для максимальной популяции. Мы видим, насколько результат может зависеть от выбора модели!

Ферхюльст стал президентом Бельгийской королевской академии в 1848 году, но на следующий год умер в Брюсселе, вероятно, от туберкулеза. Несмотря на колебания Ферхюльста между уравнениями модели, логистическое уравнение было вновь независимо введено в оборот несколько десятилетий спустя разными людьми. Робертсон использовал его в 1908 году для моделирования индивидуального роста животных, растений, человека и органов тела. МакКендрик и Кесава Пай использовали его в 1911 году для роста популяций микроорганизмов. Пирл и Рид использовали его в 1920 году для моделирования роста численности населения США, который начал замедляться. В 1922 г. Пирл, наконец, заметил работу

Ферхюльста. С тех пор логистическое уравнение вдохновляло на многие работы (см. главы 13, 20 и 24). Максимальная популяция K в конечном итоге стала известна как «ёмкость среды».

Дополнительное чтение

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 (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

Глава 7

Бьенеме, Курно и исчезновение фамилий (1845–1847)

Французский статистик Бьенеме в 1845 году понял, как вычислить вероятность исчезновения фамилии, если у каждого мужчины есть несколько сыновей в соответствии с заданным распределением вероятностей. Если среднее число сыновей меньше или равно единице, то фамилия вымрет. Если среднее число сыновей больше единицы, то вероятность исчезновения строго меньше единицы. Доказательство его результата было опубликовано два года спустя в книге, написанной его другом Курно. Эти работы были заново открыты только недавно.

Ирене-Жюль Бьенеме (Bienaymé) родился в 1796 году в Париже. Учился в Политехнической школе и сделал карьеру в Министерстве финансов, достигнув высокого уровня генерального инспектора. Под влиянием книги *Аналитическая теория вероятности*, написанной Лапласом, Бьенеме также нашел время для публикации статей о приложениях теории вероятностей во многих областях, таких как демографическая и медицинская статистика (младенческая смертность, количество рождений, ожидаемая продолжительность жизни), вероятность ошибок в правосудии, теория страхования и репрезентативность избирательных систем.

В 1845 году Бьенеме написал небольшую записку *О законе умножения и продолжительности жизни семей*, которая была опубликована в бюллетене Парижского Филоматического общества. Еще до того, ряд авторов уже писали на эту тему. Во втором издании *Очерка о законе народонаселения* (1803), Мальтус включил главу о населении Швейцарии и заметил, что

> «... в городе Берне, с 1583 по 1654 годы, суверенный совет принял в буржуазию 487 семей, 379 из которых вымерли на протяжении двух столетий, а в 1783 году их осталось только 108.»

В 1842 году Томас Дублдей в более общем плане утверждал, что дворянские или буржуазные семьи высшего класса имеют большую

Рис. 7.1:
Бьенеме (1796–1878)

склонность к исчезновению, чем семьи низшего класса. Аналогичные идеи были выдвинуты во Франции Эмилем Литтре в 1844 году в тексте введения в позитивистскую философию Огюста Конта и Бенуастоном де Шатонефом, другом Бьенеме, который в 1845 году опубликовал эссе *О продолжительности жизни дворянских родов во Франции.*

Именно в этом контексте Бьенеме пытался объяснить, как может быть так, что население страны имеет тенденцию к геометрическому росту, в то время как большое количество семей исчезает. Для решения этой проблемы, он рассмотрел упрощенный случай, когда все мужчины имеют одинаковые шансы иметь 0, 1, 2, 3,... сыновей, достигающих совершеннолетия. Точнее, он задался вопросом, какова вероятность того, что у мужчины будет потомство, носящее его фамилию через n поколений. Если среднее число сыновей меньше одного, то понятно, что эта вероятность должна стремиться к нулю по мере того, как n вырастет до бесконечности. Бенеме заметил, что тот же вывод останется верным[1], если среднее число сыновей будет точно равным единице, например, если существует вероятность $1/2$ отсутствия сына и вероятность $1/2$ рождения двух сыновей (Рис. 7.2). Но в этом случае вероятность иметь потомство в поколении n медленнее стремится к нулю: в примере это все еще будет 5 % через 35 поколений, т.е. через одиннадцать-двенадцать ве-

[1]За исключением того, случая, когда у каждого человека есть ровно один сын.

ков, если три поколения составляют одно столетие[2]. Наконец, Бьенеме, заметил, что, если среднее число сыновей больше единицы, то вымирание линии семейства не гарантировано: его вероятность можно вычислить, решив некоторое алгебраическое уравнение.

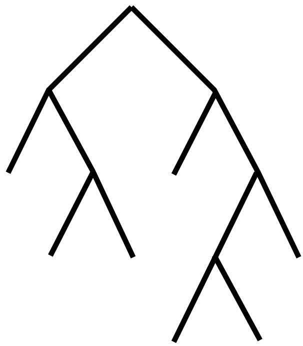


Рис. 7.2: Искусственный пример семейного древа. Предок находится на вершине дерева. В каждом поколении мужчины имеют вероятность 1/2 отсутствия сына и вероятность 1/2 рождения двух сыновей.

Статья Бьенеме не содержала других объяснений. В 1847 году его друг Антуан-Огюстен Курно (Cournot), математик и экономист, включил некоторые подробности в книгу под названием *О происхождении и пределах сходства между алгеброй и геометрией.* Он представил эту проблему в форме азартной игры, но признал, что она идентична исследованию Бьенеме, посвященному исчезновению фамилий. Если сохранить интерпретацию в терминах фамилий, то Курно сначала рассмотрел особый случай, когда у мужчин максимум два сына, предполагая p_0, p_1 и p_2, соответственно, как вероятность иметь 0, 1 или 2 сына. Очевидно, $p_0+p_1+p_2=1$. Начиная с одного предка, вероятность исчезновения после всего лишь одного поколения, обозначим ее x_1, очевидно, равна p_0. Вероятность исчезновения в течение двух поколений составляет $x_2 = p_0+p_1\,x_1+p_2\,x_1^2$: либо фамилия вымирает в первом поколении (вероятность p_0), либо в первом поколении был только один сын, у которого не было потомства мужского пола (вероятность $p_1\,x_1$), либо в первом поколении было два сына, и у обоих не было потомства мужского пола (вероятность $p_2\,x_1^2$). В общем случае, вероятность исчезновения в

[2]Как будет показано ниже, эта вероятность равна $1-x_{35}$ с $x_{n+1}=\frac{1}{2}+\frac{1}{2}\,x_n^2$ и $x_0=0$.

течение n поколений составляет

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

Действительно, если в первом поколении, например, два сына (вероятность p_2), то фамилия вымрет еще через $n-1$ поколений (т.е. в поколении n) с вероятностью $(x_{n-1})^2$. Курно заметил, что x_n - это увеличивающаяся последовательность с $x_n \leqslant 1$ для всех n. Таким образом, x_n имеет предел $x_\infty \leqslant 1$, который является решением уравнения $x = p_0 + p_1\, x + p_2\, x^2$. После подставновки $p_1 = 1 - p_0 - p_2$, это уравнение эквивалентно $0 = p_2(x-1)(x-p_0/p_2)$. Таким образом, есть два корня: $x = 1$ и $x = p_0/p_2$. В зависимости от среднего количества сыновей $p_1 + 2p_2$, которое также равно $1 - p_0 + p_2$ и которое мы обозначим $\mathscr{R}_0$, можно выделить три случая. Если $\mathscr{R}_0 < 1$, то $p_0/p_2 > 1$. Таким образом, $x = 1$ - единственное возможное значение для предела x_∞. Фамилия гарантированно вымрет. Если $\mathscr{R}_0 = 1$, то оба корня будут равны 1, и вывод будет тот же. Если же $\mathscr{R}_0 > 1$, то Курно утверждал, что x_∞ должен быть равен второму корню p_0/p_2, так как вероятность исчезновения, очевидно, должна быть равна 0 в особом случае, когда $p_0 = 0$.

Курно кратко упомянул более общий случай, когда мужчины могут иметь максимум m сыновей с вероятностью p_0, p_1,…, p_m. Аналогично предыдущему случаю, вывод зависит от значения

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

среднего количества сыновей по сравнению с 1. Уравнение для x_∞, которое имеет вид $x = p_0 + p_1\, x + \cdots + p_m\, x^m$, всегда имеет корень $x = 1$. Оно имеет еще только один положительный корень, который дает вероятность исчезновения x_∞, когда $\mathscr{R}_0 > 1$.

К сожалению, статья Бьенеме и несколько страниц книги Курно в свое время остались совершенно незамеченными. Статья была замечена только в 1970-х годах, а страницы книги - еще двадцать лет спустя! Тем временем проблема и ее решение были заново открыты другими, и тема значительно развилась. Мы вернемся к этому в главах 9, 17 и 18.

После революции 1848 года Бьенеме был вынужден уйти с работы в Министерстве финансов. Кафедра теории вероятностей в Парижском университете, для которой он, безусловно, был лучшим кандидатом, была также передана кому-то другому. Тем не менее, Бьенеме смог снова работать в Министерстве финансов после 1850 года, но ушел в отставку в 1852 году. Позднее в том же году, он был избран в Академию наук, где был специалистом в области статистики. В 1853 году он доказал неравенство, известное в современных учебниках как неравенство Чебышёва (реже - Бьенеме-Чебышёва). В 1875 году он стал президентом вновь созданного Математического общества Франции. Он умер в Париже в 1878 году.

Дополнительное чтение

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

Глава 8

Мендель и наследственность (1865)

В 1865 году Мендель опубликовал результаты своих новаторских экспериментов по гибридизации гороха. В своем анализе он использовал некоторые элементы теории вероятности. Он также рассмотрел динамическую модель популяции самоопыляюяющихся растений. Его работа, которая была заново открыта только в 1900 году, является важной вехой в истории генетики.

Иоганн Мендель (Mendel) родился в 1822 году в Моравии, которая в то время являлась частью Австрийской империи, а ныне Чехии. Второй ребенок в бедной крестьянской семье, с хорошими результатами в средней школе, но слабым здоровьем, Мендель предпочитал продолжать учебу, а не наследовать семейную ферму. Но у него не было средств для учебы в университете. Поэтому в 1843 году он поступил в аббатство Святого Фомы в Брюне (ныне Брно), где принял имя Грегора. Он изучал теологию, а также посещал некоторые курсы по сельскому хозяйству. В 1847 г. был рукоположен в священники. В течение нескольких лет он преподавал в гимназии, но не выдержал экзамены на получение звания профессора. С 1851 года, благодаря поддержке его иерархии, он все же смог продолжить свое обучение в Венском университете, где посещал лекции Кристиана Допплера по физике, а также курсы по математике, химии и естественным наукам. В 1853 году он вернулся в Брюн и преподавал физику в технической школе.

В период с 1856 по 1863 год Мендель провел серию экспериментов на большом количестве растений в саду своего аббатства. В 1865 году он представил результаты экспериментов на двух заседаниях Общества естесственных наук Брюна, членом которого он являлся. В следующем году его работа, *Эксперименты по гибридизации растений* на немецком языке была опубликована в материалах Общества. Мендель объяснил, как он пришел к изучению вариации гороха, растений, которые естественным образом размножаются путем самоопыления и семена которых могут принимать различные легкоузнаваемые формы: гладкие или морщинистые, желтые

Рис. 8.1:
Мендель (1822–1884)

или зеленые и т.д. Скрещивая растение из родословной с гладкими семенами и растение из родословной с морщинистыми семенами, он заметил, что всегда получал гибриды, дающие гладкие семена. Он назвал доминантным признаком «гладкие семена», а признак «морщинистые семена» - рецессивным. Он также показал, что доминирует признак «желтые семена», а признак «зеленые семена» - рецессивный.

Мендель затем заметил, что самоопыление растений, выращенных из гибридных семян, дали в первом поколении новые семена, которые имели либо доминантный, либо рецессивный характер в кажущихся случайными пропорциях. Кроме того, он заметил, что, многократно повторяя эксперимент, он получил в среднем семян с доминантным признаком в три раза больше, чем семян с рецессивным признаком. Например, в первом эксперименте он получил в общей сложности 5 474 гладких семян и 1 850 морщинистых семян, что дает соотношение 2,96 к 1. Второй эксперимент дал 6 022 желтых и 2 001 зеленых семян, что соотносится как 3,01 к 1 [1].

Мендель также заметил, что среди растений, выращенных из семян первого поколения с доминантным признаком, тех, которые в результате самоопыления дали семена с доминантным или рецессивным признаком, было примерно в два раза больше, чем тех,

[1]Как позднее заметил Р. А. Фишер (см. главу 14), вероятность получения экспериментальных результатов, настолько близких к теоретической величине, весьма мала. Вероятно, Мендель упорядочил свои данные. Например, во втором эксперименте относительно $n = 6\,022 + 2\,001 = 8\,023$ семян, вероятность того, что полученное в результате соотношение будет отличаться от 3 менее чем на 0,01, составляет лишь около 10 %.

которые дали семена только с доминантным признаком. Например, среди 565 растений, выращенных из гладких семян первого поколения, 372 дали как гладкие, так и морщинистые семена, в то время как 193 дали только гладкие семена; соотношение равно 1,93. Аналогично, из 519 растений, выращенных из жёлтых семян первого поколения, 353 дали и жёлтые, и зелёные семена, в то время как 166 дали только жёлтые семена; соотношение равно 2,13.

Чтобы объяснить эти результаты, Мендель предложил блестящую идею рассмотреть очевидный признак семени как результат ассоциации двух скрытых факторов, каждый из которых является либо доминантным (обозначается A) либо рецессивным (обозначается a). Таким образом, существует три возможных комбинации: AA, Aa и aa. Семена с факторами AA или Aa имеют один и тот же доминантный символ A. Семена с факторами aa имеют рецессивный символ a. Кроме того, Мендель предположил, что во время опыления пыльца и яйцеклетка (гамета) передают только один из двух факторов, каждый из которых с вероятностью $1/2$.

Таким образом, скрещивание чистых линий AA и aa дает только гибриды Aa с доминантным признаком A. Гаметы гибрида Aa передают фактор A с вероятностью $1/2$ и фактор a с вероятностью $1/2$. Поэтому самоопыление растения, выращенного из гибридного семени Aa, дает AA с вероятностью $1/4$, Aa с вероятностью $1/2$ и aa с вероятностью $1/4$, как показано в таблице 8.1.

Таблица 8.1: Возможные результаты самоопыления гибрида Aa и их вероятности как функции факторов, передаваемых мужскими гаметами (в линиях) и женскими гаметами (в столбцах).

Фактор Вероятность	A $1/2$	a $1/2$
A $1/2$	AA $1/4$	Aa $1/4$
a $1/2$	Aa $1/4$	aa $1/4$

Мендель заметил, что пропорции $AA : Aa : aa$, составляющие $1 : 2 : 1$, также могут быть получены формальным вычислением $(A + a)^2 = AA + 2\,Aa + aa$. Так как семена AA и Aa имеют доминантный признак A, в то время как только семена aa имеют рецессивный признак a, то, действительно, семян с признаком A в три раза больше, чем семян с признаком a. Кроме того, в среднем, семян Aa в два раза больше, чем семян AA. Самоопыление растений, выращенных из гибридов Aa, дает семена либо с доминантным признаком (AA или Aa) или рецессивным признаком (aa). Что касается самоопыления растений, выращенных из семян AA, то они всегда дают семена AA с доминантным признаком. Таким образом, все наблюдения объясняются.

Мендель также рассмотрел следующие поколения. Начиная с N гибридных семян Aa и предполагая для простоты, что каждое растение дает самоопылением только четыре новых вида семян, он вычислил, что среднее количество семян $(AA)_n$, $(Aa)_n$ и $(aa)_n$ в поколении n будет дано таблицей 8.2, где для наглядности презентации результаты были разделены на N.

Таблица 8.2: Последующие поколения.

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
всего	1	4	16	64	256	1 024

Эти числа просто получаются из формул

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

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

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

которые показывают, что AA дает после самоопыления четыре семени AA, что aa дает четыре семени aa и что Aa дает в среднем одно семя AA, два семени Aa и одно семя aa. Мендель также заметил, что

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

Действительно, это следует из уравнения (8.2) и из исходного условия $(Aa)_0 = 1$, что $(Aa)_n = 2^n$. Заменив это в уравнении (8.1), мы получим, что $(AA)_{n+1} = 4\,(AA)_n + 2^n$. Легко понять, что $(AA)_n = c\,2^n$ - это конкретное решение, когда $c = -1/2$. Общее решение уравнения «гомогенности» $(AA)_{n+1} = 4\,(AA)_n$

равно $(AA)_n = C\,4^n$. Наконец, добавляя эти два решения, мы видим, что $(AA)_n = C\,4^n - 2^{n-1}$ удовлетворяет начальному условию $(AA)_0 = 0$, если $C = 1/2$. Что касается последовательности $(aa)_n$, то она удовлетворяет тому же рекуррентному соотношению и тому же начальному условию, что и $(AA)_n$. Таким образом, $(aa)_n = (AA)_n$.

В заключение следует отметить, что доля гибридов Aa в общей популяции, составляющая $2^n/4^n = 1/2^n$, при каждом поколении делится на два путем самоопыления.

Работы Менделя оставались совершенно неизвестными в течение его жизни. Несколько лет спустя Мендель также пробовал проводить подобные эксперименты с другими видами растений, опубликовал несколько статей по метеорологии и исследовал наследственность пчел. Став аббатом в 1868 году, он большую часть времени занимался решением административных проблем. Он умер в 1884 году.

Лишь в 1900 году творчество Менделя было наконец-то открыто заново независимо и почти одновременно Хуго де Фризом в Амстердаме, Карлом Корренсом в Тюбингене и Эрихом фон Чермаком в Вене. Это положило бы начало новой эре в том, что мы сейчас называем генетикой.

Дополнительное чтение

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

Глава 9

Гальтон, Ватсон и проблема вымирания (1873)

В 1873 году британский статистик Гальтон и его соотечественник математик Ватсон рассмотрели проблему исчезновения фамилий, не зная работы Бьенеме. Ватсон заметил, что генерирующая функция, связанная с вероятностным распределением числа мужчин в каждом поколении, может вычисляться рекурсивно. Но он неправильно проанализировал вероятность исчезновения.

Фрэнсис Гальтон (Galton) родился в 1822 году, в том же году, что и Мендель, недалеко от Бирмингема в Англии. Он был младшим из семи детей. Его отец был богатым банкиром. Через свою мать он был кузеном Чарльза Дарвина. Гальтон начал изучать медицину в 1838 году, сначала в больнице в Бирмингеме, а затем в Лондоне. Летом 1840 года он совершил свое первое долгое путешествие по Европе до Стамбула. Затем в течение четырех лет он учился в Тринити-колледже Кембриджского университета. Но его отец умер в 1844 году, оставив значительное состояние. Гальтон отказался от идеи стать врачом. Он путешествовал в Египет, Судан и Сирию. В течение следующих нескольких лет он вёл праздный образ жизни, проводя время на охоте, путешествуя на воздушных шарах и лодках или пытаясь усовершенствовать электрический телеграф. В 1850 году он организовал исследовательскую экспедицию в Юго-Западную Африку (ныне Намибия). По возвращении в Англию в 1852 году он был избран в Королевское географическое общество. Там он мог следить за новостями экспедиций в Восточную Африку в поисках истока Нила. Он поселился в Лондоне и написал путеводитель для путешественников, который стал бестселлером. В 1856 году он был избран в Королевское общество. Затем он заинтересовался метеорологией и изобрел слово «антициклон». После публикации в 1859 году его двоюродным братом Дарвином книги *Происхождение видов*, Гальтон обратился к изучению наследственности. Он опубликовал работу *Наследственный гений* в 1869 году, в которой утверждал, что интеллектуальные способности могут передаваться по наследству.

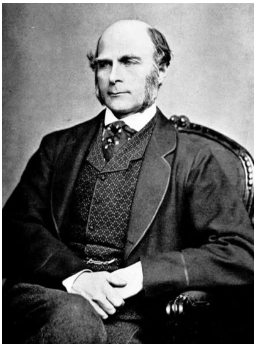


Рис. 9.1: Гальтон (слева) и Ватсон (справа)

В 1873 году швейцарский ботаник Альфонс Декандоль опубликовал книгу под названием *История науки и ученых в последние два столетия*, в которой содержалось также эссе *Влияние наследственности, изменчивости и отбора на развитие человеческого вида и на вероятное будущее этого вида*. Там он сделал следующие замечания:

> «Среди точной информации и очень здравомыслящих мнений г-на Бенуастона де Шатонефа, Гальтона и других статистиков я не увидел того важного замечания, которое они должны были сделать в отношении неизбежного исчезновения фамилий. Конечно, каждая фамилия должна исчезнуть [...] Математик мог бы вычислить, как произойдет уменьшение фамилий или титулов, зная о вероятности рождения детей женского или мужского пола и о вероятности того, что у любой данной пары не будет детей.»

Это та же проблема, которую Бьенеме изучал в 1845 году. Но Декандоль, который не знал о творчестве Бьенеме, считал, что все семьи обречены на вымирание. Гальтон заметил вышеупомянутый параграф в книге Декандоля. Поскольку он также не знал о работе Бьенеме, Гальтон поставил ее как открытую проблему для читателей *Educational Times*:

> «Проблема 4001: Большая нация, в которой мы будем заботиться только о взрослых мужчинах, N в количестве, каждый из которых носит отдельную фамилию, колонизирует округ. Их закон народонаселения таков, что в каждом поколении a_0 % взрослых мужчин не имеют детей мужского пола, которые достигают взрослой жизни; a_1 имеют одного такого ребенка мужского пола; a_2 имеют двух; и так далее до a_5, которые имеют пять.
>
> Узнайте: 1) какая доля фамилий исчезнет через r поколений; и 2) сколько будет случаев, когда носителями фамилии будет m лиц.»

Заметьте, что вторая часть проблемы не была решена Бьенеме. Гальтон не получил удовлетворительного ответа от читателей журнала и, по-видимому, не смог найти решение проблемы самостоятельно. Поэтому он попросил своего друга, математика Генри Уильяма Ватсона, попытаться решить ее.

Ватсон (Watson) родился в Лондоне в 1827 году. Его отец был офицером британского флота. Сначала он учился в Королевском колледже в Лондоне, а затем обратился к математике в Тринити-колледже Кембриджского университета, с 1846 по 1850 годы, всего через несколько лет после Гальтона. Он стал последовательным учеником Тринити-колледжа, помощником магистра в Школе Лондонского Сити, лектором по математике в Королевском Колледже и профессором математики в школе Хэрроу между 1857 и 1865 годами. Любитель альпинизма, он участвовал в экспедиции, которая достигла вершины горы Роза в Швейцарии в 1855 году. В 1856 году он был рукоположен в сан дьякона, а два года спустя - в сан англиканского священника. С 1865 года до выхода на пенсию он был приходским священником Берксвелла с Барстоном около Ковентри, - должность, которая оставляла достаточно времени для исследований.

Гальтон и Ватсон вместе написали статью под названием *О вероятности исчезновения семей*, которая была опубликована в 1875 году в журнале Королевского Антропологического Института. Гальтон представил проблему, а Ватсон объяснил свои вычисления и выводы, к которым он пришел. Они предположили, что мужчины имеют максимум q сыновей, p_k - это вероятность иметь k сыновей ($k = 0, 1, 2, \dots, q$). Другими словами, $p_k = a_k/100$, если мы используем обозначения Гальтона. Так что $p_0 + p_1 + \dots + p_q = 1$. Рассмотрим ситуацию, когда поколение 0 состоит из одного человека. Поколение 1 состоит из s мужчин с вероятностью p_s. Используя хорошо известный в свое время трюк, который задолго до Ватсона был представлен Абрахамом де Муавром, Ватсон рассматривал генерирующую функцию

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

связанную с вероятностями $p_0,\ldots,p_q$. Аналогично, пусть $f_n(x)$ будет полиномом, для которого коэффициентом при x^s является вероятность наличия s мужчин в поколении n, начиная с одного человека в поколении 0. Тогда $f_1(x) = f(x)$. Ватсон заметил, что

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

формулу, позволяющую вычислить $f_n(x)$ рекурсивно.

Действительно, положим

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

Обратите внимание, что в поколении n есть максимум q^n мужчин. Если в поколении $n-1$ есть s мужчин, пронумерованных от 1 до s, обозначим $t_1,\ldots,t_s$ количество их мужского потомства. В таком случае в поколении n будет t мужчин с вероятностью равной

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

Когда $s = 0$, следует понимать, что эта вероятность равна 1, если $t = 0$, и 0, если $t \geqslant 1$. Поэтому

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

Из этого следует, что

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

В частности, вероятность x_n вымирания фамилии в пределах n поколений равна $p_{0,n}$, что равнозначно $f_n(0)$. В качестве первого примера Ватсон взял

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

т.е. $q = 3$ и $p_0 = p_1 = p_2 = 1/3$. Он вычислил полиномы $f_n(x)$ для $n = 1, \ldots, 4$ с помощью уравнения (9.2). Он получил, например,

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

и $f_2(0) = 13/27 \approx 0{,}481$. Вычисление $f_n(x)$ для $n \geqslant 3$ становится очень утомительным, настолько утомительным, что Ватсон допустил ошибку уже для $n = 4$. Поскольку $x_5 = f_5(0) = f_4(f(0))$ может быть получено без вычисления $f_5(x)$, он получил следующий список вероятностей исчезновения $x_n = f_n(0)$:

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

Корректные значения $x_4 \approx 0{,}632$ и $x_5 \approx 0{,}677$, как можно проверить по простой формуле $x_n = f(x_{n-1})$, полученной Бьенеме. Как мы увидим в главе 17, последнюю формулу также можно вывести из уравнения (9.2).

Ватсон заметил, что у каждого человека в среднем

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

сыновей и что $\mathscr{R}_0 = 1$ в его первом примере. Таким образом, можно было подумать, что если исходное количество членов семьи мужского пола будет достаточно большим, то размер семьи останется примерно постоянным. Тем не менее, Ватсон утверждал, что вероятность вымирания x_n сходится к 1, когда $n \to +\infty$, хотя и довольно медленно. Другими словами, вся фамилия достигнет вымирания, как и предполагал Декандоль. Рисунок 9.2a, который не нарисован в оригинальной статье, и результаты Бьенеме подтверждают, что этот вывод для первого примера верен.

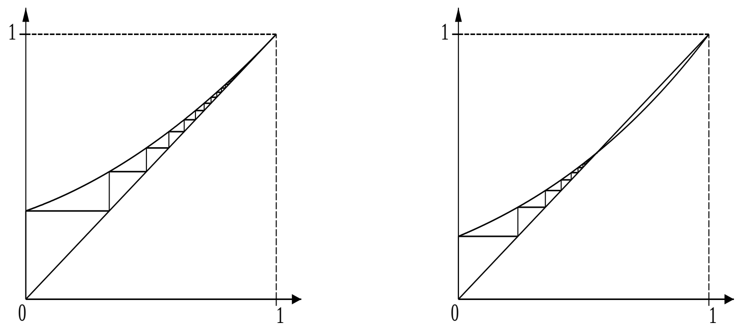


Рис. 9.2: График функций $y = f(x)$ и $y = x$. Вероятность исчезновения $x_n = f(x_{n-1})$ в пределах n поколений - это высота n-го «шага лестницы». (a): $f(x) = (1 + x + x^2)/3$. (b): $f(x) = (3 + x)^5/4^5$.

В качестве второго примера Ватсон рассмотрел биномиальное распределение вероятностей

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

для которого генерирующая функция (9.1) равна

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

Он вычислил $f_2(x)$ и $x_2 = f_2(0)$. В этот момент он понял, что $x_2 = f(x_1)$, и что $x_n = f(x_{n-1})$ для всех n. Но он думал, что эта формула верна только для специального биномиального случая (9.3). Применив ее к случаю, где $q = 5$, $a = 3$ и $b = 1$, он получил

$$x_1 \approx 0{,}237, \quad x_2 \approx 0{,}347, \quad x_3 \approx 0{,}410, \quad \ldots$$

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

Ватсон понял, что x_n должен сойтись с пределом x_∞ при $n \to +\infty$, который удовлетворяет условию

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

Он заметил, что $x = 1$ является решением этого уравнения, но не понял, что могут быть и другие решения, когда $\mathscr{R}_0 > 1$. Поэтому он ошибочно заключил, введенный в заблуждение Декандолем, что в каждом случае происходит вымирание ($x_\infty = 1$), в том числе и в только что рассмотренном числовом примере. Рисунок 9.2b показывает, что это не так!

Ватсон заметил, что среднее число сыновей в этом численном примере было больше 1 (можно показать, что $\mathscr{R}_0 = qb/(a+b) = 5/4$), то есть население имеет тенденцию к экспоненциальному росту. Но это не помогло ему обнаружить его ошибку. Он даже предположил, что для любого распределения вероятностей (p_k), т.е. не только для биномиального случая, предопределено вымирание фамилии. К этой проблеме мы вернемся в главах 17 и 18.

Гальтон продолжил свое статистическое исследование семей книгой под названием *Английские мужчины науки, их природа и воспитание*, в которой основное внимание было уделено генеалогии членов Королевского общества. Он также заинтересовался антропометрией - измерениями человеческого тела. Он воспользовался международной выставкой в 1884 году в Лондоне для сбора данных о большом количестве людей. Его результаты были опубликованы в 1889 году в книге под названием *Естественная наследственность*, в приложении к которой была воспроизведена статья, написанная в сотрудничестве с Ватсоном. В этой книге также были введены некоторые новые статистические термины, такие как «перцентиль» и «квартиль», а также слово «евгеника», т.е. улучшение человеческого вида с точки зрения наследственных признаков. После 1888 года Гальтон разработал методику распознавания отпечатков пальцев, которой через несколько лет предстояло быть использованой британской полицией. Он также продолжал изучать роль наследственности (природа) и окружающей среды (воспитание) в физических и интеллектуальных характеристиках близнецов, в размерах гороха, выращиваемого в течение нескольких поколений, или в цвете мышей, выращенных в лаборатории. Это привело его к понятию «коэффициента корреляции» между двумя переменными. В 1904 году при Университетском колледже в Лондоне была основана лаборатория Гальтона. Гальтон был посвящен в рыцари в 1909 году и умер в 1911 году.

Ватсон опубликовал несколько книг, в частности трактат по кинетической теории газов в 1876 году и трактат по математической теории электричества и магнетизма в двух томах (1885 и 1889 гг.). Он был избран в Королевское общество в 1881 году и умер в Брайтоне в 1903 году.

В 1924 году, во втором томе своей биографии Гальтона Карл, Пирсон резюмировал статью об исчезновении фамилий, не заметив ошибки. В конце концов, эта ошибка была замечена в 1930 году (см. главу 18).

Дополнительное чтение

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

Глава 10

Лотка и теория стабильного населения (1907)

В 1907 году американский химик Альфред Лотка начал изучать связь между рождаемостью, возрастной смертностью и темпом роста численности населения в модели с непрерывным временем. В 1911 году он опубликовал еще одну статью на ту же тему совместно с Ф.Р. Шарпом, в которой также рассматривались возрастные коэффициенты рождаемости. Неявное уравнение, дающее темпы роста населения, часто называют «уравнением Лотки».

Альфред Джеймс Лотка (Lotka) родился от американских родителей в 1880 году в Лемберге, который входил в состав Австро-Венгерской империи (ныне Львов в Украине). Сначала он учился во Франции и Германии, а в 1901 году получил степень бакалавра физики и химии в Бирмингемском университете в Англии. Затем он провел один год в Лейпциге, где роль термодинамики в химии и биологии была подчеркнута Вильгельмом Оствальдом, Нобелевским лауреатом по химии в 1909 году. Лотка поселился в Нью-Йорке в 1902 году и начал работать на Генеральную химическую компанию.

Рис. 10.1:
Лотка (1880–1949)

В 1907 и 1911[1] годах Лотка занялся изучением динамики возрастно-

[1]Вторая статья была написана в сотрудничестве с Ф.Р. Шарпом, математи-

структурированных популяций, не зная о работе Эйлера на эту же тему (см. главу 3). В отличие от Эйлера, он предполагал, что время и возраст являются непрерывными переменными. Пусть $B(t)$ - это интенсивность рождений мальчиков (число рождений мальчиков в единицу времени) в момент времени t, $p(x)$ - вероятность быть еще живым в возрасте x (функция дожития) и $h(x)$ - рождаемость в возрасте x: $h(x)\,dx$ - это вероятность для мужчины иметь одного новорожденного сына в возрасте от x до $x+dx$, если dx бесконечно мало. Тогда

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

это ожидаемая продолжительность жизни при рождении. Кроме того, $B(t-x)\,p(x)\,dx$ - это количество мужчин, родившихся между временем $t-x$ и $t-x+dx$, которые еще живы во время t. У этих мужчин есть $B(t-x)\,p(x)\,h(x)\,dx$ сыновей за единицу времени в момент времени t. Таким образом, интенсивность рождения мальчиков во время t равно

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

В поиске экспоненциального решения для этого интегрального уравнения в форме $B(t) = b\,e^{r\,t}$, Лотка получил делением обеих сторон на $B(t)$ уравнение

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

которое демографы теперь называют «уравнением Лотки» [2]. Эйлер получил аналогичное неявное уравнение (3.1) для скорости роста, когда время и возраст являются дискретными переменными. Лотка заметил, что правая часть (10.1) - это уменьшающаяся функция r, которая стремится к $+\infty$ при $r \to -\infty$ и к 0 при $r \to +\infty$. Таким образом, существует единственное решение уравнения (10.1) r, обозначим его r^*. Кроме того, $r^* > 0$ тогда и только тогда, когда

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

ком из Корнелльского университета.

[2]Р.А. Фишер самостоятельно пришел к этому же уравнению в 1927 г. и позже интерпретировал корень r^* как меру «дарвиновского фитнеса» в теории эволюции путем естественного отбора.

Параметр $\mathscr{R}_0$ (обознаяение было введено Дублином и Лоткой в 1925 году) - это ожидаемое количество сыновей, которое родится у одного мужчины на протяжении всей его жизни.

Лотка предположил[3], что, независимо от первоначальной возрастной структуры населения, интенсивность рождений мальчиков такова, что $B(t) \sim b\, e^{r^* t}$ при $t \to +\infty$, где b - это константа. Общая численность населения дается соотношением

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

Из этого следует, что $P(t)$ также увеличивается или уменьшается как $e^{r^* t}$, при $t \to +\infty$, т.е. асипмтотически темп роста численности населения равен r^*. Более того, возрастная структура населения, выраженная как $B(t-x)\, p(x)/P(t)$, стремится к

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

Это то, что Лотка назвал «стабильным населением»: возрастная пирамида со временем сохраняет свою форму, но общая численность населения увеличивается или уменьшается в геометрической прогрессии. Таким образом, вывод такой же, как и в модели дискретного времени Эйлера. Но в теории Лотки учитывается возрастная зависимость рождаемости. Так что она в некотором смысле более общая, чем у Эйлера.

Лотка продолжал работать над этой темой на протяжении всей своей жизни. В 1908-1909 годах он возобновил учебу в Корнельском университете, чтобы получить степень магистра. Он работал в Национальном бюро стандартов с 1909 по 1911 годы и в качестве редактора журнала *Scientific American Supplement* с 1911 по 1914 годы. В 1912 году он получил докторскую степень в Университете Бирмингема по совокупности статей, которые он опубликовал с 1907 года о динамике населения и демографии. Во время Первой мировой войны он снова работал на Генеральную химическую компанию, занимаясь проблемой получения азота из атмосферы. В 1920 году одна из его статей о биологических колебаниях (см. главу

[3]Это было строго доказано в 1941 году Вилли Феллером, который в то время был профессором математики в Брауновском университете в США. Вероятностный подход был разработан в 1968 году Крампом, Модом и Джейгером.

13) произвела глубокое впечатление на Рэймонда Перла, профессора биометрии Университета Джонса Хопкинса, который только что «открыл» логистическое уравнение (см. главу 6). В надежде найти работу в Институте медицинских исследований Рокфеллера в Нью-Йорке, Лотка работал над математическими моделями, разработанными Россом для борьбы с малярией (см. главу 12). Наконец, он получил двухлетнюю стипендию от Университета Джонса Хопкинса, что позволило ему написать книгу под названием *Элементы физической биологии*, опубликованную в 1925 году. Затем он стал руководителем исследовательского отдела Метрополитен - страховой компании по страхованию жизни в Нью-Йорке. Он сосредоточился на математическом анализе демографических вопросов и опубликовал несколько книг в сотрудничестве с коллегой, специалистом по статистике и вице-президентом компании Луи Исраэлем Дублином: *Денежная стоимость человека* (1930), *Длительность жизни* (1936) и *Двадцать пять лет прогресса в области здравоохранения* (1937). Был избран президентом Ассоциации населения Америки в 1938-1939 гг. Одна из множества его статистических моделей, «Закон Лотки» (восходит к 1926 г.) утверждает, что число авторов, написавших n статей в данной научной области, уменьшается примерно как $1/n^2$ по мере увеличения n.

Лотка также опубликовал книгу на французском языке под названием *Аналитическая теория биологических ассоциаций.* Первая часть, более философская, появилась в 1934 году. Вторая, более техническая, часть, опубликованная в 1939 году, обобщила все его исследования в области демографии человека, начиная с 1907 года. В своей книге Лотка также представил свой вклад в проблему исчезновения фамилий. После публикации в 1930 году первой статьи Штеффенсена на эту тему (см. главу 18), он применил теорию к данным, содержащимся в переписи белого населения США 1920 года. Он заметил, что наблюдаемое распределение $(p_k)_{k\geqslant 0}$ числа сыновей хорошо аппроксимируется уменьшающимся геометрическим законом для всех $k \geqslant 1$:

$$p_0 = a, \quad p_k = b\, c^{k-1}\ (k \geqslant 1),$$

с $a = 0{,}4825$, $b = 0{,}2126$ и $c = 1-b/(1-a)$. Таким образом, $\sum_{k\geqslant 0} p_k = 1$. Связанная с этим генерирующая функция имеет вид

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

Два решения уравнения $x = f(x)$ - это $x = 1$ и $x = a/c$. Вероятность исчезновения x_∞ является наименьшим из этих двух решений (см. главу 7). При численных значениях для США, он нашел $x_\infty \approx 0{,}819$, в то время как среднее число сыновей составило $\mathscr{R}_0 = f'(1) = (1-a)^2/b \approx 1{,}260$. Несмотря на то, что среднее число детей (включая сыновей и дочерей) близко к 2,5, вероятность исчезновения фамилии выше 80 %.

Лотка был избран президентом Американской статистической ассоциации в 1942 году. Он ушел на пенсию в 1947 году и умер в 1949 году в Нью-Джерси. В 1956 году вышло новое издание его книги 1925 года под немного другим названием *Элементы математической биологии.*

Дополнительное чтение

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, S. 512. Scribner, New York (1981)
6. Lotka, A.J.: Relation between birth rates and death rates. *Science* 26, 21–22 (1907) → Smith & Keyfitz (1977).
7. Lotka, A.J.: *Théorie analytique des associations biologiques*, 2$^{\text{e}}$ partie. Hermann, Paris (1939) gallica.bnf.fr
8. Sharpe, F.R., Lotka, A.J.: A problem in age-distribution. *Philos. Mag. Ser.* 6, 21, 435–438 (1911) → Smith & Keyfitz (1977).
9. Smith, D.P., Keyfitz, N.: *Mathematical Demography.* Springer, Berlin (1977)
10. Tanner, A.: *Von Molekülen, Parasiten und Menschen – A. J. Lotka und die Mathematisierung des Lebens.* ETH Zürich (2014) doi:10.3929/ethz-a-010209129

Глава 11

Закон Харди-Вайнберга (1908)

В 1908 году британский математик Харди и немецкий врач Вайнберг независимо обнаружили, что в бесконечно большой популяции, которая размножается случайным образом по законам Менделя, частоты генотипов, полученных из двух аллелей, остаются постоянными на протяжении поколений. Их математическая модель была одной из отправных точек для популяционной генетики.

Годфри Гарольд Харди (Hardy) родился в 1877 году в Суррее в Англии. Его родители были учителями. Он изучал математику в Тринити-колледже в Кембриджском университете с 1896 года, стал научным сотрудником своего колледжа в 1900 году и преподавателем математики в 1906 году. После выхода своей первой книги *Интегрирование функций одной переменной* (1905 г.), он опубликовал в 1908 г. *Курс чистой математики*, который был много раз переиздан и переведен на многие иностранные языки.

Рис. 11.1:
Харди (1877–1947)

В то время повторное открытие творчества Менделя вызвало некоторые сомнения. Некоторые биологи задавались вопросом, почему доминирующие особи не встречались все чаще из поколения в поколение. Реджинальд Паннетт, написавший в 1905 году книгу под названием *Менделизм*, задал этот вопрос Харди, с которым он играл в крикет в Кембридже. Харди привел свое решение в статье *Менделевы пропорции в смешанном населении*, которая была опубликована в 1908 году. Чтобы упростить анализ, он рассмотрел большую популяцию, в которой выбор сексуального партнера будет случайным. Более того, он ограничил свое внимание только двумя факторами (или «аллелями») A и a, A является доминирующим и a рецессивным. Для поколения n пусть p_n будет частотой «генотипа» AA, $2\,q_n$ - частотой Aa и r_n - частотой aa. Легко видеть, что $p_n + 2\,q_n + r_n = 1$. Харди также предположил, что ни один из этих генотипов не приводит к избыточной смертности или снижению фертильности по сравнению с двумя другими генотипами. Частоты в поколении $n+1$ можно легко вычислить, заметив, что один случайно выбранный индивидуум в поколении n передает аллель A с вероятностью $p_n + q_n$: либо генотип равен AA и аллель A передается с вероятностью 1, либо генотип равен Aa и аллель A передается с вероятностью 50%. Аналогично, с вероятностью $q_n + r_n$ передается аллель a. Таким образом, можно построить таблицу 11.1 так же, как и таблицу 8.1.

Таблица 11.1: Вычисление частот генотипов в поколении $n+1$ в зависимости от частот аллелей родителей (строки для матери, колонки для отца).

Аллель Частота	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 $(q_n + r_n)^2$

Частоты генотипов AA, Aa и aa в поколении $n+1$ составляют, соответственно, p_{n+1}, $2q_{n+1}$ и r_{n+1}. Итак, Харди нашел, что

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

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

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

Затем он исследовал, при каких условиях частоты генотипов могут оставаться постоянными через поколения, равными p, $2q$ и r. Так как по определению $p + 2q + r = 1$, то мы видим, что все уравнения (11.1)-(11.3) дают одно и то же условие $q^2 = p\,r$.

Например, первое уравнение дает $p = (p + q)^2 = p^2 + 2pq + q^2$, что эквивалентно $p(1 - p - 2q) = q^2$ и, наконец, $p\,r = q^2$.

Начиная с произвольных начальных условий $(p_0, 2q_0, r_0)$ с $p_0 + 2q_0 + r_0 = 1$, Харди заметил, что $q_1^2 = (p_0 + q_0)^2(q_0 + r_0)^2 = p_1\,r_1$. Поэтому состояние $(p_1, 2q_1, r_1)$ уже является равновесием. Поэтому $(p_n, 2q_n, r_n)$ остается равным $(p_1, 2q_1, r_1)$ для всех $n \geqslant 1$. Если для частоты аллели A в поколении 0 установить $x = p_0 + q_0$, то $1 - x = q_0 + r_0$ - это частота аллели a. Используя систему (11.1)–(11.3) еще раз получаем

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

при всех $n \geqslant 1$ (рис. 11.2).

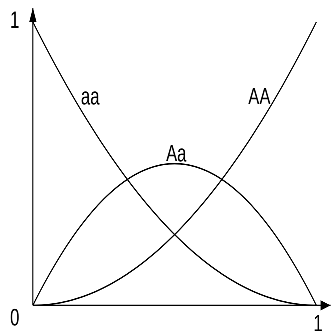


Рис. 11.2: Графики функций x^2, $2\,x\,(1-x)$ и $(1-x)^2$, соответствующие равновесным частотам генотипов AA, Aa и aa.

В заключение, приведенные выше гипотезы приводят к закону, согласно которому частоты генотипов AA, Aa и aa остаются неизменными на протяжении поколений. Теория Менделя не приводит к постепенному увеличению частоты доминирующего признака, как считалось вначале.

Несколько лет спустя Фишер почеркивал важное следствие этого закона: в первом приближении (т.е. в предположении, что гипотезы модели реалистичны) популяция сохраняет постоянную генетическую дисперсию. Это наблюдение решает одну из проблем, поднятых Дарвиновской теорией эволюции путем естественного отбора. Действительно, Дарвин, как и его современники, считал, что в каждом поколении физиологические характеристики детей являются своего рода усредненными характеристиками двух родителей, каждый из которых вносит свой вклад. Позже эта идея была тщательно изучена с использованием статистических данных Фрэнсиса Галтона и его преемника в лаборатории биометрии Карла Пирсона. Если бы это было так, то дисперсия этих характеристик в популяции должна была бы делиться на две в каждом поколении, и вскоре была бы настолько однородна, что естественный отбор, призванный объяснить эволюцию, был бы невозможен. Тем не менее, для того, чтобы этот механизм усреднения был отвергнут, понадобилось еще несколько лет, поскольку биологи, защищающие точку зрения Дарвина, не желали признавать, что законы Менделя неизбежны для понимания эволюции.

После этой работы в 1908 году Харди вернулся к чистой математике. В своей автобиографии *Извинения математика* он даже с гордостью утверждал, что избегал открытий, имеющих практическое значение. В 1910 он был избран в Королевское общество. В 1913 году он обнаружил индийского вундеркинда Рамануджана и пригласил его работать в Кембридже. После Первой мировой войны он стал профессором Оксфордского университета и продолжил плодотворное сотрудничество со своим соотечественником Литтлвудом. В 1931-1942 гг. он снова стал профессором в Кембридже. Он опубликовал много книг, часто в сотрудничестве: *Orders of Infinity* (1910), *Общая теория рядов Дирихле* с Марселем Рисом (1915), *Неравенство* с Литтлвудом и Pólya (1934), *Введение в теорию чисел* с Э. М. Райтом (1938), *Рамануджан* (1940), *Ряд Фурье* с Рогосинским (1944) и *Divergent Series.* (1949). Умер в Кембридже в 1947 году.

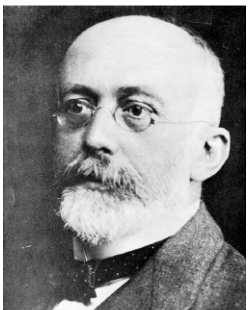


Рис. 11.3:
Вайнберг (1862–1937)

Несколько десятилетий спустя было обнаружено, что закон Харди о частотах генов был также выведен в том же 1908 году немецким врачом Вильгельмом Вайнбергом. Вайнберг (Weinberg) родился в Штутгарте в 1862 году. После учебы в Тюбингене и Мюнхене до получения докторской степени по медицине он проработал несколько лет в больницах Берлина, Вены и Франкфурта. В 1889 году он поселился в Штутгарте в качестве терапевта и акушера. Несмотря на то, что он был очень занят своей работой, он нашел время, чтобы написать много статей в немецких научных журналах. В 1901 году он изучал со статистической точки зрения частоту появления близнецов одного пола. Статья 1908 года, в которой он объяснял тот же закон, который нашел Харди, была опубликована в местном научном журнале и не была замечена. Но в отличие от Харди, он продолжил это исследование в последующие годы, открыв, например, обобщение, в котором есть более двух аллелей. Он также внес свой вклад в область медицинской статистики. Вайнберг умер в 1937 году. После открытия заново его статьи 1908 года генетики назвали закон стабильности частот генотипа «законом Харди-Вайнберга».

В настоящее время этот закон часто используется следующим образом. Если редкая рецессивная аллель a не влияет на выживание или фертильность, и если мы знаем частоту x^2 генотипа aa, потому что aa производит определенный фенотип, то мы можем вычислить x и оценить частоту $2x(1-x) \approx 2x$ генотипа Aa. В качестве примера, если частота aa составляет $1/20\,000$, то мы получим $x \approx 1/140$. Так что $2x \approx 1/70$ - это частота генотипа Aa. Рецессивный аллель a, который может показаться очень редким при осмотре фенотипов, на самом деле не так уж и редок.

Дополнительное чтение

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 (1940). archive.org
3. Punnett, R.C.: *Mendelism*, 2nd edn. Cambridge (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

Глава 12

Росс и малярия (1911)

В 1911 году британский врач Рональд Росс, уже получивший в 1902 году Нобелевскую премию за свою работу по малярии, изучил систему дифференциальных уравнений, моделирующих распространение этой болезни. Он показал, что малярия может сохраняться только в том случае, если количество комаров превышает определенный порог. Поэтому для искоренения малярии не обязательно убивать всех комаров - достаточно убить лишь определенную фракцию. Подобные модели эпидемии впоследствии были разработаны Кермаком и Мак-Кендриком.

Рональд Росс (Ross) родился в 1857 году на севере Индии, где его отец был офицером британской армии. Он изучал медицину в Лондоне, но предпочитал писать стихи и драмы. Проработав год на корабле хирургом, в 1881 году он поступил на индийскую медицинскую службу. Его медицинская работа в Индии оставила ему много свободного времени, во время которого он писал литературные произведения и изучал математику. В 1888 году, в отпуске в Англию он получил диплом в области общественного здравоохранения и изучал бактериологию - новую науку, созданную несколькими годами ранее Пастером и Кохом. Вернувшись в Индию, Росс начал изучать малярию. Во время второго отпуска в 1894 году он встретился в Лондоне с Патриком Мэнсоном, специалистом по тропической медицине, который показал ему под микроскопом то, что заметил в 1880 году французский военный врач Альфонс Лаверан: в крови больных малярией содержатся паразиты. Мэнсон предположил, что паразиты могут появляться от комаров, поскольку он обнаружил в Китае паразита от другой тропической болезни (филяриатоза) у этих насекомых. Однако он считал, что люди заражаются паразитами при питье воды, зараженной комарами. С 1895 по 1898 год Росс продолжал свои исследования в Индии и проверял идею Мэнсона. В 1897 г. он обнаружил в желудке некоторых видов комаров, которых он не изучал до этого (анофели), некоторые паразиты, похожие на тех, что наблюдал Лаверан. Его начальство отправило его в Калькутту в сезон, когда случаи малярии были редки, и он решил изучить малярию у птиц в клетках. Он обнаружил этого паразита в слюнных железах комаров-анофелей и сумел заразить экспериментально здоровых птиц, позволив комарам укусить их: это доказало, что малярия передается при укусах комаров, а не при поступлении в организм загрязненной воды. В 1899 году Росс покинул индийскую медицинскую службу, чтобы преподавать в Ливерпульской школе тропической медицины, которая была создана за год до этого. Он был избран в Королевское общество в 1901 году и получил в 1902 году Нобелевскую премию по физиологии или медицине за свою работу по борьбе с малярией. Он ездил в Африку, на Маврикий и в Средиземноморье, чтобы популяризировать борьбу с комарами. Метод был успешно применен в Египте вдоль Суэцкого канала, вдоль строящегося Панамского канала, на Кубе и в Малайзии. В некоторых других областях он оказался менее успешным. Росс опубликовал *Доклад о профилактике малярии на Маврикии* в 1908 году и *Профилактика малярии* в 1910 году.

Рис. 12.1:
Росс (1857–1932)

Несмотря на свои доказательства роли некоторых комаров в передаче малярии, Росс встретил скептицизм, когда заявил, что малярию можно искоренить, просто сократив число комаров. Во втором издании своей книги *Профилактика малярии*, опубликованной в 1911 году, он попытался построить математические модели передачи малярии в поддержку своего заявления. Одна из его моделей состояла из системы двух дифференциальных уравнений. Введем следующие обозначения:

- N: общее население в данной местности;
- $I(t)$: число людей, зараженных малярией во время t;

- n: общая численность комаров (предполагаемая постоянной);
- $i(t)$: количество комаров, заражённых малярией;
- b: частота укусов комаров;
- p (соответственно p'): вероятность передачи малярии от человека к комару (соответственно от комара к человеку) при одном укусе;
- a: скорость выздоровления человека от малярии;
- m: смертность москитов.

В течение небольшого промежутка времени dt каждый зараженный комар кусает $b\,dt$ людей, среди которых доля, равная $\frac{N-I}{N}$ еще не заражена. Учитывая вероятность передачи p', есть $b\,p'\,i\,\frac{N-I}{N}\,dt$ новых инфицированных людей. За тот же промежуток времени количество людей, которые выздоравливают, составляет $a\,I\,dt$. Следовательно,

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

Точно так же, каждый незараженный комар кусает $b\,dt$ людей, среди которых доля, равная I/N, уже инфицирована. С учетом вероятности передачи p, есть $b\,p\,(n-i)\,\frac{I}{N}\,dt$ новых зараженных комаров. Между тем, если предположить, что инфекция не влияет на смертность, то количество умирающих комаров составляет $m\,i\,dt$. Итак,

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

Поскольку малярия существует постоянно в большинстве инфицированных стран, Росс рассматривал только устойчивые состояния своей системы из двух уравнений: число инфицированных людей $I(t)$ и число инфицированных комаров $i(t)$ остаются постоянными во времени ($dI/dt = 0$ и $di/dt = 0$). Во-первых, всегда есть устойчивое состояние с $I = 0$ и $i = 0$, что соответствует отсутствию малярии. Во-вторых, Росс искал устойчивое состояние с $I > 0$ и $i > 0$ и обнаружил, что

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

Разделяя уравнения устойчивого состояния на произведение $I \times i$, получим линейную систему двух уравнений с двумя неизвестными $1/I$ и $1/i$,

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

Его решение легко получить.

Можно заметить, что $I > 0$ и $i > 0$, если количество комаров превышает критический порог:

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

В этом случае устойчивое состояние соответствует ситуации, когда болезнь является эндемической, т.е. постоянно присутствует. Росс пришел к выводу, что если количество комаров n снижается ниже критического порога n^*, то единственным оставшимся устойчивым состоянием является $I = 0$ и $i = 0$, поэтому малярия должна исчезнуть. В частности, для искоренения малярии не обязательно уничтожать всех комаров. Именно этот момент Росс хотел подчеркнуть своей моделью.

Для иллюстрации своей теории Росс искал разумные числовые значения параметров модели. Он предполагал, что

- смертность комаров такова, что только треть из них жива через десять дней; поэтому $e^{-10\,m} = \frac{1}{3}$ и $m = (\log 3)/10$ в день;
- половина людей все еще инфицирована через три месяца; так что $e^{-90\,a} = 1/2$ и $a = (\log 2)/90$ в день;
- один из восьми комаров кусают за день; так что $e^{-b} = 1 - 1/8$ и $b = \log(8/7)$ в день.
- Зараженные комары обычно не заразны в течение первых десяти дней после заражения, так как паразиты должны пройти через несколько стадий трансформации. Поскольку треть комаров может выжить десять дней, Росс предположил, что есть также около трети всех зараженных комаров, которые являются инфекционными: $p' = 1/3$;
- $p = 1/4$.

Затем Росс мог вычислить по формуле (12.1) зараженную фракцию I/N в человеческой популяции в зависимости от соотношения

n/N между комаром и человеческой популяцией. Он показал свои результаты в таблице, эквивалентной рисунку 12.2.

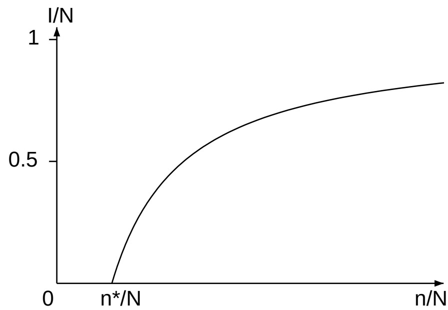


Рис. 12.2: Доля I/N инфицированных людей в зависимости от соотношения n/N между комаром и человеческой популяцией.

Форма кривой показывает, что доля инфицированных людей уже выше, чем 50 %, если соотношение n/N чуть выше критического значения n^*/N. Но эта доля не сильно меняется, когда соотношение n/N еще больше увеличивается. Это объясняет, почему корреляция между числом комаров и наличием малярии никогда раньше не была замечена. Росс, однако, заметил, что числовое значение порога n^*/N очень чувствительно к небольшим изменениям коэффициента частоты укусов b, но это не изменило общую форму кривой на рисунке 12.2. Его качественное объяснение более важно, чем количественные результаты, которые страдают от неопределенности числовых значений параметров.

Для интерпретации критического порога n^*, обнаруженного Россом[1], рассмотрим одного инфицированного человека, введенного в популяцию, тогда как остальные люди и комары не заражены. Этот человек остается инфицированным в среднем в течение периода времени, равного $1/a$. Он получает $b\,n/N$ укусов за единицу времени, так что в среднем $b\,n/(a\,N)$ укусов в общей сложности в течение периода времени, когда он является инфицированным. Таким образом, он заражает в среднем $b\,p\,n/(a\,N)$ комаров. Каждый из этих зараженных комаров живет в среднем период времени, равный $1/m$, кусает b/m людей и заражает $b\,p'/m$ людей. В целом, после передачи заболевания от первого инфицированного человека к

[1]Эта интерпретация была подчеркнута лишь спустя долгое время после работы Росса.

москитам и от этих москитов к другим людям, среднее число вновь инфицированных людей является произведением двух предыдущих формул, т.е.

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

Здесь $\mathscr{R}_0$ - количество вторичных случаев заражения человека, вызванных первичным заражением. Таким образом, процесс заражения, происходящий непрерывно во времени, может рассматриваться и через последующие поколения. Малярия может «заразить» население только в том случае, если $\mathscr{R}_0 > 1$. Это условие точно эквивалентно $n > n^*$.

В заключение Росс в более общем плане высказался в пользу математического моделирования в эпидемиологии:

> «В сущности, вся эпидемиология, интересующаяся развитием заболеваний в зависимости от времени и от места, должна рассматриваться с математической точки зрения для того, чтобы считаться наукой. Сказать, что болезнь зависит от определенных факторов, будет не слишком полезно до тех пор, пока мы не сможем оценить, насколько сильно каждый фактор влияет на весь результат. А математический метод, на самом деле, является ничем иным, как применением тщательной аргументации к рассматриваемым проблемам.»

Росс был посвящен в рыцари в 1911 году. Он переехал в Лондон и стал консультантом британской армии во время Первой мировой войны. В 1923 году он опубликовал свою автобиографию, *Мемуары с полным отчетом о большой проблеме малярии и ее решении.* В 1926 г. был открыт Институт тропических болезней Росса (в настоящее время входит в состав Лондонской школы гигиены и тропической медицины), директором которого он стал. Росс умер в Лондоне в 1932 году.

Дополнительное чтение

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. (1911) archive.org
3. Ross, R.: *Memoirs.* John Murray, London (1923) archive.org
4. Rowland, J.: *The Mosquito Man.* Roy Publishers (1958)

Глава 13

Лотка, Вольтерра и система хищник-жертва (1920–1926)

В 1920 г. Альфред Лотка изучил модель «хищник-жертва» и показал, что численность популяций может постоянно колебаться. Он разработал это исследование в своей книге 1925 года *Элементы физической биологии*. В 1926 году итальянский математик Вито Вольтерра случайно заинтересовался той же моделью, чтобы ответить на вопрос биолога Умберто д'Анкона: почему во время Первой мировой войны, когда промысловое усилие было низким, в Адриатическом море рыбаки поймали больше хищников?

В 1920 году Лотка опубликовал статью под названием *Аналитические записки об определенных ритмических отношениях в органических системах*. К тому времени уже в течение нескольких лет его интересовали некоторые химические реакции, проявляющие в лабораторных экспериментах необычные переходные колебания. Цель его статьи состояла в том, чтобы показать, что система, состоящая из двух биологических видов, может испытывать постоянные колебания. В качестве примера он привел популяцию травоядных, питающихся растениями. По аналогии с уравнениями, используемыми в химической кинетике, пусть $x(t)$ - суммарная масса растений и $y(t)$ - суммарная масса травоядных в момент времени t. В качестве модели Лотка использовал следующую систему дифференциальных уравнений

$$\frac{dx}{dt} = a\,x - b\,x\,y\,, \tag{13.1}$$

$$\frac{dy}{dt} = -c\,y + d\,x\,y\,, \tag{13.2}$$

где a, b, c и d - положительные параметры. Параметр a - это скорость роста растений при отсутствии травоядных, c - это скорость уменьшения популяции травоядных при отсутствии растений. Выражения $-b\,x\,y$ и $d\,x\,y$ означают, что чем больше животных и растений, тем выше массовый переход от растений к животным (перенос включает в себя некоторую потерю массы, $d \leqslant b$). Полагая $dx/dt = 0$ и $dy/dt = 0$, Лотка заметил, что существуют два устойчивых состояния:

- $(x = 0,\ y = 0)$, Популяция травоядных вымерла, и растений больше нет;
- $(x = c/d,\ y = a/b)$, сосуществуют травоядные и растения.

Он также написал без доказательств, что если в момент времени $t = 0$, $(x(0), y(0))$ не является одним из этих двух устойчивых состояний, то функции $x(t)$ и $y(t)$ периодически колеблются: есть число $T > 0$ такое, что $x(t + T) = x(t)$ и $y(t + T) = y(t)$ для всех $t > 0$ (рисунок 13.1)[1]. Если, например, растений очень много, то популяция травоядных увеличится, что приведет к уменьшению общей массы растений. Когда этой массы становится недостаточно для того, чтобы накормить травоядных, некоторые животные умирают от голода, и общая масса растений снова начнет расти до тех пор, пока не достигнет уровня, равного их первоначальному значению. Этот процесс повторяется.

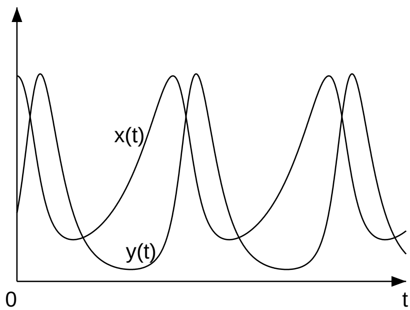


Рис. 13.1: Колебания общей массы растений $x(t)$ и общей массы травоядных $y(t)$ как функции времени.

Лотка более подробно изучил эту модель во второй статье, опубликованной в 1920 году и озаглавленной «Незатухающие колебания, вытекающие из закона действующих масс». Он объяснил, почему система может периодически колебаться. Это вытекает из того,

[1] Период T зависит от исходных условий, но Лотка осознал этот факт только в 1925 году.

что точка $(x(t), y(t))$ должна оставаться на замкнутой траектории в плоскости с x на горизонтальной оси и y на вертикальной оси; точнее, в квадранте, где $x \geqslant 0$ и $y \geqslant 0$ (рисунок 13.2).

Действительно, деля уравнение (13.1) на уравнение (13.2), мы получаем после некоторых преобразований:

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

В результате интегрирования получаем равенство

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

где K - это константа, которая зависит только от исходного состояния. Следовательно, точка $(x(t), y(t))$ остается на замкнутой кривой $d\,x-c\,\log x=a\,\log y-b\,y+K$ (рисунок 13.2).

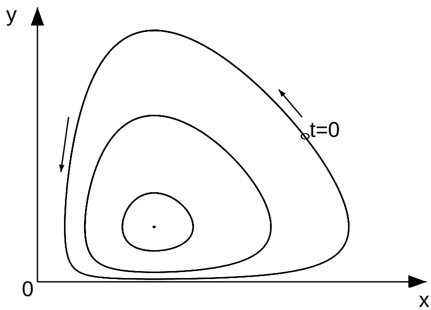


Рис. 13.2: Диаграмма с общей массой растений $x(t)$ по горизонтальной оси и общей массой травоядных $y(t)$ по вертикальной оси. Три замкнутые кривые вокруг устойчивого состояния соответствуют различным исходным условиям.

Траектория $(x(t), y(t))$ вращается вокруг устойчивого состояния $(c/d, a/b)$ против часовой стрелки, что легко видно по знаку dx/dt и dy/dt. Рядом с устойчивым состоянием система подвергается небольшим колебаниям с периодом равным $2\pi/\sqrt{a\,c}$.

Действительно, положим $x=\frac{c}{d}+x^*$ и $y=\frac{a}{b}+y^*$ где $|x^*|\ll\frac{c}{d}$

и $|y^*| \ll \frac{a}{b}$. Тогда

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

Из этих двух уравнений мы получаем

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

Эти уравнения получаются такими же, как и для колебаний простого маятника в физике. Период равен $2\pi/\sqrt{a\,c}$.

Рэймонд Перл, который передал первую статью 1920 года в журнал *Proceedings of the National Academy of Sciences*, помог Лотке получить двухлетнюю стипендию от Университета Джона Хопкинса для написания книги под названием *Элементы физической биологии.* Книга была опубликована в 1925 году. В разделе, подводящем итоги работы 1920 года, упоминалось, что системы двух видов, хозяин-паразит и хищник-жертва могут быть описаны одной и той же моделью (13.1)–(13.2). К сожалению, книга Лотки не привлекла к себе особого внимания, когда она была опубликована. Однако известный математик Вольтерра вскоре после этого самостоятельно открыл для себя эту же модель при изучении проблемы рыболовства.

Вито Вольтерра (Volterra) родился в еврейском гетто Анконы в 1860 году незадолго до объединения Италии, когда город еще принадлежал папским государствам. Он был единственным ребенком. Его отец, торговец тканями, умер, когда Вито было два года, и оставил семью без денег. Хороший ученик средней школы, Вольтерра сумел продолжить обучение, несмотря на бедность, сначала в университете Флоренции, а затем в Scuola Normale Superiore в Пизе. В 1882 году он получил докторскую степень по физике, а на следующий год стал профессором механики в Пизанском университете. Он присоединился к университету Турина в 1892 году и переехал на кафедру математической физики в университете La Sapienza в Риме в 1900 году. Он стал сенатором в 1905 году. Многие из лекций, которые он дал в Риме или в зарубежных университетах были опубликованы в виде книги: *Три лекции о некоторых последних достижениях в области математической физики* (Университет Кларка, 1909 г.), *Лекции по интегральным и интегрально-дифференциальным уравнениям* (Рим, 1910), *Лекции по линейным функциям* (Париж, 1912 г.), *Теория перестановочных функций* (Принстон, 1912). Во время Первой мировой войны он служил офицером в итальянской армии и возглавлял бюро военных изобретений. После войны активно участвовал в создании Итальянского математического союза (1922 г.) и Итальянского национального исследовательского совета (1923 г.), став первым председателем последнего. Он также стал президентом Международной комиссии по научным исследованиям Средиземного моря (1923 г.) и президентом Академии деи Линчеи (1924 г.). Другая монография, написанная в сотрудничестве с Ж. Пересом, *Лекции по композиции и пермутирующим функциям*, была опубликована в 1924 году.

Рис. 13.3: Вольтерра (1860-1940) получил почетную докторскую степень в Кембриджском университете в 1900 году.

В 1925 году, в возрасте 65 лет Вольтерра заинтересовался исследованием зоолога Умберто Д'Анконы, который впоследствии стал его зятем, о доле хрящевых рыб (таких как акулы и скаты), высаживавшихся в рыбных хозяйствах в течение 1905-1923 годов в трех гаванях Адриатического моря: Триест, Фиуме[2] и Венеция. Д'Анкона заметил, что доля этих рыб увеличилась во время Первой мировой войны, когда промысел рыбы сократился (таблица 13.1).

Хрящевая рыба является хищницей мелких рыб, и снижение промысла рыбы благоприятствует хищникам. Вольтерра, который не знал о работе Лотки, объяснил это наблюдение использованием

[2]Теперь Риека в Хорватии.

Таблица 13.1: Доля хрящевой рыбы в рыбном промысле в Триесте, Фиуме и Венеции до, во время и после Первой мировой войны.

год	1910	1911	1912	1913	1914	1915	1916
Триест	5,7	8,8	9,5	15,7	14,6	7,6	16,2
Фиуме	-	-	-	-	11,9	21,4	22,1
Венеция	21,8	-	-	-	-	-	-

год	1917	1918	1919	1920	1921	1922	1923
Триест	15,4	-	19,9	15,8	13,3	10,7	10,2
Фиуме	21,2	36,4	27,3	16,0	15,9	14,8	10,7
Венеция	-	-	30,9	25,3	25,9	25,8	26,6

той же модели:

$$\frac{dx}{dt} = a\,x - b\,x\,y\;,\quad \frac{dy}{dt} = -c\,y + d\,x\,y\,,$$

где $x(t)$ обозначает количество добычи, а $y(t)$ - количество хищников. Он заметил, как и Лотка, что эта система может периодически колебаться с периодом T, который зависит от начального условия (x_0, y_0). Он также заметил, что

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

Интегрируя за один период T (так что $x(0) = x(T)$ и $y(0) = y(T)$), он получил

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

Таким образом, среднее значение за один период как количества добычи, так и количества хищников не зависит от исходных условий. Более того, если промысловое усилие снижается, то темп прироста добычи в размере a увеличивается, в то время как темп вымирания хищников в размере c уменьшается. Поэтому среднее значение $x(t)$ уменьшается, а среднее значение $y(t)$ увеличивается: увеличивается доля хищников. Именно это наблюдалось в отношении статистики рыболовства в Адриатическом море.

Вольтерра впервые опубликовал свою статью на итальянском языке в 1926 году. Английское резюме появилось несколько месяцев спустя в *Nature*. Лотка проинформировал Вольтерру и других ученых о своих работах по изучению систем хищник-добыча, но его статья 1920 года и книга 1925 года не всегда цитировались. Тогда Лотка уже работал в страховой компании, поэтому его работа была посвящена демографии. Вольтерра продолжал работать над вариантами системы «хищник-жертва» в течение десятилетия. В 1928-1929 годах он читал серию лекций в недавно созданном Институте Анри Пуанкаре в Париже. Конспекты этих лекций были опубликованы в 1931 г. под названием *Уроки математической теории борьбы за жизнь*. В 1935 г. Вольтерра совместно с Умберто Д'Анконой опубликовал еще одну книгу под названием *Биологические ассоциации с математической точки зрения*.

Хотя модель «хищник-жертва» правильно объясняет данные о рыбном промысле, дебаты о реализме упрощенных моделей в экологии только начинаются и до сих пор являются предметом научного спора. В настоящее время модель «хищник-жертва» также известна как модель Лотки-Вольтерры и является одной из наиболее часто упоминаемых в экологии.

В 1931 году Вольтерра отказался поддерживать власть Муссолини. Он потерял профессуру в университете в Риме и был исключен из итальянских научных академий, одним из самых известных членов которых он был. С тех пор он оставался в основном за пределами Италии, путешествовал по Европе и читал лекции. Совместно с Ж. Пересом он опубликовал первый том книги *Общая теория функций* (1936) и книгу с Б. Хостинским *Бесконечно малые линейные операции* (1938). Он умер в Риме в 1940 году.

Дополнительное чтение

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*, S. 283–285. University of Chicago Press (1991)
10. Volterra, V.: *Leçons sur la théorie mathématique de la lutte pour la vie.* Gauthier-Villars, Paris (1931)
11. Volterra, V., D'Ancona, U.: *Les Associations biologiques au point de vue mathématique.* Hermann, Paris (1935)
12. Whittaker, E.T.: Vito Volterra 1860–1940. *Obit. Not. Fellows R. Soc.* 3, 690–729 (1941)

Глава 14

Фишер и естественный отбор (1922)

В 1922 году британский математический биолог Рональд Фишер опубликовал существенную работу о генетике популяции. В этой главе рассматривается только один раздел статьи, который посвящен варианту модели Харди-Вайнберга, включая естественный отбор. Фишер показал, что если предпочтение отдается гетерозиготе, то могут сосуществовать обе аллели. Если предпочтение отдается одной из двух гомозигот, то вторая аллель исчезает. Основной проблемой является объяснение того, почему у некоторых генов может быть несколько аллелей.

Рональд Эйлмер Фишер (Fisher) родился в Лондоне в 1890 году последним из шести детей. Его отец был аукционистом, но позже объявил о банкротстве. Фишер изучал математику и физику в Гонвилле и Кайус колледже Кембриджского университета между 1909 и 1913 годами. Генетика в то время быстро развивалась. Начиная с 1911 года Фишер участвовал в заседаниях Общества евгеники, инициированных Галтоном. Он начал фокусироваться на статистических проблемах, связанных с работой Галтона и Менделя. После окончания учебы в университете он провел одно лето, работая на ферме в Канаде, а затем работал в компании Mercantile and General Investment Company в лондонском Сити. Из-за своей крайней близорукости он не смог принять участие в Первой мировой войне, несмотря на то, что был добровольцем. Он провел эти годы, преподавая в средних школах. В свободное время он заботился о ферме и продолжал свои исследования. Он получил новые важные результаты, связывающие коэффициенты корреляции с генетикой Менделя. В 1919 году он начал работать в качестве статистика на экспериментальной станции Ротамстед, которая занималась сельским хозяйством.

В 1922 году Фишер опубликовал статью под названием *Отношение доминирования*. Среди ряда других важных новых идей в этой статье была рассмотрена математическая модель, сочетающая законы Менделя и идею естественного отбора, подчеркнутую Дарвином для теории эволюции. Фишер рассматривал ту же ситуацию,

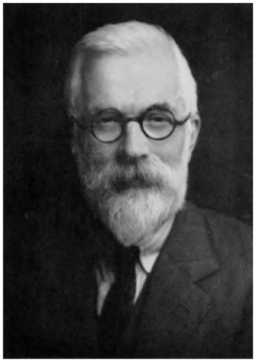


Рис. 14.1:
Фишер (1890–1962)

что и Харди с двумя аллелями A и a и со случайной гипотезой спаривания. Но он предполагал, что индивидуумы с генотипами AA, Aa и aa имеют различную смертность до достижения совершеннолетия, имитируя тем самым естественный отбор. Обозначая p_n, $2q_n$ и r_n частоты трех генотипов среди взрослых индивидуумов в поколении n, получим $(p_n + q_n)^2$, $2(p_n + q_n)(q_n + r_n)$ и $(q_n + r_n)^2$ новорожденных в поколении $n + 1$, имеющих эти генотипы, соответственно. Пусть u, v и w будут соответствующими вероятностями выживания от рождения до взрослого возраста. Тогда частоты генотипов среди взрослых индивидуумов в поколении $n + 1$, p_{n+1}, $2q_{n+1}$ и r_{n+1}, могут быть найдены по формулам

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

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

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

где мы для удобства полагаем

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

Принимая во внимание, что $p_n + 2q_n + r_n = 1$, мы видим, что когда $u = v = w$ (т.е. когда нет естественного отбора), система (14.1)-(14.3) сводится к системе (11.1)-(11.3), рассматриваемой Харди.

Пусть $x_n = p_n + q_n$ будет частотой аллели A среди взрослых людей в поколении n. Тогда $q_n + r_n = 1 - x_n$ - это частота аллели *a*. Складывая (14.1) и (14.2), мы получим

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

Это уравнение может быть переписано в виде

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

Отсюда можно заключить, что всегда имеется как минимум два устойчивых состояния, в которых частота x_n остается постоянной: $x = 0$ (население полностью состоит из гомозиготных aa) и $x = 1$ (население полностью состоит из гомозиготных AA).

Используя уравнение (14.4), можно показать, что если у гомозиготных AA больше шансов на выживание, чем у двух других генотипов ($u > v$ и $u > w$), то аллель a постепенно исчезнет из популяции. Этот случай не должен быть очень распространен в природе, если мы знаем, что обе аллели сосуществуют. Если же гетерозиготный Aa имеет эволюционное преимущество перед гомозиготными AA и aa ($v > u$ и $v > w$), то в популяции могут сосуществовать три генотипа. Это наиболее распространенный случай, который может объяснить выживаемость гибридов, наблюдаемую в природе.

Действительно, устойчивое состояние $x = 1$ стабильно, когда $u > v$, потому что $x_{n+1} - x_n \approx (1-x_n)\,(u-v)/u$, когда x_n близко к 1. Популяция сходится к этому устойчивому состоянию. Состояние $x = 1$ нестабильно, когда $u < v$, в этом случае есть третье устойчивое состояние $x^* = (v-w)/(2\,v-u-w)$ с $0 < x^* < 1$. Более того, мы можем проверить, что оно устойчиво. Устойчивое состояние x^* соответствует смеси между тремя генотипами.

Таким образом, объединяя законы Менделя и гипотезу естественного отбора (здесь - различные вероятности выживания для трех генотипов), мы можем объяснить две ситуации сосуществования или исчезновения генотипов. После Фишера эта модель была также изучена Халданом (J.B.S. Haldane, см. главу 17) и Райтом (Sewall Wright, см. главу 19).

В преддверии главы 20, отметим, что если A полностью доминирует и гомозиготный генотип aa находится в невыгодном положении по сравнению с двумя другими генотипами, то числа $u : v : w$

находятся в соотношении $1 : 1 : 1 - \varepsilon$, и уравнение (14.4) записывается в виде

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

при $\varepsilon \ll 1$. Если выживание гетерозиготных Aa равно полусумме выживаемости гомозиготных генотипов , то числа $u : v : w$ находятся в соотношении $1 : 1 - \varepsilon/2 : 1 - \varepsilon$ и

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

при $\varepsilon \ll 1$.

В компании Rothamsted Фишером были проанализированы долгосрочные данные по урожайности и метеорологии. Но он также внес большой вклад в методологию статистики. В 1925 году он опубликовал книгу *Статистические методы для научных работников*, которая пользовалась большим успехом и была многократно переиздана. В 1929 году он стал членом Королевского общества. В 1930 году Фишер опубликовал книгу *Генетическая теория естественного отбора*, которая стала важной вехой в истории генетики популяции. Он стал профессором евгеники в университетском колледже в Лондоне в 1933 году, преемником Карла Пирсона в лаборатории Галтона. В 1943 году он перешел на кафедру генетики в Кембриджском университете, на этот раз сменив на этом посту Р.К. Паннета (см. главу 11). Он также опубликовал несколько книг: *Конструирование экспериментов* (1935 г.), *Теория инбридинга* (1949 г.) и *Статистические методы и научные выводы* (1956 г.). Ставши Рыцарем в 1952 году, он поселился в Австралии после выхода на пенсию в 1959 году и умер в Аделаиде в 1962 году. К другой части его работы мы вернемся в главе 20.

Дополнительное чтение

1. Fisher Box, J.: R.A. Fisher, *The Life of a Scientist.* Wiley (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)

Глава 15

Юл и эволюция (1924)

В 1924 году британский статистик Юл изучал модель эволюции, в которой виды могут производить новые виды путем небольших мутаций, а рода могут производить новые рода путем больших мутаций. Его целью было объяснить распределение количества видов внутри родов: большинство родов содержит только один вид и несколько родов содержит большое количество видов. Стохастический «процесс рождения», который Юл ввел в свою модель, до сих пор является основным инструментом в изучении филогенетических деревьев и многих других областей.

Джордж Удни Юл (Yule) родился в Шотландии в 1871 году, его отец занимал высокий пост в британской администрации в Индии. В возрасте 16 лет Юл начал учиться в университетском колледже в Лондоне, чтобы стать инженером. В 1892 году он изменил свою специализацию и провел один год, занимаясь исследованиями в Бонне под руководством физика Генриха Герца, который несколько лет назад продемонстрировал существование электромагнитных волн. Когда Юл вернулся в Англию, Карл Пирсон предложил ему должность доцента по прикладной математике в университетском колледже. Юл, вслед за Пирсоном, начал фокусироваться на статистике. В 1911 он опубликовал *Введение в теорию статистики*, которое было перепечатано 14 раз. На следующий год он переехал в Кембриджский университет. Его исследовательская работа была посвящена теоретическим аспектам статистики, а также прикладным аспектам сельского хозяйства и эпидемиологии. В 1922 году он стал членом Королевского общества.

В 1924 году Юлом была опубликована статья *Математическая теория эволюции, основанная на выводах доктора Дж.К. Уиллиса.* Уиллис был его коллегой по Королевскому обществу, который опубликовал в 1922 году книгу под названием *Возраст и ареал, исследование географического распределения и происхождения видов.* Он изучал распределение видов между различными родами при классификации растений и животных. Собранные им данные сви-

Рис. 15.1:
Джордж Юл (1871-1951)

детельствуют о том, что большинство родов содержат только один вид, что все меньше и меньше родов содержат большее число видов и что все еще существует несколько родов, содержащих большое число видов. В таблице 15.1 приведены данные по змеям, ящерицам и двум семействам жуков (Chrysomelidae и Cerambycinae). Известные в то время 1580 видов ящериц были отнесены к 259 родам, 105 родов содержали только один вид, 44 - только два вида, 23 - только три вида и т.д., а два рода - более ста видов. Для других семейств животных и растений распределение родов по количеству содержащихся в них видов имело очень схожую форму.

Юл предложил Виллису попытаться построить график с логарифмическими шкалами. Это дало поразительный результат (рисунок 15.2): логарифм числа Q_n родов, содержащих n виды, уменьшается более или менее линейно с $\log(n)$. Другими словами, существуют константы $\alpha > 0$ и $\beta > 0$, такие, что $Q_n \approx \alpha\, n^{-\beta}$: распределение следует «степенному закону». В своей статье 1924 года Юл искал математическую модель эволюции, которая могла бы объяснить такое статистическое распределение.

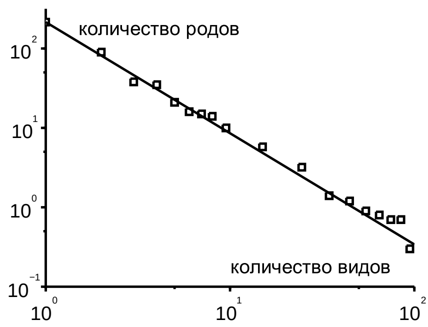
Для этого он сначала представил себе континуальную стохастическую модель [1] роста числа видов в пределах одного рода (рис. 15.3). Начиная только с одного вида в момент времени $t = 0$, он предполагал, что вероятность рождения вида путем мутации к новому виду того же рода в течение небольшого временного интервала dt (на шкале времени эволюции) равна $r\,dt$ с $r > 0$.

[1]МакКендрик (см. главу 16) уже начал изучать такие модели в динамике населения в статье, опубликованной в 1914 году.

Таблица 15.1: Данные собраны Уиллисом. (1) Хрисомелиды, (2) Керамбицинае, (3) Змеи, (4) Ящерицы.

Количество видов	Количество родов (1)	(2)	(3)	(4)
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
всего	627	1024	293	259

Рис. 15.2: Количество родов в зависимости от количества содержащихся в них видов, с десятичной логарифмической шкалой. Данные для Хрисомелид. Для сглаживания колебаний, когда n (количество видов) велико, число родов подсчитывали для диапазонов n-значений, как показано в таблице 15.1. Таким образом, среднее число родов для одного значения n может быть меньше 1.



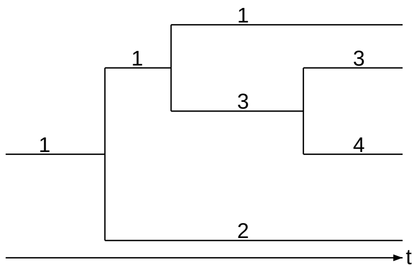


Рис. 15.3: Моделирование эволюции числа видов в рамках одного рода. Вид 1 генерирует виды 2 и 3. Вид 3 генерирует вид 4.

Пусть $p_n(t)$ - это вероятность того, что в момент времени t существует вид n (n - целое число, но t - вещественное число). Для вычисления $p_n(t + dt)$ Юл рассмотрел несколько случаев:

- если в момент времени t существует $n-1$ вид, то каждый вид имеет вероятность $r\,dt$ генерации одного нового вида между t и $t + dt$; в пределе $dt \to 0$, в момент времени $t + dt$ будет иметься n вид с вероятностью $(n-1)\,r\,dt$;
- если в момент времени n видов будет t, то в момент времени $t + dt$ видов будет $n + 1$ с вероятностью $n\,r\,dt$.

Таким образом, $p_n(t)$ дается следующей системой дифференциальных уравнений

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

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

при всех $n \geqslant 2$. Из первого уравнения мы получаем $p_1(t) = e^{-r\,t}$, потому что $p_1(0) = 1$. Можно показать, что решение второго уравнения, удовлетворяющего начальному условию $p_n(0) = 0$, равно

$$p_n(t) = e^{-r\,t}\,(1 - e^{-r\,t})^{n-1} \tag{15.3}$$

при всех $n \geqslant 2$ (рисунок 15.4). Таким образом, в некоторое фиксированное время t распределение вероятностей $(p_n(t))_{n\geqslant 1}$ является геометрическим с соотношением между двумя последовательными членами, равным $1 - e^{-r\,t}$.

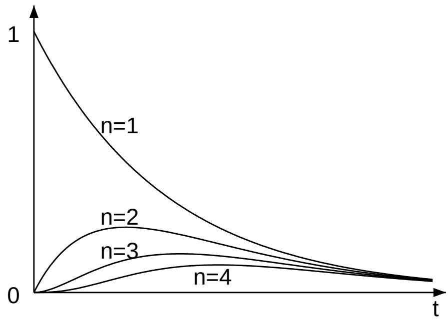


Рис. 15.4: Вероятность $p_n(t)$ того, что существуют виды одного и того же рода для $1 \leqslant n \leqslant 4$.

Действительно, сначала заметим, что уравнение (15.2) эквивалентно уравнениям

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

из которых можно последовательно вычислить $p_2(t)$, $p_3(t)$..... Мы получаем $p_2(t) = e^{-r\,t}\,(1-e^{-r\,t})$, затем $p_3(t) = e^{-r\,t}\,(1-e^{-r\,t})^2$, что предполагает формулу (15.3) для общего решения. Наконец, можно проверить, что эта формула является решением уравнения (15.4).

Из формулы (15.3) Юль также вывел, что ожидаемое количество видов растет экспоненциально с течением времени:

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

Действительно, сначала мы замечаем, что при $|x| < 1$,

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

Тогда

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

В частности, если T - это время удвоения, определяемое соотношением $e^{rT} = 2$, то распределение вероятности $(p_n(t))_{n \geqslant 1}$ числа видов в момент времени $t = T$ является геометрическим с коэффициентом $1/2$:

$$\frac{1}{2}\,, \quad \frac{1}{4}\,, \quad \frac{1}{8}\,, \quad \frac{1}{16} \quad \ldots$$

В момент времени $t = kT$, он геометрический с коэффициентом $1 - 1/2^k$ и $p_1(kT) = 1/2^k$.

Далее Юл рассматривал параллельно с ростом числа видов, принадлежащих к одному и тому же роду, аналогичный процесс, обусловленный более крупными мутациями, приводящими к созданию новых родов. Пусть $s\,dt$ будет вероятность того, что существующий род сгенерирует новый род за небольшой промежуток времени dt. Как и прежде, если предположить, что в момент времени $t = 0$ существует только один род, то ожидаемое количество родов в момент времени t составит $e^{s\,t}$. Среднее число родов, созданных за единицу времени в момент времени t, является производной $s\,e^{s\,t}$. В пределе[2], где $t \to +\infty$, среднее число родов, которые в момент времени t существовали между x и $x + dx$, равно $s\,e^{s\,(t-x)}\,dx$. Вероятность того, что случайно выбранный род в моменте времени t имеет продолжительность существования в промежутке между x и $x + dx$ составляет $s\,e^{-s\,x}\,dx$.

Если род, выбранный случайным образом в момент времени t, имеел продолжительность существованиямежду x и $x + dx$, то вероятность того, что этот род содержит n видов, по формуле (15.3), равна $e^{-r\,x}\,(1 - e^{-r\,x})^{n-1}$ для всех $n \geqslant 1$. Таким образом, вероятность q_n что случайно выбранный род в момент времени t имеет n видов, равна

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

[2] Юл также рассматривал случай, когда t нельзя считать очень большим по сравнению со временем удвоения $e^{s\,t}$. Вычисления немного сложнее, но конечные результаты не сильно отличаются.

Положим $u = r/s$. Легко вычислить, что $q_1 = 1/(1+u)$ и что

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

при всех $n \geqslant 2$.

Действительно, имеем $(1-e^{-r\,x})^{n-1} = (1-e^{-r\,x})^{n-2}\,(1-e^{-r\,x})$. Следовательно,

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

Интегрируя по частям, получаем

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

Формула (15.5) показывает, что последовательность вероятностей $(q_n)_{n\geqslant 1}$ уменьшается. Таким образом, максимум достигается при $n = 1$: большинство родов содержат только один вид. Именно это показали данные. Более того, уменьшение q_n до 0, когда n стремится к бесконечности, происходит относительно медленно, поскольку $q_n/q_{n-1} \to 1$. Это может объяснить, почему некоторые роды содержат большое количество видов. Точнее, Юл показал, что $\log q_n$ уменьшается линейно с $\log(n)$.

Введем гамма-функцию Эйлера $\Gamma(z) = \int_0^{+\infty} t^{z-1}\,e^{-t}\,dt$. Тогда $\Gamma(n+1) = n! = n\times(n-1)\times\cdots\times 2\times 1$, если n - целое число, и $\Gamma(z+1) = z\,\Gamma(z)$. Следовательно, (15.5) принимает форму

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

Но приближение Стерлинга дает $\log\Gamma(n) \approx n\log n - n - \frac{1}{2}\log n +$ c. Точно так же, $\log\Gamma(n+1+1/u) \approx n\log n - n + \left(\frac{1}{u}+\frac{1}{2}\right)\log n +$c. Таким образом, $\log q_n \approx -(1+\frac{1}{u})\log n +$ c.

Рассмотрим, например, случай ящериц. Параметр u можно оценить из пропорции $q_1 = 1/(1+u)$ родов, которые содержат только один вид. Согласно таблице 15.1, мы имеем $q_1 = 105/259$, поэтому $u \approx 1{,}467$. Затем мы можем вычислить теоретическую вероятность

q_n и ожидаемое число Q_n родов, содержащих n видов, умножив q_n на общее число видов, равное 259 (таблица 15.2). Юл заметил, что согласие между наблюдениями и расчетами относительно хорошее, учитывая простоту модели[3], которая не учитывает, например, катаклизмы, просхидившие с видами за миллионы лет эволюции.

Таблица 15.2: Сравнение данных и теории в случае ящериц (1 580 видов, классифицированных в 259 родов).

Количество видов на род	Наблюдаемое количество родов	Расчетное число родов
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
всего	259	259

После 1931 года Юл постепенно ушел на пенсию из Кембриджского университета. Он заинтересовался статистическим распределением длины предложений для выявления авторов книг. Он применил это, в частности, к книге, опубликованной Джоном Граунтом (см. главу 2), но, возможно, написанной Уильямом Петти. В 1944

[3]Для числа родов, содержащих более 100 видов, результаты получаются лучше, чем в таблице 15.2, учитывая, что t не велико по сравнению с удвоением времени удвоения e^{st}.

году он опубликовал книгу *Статистическое исследование литературной лексики.* Он умер в 1951 году.

В настоящее время модель Юля до сих пор используется для анализа филогенетических деревьев (генеалогических деревьев видов). Эти деревья, похожие на те, что показаны на рисунке 15.3, более известны благодаря новым данным, полученным из молекулярной биологии. Но применение стохастического процесса, определяемого уравнениями (15.1)-(15.2), не ограничивается теорией эволюции. Этот процесс является составным блоком многих моделей в динамике популяций, начиная с микроскопического уровня (например, для моделирования колоний бактерий) и заканчивая макроскопическим уровнем (для моделирования начала эпидемии). Он называется процессом чистого рождения или процессом Юля. Простой вариант этого процесса включает вероятность $m\,dt$ смерти в течение любого небольшого промежутка времени dt: ожидаемый размер популяции в момент времени t для этого процесса рождения и смерти составляет тогда $e^{(r-m)t}$. Что касается распределения вероятности (15.5), то его иногда называют распределением Юля. Распределения с хвостами, удовлетворяющими степенному закону, привлекли большое внимание в различных областях науки. Изучение эпидемий в случайных сетях с распределением по степенному закону - лишь один из примеров.

Дополнительное чтение

1. Aldous, D.J.: Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today. *Stat. Sci.* 16, 23–34 (2001) project-euclid.org
2. Edwards, A.W.F.: George Udny Yule. In: Heyde, C.C., Seneta, E. (eds.) *Statisticians of the Centuries*, 292–294. Springer, New York (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

Глава 16

МакКендрик и Кермак о моделировании эпидемии (1926–1927)

> В 1926 году МакКендрик изучил стохастическую модель эпидемии и нашел метод расчета вероятности того, что эпидемия достигнет определенного конечного размера. Он также вывел дифференциальное уравнение в частных производных, определяющее возраст-структурированные популяции в непрерывных временных рамках. В 1927 г. Кермак и МакКендрик исследовали детерминистскую модель эпидемии и получили уравнение для конечного размера эпидемии, которое подчеркивает определенный порог для плотности населения. Крупные эпидемии могут происходить выше, но не ниже этого порога. Эти работы до сих пор широко используются в современной эпидемиологии.

Андерсон Грэй МакКендрик (McKendrick) родился в 1876 году в Эдинбурге последним из пяти детей. Он изучал медицину в Университете Глазго, где его отец был профессором физиологии. В 1900 году он поступил на работу в индийскую медицинскую службу. Перед тем, как отправиться в Индию, он сопровождал Рональда Росса в миссии по борьбе с малярией в Сьерра-Леоне. Затем он служил в армии 18 месяцев в Судане. По прибытии в Индию он был назначен врачом в тюрьме в Бенгалии, где он пытался контролировать дизентерию. В 1905 году он поступил на работу в новый Центральный институт медицинских исследований в Касаули (на севере Индии). Он работал над исследованием бешенства, но также изучал математику. В 1920, заразившись тропической болезнью, он вернулся в Эдинбург и стал суперинтендантом лаборатории Королевского колледжа врачей.

В 1926 г. Маккендрик опубликовал статью «Применение математики к медицинским проблемам», в которой содержалось несколько новых идей. Он представил, в частности, математическую модель для эпидемий с учетом стохастического аспекта инфекции и выздоровления.

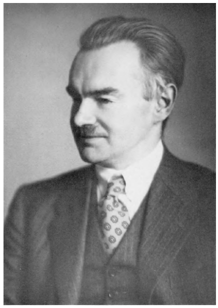


Рис. 16.1: МакКендрик (1876-1943) и Кермак (1898-1970).

Рассмотрим популяцию размером в N с изначально только одним зараженным человеком. Люди могут последовательно проходить через три состояния: восприимчивое S, инфицированное I и восстановленное R (рисунок 16.2)[1].

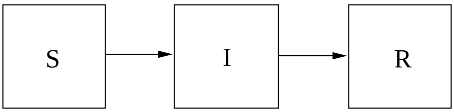


Рис. 16.2: Возможные состояния: восприимчивость (S), инфицированность (I), восстановление (R).

Пусть $p_{i,r}(t)$ - это вероятность того, что население содержит в момент времени t ровно i людей в состоянии I и r людей в состоянии R, где i и r являются целыми числами, такими, что $1 \leqslant i+r \leqslant N$. В этом случае будем говорить, что население находится в состоянии (i, r). Количество восприимчивых людей равно $s = N - i - r$. После работы Росса по малярии (см. главу 12) МакКендрик предположил, что в течение небольшого промежутка времени dt вероятность возникновения нового случая заражения равна $a\,s\,i\;dt$ (т.е. пропорциональна как количеству восприимчивых людей, так и количеству инфицированных). Вероятность одного нового выздоровления равна $b\,i\;dt$ (т.е. пропорциональна как количеству восприимчивых людей, так и числу инфицированных). И a, и b являются положительны-

[1]Модель Даниэля Бернулли (см. главу 4) включала состояния S и R, но не I, при этом длительность заражения была значительно меньше средней продолжительности жизни.

ми параметрами. Для вычисления $p_{i,r}(t+dt)$ следует различать несколько случаев:

- население находится в состоянии $(i-1,r)$ в момент времени t и одна новая инфекция переводит население в состояние (i,r) между t и $t+dt$; вероятность этого события составляет $a\,s\,(i-1)\,dt$ с $s=N(i-1)-r$;
- население находится в состоянии (i,r) в момент времени t и одна новая инфекция переводит население в состояние $(i+1,r)$ между t и $t+dt$; вероятность этого события составляет $a\,s\,i\,dt$ с $s=N-i-r$;
- Население находится в состоянии $(i+1,r-1)$ в момент времени t и одно новое восстановление перемещает население в состояние (i,r) между t и $t+dt$; вероятность этого события составляет $b\,(i+1)\,dt$;
- Население находится в состоянии (i,r) в момент времени t и одно новое восстановление перемещает население в состояние $(i-1,r+1)$ между t и $t+dt$; вероятность этого события составляет $b\,i\;dt$.

Следовательно, МакКендрик получил уравнения

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

за $1 \leqslant i+r \leqslant N$. Первое условие справа отсутствует, если $i=0$, а третье условие отсутствует, если $r=0$. Начальные условия - $p_{i,r}(0)=0$ для всех (i,r), кроме $p_{1,0}(0)=1$.

С помощью этой модели МакКендрик сумел вычислить вероятность того, что эпидемия закончится с заражением n, что является пределом $p_{0,n}(t)$ при $t\to+\infty$. Действительно, в решении системы нет необходимости (16.1). Достаточно заметить, что пока есть зараженные i и восстановленные r, вероятность нового заражения в течение небольшого промежутка времени dt составляет $a\,(N-i-r)\,i\,dt$, а вероятность нового восстановления $b\,i\;dt$. Таким образом, вероятности перехода (как их обычно называют в теории цепей Маркова) из состояния (i,r) в состояние $(i+1,r)$ или состояние $(i-1,r+1)$ соответственно равны

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

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

для всех $i \geqslant 1$ (рисунок 16.3).

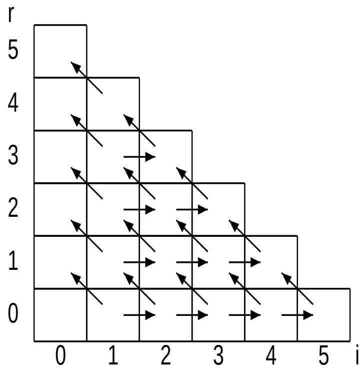


Рис. 16.3: Диаграмма, показывающая возможные состояния популяции с $N = 5$ (i по горизонтальной оси, r по вертикальной оси) и возможные переходы из-за инфекции (горизонтальные стрелки) или к выздоровлению (другие стрелки).

Пусть $q_{i,r}$ - это вероятность того, что во время эпидемии население пройдет через состояние (i,r). Поскольку $i = 1$ и $r = 0$, когда $t = 0$, у нас есть $q_{1,0} = 1$. Остальные состояния достигаются либо после заражения, либо после выздоровления:

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

Первый член выражения с правой стороны пропущен, если $i = 0$ или $i = 1$. Второе условие отсутствует, если $r = 0$. Из этой формулы можно вычислить сначала $(q_{i,0})_{2\leqslant i\leqslant N}$, затем $(q_{i,1})_{0\leqslant i\leqslant N-1}$, затем $(q_{i,2})_{0\leqslant i\leqslant N-2}$ и др. Вероятность того, что эпидемия наконец охватит n людей, составляет $q_{0,n}$. В 1926 году такие вычисления были довольно трудоемкими. Поэтому МакКендрик ограничился примерами, касающимися очень маленьких популяций, например, семьи. При $N = 5$ человек и $b/a = 2$ он получил таблицу 16.1. Наибольшие вероятности соответствуют случаю, когда заражен только один человек в семье, и случаю, когда заражена вся семья.

В той же статье 1926 года содержится и новая формулировка демографических проблем, когда время рассматривается как непрерывная переменная. Для dx бесконечно малой, пусть $P(x,t)\,dx$ -

Таблица 16.1: Вероятность эпидемии в семье из пяти человек заразить n, если $b/a = 2$.

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

это население с возрастом от x до $x + dx$ в момент времени t. Пусть $m(x)$ будет смертность в возрасте x. Затем $P(x+h, t+h) \approx P(x,t) - m(x)\, P(x,t)\, h$ для h бесконечно мал. Вводим частные производные функции $P(x,t)$:

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

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

Используя это

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

МакКендрик получил следующее дифференциальное уравнение:

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

Такое уравнение естественным образом появляется в демографических задачах, структурированных непрерывной переменной, такой как возраст в демографии (см. главу 25) или время, прошедшее с момента заражения в эпидемиологии.

В 1921 году Уильям Огилви Кермак (Kermack) был назначен ответственным за химическую секцию лаборатории Королевского колледжа врачей в Эдинбурге. Кермак родился в 1898 году в небольшом городке в Шотландии. Он учился в Абердинском университете и начал заниматься исследованиями в области органической химии в промышленной лаборатории в Оксфорде. Несмотря на то, что после взрыва в своей эдинбургской лаборатории в 1924 году он полностью ослеп, он продолжил свою химическую работу с помощью коллег и студентов. Кроме того, Кермак начал сотрудничать с МакКендриком в области математического моделирования эпидемий. Начиная с 1927 г., они вместе опубликовали серию

«Вклад в математическую теорию эпидемий», в которой изучали детерминированные модели эпидемий. Пусть N будет размер популяции с N достаточно большой. Предположим, как в статье 1926 года, что люди могут быть либо восприимчивыми, либо инфицированными, либо выздоровевшими. Если болезнь смертельна, то третьим состоянием на самом деле является смерть. Пусть $S(t)$, $I(t)$ и $R(t)$ будут числом людей в каждом из трех состояний. Модель является (в упрощенном виде) системой трех дифференциальных уравнений:

$$\frac{dS}{dt} = -a\,S\,I, \tag{16.2}$$

$$\frac{dI}{dt} = a\,S\,I - b\,I, \tag{16.3}$$

$$\frac{dR}{dt} = b\,I. \tag{16.4}$$

Таким образом, количество новых случаев инфицирования за единицу времени, как и в стохастической модели 1926 года, пропорционально как количеству восприимчивых людей, так и количеству инфицированных. В начале эпидемии, в момент времени $t = 0$, заражено определенное количество людей: $S(0) = N - I_0$, $I(0) = I_0$ и $R(0) = 0$, предполагая $0 < I_0 < N$.

Хотя система (16.2)-(16.4) не имеет явного решения, некоторые ее свойства могут быть доказаны:

- общая численность населения $S(t) + I(t) + R(t)$ остается постоянной и равна N;
- $S(t)$, $I(t)$ и $R(t)$ остаются неотрицательными (как и должно быть, поскольку речь идет о популяциях);
- когда $t \to +\infty$, $S(t)$ уменьшается до предела $S_\infty > 0$, $I(t)$ стремится к 0, а $R(t)$ увеличивается до предела $R_\infty < N$;
- более того, формула

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

косвенно дает S_∞ и, следовательно, также окончательный размер эпидемии $R_\infty = N - S_\infty$.

Действительно, сначала мы видим, что $\frac{d}{dt}(S+I+R) = 0$. Таким образом, $S(t)+I(t)+R(t) = S(0)+I(0)+R(0) = N$. Уравнения (16.2) и (16.3) могут быть переписаны как

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

С одной стороны следует, что

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

а с другой стороны, что

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

Уравнения (16.2) и (16.4) показывают, что функция $S(t)$ уменьшается, а функция $R(t)$ увеличивается (в частности, $R(t) \geqslant 0$). Так как $S(t) \geqslant 0$ и $R(t) \leqslant N$, функции $S(t)$ и $R(t)$ имеют ограничения при $t \to +\infty$. Начиная с $I(t) = N - S(t) - R(t)$, $I(t)$ также имеет лимит, когда $t \to +\infty$, который может быть только нулем, как видно из интеграции (16.4). Уравнение (16.2) также показывает, что

$$-\frac{d}{dt}[\log S] = a\,I.$$

Интегрируя от $t = 0$ до $t = +\infty$, мы находим :

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

Уравнение (16.3) можно переписать как

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

Интегрируя от $t = 0$ до $t = +\infty$, мы получаем

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

Комбинируя два результата, мы получаем формулу (16.5), которая показывает, что $S_\infty > 0$.

Когда начальное количество инфицированных I_0 невелико по сравнению с численностью населения N, что часто бывает в начале эпидемии в городе, формулу (16.5) можно переписать, используя $S_\infty = N - R_\infty$ в качестве

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

где по определению

$$\mathscr{R}_0 = \frac{a\,N}{b}\,.$$

Уравнение (16.6) имеет положительное решение, только если $\mathscr{R}_0 > 1$. Таким образом, Кермак и МакКендрик приходят к следующему выводу: эпидемия разовьется и заразит значительную часть популяции только в том случае, если $\mathscr{R}_0 > 1$. Существует порог плотности населения $N^* = b/a$, ниже которого эпидемия не может произойти.

Когда размер популяции N чуть выше этого порога ($N = N^* + \varepsilon$), происходит эпидемия малой амплитуды. Из (16.6) следует, что $R_\infty \approx 2\,\varepsilon$. Так что $S_\infty \approx N^* - \varepsilon$.

Действительно, используя приближение $-\log(1-x) \approx x + \frac{x^2}{2}$, уравнение (16.6) становится

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

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

Как и в малярийной модели Росса (глава 12), условие $\mathscr{R}_0 > 1$ имеет простую интерпретацию. Поскольку $a\,N$ - это количество людей, которое один зараженный человек заражает за единицу времени в начале эпидемии, а $1/b$ - это средний инфекционный период, то $\mathscr{R}_0 = a\,N/b$ - это среднее количество вторичных случаев заболевания, обусловленное одним зараженным человеком в начале эпидемии.

Для смертельных болезней $R(t)$ - это совокупное число смертей с начала эпидемии, а dR/dt - число смертей за единицу времени.

Кермак и МакКендрик заметили, что график функции dR/dt в их математической модели действительно имеет ту форму колокольчика, которую можно ожидать от кривой эпидемии (рисунок 16.4).

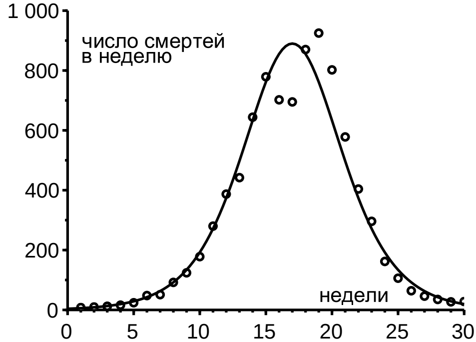


Рис. 16.4: Кривая dR/dt в зависимости от времени и данных о количестве смертей в неделю во время эпидемии чумы в Бомбее в 1905-1906 гг.

Для получения dR/dt они разделили (16.2) на (16.4), чтобы получить $dS/dR = -a\,S/b$. Так

$$S(t) = S(0)\,\exp(-a\,R(t)/b).$$

Заменив это на уравнение (16.4) и используя $S(t)+I(t)+R(t) = N$, они получили уравнение.

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

которое до сих пор не может быть решено однозначно. Тем не менее, если $\frac{a}{b}\,R(t)$ останется небольшим на протяжении всей эпидемии, то аппроксимация $\exp(-u) \approx 1 - u + u^2/2$ даст

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

Это так называемое уравнение Риккати с двумя постоянными решениями, одно положительное R_+ и одно отрицательное R_-, заданное корнями полинома второго порядка в R справа (16.8). Пусть $\tilde{R}(t)$ будет точным решением (16.8), а $Q(t) = \tilde{R}(t) -$

R_+. Затем $Q(t)$ удовлетворяет дифференциальному уравнению Бернулли, аналогичному уравнению Даниэля Бернулли и Ферхюльста (см. (4.5) и (6.1))). Таким образом, можно напрямую адаптировать формулу (6.2), чтобы получить $Q(t)$. Не сложное, но длинное вычисление показывает, что dQ/dt имеет форму

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

где α, β и γ являются константами, которые сложным образом зависят от параметров модели. Поскольку $dR/dt \approx d\tilde{R}/dt = dQ/dt$, Кермак и МакКендрик могли выбрать (α, β, γ) в соответствии со своими данными. Конечно, современные компьютеры и программы могут легко решить численное дифференциальное уравнение (16.7), не проходя через эти аппроксимации.

Таким образом, кривая для dR/dt хорошо подогнана под данные о количестве смертей в неделю во время эпидемии plague в Бомбее в период с декабря 1905 года по июль 1906 года (рис. 16.4).

Кермак и МакКендрик также рассмотрели более общую модель, где инфекционность $a(x)$ зависит от времени x с момента заражения, и где скорость восстановления $b(x)$ также зависит от x. Уравнение, дающее окончательную величину эпидемии (когда начальное число инфицированных случаев небольшое), все еще (16.6), но с

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

Параметр $\mathscr{R}_0$ имеет ту же интерпретацию, что и в предыдущем случае: это среднее число вторичных случаев, вызванных одним инфицированным в начале эпидемии. Обратите внимание на сходство (16.9) и формулы Лотки (10.2) для $\mathscr{R}_0$ в демографии: возраст заменяется временем с момента инфицирования, выживаемость - вероятностью $e^{-\int_0^x b(y)\,dy}$ быть все еще инфицированным, фертильность - степенью контакта $N\,a(x)$.

Кермак и МакКендрик разработали несколько других математических моделей эпидемий в течение 1930-х годов. Они до сих пор являются строительными блоками для большинства более сложных моделей, используемых в настоящее время в эпидемиологии. Параметр $\mathscr{R}_0$ по-прежнему играет центральную роль в анализе модели.

МакКендрик ушел в отставку в 1941 году и умер в 1943 году. Между 1930 и 1933 Кермак был соавтором нескольких статей по математической физике с Уильямом МакКреа и Эдмундом Уиттакером, оба с математического факультета Эдинбургского университета. В 1930-х и 1940-х годах команда химиков Кермака пыталась синтезировать новые молекулы с противомалярийной активностью, но с ограниченным успехом. В 1938 году Кермак вместе с Филиппом Эгглтоном выступил соавтором популярной книги по элементарной биохимии, *Вещи, из которых мы сделаны.* В 1944 году он был избран членом Королевского общества, а в 1949 году занял кафедру биохимии в Абердинском университете. Позднее он работал деканом факультета естественных наук. В 1968 году он ушел на пенсию и умер в 1970 году.

Дополнительное чтение

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

Глава 17

Халдейн и мутации (1927)

В другом разделе своей статьи 1922 года Фишер рассмотрел проблему мутирующего гена, который может быть передан случайному числу потомства с заданным распределением вероятностей. Формально проблема была такой же, как и проблема исчезновения семейств, но в генетическом контексте. Фишер показал, что если бы распределение вероятности было распределением Пуассона и если бы у мутирующего гена не было селективного преимущества, то мутирующий ген мог бы очень медленно исчезнуть из популяции. В 1927 году британский биолог Халдейн подтолкнул изучение этой модели и показал, что вероятность сохранения свойства гена мутировать в два раза выше его селективного преимущества. Он также дал более строгий подход к проблеме исчезновения.

Джон Бурдон Сандерсон Халдейн (Haldane) родился в 1892 году в Оксфорде, где его отец был профессором физиологии в университете. Халдейн учился в Итонском колледже и после 1911 года в Новом колледже Оксфордского университета. Сосредоточившись на математике на первом курсе, позже он обратился к гуманитарным наукам. Его обучение было прервано Первой мировой войной, во время которой он служил во Франции и Ираке. Будучи раненым, он был отправлен в качестве военного инструктора в Индию. В 1915 году он опубликовал первую статью о генетических экспериментах на мышах, которые он начал до войны. В 1919 г. он стал стипендиатом Нового колледжа, преподавая физиологию и изучая дыхание, как и его отец. В 1923 году он присоединился к лаборатории биохимии Ф. Г. Хопкинса[1] в Кембриджском университете, где он сосредоточился на кинетике ферментов. Он также опубликовал научно-фантастический роман *Дедал или наука и будущее* (1923 г.) и эссе под названием *Каллиник, защита химической войны* (1925 г.). В период с 1924 г. по 1934 г. он написал серию из десяти статей

[1] Фредерик Гоуланд Хопкинс, получивший Нобелевскую премию по физиологии и медицине в 1929 году за работу над витаминами.

под названием *Математическая теория естественного и искусственного отбора*.

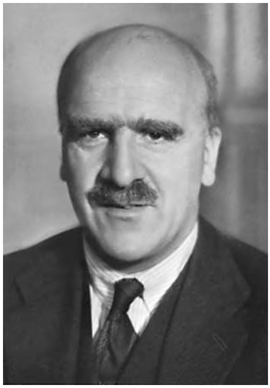


Рис. 17.1:
Халдейн (1892–1964)

В пятой статье этой серии, опубликованной в 1927 году, Халдейн пересмотрел другую генетическую модель, изученную Фишером в 1922 году, модель, которая учитывает мутации. Фишер изучил вероятность того, что мутирующий ген вторгнется в популяцию или исчезнет. Эта проблема формально такая же, как и у Бьенэме, Галтона и Ватсона, касающаяся исчезновения семейств. Но Фишер не делал никаких ссылок на эти работы, хотя, возможно, он читал статью Галтона и Ватсона, воспроизведенную в приложении к книге Галтона 1889 года *Естественное наследие*. Как и в главе 9, обозначим p_k вероятность передачи гена потомству k первого поколения ($k \geqslant 0$). Фишер как и Ватсон рассматривал генерирующую функцию

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

за исключением того, что он не зафиксировал верхнюю границу для k: сумма может включать в себя бесконечное количество элементов. Он понял, что начиная с одной особи с мутирующим геном в поколении 0, вероятность нахождения этого гена в k индивидуумах составляет коэффициент x^k в $f_1(x) = f(x)$ для поколения 1, в $f_2(x) = f(f(x))$ для поколения 2, в $f_3(x) = f(f(f(x))$ для поколения 3 и др. Таким образом, становится понятно, что следующее уравнение выполняется.

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

Это уравнение намного более практично, чем уравнение $f_n(x) = f_{n-1}(f(x))$, полученное Ватсоном. В частности, из (17.1) следует, что вероятность исчезновения в пределах n поколений $x_n = f_n(0)$ удовлетворяет итерационной формуле $x_n = f(x_{n-1})$, как уже заметил Бьенэме.

В качестве примера, Фишер рассмотрел случай растения с мутирующим геном, который может производить N семян, каждое семя имеет вероятность q выжить, чтобы произвести новое растение. Вероятность p_k получения k потомства с мутирующим геном является биномиальной:

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

для всех $0 \leqslant k \leqslant N$ и $p_k = 0$ для $k > N$. Таким образом, генерирующая функция записывается следующим образом: $f(x) = (1 - q + q\,x)^N$. Пусть $\mathscr{R}_0 = Nq$ будет средним количеством семян, которые выживут, чтобы произвести новое растение. Если N большое, а q маленькое, то

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

Вероятностное распределение (p_k) стремится к $e^{-\mathscr{R}_0}\,(\mathscr{R}_0)^k/k!$, которое называется пуассоновским распределением. Затем Фишер вычислил вероятность исчезновения мутирующего гена в пределах поколений n, используя $x_0 = 0$, $x_n \approx e^{\mathscr{R}_0(x_{n-1}-1)}$ и числовые значения $N = 80$ и $q = 1/80$. В данном случае $\mathscr{R}_0 = Nq = 1$. Громоздкие вычисления показывают, что $x_{100} \approx 0{,}98$: мутирующий ген без селективного преимущества ($\mathscr{R}_0 = 1$) исчезает очень медленно. По-прежнему существует 2 % шанс, что этот ген будет присутствовать в популяции через 100 поколений. В 1922 году Фишер остановил дальнейшие исследования этой модели.

Продолжая работу Фишера, Халдейн впервые заметил в своей статье 1927 года, что при любом вероятностном распределении (p_k) таком, что $p_0 > 0$, уравнение $x = f(x)$ имеет ровно два корня в интервале $(0, 1]$, когда среднее число потомства, несущего мутирующий ген $\mathscr{R}_0$ строго больше 1, т.е. когда у мутирующего гена есть селективное преимущество. Более того, вероятность исчезновения x_∞, которая является пределом x_n при $n \to +\infty$, является наименьшим из двух корней $x = f(x)$: ген имеет ненулевую вероятность оседания в популяции. В отличие от Бьенэме и Курно, Халдейн представил доказательство этого вывода.

Действительно, $f'(x) \geqslant 0$ и $f''(x) \geqslant 0$ на интервале $[0, 1]$. Другими словами, функция $f(x)$ является растущей и выпуклой. Предположения $f(0) = p_0 > 0$ и

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

подразумевают, что уравнение $f(x) = x$ имеет ровно два решения на интервале $(0, 1]$: $x = 1$ и x^* таким образом, что $0 < x^* < 1$. Затем Халдейн сослался на статью Габриэля Кёнигса 1883 года, в которой показано, что если $x_n = f(x_{n-1})$ и $x_n \to x_\infty$, то $x_\infty = f(x_\infty)$ и $|f'(x_\infty)| \leqslant 1$. Когда $f'(1) > 1$, единственная возможность - это $x_\infty = x^*$.

Для случая пуассонского распределения с $f(x) = e^{\mathscr{R}_0(x-1)}$ и $\mathscr{R}_0$ чуть больше 1, вероятность исчезновения x_∞ очень близка к 1. Уравнение $f(x_\infty) = x_\infty$ тогда эквивалентно

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

Из этого следует, что

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

Халдейн пришел к выводу, что *вероятность того, что мутантный ген не вымрет, в два раза превышает его селективное преимущество* $\mathscr{R}_0 - 1$. Не цитируя Халдейна, Фишер взял в качестве примера в своей книге 1930 года случай, когда $\mathscr{R}_0 = 1{,}01$, что дает 2 % вероятность того, что мутантный ген не вымрет.

Халдейн стал членом Королевского общества в 1932 году. Он покинул Кембридж, чтобы стать профессором генетики, а затем биометрии в университетском колледже в Лондоне. Тогда его особенно интересовала генетика человека: оценка частоты мутаций, генетические карты хромосом и др. Помимо его научных книг (*Биология животных* в 1927 г. с Джулианом Хаксли, *Энзимы* в 1930 г. и *Причины эволюции* в 1932 г., *Биохимия генетики* в 1954 г.), он опубликовал большое количество статей по науке в прессе (например, о происхождении жизни) и некоторые эссе (*Неравенство человека* в 1932 г., *Философия биолога* в 1935 г., *Марксистская философия и наука* в 1938 г., *Наследственность и политика* в 1938 г. и *Научные достижения* в 1947 г.). После нескольких визитов в Испанию во время гражданской войны он пытался убедить свою страну построить укрытия от воздушных бомбардировок. Во время Второй мировой войны он работал над проблемами дыхания на подводных лодках. Член коммунистической партии с 1942 г.,он ушел в отставку в 1950 г. из-за официального отказа от генетической теории Менделя по идеологическим причинам в СССР под влиянием Лысенко. В 1957 году он поселился в Индии, где продолжил свои исследования, сначала в Индийском статистическом институте в Калькутте, а затем в Бхубанесваре. Став гражданином Индии, он умер в 1964 году.

Дополнительное чтение

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)

Глава 18

Эрланг и Штеффенсен о проблеме вымирания (1929–1933)

В 1929 году датский телефонный инженер Эрланг в очередной раз рассмотрел проблему исчезновения фамилий. Его соотечественник статистик Штеффенсен разработал полное решение этой проблемы. Он показал, в частности, что математическое ожидание числа потомков в каждом поколении растет в геометрической прогрессии, тем самым превращаясь в мост между стохастической и детерминистической моделями населения.

Агнер Краруп Эрланг (Erlang) родился в 1878 году в Лонборге, Дания. Его отец был школьным учителем. Между 1896 и 1901 годами молодой Эрланг изучал математику, физику и химию в Копенгагенском университете. Затем он преподавал несколько лет в средних школах, сохраняя при этом интерес к математике, особенно к теории вероятностей. Он познакомился с Йенсеном, главным инженером Копенгагенской телефонной компании и математиком-любителем, который убедил его в 1908 году присоединиться к новой научно-исследовательской лаборатории компании. Эрланг начал публиковать статьи о применении теории вероятностей для управления телефонными звонками. В 1917 году он открыл формулу времени ожидания, которая быстро стала использоваться телефонными компаниями по всему миру. Его статьи, впервые опубликованные на датском языке, были затем переведены на несколько других языков.

В 1929 году Эрланг заинтересовался той же проблемой вымирания, которую Бьенеме, Гальтон и Ватсон изучали до него применительно к фамилиям, а Фишер и Халдан - применительно к мутантным генам. Как и его предшественники, он не знал обо всех опубликованных работах. Снова назвав p_k вероятностью для одного человека иметь потомство k, он заметил, что вероятность x_n вымирания в пределах n поколений удовлетворяет уравнению

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

Рис. 18.1:
Эрланг (1878–1929)

с $x_0 = 0$. Он также заметил, что общая вероятность вымирания x_∞, которая является пределом x_n при $n \to +\infty$, является решением уравнения $x_\infty = f(x_\infty)$. Он понял, что $x = 1$ всегда является решением, и что существует другое решение между 0 и 1, когда среднее количество потомства $\mathscr{R}_0 = f'(1)$ больше 1. Но, похоже, он не смог понять, какое из этих двух решений является правильным. Как и Гальтон, он представил эту проблему в 1929 году в датском математическом журнале *Matematisk Tidsskrift*:

> «Вопрос 15. Когда вероятность того, что у человека есть k детей, равна p_k, где $p_0 + p_1 + p_2 + \cdots = 1$, найдите вероятность того, что его семья вымрет.»

К сожалению, Эрланг умер в том же 1929 году в возрасте 51 года. На самом деле, он умер бездетным[1].

Профессор актуарной математики Копенгагенского университета Йохан Фредерик Штеффенсен (Steffensen) ответил на вопрос Эрланга. Он опубликовал в 1930 году свое решение в том же датском журнале: вероятность исчезновения x_∞ всегда наименьший корень уравнения $x = f(x)$ в замкнутом интервале $[0, 1]$, как уже заметили Бьенеме и Холдейн. Доказательство Штеффенсена можно найти в современных учебниках.

[1]В память о нем Международный телефонный консультативный комитет в 1946 году решил назвать «эрланг» единицей измерения интенсивности телефонного трафика. «Эрланг» - это также название языка программирования, данное компанией Ericsson.

Действительно, мы видели, что вероятность исчезновения x_∞ - это решение $x = f(x)$ на закрытом интервале $[0, 1]$. Пусть x^* будет наименьшим таким решением. По определению $x^* \leqslant x_\infty$. Сначала Штеффенсен заметил, что $x^* = f(x^*) \geqslant p_0 = x_1$. Предположим по индукции, что $x^* \geqslant x_n$. Тогда $x^* = f(x^*) \geqslant f(x_n) = x_{n+1}$, так как функция $f(x)$ растущая. Так что $x^* \geqslant x_n$ для всех n. Взяв предел, имеем $x^* \geqslant x_\infty$. Так что $x_\infty = x^*$.

Штеффенсен также дал более формальное объяснение, почему $x = 1$ является единственным корнем $x = f(x)$, когда среднее число потомства $\mathscr{R}_0 = f'(1)$ меньше или равно 1 (рисунок 18.2a), и почему есть только один корень, отличный от $x = 1$, в случае, когда $\mathscr{R}_0 > 1$ (рисунок 18.2b). Обратите внимание, что $\mathscr{R}_0 = f'(1)$ - это наклон функции $f(x)$ при $x = 1$.

Он заметил, что для любого корня $x = f(x)$,

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

Предположив $x \neq 1$ и разделив на $1 - x$, мы получим

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

Когда x увеличивается с 0 до 1, правая часть уравнения (18.1) увеличивается с $1 - p_0$ до $\mathscr{R}_0 = f'(1)$. Если $\mathscr{R}_0 < 1$, то уравнение (18.1) не имеет решения. Если $\mathscr{R}_0 \geqslant 1$ и если мы исключим тривиальный случай $p_1 = 1$, то правая часть уравнения (18.1) является строго возрастающей функцией x. Иначе не было бы $k \geqslant 2$, такого, что $p_k \neq 0$, и $\mathscr{R}_0$ было бы равно $p_1 < 1$. Как следствие, когда $\mathscr{R}_0 \geqslant 1$, у уравнения (18.1) в интервале $[0, 1]$ есть одно и только одно решение.

Штеффенсен, который также был президентом Датского актуарного общества и Датского математического общества, был приглашен в Лондонский университет в 1930 году. Его британский коллега У.П. Элдертон рассказал ему о работе Гальтона и Ватсона. В 1933 году Штеффенсен опубликовал новую статью в анналах Института Анри Пуэнкаре, где в 1931 году он провел конференцию. Он резюмировал результаты своей статьи на датском языке и сравнил

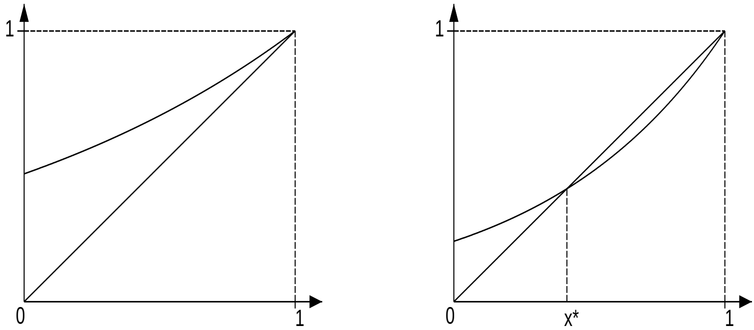


Рис. 18.2: График функций $y = x$ и $y = f(x)$ в примере главы 17, $f(x) = e^{\mathscr{R}_0(x-1)}$, с $\mathscr{R}_0 = 0{,}75 < 1$ (a) или $\mathscr{R}_0 = 1{,}5 > 1$ (b).

их с результатами Ватсона. Он также показал, что математическое ожидание числа потомков в поколении n равно $(\mathscr{R}_0)^n$.

Действительно, пусть $p_{k,n}$ - это вероятность того, что в поколении n есть k потомков, начиная с одного человека в поколении 0. В своей статье 1930 года Штеффенсен, как и его предшественники, заметил, что генерирующая функция

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

для поколения n удовлетворяет $f_1(x) = f(x)$ и

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

Пусть M_n будет математическим ожиданием количества потомков в поколении n. Тогда

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

Дифференцируя (18.2), мы получаем $f_n'(x) = f'(f_{n-1}(x)) \times$

$f'_{n-1}(x)$. Итак,

$$M_n = f'_n(1) = f'(f_{n-1}(1)) \times f'_{n-1}(1) = f'(1) \times M_{n-1} = \mathscr{R}_0 \times M_{n-1}.$$

Поскольку $M_1 = f'_1(1) = f'(1) = \mathscr{R}_0$, из этого следует, что $M_n = (\mathscr{R}_0)^n$ для всех n.

Следовательно, ожидаемое количество потомства увеличивается или уменьшается в геометрической прогрессии в зависимости от того, больше или меньше $\mathscr{R}_0$ по сравнению с 1. Ожидаемое количество потомства ведет себя как в детерминистических моделях роста населения, рассмотренных Эйлером, Мальтусом и др. Однако, даже когда $\mathscr{R}_0 > 1$, существует ненулевая вероятность x_∞, что семейство вымрет. В детерминистических моделях такая возможность не возникает.

Изученный Штеффенсеном и его предшественниками стохастический процесс до сих пор является основным элементом многих других реалистичных моделей динамики населения. Мы еще, в последний раз, упомянем об этой проблеме в главе 20. Что касается Штеффенсена, то он оставался профессором Копенгагенского университета до 1943 года и умер в 1961 году.

Дополнительное чтение

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

Глава 19

Райт и случайный генетический дрейф (1931)

В 1931 году американский биолог Сьюалл Райт разработал исследование стохастической модели в популяционной генетике, которое основано на тех же предположениях, что и в законе Харди-Вайнберга, за исключением того, что популяция не считается бесконечно большой. Частоты генотипов перестали быть постоянными. Одна из двух аллелей на самом деле исчезнет, но, возможно, через очень долгое время. Толкование этой модели оставалось предметом спора между Райтом и Фишером, который считал, что естественный отбор играет в эволюции более важную роль, чем стохастика.

Сьюалл Райт (Wright) родился в Массачусетсе в 1889 году. Он учился в небольшом колледже в Иллинойсе, где его отец преподавал экономику. После получения степени магистра биологии в Университете Иллинойса в Урбане и летней школы в лаборатории Cold Spring Harbor, Райт получил степень доктора философии в Гарвардском университете по вопросу о наследовании цвета шерсти у морских свинок. С 1915 по 1925 год он продолжал работать над экспериментами по скрещиванию морских свинок в отделе животноводства Министерства сельского хозяйства США в Вашингтоне. Для анализа этих экспериментов он разработал «Метод коэффициентов пути». Затем он поступил на факультет зоологии Чикагского университета.

Под влиянием статьи Фишера 1922 года о генетике популяции (см. главу 14) Райт написал в 1925 году длинную статью под названием *Эволюция в популяциях Менделя*, которая была окончательно опубликована в 1931 году. В частности, он изучал математическую модель, которая также косвенно появилась в книге Фишера в 1930 году *Генетическая теория естественного отбора*. Как и в законе Харди-Вайнберга, в этой модели рассматривается случай, когда существует только две возможные аллели A и a для одного локуса, но предполагается, что популяция не бесконечно велика. Смысл состоит в том, чтобы понять, как влияет это ограничение на генетический состав популяции. Так пусть N будет общим числом ин-

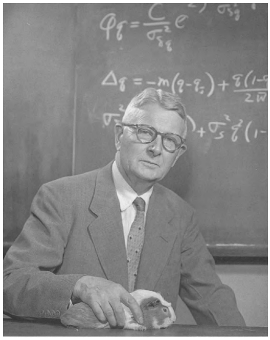


Рис. 19.1:
Райт (1889–1988)

дивидов, которое предполагается одинаковым для всех поколений. У каждого индивида есть две аллели. Таким образом, в популяции в каждом поколении есть в общей сложности $2N$ аллелей. Модель также предполагает, что спаривание происходит случайным образом. Если есть i аллелей A и $2N - i$ аллелей a в поколении n, то аллель, выбранная случайным образом среди индивидуумов в поколении $n+1$, будет A с вероятностью $\frac{i}{2N}$ и a с вероятностью $1-\frac{i}{2N}$. Таким образом, количество аллелей A в поколении $n+1$ будет равно j с вероятностью :

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

где $\binom{2N}{j}$ - биномиальный коэффициент, равный $(2N)!/j!/(2N-j)!$. Пусть X_n будет числом аллелей A в генерации n: это случайная переменная (рисунок 19.2).

Зная, что $X_n = i$, можно показать, что ожидание X_{n+1} равно i: это напоминает Харди-Вайнбергский закон, где частота аллели A оставалась неизменной на протяжении поколений.

Действительно, рассмотрим генерирующую функцию

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

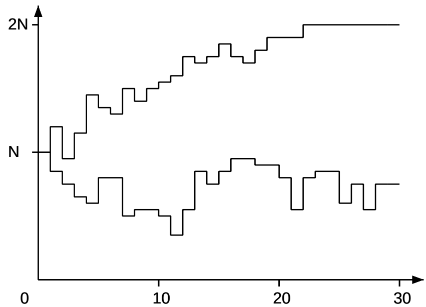


Рис. 19.2: Две симуляции, показывающие вариации числа X_n аллелей A в течение 30 поколений, если $N = 20$ и $X_0 = 10$.

Ожидание X_{n+1}, зная, что $X_n = i$ будет тогда

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

Однако в данной модели возможно, что, начиная с начального условия $X_0 = i$ для $0 < i < 2N$, событие $X_n = 0$ происходит случайно через некоторое количество поколений. В таком случае все аллели будут иметь тип a и X_n останутся равными 0 во всех последующих поколениях. Такая же фиксация произойдет и с аллелью A, если $X_n = 2N$ через некоторое количество поколений. Таким образом, когда предполагается бесконечно большая популяция, как в модели Харди-Вайнберга, две аллели не могут исчезнуть, так как их частоты остаются постоянными. При учете конечного размера популяции, как в модели Фишера-Райта, частоты двух аллелей колеблются, и одна из аллелей может исчезнуть и исчезнет.

Начиная с $X_0 = i$, можно легко вычислить вероятность Q_i для популяции, которая будет зафиксирована в состоянии $X = 0$. Действительно, Q_i должно удовлетворять «граничным условиям».

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

Более того,

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

потому что $p_{i,j}\, Q_j$ - это вероятность, зафиксированная в состоянии $X = 0$, начиная с $X_0 = i$ и проходя через $X_1 = j$. Поскольку

$\sum_{j=0}^{2N} p_{i,j} = 1$, то мы видим (19.2), что $Q_i = 1 - \frac{i}{2N}$ - это решение системы (19.3)-(19.4). Отсюда следует, что вероятность того, что, начиная с аллелей i типа A в популяции размером N, система развивается в сторону популяции, содержащей только аллель a, равна $1 - \frac{i}{2N}$. Аналогично вероятность того, что она эволюционирует в сторону совокупности, содержащей только аллель A, равна $\frac{i}{2N}$.

Райту удалось показать, что количество поколений, которые проходят до фиксации в одном из двух крайних состояний, составляет порядка $2N$ поколений (рисунок 19.3). Для населения в несколько миллионов индивидуумов это время было бы настолько длительным, что частоты аллелей можно было бы считать почти постоянными, как в законе Харди-Вайнберга.

Действительно, предположим, что в популяции в поколении 0 имеется i_0 аллелей типа A. Пусть $u_i^{(n)}$ будет вероятность того, что в населении в поколении n присутствуют i аллелей типа A. Затем .

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

для всех $j = 0, \ldots, 2N$. Мы уже видели это, когда $n \to +\infty$,

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

для всех $0 < i < 2N$. Райт заметил, что если $u_i^{(n)} = v$ для всех $i = 1, \ldots, 2N-1$, тогда

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

для всех $1 < j < 2N$, потому что $p_{0,j} = p_{2N,j} = 0$. Когда N достаточно велико,

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

значение интеграла, получаемого последовательным интегрированием по частям. Объединяя (19.5) и (19.6), мы получаем, наконец, что при $0 < j < 2N$

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

Таким образом, вероятности $u_j^{(n)}$ для всех $0 < j < 2N$ уменьшаются со скоростью около $1/2N$ за поколение. Эта скорость очень медленная, если N большое. Уменьшения практически не происходит, если, например, N равно миллиону.

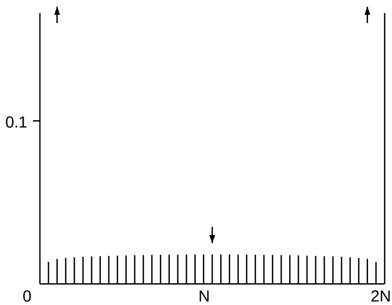


Рис. 19.3: Вероятность того, что в популяции есть i аллелей A ($i = 0, \ldots, 2N$ по горизонтальной оси) после 30 поколений, если $N = 20$ и $X_0 = 10$.

В 1922 году Фишер уже пытался оценить этот коэффициент фиксации ($1/2N$), но потерял множитель 2. В любом случае, два ученых разошлись во мнениях по поводу типичного для размножающихся популяций размера N. Для теории эволюции, работа Райта показала, что случайный генетический дрейф в небольшой популяции может быть механизмом происхождения видов. Биологи, работающие над классификацией видов, действительно заметили, что различия между видами или подвидами часто не имеют очевидного объяснения с точки зрения естественного отбора. В 1940-х-1950-х годах Фишер и его коллега Э.Форд отвергали эту идею. Они оба думали, что случайный генетический дрейф был пренебрежимо мал по сравнению с естественным отбором. Они ссылались, в частности, на свое исследование колебаний частоты генов в небольшой изолированной популяции мотыльков (*Panaxia dominula*) близ Оксфорда, где три генотипа для определенного гена (общая гомозигота, гетерозигота и редкая гомозигота) можно было различить в поле зрения. Еще одна дискуссия по поводу соответствующего влияния естественного отбора и случайного дрейфа касалась эволюции улиток рода *Cepaea*. Более реалистичные модели эволюции теперь сочетают в себе случайный дрейф, селекцию, мутацию, миграцию, неслучайное спаривание и др. Позднее роль случайного дрейфа была вновь подчеркнута японским ученым Мото Кимура с его «нейтральной теорией молекулярной эволюции». Другим результатом стало развитие коалесцентной теории (введенной Джоном Кингманом в 1982 г.), в которой прослеживается происхождение генов в прошлом до того момента, когда у них появился единый общий предок.

Райт стал членом Национальной академии наук в 1934 году. В течение многих лет он работал с Феодосием Добжанским над генетикой естественных популяций мух (*Drosophila pseudoobscura*) в районе Долины Смерти (США). В 1955 г. он ушел на пенсию из Чикагского университета, но еще пять лет проработал профессором в Университете Висконсина-Мэдисона. В период с 1968 по 1978 гг. он опубликовал четырехтомный трактат, обобщающий его работу *Эволюция и генетика населения*. Он получил премию Бальзана в 1984 г. и умер в 1988 г. в возрасте 98 лет.

Дополнительное чтение

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

Глава 20

Распространение генов (1937)

В 1937 г. Рональд Фишер и три русских математика - Колмогоров, Петровский и Пискунов - самостоятельно исследовали уравнение в частных производных для географическое распределение выгодного гена. Они показали, что частота гена ведет себя как волна, движущаяся с четко определенной скоростью в зависимости от преимущества гена и коэффициента диффузии. Их работы послужили отправной точкой для теории реакционно-диффузионных уравнений.

В 1937 году были опубликованы две статьи, знакомящие с новым подходом к изучению пространственной неоднородности в динамике населения. Фишер был автором первой статьи под названием *Волна продвижения выгодных генов*, появившейся в *Annals of Eugenics*. Он изучал пространственное распространение выгодного гена в популяции. В качестве упрощения он рассматривал пространство, сведенное к одному измерению и назвал $u(x,t)$ пропорцией популяции, находящейся в точке x в момент времени t, обладающей выгодным геном. Так что $0 \leqslant u(x,t) \leqslant 1$. Для включения естественного отбора он использовал уравнение (14.6) с непрерывной временной переменной

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

где a - положительный параметр. Для заданного значения x мы распознаем логистическое уравнение Верхульста (см. главу 6) с решением $u(x,t)$, которое стремится к 1 при $t \to +\infty$. Кроме того, Фишер предположил, что потомство особи, находящейся в точке x с выгодным геном, не остается в той же точке, а случайным образом рассеивается в окрестностях x. По аналогии с физикой он утверждал, что к уравнению для $u(x,t)$ необходимо добавить диффузионный член, что приведет к дифференциальному уравнению в частных производных

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

Когда коэффициент отбора a равен нулю, это сводится к уравнению диффузии, введенному Фурье в свою теорию тепла и впоследствии использованному Фиком для диффузии физических частиц. В 1904 году Рональд Росс начал рассматривать случайное рассеяние в демографической динамике. Тогда он задавался вопросом, как плотность комаров уменьшается с увеличением расстояния от места размножения. Эта проблема привлекла внимание Карла Пирсона и лорда Рейли. К 1937 году значительно вырос объем научной литературы по уравнениям диффузии, в частности, после работы Эйнштейна о броуновском движении.

Фишер показал, что существуют решения уравнения (20.1) вида $u(x,t) = U(x + v\,t)$, удовлетворяющие трем условиям

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

при условии, что $v \geqslant v^*$, где

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

Эти решения связывают устойчивое состояние $u = 1$ с выгодным геном, а устойчивое состояние $u = 0$ - с отсутствием такого гена. Они представляют собой волны, распространяющиеся со скоростью v в направлении уменьшения значений x. Действительно, $u(x - v\,T, t + T) = u(x,t)$: часть волны, которая находилась в позиции x в момент времени t, перемещается в позицию $x - v\,T$ в момент времени $t + T$.

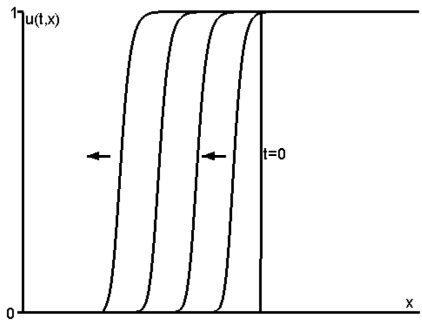


Рис. 20.1: Распространение справа налево выгодного гена со скоростью v^*. Частота гена $u(t,x)$ при $t = 0$ является ступенчатой функцией.

Действительно, установив $z = x + v\,t$, Фишер заметил, что если $u(x,t) = U(z)$, то

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

Если u является решением уравнения (20.1), то

$$v\,U'(z) = a\,U(z)\,(1 - U(z)) + D\,U''(z)\,. \tag{20.2}$$

Когда u близко к 0, т.е. когда $z \to -\infty$, Фишер ожидал, что $U(z) \to 0$ и $U'(z) \to 0$. Обозначив k предел $U'(z)/U(z)$, когда $z \to -\infty$, мы знаем из правила Лопиталя, что $U''(z)/U'(z)$ также стремится к k. Поэтому,

$$U''(z)/U(z) = [U''(z)/U'(z)] \times [U'(z)/U(z)] \longrightarrow k^2.$$

Разделив уравнение (20.2) на $U(z)$ и положив z стремящимся к $-\infty$, мы получаем уравнение второго порядка $D\,k^2 - v\,k + a = 0$. Но k должно быть действительным числом. Поэтому дискриминант этого уравнения должен быть положительным: $v^2 - 4\,a\,D \geqslant 0$, или

$$v \geqslant 2\sqrt{a\,D} = v^*.$$

Следовательно, $v \geqslant v^*$ является необходимым условием существования волны, распространяющейся со скоростью v. Также это условие является достаточным, как поясняется ниже.

Фишер заметил, что для большого класса начальных условий, например, для функции с начальными условиями в виде ступеньки: $u(x,0) = 0$ для $x < 0$, $u(x,0) = 1$ для $x \geqslant 0$, выбирается только та волна, которая распространяется точно со скоростью v^*. На рисунке 20.1 показано, как это прерывистое начальное условие постепенно превращается в гладкую волну, распространяющуюся в направлении уменьшения x со скоростью v^*.

В том же 1937 году, независимо от работ Фишера, Андрей Николаевич Колмогоров, Иван Георгиевич Петровский и Николай Семенович Пискунов изучали одну и ту же проблему распространения доминантного гена.

Колмогоров родился в 1903 году в Тамбове, Россия. Изучая математику в Московском государственном университете он выполнил ряд важных работ по тригонометрическим рядам. Он стал научным сотрудником Института математики и механики в 1929 году и профессором университета в 1931 году. Он работал над стохастическими процессами и их связью с дифференциальными уравнениями и уравнениями в частных производных. В 1933 году он опубликовал трактат, заложивший современные основы теории вероятностей. Его исследовательские интересы включали топологию, теорию аппроксимации, марковские цепи, броуновское движение, а также приложения к биологическим проблемам. В 1935 году он опубликовал статью по генетике, в которой обсуждались результаты работ Харди, Фишера и Райта. В 1936 г. опубликовал статью об обобщении системы Лотка-Вольтерра.

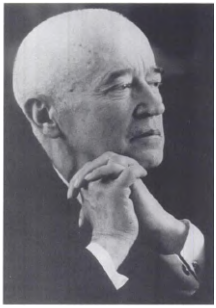



Рис. 20.2: Колмогоров (1903-1987) и Петровский (1901-1973)

Петровский родился в 1901 году в Севске. Он также изучал математику в Московском государственном университете, где он стал профессором в 1933. Он работал в основном над теорией дифференциальных уравнений в частных производных и топологией действительных алгебраических кривых, а также написал ряд статей по обычным дифференциальным уравнениям и по теории вероятности. Пискунов, родившийся в 1908 году, также закончил факультет математики Московского государственного университета.

В течение 1930-х годов Колмогоров имел контакты с А. С. Серебровским, основоположником популяционной генетики в Москве. В то время защита генетики в СССР становилась все более опасной из-за взлета Лысенко, агронома, которому удалось убедить Сталина в том, что генетика Менделя - это всего лишь «буржуазная псевдонаука». Седьмой Международный конгресс по генетике, первоначально запланированный на 1937 год в Москве, был отменен. Многие советские генетики были казнены или отправлены в трудовые лагеря.

В своей статье 1937 г. под названием *Исследование уравнения диффузии с ростом количества вещества и его применение к биологической проблеме*, опубликованной в Вестнике МГУ, Колмогоров, Петровский и Пискунов тем не менее использовали математическую модель, основанную на генетике Менделя. Их модель представляла собой дифференциальное уравнение в частных производных вида

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

где $u(x,t)$ - это частота выгодного гена в точке x и в момент времени t. Предполагается, что функция $f(u)$ удовлетворяет нескольким условиям: $f(0) = f(1) = 0$, $f(u) > 0$ если $0 < u < 1$, $f'(0) > 0$ и $f'(u) < f'(0)$ если $0 < u \leqslant 1$. Авторы показали результат, аналогичный Фишеру, но с более строгим доказательством: если исходное условие таково, что $0 \leqslant u(x,0) \leqslant 1$, $u(x,0) = 0$ для всех $x < x_1$ и $u(x,0) = 1$ для всех $x > x_2 \geqslant x_1$, то ген распространяется со скоростью

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

Поиск решения $u(x,t) = U(z)$, где $z = x+vt$ приводит к очевидному обобщению уравнения (20.2) $vU'(z) = f(U(z)) + D\,U''(z)$. Это дифференциальное уравнение второго порядка можно переписать как систему дифференциальных уравнений первого порядка.

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

Напомним, что $U(z)$ должно быть таким, что $U(z) \to 0$ при $z \to -\infty$ и $U(z) \to 1$ при $z \to +\infty$. В окрестности устойчивого состояния $(U = 0, p = 0)$ системы (20.4), имеем $f(U) \approx f'(0)\,U$. Так что (20.4) можно аппроксимировать линейной системой

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

Поиск экспоненциальных решений в виде $U(z) = U_0\,e^{kz}$ и $p(z) = p_0\,e^{kz}$ дает характеристическое уравнение $Dk^2 - vk + f'(0) = 0$, как в статье Фишера. Опять же k должно быть действительным (иначе u колебалось бы и приняло бы отрицательные значения). Таким образом, $v \geqslant 2\sqrt{f'(0)D} = v^*$. Тогда два корня для k будут действительными и положительными. Если $v > v^*$, то оба корня разные, а устойчивое состояние $(U = 0, p = 0)$ - нестабильный узел. Если $v = v^*$, то два корня идентичны и $(U = 0, p = 0)$ - нестабильный дегенеративный узел, как показано на рисунке 20.3. Аналогично, система (20.4) в окрестности устойчивого состояния $(U = 1, p = 0)$ приводит к линейной системе

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

и к характеристическому уравнению $Dk^2 - vk + f'(1) = 0$. Дискриминант $v^2 - 4Df'(1) \geqslant 0$ с $f'(1) \leqslant 0$. Если $f'(1) < 0$, то существуют два вещественных корня противоположного знака и $(U = 1, p = 0)$ является седловидной точкой. Если $f'(1) = 0$, то один корень равен нулю, а другой положительный (см. рисунок 20.3). Детальный анализ показывает, что для всех

$$v \geqslant 2\sqrt{f'(0)D}$$

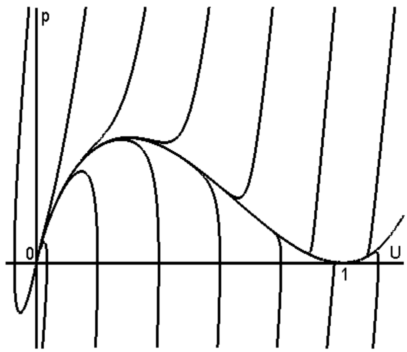
существует единственная интегральная кривая, соединяющая два устойчивых состояния $(U = 0, p = 0)$ и $(U = 1, p = 0)$, как в особом случае рисунка 20.3.

Колмогоров, Петровский и Пискунов неукоснительно показывали, что дифференциальное уравнение в частных производных (20.3) имеет единственное решение $u(x,t)$, удовлетворяющее исходному условию, такое, что $0 < u(x), t) \leqslant 1$ для всех x и $t > 0$, что $u(x,t)$ остается растущей функцией x, если она такова при $t = 0$, и, наконец, что $u(x,t)$ действительно сходится в направлении волнового профиля, распространяющегося со скоростью v^*. Доказательства здесь слишком длинные, чтобы их можно было привести.

Обратите внимание, что функция $f(u) = a\,u\,(1-u)$, используемая Фишером, действительно удовлетворяет всем этим условиям с $f'(0) = a$. Опираясь на уравнение (14.5), Колмогоров, Петровский и Пискунов рассмотрели функцию $f(u) = a\,u\,(1-u)^2$, которая удовлетворяет тем же условиям и дает ту же скорость распространения.

Статьи Фишера и Колмогорова, Петровского и Пискунова стали отправной точкой для построения многих математических моделей с географической диффузией в генетике, экологии и эпидемиоло-

Рис. 20.3: На диаграмме (U, p) показаны некоторые интегральные кривые системы (20.5) и, в частности, единственная кривая, соединяющая $(U = 1, p = 0)$ с $(U = 0, p = 0)$, дающая форму распространяющейся волны. Здесь $f(u) = au(1 - u)^2$, $a = 1$, $D = 1$ и $v = v^* = 2$.



гии. Эти модели известны как «реакционно-диффузионные системы».

Что касается Колмогорова, то, начиная с 1938 года, он изучал также проблему исчезновения фамилий, рассматриваемых Бьенеме, Гальтоном, Ватсоном, Фишером, Халдейном, Эрлангом и Штеффенсеном: стохастический процесс, который является общим для всех этих работ, он назвал «ветвящийся процесс». В 1939 году он стал членом Академии наук СССР. Позднее он внес важный вклад в проблему турбулентности в механике жидкостей (1941), в теорию динамических систем, связанных с механикой небесных тел (1953) и в теорию информации (начиная с 1956). Он также внес вклад в написание энциклопедии и школьных и университетских учебников, помог основать экспериментальную среднюю школу и редактировал научно-популярный журнал. Он получил множество международных премий (в том числе премию Бальцана в 1963 г. и премию Вольфа в 1980 г.) и умер в 1987 г. в Москве.

Петровский стал деканом механико-математического факультета МГУ в 1940 году. С 1951 года и до своей смерти в 1973 году он был ректором университета. С 1946 года он был действительным членом Академии наук СССР и президентом Международного конгресса математиков, который проходил в Москве в 1966 году. Он также написал учебники по обычным дифференциальным уравнениям, дифференциальным уравнениям в частных производных и интегральным уравнениям.

Пискунов стал профессором военной академии. Его учебник по дифференциальному и интегральному исчислению использовался во многих технических университетах. Умер он в 1977 году.

Дополнительное чтение

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)
   → 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, S. 4–30. Gordon and Breach (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)

Глава 21

Матрица Лесли (1945)

В 1945 году британский эколог П.Х. Лесли проанализировал матричную модель возрастно-структурированной популяции грызунов, адаптировав работу Лотки к дискретно-временным рамкам. Он показал, что скорость роста соответствует собственному значению и стабильная возрастная структура собственному вектору матрицы. Он также численно оценил чистый коэффициент воспроизводства $\mathscr{R}_0$ серой крысы.

Патрик Холт Лесли (Leslie) родился в 1900 году недалеко от Эдинбурга в Шотландии. Он учился в колледже Крайст-Черч Оксфордского университета и в 1921 году получил степень бакалавра физиологии. Но он не смог закончить свои медицинские исследования из-за проблем со здоровьем. После нескольких лет работы в качестве ассистента по бактериологии в департаменте патологии, он обратился к статистике и в 1935 году присоединился к Бюро по вопросам животных популяций, новом исследовательском центре, созданном Чарльзом Элтоном. Целью этого центра было изучение колебаний популяций животных с помощью полевых исследований и лабораторных экспериментов. Большая часть исследований проводилась на грызунах: анализ циклов зайца и его хищника рыси с использованием архивов Компании Гудзонова залива в Канаде, наблюдение за территориальным расширением популяции серой белки за счет рыжей белки в Англии, сбор данных о полёвках в окрестностях Оксфорда и т.д. Лесли применил к данным о полёвках методы, разработанные Лоткой для демографии человека. Во время Второй мировой войны исследования центра были сосредоточены на методах борьбы с крысами и мышами в шахтах.

В 1945 году Лесли опубликовал свою самую известную статью в *Biometrika*, журнале, который был основан Гальтоном, Пирсоном и Уэлдоном в 1901 году. Статья была озаглавлена *Об использовании матриц в математическом моделировании некоторых популяций.* Лесли рассматривал модель роста числа самок в популяции животных, например, в популяции крыс (но это могут быть и люди). Популяция разделена на $K+1$ возрастную группу : $P_{k,n}$ - это число са-

Рис. 21.1: Лесли (1900–1972)

мок в возрасте k на момент времени n ($k = 0, 1, \ldots, K$; $n = 0, 1, \ldots$). Обозначим через f_k рождаемость в возрасте k, а точнее количество дочерей, родившихся на одну женщину в промежутке между n и $n+1$. Тогда K - это максимальный возраст с ненулевой рождаемостью ($f_K > 0$). Обозначим s_k вероятность того, что животное в возрасте k выживет, по крайней мере, до тех пор, пока не достигнет возраста $k+1$. Тогда возрастная структура популяции будет задаваться следующим набором уравнений:

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

Все цифры f_k являются неотрицательными, в то время как s_k удовлетворяет неравенствам $0 < s_k < 1$. В конце девятнадцатого - начале двадцатого века у математиков появилась традиция писать такие системы уравнений в сокращенном виде[1]

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

где P_n - вектор-столбец $(P_{0,n}, \ldots, P_{K,n})$, а M - квадратная матрица (т.е. таблица чисел с $K+1$ строкой и $K+1$ столбцом)

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

[1]Это значит, что $P_{k,n+1} = M_{k,0}\, P_{0,n} + M_{k,1}\, P_{1,n} + \cdots + M_{k,K}\, P_{K,n}$ для всех k.

Чтобы понять поведение системы (21.1) как функции времени, Лесли искал геометрически возрастающее или убывающее решение $P_n = r^n V$. Число r и вектор V должны соответствовать условию

$$M V = r V . \qquad (21.2)$$

В этом случае r называется собственным значением, а V - собственным вектором матрицы M. Другими словами, задача состоит в том, чтобы найти возрастное распределение V, которое на каждом шаге умножается на постоянную r. Следуя терминологии Лотки, такие распределения называются стабильными. Возвращаясь к более обычным обозначениям, уравнение (21.2) можно переписать как

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

Из последних K уравнений следует, что

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

Подставив это в первое уравнение, сократив на V_0 и умножив на r^K, Лесли получил характеристическое уравнение

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

Это полиномиальное уравнение степени $K+1$ относительно r . Таким образом, существуют $K+1$ действительных или комплексных корня $r_1, \ldots, r_{K+1}$. Более того, Лесли заметил (используя правило знаков Декарта для полиномов), что существует только один действительный положительный корень. Назовем его r_1.

Лесли также предположил, что при наиболее биологически реалистичных условиях (которые можно формализовать, используя теорию Перрона-Фробениуса для неотрицательных матриц), собственное значение r_1 строго больше модуля всех остальных действительных или комплексных собственных значений ($r_2, \ldots, r_{K+1}$). Кроме того, все корни (21.3) обычно различны. Для каждого собственного значения r_i можно найти ассоциированный собственный вектор. Пусть Q будет квадратной матрицей размера $(K+1, K+1)$, столбцы которой содержат собственные векторы, соответственно связанные с $r_1, \ldots, r_{K+1}$, тогда $M Q = Q D$, где D - диагональная матрица $[r_1, \ldots, r_{K+1}]$.

Так что $M = Q\,D\,Q^{-1}$, и

$$P_n = M^n\,P_0 = Q\,D^n\,Q^{-1}\,P_0\,.$$

Обратите внимание, что D^n - это диагональная матрица

$$[(r_1)^n, \dots, (r_{K+1})^n],$$

и что

$$D^n/r_1^n \longrightarrow \mathscr{D} = [1, 0, \dots, 0],$$

когда $n \to +\infty$, потому что $r_1 > |r_i|$ для всех $i \neq 1$. Поэтому $P_n/(r_1)^n$ сходится к $Q\,\mathscr{D}\,Q^{-1}\,P_0$.

Каждый компонент вектора возрастной структуры P_n увеличивается или уменьшается подобно $(r_1)^n$. Если $r_1 > 1$, то население увеличивается экспоненциально. Если $r_1 < 1$, то оно уменьшается по экспоненте. Из уравнения (21.3) можно легко показать, что условие $r_1 > 1$ истинно, если и только если параметр $\mathscr{R}_0$, заданный как

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

строго больше 1. Обратите внимание, что $s_0\,s_1 \cdots s_{k-1}$ - это вероятность дожить, по крайней мере, до возраста k. Таким образом, параметр $\mathscr{R}_0$ - это среднее число дочерей, рожденных от одной самки за всю ее жизнь, и является аналогом формул (10.2), (12.2) и (16.9). Приведенная модель является своего рода дискретно-временным аналогом работы Лотки (см. главу 10) и обобщением, включающим в себя возраст-зависимую рождаемость из работы Эйлера (см. главу 3).

Лесли проиллюстрировал свой метод данными, опубликованными американским коллегой, о коэффициентах рождаемости и выживаемости f_k и s_k для серой крысы. После некоторых статистических операций по обоснованному восполнению данных, он получил $\mathscr{R}_0 \approx 26$.

В настоящее время, матричная формулировка Лесли проблемы динамики популяции используется многими биологами. Вычисления значительно упрощаются современными компьютерами и научным программным обеспечением, способным вычислять собственные значения и собственные векторы любой матрицы. Можно легко вычислить как параметр $\mathscr{R}_0$, так и скорость роста r_1.

После Второй мировой войны Лесли использовал свой метод для расчета скорости роста других видов животных: птиц, жуков и т.д. Он также работал над стохастическими моделями, над моделями конкуренции между видами и над анализом данных метода мечения-повторного отлова. Ушел на пенсию в 1967 году. В том же году, когда Чарльз Элтон также вышел на пенсию, Бюро по вопросам животных популяций прекратило свое существование в качестве независимого исследовательского центра и стало частью кафедры зоологии Оксфордского университета. Лесли умер в 1972 году.

Дополнительное чтение

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)

Глава 22

Перколяция и эпидемии (1957)

В 1957 г. Хаммерсли и Бродбент рассматривали распространение «жидкости» в бесконечной регулярной квадратной сети, где с заданной вероятностью связаны два соседних узла. Среди возможных примеров они упомянули распространение эпидемии во фруктовом саду. Они показали, что существует критическая вероятность, ниже которой не может произойти крупная эпидемия, и выше которой с положительной вероятностью происходят крупные эпидемии. Их статья стала отправной точкой теории перколяции.

Джон Майкл Хаммерсли (Hammersley) родился в 1920 году в Шотландии, где его отец работал на американскую компанию по экспорту стали. Он начал учиться в Эммануэль-Колледже Кембриджского университета, но в 1940 году был вынужден поступить на службу в армию. Он работал над усовершенствованием вычислений для артиллерии. После окончания учебы в 1948 году он стал ассистентом в Оксфордском университете в группе, работавшей над проектированием и анализом экспериментов. В 1955 году он поступил на работу в Центр научных исследований по атомной энергии в Харвелле недалеко от Оксфорда.

Рис. 22.1:
Хаммерсли (1920–2004)

Саймон Ральф Бродбент родился в 1928 году. Он изучал инженерное дело в Кембридже, математику в Магдаленском колледже в Оксфорде (где он также писал стихи) и начал работу над докторской диссертацией «Некоторые тесты по отклонению от равномерной дисперсии» по статистике в Имперском колледже в Лондоне. Для получения степени доктора философии он получил поддержку Британской ассоциации по исследованию угля для изучения статистических проблем, которые могут быть связаны с добычей угля.

В 1954 году в Королевском статистическом обществе в Лондоне был проведен симпозиум по методам Монте-Карло, спонсором которого выступил Центр научных исследований по атомной энергии. В этих методах, инициированных в 1940-х годах Джоном фон Нейманом, Станиславом Уламом и Николасом Метрополисом в Лос-Аламосской лаборатории, используются стохастические компьютерные модели для оценки неизвестных математических величин. Хаммерсли представил на Лондонском симпозиуме доклад, который он подготовил в сотрудничестве с Мортоном, коллегой из Харвелла. Этот доклад был также опубликован в Журнале Королевского статистического общества. В ходе дискуссии, последовавшей за выступлением на симпозиуме, Бродбент упомянул интересную проблему, которую можно было бы изучить по методу Монте-Карло: учитывая, что обычная сеть пор в двух или трех измерениях такова, что две соседние поры соединены с вероятностью p, какая часть сети будет заполнена газом, если он будет введен через одну из этих пор? На самом деле Бродбент думал о проектировании противогазов для шахтеров и, в частности, о размерах пор, которые необходимы для их эффективного функционирования.

Затем Хаммерсли начал работать с Бродбентом над проблемой противогаза. Они поняли, что это всего лишь прототип еще не изученного семейства проблем: распространение «жидкости» (значение в зависимости от контекста) в случайной среде. Хаммерсли назвал это «перколяцией» по аналогии с тем, что происходит в кофеварке. В Центре исследований по атомной энергии Хаммерсли также имел доступ к некоторым из самых мощных компьютеров своего времени для приложения методов Монте-Карло к проблемам перколяции.

В 1957 году Бродбент и Хаммерсли, наконец, опубликовали первую статью по математической теории перколяции. Среди рассмотренных ими примеров была модель динамики населения, а именно распространение эпидемии в саду. Предполагается, что деревья очень большого фруктового сада располагаются в узлах квадратной сети. Каждое из четырех ближайших деревьев данного зараженного дерева имеет вероятность заражения p. Вопрос в том, будет ли заражено большое количество деревьев или эпидемия останется локальной. Это, конечно же, зависит от вероятности p, которая, в свою очередь, связана с расстоянием, разделяющим деревья, т.е. шириной ячейки сети.

Бродбент и Хаммерсли посмотрели на предельный случай, когда фруктовый сад бесконечен и покрывает всю плоскость, а в начале - только одно зараженное дерево. Пусть $f(p)$ - это вероятность того, что бесконечное количество деревьев будет заражено из этого источника. Предполагается, что $f(p)$ будет увеличивающейся функцией p с $f(0) = 0$ и $f(1) = 1$. Их основным результатом было утверждение, что существует критическая вероятность p^*, $0 < p^* < 1$ такая, что:

- если $p < p^*$, то $f(p) = 0$, поэтому заражено только конечное число деревьев;
- если $p > p^*$, то $f(p) > 0$ и тогда бесконечное количество деревьев может быть заражено.

Доказательство проводится на примере рассмотрения «произвольных перемещений» по бесконечному саду, начиная с источника инфекции. Эти перемещения проходят через определенное количество соседних деревьев (напомним, что у каждого дерева есть четыре соседа), не посещая ни одного дерева более одного раза. Произвольное перемещение с шагом в n - это путь заражения с вероятностью p^n, так как инфекция может передаваться от каждого посещаемого дерева к следующему с вероятностью p. Теперь пусть $q(j, n)$ будет вероятность того, что среди всех n-шаговых произвольных перемещений есть именно j таких перемещений, которые являются путями заражения. Если есть бесконечное количество зараженных деревьев, то для всех целых n существует хотя бы одно n-шаговое произвольное перемещение, которое является путём заражения. Итак,

$$0 \leqslant f(p) \leqslant \sum_{j=1}^{+\infty} q(j, n) \leqslant \sum_{j=1}^{+\infty} j\ q(j, n)$$

для всех n. Но $\sum_{j=1}^{+\infty} j\, q(j,n)$ - это ожидаемое количество произвольных перемещений в n шагов, которые являются путями заражения. Это число равно $p^n\, s(n)$, где $s(n)$ - общее число n-шаговых произвольных перемещений. Хаммерсли может показать в сопроводительной статье, что $s(n)$ растет как $e^{\kappa n}$ при $n \to +\infty$, где κ называется соединительной константой. Если $p < e^{-\kappa}$, то $p^n s(n)$ стремится к 0 при $n \to +\infty$ и $f(p) = 0$. Таким образом, $p^* \geqslant e^{-\kappa} > 0$.

Поэтому на практике лучше, если деревья располагаются не слишком близко, чтобы держать p ниже p^* в случае эпидемии. Но чем ближе деревья, тем выше производительность с гектара. Таким образом необходимо найти компромисс.

Как заметили Бродбент и Хаммерсли, существует определенное сходство между существованием критической вероятности в перколяционных процессах и существованием порога в ветвистых процессах (см. главу 7).

Можно попробовать численно оценить критическую вероятность p^*. Для этого зафиксируем значение p и аппроксимируем бесконечную сеть сетью конечных квадратов размером $N \times N$ с достаточным размером N. Предположим, например, что заражено дерево в середине сети. С помощью компьютера можно случайным образом выбирать, какие деревья могут заразить другие. На рисунке 22.2 и рисунке 22.3 показаны случайно выбранные пути заражения краями, как на графике. На рисунке 22.2, p меньше чем p^*. На рисунке 22.3, p больше p^*. Легко определить, какие деревья могут быть заражены, а именно те, до которых можно добраться по пути следования краев, начинающихся от зараженного дерева в центре. На рисунках они обозначены маленькими черными квадратами.

Затем можно проверить, достигла ли эпидемия хотя бы границы сети $N \times N$. Если это так и если N достаточно большое, то можно считать, что количество зараженных деревьев «почти бесконечно». Многократно повторяя подобную симуляцию, можно найти приблизительное значение вероятности $f(p)$ такое, что количество зараженных деревьев бесконечно (это метод Монте-Карло). Наконец, позволив p варьироваться от 0 до 1, можно получить аппроксимацию порога p^*, наименьшего значения, равного $f(p) > 0$, если $p > p^*$.

В статье Бродбента и Хаммерсли приводились только доказательства существования порога p^*. В последующие годы Хаммерс-

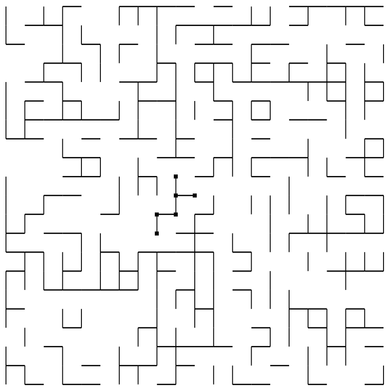


Рис. 22.2: Перколяция с $p = 0{,}4$.

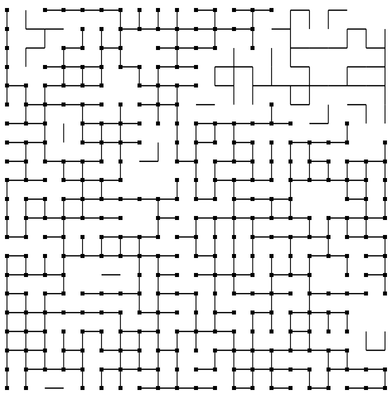


Рис. 22.3: Перколяция с $p = 0{,}55$.

ли продолжал развивать математическую теорию перколяции, в то время как Бродбент обратился к другим предметам. С развитием компьютеров в 1970-х годах стало легче проводить симуляции, описанные выше (рисунок 22.4). Затем появилось предположение, что $p^* = 1/2$. Этот результат был окончательно подтвержден в 1980 году Гарри Кестеном из Корнелльского университета.

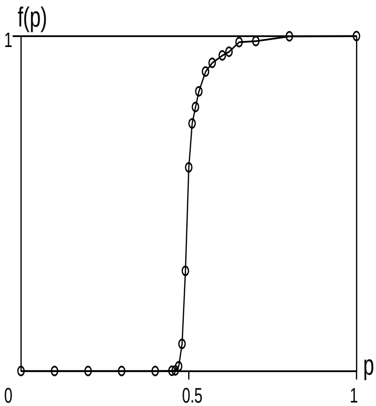


Рис. 22.4: Вероятность $f(p)$ того, что бесконечно большое количество деревьев заразится как функция p. Кривая получена путем проведения 1000 симуляций в сети размера 200×200.

В 1959-1969 гг. Хаммерсли работал в Институте экономики и статистики Оксфордского университета. Он стал стипендиатом Тринити-колледжа. В 1964 году Хаммерсли совместно с Дэвидом Хэндскомбом опубликовал книгу под названием *Методы Монте-Карло*. В 1976 году он был избран в Королевское общество. В 1987 году он ушел на пенсию, но продолжил посещать Оксфордский центр промышленной и прикладной математики. Скончался Хаммерсли в 2004 году.

Бродбент получил докторскую степень в Имперском колледже в 1957 году. Он нашел работу в промышленной компании «Объединенные производители стеклянных бутылок». После десяти лет работы в промышленности он начал работать в информационном агентстве «Лондонская Биржа Пресс», которое проводило научные исследования читательской аудитории. Агентство было куплено в 1969 году американской рекламной компанией Лео Бернетт. Бродбент работал над тем, как измерить эффективность рекламы и опубликовал несколько книг на эту тему: *Расходы на рекламу* (1975), *Рекламный бюджет* (1989), *Ответственная реклама* (1997) и *Когда рекламировать* (1999). В 1980 году Бродбент помог учредить премию «За эффективность рекламы». Несколько лет он проработал в головном офисе Лео Бернетта в Чикаго в качестве директора по экономике бренда. Он также руководил собственной консалтинговой компанией BrandCon Limited. Умер Бродбент в 2002 году.

Дополнительное чтение

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)

Глава 23

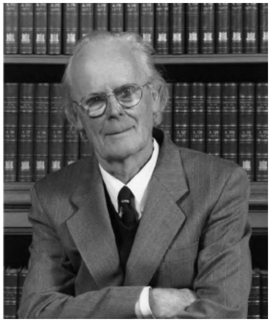
Теория игр и эволюция (1973)

В 1973 году Мэйнард Смит и Прайс опубликовали статью, в которой анализировали, почему животные избегают применения своего самого опасного оружия во внутривидовых конфликтах. Их модель использовала теорию игр и была одной из тех, которые положили начало применению этой математической теории в эволюционных задачах.

Джон Мэйнард Смит (Maynard Smith) родился в Лондоне в 1920 году. Его отец был хирургом и умер, когда сыну было восемь лет. Мэйнард Смит учился в Итонском колледже и обратился к инженерным специальностям в Тринити-колледже Кембриджского университета. В то время он был членом Коммунистической партии Великобритании. В 1939 году, когда началась война, он попытался поступить добровольцем на службу в армию, но был отвергнут из-за плохого зрения. Он закончил инженерное училище и несколько лет работал над конструированием военных самолетов. В конце концов, он решил обратиться к биологии, изучая генетику в университетском колледже в Лондоне с Халдейном в качестве руководителя. В 1952 году он стал преподавателем зоологии. После событий 1956 года в Венгрии он вышел из коммунистической партии. Его первая книга, озаглавленная *Теория эволюции*, была опубликована в 1958 году. В 1965 году он стал профессором биологии во вновь созданном Сассекском университете. Затем он опубликовал еще две книги: *Математические идеи в биологии* (1968) и *Об эволюции* (1972).

Джордж Р. Прайс родился в 1922 году в США. Изучал химию в Чикагском университете, получив докторскую степень в 1946 году после работы над Манхэттенским проектом по созданию атомной бомбы. В 1950 году он стал младшим научным сотрудником по медицине в Университете Миннесоты. Позже он работал независимым журналистом в различных журналах, а затем вернулся к исследованиям в компании IBM. В 1967 году, после лечения рака щитовидной железы, он обосновался в Англии и обратился к изучению совершенно другой темы: эволюционной биологии. С 1968 года

Рис. 23.1: Мэйнард Смит (1920–2004)



он работал в Лондоне в лаборатории Гальтона в университетском колледже. Его первая работа в этой новой области, *Избранное и ковариативность*, была опубликована с помощью У. Д. Гамильтона в 1970 г. в номере *Nature* и содержала то, что сейчас называется уравнением Прайса.

Прайс также представил еще одну работу в журнал *Nature*, на этот раз о конфликтах животных. Но она не имела подходящего формата для этого журнала. Поэтому Мэйнард Смит, который был рецензентом, предложил подготовить более короткую версию. Прайс начал работать над чем-то другим, в то время как Мэйнард Смит начал разрабатывать идею Прайса самостоятельно. Наконец, Мэйнард Смит и Прайс опубликовали совместную статью под названием *Логика конфликта животных*, которая в *Nature* была опубликована в 1973 году. Статья внесла интересный вклад в использование теории игр в эволюционной биологии. До этого теория игр разрабатывалась в основном для экономики и политики, особенно после книги Джона фон Неймана и Оскара Моргенштерна 1944 года под названием *Теория игр и экономическое поведение.* Отправной точкой для Мейнарда Смита и Прайса стал следующий вопрос: как получилось, что в конфликтах между животными одного вида «оружие», находящееся в их распоряжении (рога, когти, яд и т.д.), редко используется для убийства? Следуя идеям Дарвина о борьбе за жизнь, более агрессивные животные должны выигрывать больше сражений и иметь в своем распоряжении более многочисленное потомство, что приведет к эскалации использования «оружия». Заметим, что это было время Холодной войны, так что эта тема также имела и политическую окраску.

Мэйнард Смит и Прайс представили последовательность игр, в которых два животного могут участвовать в борьбе за ресурс, например, территорию в благоприятной среде обитания. В упрощенной презентации, которую Мэйнард Смит будет использовать в своей книге 1982 года *Эволюция и теория игр*, каждое животное принимает либо «стратегию ястреба», либо «стратегию голубей». Далее мы говорим просто о ястребах и голубях, но имеем в виду стратегии, принятые животными одного вида. Пусть $V > 0$ будет значением ресурса, то есть если $\mathscr{R}_0$ - нормальное среднее количество потомства животного, то у победителя конкурса в среднем $\mathscr{R}_0 + V$ потомство.

Если ястреб встречает другого ястреба, то он борется за ресурс: победитель получает ресурс стоимостью V, проигравший довольствуется ресурсом стоимости $C > 0$. Каждый из двух ястребов имеет вероятность выигрыша в соревновании, равную $1/2$ и такую же вероятность проигрыша. Таким образом, стоимость ожидаемого выигрыша от борьбы между двумя ястребами составляет $\frac{1}{2}(V - C)$ для двух участников. Если же ястреб встречает голубя, то ястреб получает ресурс V, голубь убегает без боя, и его выигрыш равен 0. Наконец, если встречаются два голубя, то один из них получает ресурс V, другой бежит без боя и безвозмездно. Каждый из двух голубей с одинаковой вероятностью выигрыша $1/2$, поэтому стоимость ожидаемого выигрыша при встрече двух голубей составляет $V/2$. Сумму всевозможных выигрышей можно резюмировать в таблице 23.1.

Таблица 23.1: Ожидаемый выигрыш в игре «Ястреб - голубь».

	ястреб	голубь
выигрыш ястреба против...	$\frac{1}{2}(V - C)$	V
выигрыш голубя против...	0	$V/2$

В более общем плане, можно представить себе бои между людьми, которые могут принять одну из двух стратегий, называя их 1 и 2, с матрицей ожидаемых выплат $(G_{i,j})_{1 \leqslant i,j \leqslant 2}$. В приведенном выше примере ястребы следуют за стратегией 1, голуби - за стратегией 2, $G_{1,1} = \frac{1}{2}(V - C)$, $G_{1,2} = V$, $G_{2,1} = 0$ и $G_{2,2} = V/2$. В вышеупомянутой статье 1973 года Мэйнард Смит и Прайс фактически уже использовали компьютерное моделирование для проверки более чем двух возможных стратегий (они назывались «ястребом», «мышью», «булли», «реталятором» и «пробером-реталятором»).

Представьте себе сейчас большую популяцию животных одного вида с пропорцией x_n ястребов и пропорцией $1 - x_n$ голубей в поколении n. Ястребы в поколении n имеют среднее количество потомства, равное

$$R_1(n) = \mathscr{R}_0 + x_n\, G_{1,1} + (1 - x_n)\, G_{1,2}\,. \tag{23.1}$$

Аналогичным образом, у голубей среднее количество потомства равно

$$R_2(n) = \mathscr{R}_0 + x_n\, G_{2,1} + (1 - x_n)\, G_{2,2}\,. \tag{23.2}$$

Таким образом, средняя численность потомства во всем населении составляет

$$R(n) = x_n\, R_1(n) + (1 - x_n)\, R_2(n)\,.$$

Не принимая во внимание возможные тонкости, связанные с половым размножением, мы видим, что доля ястребов в следующем поколении составляет

$$x_{n+1} = x_n\, R_1(n)/R(n)\,. \tag{23.3}$$

Следовательно, $x_{n+1} > x_n$ если $R_1(n) > R(n)$ и $x_{n+1} < x_n$ если $R_1(n) < R(n)$. Существует три возможных устойчивых состояния: $x = 0$, $x = 1$ и

$$x^* = \frac{G_{1,2} - G_{2,2}}{G_{2,1} - G_{1,1} + G_{1,2} - G_{2,2}}$$

при условии $0 < x^* < 1$. В игре ястреб - голубь, $x^* = V/C < 1$ при условии $V < C$.

> Действительно, $x = 0$ - это очевидное устойчивое состояние (23.3). Если $x \neq 0$ - еще одно устойчивое состояние, то $R_1 = R = x\, R_1 + (1 - x) R_2$. Так что либо $x = 1$, либо $R_1 = R_2$. Последняя возможность эквивалентна
>
> $$x\, G_{1,1} + (1 - x)\, G_{1,2} = x\, G_{2,1} + (1 - x)\, G_{2,2},$$
>
> что дает устойчивое состояние x^*.

Стабильное состояние $x = 1$ соответствует населению с 100% лиц, придерживающихся стратегии 1. Это устойчивое состояние является стабильным, если в него не могут вторгнуться несколько человек, следующих стратегии 2. Из (23.3) видно, что это условие эквивалентно тому, что $R_1(n) > R(n)$ для всех x_n достаточно близких к 1. Так как $R(n) = x_n\, R_1(n) + (1 - x_n)\, R_2(n)$, то условие становится $R_1(n) > R_2(n)$ для всех x_n достаточно близких к 1. Глядя на выражения (23.1)-(23.2) R_1 и R_2, мы приходим к выводу, что $x = 1$ является стабильным, если и только если выполнено одно из двух следующих условий:

- $G_{1,1} > G_{2,1}$;
- $G_{1,1} = G_{2,1}$ и $G_{1,2} > G_{2,2}$.

Если это так, то, как говорят, стратегия 1 - это эволюционно стабильная стратегия. В игре hawk–dove условие $G_{1,2} > G_{2,2}$ всегда истинно. Таким образом, стратегия ястреба является эволюционно стабильной, если и только если $G_{1,1} \geqslant G_{2,1}$, т.е. $V \geqslant C$.

Стабильное состояние $x = 0$ соответствует населению, все физические лица которого следуют стратегии 2. Эта ситуация симметрична с предыдущей, если мы обменяемся индексами 1 и 2. В игре «Ястреб - голубь» мы имеем $G_{1,2} = V > G_{2,2} = V/2$, так что устойчивое состояние $x = 0$ всегда нестабильно. Введение небольшого количества ястребов в популяцию голубей приведет к прогрессирующему вторжению ястребов.

Точно так же можно показать, что третье устойчивое состояние x^*, при условии $0 < x^* < 1$, всегда стабильно. В игре «Ястреб - голубь» $x^* = V/C$ соответствует смешанной популяции как с ястребами, так и с голубями.

В заключение, есть два случая в игре ястреб - голубь. Если $V \geqslant C$, т.е. если стоимость ресурса V больше возможной стоимости ресурса C, то популяция имеет тенденцию к устойчивому состоянию с ястребами, но без голубей, независимо от исходного условия $x(0)$ с $0 < x(0) < 1$. Тогда стратегия «ястреба» является эволюционно стабильной стратегией. Если же, напротив, $V < C$, то население стремится к смешанному устойчивому состоянию с пропорцией x^* ястребов и пропорцией $1 - x^*$ голубей. Таким образом, модель дает объяснение того, почему люди с менее агрессивным поведением могут выжить при $V < C$. Кроме того, формула $x^* = V/C$ показывает, что чем выше стоимость C для проигравших, тем меньше доля x^* ястребов в популяции. Поэтому виды с самым опасным «оружием» редко используют его для внутривидовых боев: они предпочитают безобидные ритуальные бои, в которых соперничающие животные пытаются произвести впечатление друг на друга, но избегают реальных боев, которые могут нанести травмы.

В исходной статье Мэйнарда Смита и Прайса 1973 года обсуждалась концепция эволюционно стабильной стратегии и использовалось в основном компьютерное моделирование конфликтов животных, фиксировавшее отдачу от различных стратегий. Подход, использующий динамические уравнения типа (23.3), был разработан несколько позже, в частности Тейлором и Джонкером. С тех пор многие авторы применили идеи теории игр к вопросам эволюционной биологии или, наоборот, применили динамические эволюционные подходы к более классическим задачам теории игр. Помимо вопросов, касающихся конфликтов животных, можно привести, например, проблемы родительского инвестирования или соотношения полов (соотношение между числом самцов и самок при рождении), последнее уже изучалось Карлом Дюзингом в 1884 году и Рональдом Фишером в его книге 1930 г. *Генетическая теория естественного отбора*. Некоторые другие модели фокусируются на динамических аспектах «дилеммы заключенного» или игры в «камень-ножницы-бумага». Также было осознано, что концепция эволюционно устойчивой стратегии тесно связана с концепцией равновесия Нэша в теории игр.

Прайс, который был убежденным атеистом, пережил мистический опыт в 1970 году и обратился в христианскую веру. Он отказался от своих исследований в 1974 году, потому что считал, что «теоретическая математическая генетика, которой он занимался, не имеет большого отношения к человеческим проблемам». Он отдал все свои вещи бездомным и через несколько месяцев покончил жизнь самоубийством.

Мэйнард Смит, напротив, продолжил эту линию мысли и был избран в Королевское общество в 1977 году. Он опубликовал много книг: *Модели в экологии* (1974), *The Evolution of Sex* (1978), *Эволюция и теория игр* (1982), *Проблемы биологии* (1986), *Дарвин правильно понял?* (1988) и *Эволюционная генетика* (1989). Он также опубликовал в сотрудничестве с E. Szathmáry *Основные переходы в эволюции* (1995) и *Происхождение жизни: От рождения жизни к происхождению языка* (1999). Он ушел на пенсию в 1985 году. В 1999 году он получил премию Крафорда в области бионаук от Шведской королевской академии наук за «фундаментальный вклад в концептуальное развитие эволюционной биологии». В 2003 году он опубликовал в сотрудничестве с Д. Харпером *Сигналы животных.* Умер Мэйнард Смит в Сассексе в 2004 году.

Дополнительное чтение

1. Charlesworth, B., Harvey, P.: John Maynard Smith, 6 January 1920–19 April 2004. *Biog. Mem. Fellows R. Soc.* 51, 253–265 (2005)
2. Edwards, A.W.F.: Carl Düsing (1884) on the regulation of the sex-ratio. *Theor. Pop. Biol.* 58, 255–257 (2000)
3. Frank, S.A.: George Price's contributions to evolutionary genetics. *J. Theor. Biol.* 175, 373–388 (1995)
4. Maynard Smith, J., Price, G.R.: The logic of animal conflict. *Nature* 246, 15–18 (1973)
5. Maynard Smith, J.: *Evolution and the Theory of Games.* Cambridge University Press (1982)
6. Schwartz, J.: Death of an altruist: Was the man who found the selfless gene too good for this world? *Lingua Franca* 10, 51–61 (2000) bio.kuleuven.be/ento/pdfs/schwartz2000.pdf
7. Sigmund, K.: John Maynard Smith and evolutionary game theory. *Theor. Pop. Biol.* 68, 7–10 (2005)
8. Taylor, P.D., Jonker, L.B.: Evolutionary stable strategies and game dynamics. *Math. Biosci.* 40, 145–156 (1978)
9. Von Neumann, J., Morgenstern, O.: *Theory of Games and Economic Behavior.* Princeton University Press (1944) archive.org

Глава 24

Хаотические популяции (1974)

В 1974 году Роберт Мэй, австралийский физик, ставший экологом, изучал дискретно-временное логистическое уравнение как модель динамики популяции. Он заметил, что в этой модели происходят неожиданные бифуркации и что асимптотическое поведение может быть даже хаотичным. Поэтому долгосрочные прогнозы могут быть невозможны даже при простой детерминистической модели. Статья Мэя была одной из тех, которые положили начало «теории хаоса».

Роберт МакКреди Мэй (May) родился в 1936 году в Австралии. После изучения теоретической физики и получения степени доктора философии в университете Сиднея в 1959 году, он провел два года на кафедре прикладной математики в Гарвардском университете. Вернувшись в Австралию, он стал профессором теоретической физики. В 1971 году, посетив Институт высших исследований в Принстоне, он изменил тематику своих исследований и стал уделять особое внимание динамике популяций животных. В 1973 году он стал профессором зоологии в Принстоне. В том же году он опубликовал книгу под названием *Стабильность и сложность в модели экосистем.*

Рис. 24.1:
Роберт М. Мэй (1936-2020)

В 1974 году в журнале *Science* Мэй опубликовал статью под названием *Биологические популяции с непересекающимися поколениями: стабильные точки, стабильные циклы и хаос*, в которой он показал, что очень простые математические модели в динамике популяции могут вести себя хаотично.

Чтобы понять происхождение этой проблемы, нужно вернуться примерно на десять лет назад. В 1963 году американский метеоролог Эдвард Лоренц, работающий в Массачусетском технологическом институте (M.I.T.), заметил при проведении численного моделирования на своем компьютере, что упрощенная модель атмосферы, содержащая всего три дифференциальных уравнения, может вести себя очень удивительным образом: малейшее изменение исходных условий может полностью изменить конечный результат моделирования, а значит и метеорологические прогнозы. Математик Анри Пуанкаре, изучая движение планет в Солнечной системе, фактически уже предсказывал эту возможность в начале двадцатого века, задолго до компьютерной эры. Но в начале 1970-х годов лишь немногие исследователи начали более внимательно изучать это странное свойство. В Университете Мэриленда Джеймс Йорк размышлял о работе Лоренца и ввел в этот контекст термин «хаос». Статья, которую он написал вместе со своим студентом Тянь-Иен Ли, озаглавленная *Третий период подразумевает хаос*[1], появилась в 1975 году.

Со своей стороны, Мэй сосредоточился на модели вида

$$p_{n+1} = p_n + a\,p_n(1 - p_n/K), \tag{24.1}$$

где a и K - положительные параметры, а p_n - размер популяции животного в год n. Когда p_n мал по сравнению с несущей способностью K, динамика близка к геометрическому росту $p_{n+1} \approx (1+a)\,p_n$. Полное уравнение является своего рода дискретно-временным аналогом логистического уравнения, введенного Ферхюльстом (см. главу 6). Но в отличие от последнего, Май показал, что дискретно-временное уравнение может иметь гораздо более удивительное поведение, которое легко наблюдать с помощью простого карманного калькулятора, делающего сложения и умножения (рисунок 24.2). Мэйнард Смит уже рассматривал уравнение (24.1) в своей книге 1968 года *Математические идеи в биологии.* Но он попробовал всего лишь несколько числовых значений для a и не понял, что есть особенности поведения уравнения.

[1] Примечательно, что более общий результат был доказан О.М. Шарковским в 1964 году, но его статья, опубликованная в украинском математическом журнале, была малоизвестна.

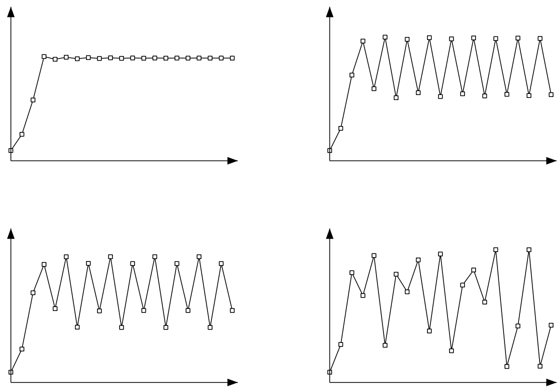


Рис. 24.2: Во всех цифрах: n на горизонтальной оси, p_n на вертикальной оси и $p_0 = K/10$. Линии получаются путем сшивания точек с координатами (n, p_n). Слева вверху: $0 < a < 2$ (устойчивое состояние). Сверху справа: $2 < a \leqslant 2{,}449$ (период 2 цикла). Слева внизу: $2{,}450 \leqslant a \leqslant 2{,}544$ (период 4 цикла). Справа внизу: $2{,}570 \leqslant a \leqslant 3$ (возможно хаос).

На рисунке 24.2, аналогичном статье Мэя 1974 года, видно, что население p_n переходит в устойчивое состояние, когда $0 < a < 2$. Когда $2 < a \leqslant 2{,}449$ (верхняя граница 2,449 - аппроксимация), население p_n имеет тенденцию к циклу периода 2. Когда $2{,}450 \leqslant a \leqslant 2{,}544$, население p_n имеет тенденцию к циклу периода 4. Когда $2{,}545 \leqslant a \leqslant 2{,}564$, p_n имеет тенденцию к циклу периода 8, и т.д. Интервалы параметра a, для которого p_n имеет тенденцию к циклу периода 2^n, уменьшаются с увеличением n и никогда не превышают 2,570. Когда $a \geqslant 2{,}570$, p_n может вести себя «хаотически».

В 1976 г. Мэй опубликовал обзор проблемы в *Nature* под названием *Простые математические модели с очень сложной динамикой.* Там он собрал не только собственные результаты, но и результаты других исследователей. Во-первых, установив $x_n = \frac{a\,p_n}{K\,(1+a)}$ и $r = 1 + a$ (так что $r > 1$), мы видим, что уравнение (24.1) можно переписать в более простом виде.

$$x_{n+1} = r\,x_n\,(1 - x_n)\;. \tag{24.2}$$

Для того, чтобы это уравнение имело смысл в динамике населения, x_n должен быть неотрицательным для всех n. Поэтому мы предполагаем, что начальное условие x_0 удовлетворяет $0 \leqslant x_0 \leqslant 1$, а $r \leqslant 4$. Последнее условие гарантирует, что правая часть (24.2) останется между 0 и 1. Примечательно, что хаотический случай $r = 4$ уже использовался в качестве генератора случайных чисел Станиславом Уламом и Джоном фон Нейманом еще в 1947 году. Если мы введем функцию

$$f(x) = r\,x(1-x),$$

то уравнение (24.2) можно переписать как $x_{n+1} = f(x_n)$, а устойчивые состояния - это решения $x = f(x)$. Графически это пересечения кривых $y = f(x)$ и $y = x$ (рисунок 24.3). Обратите внимание, что $x = 0$ - это всегда устойчивое состояние. Поскольку $r > 1$, существует еще одно устойчивое состояние $x^* > 0$, такое, что $x^* = r\,x^*\,(1 - x^*)$, т.е. $x^* = 1 - 1/r$.

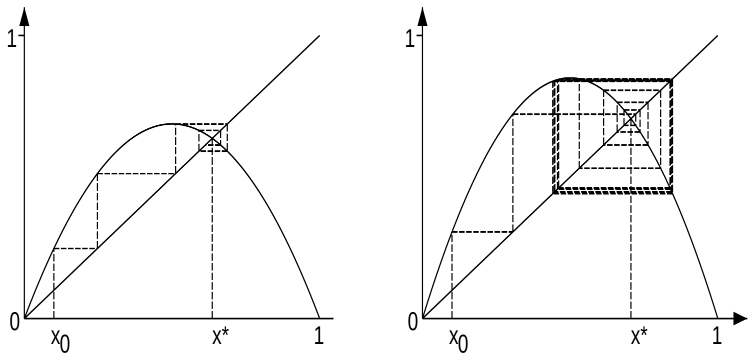


Рис. 24.3: Функция $y = f(x)$, прямая $y = x$, устойчивое состояние x^* и последовательность, определяемая $x_{n+1} = f(x_n)$. (a) : $r = 2{,}75$, последовательность стремится к x^*. (b) : $r = 3{,}4$, устойчивое состояние x^* нестабильно и последовательность имеет тенденцию к циклу периода 2.

Для всех $r > 1$, устойчивое состояние $x = 0$ нестабильно. Действительно, когда x_n близок к 0, у нас есть $x_{n+1} \approx r\,x_n$. Таким образом, x_n имеет тенденцию отклоняться от 0. Что касается устойчивого состояния x^*, то оно локально стабильно только для $1 < r < 3$.

Действительно, обозначим $y_n = x_n - x^*$. Тогда легко увидеть, что уравнение (24.2) эквивалентно $y_{n+1} = (2-r-r\,y_n)\,y_n$. Если x_n близко к x^*, то y_n близко к 0 и $y_{n+1} \approx (2-r)\,y_n$. Но для уравнения вида $y_{n+1} = k\,y_n$ находим, что $y_n = k^n\,y_0$ и $y_n \to 0$ при $n \to \infty$ если и только если $-1 < k < 1$. Здесь устойчивое состояние x^* локально стабильно тогда и только тогда, когда $-1 < 2 - r < 1$, т.е. $1 < r < 3$.

Если $1 < r < 3$, то можно показать, что при всех начальных условиях $0 < x_0 < 1$, последовательность x_n действительно стремится к x^* (рисунок 24.3a). Но что произойдет, когда $3 < r \leqslant 4$? Чтобы ответить на этот вопрос, обратим внимание на то, что $x_{n+2} = f(x_{n+1}) = f(f(x_n))$. Введем функцию

$$f_2(x) = f(f(x)) = r^2\,x\,(1-x)\big(1 - r\,x\,(1-x)\big)$$

и рассмотрим решения уравнения $x = f_2(x)$, которые называются фиксированными точками функции $f_2(x)$. Графически - это пересечения кривых $y = f_2(x)$ и $y = x$ (рисунок 24.4).

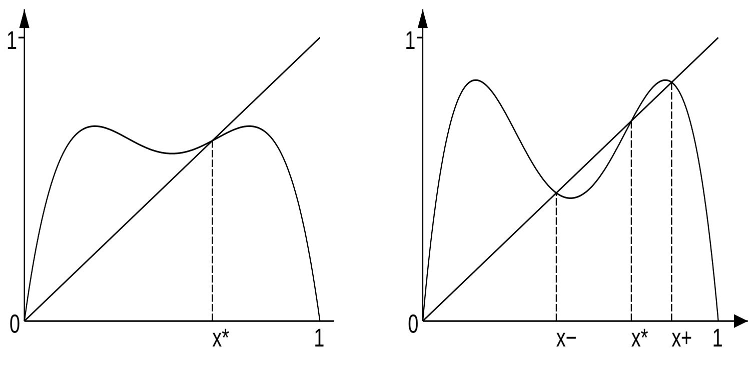


Рис. 24.4: Кривые $y = f_2(x) = f(f(x))$ и $y = x$ и устойчивое состояние x^*. (a) : $r = 2{,}75$. (b) : $r = 3{,}4$ и два других решения x_- и x_+ уравнения $x = f_2(x)$.

Если $x = f(x)$, то $x = f(f(x)) = f_2(x)$. Таким образом, $x = 0$ и $x = x^*$ также являются фиксированными точками функции $f_2(x)$.

Но когда $r > 3$, функция $f_2(x)$ имеет еще две фиксированные точки, x_- и x_+, так что $f(x_-) = x_+$ и $f(x_+) = x_-$.

Действительно, мы замечаем, что $f_2'(x) = f'(f(x))\, f'(x)$ так что $f_2'(x^*) = [f'(x^*)]^2$. Но $f'(x) = r(1-2x)$ и $x^* = 1 - 1/r$. Так что $f'(x^*) = 2 - r$ и $f_2'(x^*) = (2-r)^2$. Следовательно, наклон функции $f_2(x)$ при $x = x^*$ таков, что $f_2'(x^*) > 1$ если $r > 3$. Но поскольку $f_2(0) = 0$, $f_2'(0) = r^2 > 1$ и $f_2(1) = 0$, то на рисунке 24.4b видно, что обязательно есть еще два решения x_- и x_+ уравнения $x = f_2(x)$, где $0 < x_- < x^*$ и $x^* < x_+ < 1$. Другой способ прийти к такому же выводу заключается в решении уравнения $x = f_2(x)$, которое является полиномиальным уравнением степени 4 с двумя известными корнями: $x = 0$ и $x = x^*$. Два других решения x_- и x_+ являются корнями многочлена

$$x^2 - \frac{1+r}{r}\,x + \frac{1+r}{r^2} = 0\,. \qquad (24.3)$$

Они являются действительными, если дискриминант положителен, т.е. если $r > 3$. Поскольку $f_2(f(x_-)) = f(f(f(x_-))) = f(f_2(x_-)) = f(x_-)$, то точка $f(x_-)$ также является фиксированной точкой $f_2(x)$. Но $f(x_-) \neq x_-$, потому что x_- не является фиксированной точкой $f(x)$. А $f(x_-) \neq x^*$, иначе мы бы имели $x_- = f(f(x_-)) = f(x^*) = x^*$. Начиная с $f(x_-) \neq 0$, мы делаем вывод, что $f(x_-) = x_+$. Аналогично $f(x_+) = x_-$.

Следовательно, для $r > 3$ мы видим, что если, например, $x_0 = x_-$, то $x_1 = x_+$, $x_2 = x_-$, $x_3 = x_+$, и т.д. Можно также показать, что почти для каждого начального условия $0 < x_0 < 1$, последовательность x_n стремится при $n \to +\infty$ к циклу периода 2 x_-, x_+, x_-, x_+ и др. (рисунок 24.3b и 24.4b). Данный цикл остается стабильным до тех пор, пока r находится ниже критического значения $r_1 = 1 + \sqrt{6} \approx 3{,}449$, где $f_2'(x_-) = -1$.

Действительно, мы видим, используя (24.3), что

$$\begin{aligned} f_2'(x_-) &= f'(f(x_-))\, f'(x_-) = f'(x_+)\, f'(x_-) \\ &= r^2\,(1 - 2\,x_+)(1 - 2\,x_-) = r^2\,(1 - 2(x_+ + x_-) + 4\,x_+ x_-) \\ &= r^2\Big(1 - 2\,\frac{1+r}{r} + 4\,\frac{1+r}{r^2}\Big) = -r^2 + 2r + 4\,. \end{aligned}$$

Так что $f_2'(x_-) = -1$ если $-r^2 + 2r + 5 = 0$ и, в частности, если $r = 1 + \sqrt{6}$.

Для $r_1 < r < r_2$ цикл периода размера 4 становится стабильным: появляются четыре новые фиксированные точки функции

$$f_4(x) = f_2(f_2(x)) = f(f(f(f(x))))$$

(рисунок 24.5). Для $r_2 < r < r_3$ это цикл размера 8 и так далее. Цифры r_n имеют тенденцию к ограничению $r_\infty \approx 3{,}570$, когда $n \to +\infty$. Когда $r_\infty < r \leqslant 4$, система может быть даже хаотичной! На рисунке 24.6 показана диаграмма[2] бифуркации, дающая представление о сложности динамики.

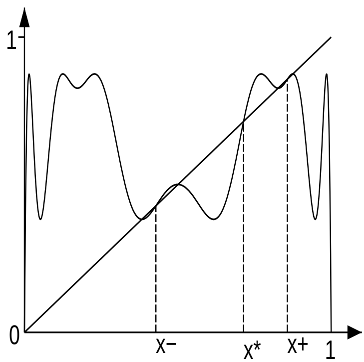


Рис. 24.5: Кривая $y = f_4(x)$ при $r = 3{,}5$ и биссектриса $y = x$. Кроме x^*, x_+ и x_- есть еще четыре фиксированные точки, которые не так просто различить.

Р. М. Мэй в заключение подчеркнул, что даже очень простые динамические системы могут иметь очень сложное поведение. Это не относится к уравнению $x_{n+1} = r\,x_n\,(1 - x_n)$. Такой же «каскад удвоения периодов», приводящий к хаосу, появляется и для других уравнений с функцией $f(x)$, имеющей форму «горба». Так, например, обстоит дело с другим уравнением, используемым в популяционной биологии: $x_{n+1} = x_n \exp(r(1 - x_n))$.

[2]Эта диаграмма была получена путем построения графиков для каждого заданного значения r точек с координатами (r, x_{200}), (r, x_{201}),...,(r, x_{220}), где $x_{n+1} = f(x_n)$ и $x_0 = 0{,}1$. Если x_n стремится к стабильному состоянию, то на диаграмме мы видим только одну точку. Если x_n имеет тенденцию к циклу периода 2, то мы видим две точки и так далее.

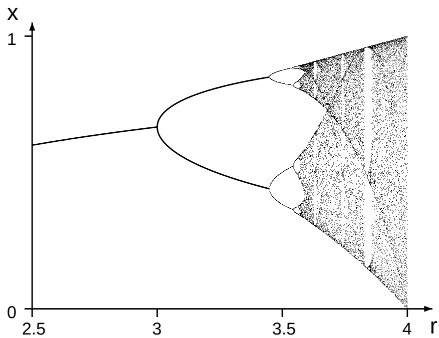


Рис. 24.6: Бифуркационная диаграмма уравнения (24.2).

Данное исследование предполагает, что не следует удивляться, если многие наборы данных о динамике населения трудно поддаются анализу. Модель также показывает, что различие между детерминистическими и стохастическими моделями не так ясно, как считалось ранее: даже при простой детерминистической модели невозможно делать долгосрочные прогнозы, если параметры находятся в хаотическом режиме.

В 1979 году Мэй был избран в Королевское общество. С 1988 по 1995 год он был профессором Оксфордского университета и Императорского колледжа в Лондоне, с 1995 по 2000 год был главным научным советником британского правительства. В 1996 году Мэй получил премию Crafoord «за новаторские экологические исследования, касающиеся теоретического анализа динамики популяций, сообществ и экосистем». Из экологии он обратился к эпидемиологии и иммунологии, опубликовав две книги: *Инфекционные заболевания человека* (1991, с Рой Андерсоном) и *Вирусная динамика, Математические основы иммунологии и вирусологии* (2000, с Мартином Новаком). В последней книге анализируется взаимодействие между клетками иммунной системы и ВИЧ (вирусом, вызывающим СПИД) как некой системы хищник-жертва (см. главу 13). С 2000 по 2005 год Мэй был президентом Королевского общества. Он был посвящен в рыцари в 1996 году и принят в Палату Лордов в 2001 году. Умер Мэй в 2020 году.

Дополнительное чтение

1. Gleick, J.: *Chaos, Making a New Science.* Viking Penguin (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

Глава 25

Политика «одного ребенка» в Китае (1980)

В 1980 году Сун Цзянь и его коллеги, которые были специалистами по теории управления, применявшейся в авиационной технике, подсчитали, что, если уровень рождаемости в Китае останется на текущем уровне, то в XXI веке население страны достигнет более двух миллиардов человек. Их результаты, основанные на возрастно-структурированной математической модели, способствовали принятию правительством решения о переходе к политике «одна семья - один ребенок».

Сун Цзянь[1] (Song Jian) родился в 1931 году в городе Жунчэн в китайской провинции Шаньдун. В 1950-х годах учился в Советском Союзе в Московском государственном техническом университете имени Н.Э. Баумана и на механико-математическом факультете МГУ. Затем он вернулся в Китай и стал руководителем Бюро кибернетических исследований в Институте математики Китайской академии наук. Он был специалистом по применению теории управления к наведению ракет. Он также работал в Седьмом машиностроительном министерстве, которое позднее было переименовано в Министерство аэрокосмической промышленности. В 1978 году он начал уделять внимание связям между теорией управления и демографией.

Рис. 25.1: Сун Цзянь

[1]Сун - фамилия, она всегда пишется первой на китайском языке.

Чтобы понять контекст работы Сун Цзяня над динамикой народонаселения, сначала нужно дать представление о том, что такое теория контроля. Это - изучение динамических систем, поведение которых зависит от некоторых параметров, которые могут быть модифицированы с течением времени с целью оптимизации заданного критерия. Эта теория была особенно развита в связи с космическими программами в США и СССР. Действительно, инженерам приходилось контролировать траекторию движения космических челноков, чтобы вывести спутники на орбиту вокруг Земли. Но применение не ограничивалось физическими или инженерными проблемами. Политику контроля рождаемости также можно было рассматривать как некую оптимальную проблему контроля в математическом смысле.

Следует также упомянуть эссе под названием *Пределы роста - Доклад по проекту Римского клуба «Проблемы человечества»*, опубликованный в 1972 году и написанный группой из Массачусетского технологического института (M.I.T.). Это исследование было основано на математической модели мирового экономического роста, учитывающей природные ресурсы, численность населения и загрязнение окружающей среды. В докладе высказывается мысль о том, что мировая экономика движется к катастрофе, вызванной истощением невозобновляемых ресурсов, нехваткой продовольствия для населения или чрезмерным загрязнением. Одним из предлагаемых решений является добровольное ограничение рождаемости. Подводя итог, можно сказать, что это была своего рода современная версия тезисов Мальтуса. Доклад получил большой резонанс на Западе в 1970-е годы.

С момента основания Народной Республики в 1949 году рождаемость в Китае была очень высокой, за исключением катастрофического «Большого скачка». В середине 1970-х годов Китай медленно восстанавливался после Культурной революции. Планирование семьи настоятельно призывало женщин откладывать роды, увеличивать время между двумя последовательными родами и рожать меньше детей. Дэн Сяопин, ставший новым лидером после смерти Мао Цзэдуна в 1976 году, в 1978 году начал политику «Четырех модернизаций»: сельское хозяйство, промышленность, наука и техника, национальная оборона. Размер и рост населения Китая тогда воспринимались как важные препятствия на пути этих модернизаций. Ученым, которые до этого времени работали над военными приложениями, было предложено найти решения этой трудной проблемы.

На этом фоне, Сун Цзянь в 1978 году отправился в Хельсинки на конгресс Международной федерации автоматического управления. Там он заметил, что некоторые исследователи в Европе пытались применить теорию контроля к проблемам народонаселения с мыслью о том, что строгий контроль рождаемости в конечном итоге может предотвратить катастрофы, представленные в докладе *Пределы роста*. Вернувшись в Китай, он создал небольшую команду, в которую вошли его коллега Юй Цзиньцзюань и компьютерный эксперт Ли Гуанюань, для применения такого рода математического моделирования к данным, касающимся населения Китая. В то время научная связь между Китаем и остальным миром была скудной. Команда заново разработала уравнения, описывающие эволюцию возрастной структуры населения, точно так же, как это делали Лотка и МакКендрик (см. главы 10 и 16). Используя модель непрерывного времени, введем обозначения:

- $P(x,t)$ - население в возрасте x во время t;
- $m(x)$ - смертность в возрасте x;
- $P_0(x)$ - возрастная структура населения на момент $t = 0$;
- $b(t)$ СКР (суммарный коэффициент рождаемости), т.е. среднее число детей, которое женщина имела бы в течение своей жизни, если бы ее возрастная рождаемость оставалась на том же уровне, что и во время t;
- f доля девочек среди новорожденных;
- $h(x)$ плотность распределения возраста матери при рождении ребенка ($\int_0^{+\infty} h(x)\, dx = 1$).

С помощью этих обозначений и гипотез эволюция возрастной структуры может быть смоделирована с помощью дифференциального уравнения в частных производных

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

с начальным условием $P(x,0) = P_0(x)$ и граничным условием

$$P(0,t) = b(t)\, f \int_0^{+\infty} h(x)\, P(x,t)\, dx\,,$$

где $b(t)$ - параметр, которым нужно управлять. Если СКР постоянен и превышает критический порог

$$b^* = 1/\left[f\int_0^{+\infty} h(x)\, e^{-\int_0^x m(y)\,dy}\,dx\right],$$

то население увеличивается в геометрической прогрессии. Этот критерий аналогичен критерию, полученному Лоткой в формуле (10.2). Команда Сун Цзянь также рассмотрела дискретно-временную версию модели, аналогичную модели Лесли (см. главу 21). Обозначим $P_{k,n}$ население в возрасте k в год n. Аналогично введем m_k, b_n и h_k. Тогда

$$P_{k+1,n+1} = (1-m_k)\,P_{k,n}\ ,\quad P_{0,n+1} = b_n\, f \sum_{k\geqslant 0} h_k\, P_{k,n}\,.$$

Зная из выборочных обследований смертность m_k (рисунок 25.2), долю девочек среди новорожденных $f \approx 0{,}487$, возрастное распределение матерей h_k (рисунок 25.3), начальное условие $P_{k,0}$ - возрастную структуру населения в 1978 г. (рисунок 25.4) и варьируя общую рождаемость b (предполагаемую постоянной во всех расчетах), команда Сун Цзянь могла делать демографические прогнозы для своей страны со временным горизонтом в сто лет, с 1980 по 2080 годы (рисунок 25.5). С учетом требуемых тысяч операций сложения и умножения (год n колеблется от 0 до 100 лет, возраст k от 0 до 90 лет), необходим был компьютер. В то время в Китае доступ к такому оборудованию имели лишь немногие, за исключением тех, кто работал на военных. Одним из них был Сун Цзянь, ведущий эксперт в области наведения ракет.

Согласно прогнозам, даже если бы Китай сохранил рождаемость на уровне 1978 года, 2,3 на одну женщину, что чуть выше критического порога, оценивавшегося в $b^* = 2{,}19$, население выросло бы с 980 миллионов человек в 1980 году до 2,12 миллиарда человек к 2080 году. Но Китай уже использовал почти все земли, которые могли служить для ведения сельского хозяйства. Он даже имел тенденцию терять часть этих земель из-за опустынивания и урбанизации. Как прокормить такое население, если прогресс в повышении урожайности фермерских хозяйств недостаточен? Этот же вопрос Мальтус рассматривал два столетия назад. При рождаемости 1975 года в 3,0, население могло даже достичь 4,26 миллиарда человек в 2080 году. При $b = 2$, численность населения могда

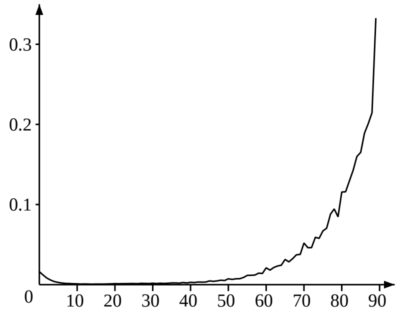


Рис. 25.2: Смертность (за год) в зависимости от возраста в 1978 году.

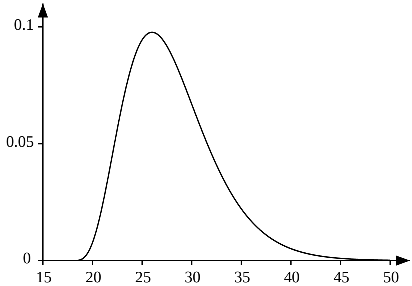


Рис. 25.3: Сглаженная форма функции рождаемости (за год) в зависимости от возраста в 1978 году.

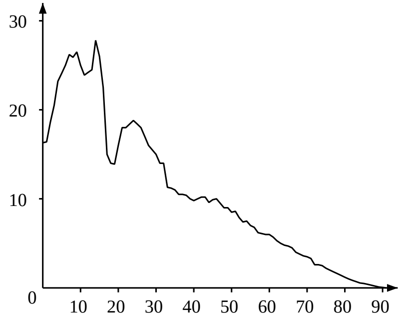
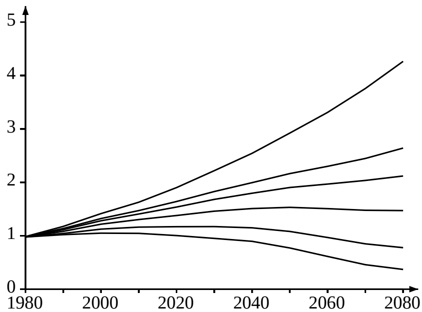


Рис. 25.4: Возрастная пирамида населения Китая в 1978 году. Горизонтальная ось: возраст. Вертикальная ось: население (в миллионах).

достигнуть максимума в 1,53 миллиарда человек примерно в 2050 году, прежде чем началось бы ее незначительное сокращение. При $b = 1{,}5$, максимум 1,17 миллиарда человек был бы достигнут примерно в 2030 году. При $b = 1{,}0$, максимум составил бы всего 1,05 миллиарда человек и был бы достигнут примерно в 2000 году. При таком предположении население вернулось бы на уровень 1978 года только к 2025 году.

Рис. 25.5: Демографические прогнозы (в миллиардах) для Китая, основанные на различных гипотезах о среднем числе детей на одну женщину. Снизу вверх: 1,0, 1,5, 2,0, 2,3, 2,5, 3,0.



Самым удивительным в этой работе было ее практическое значение, по сути, непревзойденное в истории математической динамики населения. Действительно, Ли Гуанюань показал результаты моделирования команды в декабре 1979 года во время симпозиума по народонаселению в Чэнду, провинция Сычуань.[2]. В январе 1980 года Сун Цзянь, Юй Цзиньюань и Ли Гуанюань опубликовали эти результаты в китайском экономическом журнале, выступая, кстати, за политику в отношении одного ребенка. Они также прислали свою статью – *Отчет о количественных исследованиях по вопросу развития населения Китая* – ведущему китайскому ученому Цянь Сюэню, который направил ее с рекомендацией руководителю управления планирования рождаемости. Результаты работы команды Сун Цзянь произвели глубокое впечатление на большинство политических лидеров. Они уже были убеждены в необходимости усиления контроля над рождаемостью, несмотря на то, что написал Маркс (см. главу 5), но все еще колебались в

[2]Здесь и далее мы подводим итоги подробного отчета Сьюзен Гринхалг [1,2].

отношении уровня контроля. В феврале 1980 года Госсовет и Центральный комитет партии поставили перед китайским населением цель в 1,2 миллиарда человек на горизонте 2000 года. В марте 1980 года результаты работы команды Сун Цзянь были опубликованы в *Жэньминь жибао*. В апреле комиссия, состоящая из политических лидеров и специалистов по народонаселению, изучила экологические и экономические последствия роста населения и пришла к выводу, что для достижения цели, поставленной Дэн Сяопином в отношении дохода на душу населения в 2000 году, необходима политика в отношении одного ребенка. Политика стала официальной в сентябре того же года, а на первой странице *Жэньминь жибао* было опубликовано открытое письмо, разъясняющее ее населению.

К 1983 году будет еще много несанкционированных родов. Было решено, что один из супругов каждой пары, имеющей уже двух детей, будет стерилизован и что каждая запрещенная беременность будет прервана. Однако, начиная с 1984 года, сельским парам, имеющим только дочь, было разрешено рожать второго ребенка. Политика в отношении одного ребенка оставалась в силе до 2015 года. В последние годы были внесены определенные коррективы: если и мужчина, и женщина были одинокими детьми, то у них может быть двое детей. Репрессивные меры в отношении пар, имеющих более одного ребенка, являлись жестокими: государственные служащие могли потерять работу, дорогостоящие штрафы должны были быть уплачены за получение административных документов на второго ребенка и т.д. Подводя итог, трудно найти в истории математического моделирования еще один пример, имевший столь сильное социальное воздействие. Конечно, работа Сун Цзяня и его соавторов была лишь одним из элементов, приведших к выбору политики в отношении одного ребенка. Но, похоже, она сыграла важную роль.

Как и в предыдущих главах, роль математического моделирования может быть предметом озабоченности. Начиная с реальной ситуации, строится модель. Она может быть проанализирована математически или смоделирована с помощью компьютера. Затем можно понять, как модель ведет себя при изменении некоторых параметров. Однако математика не говорит, является ли модель верной картиной реальной жизни. Некоторые очень важные аспекты, возможно, были проигнорированы. Некоторые модели также содержат целевую функцию, например, удержание населения Китая в пределах 1,2 миллиарда человек. Математика не говорит, была ли эта цель уместной[3].

В 1980 году Сун Цзянь также был соавтором нового издания книги *Инженерная кибернетика* Цяня Сюсэня, «отца» китайской космической программы. Затем он занимал различные политические должности высокого уровня: заместитель министра и главный инженер-ученый министерства аэрокосмической промышленности (1981-1984 гг.), член Центрального комитета Коммунистической партии Китая (1982-2002 гг.), председатель Государственной комиссии по науке и технике (1985-1998 гг.), Государственный советник (1986-1998 гг.) и др. Он также опубликовал две другие книги, которые были переведены на английский язык: *Контроль за численностью населения в Китае* (1985 год, в соавторстве с Туан Чи-Сянь и Юй Цзинъюань) и *Управление системой народонаселения* (1988 год, в соавторстве с Юй Цзинъюань). В этих книгах развивается теория оптимального управления применительно к динамике населения. Сун Цзянь был избран в 1991 году в Китайскую Академию Наук, а в 1994 году в Академию Инженеров, президентом которой он являлся с 1998 по 2002 годы.

Дополнительное чтение

1. Greenhalgh, S.: Missile science, population science: The origins of China's one-child policy. *China Q.* 182, 253–276 (2005)
2. Greenhalgh, S.: *Just One Child, Science and Policy in Deng's China.* University of California Press (2008)
3. Meadows, D.H., Meadows, D.L., Randers, J., Behrens, W.W.: *The Limits to Growth, A Report for the Club of Rome's Project on the Predicament of Mankind*, 2nd edn. Universe Books, New York (1974)
4. Song, J.: Selected Works of J. Song. *Science Press*, Beijing (1999)
5. Song, J.: Some developments in mathematical demography and their application to the People's Republic of China. *Theor. Popul. Biol.* 22, 382–391 (1982)
6. Song, J., Yu, J.: *Population System Control.* Springer, Berlin (1988)

[3] Численность населения в 2000 году оценивалась в 1,264 миллиарда человек. В период с 1980 по 2000 год доход на душу населения вырос примерно с \$200 до \$1000. В то же время соотношение полов стало крайне смещенным в сторону большей доли мальчиков, главным образом из-за абортов по признаку пола.

Глава 26

Некоторые современные проблемы

В этой главе дается краткий обзор некоторых современных проблем в математической динамике народонаселения: старение населения в демографии; возникающие болезни (СПИД, атипичная пневмония, переносчики инфекционных заболеваний...) и политика в области вакцинации в эпидемиологии; политика в области рыболовства в экологии; рассеивание генетически модифицированных организмов в генетике популяции. Упоминаются специализированные учреждения, работающие во Франции над моделированием этих проблем. Также подчеркиваются различные аспекты исследовательской работы.

В данной главе мы рассмотрим некоторые вопросы современных исследований по математическому моделированию динамики населения. Тема достаточно обширна и здесь приводятся лишь несколько примеров исследований, проводимых французскими учеными.

В демографии в последние десятилетия появилась относительно новая проблема: старение населения. Эта проблема актуальна не только во Франции (рис. 26.1), но и во многих других европейских странах, а также в Японии. Она имеет важные экономические и социальные последствия: пенсионные системы, иммиграционная политика и т.д. Во Франции математические модели, пытающиеся проанализировать феномен старения, разрабатываются Национальным институтом демографических исследований (INED) и Национальным институтом статистики и экономических исследований (INSEE). Одна из трудностей демографических прогнозов заключается в том, что уровень рождаемости может значительно варьироваться во времени и не может быть предсказуемым даже на одно десятилетие вперед. Это особенно бросается в глаза, если вспомнить прогнозы, сделанные в 1968 г. в отношении населения Франции в 1985 г.: эти прогнозы не предвещали снижения рождаемости, которое происходило в течение 1970-х годов. Было бы интересно пересмотреть все прогнозы, основанные на математических моделях, которые оказались ошибочными, особенно те, которые нашли отклик в средствах массовой информации. Это уравновесило бы впечатление «прогресса», созданное настоящей книгой, впечатление, которое, возможно, уже показалось читателю подозрительным после прочтения главы, посвященной китайской политике в отношении одного ребенка.[1]. Что касается последней темы, то в настоящее время возникает новая проблема: как смягчить политику, чтобы избежать явления быстрого старения, которое ожидается в ближайшие несколько десятилетий. Математические модели вновь вносят свой вклад в дискуссию.

Рис. 26.1: Возрастная пирамида французского населения на 1 января 2010 года. Источник: www.insee.fr.

В эпидемиологии среди новых проблем, возникших во всем мире за последние два десятилетия, развитие эпидемии СПИДа является особенно поразительным. Некоторые модели пытаются угадать будущее эпидемии в таких недавно инфицированных странах, как Россия, Индия или Китай. Трудно предсказать, замедлится ли эпидемия, как в Западной Европе и Северной Америке, или она достигнет значительной доли населения, как в некоторых странах, расположенных к югу от Сахары. Другие возникающие заболевания, такие как лихорадка Эбола в Африке, лихорадка Западного Нила в Северной Америке, атипичная пневмония (SARS) (острый респираторный синдром), птичий грипп, чикунгунья или грипп H1N1 - все они были тщательно изучены с помощью математических моделей, хотя, по общему признанию, они оказались малоуспешными.

Применительно к атипичной пневмонии одна из трудностей, связанных с моделированием, заключалась в том, что эпидемия оста-

[1]Эта политика часто критикуется на Западе, но, похоже, относительно хорошо принимается многими китайцами.

валась относительно ограниченной в каждой стране, но могла очень быстро распространяться из страны в страну (Гонконг и Китай, Сингапур, Канада ...). Нельзя пренебрегать случайным характером кривых эпидемии в каждом новом направлении. Как мы видели в главах 16 и 22, со стохастическими моделями, как правило, сложнее работать.

В отношении эпидемии чикунгуньи, которая имела место в 2005-2006 годах на острове Реюньон (французская заморская территория в Индийском океане), некоторые модели были ранее разработаны Россом для борьбы с малярией (см. главу 12), причем эти две болезни передаются комарами. Важным аспектом, который следует принимать во внимание, является влияние сезонности. Действительно, популяция комаров уменьшается в течение южной зимы, тем самым снижая передачу болезни. Это можно увидеть на рисунке 26.2, где показано количество новых случаев, сообщаемых каждую неделю небольшой сетью из примерно тридцати врачей общей практики, охватывающей лишь малую часть населения острова. В течение нескольких недель в сентябре и октябре 2005 г. сеть не обнаружила никаких новых случаев заболевания, но его передача все еще продолжалась. Математические модели эпидемии были разработаны в Национальном институте здравоохранения и медицинских исследований (INSERM) и в Тропическом научно-исследовательском институте (IRD). Несмотря на эти модели, никто не мог предвидеть, что эпидемия не исчезнет до конца южной зимы 2005 г., когда она заразила всего несколько тысяч человек. В результате, заразилась почти треть населения острова, то есть около 266 тыс. человек. Это говорит о том, что, если в этом есть необходимость, предсказать будущее эпидемии может быть довольно трудно, и что в первые дни эпидемии не так легко отличить, будет ли это незначительная или серьезная эпидемия. Можно провести параллель с прогнозированием погоды. Такой вид прогнозирования в настоящее время опирается на интенсивное компьютерное моделирование сложных математических моделей океана и атмосферы. Тем не менее, прогнозы, выходящие за рамки нескольких дней, не являются надежными.

С более теоретической точки зрения эпидемия чикунгунья поставила вопрос о том, как адаптировать понятие базового числа воспроизводства $\mathscr{R}_0$ в моделях, которые предполагают, что окружающая среда имеет сезонные (например, периодические) колебания. Адаптация не так проста, и это вызывает определенную озабочен-

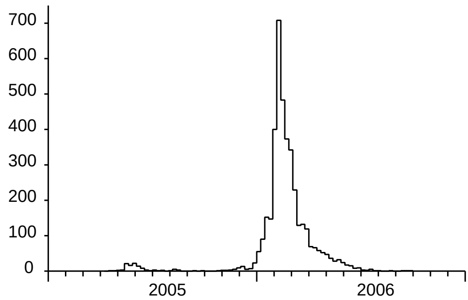


Рис. 26.2: Эпидемия чикунгунья на острове Реюньон в 2005-2006 годах. Количество новых случаев заболевания, регистрируемых в неделю небольшой сетью врачей, в зависимости от времени. Первый небольшой пик был достигнут в мае 2005 года, второй большой пик - в феврале 2006 года. Чтобы получить реальные масштабы эпидемии, цифры на этом рисунке должны быть умножены примерно на 67. Источник: www.invs.sante.fr.

ность в связи с тем, как параметр $\mathscr{R}_0$ был использован для других эпидемий, подверженных влиянию сезонных колебаний, таких как пандемия гриппа H1N1 2009 года.

Еще одной проблемой, вызывающей все большую озабоченность, которую ученые пытались проанализировать, является лекарственная резистентность (антибиотики, противомалярийные препараты). До сих пор в эпидемиологии повторяющийся со времен Даниэля Бернулли и Д'Аламбера вопрос о том, как сбалансировать затраты и выгоды, когда инъекция вакцины несет в себе потенциальный риск, все еще является предметом споров и может остаться таковым, так как чувствительность к риску меняется. Поэтому, следуя некоторым предположениям о том, что вакцина против гепатита B может вызвать серьезные осложнения в очень небольшом числе случаев, Министерство здравоохранения Франции в 1998 г. прекратило свою кампанию по вакцинации в школах, хотя риск оказался незначительным по сравнению с риском смерти после инфицирования вирусом гепатита B.

В экологии изучение динамики популяций рыб все еще создает множество проблем. Тем не менее, оно должно служить научной базой для выбора квот вылова и других ограничений. Превышение вылова анчоуса в Бискайском заливе и красного тунца в Средиземном море - лишь два недавних примера. Поскольку оценка рыбных запасов часто бывает неточной, к моделям, использующим такие данные, следует относиться с осторожностью. Во Франции такого рода исследования в основном проводятся Научно-исследовательским институтом по эксплуатации морских ресурсов (ИФРЕМЕР). Некоторые математические модели также играли определенную роль в предыдущих решениях Международной китобойной комиссии.

В популяционной генетике рассеяние генетически модифицированных организмов также является предметом споров, которые ученые пытались исследовать и разрешить, используя модели «реакции-диффузии», вдохновленные моделью Фишера (см. главу 20). Это область деятельности Национального института сельскохозяйственных исследований (INRA).

С более теоретической точки зрения, можно упомянуть:

- работы по уравнениям в частных производных, таким как уравнения диффузии (см. главу 20) или уравнения с возрастной структурой (см. главу 16);

- работы по стохастическим моделям с пространственным распределением или без него (см. главы 16 и 22), в том числе по случайным сетям, моделирующим распространение эпидемий, а также по детерминистическим приближениям этих моделей.

Этот вид исследований в основном проводится прикладными математиками. В последние годы во французских университетах и других высших учебных заведениях было введено несколько магистерских курсов по математической биологии.

Как и в других научных областях, математическое изучение динамики населения организуется в основном по следующим направлениям:

- «обучение общества»: Общество математической биологии (с 1973 г.), Французское общество теоритической биологии (1985 г.), Японское общество математической биологии (1989 г.), Европейское общество математической и теоретической биологии (1991 г.) и др.

- специализированные журналы: *Bulletin of Mathematical Biology* (с 1939 года), *Mathematical Biosciences* (1967), *Journal of Mathematical Biology* (1974), *Mathematical Medicine and Biology* (1984), *Mathematical Population Studies* (1988), *Mathematical Biosciences and Engineering* (2004) etc.

- конференции (ежегодная встреча Общества математической биологии, математической и вычислительной динамики населения, Европейская конференция по математической и теоретической биологии и др.)

Мы упомянули только те элементы, которые прямо указывают на то, что моделирование находится на границе между математикой и ее применением в динамике народонаселения. Но для каждой конкретной области (демография, экология, генетика популяции, эпидемиология и т.д.) можно найти определенные вопросы, для исследования которых в той или иной степени можно применить математическое моделирование.

В заключение, заинтересованному читателю предлагается ознакомиться с оригинальными статьями, доступными во Всемирной паутине. Адреса приводятся в ссылках в конце каждой главы. Как однажды писал Рональд Фишер о Менделе:

> «История науки сильно пострадала от использования учителями устаревшего материала и, как следствие, исчезновение среды и интеллектуальной атмосферы, в которой происходили великие открытия прошлого. Изучение из первых рук всегда поучительно и часто ... полно сюрпризов.»

Дополнительное чтение

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/

Рисунки

- С . 6. Портрет Томаса Мюррея (около 1687 г.), сделанный Королевским обществом в Лондоне. Chapman, S.: Edmond Halley, F.R.S. 1656–1742. *Notes Rec. R. Soc. Lond.* 12, 168–174 (1957) © The Royal Society.
- С. 13. Портрет Эмануэля Хандманна (1753 г.), хранящийся в музее Кунстмузея в Базеле. *Leonhard Euler 1707–1783, Beiträge zu Leben und Werk*. Birkhäuser, Basel (1983)
- С. 20. Портрет, когда-то принадлежавший Петри-Кирхе, вероятно, был разрушен во время Берлинской битвы в 1945 году. Reimer, K.F.: Johann Peter Süssmilch, seine Abstammung und Biographie. *Arch. soz. Hyg. Demogr.* 7, 20–28 (1932)
- С. 26. Портрет Иоганна Никлауса Грута (ок. 1750-1755), хранящийся в Натур-историческом музее в Базеле. Speiser, D.: *Die Werke von Daniel Bernoulli*, Band 2. Birkhäuser, Basel (1982)
- С. 34. Портрет Мориса Квентина Делатура (1753 г.), представленный на сайте Musée du Louvre в Париже.
- С. 38. Портрет Джона Линнелла (1833 г.), хранящийся в колледже Хейлибери, Англия. Habakkuk, H.J.: Robert Malthus, F.R.S. (1766-1834). *Notes Rec. R. Soc. Lond.* 14, 99–108 (1959)
- С. 42. Гравюра Фламенга (1850). Quetelet, A.: Pierre-François Verhulst. *Annu. Acad. R. Sci. Lett. B.-Arts Belg.* 16, 97–124 (1850)
- С. 49. Heyde, C.C., Seneta, E.: I. J. Bienaymé, *Statistical Theory Anticipated*. Springer-Verlag, New York (1977) © Académie des sciences, Institut de France.
- С. 54. Bateson, W.: *Mendel's Principles of Heredity*. Cambridge University Press (1913)
- С. 59. Pearson, K.: *The Life, Letters, and Labors of Francis Galton*, vol. 1. Cambridge University Press (1914)
- С. 59. Портрет Уотсона в библиотеке Тринити-колледжа Кембриджского университета. Kendall, D.G.: Branching processes since 1873. *J. Lond. Math. Soc.* 41, 385–406 (1966)
- С. 66. Альфред Дж. Лотка Пейперс. Документы по государственной политике. Отдел редких книг и специальных коллекций. © Princeton University Library.
- С. 71. Titchmarsh, E. C.: Godfrey Harold Hardy 1877–1947. *Obit. Not. Fellows R. Soc.* 6, 446–461 (1949)

- С. 74. Stern, C.: Wilhelm Weinberg. *Genetics* 47, 1–5 (1962)
- С. 77. G.H.F.N.: Sir Ronald Ross 1857–1932. *Obit. Not. Fellows R. Soc.* 1, 108–115 (1933) © The Royal Society.
- С. 86. Whittaker, E.T.: Vito Volterra 1860–1940. *Obit. Not. Fellows R. Soc.* 3, 690–729 (1941)
- С. 91. Yates, F., Mather, K.: Ronald Aylmer Fisher, 1890–1962. *Biog. Mem. Fellows R. Soc.* 9, 91–120 (1963) © The Royal Society/Godfrey Argent Studio.
- С. 95. Yates, F.: George Udny Yule. *Obit. Not. Fellows R. Soc.* 8,308–323 (1952)
- С. 104. Heyde, C.C., Seneta, E. (eds.): *Statisticians of the Centuries.* Springer, New York (2001)
- С. 115. britannica.com/EBchecked/topic/252257/J-B-S-Haldane © Bassano and Vandyk Studios.
- С. 125. Hill, W.G.: Sewall Wright, 21 December 1889-3 March 1988. *Biog. Mem. Fellows R. Soc.* 36, 568–579 (1990) © Llewellyn Studios, Chicago.
- С. 120. Nybølle, H.C.: Agner Krarup Erlang f. 1. Januar 1878 - d. 3. Februar 1929. *Mat. Tidsskr. B*, 32–36 (1929)
- С. 133. 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)
- С. 133. *I. G. Petrowsky Selected Works Part I.* Gordon and Breach, Amsterdam (1996) © Taylor and Francis Books UK.
- С. 139. Фотография Дениса Кемпсона. Crowcroft, P.: *Elton's Ecologists, a History of the Bureau of Animal Population.* University of Chicago Press (1991)
- С. 143. © Geoffrey Grimmett.
- С. 151. 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.
- С. 157. © Samuel Schlaefli / ETH Zürich.
- С. 166. Selected works of J. Song. *Science Press*, Beijing (1999) © Song Jian.

Оглавление

В этой книге прослеживается история динамики населения - теоретической дисциплины, тесно связанной с генетикой, экологией, эпидемиологией и демографией, - где математика принесла значительные познания. В ней представлен обзор генезиса нескольких важных тем: экспоненциальный рост, от Эйлера и Мальтуса до китайской политики в отношении одного ребенка; разработка стохастических моделей, от законов Менделя и вопроса о вымирании фамилий до теории перколяции для распространения эпидемий, и хаотические популяции, где переплетаются детерминизм и случайность.

С последними достижениями в области машинного перевода виртуальная монополия одного языка в научной литературе больше не оправдана. Растущее лингвистическое отчуждение в университетах может быть обращено вспять. Этим переводом на русский язык, тщательно отредактированным двумя специалистами по динамике населения, мы поощряем этот новый путь.

ISBN : 979-10-343-8016-9

15€

Николя Бакаэр   Краткая история математической динамики населения