니콜라 바카에르
(Nicolas Bacaër)

나경아　김광수

# 수학적 인구 역학의 짧은 역사

# 수학적 인구 역학의 짧은 역사

니콜라 바카에르
(Nicolas Bacaër)

나경아　김광수

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

나경아
국가수리과학연구소
knah@nims.re.kr

김광수
국립부경대학교
kwangsukim@pknu.ac.kr

이 책의 종이책을 구매하고 싶은 독자는 이메일(nicolas.bacaer@ird.fr)로 문의하시기 바랍니다.

표지 사진: 책거리 (국립 기메 동양 박물관)

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

Pour l'édition coréenne :
© Nicolas Bacaër, Paris, 2024
ISBN : 979-10-426-1860-5
Dépôt légal : 2024

# 소개

인구 역학은 인간, 동물, 식물 또는 미생물과 같은 생물학적 집단의 크기와 구성의 시간적 변화를 단순한 기계적 방식으로 설명하려는 과학 분야입니다. 이는 인구 통계의 보다 설명적인 영역과 관련이 있지만 여전히 상당히 구별됩니다. 한 가지 공통점은 수학적 언어를 광범위하게 사용한다는 점입니다.

인구 역학은 수학, 사회과학(인구학), 생물학(인구 유전학 및 생태학), 의학(역학) 등 다양한 분야의 교차점에 있습니다. 그 결과 다양한 응용 분야에서 나타나는 문제들 간의 유사성에도 불구하고 전체적으로 제시되는 경우는 많지 않습니다. 프랑스어로 된 주목할 만한 예외는 알랭 힐리온의 ≪수학적 인구 이론≫(1986)이라는 책입니다. 이 책은 수학자의 관점에서 이산 시간 모델($t = 0, 1, 2...$)과 연속 시간 모델($t$는 실수), 결정론적 모델(현재 상태를 정확히 알면 미래 상태를 정확히 알 수 있음)과 확률론적 모델(확률이 중요한 역할을 함)과 같은 다양한 유형의 모델을 구분하여 이 주제를 제시합니다. 그런 다음 논리적으로 이산 결정론적 모델, 연속 결정론적 모델, 이산 확률론적 모델 및 연속 확률론적 모델을 고려합니다.

이 책에서는 동일한 주제를 역사적 관점에서 논의하려고 노력했습니다. 연구는 그 맥락에서 설명됩니다. 과학자들의 짧은 전기가 포함되어 있습니다. 이것은 수학에 익숙하지 않은 사람들이 책을 더 쉽게 읽을 수 있도록 하고 일반적으로 연구 중인 문제의 기원을 이해하는 데 도움이 될 수 있습니다. 하지만 이 책은 역사에 관한 책만이 아닙니다. 수학적 모델링에 대한 입문서 역할도 할 수 있습니다. 독자가 모델의 한계를 실제로 볼 수 있도록 대부분의 계산에 대한 세부 사항을 포함하는 것이 중요하다고 생각했습니다. 기술적인 부분은 박스 안에 강조 표시되어 있으며 처음 읽을 때는 건너뛸 수 있습니다. 마지막 장에서는 수학적 관점에서 분석할 수 있는 인구 역학 분야의 다양한 현대적 문제에 초점을 맞췄습니다. 더 많은 것을 알고 싶은 분들을 위해 각 장의 마지막에 있는 참고 문헌 목록에는 원본을 다운로드할 수 있는 웹 사이트도 포함되어 있습니다.

이 정도 분량의 책에서 지금까지 개발된 모든 연구 결과를 완벽하게 설명하거나 이 주제에 기여한 모든 과학자에 대해 이야기하는 것은 불가능했습니다. 특히 최근 수십 년 동안의 선택에는 필연적으로 임의적인 요소가 포함되어 있습니다. 그럼에도 불구하고 선택한 샘플이 충분히 대표성을 지니고 있기를 기대하며, 언급되지 않은 분야에서 활동하는 사람들을 노하게 만들지 않기를 바랍니다.

이 책의 이상적인 독자는 다음과 같습니다.

- 자신이 수강해야 하는 수학 과목과 주변 세계 사이에 어떤 연관성이 있는지 궁금해 하는 고등학생과 대학생. 또는 인구 역학과 관련된 주제로 개인 과제를 준비하는 학생.
- 자신의 수업을 더 매력적으로 만들려고 노력하는 수학 교사. 사칙연산에 대한 지식만 있으면 1, 2, 5장의 대부분을 이해할 수 있습니다. 3장은 로그의 응용에 대한 입문서 역할을 할 수 있습니다. 이 책에서는 또한 다음을 다룹니다.
  - 점화식: 1, 3, 8, 11, 14, 21, 23, 24장
  - 미분 방정식: 4, 6, 12, 13, 16장
  - 편미분 방정식: 20, 25장
  - 적분 방정식: 10장
  - 확률론의 응용: 2, 7, 8, 9, 15, 16, 17, 18, 19, 22장
- 인구학, 역학, 유전학 또는 생태학에 이미 익숙하고 자신이 좋아하는 분야를 유사한 수학적 모델을 포함하는 다른 분야와 비교해보고자 하는 사람들.
- 과학의 역사에 관심이 있는 독자.

이 책은 첫 단계로 DeepL 소프트웨어를 통해 번역되었습니다. 이 기계 번역의 교정을 도와주신 나경아, 김광수 님께도 감사드립니다.

# 목차


1 피보나치 수 (1202) 1

2 핼리의 생명표 (1693) 5

3 오일러와 인구의 기하급수적 성장 (1748–1761) 12

4 다니엘 베르누이, 달랑베르와 천연두 예방 접종 (1760) 26

5 맬서스와 기하급수적 성장의 장애물 (1798) 39

6 베르휼스트와 로지스틱 방정식 (1838) 44

7 비에나메, 쿠르노, 그리고 가족 성의 멸종 (1845–1847) 51

8 멘델과 유전형질 (1865) 56

9 골턴, 왓슨, 그리고 소멸 문제 (1873–1875) 61

10 로트카와 안정적인 인구 이론 (1907–1911) 69

11 하디-바인베르크 법칙 (1908) 75

12 로스와 말라리아 (1911) 81

13 로트카, 볼테라와 포식자-피식자 시스템 (1920–1926) 88

14 피셔와 자연 선택 (1922) 97

15 율과 진화 (1924) 102

16 전염병 모델링의 맥켄드릭과 커맥 (1926–1927) 112

17 할데인과 돌연변이 (1927) 123

18 멸종 문제에 대한 얼랑과 스테펜센의 견해 (1929–1933) 128

19 라이트 및 무작위 유전적 드리프트 (1931) 134

20 유전자의 확산 (1937) 141

21 레슬리 매트릭스 (1945) 150
22 퍼콜레이션 및 전염병 (1957) 155
23 게임 이론과 진화 (1973) 162
24 카오스적인 개체수 (1974) 169
25 중국의 한 자녀 정책 (1980) 179
26 현대의 몇 가지 문제 188

1장

# 피보나치 수 (1202)

1202년 피사의 레오나르도(피보나치라고도 불림)는 아랍 수학자들도 채택한 인도의 십진수 체계를 유럽에 널리 알린 책을 출간했다. 이 책에 제시된 많은 예 중 하나는 토끼의 개체 수 증가에 관한 것이다. 이 예는 인구의 역학에 대한 수학적 모델의 가장 오래된 예 중 하나이다.

사후 피보나치(Fibonacci) 라는 이름이 붙여진 피사의 레오나르도는 지중해세계에서 상업과 군사력이 절정에 달했던 1170년경에 피사 공화국에서 태어났다. 1192년경 피보나치의 아버지는 공화국에 의해 현재 알제리에 있는 베자이아 항구로 파견되어 무역소를 이끌었다. 그의 아들은 얼마 지나지 않아 상인이 될 준비를 하기 위해 아버지가 있는 곳으로 갔다. 레오나르도는 아랍인들이 인도에서 가져온 십진수 체계를 배우기 시작했다. 이는 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 처럼 오늘날에도 거의 동일한 형태를 갖고 있다. 그는 지중해 주변을 여행하면서 다양한 숫자 체계를 비교하고 아랍 수학을 연구했다. 피사로 돌아온 그는 1202년에 새로운 수 체계를 설명하고 회계, 무게 및 통화 변환, 이자율 및 기타 여러 응용 분야에의 적용 방법을 보여주는 라틴어로 된 ≪Liber Abaci≫ (계산의 책) 이라는 책을 완성했다. 또한 아랍인들에게 알려진 대수학과 산술에 대한 대부분의 결과를 정리했다.

피보나치는 그의 책에서 오늘날 인구 역학의 문제라고 할 만한 것을 다뤘다. 이 책의 앞부분은 $28 = 14 + 7 + 4 + 2 + 1$과 같이 인자의 합인 완전수에 관한 문제이고, 뒷부분은 네 사람이 돈을 나누는 문제로서 네 개의 방정식으로 이루어진 선형 시스템과 동일하다. 다음은 인구 문제를 라틴어로 번역한 것이다.

> "한 남자가 밀폐된 장소에 토끼 한 쌍을 갖고 있다. 토끼 한 쌍이 생후 두 번째 달부터 매 달 다른 한 쌍을 낳을 때 1년에 몇 쌍의 토끼가 태어나는지 알고자 한다."

첫 달 초에 한 쌍의 갓 태어난 토끼가 있다면, 이 쌍은 한 달이 지나도 아직 번식하지 못하고 두 번째 달 초에는 여전히 한 쌍의 토끼만 있을 것이다. 이 토끼 한 쌍은 세 번째 달이 시작될 때 다른 한 쌍을 낳아 총 두 쌍이 될 것이다. 첫 번째 토끼 한 쌍은 네 번째 달 초에 다시 다른 한 쌍을 낳는다. 그러나 두 번째 토끼 쌍은 아직 새끼를 낳지 못한다. 토끼는 세 쌍만 남게 된다.

현대식 표기법을 사용하여 $P_n$을 $n$번째 달 초의 토끼 쌍의 수라고 하자. 그러면 $n+1$번째 달의 토끼 쌍의 수 $P_{n+1}$는 $n$번째 달에 있는 쌍의 수 $P_n$과 $n+1$번째 달에 새로 태어난 쌍의 수 의 합이다. 그러나 생후 2개월 이상 된 토끼 쌍만 $n+1$ 달에 새로운 토끼 쌍을 낳는다. 이들은 $n-1$개월에 이미 존재했던 쌍이며, 그 수는 $P_{n-1}$이다. 따라서

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

이 성립한다. 이것은 점화식으로, 이전 두 달의 인구 수로 $n+1$째 달의 인구를 표현한다. 이를 이용해 피보나치는 $1+1=2$, $1+2=3$, $2+3=5$, $3+5=8$ 등과 같은 아래의 표를 쉽게 만들 수 있었다.

| $n$ | 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 |

실제로 피보나치는 2개월째의 상황을 초기 조건으로 간주했다. $P_{14} = 144 + 233 = 377$이므로, 피보나치는 시작 시점으로부터 12개월 후 마침내

377쌍의 토끼를 얻게 된다. 그는 이러한 수열이 무한히 계속될 수 있다는 것을 알아냈다.

1202년 이후 피보나치는 1220년에 ≪Practica geometriae≫ (실용 기하학), 1225년에 ≪Liber quadratorum≫ (사각형의 책) 등 여러 권의 책을 저술했다. 그의 명성은 과학을 높이 평가하는 황제 프레데릭 2세 (Friedrich II)와의 만남으로 이어졌다. 1240년 피사 공화국은 피보나치에게 연금을 수여했다. 그의 사망 연도는 알려지지 않았다.

그 후 몇 세기 동안 피보나치의 토끼 문제는 잊혀졌고 인구 역학에 대한 수학적 모델 개발에도 영향을 미치지 못했다. 몇몇 과학자들은 연구에서 동일한 수열을 발견했지만 피보나치나 인구에 대한 언급은 없었다. 케플러(Kepler)의 여러 저서에는 $n$이 무한대로 갈 때 비율 $P_{n+1}/P_n$이 황금비 $\phi = (1+\sqrt{5})/2$에 수렴한다는 내용이 포함되어 있다. 이는 대부분의 인구 모델에 공통적으로 나타나는 특성인 기하학적으로 증가하는 경향의 특정 경우이다 (3장 및 21장 참조). 1728년 다니엘 베르누이 (Daniel Bernoulli)는 점화식에 대해 연구하여 다음과 같은 일반항에 대한 정확한 공식을 얻었다.

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

피보나치의 업적에 대한 완성본은 19세기에 출판되었다. 그 이후로 피보나치 수열 $(P_n)$은 피보나치 수열이라는 이름으로 교양 수학 책에서 찾아볼 수 있게 되었다.

토끼 개체군 모델링에서 피보나치 수열을 도출할 때에 토끼의 사망이나 암수 구분 등은 고려되지 않았으며, 이러한 가설들이 현실과 거리가 먼 것은 사실이다. 최근 수십 년 동안 생물학계에서 이 수열에 관심을 갖게 된 것은 솔방울의 8과 13, 해바라기의 34와 55와 같이 일부 식물에 $P_n$의 몇몇 숫자들이 포함된 구조가 존재한다는 사실에서 비롯되었다. 심지어 피보나치 수열의 특성과 응용에 관한 과학 저널인 ≪The Fibonacci Quarterly≫는 전적으로 피보나치 수열에 관한 내용만을 다루고 있다!

# 참고문헌

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

2장

# 핼리의 생명표 (1693)

1693년 영국의 유명한 천문학자 에드먼드 핼리는 카스파르 노이만이 왕립학회에 전달한 브레슬라우 시의 출생 및 사망 기록을 연구했다. 그는 같은 해에 태어난 코호트에서 일정 나이까지 생존한 사람들의 수를 보여주는 생명표를 만들었다. 그는 또한 이 표를 사용하여 종신연금의 가격을 계산했다. 이 장에서는 이 작업을 회상하며 핼리의 생애와 그레이언트, 페티, 드 비트, 후드, 후이겐스, 라이프니츠, 드 모브르와 같은 사람들이 관심을 가졌던 '정치 산술'과 확률이론의 초기 발전에 대해 알아본다.

에드먼드 핼리(Halley)는 1656년 런던 근처에서 태어났다. 그의 아버지는 부유한 비누 제조업자였다. 에드먼드는 어린 나이에 천문학에 관심을 갖게 되었다. 그는 옥스퍼드 대학교 퀸즈 칼리지에서 공부를 시작했다. 1675년 그리니치 천문대가 개장했을 때 핼리는 이미 왕립 천문학자인 플램스티드(Flamsteed)를 방문할 수 있었다. 그는 1676년부터 1678년까지 학업을 중단하고 세인트 헬레나 섬으로 가서 남반구에서 볼 수 있는 별들의 카탈로그를 만들었다. 영국으로 돌아온 그는 왕립학회 회원이 되었다. 그는 세인트 헬레나로 여행하는 동안 바람의 순환에 대해 관찰한 내용을 발표하기도 했다. 1684년에는 케플러(Kepler)의 행성 운동 법칙과 태양이 가하는 인력 사이의 연관성을 논의하기 위해 케임

브리지에 있는 뉴턴을 방문했다. 그는 뉴턴에게 유명한 ≪자연철학의 수학적 원리≫를 쓰도록 권유했고, 뉴턴은 마침내 자비로 이 책을 출판했다. 당시 그는 왕립학회 서기로 일하고 있었다. 1689년에는 수중 잠수용 종을 설계하여 직접 테스트하기도 했다.

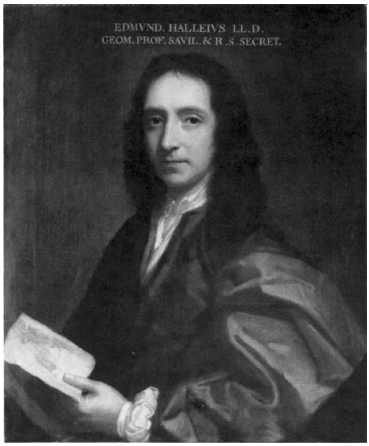


그림 2.1: 핼리 (1656–1742)

비슷한 시기에 브레슬라우에 살던 신학자 카스파르 노이만(Neumann)은 도시의 출생 및 사망자 수에 대한 데이터를 수집하고 있었다. 브레슬라우는 합스부르크 제국에 속해 있었다 (현재는 폴란드에 속해 있으며 브로츠와프라고 불린다). 데이터에는 사람들이 사망한 나이가 포함되어 있었다. 따라서 특정 연령까지 생존할 확률을 보여주는 생명표를 작성하는 데 사용할 수 있었다.

최초의 생명표는 1662년 런던에서 ≪사망률에 관한 자연적, 정치적 관찰≫이라는 제목의 책으로 출판되었다. 이 책은 일반적으로 통계학과 인구학의 기초가 되는 책으로 간주되며, 특이한 점은 오늘날까지도 이 책이 런던의 상인이자 책 표지에 표시된 저자인 존 그라운트(Graunt)가 쓴 것인지 아니면 그의 친구이자 왕립학회 창립자 중 한 명인 윌리엄 페티(Petty)가 쓴 것인지 모른다는 것이다. 이 책에 포함 된 생명표는 17세기 초부터 런던에서 정기적으로 장례와 세례를 보고하던 회보를 활용하였다. 이 회보는 사람들에게 반복되는 페스트의 유행에 대해 알리는 데 주로 사용되었다. 이러한 이유로 회보에는 사람들이 사망한 연령이 아닌

사망 원인이 표시되었다. 연령에 따른 생존 확률을 나타내는 생명표를 얻기 위해 그라운트나 페티는 연령대에 따라 사망 원인이 어떻게 다른지 추측해야 했다. 따라서 생명표에는 큰 오차가 있을 수 있다. 그럼에도 불구하고 이 책은 1662년부터 1676년까지 다섯 차례에 걸쳐 출간되며 큰 성공을 거두었다. 이후 유럽의 여러 도시에서 런던과 유사한 회보를 발행하기 시작했다.

노이만이 라이프니츠(Leibniz)의 제안대로 왕립학회 서기 헨리 저스텔(Justel)에게 1687년부터 1691년까지 브레슬라우 시의 인구 통계 데이터를 보낸 것은 첫 생명표가 나온 지 거의 30년이 지나서의 일이다. 저스텔은 얼마 지나지 않아 사망했고, 핼리는 이 데이터를 입수하여 분석한 후 1693년 ≪왕립학회지≫에 자신의 결론을 발표했다. 그의 논문 제목은 ≪브레슬라우 시의 출생과 장례식에 대한 흥미로운 표를 통해 도출한 인류의 사망률에 대한 추정과 생명에 대한 연금 가격 책정의 시도≫였다.

5년의 연구기간 동안 핼리는 브레슬라우의 출생자 수가 사망자 수와 거의 같아서 총 인구가 거의 일정하다는 것을 발견했다. 분석을 단순화하기 위해 그는 인구가 정확히 정상상태에 있다고 가정했다. 즉, 연간 출생 수($P_0$ 라고 함), 총 인구, $k$세의 인구($P_k$) 및 $k$세의 연간 사망자 수($D_k$)는 시간이 지남에 따라 모두 일정하다는 것이다. 런던과 같이 빠르게 성장하는 도시에서는 시골에서 유입되는 인구로 인해 통계가 편향되었을 것이기 때문에 이러한 단순화가 불가능했을 것이고, 이는 브레슬라우 데이터의 또 다른 흥미로운 속성을 강조한다.

브레슬라우의 데이터에는 연간 평균 1,238명의 출생이 있었고 핼리는 이 값을 $P_0$로 취했다. 원칙적으로 그는 데이터로부터 모든 $k \geqslant 0$에 대해 $k$세 인구의 연간 사망자 수 $D_k$의 평균도 계산할 수 있었다. 다음의 공식

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

를 사용하면 $P_k$를 제공하는 테이블 2.1을 만들 수 있다. 반대로 $D_k = P_k - P_{k+1}$ 공식에서 그가 사용한 $D_k$의 값을 찾을 수 있다. 이는 $D_0 = 238$, $D_1 = 145$, $D_2 = 57$, $D_3 = 38$ 등이다. 실제로 핼리는 반올림한 값을 얻기 위해 ($D_1$의 경우 $P_1 = 1,000$이 되도록 약간 변경됨), 또는 5년의 연구기간

표 2.1: $k$세의 인구 $P_k$를 나타내는 핼리의 생명표.

| $k$ | $P_k$ | $k$ | $P_k$ | $k$ | $P_k$ | $k$ | $P_k$ | $k$ | $P_k$ | $k$ | $P_k$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 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 |

동안 노인 사망자 수가 적기 때문에 불규칙성을 완화하기 위해 결과를 약간 조정했다. 핼리는 표에 있는 모든 숫자 $P_k$의 합을 구하여 브레슬라우의 총 인구를 34,000명에 가깝게 추산했다.[1] 요약하면, 이 방법은 일반적인 인구 조사가 필요하지 않고 몇 년 동안 출생과 사망의 수와 사망 연령만 알면 된다는 큰 장점이 있었다.

핼리의 생명표는 18세기의 다양한 연구에서 참고 자료로 사용되었다 (4장 참조). 비록 $P_k$의 값은 브레슬라우시에 한정된 것이지만, $P_{k+1}/P_k$의 비율은 이미 $k$ 나이에 도달했을 때 $k+1$ 나이까지 생존할 확률이라고 생각할 수 있다. 이 확률은 당시 다른 유럽 도시의 인구에도 합리적으로 사용할 수 있다. 예를 들어, 1세 어린이가 10세에 도달할 확률은 1,000분의 661, 20세에 도달할 확률은 1,000분의 598일 것으로 예상할 수 있다.

핼리는 또한 생명표를 사용하여 종신연금의 가격을 계산했다. 16세기와 17세기에 여러 도시와 주에서 시민들에게 이러한 연금을 판매하여 자금을 모았다. 구매자는 사망 할 때까지 매년 처음 지불 한 금액의 일정 비율에 해당하는 고정 금액을 받았으며, 이는 종종 당시 이자율의 두 배에 해당하지만 구매자의 나이와는 무관하게 지급되었다. 물론 기대 수명이 매우 긴 사람들이 너무 많이 연금을 구입하면 기관은 파산 할 위험이 있

[1] 84세 이상의 경우 핼리는 그 수가 107명이라고만 언급했다.

었다. 신뢰할 수 있는 생명표 없이는 이 문제를 제대로 해결할 수 없었다.

홀란드 총리 요한 드 비트(de Witt)와 암스테르담 시장 중 한 명인 요하네스 후데(Hudde)는 1671년 이미 종신연금의 가격을 계산하는 문제에 대해 생각해 보았다. 프랑스 군대의 침공을 두려워한 그들은 군대를 강화하기 위한 자금을 모으고자 했다. 그들은 수십 년 전에 연금을 구입한 사람들에 관한 데이터, 특히 연금을 구입한 나이와 사망한 나이에 관한 데이터를 가지고 있었다. 그들은 연금의 가격을 어느 정도 정확하게 계산하는 데 성공했지만 그 방법은 나중에 잊혀졌다. 이듬해 홀란드는 침략을 당했고 드 비트는 군중에게 린치를 당했다.

핼리는 1693년 브레슬라우의 생명표를 사용하여 이자율을 6% 가정하고 이 문제를 새롭게 고려했다. 계산 방법은 간단하다. $i$를 이자율이라고 하고, $R_k$를 나이 $k$인 사람이 연간 1파운드의 연금을 구입할 수 있는 가격이라고 하자. 이 사람이 $k+n$의 나이에 여전히 살아있을 확률은 $P_{k+n}/P_k$ 이다. 이 사람이 이 나이에 도달하면 국가가 지급하겠다고 약속한 파운드는 초기 가격에 $1/(1+i)^n$ 파운드를 곱하면 구할 수 있다 (이자율이 $i$일때). 따라서 초기 금액이 연금을 지불하는 데만 사용된다는 단순한 가정을 하면 가격은 다음과 같아야 한다.

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

이렇게 해서 핼리는 초기 연금 가격을 구하기 위해 원하는 연금액에 곱해야 하는 계수 $R_k$를 보여주는 표 2.2를 얻었다. 따라서 20세 남성은 초기 금액의 $1/12.78 \approx 7.8\%$를 매년 받게 된다. 그러나 50세 남성은 생존기간이 더 적기 때문에 $1/9.21 \approx 10.9\%$를 받게 된다. 이자율이 두배가 되면 연금액이 초기 금액의 12%, 혹은 연금 가격이 연금액의 8.33배에 해당한다는 것을 알 수 있다.

표 2.2: 종신 연금의 가격을 계산해 주는 배율.

| $k$ | $R_k$ | $k$ | $R_k$ | $k$ | $R_k$ | $k$ | $R_k$ | $k$ | $R_k$ |
|---|---|---|---|---|---|---|---|---|---|
| 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 |

물론 계산은 상당히 지루하다. 그럼에도 불구하고 핼리는 상용로그표를 사용하여 $P_{k+n}/(1+i)^n$ 이라는 일반항을 더 빨리 구할 수 있었다. 그는 84세 이상의 $P_k$ 에 대한 값을 보여주지 않았기 때문에 그의 계산을 정확히 확인할 수는 없다. 마지막으로, 핼리의 연구는 즉각적인 영향을 미치지 못했다. 수십 년 동안 영국과 다른 지역에서 종신 연금은 구매자의 나이와 무관하게 연금액의 7배와 같이 훨씬 낮은 가격으로 계속 판매되었다.

생명표에서 파생된 질문은 핼리 시대에 많은 과학자의 관심을 끌었다. 1657년 확률론에 관한 최초의 책자를 저술한 네덜란드의 크리스티안 후이겐스(Huygens)는 1669년 동생과의 서신에서 그라운트의 생명표와 기대 수명의 계산에 대해 논의했다.[2] 라이프니츠는 노이만을 왕립학회에 소개하기 몇 년 전에도 기대 수명 계산에 관한 에세이를 썼는데, 이 글은 출판되지 않았다. 1709년에는 니콜라우스 베르누이 1세(Nikolaus I Bernoulli)의 차례였다. 1725년 아브라함 드 모브르(de Moivre)는 ≪연금에 관한 논문≫ 전체를 출판했다. 그는 특별히 고연령의 경우 공식(2.2)에 몇 개의 항만 포함되어 있기 때문에 연금 가격 $R_k$ 를 쉽게 계산할 수 있다는 사실을 발견했다. 그리고는 다음의 역 점화식

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

을 사용할 수 있었다. 이 식은 방정식(2.2)로부터 쉽게 증명할 수 있다. 따라서 핼리가 70세의 가격에 대해 계산한 값을 사용하여 표 2.2의 다른 값들을 확인해 볼 수 있다.[3]

인구학에 초점을 맞춘 이 연구 이후 핼리는 자신의 주요 연구 주제로 돌아갔다. 1698년에서 1700년 사이에 대서양을 항해하며 지구 자기장 지도를 그렸다. 1704년에는 옥스퍼드 대학교의 교수가 되었다. 이듬해 혜성에 관한 책을 출판하고 케플러가 1607년에 관측한 1682년의 혜성이 1758년에 다시 올 것이라고 예측했는데, 이 혜성은 '핼리의 혜성'으로 알려지게 되었다. 또한 페르게의 아폴로니우스(Απολλώνιοσ ο Περγαίοσ)

[2] 나이 $k$의 기대 수명은 $i=0$ 일때의 공식 (2.2)에 의해 주어진다.
[3] 5세와 15세를 포함해 표에 몇 가지 오류가 있는 것 같다.

가 쓴 원뿔에 관한 책을 번역하여 출판했다. 1720년 그는 플램스티드의 뒤를 이어 왕실천문관으로 임명되었다. 그는 항해에 매우 실용적인 문제인 달의 관측을 통해 바다에서 경도를 정확하게 구하는 문제를 해결하려고 노력했다. 그는 1742년 그리니치에서 86세의 나이로 사망했다.

# 참고문헌

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

3장

# 오일러와 인구의 기하급수적 성장 (1748–1761)

오일러는 인구 역학에 대해 여러 차례 글을 썼다. 1748년 논문인 ≪무한대 분석 입문≫에서 지수 함수를 다룬 장에는 인구의 기하급수적 증가에 대한 네 가지 예가 포함되어 있다. 1760년에는 이 기하급수적 증가를 인구의 연령 구조와 결합한 논문을 발표했다. 이 연구는 20세기에 개발되어 인구학에서 중요한 역할을 하는 '안정적' 인구 이론의 초석이다. 1761년 오일러는 인구학에 관한 논문의 두 번째 판을 집필하는 데에도 쥐스밀히를 도왔다. 그는 피보나치 수열의 일종의 변형인 흥미로운 모델을 연구했지만 자세한 분석은 발표하지 않았다.

레오나르드 오일러(Euler)는 1707년 스위스 바젤에서 태어났다. 그의 아버지는 개신교 목사였다. 1720년 오일러는 대학에서 공부를 시작했다. 또한 라이프니츠(Leibniz)와 뉴턴(Newton) 이후 당대 가장 유명한 수학자 중 한 명인 요한 베르누이(Johann Bernoulli)에게 개인 수학 레슨을 받았다. 그는 요한 베르누이의 두 아들 니콜라우스 2세(Nikolaus II)와 다니엘(Daniel)과 친구가 되었다. 1727년 오일러는 다니엘을 따라 상트페테르부르크에 새로 설립된 과학 아카데미에 들어갔다. 그는 수학

외에도 물리학 및 기타 여러 과학 기술 분야에도 관심이 많았다. 1741년 프로이센의 프레데릭 2세(Friedrich II)는 그를 베를린 과학 아카데미의 수학 부문 책임자로 초빙했다. 오일러는 역학(천문학, 탄성, 유체, 고체)과 수학(정수론, 대수학, 무한 급수, 초등 함수, 복소수, 미적분, 미분 및 편미분 방정식, 최적화, 기하학)뿐만 아니라 인구학에 대해서도 상당한 수의 논문과 저서를 발표했다. 그는 당대 가장 다작을 한 수학자였다.

그림 3.1: 오일러 (1707–1783)

1748년 오일러는 ≪무한대 분석 입문≫이라는 제목의 논문을 라틴어로 출판했다. 그는 이 책에서 지수 함수와 로그에 관한 여섯 가지 예제를 고려했다. 음악 음계의 수학적 이론에 관한 한 가지 예, 이자가 있는 대출 상환에 관한 또 다른 예, 인구 역학에 관한 네 가지 예 등이 그것이다. 후자의 예제에서 오일러는 연도 $n$의 인구 $P_n$이 모든 정수 $n$에 대해 다음과 같이 주어진다고 가정했다.

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

여기서 성장률 $x$는 양의 실수이다. 초기 조건 $P_0$에서 시작하여, 연도 $n$의 인구는

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

으로 주어진다. 이를 기하급수적 또는 지수적 성장이라고 한다. 첫 번째

예에서는 다음과 같은 질문을 던진다.

> "처음에 10만명 이었던 어떤 지역의 인구가 매년 1/30씩 증가할 경우 100년 후의 인구를 알고자 한다."

답은 $P_{100} = (1+1/30)^{100} \times 100,000 \approx 2,654,874$이다. 이 예에서 오일러는 1747년에 실시된 베를린 인구 조사에서 107,224명의 인구를 추산한 데서 영감을 얻었다. 그의 계산은 인구가 한 세기 안에 10배 이상 증가할 수 있다는 것을 보여준다. 이는 당시 런던에서 관찰된 증가율과 정확히 일치한다.

$(1+1/30)^{100}$을 계산하는 것은 현대의 휴대용 계산기를 사용하면 매우 쉽다. 하지만 오일러의 시대에는 수많은 곱셈을 손으로 하는 것을 피하고 결과를 빠르게 얻기 위해 대수를 사용해야 했다. 먼저 $P_{100}$의 상용로그(밑이 10인 로그)를 계산한다. 로그의 기본 속성인 $\log(ab) = \log a + \log b$를 이용하면 다음의 식을 얻게 된다.

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

로그 함수는 1614년 스코틀랜드인 존 네이피어(Napier)에 의해 소개되었다. 그의 친구 헨리 브릭스(Briggs)는 1617년에 첫번째 상용로그표를 발표했다. 1628년 네덜란드의 아드리안 블라크(Vlacq)는 1부터 100,000까지 10의자리 단위의 정수에 대한 상용로그표를 발표하여 브릭스의 작업을 완성했다. 오일러가 $\log 30 \approx 1.477121255$, $\log 31 \approx 1.491361694$, 그리고 마지막으로 $\log P_{100} \approx 6.4240439$를 얻기 위해 사용한 테이블이 바로 이 테이블이다. 이제 그 로그값을 알고 있는 숫자 $P_{100}$를 찾아야 한다. 1에서 100,000까지 정수의 상용로그는 0에서 5까지의 범위이므로 대신 $P_{100}/100$의 로그, 즉 4.4240439를 찾는다. 로그 테이블에서 $\log 26,548 \approx 4.424031809$ 및 $\log 26,549 \approx 4.424048168$을 확인할 수 있다. 로그 함수를 26,548과 26,549 사이의 직선으로 바꾸어 오일러는 다음과 같은 결과를 얻는다.

$$\frac{P_{100}}{100} \approx 26,548 + \frac{4.4240439 - 4.424031809}{4.424048168 - 4.424031809} \approx 26,548.74\,.$$

따라서 $P_{100} \approx 2,654,874$ 이다.

오일러의 책에 나오는 인구 역학에 관한 두 번째 예는 다음과 같다.

> "홍수 이후 6명에서 시작한 집단의 후손이 200년 후 1,000,000 명이라고 가정할 때 인구의 연간 증가율을 구하고자 한다."

$10^6 = (1+x)^{200} \times 6$ 이므로 휴대용 계산기를 사용하면 $x = (10^6/6)^{1/200} - 1 \approx 0.061963$ 을 구할 수 있다. 로그 테이블을 사용하면 $\log(10^6) = 200 \times \log(1+x) + \log 6$ 에서 $\log(1+x) = (6 - \log 6)/200 \approx 0.0261092$ 이 되고 $1+x \approx 1.061963$ 이 된다. 따라서 오일러는 인구가 매년 $x \approx 1/16$ 씩 증가한다는 결론을 내릴 수 있었다. 이 예의 기원을 이해하려면 현대 철학자들이 성경 이야기의 진실성을 부정하기 시작했다는 사실을 기억해야 한다. 성경에 따르면 홍수의 시기는 기원전 2350년 경이며 생존자는 노아와 그의 세 아들, 그리고 그들의 아내로 명시되어 있다. 창세기는 이렇게 말한다.

> "노아의 이 세 아들로부터 백성이 온 땅에 퍼지니라."

대홍수 이후 매년 1/16(또는 6.25%) 인구 증가율은 오일러에게 그리 비현실적으로 보이지 않았다. 개신교 목사의 아들로 태어나 평생을 신앙생활을 해온 그는 이러한 결론을 내렸다.

> "이런 이유로 믿지 않는 사람들이 한 사람으로 시작한 인구가 그렇게 짧은 시간 안에 지구 전체를 채울 수 없다고 주장하는 것은 어처구니가 없다."[1]

오일러는 또한 홍수 이후 400년 동안 같은 속도의 성장이 계속되었다면 인구는 $(1+x)^{400} \times 6 = (10^6/6)^2 \times 6 \approx 1660$ 억이 되었을 것이라는 사실도 알아냈다.

[1] 1662년 그라운트가 출판한 책(챕터 2 참조)에서도 비슷한 언급을 찾을 수 있다.

> "아담과 하와가 성경에 따른 세계의 나이인 5,160년 동안 64년마다 인구를 두 배로 늘렸다면 지금보다 훨씬 더 많은 사람들이 생겨났을 것이다. 그러므로 세계의 나이는 어떤 사람들이 헛되이 상상하는 것처럼 10만 년이 넘지 않았고 성경이 말하는 것보다도 많지 않다."

"하지만 지구 전체가 그 인구를 감당할 수는 없을 것이다."

이 아이디어는 반세기 후 맬서스에 의해 비약적으로 발전하게 된다 (5장 참조).

오일러의 세 번째 예는 다음과 같이 묻는다.

"한 세기마다 인구가 두 배로 증가한다면 연간 증가율은 얼마인가?"

$(1+x)^{100}=2$ 이므로 휴대용 계산기를 사용하면 $x=2^{1/100}-1\approx 0.00695$ 를 구할 수 있다. 로그표를 사용하면 $100\log(1+x)=\log 2$ 에서 $\log(1+x)\approx 0.0030103$ 가 되고 $1+x\approx 1.00695$ 가 된다. 따라서 인구는 매년 $x\approx 1/144$ 씩 증가한다. 네 번째이자 마지막 예시도 비슷한 질문을 던진다.

"인구가 매년 $1/100$ 씩 증가한다면 인구가 10배가 되는 데 얼마나 걸리는지 알고자 한다."

$(1+1/100)^n=10$ 를 통해 $n\log(101/100)=1$ 을 구할 수 있다. 따라서 $n=1/(\log 101-2)\approx 231$ 년이다. 이것이 인구 역학에 관한 1748년의 ≪무한대 분석 입문≫에서 찾을 수 있는 전부이다. 오일러는 몇 년 후 이 주제를 더 자세히 다루게 된다.

1760년 그는 베를린 과학 아카데미 회보에 ≪인간종의 사망률과 증식에 대한 탐구≫라는 제목의 논문을 발표했다. 이 연구는 인구의 기하급수적 증가에 대한 그의 분석과 생명표에 대한 이전의 연구를 종합한 것이었다 (2장 참조). 오일러는 다음과 같은 문제를 고려했다.

"사망률에 대한 가정이 주어졌을 때 1년 동안 발생하는 출생과 사망 수를 알면 모든 생존자의 수와 인구의 연간 증가율을 구할 수 있다."

오일러는 여기서 다음의 숫자들을 이미 알고 있다고 가정했다.

- $n$ 년도의 출생 수, $B_n$.

- $n$ 년도의 사망 수, $D_n$.
- 연령 $k \geqslant 1$에 도달하는 신생아 중 $q_k$의 비율.

$P_n$을 연도 $n$의 인구라고 하자. 오일러는 추가적으로 다음의 두 가정을 하고 있다.

- 인구는 기하급수적으로 증가한다. 즉, $P_{n+1} = rP_n$ 이다. (우리의 경우 $r = 1+x$ 이다.)
- 출생과 인구의 비율은 일정하다. 즉, $B_n/P_n = m$ 이다.

이 두 가지 가정에 따르면 출생 수는 동일한 비율로 기하급수적으로 증가한다. 정확히는 $B_{n+1} = rB_n$ 이다. 오일러는 100세 이후에는 아무도 생존하지 못한다고 가정하고 100년 간격, 즉 $n = 0$에서 $n = 100$ 사이의 인구 상태를 고려했다. 명확한 표현을 위해, $n-k$ 연도에 태어나 연도 $n$ 초까지 생존한 인구를 $P_{k,n}$ $(k \geqslant 1)$이라고 한다. 연도 $n$의 출생자 수는 $P_{0,n} = B_n$ 이다. 생존 계수 $q_k$의 정의로부터 $P_{k,n} = q_k P_{0,n-k} = q_k B_{n-k}$를 구할 수 있다. 따라서

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

위의 등식을 $r^{100}P_0$으로 나누면 다음을 얻을 수 있다.

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

이는 인구 통계학에서 '오일러 방정식'이라고도 불리는 방정식이다. 출생과 사망을 따로 계산하면 다음과 같이 된다.

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

따라서 사망자 수도 $D_{n+1} = rD_n$ 으로 기하급수적으로 증가한다. 또한,

다음의 식

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

을 (3.1)의 등식에 대입하면 최종적으로 다음과 같은 방정식이 도출된다.

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

여기서 미지수는 $r$ 하나만 남는다. 다른 매개변수들의 함수로서 $r$을 표현할 수 없기 때문에 이를 일반적으로 음(형태)의 방정식이라고 부른다. 하지만 고정된 값 $r$에 대해 방정식 (3.4)의 좌변과 우변을 계산하고 $r$을 변화시켜 가며 두 변이 같아지게 만들 수 있다. 이렇게 계산한 $r$의 값으로 인구의 성장률 $x = r - 1$을 구할 수 있다. 방정식 (3.1) 및 (3.3)에서 인구 $P_n$에 대해 다음과 같은 식을 구할 수 있다.

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

인구 변화가 없는 경우($r = 1$) 이 표현식은 핼리(Halley)가 브레슬라우시의 인구를 추정하는 데 사용한 것과 동일하다 (2장 참조).

오일러는 다음의 문제도 고려했다.

> "사망률과 출산율에 대한 가설이 주어지고 총 인구를 알고 있을 때 각 연령에 몇 명의 인구가 있을지 구하려 한다."

생존 계수 $q_k$와 출산 계수 $m$을 알고 있으므로 성장률 $r$은 방정식 (3.1)에서 계산할 수 있다. $n$년도에 구한 $n-k$년에 태어난 사람의 수는 $q_k B_{n-k} = q_k B_n / r^k$ 이다 ($q_0 = 1$). 따라서 전체 인구 중 연령 $k$의 비율은

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

이다. 이 비율은 일정하다. 로트카(Lotka)의 용어를 사용하면(10장 참조) 인구는 '안정적' 이다. 즉, 연령 피라미드가 시간이 지나도 같은 모양을 유지한다.

오일러는 인구가 고정되어 있지 않고 기하급수적으로 증가할 때 생

명표를 구하는 문제를 재검토했다.

> "1년 동안의 총 인구, 출생 수와 연령별 사망자 수를 알면 사망률의 법칙을 구할 수 있다."

오일러가 말한 '사망률의 법칙' 이란 생존 계수 $q_k$의 집합을 의미한다. 핼리의 경우와는 달리 이제 총 인구는 인구 조사를 통해 알 수 있다고 가정한다 (2장 참조). 방정식 (3.2)는 성장률이

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

임을 보여준다. $D_{k,n}$를 $n$년도에 $k$세에 사망한 사람의 수라고 한다. 즉, 이 사람들은 $n-k$년에 태어났다. 따라서 $D_{k,n} = (q_k - q_{k+1})B_{n-k}$이다. 한편, $B_{n-k} = B_n / r^k$이다. 따라서 생존 계수 $q_k$는 $q_0 = 1$와 임의의 $k \geqslant 0$에 대한 점화식

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

을 사용하여 계산할 수 있다. 이 공식에 $B_n$을 곱하면 핼리가 $r = 1$의 안정적 경우에 사용한 공식 (2.1)이 나온다. 하지만 오일러는 자신이 고안한 생존 계수 $q_k$의 계산 방법은 인구가 규칙적으로 증가한다는 가정에 기반하며, 전염병 유행, 전쟁, 기근 등과 같은 사고는 고려되지 않았다는 사실을 강조했다. 오일러가 살던 시대의 인구조사에 스웨덴의 경우처럼 인구의 연령이 기록되어 있었다면 이러한 가정은 불필요했을 것이고 계수 $q_k$를 더 쉽게 계산할 수 있었을 것이다.

오일러는 생존 계수 $q_k$가 주어졌을 때 종신연금의 가격을 계산하는 방법도 보여주었다. 그는 이 주제에 대한 핼리와 모브르의 연구는 언급하지 않았다. 오일러는 5% 의 이자율과 네덜란드인 빌럼 케르세붐 (Kersseboom)이 1742년 발표한 생명표를 사용했다.

베를린 아카데미에서 인구학에 관심을 보인 과학자는 오일러뿐이 아니었다. 그의 동료인 요한 페터 쥐스밀히(Süssmilch)는 1741년에 ≪인간의 출생, 사망, 출산을 통한 인간 세대의 변화에 있어서의 신성한 질서≫라는 제목의 책을 독일어로 출판했는데, 이 책은 오늘날 인구학에

관한 최초의 책으로 여겨지고 있다. 쥐스밀히는 1752년에 ≪베를린 시의 급속한 성장에 관하여≫라는 책을 저술하기도 했다.

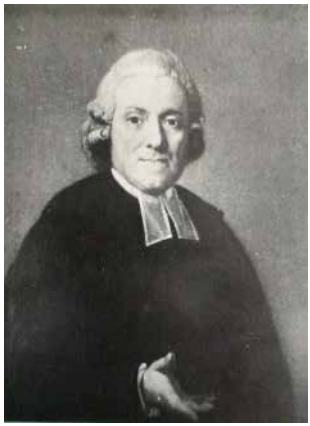

그림 3.2:
쥐스밀히 (1707–1767)

1761년 쥐스밀히는 논문의 두 번째 판을 출판했다. ≪인구의 증가율과 배가시간에 대하여≫라는 제목의 장에 그는 오일러가 자신을 위해 만든 흥미로운 수학적 모델을 포함시켰다. 이 모델은 피보나치(1장 참조)의 모델과 유사하지만 인간 인구에 대한 모델이다. 오일러는 0 년도에 20세인 한 부부 (남자 한 명, 여자 한 명)를 시작으로 사람들이 40세에 사망하고 20세에 결혼하며, 각 부부는 22세에 두 명의 자녀(남자 아이와 여자 아이)를, 24세에 또 다른 두 명을, 26세에 마지막 두 명을 낳아 총 6명의 자녀를 둔다고 가정했다. 오일러는 연도를 2씩 세어 $B_i$가 연도 $2i$ 동안의 출생 수가 되도록 하여 모든 $i \geqslant 1$에 대해

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

라는 결론을 내렸다. 초기 조건은 $B_{-12} = 0$, $B_{-11} = 0$, $B_{-10} = 2$, 그리고 모든 $-9 \leqslant i \leqslant 0$에 대한 $B_i = 0$ 이다. 따라서 오일러는 표 3.1의 두 번째 열에 표시된 것처럼 출생 수를 계산할 수 있다. 그러면 연도 $2i$의 사망자 수 $D_i$ 는 연도 $2i-40$의 출생자 수와 같다. 여기 $i \geqslant 10$ 의 경우 $D_i = B_{i-20}$ 이고, $i \leqslant 9$ 의 경우 $D_i = 0$ 이다. 연도 $2i$에 생존한 사람의 수 $P_i$ 는 연도 $2i-2$에 생존한 사람의 수에 연도 $2i$의 출생자 수를 더한 값에서 연도 $2i$

의 사망자 수를 뺀 값과 같다. 즉, $P_i = P_{i-1} + B_i - D_i$ 이다.

쥐스밀히의 책에서 이 장은 피보나치 수열에 대해서도 적용 가능했을 법한 발언으로 끝난다.

> "오일러의 표에 만연해 보이는 무질서는 출생 수가 점화 수열이라고 부르는 일종의 수열을 따르는 것을 막지 못한다 [...] 이러한 수열의 초기 불규칙 상태가 무엇이든 방해받지 않으면 기하수열로 바뀌고 시작의 불규칙 상태가 조금씩 사라지고 결국 거의 완전히 사라진다."

이 책에서는 이 인구 모델의 수학에 대해 더 자세히 설명하지 않았다. 그러나 오일러는 ≪인류의 증식에 관하여≫ 라는 제목의 원고를 통해 이 연구를 훨씬 더 발전시켰는데, 이 원고는 생전에 출판되지 않았다. 기하급수의 형태를 가진, 즉 $B_i = c\,r^i$ 형태 방정식 (3.5)의 해를 찾는 도중 그는 13차수의 다항 방정식을 구했다.

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

그는 $r = 1$에 가까운 해를 찾다가 $r^{13}$의 계산을 위해 로그표를 사용하여

$$1 + r + r^2 - r^{13} \approx \begin{cases} 0.212 & (r = 1.09)\,, \\ -0.142 & (r = 1.10) \end{cases}$$

라는 것을 알아냈다. 따라서 방정식 (3.6)의 근은 1.09에서 1.10 사이이다. 이 구간의 선분으로 함수 $1 + r + r^2 - r^{13}$을 근사화하여 오일러는

$$r \approx \frac{0.142 \times 1.09 + 0.212 \times 1.10}{0.142 + 0.212} \approx 1.0960$$

를 얻는다. 연도를 2씩 계산할 때, 출생자 수는 매년 $\sqrt{r}$배가 되는 경향이 있다. 이 수치는 $(\sqrt{r})^n = 2$인 경우, 즉 $n = 2\log 2/\log r \approx 15$년마다 두 배가 된다. 점근적으로 $B_i \approx c\,r^i$이고 $2i$년도의 사망자 수 $D_i$는 $B_{i-20}$와 같으므로 $r^{20} \approx 6.25$로 $D_i \approx B_i/r^{20}$를 구할 수 있다. 출생자 수는 사망자 수의 약 6배이다. 연도 $2i$에 생존한 사람의 수 $P_i$는 $B_i + B_{i-1} + \cdots + B_{i-19}$

표 3.1: 오일러의 표.

| $i$ | 출생 수 | 사망자 수 | 인구 규모 |
|---|---|---|---|
| 0 | 0 | 0 | 2 |
| 1 | 2 | 0 | 4 |
| 2 | 2 | 0 | 6 |
| 3 | 2 | 0 | 8 |
| 4 | 0 | 0 | 8 |
| 5 | 0 | 0 | 8 |
| 6 | 0 | 0 | 8 |
| 7 | 0 | 0 | 8 |
| 8 | 0 | 0 | 8 |
| 9 | 0 | 0 | 8 |
| 10 | 0 | 2 | 6 |
| 11 | 0 | 0 | 6 |
| 12 | 2 | 0 | 8 |
| 13 | 4 | 0 | 12 |
| 14 | 6 | 0 | 18 |
| 15 | 4 | 0 | 22 |
| 16 | 2 | 0 | 24 |
| 17 | 0 | 0 | 24 |
| 18 | 0 | 0 | 24 |
| 19 | 0 | 0 | 24 |
| 20 | 0 | 0 | 24 |
| 21 | 0 | 2 | 22 |
| 22 | 0 | 2 | 20 |
| 23 | 2 | 2 | 20 |
| 24 | 6 | 0 | 26 |
| 25 | 12 | 0 | 38 |
| 26 | 14 | 0 | 52 |
| 27 | 12 | 0 | 64 |
| 28 | 6 | 0 | 70 |
| 29 | 2 | 0 | 72 |
| 30 | 0 | 0 | 72 |
| 31 | 0 | 0 | 72 |
| 32 | 0 | 2 | 70 |
| 33 | 0 | 4 | 66 |
| 34 | 2 | 6 | 62 |
| 35 | 8 | 4 | 66 |
| 36 | 20 | 2 | 84 |
| 37 | 32 | 0 | 116 |
| 38 | 38 | 0 | 154 |
| 39 | 32 | 0 | 186 |
| 40 | 20 | 0 | 206 |
| 41 | 8 | 0 | 214 |
| 42 | 2 | 0 | 216 |
| 43 | 0 | 2 | 214 |
| 44 | 0 | 6 | 208 |
| 45 | 2 | 12 | 198 |
| 46 | 10 | 14 | 194 |
| 47 | 30 | 12 | 212 |
| 48 | 60 | 6 | 266 |
| 49 | 90 | 2 | 354 |
| 50 | 102 | 0 | 456 |
| 51 | 90 | 0 | 546 |
| 52 | 60 | 0 | 606 |
| 53 | 30 | 0 | 636 |
| 54 | 10 | 2 | 644 |
| 55 | 2 | 8 | 638 |
| 56 | 2 | 20 | 620 |
| 57 | 12 | 32 | 600 |
| 58 | 42 | 38 | 604 |
| 59 | 100 | 32 | 672 |
| 60 | 180 | 20 | 832 |
| 61 | 252 | 8 | 1,076 |
| 62 | 282 | 2 | 1,356 |
| 63 | 252 | 0 | 1,608 |
| 64 | 180 | 0 | 1,788 |
| 65 | 100 | 2 | 1,886 |
| 66 | 42 | 10 | 1,918 |
| 67 | 14 | 30 | 1,902 |
| 68 | 16 | 60 | 1,858 |
| 69 | 56 | 90 | 1,824 |
| 70 | 154 | 102 | 1,876 |
| 71 | 322 | 90 | 2,108 |
| 72 | 532 | 60 | 2,580 |
| 73 | 714 | 30 | 3,264 |
| 74 | 786 | 10 | 4,040 |
| 75 | 714 | 2 | 4,752 |
| 76 | 532 | 2 | 5,282 |
| 77 | 322 | 12 | 5,592 |
| 78 | 156 | 42 | 5,706 |
| 79 | 72 | 100 | 5,678 |
| 80 | 86 | 180 | 5,584 |
| 81 | 226 | 252 | 5,558 |
| 82 | 532 | 282 | 5,808 |
| 83 | 1,008 | 252 | 6,564 |
| 84 | 1,568 | 180 | 7,952 |
| 85 | 2,032 | 100 | 9,884 |
| 86 | 2,214 | 42 | 12,056 |
| 87 | 2,032 | 14 | 14,074 |
| 88 | 1,568 | 16 | 15,626 |
| 89 | 1,010 | 56 | 16,580 |
| 90 | 550 | 154 | 16,976 |
| 91 | 314 | 322 | 16,968 |
| 92 | 384 | 532 | 16,820 |
| 93 | 844 | 714 | 16,950 |
| 94 | 1,766 | 786 | 17,930 |
| 95 | 3,108 | 714 | 20,324 |
| 96 | 4,608 | 532 | 24,400 |
| 97 | 5,814 | 322 | 29,892 |
| 98 | 6,278 | 156 | 36,014 |
| 99 | 5,814 | 72 | 41,756 |
| 100 | 4,610 | 86 | 46,280 |
| 101 | 3,128 | 226 | 49,182 |
| 102 | 1,874 | 532 | 50,524 |
| 103 | 1,248 | 1,008 | 50,764 |
| 104 | 1,542 | 1,568 | 50,738 |
| 105 | 2,994 | 2,032 | 51,700 |
| 106 | 5,718 | 2,214 | 55,204 |
| 107 | 9,482 | 2,032 | 62,654 |
| 108 | 13,530 | 1,568 | 74,616 |
| 109 | 16,700 | 1,010 | 90,306 |
| 110 | 17,906 | 550 | 107,662 |
| 111 | 16,702 | 314 | 124,050 |
| 112 | 13,552 | 384 | 137,218 |
| 113 | 9,612 | 844 | 145,986 |
| 114 | 6,250 | 1,766 | 150,470 |
| 115 | 4,664 | 3,108 | 152,026 |
| 116 | 5,784 | 4,608 | 153,202 |
| 117 | 10,254 | 5,814 | 157,642 |
| 118 | 18,194 | 6,278 | 169,558 |
| 119 | 28,730 | 5,814 | 192,474 |

와 같으므로, 또한 다음을 구할 수 있다.

$$P_i \approx B_i \left(1 + \frac{1}{r} + \cdots + \frac{1}{r^{19}}\right) = B_i \frac{1 - r^{20}}{r^{19} - r^{20}} \approx 9.59 B_i \,.$$

총 인구는 출생아 수의 약 10배에 달한다.

표에 표시된 수열 $(B_i)$이 실제로 $r^i$처럼 점근적으로 증가한다는 것을 증명하는 것은 좀 더 복잡하다. 아브라함 드 모브르(De Moivre)의 점화수열 연구 이후 '생성 함수'

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

가 $f(x)$를 유리함수로 표현하는 데 도움이 될 수 있다는 사실이 알려졌다. 오일러는 1748년 ≪무한대 해석 입문≫에서 이 방법을 설명했다. 점화식(3.5)으로부터 다음을 얻을 수 있다.

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

따라서 다음이 성립한다.

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

오일러는 이 유리 함수가 다음과 같은 형태로 분해될 수 있다는 것을 알고 있었다.

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

숫자 $x_1, \ldots, x_{13}$은 방정식 $1 - x^{11} - x^{12} - x^{13} = 0$의 실근 또는 복소근

이다. 따라서

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

이다. $B_i$는 $f(x)$에서 $x^i$의 계수이므로, 오일러는

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

임을 구했다. 여기서 $x_k$는 가장 작은 크기를 가진 근이다. 즉, $B_i$는 $(1/x_k)^i$처럼 기하급수적으로 증가하는 경향이 있다. 알아두어야 할 것은 $r = 1/x_k$가 방정식 ((3.6)의 근인것과 $x_k$가 $1 - x^{11} - x^{12} - x^{13} = 0$ 방정식의 근인 것은 동치관계라는 점이다. 이러한 증명의 구체적인 부분들은 1916년 굼벨에 의해 마침내 명확히 증명 되었다.

쥐스밀히는 1765년에 그의 책의 세 번째 판을 출간한 뒤 1767년 베를린에서 사망했다. 프로이센 국왕과 사이가 좋지 않았던 오일러는 1766년 상트페테르부르크로 돌아왔다. 시력을 잃었음에도 불구하고 그는 아들과 동료들의 도움을 받아 대수학, 적분, 광학 및 조선에 관한 많은 저작을 계속 발표했다. 1760년에서 1762년 사이에 베를린에서 쓴 ≪독일 공주에게 보내는 자연철학의 여러 주제에 관한 편지≫는 1768년에서 1772년 사이에 출판되어 유럽 전역에서 베스트셀러가 되었다. 오일러는 1783년 상트페테르부르크에서 사망했다. 수학적 인구학에 대한 그의 공헌, 특히 기하급수적으로 증가하는 인구에서 '안정된' 연령 피라미드에 대한 그의 분석은 20세기에 이르러서야 재발견 되었다 (10장 및 21장 참조).

# 참고문헌

1. Euler, L.: Recherches générales sur la mortalité et la multiplication du genre humain. Hist. Acad. R. Sci. B.-Lett. Berl. 16, 144–164 (1760). 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 (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)

4장

# 다니엘 베르누이, 달랑베르와 천연두 예방 접종 (1760)

1760년 다니엘 베르누이는 천연두 모델링에 대한 논문을 썼다. 당시에는 사람을 보호할 수 있지만 치명적일 수도 있는 예방접종에 대해 많은 논란이 있었다. 그는 핼리의 생명표와 천연두에 관한 몇 가지 데이터를 사용하여 사망 위험이 11% 미만인 경우 접종이 유리하다는 것을 보여주었다. 접종으로 출생 시 기대 수명을 최대 3년까지 늘릴 수 있다는 것이었다. 달랑베르는 역학 분야에서 최초의 수학적 모델인 베르누이의 연구를 비판했다.

다니엘 베르누이(Daniel Bernoulli)는 1700년 네덜란드 흐로닝언에서 태어났다. 그의 가족에는 이미 두 명의 유명한 수학자, 즉 아버지 요한 베르누이(Johann Bernoulli)와 삼촌 야콥 베르누이(Jakob Bernoulli)가 있었다. 1705년 요한은 스위스 바젤로 이주하여 야콥의 사망으로 공석이 된 교수직을 맡았다. 요한은 아들이 수학을 공부하는 것을 원하지 않았다. 그래서 다니엘은 의학으로 전향하여 1721년 호흡에 관한 논문으로 박사 학위를 취득했다. 그는 베니스로 이주하여 수학에 집중하기 시작했고 1724년에 책을 출판했다. 같은 해 파리 과학 아카데미에서 ≪바다에 떠 있는 배 위의 모래시계의 완벽성에 관하여≫라는 에세이로 상을 받은

그는 새로운 상트페테르부르크 아카데미에서 교수직을 얻었다. 이 기간 동안 그는 점화식이나 확률론의 '상트페테르부르크의 역설'에 대해 집중적으로 연구했다. 1733년 다니엘 베르누이는 바젤 대학으로 돌아와 식물학, 생리학, 물리학을 연속으로 가르쳤다. 1738년에는 물리학 역사에 길이 남을 유체 역학에 관한 책을 출간했다. 1753년경 그는 오일러(Euler), 달랑베르(d'Alembert)와 동시에 진동하는 끈의 문제에 관심을 갖게 되었고, 이는 중요한 수학적 논쟁을 불러일으켰다.

그림 4.1:
다니엘 베르누이 (1700–1782)

1760년 그는 파리의 과학 아카데미에 ≪천연두로 인한 사망률과 예방 접종의 이익에 대한 새로운 분석 시도≫라는 제목의 논문을 제출했다. 문제는 접종(감염의 예방을 위해 독성이 덜한 천연두 바이러스를 자발적으로 체내에 소량 주입하는 것)이 죽음에 이르게 할 위험이 있음에도 장려되어야 하는지에 대한 것이었다. 이 기술은 아시아에서 오랫동안 알려져 왔으며 1718년 영국에서 오스만 제국 주재 영국 대사의 아내인 몬타구 여사에 의해 소개되었다. 프랑스는 1711년 루이 14세의 장남이 천연두로 사망했음에도 불구하고 예방 접종이 선뜻 도입되지 않았다. 1723년 천연두에서 살아남아 영국에서 수년간 망명 생활을 하며 최신의 혁신 기술을 관찰했던 볼테르(Voltaire)는 1734년 ≪철학 편지≫에서 예방 접종의 도입을 옹호했다. 천연두에서 살아남은 또다른 프랑스 과학자 라 콩다민(La Condamine)도 1754년 파리의 과학 아카데미에서 접종을

청원했다.

1759년 바젤에서 죽기 전, 모페르튀이(Maupertuis)는 다니엘 베르누이에게 수학적 관점에서 예방 접종 문제를 연구해 보라고 권유했다. 더 정확하게 말하면, 접종의 장기적인 이득과 즉각적인 사망 위험을 비교하는 방법을 찾는 것이 과제였다. 문제를 단순화 하기 위해 베르누이는 다음과 같은 가정을 세웠다.

- 천연두에 처음 감염된 사람은 (나이와 무관하게) $p$의 확률로 사망하고 $1-p$의 확률로 생존한다.
- 모든 사람은 매년 감염될 확률 $q$를 가지며, 더 정확하게는 한 개인이 $x$세에서 $x+dx$세 사이에 감염될 확률은 $q\,dx$이며, 여기서 $dx$는 무한소의 시간을 의미한다.
- 천연두에서 살아남은 사람들은 평생 동안 새로운 감염으로부터 보호받는다 (획득된 면역으로 인해).

천연두 이외의 원인으로 인한 연령 $x$의 사망률을 $m(x)$라고 하면, 한 개인이 연령 $x$와 연령 $x+dx$ 사이의 무한소 기간 $dx$에 사망할 확률은 $m(x)\,dx$이다. 같은 해에 태어난 $P_0$의 사람들 그룹을 고려할 때, 다음을 정의한다.

- $S(x)$는 천연두에 감염된 적이 없이 $x$세까지 생존해 있는 '감수성' 인구의 수이다.[1]
- $R(x)$는 천연두를 앓고 생존한 $x$세의 사람 수이다.
- $P(x)=S(x)+R(x)$는 나이 $x$에 생존한 총 인구 수이다.

출생은 나이 $x=0$에 해당한다. 따라서 $S(0)=P(0)=P_0$ 및 $R(0)=0$이다. 다니엘 베르누이는 17세기 말 뉴턴(Newton)과 라이프니츠(Leibniz), 그리고 그 후 아버지가 개발한 미적분법을 적용하여 나이 $x$와 나이 $x+dx$ ($dx$는 무한히 작음) 사이에서 각 감수성 있는 개인이 천연두에 감염될 확률 $q\,dx$와 다른 원인으로 사망할 확률 $m(x)\,dx$가 있음을 알아냈다. 따

[1] 더 정확하게는 정수 단위가 아닌 연속 함수처럼 변할 수 있는 이 숫자의 수학적 기대치이다.

라서 감수성자 수의 변화는 $dS = -Sq\,dx - Sm(x)\,dx$ 이며, 이는 다음의 미분 방정식으로 나타낼 수 있다.

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

이 방정식에서 $dS/dx$는 함수 $S(x)$의 미분이라고 한다. 동일한 짧은 시간 간격 동안 천연두로 사망하는 사람의 수는 $pSq\,dx$이고 천연두에서 생존하는 사람의 수는 $(1-p)Sq\,dx$이다. 또한 천연두가 아닌 다른 원인으로 사망하는 사람은 $Rm(x)\,dx$ 이다. 이것은 두 번째 미분 방정식으로 이어진다.

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

두 방정식을 더하면 다음과 같다.

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

베르누이는 방정식 (4.1)과 (4.3)을 통해 $x$ 나이에 여전히 감수성자인 사람의 비율을 다음처럼 나타낼 수 있었다.

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

공식 (4.4)을 구하기 위해 베르누이는 방정식 (4.1)와 (4.3)에서 $m(x)$를 제거했다.

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

식을 재정리하면 다음의 식을 얻을 수 있다.

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

왼쪽은 $f(x) = S(x)/P(x)$의 도함수이며, 이는 $x$세 인구 중 감염에

취약한 사람의 비율이다. 따라서 다음과 같다.

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

이러한 유형의 방정식의 해는 당시로부터 수십년 전 다니엘의 삼촌인 야콥 베르누이(Jakob Bernoulli)의 연구 덕분에 알려져 있었다. 방정식을 $f^2$로 나누고 $g(x) = 1/f(x)$로 설정하면 다음과 같이 된다. $dg/dx = qg - pq$이며 $g(0) = 1/f(0) = 1$이라는 것을 알 수 있다. $h(x) = g(x) - p$로 설정하면 $dh/dx = qh$이다. 따라서 $h(x) = h(0)e^{qx} = (1-p)e^{qx}$가 된다. 마지막으로 $g(x) = (1-p)e^{qx} + p$ 및 $f(x) = 1/g(x)$이다.

베르누이는 자신의 이론을 적용하기 위해 핼리(Halley)의 생명표를 사용했다(2장 참조). 이 표는 0년도에 태어난 1,238명의 코호트 집단 중 $x$년($x = 1, 2...$)이 시작될 때 아직 살아있는 사람의 수를 나타낸다. 그러나 베르누이의 모델에서는 실제로 $x$세까지 도달하는 사람 수 $P(x)$가 필요했는데, 이는 약간 달랐다. 대부분의 동시대 사람들과 마찬가지로 베르누이는 그 차이를 깨닫지 못했기 때문에 (핼리의 논문은 실제로 명확하지 않다.) 핼리의 표에서 첫 번째 숫자 1,238을 제외한 나머지 숫자를 그대로 유지했으며 이를 1,300으로 대체하여 생후 첫해의 현실적인 사망률을 얻었다. 이 숫자는 표 4.1의 두 번째 열에 나와 있다.

베르누이는 천연두로 사망할 확률로 $p = 1/8 = 12.5\%$를 선택했는데, 이는 당시의 관측과 일치하는 수치이다. 연간 천연두에 걸릴 확률 $q$는 직접 추정할 수 없었다. 따라서 베르누이는 $q$에 대해 여러 가지 값을 시도해 보았을 것이고, 마침내 아래의 모든 계산을 거친 후 천연두로 인한 사망자 수가 전체 사망자 수의 약 1/13이 되는 값을 선택했는데, 이는 당시 유럽의 여러 도시에서 관찰되었던 비율이었다. 앞으로 살펴보겠지만 연간 $q = 1/8$로 선택한 것이 잘 맞는 것으로 판명되었다.[2]

공식(4.4)와 표의 두 번째 열에 있는 $P(x)$ 값을 사용하면 $x$ 연령의 감수성자의 수 $S(x)$를 계산할 수 있다. 이것을 가장 가까운 정수로 반올림한 값이 표의 세 번째 열에 나타나 있다. 네 번째 열은 천연두를 앓고 생존한

[2] $p$와 $q$가 같은 것은 우연의 일치이다.

표 4.1: 핼리의 생명표와 베르누이의 계산.

| 나이 $x$ | 생존 인구 $P(x)$ | 감수성 인구 $S(x)$ | 면역 인구 $R(x)$ | 천연두로 인한 사망 | 천연두 없음 $P^*(x)$ |
|---|---|---|---|---|---|
| 0 | 1300 | 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 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

$x$세 인구의 수 $R(x) = P(x) - S(x)$를 보여준다. 나이 $x$에 해당하는 행의 다섯 번째 열은 나이 $x$에서 나이 $x+1$ 사이의 천연두로 인한 사망자 수를 표시한다. 이론적으로 이 숫자는 적분값 $pq\int_x^{x+1} S(t)\,dt$ 이어야 하지만 $pq[S(x)+S(x+1)]/2$ 공식을 사용하면 그림 4.2에서 처럼 좋은 근사치를 얻을 수 있다. 사다리꼴의 면적은 곡선 아래의 면적, 즉 함수의 적분값에 가깝다.

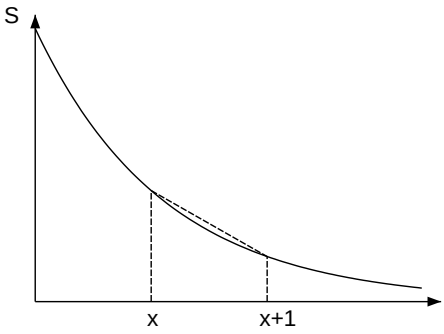


그림 4.2: 점선으로 표시된 사다리꼴의 면적으로 $x$와 $x+1$ 사이의 함수 $S$의 적분을 근사할 수 있다.

베르누이는 다섯 번째 열에 있는 모든 숫자를 합하여 24세 이전에 천연두로 사망한 사람이 98명이라는 것을 알 수 있었다. 이 표를 이후의 연령까지 계속 확장하면 24세에도 여전히 천연두에 감수성이 있는 32명 중 천연두로 인한 사망자는 3명이 될 것이라는 것을 알 수 있다. 즉, 1,300명의 출생자 중 101명이 천연두로 사망할 운명이다. 이는 기대한 비율인 1/13과 거의 일치한다.

베르누이는 태어날 때 모든 사람에게 천연두를 접종하여 사망자가 발생하지 않는 상황을 고려했다. 천연두는 박멸될 것이고 이 경우 문제는 증가할 기대 수명을 추정하는 것이다. 동일한 출생 수 $P_0$에서 시작하여, 천연두가 사라졌을 때 $x$세가 된 사람들의 수를 $P^*(x)$라고 하자. 그러면

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

이다. 베르누이는 다음의 등식이 성립함을 보였다.

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

여기서 $P(x)$는 앞에서와 같이 천연두가 존재할 때 $x$ 연령의 인구이다.

실제로 방정식 (4.6)와 (4.3) 에서 이전과 같이 $m(x)$를 제거하여 재배열 하면 베르누이가 얻은 수식이 된다.

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

그는 $h(x) = P(x)/P^*(x)$로 설정했다. 공식 (4.4)에서 분자와 분모에 $e^{-qx}$을 곱한 식을 사용하면

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

가 되고, 이것은 $\frac{d}{dx}\log h = \frac{d}{dx}\log(1-p+pe^{-qx})$와 동일하다. 여기서 log는 10진수가 아닌 자연 로그를 나타낸다. 여기서 $h(0)=1$ 이므로 $h(x) = 1-p+pe^{-qx}$ 이다.

연령 $x$가 충분히 높을 때 $P(x)/P^*(x)$ 비율은 $1-p$가 되는 경향이 있음을 알 수 있다. 표의 여섯 번째 열은 $P^*(x)$를 보여준다. $P(x)$와 $P^*(x)$를 비교하는 한 가지 방법은 출생 시 기대 수명을 추정하는 것인데, 천연두가 존재할 때의 기대수명은 이론적으로 다음과 같다.

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

천연두가 존재하지 않을 때에는 $P^*(x)$를 $P(x)$로 대체한 식이 기대수명이 된다. 베르누이는 근사화한 공식 $[\frac{1}{2}P(0)+P(1)+P(2)+\cdots]/P_0$을 사용했는데, 이는 사다리꼴 방법(그림 4.2)에 의해 주어진 공식이다. 24세 이후 84세까지 표를 계속 이어가면 (표 2.1 참조), 천연두가 있는 경우의 기대수명 $E$는 $[\frac{1}{2}1{,}300+1{,}000+\cdots+20]/1{,}300 \approx 26.57$ 년, 즉 26년 7개월이 된다. 천연두가 없다면 기대 수명 $E^*$는 $[\frac{1}{2}1{,}300+1{,}015+\cdots+23]/1{,}300 \approx$

29.65 년, 즉 29년 8개월이 된다. 출생시 예방접종을 하면 기대 수명이 3년 이상 늘어나는 것이다.

베르누이가 이 공식을 구하는 데 사용한 방법보다 더 간단하고 빠른 방법이 있다. $S(x)$에 대한 미분 방정식 (4.1)을 살펴보면 다음과 같은 사실을 알 수 있다.

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

이 식을 $R(x)$에 대한 방정식 (4.2)에 사용하면 다음을 얻는다.

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

$P^*(x)$ 에 대한 방정식 (4.6)에 의해

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

이다. 또한, 공식 (4.4) 와 (4.7)도 성립한다!

물론 독성이 덜한 천연두 균주를 접종한다고 해서 완전히 안전한 것은 아니다. $p'$가 접종 직후 천연두로 사망할 확률 $(p' < p)$이라면, 모든 사람이 출생 시 접종을 받을 경우의 기대 수명은 $(1-p')E^*$가 된다. 이 기대 수명은 $p' < 1 - E/E^*$, 또는 약 11% 일때 자연 기대 수명 $E$보다 높게 유지된다. 당시에는 $p'$에 관한 데이터를 구하기 어려웠다. 하지만 베르누이는 위험 $p'$가 1% 미만이라고 추정했다. 국가가 예방 접종을 장려해야 한다는 데는 의심의 여지가 없었다. 그는 이렇게 끝맺었다.

> "저는 인류의 안녕과 밀접한 관련이 있는 문제에서 약간의 분석과 계산을 통해 얻을 수 있는 지식 없이 어떤 결정도 내려지지 않기를 바랄 뿐이다."

베르누이의 연구는 1760년 4월 파리의 과학 아카데미에서 발표되었다. 11월에 달랑베르는 ≪천연두 예방 접종에 대한 확률 이론의 적용에

관하여≫라는 제목의 논평을 발표했다. 이 논평은 얼마 지나지 않아 그의 ≪수학 소논문≫의 두 번째 권에 더 자세한 계산과 함께 ≪접종의 수학적 이론≫이라는 제목의 다른 작품과 함께 출판되었다. 달랑베르는 천연두 감염 확률과 천연두로 사망할 확률이 나이와 무관하다는 베르누이의 가정을 비판했다. 그는 이러한 가정이 필요 없는 다른 해결책을 제안했다. 나이 $x$에서 천연두로 인한 사망률을 $v(x)$, 다른 원인으로 인한 사망률을 $m(x)$, 아직 생존한 사람의 수를 $P(x)$라고 부른다. 따라서,

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

이다. 방정식 (4.3)과 비교하면 실제로 $v(x) = pqS(x)/P(x)$ 임을 알 수 있다. 그리고고 다음을 얻는다.

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

여기서 $P^*(x)$는 천연두가 사라졌을 때 나이 $x$에 생존한 사람의 수를 나타낸다.

그림 4.3:
달랑베르 (1717–1783)

실제로 방정식 (4.6)와 (4.9) 사이에 함수 $m(x)$를 대입하거나 (4.8)의 공식을 $P^*(x)$에 사용하면 (4.9)의 해가 다음과 같이 주어지는 것을

확인할 수 있다.

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

달랑베르가 제시한 공식(4.10)은 베르누이의 공식(4.7)과 모순되지 않는다. 다만 사망자 명부에 사망 원인은 포함되지만 나이는 포함되지 않았기 때문에 당시에는 사용할 수 없었던 다른 유형의 정보 $v(x)$를 사용할 뿐이다. 달랑베르는 이러한 유형의 데이터가 나오기 전에는 예방 접종이 유용한지 여부를 결론 내릴 수 없다고 했다.

또한 달랑베르는 기대수명을 결정의 기준으로 사용하는 것이 유용하지 않다고 비판하였는데, 이는 기대수명이 가까운 미래든 먼 미래든 모든 연도에 동일한 가중치를 부여하기 때문이다. 그는 개인이나 국가의 관점에서 볼 때 모든 연도가 동일한 '유용성'을 갖는 것은 아니며, 젊은 나이와 노년기는 그 사이의 연령보다 가치가 낮다는 점에 주목했다. 이러한 모든 비판에도 불구하고 달랑베르는 예방 접종에 찬성한다고 선언했다.

출판 지연으로 인해 베르누이의 연구는 1766년에야 출판된 반면, 달랑베르는 자신의 연구를 매우 빠르게 출판할 수 있었다. 베르누이는 오일러에게 보낸 편지에서 씁쓸함을 표현했다.

> "확률에 대한 위대한 달랑베르의 엄청난 미사여구에 대해 어떻게 생각하십니까? 그의 출판물에서 내 성과가 너무 자주 부당하게 취급되는 것을 알고부터는 그의 펜에서 나온 어떤 것도 더 이상 읽지 않기로 결정한지가 좀 되었습니다. 제가 8년 전 파리 아카데미에 보낸 접종에 관한 논문이 결정의 계기가 되었습니다. 그 논문은 분석의 참신함 때문에 높이 평가되었습니다. 감히 말하자면 새로운 주제를 수학의 본체에 결합시키는 것과 같았습니다. 이 새로운 분석의 성공은 그에게 마음의 고통을 안겨준 것 같습니다. 그는 수천 가지 방법으로 모두 똑같이 우스꽝스럽게 비판했고, 그것들에 대한 비판을 잘 받은 후 그는 오직 언급되지 않았을 뿐인 이론의 첫 번째 저자인척 했습니다. 그는 내 원고가 7~8년 후에야 나올 수 있다는 것을 알고 있었습니다. 그는 오직 아카데미 회원의

> 자격을 통해 그 정보를 알 수 있었습니다. 이런 점에서 제 원고는 공개될 때까지 비밀로 유지되어야 했습니다. Dolus an virtus quis in hoste requirat?[3]"

베르누이와 달랑베르의 연구에도 불구하고 프랑스에서는 대규모 예방 접종이 실시되지 않았다. 루이 15세는 1774년 천연두로 사망했다. 궁정의 의사들은 얼마 지나지 않아 나머지 왕실 가족들에게도 접종을 실시했다. 에드워드 제너(Jenner)가 인간에게 우두를 접종하는 것('백신접종')이 천연두를 예방하고 안전하다는 사실을 발견하면서 이 문제는 더 이상 중요하지 않게 되었다. 그의 연구 ≪천연두 백신의 원인과 효과에 대한 탐구≫는 1798년에 발표되었다. 백신 접종은 유럽 전역으로 빠르게 확산되었다. 그럼에도 불구하고 한 가지 사망 원인을 제거할 경우 기대수명의 증가를 계산하기 위해 개발된 방법은 오늘날에도 여전히 사용되고 있다.

그 후 수십 년이 지나고 천연두로 사망하는 연령에 관한 데이터를 사용할 수 있게 되었다. 이 문제는 다음의 사람들에 의해 집중적으로 재검토되었다.

- 1772년 베를린 아카데미 출신의 수학자 요한 하인리히 람베르트 (Lambert)
- 당시 파리 내무부에서 인구 통계를 담당했던 엠마뉴엘-에티엔느 뒤빌라르 (Duvillard)의 ≪연령별 사망률에 대한 천연두의 영향 분석 및 표≫ (1806)
- 피에르-시몽 라플라스 (Laplace)의 ≪확률의 분석 이론≫ (1812)

일례로, 뒤빌라드와 라플라스는 모수 $p$와 $q$가 나이에 따라 달라지는 경우 공식(4.7)을 어떻게 수정할지 보였다.

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

[3] 적을 물리치는데 용맹에 의해서든 전략에 의해서든 무엇이 중요한가! 베르길리우스: ≪아이네이스 2≫.

여기서 $p(x)$는 나이 $x$에 감염되었을 때 천연두로 사망할 확률이고, $q(x)$는 나이 $x$에 천연두에 감염될 확률이다.

천연두에 대한 이 연구 이후, 다니엘 베르누이는 인구 역학의 다른 어떤 문제도 다루지 않았다. 그는 1782년 바젤에서 사망했다. 달랑베르는 1년 후 파리에서 사망했다.

# 참고문헌


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

5장

# 맬서스와 기하급수적 성장의 장애물 (1798)

1798년 맬서스는 ≪인구론≫에서 식량 공급이 기하급수적으로 증가하는 인구의 자연적 경향을 오랫동안 따라갈 수 없다고 주장했다. 인구가 비교적 일정하게 유지되는 것은 인류의 상당수가 식량 부족으로 고통받고 있기 때문이었다. 맬서스는 ≪인구론≫을 인간 사회의 진보를 강조한 고드윈과 콩도르세의 저술에 대한 반론으로 간주했다. 맬서스의 글은 다윈과 월리스의 진화론에 영향을 미쳤고 마르크스에게 비판을 받았지만 중국의 한자녀 정책으로 실천에 옮겨졌다.

토머스 로버트 맬서스(Malthus)는 1766년 런던 근교에서 7남매 중 여섯째로 태어났다. 장자크 루소(Rousseau)의 친구이자 추종자였던 아버지가 그의 첫 스승이었다. 1784년 맬서스는 케임브리지 대학에서 수학을 공부하기 시작했다. 1791년에 학위를 취득하고 1793년에 지저스 칼리지의 펠로우가 되었으며 1797년에는 영국 성공회의 사제가 되었다.

1798년 맬서스는 ≪인구의 원리에 관한 일론(一論), 그것이 장래의 사회개량에 미치는 영향을 G.W.고드윈·M.콩도르세 그리고 그 밖의 저작가들의 사색에 언급하며 논함≫이라는 제목의 책을 익명으로 출간했

그림 5.1: 맬서스 (1766–1834)

다. 이는 고드윈(Godwin)의 ≪정치적 정의와 그것이 일반 미덕과 행복에 미치는 영향에 관한 고찰≫(1793)과 콩도르세(Condorcet)의 ≪인간 정신의 진보에 관한 역사적 개요≫(1794)에 대한 반작용으로 나온 것이다. 프랑스 혁명이 진보라는 이름으로 행한 끔찍한 일에도 불구하고 두 저자는 사회의 진보는 불가피하다고 주장했다. 맬서스는 같은 낙관론을 갖지 않았다. 그는 또한 자녀가 많은 가난한 가정을 도왔던 영국 빈민법이 인구 성장만을 촉진시킬 뿐 이에 상응하는 식량 생산의 증가를 유도하지 못한다고 비판했다. 그에게는 이러한 법이 실제로 가난한 사람들을 구제하는 것이 아니라 오히려 그 반대인 것처럼 보였다. 통상 인구는 식량 생산보다 항상 빠르게 증가하는 경향이 있으며, 사회의 일부는 곤궁, 기아 또는 전염병으로부터 구제받지 못한 것처럼 보였다. 이들은 인구 증가를 늦추는 재앙이며 맬서스의 의견으로는 사회 발전의 주요 장애물이다. 진보를 약속하는 모든 이론은 유토피아적일 뿐이다. 이러한 생각을 바탕으로 맬서스는 1798년에 저서를 출간했다. 맬서스는 자신의 논제를 이렇게 요약했다.

> [...] "인구의 힘은 지구가 인간에게 필요한 자원을 생산할 수 있는 힘보다 무한히 크다. 인구의 자연적 증가는 기하급수적이지만 식량은 산술급수적으로 밖에 증가하지 않는다. 산수를 조금이라도 할 줄 아는 사람이라면 단순한 곱에 비해 거듭제곱이 얼마나 엄청난지 알 것이다. 생존에 음식이 필수인

> 인간 본성의 법칙에 따라 이 두 가지 불균형한 힘의 결과는 동일하게 유지되어야 한다. 이것은 생존의 어려움으로부터의 인구에 대한 강력하고 지속적인 견제를 의미한다. 이 어려움은 어딘가에는 반드시 닥치는 것이며, 인류의 상당수가 심각한 부담을 질 수 밖에 없다."

맬서스의 책은 매우 성공적이었다. 여기에는 데이터가 거의 포함되어 있지 않았다. 예를 들어 맬서스는 18세기에 미국 인구가 25년마다 두 배씩 증가했다는 사실을 발견했다. 맬서스는 자신의 논제를 수학적 모델로 번역하려 하지 않았지만, 다음 장의 주제가 될 아돌프 케틀레(Quetelet)와 피에르-프랑수아 베르홀스트(Verhulst)의 후대 연구를 위한 길을 닦아 주었다.

맬서스는 책을 출간한 후 친구들과 함께 독일, 스칸디나비아, 러시아를 거쳐 프랑스와 스위스로 여행을 떠났다. 여행 중에 수집한 정보를 종합하여 1803년에 ≪인구의 원리에 관한 에세이, 또는 인구가 인간의 행복에 미치는 과거와 현재의 영향에 대한 견해, 그리고 인구가 초래하는 악의 제거 또는 완화에 관한 우리의 전망에 대한 탐구≫라는 다른 부제를 붙인 매우 증보된 2판을 자신의 이름으로 출간했다. 이 새로운 판에서는 결혼 지연, 낙태, 영아 살해, 기근, 전쟁, 전염병, 경제적 요인 등 여러 국가에서 인구 증가의 장애물에 대해 자세히 논의했다. 맬서스에게 결혼을 늦추는 것은 인구 안정화를 위한 최선의 선택이었다. 이 책의 다른 판본은 1806년, 1807년, 1817년, 1826년에 네 차례 더 출간되었다. 1805년 맬서스는 서인도 회사가 직원들을 위해 설립한 새 학교의 역사 및 정치경제학 교수가 되었다. 그는 또한 ≪지대의 성질과 추이에 대한 탐구≫(1815)와 ≪정치경제학의 원리≫(1820)를 출간했다. 1819년 맬서스는 왕립학회 회원으로 선출되었다. 1834년에는 통계학회의 창립 멤버 중 한 명이 되었다. 그는 같은 해 바스 근처에서 사망했다.

맬서스의 연구는 진화론의 발전에 큰 영향을 미쳤다. 비글호 여행을 마치고 돌아온 찰스 다윈(Darwin)은 1838년 맬서스의 인구에 관한 책을 읽었다. 1859년에 출간된 그의 유명한 저서 ≪종의 기원에 대하여≫의 서문에서 그가 쓴 글은 다음과 같다.

> "다음 장에서는 세계의 모든 생물이 높은 기하급수적 비율로 증식하는 결과 필연적으로 일어나는 생존을 위한 투쟁에 대해 다루려 한다. 이는 맬서스의 원리를 모든 동식물계에 적용한 것이다."

다윈과 동시에 진화론을 발전시킨 앨프리드 러셀 월리스(Wallace)도 맬서스의 책을 읽은 후 그의 아이디어가 나왔다고 말했다.

이와는 대조적으로 카를 마르크스(Marx)는 ≪자본론≫의 각주에서 맬서스의 책의 성공에 대한 다른 관점을 보여준다.

> "만약 독자가 1798 년에 출판된 맬서스의 ≪인구론 에세이≫를 떠올린다면, 저는 이 작품의 초판이 디포, 제임스 스튜어트경, 타운센드, 프랭클린, 월리스에 등의 작품에서 발췌한 일종의 초등학교 수준의 표절이며 자신이 생각한 문장이 하나도 포함되어 있지 않다는 것을 상기시키겠다. 이 에세이가 큰 센세이션을 일으킨 것은 단지 정당의 이익 때문이었다. 프랑스 혁명은 영국에서 열정적인 지지자를 얻었고, 18세기에 천천히 발전된 ≪인구론≫은 큰 사회적 위기의 한가운데 드럼과 트럼펫 소리와 함께 콩도르세 등의 교훈에 대한 절대적인 해독제로 선언되었으며, 이는 영국 귀족들에게 인간의 발전에 대한 갈망의 대폭파자로 환영받았다. 예상치 못한 성공에 놀란 맬서스는 자신의 책에 미처 발견하지 못한 새로운 소재를 추가하면서 겉보기만으로 조합한 자료를 채우기 시작했다"

물론 맬서스의 논제가 완전히 새로운 것은 아니었다. 예를 들어, 인구가 기하급수적으로 증가하는 경향이 있다는 것은 3장에서 봤듯 오일러가 반세기 전에 이미 했던 생각이지만 맬서스[1]의 아이디어로 자주 언급된다. 그러나 맬서스는 이 아이디어를 실제 입법 문제와 논쟁적인 방식으로 연결하여 대중에게 알렸다. 아이러니하게도 맬서스의 출산 제한 제안이 가장 눈에 띄게 적용될 곳은 공산당 집권 하의 중국이었다 (25장 참조).

[1] 피셔(14장 및 20장 참조)는 인구 증가율을 '맬서스 인구 증가율' 이라고 불렀다. 맬서스의 저서에서 그는 쥐스밀히의 연구를 언급했다.

# 참고문헌

1. Condorcet: Esquisse d'un tableau historique des progrès de l'esprit humain. Agasse, Paris (1794). gallica.bnf.fr
2. Darwin, C.: On the Origin of Species by Means of Natural Selection. John Murray, London (1859). darwin-online.org.uk
3. Godwin, W.: An Enquiry Concerning Political Justice. Robinson, London (1793). archive.org
4. Malthus, T.R.: An Essay on the Principle of Population, 1st edn. London (1798). econlib.org
5. Marx, K.: Das Kapital, Erster Band. In : Marx, K., Engels, F., Werke, Band 23, Dietz Verlag, Berlin (1962) kulturkritik.net
6. Simpkins, D.M.: Malthus, Thomas Robert. In: Gillespie, C.C. (ed.) Dictionary of Scientific Biography, vol. 9, 67–71. Scribner, New York (1974)

6장

# 베르휼스트와 로지스틱 방정식 (1838)

1838년 벨기에의 수학자 베르휼스트는 기하급수적 성장 방정식을 일반화하되 인구의 최대치를 고려한 로지스틱 방정식을 도입했다. 그는 여러 국가, 특히 벨기에의 데이터를 사용하여 미지의 매개변수를 추정했다. 베르휼스트의 연구는 1920년대에 이르러서야 재발견되었다.

피에르-프랑수아 베르휼스트(Verhulst)는 1804년 브뤼셀에서 태어났고, 1825년 겐트대학교에서 수학 박사 학위를 취득했다. 그는 정치에도 관심이 많았다. 결핵 치료를 위해 이탈리아에 머무는 동안 그는 교황청 헌법에 찬성하는 탄원서를 제출했지만 성공하지 못했다. 1830년 혁명과 벨기에 독립 이후에는 18세기 애국자에 관한 역사 에세이를 발표했다. 1835년에는 브뤼셀에 새로 설립된 자유 대학교의 수학 교수가 되었다.

같은 해인 1835년, 통계학자이자 브뤼셀 천문대 소장인 동포 아돌프 케틀레(Quetelet)는 ≪인간과 인간의 능력 개발에 관한 논의≫를 발표했다. 케틀레는 맬서스(Malthus)가 언급한 장애물이 일종의 '저항'을 형성하기 때문에 인구가 오랜 기간 동안 기하급수적으로 증가할 수 없다고 주장했고, 그 저항이 유체역학에서처럼 인구 증가 속도의 제곱에 비례한다고 생각했다. 이 비유는 실제 근거는 없었지만 베르휼스트에게 영감을 주었다.

그림 6.1:
베르휼스트 (1804–1849)

실제로 베르휼스트는 1838년에 ≪인구 증가 법칙에 관한 논평≫를 발표했다. 다음은 몇 가지 발췌문이다.

> "우리는 유명한 맬서스가 인구가 기하급수적으로 증가하여 일정 기간, 예를 들면 25년마다 두 배로 증가하는 경향이 있다는 원리를 보여 주었다는 것을 알고 있다. 이 명제는 식량을 구하기가 점점 더 어려워지는 것을 고려할 때 논쟁의 여지가 없다 [...].
>
> 따라서 인구의 실질적인 증가는 국가의 규모와 출산율에 의해 제한된다. 결과적으로 인구는 점점 더 안정 상태에 가까워진다."

베르휼스트는 케틀레의 유체역학에 대한 비유가 합리적이지 않다는 것을 깨닫고는 대신 시간 $t$에서의 인구 $P(t)$에 대한 다음과 같은 (여전히 다소 자의적인) 미분 방정식을 제안한 것으로 보인다.

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

인구 $P(t)$가 매개변수 $K$에 비해 작으면 방정식을 다음과 같이 근사할 수 있다.

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

위 방정식의 해는 $P(t) \approx P(0)e^{rt}$ 로 지수함수적으로 증가한다.[1] $P(t)$ 가 $K$ 에 가까워질수록 성장률은 감소한다. 심지어 $P(t)$ 가 $K$ 를 초과할 경우 성장률은 음수가 될 수도 있다. 방정식 (6.1)의 해의 수식을 구하려면 다니엘 베르누이(Daniel Bernoulli)가 방정식 (4.5)을 풀었던 것과 같은 방식으로 진행하면 된다.

> 방정식(6.1)을 $P^2$ 로 나누고 $p = 1/P$ 로 두면 $dp/dt = -rp + r/K$ 를 구할 수 있다. $q = p - 1/K$ 를 사용하면 $dq/dt = -rq$ 와 $q(t) = q(0)e^{-rt} = (1/P(0) - 1/K)e^{-rt}$ 를 얻는다. 따라서 $p(t)$ 와 $P(t)$ 를 구할 수 있다.

수식을 약간 정리하면 마침내 다음과 같은 공식을 얻는다.

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

총 인구는 $t = 0$ 시점의 $P(0)$ 에서 $t \to +\infty$ 시점의 극한값 $K$ 까지 점진적으로 증가한다 (그림 6.2). 미지의 매개 변수 $r$ 및 $K$ 에 사용한 값을 밝히지 않은 채 베르휼스트는 1817년에서 1831년 사이의 프랑스, 1815년에서 1833년 사이의 벨기에, 1811년에서 1831년 사이의 영국 에식스주, 1796년에서 1827년 사이의 러시아 인구에 관한 데이터와 자신의 결과를 비교했다. 그의 모델은 데이터의 추세와 꽤 잘 맞았다.

1840년 베르휼스트는 브뤼셀 왕립 군사학교의 교수가 되었다. 이듬해에는 ≪타원 함수의 기초 논문≫을 발표하고 벨기에 왕립 아카데미에 선출되었다. 1845년에는 ≪인구 증가 법칙에 관한 수학적 탐구≫라는 제목의 논문으로 인구 연구를 계속했다. 그는 미국의 인구가 25년마다 두배씩 증가해 왔다는 맬서스의 발언을 회상했다(표 6.1). 연도 $n + 10$ 의 인구와 연도 $n$ 의 인구의 비율을 계산하면 각각 1.350, 1.364, 1.331, 1.335 및 1.326으로 상당히 일정하다는 것을 알 수 있다. 따라서 인구는 평균적으로 10년마다 1.34씩, 25년마다 $1.34^{25/10} \approx 2.08$ 씩 곱해졌다. 따라서 거의 반세기 전 맬서스의 논문 이후 25년마다 두 배씩 계속 증가해

[1] 일반적으로 불연속 시간 모델에서는 기하급수적 성장, 연속 시간 모델에서는 지수함수적 성장이라고 하지만 본질적으로 같은 개념이다.

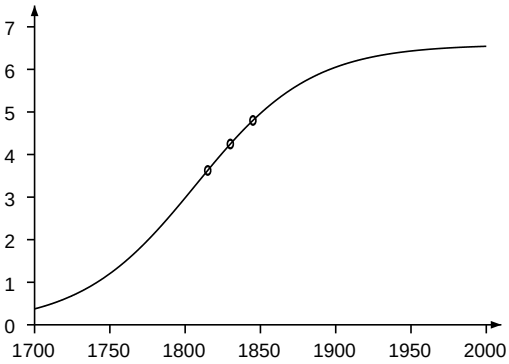


그림 6.2: 벨기에 인구(백만 명 단위)와 로지스틱 곡선. 세 개의 점은 1815년, 1830년, 1845년에 해당한다. 매개변수 값은 1845년 기록에 따른 값이다.

표 6.1: 미국의 공식 인구 조사 표.

| 연도 | 인구 규모 |
|---|---|
| 1790 | 3,929,827 |
| 1800 | 5,305,925 |
| 1810 | 7,239,814 |
| 1820 | 9,638,131 |
| 1830 | 12,866,020 |
| 1840 | 17,062,566 |

온 것이다. 그러나 베르휼스트는 다음과 같은 말을 덧붙였다.

> "기하급수적 인구 증가 가설은 특수한 상황에서만 성립하는 것으로 일반적으로 이를 만족할 것으로 주장할 수는 없다. 예를 들어 이 가설은 최초의 미국 식민지의 경우처럼 거의 무한한 크기의 비옥한 영토에 선진 문명을 가진 사람들이 거주하게 되는 경우처럼 매우 특수한 상황에만 성립하게 된다."

그의 논문에서 베르휼스트는 '로지스틱'이라고 부르는 방정식(6.1)을 다시 언급했다. 그는 곡선 $P(t)$가 $P(t) < K/2$까지는 양의 곡률(아래로 볼록)로 $K$를 향해 계속 증가하지만 $P(t) > K/2$인 순간 음의 곡률(위로 볼록)로 증가한다는 것을 발견했다. 따라서 곡선은 뒤틀린 문자 S의 모양을 갖는다 (그림 6.2).

실제로 $d^2P/dt^2 = r(1-2P/K)\,dP/dt$이다. 따라서 $P < K/2$인 경우 $d^2P/dt^2 > 0$이고, $P > K/2$인 경우 $d^2P/dt^2 < 0$이다.

또한 베르휼스트는 서로 다르지만 동일한 간격의 세 해의 인구 $P(t)$로부터 매개 변수 $r$과 $K$를 추정할 수 있는 방법을 설명했다. 만약 $P_0$이 $t = 0$ 시점의 인구, $P_1$이 $t = T$ 시점의 인구, $P_2$가 $t = 2T$ 시점의 인구라면, 방정식 (6.2)에서 시작하는 지루한 계산을 통해 다음이 성립함을 알 수 있다.

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

1815년, 1830년, 1845년의 벨기에 인구 추정치(각각 362만 7천명, 424만 7천명, 480만 천명)를 사용하여 $K = 6.584$백만, 과 연간 $r = 2.62\%$를 얻었다. 그런 다음 방정식(6.2)을 사용하여 벨기에의 인구가 1851년 초에는 499만 8천명, 1900년 초에는 606만 4천명이 될 것이라고 예측할 수 있었다 (그림 6.2). 베르휼스트는 프랑스에 대해서도 비슷한 연구를 수행했다. 그는 $K = 39.685$백만, 연간 $r = 3.2\%$를 얻었다. 그 사이 벨기에와 프랑스의 인구가 이 값인 $K$를 크게 초과했기 때문에, 로지스틱 방정식은 1838년의 논문에서와 같이 수십 년의 기간 동안만 현실적인 모델이 될

수 있을 뿐 그 이상 기간 동안은 현실적이지 않다는 것을 알 수 있다.

1847년에 ≪인구 증가 법칙에 대한 두 번째 탐구≫가 나왔고, 여기서 베르휼스트는 로지스틱 방정식 대신 다음과 같은 형식으로 쓸 수 있는 미분 방정식을 선택했다.

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

그는 이 방정식이 인구 $P(t)$가 특정 임계값 이상일 때 성립할 것이라고 생각했다. 방정식의 해는 다음과 같다.

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

벨기에에 대한 동일한 인구 통계 데이터를 사용하여 베르휼스트는 매개변수 $r$과 $K$를 새롭게 추정했다. 이번에는 최대 인구로 $K = 9.4$백만을 찾았다. 이는 모델 선택에 따라 결과가 얼마나 달라질 수 있는지를 보여준다!

베르휼스트는 1848년 벨기에 왕립 아카데미의 회장이 되었지만 이듬해 브뤼셀에서 결핵으로 사망했다. 베르휼스트는 여러 모델 방정식 사이에서 망설였지만, 로지스틱 방정식은 수십 년 후 다른 사람들에 의해 독립적으로 개발 되었다. 로버트슨(Robertson)은 1908년에 동물, 식물, 인간 및 신체 기관의 개별 성장을 모델링하는 데 이 방정식을 사용했다. 맥켄드릭(McKendrick)과 케사바 파이(Pai)는 1911년에 미생물 개체군의 성장에 사용했고, 펄(Pearl)과 리드(Reed)는 1920년 성장세가 둔화되기 시작한 미국 인구의 증가에 이 방정식을 사용했다. 1922년 펄은 마침내 베르휼스트의 연구에 주목했다. 그 이후로 로지스틱 방정식은 많은 연구에 영감을 주었다(13, 20, 24 장 참조). 최대 인구 $K$는 이후 '수용력'으로 불리게 되었다.

# 참고문헌


1. Lloyd, P.J.: American, German and British antecedents to Pearl and Reed's logistic curve. Pop. Stud. 21, 99–108 (1967)
2. McKendrick, A.G., Kesava Pai, M.: The rate of multiplication of micro-organisms: A mathematical study. Proc. R. Soc. Edinb. 31, 649–655 (1911)
3. Pearl, R.: The Biology of Death. Lippincott, Philadelphia (1922). archive.org
4. Pearl, R., Reed, L.J.: On the rate of growth of the population of the United States since 1790 and its mathematical representation. Proc. Natl. Acad. Sci. 6, 275–288 (1920). pnas.org
5. Quetelet, A.: Sur l'homme et le développement de ses facultés. Bachelier, Paris (1835). gallica.bnf.fr
6. Quetelet, A.: Pierre-François Verhulst. Annu. Acad. R. Sci. Lett. B.-Arts Belg. 16, 97–124 (1850). archive.org
7. Quetelet, A.: Sciences mathématiques et physiques au commencement du XIX$^{e}$ siècle. Mucquardt, Bruxelles (1867). gallica.bnf.fr
8. Robertson, T.B.: On the normal rate of growth of an individual and its biochemical significance. Arch. Entwicklungsmechanik Org. 25, 581–614 (1908)
9. Verhulst, P.-F.: Notice sur la loi que la population poursuit dans son accroissement. Corresp. Math. Phys. 10, 113–121 (1838). archive.org
10. Verhulst, P.-F.: Recherches mathématiques sur la loi d'accroissement de la population. Nouv. Mém. Acad. R. Sci. B.-lett. Brux. 18, 1–45 (1845). uni-goettingen.de
11. Verhulst, P.-F.: Deuxième mémoire sur la loi d'accroissement de la population. Mém. Acad. R. Sci. Lett. B.-Arts Belg. 20 (1847). archive.org

7장

# 비에나메, 쿠르노, 그리고 가족 성의 멸종 (1845–1847)

프랑스의 통계학자 비에나메는 1845년 각 남성이 주어진 확률 분포에 따라 아들 수를 가질 경우 성씨가 소멸될 확률을 계산하는 방법을 알아냈다. 아들 수의 평균이 1보다 작거나 같으면 성이 멸종된다. 평균이 1보다 크면 멸종 확률은 엄밀히 말해 1보다 작다. 그의 결과에 대한 증거는 2년 후 그의 친구 쿠르노가 쓴 책에서 발표되었다. 이 업적은 최근에야 재발견되었다.

이레네 쥘 비에나메(Bienaymé)는 1796년 파리에서 태어났다. 에콜 폴리테크니크에서 공부한 후 재무부에서 경력을 쌓아 고위 감사관의 자리에 올랐다. 라플라스(Laplace)가 쓴 ≪확률의 분석 이론≫이라는 책에 영향을 받아 인구 통계 및 의학 통계(영아 사망률, 출생 수, 기대 수명), 사법 오류의 확률, 보험 이론, 투표 시스템의 대표성 등 확률 이론의 다양한 응용 분야에 대한 논문을 발표하기도 했다.

1845년 비에나메는 ≪곱의 법칙과 가족의 존속기간에 관하여≫라는 짧은 논문을 썼는데, 이 글은 파리의 철학 학회 회보에 게재되었다. 이미 많은 저자들이 이 주제에 대해 글을 썼다. 맬서스(Malthus)는 ≪인구론≫(1803)의 두 번째 판에서 스위스의 인구에 관한 장을 포함시켰고, 이에

그림 7.1: 비에나메 (1796–1878) 과 쿠르노 (1801-1877).

주목했다.

> “베른 마을에서 1583년부터 1654년까지 주권 의회는 부르주아지에 487개의 가문을 인정했는데, 이 중 379개 가문이 2세기 동안 멸종했고 1783년에는 108개 가문만이 남았다.”

1842년 토마스 더블데이(Doubleday)는 더 나아가 귀족이나 부르주아 출신의 상류층 가문이 하류층 가문보다 사라지는 경향이 더 크다고 주장했다. 프랑스에서도 1844년 에밀 리트르(Littré)가 오귀스트 콩트(Comte)의 실증주의 철학을 소개하는 글에서 비슷한 아이디어를 제시했고, 1845년 비에나메의 친구인 베누이스탕 드 샤토네프(Châteauneuf)가 ≪프랑스 귀족 가문의 존속 기간에 관하여≫라는 에세이를 발표했다.

이러한 맥락에서 비에나메는 한 국가의 인구가 기하급수적으로 증가하는 경향이 있는 반면 많은 수의 가문이 사라지는 이유를 설명하려고 했다. 이 문제를 해결하기 위해 그는 모든 남성이 동일하게 0, 1, 2, 3, ...명의 성인이 될 아들을 가질 확률을 갖는 단순한 경우를 고려했다. 더 정확하게는 한 남성이 $n$ 세대 후에 자신의 성을 가진 자손을 낳을 확률은 얼마인가 하는 질문을 생각했다. 평균 아들 수가 1보다 작다면, 이 확률은 $n$이 무한대로 증가함에 따라 0이 되는 경향이 있다는 것이 분명하다. 비에나메는 아들이 없을 확률이 1/2이고 아들이 둘일 확률이 1/2인 경우와 같이 아들의 평균 수가 정확히 한 명인 경우에도 같은 결론이 유지된다는

것을 알아냈다(그림 7.2).[1] 그러나 이 경우 $n$ 세대에 자손이 있을 확률은 더 천천히 0이 되는 경향이 있다. 이 예에서는 35세대가 지난 후에도 5% 확률이다. 즉, 1세기당 3세대가 있는 경우 11세기 또는 12세기가 지난 후에도 자손이 있을 확률은 5% 이다.[2] 비에나메는 마침내 평균 아들 수가 한 명보다 많으면 가계의 멸종이 확실하지 않으며 그 확률은 대수 방정식을 풀어서 계산할 수 있다는 것을 발견했다.

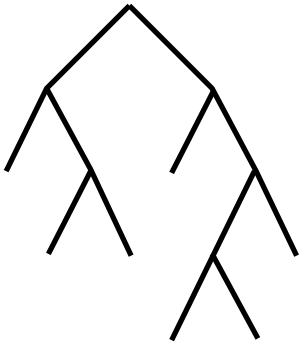

그림 7.2: 인위적으로 만든 가계도의 예. 조상이 나무의 맨 위에 있다. 각 세대에서 남성에게 아들이 없을 확률이 1/2, 아들이 둘일 확률이 1/2이다.

비에나메의 글에는 더 자세한 설명이 포함되어 있지 않았다. 1847년 그의 친구이자 수학자이자 경제학자인 앙투안 오귀스탱 쿠르노(Cournot)는 ≪대수와 기하학 사이의 대응의 기원과 한계에 관하여≫라는 제목의 책에 몇 가지 세부 사항을 포함시켰다. 그는 이 문제를 확률 게임의 형태로 제시했지만, 이 문제가 성의 소멸에 관한 비에나메의 연구와 동일하다는 점을 인정했다. 가족 성의 관점에서 해석을 유지한다면, 쿠르노는 먼저 남성이 최대 두 명의 아들을 가진 특수한 경우를 고려했는데, $p_0, p_1$, $p_2$는 각각 아들을 0, 1, 2명 가질 확률이다. 물론 $p_0 + p_1 + p_2 = 1$이다. 한 조상으로부터 시작했을 때, 단 한 세대 만에 멸종할 확률, 즉 $x_1$은 당연히 $p_0$과 같다. 2세대 내에 멸종할 확률은 $x_2 = p_0 + p_1 x_1 + p_2 {x_1}^2$이다. 즉, 1세대에 이미 멸종했거나(확률 $p_0$), 1세대에 아들 한 명만 있었고 더

[1] 모든 남성이 정확히 한 명의 아들을 둔 경우는 예외이다.

[2] 아래에서 다루겠지만 이 확률은 $1 - x_{35}$와 같으며, 여기서 $x_{n+1} = \frac{1}{2} + \frac{1}{2} {x_n}^2, x_0 = 0$이다.

이상 아들인 자손이 없거나(확률 $p_1 x_1$), 1세대에 아들 둘이 있었는데 각 아들마다 아들 자손이 없는 경우(확률 $p_2 x_1^2$)이다. 보다 일반적으로, $n$ 세대 내에 멸종할 확률은 $x_n = p_0 + p_1 x_{n-1} + p_2 (x_{n-1})^2$ 이다. 예를 들어 첫 번째 세대에 두 아들이 있다면(확률 $p_2$), 그 가족은 $n-1$ 세대 후(즉, $n$ 세대에) $(x_{n-1})^2$ 와 같은 확률로 멸종할 것이다. 쿠르노는 $x_n$ 이 모든 $n$ 에 대해 $x_n \leqslant 1$ 인 증가 수열이라는 것을 알아냈다. 따라서 $x_n$ 은 $x_\infty \leqslant 1$ 의 극한을 가지며, 이는 $x = p_0 + p_1 x + p_2 x^2$ 방정식의 해이다. $p_1 = 1 - p_0 - p_2$ 를 사용하면 이 방정식은 $0 = p_2(x-1)(x - p_0/p_2)$ 와 동일하다. 따라서 $x = 1$ 과 $x = p_0/p_2$ 의 두 가지 근이 있다. 아들의 평균 수 $p_1 + 2p_2$ 에 따라 세 가지 경우를 구분할 수 있는데, 이는 $1 - p_0 + p_2$ 와 같으며 이를 $\mathscr{R}_0$ 이라고 부른다. 만약 $\mathscr{R}_0 < 1$ 이면 $p_0/p_2 > 1$ 이 된다. 따라서 $x = 1$ 은 극한 $x_\infty$ 이 될 수 있는 유일한 값이다. 확실히 가문은 멸종할 것이다. 만약 $\mathscr{R}_0 = 1$ 이라면 두 근은 모두 1이고 결론은 동일하다. 만약 $\mathscr{R}_0 > 1$ 이라면, 쿠르노는 $p_0 = 0$ 이라는 특수한 경우에서 소멸 확률은 분명히 0이어야 하기 때문에 $x_\infty$ 는 두 번째 근 $p_0/p_2$ 와 같아야 한다고 주장했다.

쿠르노는 남성이 $p_0$, $p_1$, ..., $p_m$ 의 확률로 최대 $m$ 의 아들을 가질 수 있는 보다 일반적인 경우를 간략하게 언급했다. 결론은 같은 방식으로 아들의 평균 수 $\mathscr{R}_0 = p_1 + 2p_2 + \cdots + mp_m$ 의 1에 대한 크기에 따라 달라진다. $x_\infty$ 에 대한 방정식 $x = p_0 + p_1 x + \cdots + p_m x^m$ 은 항상 근 $x = 1$ 을 갖는다. 다른 양의 근은 단 하나뿐이며, 이 근은 $\mathscr{R}_0 > 1$ 일 때 소멸 확률 $x_\infty$ 을 나타낸다.

안타깝게도 비에나메의 논문과 쿠르노의 책에 실린 몇 페이지는 당시에는 전혀 주목받지 못했다. 이 논문은 1970년대에야, 책은 20년이 더 지난 후에야 주목받았다! 그 사이 이 문제와 해결책은 다른 사람들에 의해 재발견되었고 이 주제는 상당히 발전했다. 이는 9장, 17장, 18장에서 다시 다룰 예정이다.

1848년 혁명 이후 비에나메는 재무부에서 일하던 직장을 그만둬야 했다. 그가 가장 적임자였던 파리 대학의 확률론 교수 자리도 다른 사람에게 돌아갔다. 그럼에도 불구하고 비에나메는 1850년 이후 재무부에서 다시 일할 수 있었지만 1852년 사임했다. 그해 말, 그는 과학 아카데미에

선출되어 통계 분야의 전문가로 활동했다. 1853년 그는 현대 교과서에서 '비에나메-체비쇼프' 부등식이라고 부르는 것을 증명했다. 1875년 그는 새로 창설된 프랑스 수학회의 회장이 되었다. 그는 1878년 파리에서 사망했다.

# 참고문헌


1. Bienaymé, I.J.: De la loi de multiplication et de la durée des familles. Extr. p. v. séances – Soc. Philomat. Paris, 37–39 (1845) biodiversitylibrary.org
2. Bru, B.: À la recherche de la démonstration perdue de Bienaymé. Math. Sci. Hum. 114, 5–17 (1991). archive.numdam.org
3. Bru, B., Jongmans, F., Seneta, E.: I.J. Bienaymé: Family information and proof of the criticality theorem. Int. Stat. Rev. 60, 177–183 (1992)
4. Brun, J., Robinet, A. (éd.): A. Cournot, études pour le centenaire de sa mort. Economica / Vrin, Paris (1978)
5. Cournot, A.-A.: De l'origine et des limites de la correspondance entre l'algèbre et la géométrie. Hachette, Paris (1847). archive.org
6. Doubleday, T.: The True Law of Population (1842). archive.org
7. Heyde, C.C., Seneta, E.: I.J. Bienaymé. Springer (1977)
8. Kendall, D.G.: The genealogy of genealogy: branching processes before (and after) 1873. Bull. Lond. Math. Soc. 7, 225–253 (1975)
9. Littré, É.: Conservation, révolution et positivisme (1852). gallica.bnf.fr
10. Malthus, T.R.: An Essay on the Principle of Population (1803). archive.org
11. Martin, T.: Antoine Augustin Cournot. In: Heyde, C.C., Seneta, E. (eds.) Statisticians of the Centuries, 152–156. Springer (2001)
12. Seneta, E.: Irenée-Jules Bienaymé. In: ibid., 132–136.

8장

# 멘델과 유전형질 (1865)

1865년 멘델은 완두콩의 교배에 관한 선구적인 실험 결과를 발표했다. 그의 분석은 기본적인 확률 이론을 사용했다. 그는 또한 자가수정 식물 집단에 대한 동역학 모델을 고안했다. 1900년에야 재발견된 그의 연구는 유전학의 역사에서 이정표가 되었다.

요한 멘델(Mendel)은 1822년 당시 오스트리아 제국의 일부이자 현재 체코의 일부인 모라바에서 태어났다. 그의 아버지는 소작농이었다. 고등학교 성적이 좋았고 건강이 좋지 않았던 멘델은 가족 농장에서 일하기보다는 계속 공부하는 것을 선호했다. 그러나 그는 대학에 갈 형편이 되지 못했다. 그래서 1843년 브륀(현재 브르노)에 있는 성 토마스 수도원에 들어가 그레고르라는 이름을 얻었다. 그는 신학을 공부했지만 농업에 관한 몇 가지 수업도 들었다. 1847년 그는 사제 서품을 받았다. 그는 몇 년 동안 고등학교에서 가르쳤지만 일반 교수가 되기 위한 시험에 떨어졌다. 1851년에서 1853년 사이에 그는 성직자단의 지원 덕분에 비엔나 대학교에서 물리학, 수학 및 자연 과학 과정을 수강하면서 학업을 계속할 수 있었다. 그 후 브륀으로 돌아와 기술학교에서 물리학을 가르쳤다.

1856년부터 1863년까지 멘델은 수도원 정원에서 수많은 식물에 대한 일련의 실험을 수행했다. 1865년 그는 자신이 회원으로 있던 브륀의 자연사학회 회의에서 두 차례에 걸쳐 자신의 연구 결과를 발표했다. 그의

그림 8.1:
멘델 (1822–1884)

연구 결과인 ≪식물의 잡종에 관한 실험≫은 이듬해 독일어로 학회 회보에 게재되었다. 멘델은 자가 수정으로 자연 번식하며 씨앗이 둥글거나 주름진, 노란색 또는 녹색 등 쉽게 식별할 수 있는 다양한 형태를 띠는 식물인 완두콩의 변이를 연구하게 된 배경을 설명했다. 그는 순종의 둥근 씨앗을 가진 식물과 순종의 주름진 씨앗을 가진 식물을 교배하면 항상 둥근 씨앗을 가진 잡종을 얻는다는 사실을 발견했다. 그는 '둥근 씨앗'의 특성을 우성, '주름진 씨앗' 특성을 열성이라고 불렀다. 그는 같은 방식으로 '노란 씨앗' 특성이 우성이고 '녹색 씨앗' 특성이 열성이라는 것을 보여주었다.

그 후 멘델은 잡종 종자에서 자란 식물의 자가 수정이 우성 또는 열성 형질을 가진 1세대 새 종자를 무작위 비율로 생성한다는 사실을 발견했다. 또한, 그는 실험을 여러 번 반복함으로써 열성 특성보다 우성 특성을 가진 씨앗을 평균적으로 약 3배 더 많이 얻었다는 사실을 발견했다. 예를 들어, 그는 첫 번째 실험에서 총 5,474 개의 둥근 씨앗과 1,850 개의 주름진 씨앗을 얻었으며 이는 2.96 대 1의 비율에 해당한다. 두 번째 실험에서는 총 6,022개의 노란색 씨앗과 2,001개의 녹색 씨앗이 생성되어 3.01대 1의 비율을 보였다.[1]

[1] 이후 피셔(R. A. Fisher)가 보였듯이(14장 참조), 실험 결과가 이론적 값에 그토록 근접할 확률은 상당히 작다. 멘델은 아마도 자신의 데이터를 정리했을 것이다. 예를 들어, $n = 6{,}022 + 2{,}001 = 8{,}023$개의 씨앗에 관한 두 번째 실험에서 비율이 3과 0.01 미만으로 다를 확률은 약 10% 불과하다.

멘델은 또한 우성 형질을 가진 1세대의 씨앗에서 자란 식물 중에서 자가 수정을 통해 우성 또는 열성 형질을 가진 씨앗을 생성한 식물의 수가 우성 형질만 가진 씨앗을 생성한 식물 수의 약 2배에 달한다는 사실을 알아냈다. 예를 들어, 1세대의 둥근 씨앗에서 자란 565개의 식물 중 372개는 둥근 씨앗과 주름진 씨앗을 모두 생성한 반면, 193개는 둥근 씨앗만 생성했다. 즉, 비율은 1.93이다. 마찬가지로 1세대의 노란색 씨앗에서 자란 519개 식물 중 353개는 노란색과 녹색 씨앗을 모두 생성한 반면 166개는 노란색 씨앗만 생성했으며, 그 비율은 2.13이다.

이러한 결과를 설명하기 위해 멘델은 씨앗의 겉으로 드러나는 특성을 두 개의 숨겨진 요소가 연관된 결과로 간주하는 기발한 아이디어를 냈는데, 각 요소는 우성($A$로 표기) 또는 열성($a$로 표기)이다. 따라서 가능한 조합은 $AA$, $Aa$ 및 $aa$의 세 가지다. $AA$ 또는 $Aa$를 가진 씨앗은 동일한 우성의 형질 $A$를 갖는다. $aa$를 가진 씨앗은 열성 형질질 $a$를 갖는는다. 또한 멘델은 수정 과정에서 꽃가루 입자와 밑씨(생식세포)가 두 가지 요소 중 하나를 전달하며, 각 요소는 1/2의 확률로 전달된다고 가정했다.

따라서 순수 계통의 $AA$와 $aa$를 교배하면 모두 $Aa$로 우성 형질 $A$를 갖는 잡종이 생성된다. 잡종 $Aa$의 생식세포는 1/2 확률로 요소 $A$를 전달하고 1/2 확률로 $a$를 전달한다. 따라서 잡종 종자 $Aa$에서 자란 식물의 자가 수정은 표 8.1에 표시된 것처럼 1/4 확률로 $AA$, 1/2 확률로 $Aa$, 1/4 확률로 $aa$를 생성한다.

표 8.1: 잡종 $Aa$를 자가 수정 했을 때 가능한 결과와 각 결과에 대해 암 생식세포(행)와 수 생식세포 (열)에 의해 전달되는 요소에 관한 함수로서의 확률.

| 요소 | $A$ | $a$ |
|---|---|---|
| 확률 | 1/2 | 1/2 |
| $A$ | $AA$ | $Aa$ |
| 1/2 | 1/4 | 1/4 |
| $a$ | $Aa$ | $aa$ |
| 1/2 | 1/4 | 1/4 |

멘델은 $1:2:1$인 $AA:Aa:aa$의 비율이 $(A+a)^2 = AA + 2Aa + aa$의 계산으로도 구할 수 있다는 것을 알아냈다. 씨앗 $AA$와 $Aa$는 형질 $A$를 가지고 있고 씨앗 $aa$만 형질 $a$를 가지고 있기 때문에 실제로 $A$ 형질을 가진

씨앗이 $a$ 형질을 가진 씨앗보다 세 배 더 많다. 또한, $AA$ 보다 평균적으로 두 배 더 많은 $Aa$ 씨앗이 있다. 전자에서 자란 식물의 자가 수정은 우성 ($AA$ 또는 $Aa$) 또는 열성 ($aa$)의 특성을 가진 씨앗을 생성한다. 씨앗 $AA$ 에서 자란 식물의 자가 수정의 경우, 항상 우성 특성을 가진 씨앗 $AA$ 을 갖게 된다. 이렇게 모든 관찰이 설명된다.

멘델은 다음 세대에 대해서도 살펴봤다. $N$ 개의 잡종 씨앗 $Aa$ 에서 시작해 각 식물이 자가 수정으로 4개의 새로운 씨앗만 생성한다는 가정을 통해 그는 $n$ 세대의 평균 씨앗 수 $(AA)_n$, $(Aa)_n$ 및 $(aa)_n$ 을 표 8.2과 같이 계산했으며, 명확성을 위해 결과를 $N$ 으로 나눴다.

표 8.2: 잇따른 세대.

| $n$ | 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 |
| 총 개수 | 1 | 4 | 16 | 64 | 256 | 1,024 |

위의 결과는 다음의 공식에서 간단히 도출된다.

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

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

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

이 공식은 $AA$ 가 자가 수정 후 4개의 씨앗 $AA$ 을 생성하고, $aa$ 는 4개의 씨앗 $aa$ 을 생성하며, $Aa$ 는 평균적으로 1개의 씨앗 $AA$, 2개의 씨앗 $Aa$, 1개의 씨앗 $aa$ 를 생성한다는 뜻이다. 멘델은 또한 다음과 같은 사실을 발견했다.

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

실제로 방정식 (8.2)과 초기 조건 $(Aa)_0 = 1$에서 $(Aa)_n = 2^n$가 얻어진다. 이를 방정식 (8.1)에 대입하면 $(AA)_{n+1} = 4(AA)_n + 2^n$을 구할 수 있다. 우리는 $c = -1/2$일 때 $(AA)_n = c2^n$이 특수해라는 것을 쉽게 알 수 있다. 제차방정식 $(AA)_{n+1} = 4(AA)_n$의 일반해는 $(AA)_n = C4^n$이다. 마지막으로 이 두 해를 더하면 $(AA)_n = C4^n - 2^{n-1}$은 $C = 1/2$일 때 초기 조건 $(AA)_0 = 0$을 만족한다는 것을 알 수 있다. 수열 $(aa)_n$의 경우, $(AA)_n$과 동일한 점화식과 초기 조건을 만족한다. 따라서 $(aa)_n = (AA)_n$이다.

결론적으로, 전체 개체군에서 잡종 $Aa$의 비율은 $2^n/4^n = 1/2^n$이며, 각 세대마다 자가 수정에 의해 2로 나뉜다.

멘델의 연구는 그의 생애 동안 전혀 주목받지 못했다. 몇 년 후 멘델은 다른 식물 종에 대해서도 비슷한 실험을 시도하고 기상학에 관한 몇 편의 논문을 발표했으며 꿀벌의 유전에 대해 조사했다. 1868년 수도원장이 된 후 그는 대부분의 시간을 행정 문제를 관리하는데 보냈다. 그는 1884년에 사망했다.

1900년에 이르러서야 멘델의 연구는 암스테르담의 휘호 더프리스(de Vries), 튀빙겐의 칼 코렌스(Correns), 비엔나의 에리히 폰 체르마크(Tschermak)에 의해 거의 동시에 독립적으로 재발견되었다. 이로써 오늘날 우리가 유전학이라고 부르는 새로운 시대가 열리게 된다.

# 참고문헌


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

9장

# 골턴, 왓슨, 그리고 소멸 문제 (1873–1875)

1873년 영국의 통계학자 골턴과 그의 동료 수학자 왓슨은 비에나메의 연구를 알지 못한 채 가족 성의 소멸 문제를 고민했다. 왓슨은 각 세대 남성 수의 확률 분포와 관련된 생성 함수가 재귀적으로 계산될 수 있다는 것을 알아냈다. 그러나 그는 절멸확률을 잘못 분석했다.

프랜시스 골턴(Galton)은 멘델(Mendel)과 같은 해인 1822년 영국 버밍엄 근처에서 태어났다. 그는 일곱 자녀 중 막내였으며 아버지는 부유한 은행가였다. 또한 찰스 다윈(Darwin)의 외사촌이었다. 골턴은 1838년 버밍엄의 한 병원에서 의학을 공부하기 시작했고, 이후 런던에서 학업을 이어나갔다. 1840년 여름, 그는 이스탄불까지 유럽을 횡단하는 첫 장거리 여행을 떠났다. 그 후 케임브리지 대학교의 트리니티 칼리지에서 4년간 공부했다. 하지만 1844년 아버지가 사망하면서 상당한 재산을 남겼고, 골턴은 의사가 되겠다는 생각을 접었다. 그는 이집트, 수단, 시리아로 여행했다. 그 후 몇 년 동안 그는 사냥을 하거나 열기구와 보트를 타고 여행하거나 전기 전신을 개선하는 데 시간을 보내며 여유로운 생활을 즐겼다. 1850년에는 남서아프리카(현 나미비아)로 탐험대를 꾸려 탐험을 떠났다. 1852년 영국으로 돌아온 그는 왕립지리학회 회원으로

선출되었다. 그곳에서 그는 나일강의 근원을 찾기 위해 동아프리카로 떠난 탐험대의 동향을 파악할 수 있었다. 런던에 정착한 그는 여행자를 위한 가이드북을 저술하였고 그 책은 베스트셀러가 되었다. 1856년에는 왕립학회 회원으로 선출되었다. 이후 그는 기상학에 관심을 갖고 '안티사이클론'이라는 단어를 새로 만들었다. 1859년 사촌 다윈이 ≪종의 기원≫을 출간한 후 골턴은 유전학 연구로 방향을 전환했다. 그는 1869년에 ≪유전적 천재≫를 출판하여 지적 능력이 유전에 의해 전달될 수 있다고 주장했다.

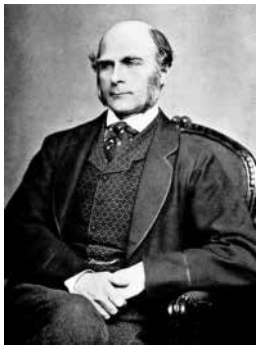

그림 9.1: 골턴(왼쪽)과 왓슨(오른쪽).

1873년 스위스의 식물학자 알퐁스 드 캉돌(Candolle)은 ≪유전, 변이, 선택이 인간 종의 발달과 미래에 미치는 각각의 영향≫에 대한 에세이를 담은 ≪과학과 과학자의 200년사≫라는 제목의 책을 출간했다. 여기서 그는 다음과 같은 발언을 했다.

> "베누아스 드 샤토네프, 골턴 및 기타 통계학자들의 정확한 정보와 합리적 의견들 중 나는 그들이 언급했어야 했던 필연적인 성씨의 소멸에 대한 발언을 보지 못했다. 물론 모든 성은 소멸될 수 밖에 없다 [...] 수학자는 한 커플이 여성 또는 남성의 자녀를 가질 확률과 자녀를 갖지 않을 확률을 알고있다면 성이나 호칭의 감소가 어떻게 일어날지 계산할 수 있다."

이것은 1845년 비에나메가 연구했던 것과 같은 문제이다. 그러나 비에나메의 연구를 알지 못했던 캉돌은 모든 가족 성은 소멸할 수 밖에 없다고 생각했다. 골턴은 캉돌의 책에서 위의 단락을 발견했다. 그 역시 비에나메의 연구에 대해 알지 못했기 때문에 골턴은 이 문제를 ≪Educational Times≫의 독자들에게 공개 문제로 제기했다.

> "문제 4001: 각각 다른 성씨를 가진 $N$ 명의 성인 남성을 가진 무리가 한 지역에 살고 있다. 이 무리의 인구 중 성인 남성만을 고려한다. 무리의 각 세대에서 성인 남성의 $a_0$ 퍼센트는 성인이 될 남자 자녀가 없고, $a_1$ 는 성인이 될 남자 자녀를 1명, $a_2$ 는 2명, 그리고 $a_5$ 는 5명 갖는다.
>
> (1) $r$ 세대가 지나면 무리의 성씨 중 몇 퍼센트가 소멸될 것인지, 그리고 (2) $m$ 명의 사람들이 가지고 있는 성씨의 수는 몇 개가 될 것인지 구하라."

문제의 두 번째 부분은 비에나메에 의해 제기되지 않은 새로운 문제임을 알 수 있다. 골턴은 저널 독자들로부터 만족스러운 답을 얻지 못했고 그 스스로도 문제의 해답을 찾을 수 없었던 것 같다. 그래서 그는 친구인 수학자 헨리 윌리엄 왓슨(Watson)에게 이 문제를 풀어달라고 부탁했다.

왓슨은 1827년 런던에서 태어났다. 아버지는 영국 해군 장교였다. 왓슨은 런던의 킹스칼리지에서 공부한 후 1846년부터 1850년까지 케임브리지 대학교 트리니티 칼리지에서 수학으로 전향하여 골턴보다 몇 년 뒤인 1846년부터 1850년까지 수학을 전공했다. 그는 트리니티 칼리지의 펠로우, 런던 시티 대학교의 조교, 킹스칼리지의 조교수, 1857년부터 1865년까지 해로우 스쿨의 수학 교수를 연이어 역임했다. 산악 등반에 관심이 많았던 그는 1855년 스위스 몬테로사산 정상에 오른 탐험대의 일원이었다. 1856년 부제 서품을 받았고 2년 후에는 성공회 사제로 서품되었다. 1865년부터 은퇴할 때까지 코번트리 인근의 버크스웰 위드바튼의 목사로 재직하며 연구에 몰두할 충분한 시간을 가질 수 있었다.

골턴과 왓슨은 1875년 ≪왕립 인류학 연구소 저널≫에 실린 ≪가족의 소멸 확률에 관하여≫라는 제목의 논문을 함께 썼다. 골턴이 문제를 제시하면 왓슨이 자신의 계산과 도달한 결론을 설명했다. 그들은 남자가

최대 $q$명의 아들을 가질 수 있다고 가정하고, $p_k$는 $k$명의 아들을 가질 확률($k=0,1,2,\ldots,q$)이라고 두었다. 즉, 골턴이 사용했던 표기대로라면 $p_k=a_k/100$이 된다. 따라서 $p_0+p_1+\cdots+p_q=1$이다. 0번째 세대가 한 남성으로 구성된 상황을 생각해 보자. 1세대는 확률 $p_s$로 $s$명의 남성으로 구성된다. 왓슨은 그의 시대에는 이미 잘 알려진, 한 세기 전 아브라암 드무아브르(de Moivre)가 제시한 트릭을 써서 확률 $p_0$, ..., $p_q$와 관련된 생성 함수를 고려했다.

$$f(x)=p_0+p_1x+p_2x^2+\cdots+p_qx^q \tag{9.1}$$

마찬가지로, $f_n(x)$를 $x^s$의 계수가 0세대의 한 남자로부터 시작하여 $n$세대에 $s$ 남성이 있을 확률을 계수로 가지는 다항식이라고 하자. 그러면 $f_1(x)=f(x)$가 된다. 왓슨은 다음과 같은 사실을 발견했다.

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

이 공식을 사용하면 $f_n(x)$를 재귀적으로 계산할 수 있다.

여기서 $f_n(x)=p_{0,n}+p_{1,n}x+p_{2,n}x^2+\cdots+p_{q^n,n}x^{(q^n)}$로 설정한다. n세대에는 최대 $q^n$ 명의 남성이 존재할 수 있다. $n-1$ 세대에 1부터 $s$까지 번호가 매겨진 $s$ 명의 남성이 있다면, 그들의 남성 자손의 수를 $t_1$, $\cdots$,$t_s$라고 부른다. 이 경우, $n$ 세대에는 다음과 같은 확률로 $t$ 명의 남성이 존재할 것이다.

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

$s=0$일 때, 이 확률은 $t=0$일 경우 1과 같고 $t\geqslant 1$일 경우 0과 같다. 그러므로

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

이고, 다음이 성립한다.

$$\begin{aligned} f_n(x) = \sum_{t\geqslant 0} p_{t,n}x^t &= \sum_{s\geqslant 0} p_{s,n-1} \sum_{t\geqslant 0} \sum_{t_1+\cdots+t_s=t} (p_{t_1}x^{t_1})\times\cdots\times(p_{t_s}x^{t_s}) \\ &= \sum_{s\geqslant 0} p_{s,n-1}\left[p_0x^0+p_1x^1+p_2x^2+\cdots\right]^s \\ &= \sum_{s\geqslant 0} p_{s,n-1}[f(x)]^s = f_{n-1}(f(x))\,. \end{aligned}$$

특히 $n$ 세대 내에 성이 소멸할 확률 $x_n$은 $p_{0,n}$과 같으며, 이는 $f_n(0)$과 동일하다. 첫번째 예로, 왓슨은 다음과 같이 두었다.

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

즉, $q=3$ 및 $p_0=p_1=p_2=1/3$ 이다. 그는 방정식(9.2)을 사용하여 $n=1,\ldots,4$에 대한 다항식 $f_n(x)$를 계산했다. 예를 들어 다음과 같은 결과를 얻었다.

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

그리고 $f_2(0)=13/27\approx 0.481$ 이다. $n\geqslant 3$에 대한 $f_n(x)$의 계산은 매우 길어져서 $n=4$의 경우 왓슨이 계산 실수를 할 정도였다. $x_5=f_5(0)=f_4(f(0))$를 이용하면 $f_5(x)$의 계산을 피할 수 있고, 다음과 같이 소멸할 확률 $x_n=f_n(0)$을 구할 수 있었다.

$$x_1\approx 0.333,\quad x_2\approx 0.481,\quad x_3\approx 0.571,\quad x_4\approx 0.641,\quad x_5\approx 0.675\,.$$

정확한 값은 $x_4\approx 0.632$와 $x_5\approx 0.677$이며, 이는 비에나메가 도출한 간단한 공식 $x_n=f(x_{n-1})$를 사용하여 확인할 수 있다. 17장에서 살펴보겠지만, 위의 공식은 방정식(9.2)에서도 도출할 수 있다.

왓슨은 첫 번째 예제에서 각 남성이 평균적으로 $\mathscr{R}_0=p_1+2p_2+\cdots+qp_q$의 아들을 가지고 있으며 그의 첫번째 예에서 $\mathscr{R}_0=1$이라는 것을 알아냈다. 따라서 초기 남성 가족 구성원의 수가 충분히 많으면 가족 규모

가 거의 일정하게 유지될 것이라고 생각할 수 있다. 그럼에도 불구하고 왓슨은 절멸확률 $x_n$이 매우 느리지만 $n \to +\infty$가 되면 1로 수렴한다고 주장했다. 즉, 캉돌이 주장한 대로 모든 가족이 멸종에 도달한다는 것이다. 원본 기사에는 그려져 있지 않은 그림 9.2(왼쪽)와 비에나메의 결과를 보면 첫 번째 예에 대한 결론이 옳다는 것을 확인할 수 있다.

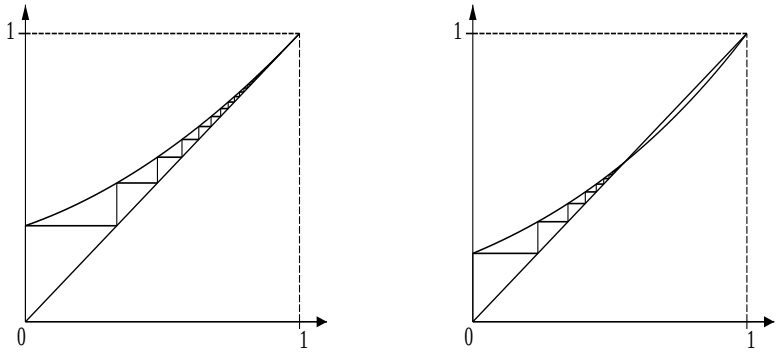


그림 9.2: 함수 $y = f(x)$와 $y = x$의 그래프. $n$ 세대 내의 소멸 확률 $x_n = f(x_{n-1})$는 $n$ 번째 '계단칸'까지의 높이이다. 왼쪽: $f(x) = (1+x+x^2)/3$. 오른쪽: $f(x) = (3+x)^5/4^5$.

두 번째 예로, 왓슨은 다음의 이항 확률 분포를 고려했다.

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

생성 함수(9.1)는 $f(x) = (a+bx)^q/(a+b)^q$ 이다. 왓슨은 $f_2(x)$와 $x_2 = f_2(0)$를 계산했다. 여기서 그는 $x_2 = f(x_1)$이고 모든 $n$에 대해 $x_n = f(x_{n-1})$이라는 것을 깨달았다. 그러나 그는 이 공식이 이항 분포의 특수한 경우(9.3)에만 적용된다고 생각했다. 이 공식을 $q = 5, a = 3, b = 1$인 경우에 적용하여 다음과 같은 결과를 얻었다.

$$x_1 \approx 0.237, \quad x_2 \approx 0.347, \quad x_3 \approx 0.410, \quad \ldots \quad x_9 \approx 0.527, \quad x_{10} \approx 0.533, \ldots$$

왓슨은 $x_n$이 $n \to +\infty$일때 극한값 $x_\infty$에 수렴해야 하며, $x_\infty = f(x_\infty) = (a+bx_\infty)^q/(a+b)$를 만족한다는 것을 깨달았다. 그는 $x=1$이 이 방정식의 해라는 것을 알았지만 $\mathscr{R}_0 > 1$일 때 다른 해가 있을 수 있다는 것을 깨닫지 못했다. 그래서 그는 캉돌의 오류의 영향으로 방금 고려한 수치적 예를 포함해 모든 경우에 소멸($x_\infty = 1$)하게 된다고 잘못 결론 내렸다. 그림 9.2(오른쪽)는 이것이 사실이 아님을 보여준다!

왓슨은 이 예시에서 평균 아들 수가 1보다 크다는 것을 알아냈다(계산해 보면 $\mathscr{R}_0 = qb/(a+b) = 5/4$가 된다.). 이는 인구가 기하급수적으로 증가하는 경향이 있다는 것을 의미한다. 그러나 이것은 그가 자신의 오류를 발견하는 데 도움이 되지 않았다. 그는 심지어 이항분포의 경우 뿐만 아니라 모든 확률 분포 ($p_k$)에 대해 성의 소멸이 확실하다고 추측했다. 이 문제는 17장과 18장에서 다시 다루겠다.

골턴은 왕립학회 회원들의 계보에 초점을 맞춘 ≪영국 과학자, 그들의 본성과 양육≫이라는 제목의 책을 내며 가족에 대한 통계 연구를 이어나갔다. 그는 또한 인체를 측정하는 인체 측정학에도 관심을 갖게 되었다. 그는 1884년 런던에서 열린 국제박람회 참가하여 많은 사람들의 데이터를 수집했다. 그의 결과는 1889년 ≪자연의 유전≫이라는 제목의 책으로 출판되었으며, 부록에는 왓슨과 공동으로 작성한 그의 논문이 실렸다. 이 책에는 '백분위수'와 '사분위수'와 같은 새로운 통계 어휘와 '우생학', 즉 유전적 특성의 관점에서 인간 종의 개량이라는 단어도 소개되었다. 1888년 이후 골턴은 몇 년 후 영국 경찰이 사용하게 될 지문 인식 기술을 개발했다. 또한 쌍둥이의 신체적, 지적 특성, 여러 세대에 걸쳐 자란 완두콩의 크기 또는 실험실에서 사육된 쥐의 색깔에 대한 유전(본성)과 환경(양육)의 각기 다른 역할에 대해 연구를 이어나갔다. 이를 통해 그는 두 변수 간의 '상관 계수'라는 개념을 고안했다. 1904년 런던 유니버시티 칼리지 내에 골턴 연구소가 설립되었다. 골턴은 1909년 기사 작위를 받았으며 1911년에 사망했다.

왓슨은 1876년에 기체의 운동 이론에 관한 논문, 1885년과 1889년 두 권에 걸친 전기와 자기의 수학적 이론에 관한 논문 등 여러 권의 저서를 출간했다. 1881년 왕립학회에 선출되었고 1903년 브라이턴에서 사망했다.

1924년, 칼 피어슨(Pearson)은 골턴 전기의 두 번째 권에서 가족 성의 소멸에 관한 논설의 오류를 알아차리지 못한 채 요약했다. 이 오류는 1930년에 마침내 발견되었다 (18장 참고).

# 참고문헌


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

10장

# 로트카와 안정적인 인구 이론 (1907–1911)

1907년 미국의 화학자 앨프리드 로트카는 연속 시간 모델을 사용하여 출생률, 연령별 사망률, 인구 증가율 간의 관계를 연구하기 시작했다. 1911년 그는 F. R. 샤프와 함께 같은 주제에 대한 또 다른 논문을 발표했는데, 여기에는 연령별 출산율도 포함되어 있었다. 인구 증가율을 나타내는 암시적 방정식은 흔히 '로트카 방정식'이라고 불린다.

앨프리드 제임스 로트카(Lotka)는 1880년 오스트리아-헝가리 제국의 일부였던 렘베르크(현재 우크라이나의 리비우)의 미국인 부모 사이에서 태어났다. 그는 프랑스와 독일에서 공부를 시작했고 이후 1901년 영국 버밍엄 대학교에서 물리학 및 화학 학사 학위를 받았다. 그 후 1909년 노벨 화학상 수상자인 빌헬름 오스트발트(Ostwald)가 화학과 생물학에서 열역학의 역할을 강조하던 라이프치히에서 1년을 보냈다. 1902년 뉴욕에 정착한 로트카는 제너럴 케미컬 컴퍼니에서 일하기 시작했다.

1907년과 1911년에 로트카는 같은 주제에 대한 오일러(Euler)의 연구를 알지 못한 채 연령구조의 인구 역학에 대한 연구를 시작했다 (3장

그림 10.1:
로트카 (1880–1949)

참조).[1] 오일러와 달리 그는 시간과 나이가 연속적인 변수라고 가정했다. 시간 $t$에서 남성 출생률(단위 시간당 남아 출생자 수)을 $B(t)$, 나이 $x$에서 아직 살아있을 확률을 $p(x)$, 나이 $x$에서 출산율을 $h(x)$라고 한다. $dx$가 무한히 작다면 $h(x)\,dx$는 한 남성이 나이 $x$에서 $x+dx$ 사이에 신생아 아들을 가질 확률이다. 출생시 기대 수명은 다음과 같다.

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

또한 $B(t-x)\,p(x)\,dx$는 시간 $t-x$와 $t-x+dx$ 사이에 태어나 시간 $t$에 여전히 살아 있는 남자의 수이다. 이 남자들은 시간 $t$에 단위 시간당 $B(t-x)\,p(x)\,h(x)\,dx$명의 아들을 가진다. 따라서 시간 $t$의 총 남성 출생률은 다음과 같다.

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

이 적분 방정식에 대해 $B(t) = b\,e^{rt}$ 형식의 지수 해를 구하기 위해 로트카는 양변을 $B(t)$로 나눈 후 다음의 방정식을 구했다.

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

[1] 두 번째 논문은 코넬 대학의 수학자 F.R. Sharpe와 공동으로 작성했다.

인구통계학자들은 이 방정식을 '로트카의 방정식'이라고 부른다.[2] 오일러는 시간과 연령이 이산 변수일 때 성장률에 대한 유사한 암시적 방정식 (3.1)을 구했다. 로트카는 (10.1)의 우변이 $r$의 감소 함수이며, $r \to -\infty$일 때는 $+\infty$로, $r \to +\infty$일 때는 0에 가까워지는 경향이 있다는 것을 알아냈다. 따라서 방정식 (10.1)이 성립하는 $r$의 고유한 값이 있다 (이를 $r^*$라고 하겠다.). 게다가, $r^* > 0$은 다음과 동치이다.

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

매개변수 $\mathscr{R}_0$(이 표기는 1925년 더블린과 로트카에 의해 도입되었다.)는 한 남성이 평생 동안 가지는 평균 아들 수이다.

로트카는 인구의 초기 연령 구조에 상관 없이 단위 시간당 남성 출생자 수는 $t \to +\infty$일 때 $B(t) \sim b e^{r^* t}$가 된다고 하였다 ($b$는 상수).[3] 따라서 총 인구는 다음과 같이 주어진다.

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

여기서 $P(t)$도 $t \to +\infty$일 때 $e^{r^* t}$처럼 증가하거나 감소하게 되며 성장률은 $r^*$와 같다. 또한 인구의 연령 분포는 $B(t-x)\, p(x) / P(t)$로 주어지며, 다음에 가까워진다.

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

로트카는 이를 '안정된 인구'라고 불렀는데, 연령 피라미드는 시간이 지나도 같은 모양을 유지하지만 총 인구는 기하급수적으로 증가하거나 감소한다. 이 결론은 오일러의 이산 시간 모델에서와 동일하다. 하지만 로트카의 연구는 연령에 따른 출산율의 차이를 고려하므로 어떤 의미에서 오일러의 연구보다 더 일반화된 것이다.

로트카는 평생 동안 이 주제에 대한 연구를 계속했다. 1908 ~1909년

[2] R. A. 피셔는 1927년에 독립적으로 동일한 방정식에 도달했으며, 이후 자연 선택에 의한 진화 이론에서 해 $r^*$를 '다윈적 적합성'의 척도로 해석했다.

[3] 이는 1941년 당시 미국 브라운 대학교의 수학과 교수였던 윌리 펠러에 의해 엄밀하게 증명되었다. 확률론적인 접근 방법은 1968년 크럼프, 모드, 예거스에 의해 개발되었다.

에는 코넬 대학교에서 석사 학위를 취득하기 위해 학업을 재개했다. 1909년부터 1911년까지 국립표준국에서 근무했으며, 1911년부터 1914년까지 ≪사이언티픽 아메리칸 서플리먼트≫ 저널의 편집자로 일했다. 1912년에는 버밍엄 대학교에서 1907년부터 발표한 인구 역학 및 인구 통계에 관한 논문을 정리하여 박사 학위를 받았다. 제1차 세계대전 중에는 다시 제너럴 케미컬 컴퍼니에서 일하며 대기 중 질소를 고정하는 방법에 대해 연구했다. 1920년에는 생물학적 진동에 관한 그의 논문 중 하나(13장 참조)가 당시 막 로지스틱 방정식(6장 참조)을 '재발견'한 존스 홉킨스 대학의 계량생물학 교수 레이몬드 펄(Pearl)에게 깊은 인상을 남겼다. 뉴욕 록펠러 의학연구소에 취직하기를 원했던 로트카는 로스(Ross)가 개발한 말라리아에 관한 수리모델을 연구했다(12장 참조). 마침내 그는 존스 홉킨스 대학에서 2년간의 장학금을 받아 1925년에 출판된 ≪물리생물학의 기초≫ 라는 책을 저술할 수 있었다. 그 후 그는 뉴욕에 있는 메트로폴리탄 생명 보험 회사의 연구 부서장이 되었다. 그는 인구통계학적 질문에 대한 수학적 분석에 중점을 두고 통계학자이자 부사장인 그의 동료 루이 이스라엘 더블린과 공동으로 ≪사람의 금전적 가치≫(1930), ≪수명≫(1936), ≪25년간의 건강증진≫(1937) 등 여러 권의 책을 출간했다. 또, 1938~1939년의 미국인구학회 회장직에 선출되었다. 그의 다양한 통계 연구 중 '로트카의 법칙'(1926년으로 거슬러 올라감)은 특정 과학 분야에서 $n$편의 논문을 쓴 저자의 수가 $n$이 증가함에 따라 대략 $1/n^2$로 감소한다는 법칙을 말한다.

로트카는 ≪생물학적 연관성 분석 이론≫이라는 제목의 프랑스어 책도 출간했다. 이 책의 첫 번째 파트는 좀 더 철학적이며 1934년에 출간되었다. 1939년에 출간된 좀 더 테크니컬한 두 번째 파트는 1907년 이후 인구학에 대한 그의 모든 연구를 요약한 것이다. 그의 저서에서 로트카는 가족 성의 소멸 문제에 대한 그의 기여도 다루었다. 1930년 이 주제에 대한 스테펜센(Steffensen)의 첫 번째 논문이 발표된 후(18장 참조), 그는 1920년 미국 백인 인구 조사에 포함된 데이터에 이 이론을 적용했다. 그는 관찰된 아들 수의 분포 $(p_k)_{k\geqslant 0}$가 모든 $k \geqslant 1$에 대해서 다음과 같이

감소하는 기하급수에 의해 잘 근사된다는 사실을 발견했다.

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

여기서 $a = 0.4825$, $b = 0.2126$, $c = 1 - b/(1-a)$ 이다. 그러면 $\sum_{k\geqslant 0} p_k = 1$ 이 된다. 여기서 생성 함수는 다음과 같다.

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

$x = f(x)$ 방정식의 두 해는 $x = 1$과 $x = a/c$ 이다. 절멸확률 $x_\infty$는 이 두 해 중 가장 작은 해이다(7장 참조). 미국에 대한 수치로 그는 $x_\infty \approx 0.819$를 구했고, 평균 아들 수는 $\mathscr{R}_0 = f'(1) = (1-a)^2/b \approx 1.260$였다. 평균 자녀 수(아들, 딸 포함)가 2.5명에 가까운데도 불구하고 가족 성이 소멸할 확률은 80% 이상이다.

로트카는 1942년 미국 통계학회 회장으로 선출되었다. 그는 1947년에 은퇴하고 1949년 뉴저지에서 사망했다. 1925년에 출간된 그의 책의 개정판은 1956년에 ≪수리생물학의 기본≫ 이라는 조금 다른 제목으로 출간되었다.

# 참고문헌


1. Crump, K.S., Mode, C.J.: A general age-dependent branching process. J. Math. Anal. Appl. 24, 494–508 (1968)
2. Dublin, L.I., Lotka, A.J.: On the true rate of natural increase. J. Amer. Stat. Assoc. 20, 305–339 (1925)
3. Feller, W.: On the integral equation of renewal theory. Ann. Math. Stat. 12, 243–267 (1941). projecteuclid.org
4. Fisher, R.A.: The actuarial treatment of official birth records. Eugen. Rev. 19, 103–108 (1927). digital.library.adelaide.edu.au
5. Gridgeman, N.T.: Lotka, Afred James. In Gillespie, C.C. (ed.) Dictionary of Scientific Biography, vol. 8, 512. Scribner, New York (1981)

6. Jagers, P.: Age-dependent branching processes allowing immigration. Theor. Probab. Appl. 13, 225–236 (1968).
7. Lotka, A.J.: Relation between birth rates and death rates. Science 26, 21–22 (1907) → Smith & Keyfitz (1977).
8. Lotka, A.J.: Théorie analytique des associations biologiques, 2$^{e}$ partie. Hermann, Paris (1939) gallica.bnf.fr
9. Sharpe, F.R., Lotka, A.J.: A problem in age-distribution. Philos. Mag. Ser. 6, 21, 435–438 (1911) → Smith & Keyfitz (1977).
10. Smith, D.P., Keyfitz, N.: Mathematical Demography. Springer (1977)
11. Tanner, A.: Von Molekülen, Parasiten und Menschen, Lotka und die Mathematisierung des Lebens. Zürich (2014) doi:10.3929/ethz-a-010209129

11장

# 하디-바인베르크 법칙 (1908)

1908년 영국의 수학자 하디와 독일의 의학자 바인베르크는 멘델의 법칙에 따라 무작위로 짝짓기를 하는 무한히 큰 집단에서 두 대립 유전자에서 얻은 유전자형의 빈도가 세대를 거쳐 일정하게 유지된다는 사실을 독자적으로 발견했다. 그들의 수학적 모델은 인구유전학의 출발점 중 하나였다.

고드프리 해럴드 하디(Hardy)는 1877년 영국 서리주에서 태어났다. 그의 부모님은 교사였다. 1896년부터 케임브리지 대학교 트리니티 칼리지에서 수학을 공부했고, 1900년에는 같은 대학의 펠로우가, 1906년에는 수학 교수가 되었다. 첫 번째 저서 ≪단일 변수 함수의 적분≫(1905) 이후 1908년 ≪순수 수학 강의≫를 출간했으며, 이 책은 수차례 재편집되어 여러 외국어로 번역되었다.

당시 멘델(Mendel)의 연구가 재발견되면서 몇 가지 의문이 제기되었다. 일부 생물학자들은 왜 우성의 특성이 세대를 거듭할수록 더 빈번하게 나타나지 않는지 궁금해했다. 1905년 ≪멘델리즘≫라는 제목의 책을 쓴 레지널드 퍼넷(Punnett)은 케임브리지에서 함께 크리켓을 하던 하디에게 이 질문을 던졌다. 하디는 1908년에 출판된 ≪혼합 집단에서의 멘델 비율≫에 관한 논문에서 자신의 해법을 제시했다. 분석을 단순화하기 위해 그는 성 파트너의 선택이 무작위로 이루어지는 대규모 집단을 상상

그림 11.1: 하디 (1877–1947)

했다. 또한 그는 $A$와 $a$, 즉 우성과 열성의 두 가지 인자(또는 '대립유전자')에만 관심을 제한했다. 세대 $n$에 대해 $p_n$을 '유전자형' $AA$의 빈도, $2q_n$은 $Aa$의 빈도, $r_n$은 $aa$의 빈도라고 하자. 물론 $p_n + 2q_n + r_n = 1$이다. 하디는 또한 이 유전자형 중 어느 것도 다른 두 유전자형과 비교했을 때 사망률의 증가나 출산율의 감소를 일으키지는 않는다고 가정했다. $n+1$ 세대의 빈도는 $n$세대에서 무작위로 선택된 한 개체가 $p_n + q_n$의 확률로 대립유전자 $A$를 전달한다는 점을 생각하면 쉽게 계산할 수 있다. 그 개체의 유전자형이 $AA$일 경우 대립유전자 $A$가 확실히 전달될 것이고 유전자형이 $Aa$인 경우에는 대립유전자 $A$가 50% 확률로 전달되기 때문이다. 마찬가지로 대립유전자 $a$는 $q_n + r_n$의 확률로 전달된다. 따라서 표 11.1를 표 8.1과 같은 방식으로 구성할 수 있다.

표 11.1: 부모의 대립유전자 빈도로부터 계산한 $n+1$ 세대의 유전자형 빈도(행은 어머니, 열은 아버지에 해당).

| 대립유전자<br>빈도 | $A$<br>$p_n + q_n$ | $a$<br>$q_n + r_n$ |
|---|---|---|
| $A$<br>$p_n + q_n$ | $AA$<br>$(p_n + q_n)^2$ | $Aa$<br>$(p_n + q_n)(q_n + r_n)$ |
| $a$<br>$q_n + r_n$ | $Aa$<br>$(p_n + q_n)(q_n + r_n)$ | $aa$<br>$(q_n + r_n)^2$ |

$n+1$ 세대에서 유전자형 $AA$, $Aa$, $aa$의 빈도는 각각 $p_{n+1}$, $2q_{n+1}$, $r_{n+1}$ 이다. 따라서 하디는 다음의 공식을 찾았다.

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

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

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

그런 다음 그는 어떤 조건에서 유전자형들의 빈도가 각각 $p$, $2q$ 및 $r$로 여러 세대에 걸쳐 일정하게 유지될 수 있는지 살펴보았다. 정의에 따라 $p+2q+r=1$ 이므로, 방정식 (11.1)-(11.3) 모두 동일한 조건 $q^2 = pr$ 하에 성립함을 알 수 있다.

> 예를 들어, 첫 번째 방정식에 의해 $p = (p+q)^2 = p^2 + 2pq + q^2$ 가 되고, 이는 $p(1-p-2q) = q^2$ 와 같고 따라서 $pr = q^2$ 가 된다.

$p_0 + 2q_0 + r_0 = 1$ 를 만족하는 임의의 초기 조건 $(p_0, 2q_0, r_0)$ 에서 시작하여 하디는 다음의 식이 성립함을 발견했다.

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

따라서 $(p_1, 2q_1, r_1)$ 상태는 이미 평형 상태이며 $(p_n, 2q_n, r_n)$ 은 모든 $n \geqslant 1$ 에 대해 $(p_1, 2q_1, r_1)$ 과 동일하게 유지된다. 0세대에서 대립유전자 $A$의 빈도를 $x = p_0 + q_0$ 로 설정하면, $1-x = q_0 + r_0$ 은 대립유전자 $a$의 빈도이다. 연립방정식 (11.1)–(11.3)을 다시 한 번 사용하면 모든 $n \geqslant 1$에 대해 다음이 성립함을 알 수 있다 (그림 11.2).

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

결과적으로, 위의 가설은 유전자형 $AA$, $Aa$ 및 $aa$의 빈도가 세대를 거쳐도 변하지 않는 법칙으로 이어진다. 멘델의 이론은 처음에 생각했던 것처럼 우성 유전자의 빈도가 점진적으로 증가하지 않는다.

몇 년 후 피셔는 이 법칙의 중요한 따름 정리를 주장했는데, 근사적

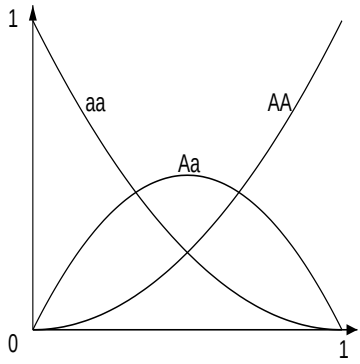


그림 11.2: 유전자형 $AA, Aa$ 및 $aa$의 평형상태의 빈도에 해당하는 함수 $x^2, 2x(1-x)$ 및 $(1-x)^2$의 그래프.

으로 (즉, 모델의 가설이 현실적이라고 가정할 때) 집단은 일정한 유전적 분산을 유지한다는 것이었다. 이 발견은 다윈(Darwin)의 자연 선택에 의한 진화 이론의 문제 중 하나를 해결한다. 실제로 다윈은 동시대 사람들과 마찬가지로 각 세대에서 자녀의 생리적 특성은 두 부모의 특성의 일종의 평균이며, 각 부모가 절반씩 기여한다고 생각했다. 이 아이디어는 후에 프랜시스 골턴(Galton)과 그의 후계자인 생체학 연구소의 칼 피어슨(Pearson)에 의해 통계학을 사용하여 철저히 연구되었다. 만약 이것이 사실이라면, 한 개체군에서 이러한 특성의 차이는 각 세대마다 2로 나뉘어져야 하며, 곧 진화를 설명하는 자연 선택이 불가능할 정도로 동질성을 갖게 될 것이다. 그럼에도 불구하고 생체학자들은 다윈의 관점을 옹호하고 멘델의 법칙이 진화를 이해하는 데 불가피하다는 사실을 인정하기를 꺼려하면서 이 평균화 메커니즘이 거부되기까지 몇 년이 필요했다.

1908년의 이 연구 이후 하디는 순수 수학으로 돌아갔다. 자서전 ≪어느 수학자의 변≫에서 그는 실용적인 것에 대한 발견을 피한 것에 대해 자부심을 가진다고 주장하기도 했다. 1910년 그는 왕립학회에 선출되었다. 1913년에는 인도의 신동 라마누잔(Ramanujan)을 발견하고 케임브

리지에서 일하도록 그를 초대했다. 제1차 세계대전이 끝난 후엔 옥스퍼드 대학의 교수가 되어 동포 리틀우드(Littlewood)와 성공적인 협업을 이어나갔다. 1931년부터 1942년까지 그는 다시 케임브리지에서 교수로 재직했다. 그는 종종 공동 작업을 통해 ≪무한의 규칙≫(1910), 마르셀 리스(Riesz)와 함께 ≪디리클레 급수의 일반이론≫(1915), 리틀우드 및 폴야(Pólya)와 함께 ≪부등식≫(1934), E. M. 라이트(Wright)와 함께 ≪정수론 개론≫(1938), ≪라마누잔≫(1940), 로고신스키(Rogosinski)와 함께 ≪푸리에 급수≫(1944), ≪발산 급수≫(1949) 등의 많은 책을 저술했다. 그는 1947년 케임브리지에서 사망했다.

그림 11.3: 바인베르크 (1862–1937)

수십 년 후, 사람들은 유전자형 빈도에 대한 하디의 법칙이 1908년 독일의 의사인 빌헬름 바인베르크(Weinberg)에 의해 독자적으로 발견되었다는 사실을 알게 되었다. 바인베르크는 1862년 슈투트가르트에서 태어났다. 튀빙겐과 뮌헨에서 의학 박사 학위를 받을 때까지 공부한 후 베를린, 비엔나, 프랑크푸르트의 병원에서 수년간 근무했다. 그는 1889년 슈투트가르트에 일반의 겸 산부인과 의사로 정착했다. 업무로 매우 바빴음에도 불구하고 그는 독일 과학 저널에 많은 논문을 쓸 시간을 확보했다. 1901년 그는 동성 쌍둥이의 빈도를 통계적 관점에서 연구했다. 하디가 발견한 것과 같은 법칙을 설명한 1908년 논문은 지역 과학 저널에

게재되었지만 주목받지 못했다. 그러나 하디와 달리 그는 이후 수년간 이 연구를 계속하여 대립유전자가 두 개보다 많은 경우의 일반화 등에 대한 결과를 얻었다. 그는 의학 통계 분야에도 기여했다. 바인베르크는 1937년에 사망했다. 그의 1908년 논문이 재발견된 후 유전학자들은 유전자형 빈도의 안정성 법칙을 '하디-바인베르크 법칙'이라고 불렀다.

오늘날 이 법칙은 다음과 같이 자주 사용된다. 생존이나 생식력에 영향을 미치지 않는 희귀 열성 대립유전자 $a$가 있다고 가정해 보자. 유전자형 $aa$가 특정 표현형을 생성하기 때문에 유전자형 $aa$의 빈도 $x^2$를 알고 있다면, 먼저 $x$를 계산한 다음 유전자형 $Aa$의 빈도 $2x(1-x) \approx 2x$를 알아낼 수 있다. 예를 들어, $aa$의 빈도가 $1/20{,}000$이면 $x \approx 1/140$이 된다. 따라서 $2x \approx 1/70$은 유전자형 $Aa$의 빈도이다. 표현형만 보았을 때는 매우 드물게 보일 수 있는 열성 대립유전자 $a$는 실제로는 그렇게 드물지 않다.

## 참고문헌


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

12장

# 로스와 말라리아 (1911)

1911년 말라리아 연구로 이미 1902년 노벨상을 수상한 영국의 의사 로널드 로스는 말라리아의 확산을 모델링하는 미분 방정식 시스템을 연구했다. 그는 모기의 수가 특정 임계값 이상일 때만 말라리아가 지속될 수 있음을 보여주었다. 따라서 말라리아를 박멸하기 위해 모든 모기를 박멸시킬 필요 없이 일부만 없애면 충분하다. 비슷한 전염병 모델이 이후 커맥과 맥켄드릭에 의해 개발되었다.

로널드 로스(Ross)는 1857년 인도 북부에서 태어났으며, 그의 아버지는 영국군 장교였다. 그는 런던에서 의학을 공부했지만 시와 드라마를 쓰는 것을 더 좋아했다. 1년 동안 배에서 외과의사로 일한 후 1881년 인도 의무성에 들어갈 수 있었다. 인도에서 의료 활동을 하면서 많은 자유 시간을 가졌고, 그 시간 동안 문학 작품을 쓰고 수학을 독학했다. 1888년 영국으로 휴가를 떠난 그는 공중보건학 학위를 취득하고 몇 년 전 파스퇴르(Pasteur)와 코흐(Koch)가 창안한 새로운 과학인 세균학을 공부했다. 인도로 돌아온 로스는 말라리아를 연구하기 시작했다. 1894년 두 번째 휴가 중 런던에서 만난 열대 의학 전문가 패트릭 맨슨(Manson)은 1880년 프랑스 군의관 알퐁스 라베랑(Laveran)이 발견한 사실인 말라리아 환자의 혈액에 기생충이 있다는 것을 현미경으로 보여줬다. 맨슨이 중국에 있을 때 모기에서 다른 열대성 질병(사상충증)의 기생충을 발견한 적이

있었기 때문에 기생충이 모기에서 왔을 수 있다고 했다. 그러나 그는 모기에 의해 오염된 물을 마실 때 인간이 기생충에 감염된다고 믿었다. 1895년부터 1898년까지 로스는 인도에서 연구를 계속하며 맨슨의 아이디어를 실험 했다. 1897년 그는 이전에 연구하지 않았던 특정 모기 종(아노펠레스)의 뱃속에서 라베랑이 관찰한 것과 유사한 기생충을 발견했다. 말라리아가 거의 발생하지 않는 계절에 상사에 의해 콜카타로 보내진 그는 조류에서 말라리아를 연구하기로 결정했다. 그는 아노펠레스 모기의 침샘에서 기생충을 발견하고 건강한 새를 모기에 물려 감염시키는 실험에 성공하여 말라리아가 오염된 물 섭취가 아니라 모기에 물려서 전염된다는 사실을 증명했다. 1899년 로스는 인도 의무성을 떠나 1년 전에 설립된 리버풀 열대의학대학에서 강의 시작했다. 1901년 왕립 학회 회원으로 선출되었고 1902년에는 말라리아 연구로 노벨 생리의학상을 수상했다. 그는 모기 퇴치를 널리 알리기 위해 아프리카, 모리셔스, 지중해 지역을 여행했다. 이 방법은 수에즈 운하를 따라 이집트, 건설중인 파나마 운하, 쿠바 및 말레이시아에서 성공했다. 다른 지역에서는 그다지 성공적이지 못했다. 로스는 1908년에 ≪모리셔스의 말라리아 예방에 관한 보고서≫를, 1910년에 ≪말라리아 예방≫을 발표했다.

그림 12.1: 로스 (1857–1932)

로스는 말라리아 전파에서 특정 모기의 역할을 증명했지만, 단순히 모기 수를 줄이는 것만으로 말라리아를 퇴치할 수 있다고 주장하면서 회의론에 부딪혔다. 1911년에 출간된 ≪말라리아 예방≫의 두 번째 판

에서 그는 자신의 주장을 뒷받침하기 위해 말라리아 전파에 대한 수학적 모델을 도입하였다. 그의 모델 중 하나는 이원 연립 미분 방정식으로 구성되었다. 다음의 표기를 보자.

- $N$: 특정 지역의 총 인구 수
- $I(t)$: $t$ 시점에 말라리아에 감염된 사람의 수
- $n$: 총 모기 개체 수 (상수라고 가정)
- $i(t)$: 말라리아에 감염된 모기 수
- $b$: 모기의 흡혈율
- $p$(또는 $p'$): 한 번의 흡혈시 말라리아가 사람에서 모기로(또는 모기에서 사람으로) 전염될 확률
- $a$: 사람이 말라리아에서 회복하는 속도
- $m$: 모기의 사망률

짧은 시간 간격 $dt$ 동안 감염된 모기 한 마리가 $b\,dt$ 명의 사람을 물며, 이 중 $\frac{N-I}{N}$ 의 비율은 아직 감염되지 않았다. 전염 확률 $p'$를 고려하면, 새로 감염된 사람 수는 $b\,p'\,i\,\frac{N-I}{N}\,dt$ 이다. 같은 시간 간격 동안 회복되는 사람 수는 $a\,I\,dt$ 이다. 따라서

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

가 된다. 마찬가지로 감염되지 않은 모기는 각각 $b\,dt$ 명의 사람을 물며, 이 중 $I/N$의 비율은 이미 감염된 모기이다. 전염 확률 $p$를 고려하면, 새로 감염된 모기의 수는 $b\,p\,(n-i)\,\frac{I}{N}\,dt$ 가 된다. 한편 감염이 사망률에 영향을 미치지 않는다고 가정하면, 사망하는 모기의 수는 $m\,i\,dt$ 이다. 따라서

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

이다. 말라리아는 대부분의 감염 국가에서 영구적으로 존재하기 때문에, 로스는 감염된 사람의 수 $I(t)$와 감염된 모기의 수 $i(t)$가 시간이 지나도

일정하게 유지되는 두 방정식의 평형점($dI/dt = 0$ 및 $di/dt = 0$)만을 고려했다. 먼저 말라리아가 없는 상태에 해당하는 $I = 0$ 및 $i = 0$의 평형점이 항상 존재한다. 다음으로 로스는 $I > 0$ 및 $i > 0$인 평형점을 찾았고 다음을 발견했다.

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

평형점에 대한 방정식들을 곱 $I \times i$로 나누면, 두 개의 미지수 $1/I$와 $1/i$로 이루어진 선형 연립 방정식을 얻게 된다.

$$\frac{bp'}{I} - \frac{a}{i} = \frac{bp'}{N}, \quad -\frac{m}{I} + \frac{bpn}{Ni} = \frac{bp}{N}.$$

방정식의 해는 쉽게 찾을 수 있다.

모기 수가 어떤 임계값을 초과하는 경우, 즉

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

이면 $I > 0$ 및 $i > 0$이 유지됨을 알 수 있다. 이 경우 평형점은 질병이 풍토병인 상황, 즉 영구적으로 존재하는 상황에 해당한다. 로스는 모기의 수 $n$이 임계치 $n^*$ 이하로 감소하면 $I = 0$과 $i = 0$의 평형점만 남게 되므로 말라리아가 사라질 것이라고 결론지었다. 특히 말라리아를 박멸하기 위해 모든 모기를 박멸할 필요는 없다. 이것이 바로 로스가 자신의 모델을 통해 강조하고 싶었던 점이다.

자신의 이론을 설명하기 위해 로스는 모델의 매개변수에 대한 합리적인 수치를 찾았다. 그는 다음과 같이 가정 했다.

- 모기의 사망률은 10일 후 1/3이 살아남으므로 $e^{-10m} = \frac{1}{3}$ 이 된다. 따라서 $m = (\log 3)/10$ 일 이다.
- 감염된 사람의 절반은 3개월 후에도 여전히 감염되어 있으므로 $e^{-90a} = 1/2$, 즉 $a = (\log 2)/90$ 일 이다.

- 매일 모기 8마리 중 1마리가 문다고 가정하면 $e^{-b} = 1 - 1/8$ 이고 $b = \log(8/7)$ 일 이다.

- 감염된 모기는 기생충이 여러 단계의 변태 과정을 거쳐야 하기 때문에 감염 후 처음 10일 동안은 보통 감염력이 없다. 모기의 3분의 1은 열흘 동안 생존할 수 있기 때문에, 로스는 감염된 모기의 3분의 1 정도는 전염성이 있다고 가정했다. 즉, $p' = 1/3$ 이다.

- $p = 1/4$ 이다.

그런 다음 로스는 공식(12.1)를 이용해 인간 인구에서 감염된 비율 $I/N$을 모기와 인간 인구 사이의 비율 $n/N$에 대한 함수로써 계산할 수 있었다. 그는 이 결과를 그림 12.2에 대응하는 표에 나타내었다.

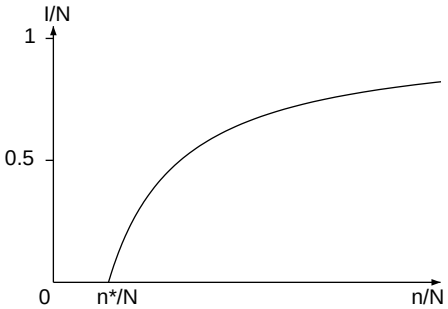


그림 12.2: 모기와 인간 인구 사이의 비율 $n/N$ 에 대한 함수로 표현된 감염 인간의 비율 $I/N$.

곡선의 모양은 비율 $n/N$이 임계값 $n^*/N$보다 약간만 더 높아도 감염된 사람의 비율이 이미 50% 를 넘는다는 것을 보여준다. 그러나 비율 $n/N$이 더 증가해도 감염된 사람의 비율은 크게 변하지 않는다. 이것으로부터 왜 이전에는 모기 수와 말라리아의 사이의 관계를 알아차리기 힘들었는지를 유추해 볼 수 있다. 그러나 로스는 임계값 $n^*/N$의 수치값이 흡혈율 $b$의 작은 변화에 매우 민감하다는 것을 알아차렸고, 그럼에도 불구하고 그림 12.2의 곡선 모양은 변하지 않는다는 사실을 발견했다. 그의 이러한 정성적 해석은 매개변수의 수치에 대한 불확실성에 영향을 받는 정량적 결과보다 더 중요한 의미를 갖고 있다.

로스 박사가 발견한 임계값 $n^*$을 해석하기 위해, 인간 인구와 모기 인구가 모두 감염되지 않은 지역에 감염된 사람 한 명이 유입되었다고 가정해 보겠다.[1] 이 감염자는 평균적으로 $1/a$의 기간 동안 감염 상태를 유지한다. 이 사람은 단위 시간당 $bn/N$의 모기에 물렸으므로 감염된 동안 평균적으로 $bn/(aN)$의 모기에 물린 것이다. 따라서 평균 $bpn/(aN)$ 마리의 모기를 감염시킨다. 이렇게 감염된 모기들은 평균 $1/m$의 기간 동안 생존하며, $b/m$의 사람을 물어서 $bp'/m$의 사람을 감염시킨다. 처음 감염된 인간에서 모기들로, 그리고 이 모기들에서 다른 인간들로 전염이 이루어진 후, 새로 감염된 인간의 평균 수는 앞의 두 결과의 곱이다.

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

이 $\mathscr{R}_0$은 초기 감염자 한 명으로 인한 2차 감염자 수이다. 따라서 시간이 지남에 따라 지속적으로 발생하는 감염 과정도 연이은 감염의 세대를 통해 고려할 수 있다. 말라리아는 $\mathscr{R}_0 > 1$일 때만 인구에 '침입'할 수 있다. 이 조건은 $n > n^*$와 정확히 일치한다.

더 나아가, 로스는 역학에서 수학적 모델링의 중요성에 대해 설파하였다.

> "질병이 시기에 따라 또는 장소에 따라 변한다는 것을 고려하면 모든 역학 문제에는 수리적 접근이 필요하다. 하지만 과학적 결론을 위해서는 많은 변수들이 포함되어야 한다. 질병이 특정 요인에 의존한다고 말하는 것은 각 요인이 전체 결과에 얼마나 큰 영향을 미치는지에 대한 추정치를 구할 수 있을 때까지는 큰 의미가 없다. 여기에의 수학적 해결 방법은 실제로는 문제에 신중한 추론을 적용하는 것일 뿐이다."

로스는 1911년에 기사 작위를 받았다. 그는 런던으로 이주하여 제1차 세계 대전 중 영국군의 자문 위원이 되었다. 1923년에는 자서전 ≪심각한 말라리아 문제와 그 해결책에 대한 전체 설명이 포함된 회고록≫을

[1] 이 해석은 로스가 연구한 지 한참 후에야 강조되었다.

출간했다. 1926년에는 로스 열대병 연구소(현재 런던 위생열대의학대학원의 일부)가 설립되었고 소장으로 취임했다. 로스는 1932년 런던에서 사망했다.

## 참고문헌

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

13장

# 로트카, 볼테라와 포식자-피식자 시스템 (1920–1926)

1920년 앨프리드 로트카는 포식자-피식자 모델을 연구하여 개체군의 수가 영구적으로 진동할 수 있음을 보여주었다. 그는 1925년 저서 ≪물리 생물학의 원리≫에서 이 연구를 발전시켰다. 1926년 이탈리아의 수학자 비토 볼테라는 생물학자 움베르토 단코나가 제기한 질문, 즉 어획량이 적었던 제1차 세계대전 당시 아드리아해에서 어부들이 잡은 포식자 물고기가 더 많았던 이유에 대한 답을 찾기 위해 동일한 모델에 관심을 갖게 되었다.

1920년 로트카(Lotka)는 ≪생태계의 주기적 관계에 대한 분석≫이라는 제목의 논문을 발표했다. 그는 이미 몇 년 전부터 실험 중에 이상한 일시적 진동을 보이는 화학 반응에 관심을 갖고 있었다. 그의 논문의 목표는 두 생물 종으로 구성된 시스템이 영구적으로 진동할 수 있다는 것을 제안하는 것이었다. 그가 고려한 예는 식물과 초식 동물의 개체군이다. 화학 반응속도론에서 사용되는 방정식과 유사하게, $x(t)$를 시간 $t$의 식물의 총 질량, $y(t)$를 초식동물의 총 질량이라고 가정한다. 로트카는

다음과 같은 연립미분방정식을 모델로 사용했다.

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

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

여기서 매개변수 $a, b, c$ 및 $d$는 모두 양수이다. 매개변수 $a$는 초식 동물이 없을 때 식물의 성장률이다. 매개변수 $c$는 식물이 없을 때 초식 동물의 개체수 감소율이다. $-bxy$ 및 $dxy$ 항은 동물과 식물이 많을수록 식물에서 동물로의 질량 이동이 증가한다는 것을 나타낸다 (이 이동에는 약간의 질량 손실이 포함되므로 $d \leqslant b$ 이다). 로트카는 $dx/dt = 0$과 $dy/dt = 0$을 통해 두 가지 정상상태가 있다는 것을 알아냈다.

- $(x = 0, y = 0)$, 초식 동물이 멸종하고 식물이 존재하지 않는다.

- $(x = c/d, y = a/b)$, 초식동물과 식물이 공존한다.

그는 또한 $t = 0$ 시점에 $(x(0), y(0))$가 이 두 정상 상태 중 하나가 아니라면 함수 $x(t)$와 $y(t)$는 주기적으로 진동한다, 즉, 모든 $t > 0$에 대해 $x(t+T) = x(t)$와 $y(t+T) = y(t)$가 되는 수 $T > 0$가 존재한다고 증명을 생략하고 적었다(그림 13.1).[1] 예를 들어 식물이 풍부하다면 초식 동물의 개체수가 증가하여 식물의 총 질량이 감소한다. 이 질량이 초식동물을 먹이기에 부족해지면 일부 동물은 굶어 죽고 식물의 총 질량은 초기 값과 같은 수준에 도달할 때까지 다시 증가하기 시작한다. 이 현상은 반복된다.

로트카는 1920년에 발표한 ≪질량보존의법칙에서 파생된 비 감쇠 진동≫이라는 제목의 두 번째 논문에서 이 모델을 좀 더 자세히 연구했다. 그는 시스템이 주기적으로 진동할 수 있는 이유를 설명했다. 이는 가로축 $x$, 세로축 $y$의 평면에서, 더 정확하게는 $x \geqslant 0$과 $y \geqslant 0$인 사분면에서 점 $(x(t), y(t))$가 닫힌 궤적 위에 있어야 한다는 사실에 따른 것이다 (그림 13.2).

[1] 기간 $T$는 초기 조건에 따라 달라지지만, 로트카는 1925년에야 이 사실을 깨달았다.

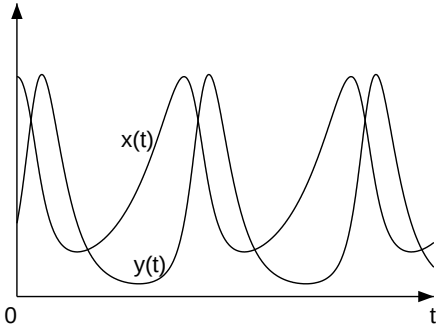


그림 13.1: 시간에 대한 함수로서 식물의 총 질량 $x(t)$ 와 초식 동물의 총 질량 $y(t)$ 의 진동.

실제로 방정식 (13.1)를 방정식 (13.2)로 나눈 뒤 약간의 순서를 바꾸면 다음과 같은 결과를 얻을 수 있다.

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

이 미분방정식을 적분하면

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

가 된다. 여기서 $K$는 초기 조건에만 의존하는 상수이다. 따라서 점 $(x(t),y(t))$은 항상 $d\,x-c\log x=a\log y-by+K$ 곡선 위에 있으며, 이는 닫힌 곡선이다(그림 13.2).

$(x(t),y(t))$의 궤적은 정상 상태 $(c/d,a/b)$를 중심으로 시계 반대 방향으로 회전하는데, 이는 $dx/dt$와 $dy/dt$의 부호를 살펴보면 쉽게 알 수 있다. 정상 상태 근처에서 시스템은 $2\pi/\sqrt{ac}$의 주기를 갖는 작은 진동을 보인다.

실제로 $x=\frac{c}{d}+x^*$ 와 $y=\frac{a}{b}+y^*$ 로 두고, $|x^*|\ll\frac{c}{d}$ 와 $|y^*|\ll\frac{a}{b}$ 라고

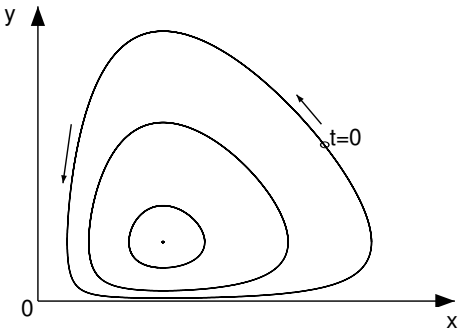


그림 13.2: 가로축에 식물의 총 질량 $x(t)$, 세로축에 초식 동물의 총 질량 $y(t)$ 가 있는 다이어그램. 정상 상태 주변의 세 개의 닫힌 곡선은 서로 다른 초기 조건에 대응된다.

하자. 그러면

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

이다. 이 두 방정식에서 다음을 얻을 수 있다.

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

이 방정식은 물리학의 단순 진자의 운동에서 진동 방정식과 동일하다. 주기는 $2\pi/\sqrt{ac}$ 이다.

1920년 로트카의 논문을 미국 국립과학원 회보에 발표했던 레이몬드 펄(Pearl)은 로트카가 존스홉킨스 대학에서 2년간 장학금을 받으며 ≪물리생물학 개론≫라는 제목의 책을 쓸 수 있도록 도왔다. 이 책은 1925년에 출판되었다. 1920년의 연구를 요약한 부분에서는 숙주 한 종과 기생충 한 종 또는 피식자 한 종과 포식자 한 종 등 두 종으로 이루어진 시스템도 (13.1)–(13.2)와 동일한 모델로 기술할 수 있다고 언급했다. 불행히도 로트카의 책은 출판 당시에는 큰 주목을 받지 못했다. 그러나 얼마 지나지 않아 유명한 수학자 볼테라가 어업 문제를 연구하던 중 같은 모델을 독자적으로 재발견했다.

비토 볼테라(Volterra)는 이탈리아가 통일되기 직전인 1860년 안코나의 유대인 게토에서 태어났고, 당시 안코나는 아직 교황청에 속해 있었다. 볼테라는 외동아들이었다. 옷감 상인이었던 그의 아버지는 비토가 두 살 때 돌아가셨고 가족에게 남긴 돈이 없었다. 고등학교 시절 공부를 잘했던 볼테라는 가난 속에서도 공부를 계속하여 처음에는 피렌체 대학교에서, 나중에는 피사의 스쿠올라 노르말레 수페리오레 대학교에서 공부했다. 1882년 물리학 박사 학위를 취득하고 이듬해 피사 대학교의 역학 교수가 되었다. 1892년 토리노 대학교에 부임한 그는 1900년 로마 라 사피엔차 대학교의 수리물리학 교수로 자리를 옮겼다. 1905년에는 상원의원이 되었다. 그가 로마와 해외 대학에서 행한 많은 강연이 책으로 출판되었다. ≪수리물리학의 최근 발전에 관한 세 강의≫(클라크 대학, 1909), ≪적분 및 적분-미분 방정식에 관한 강의≫(로마, 1910), ≪일차함수에 관한 강의≫(파리, 1912), ≪가환함수론≫(프린스턴, 1912) 등이 그것이다. 제1차 세계대전 중에는 이탈리아 장교로 복무하며 전쟁 발명 부서를 이끌었다. 전쟁이 끝난 후 이탈리아 수학연합(1922년)과 이탈리아 국립연구위원회(1923년) 설립에 적극적으로 참여하였고, 국립연구위원회의 초대 의장이 되었다. 또한 국제 지중해 과학 위원회(1923)와 린체이 아카데이(1924)의 회장을 역임했다. J. 페레스(J. Pérès)와 공동 집필한 또 다른 저서 ≪합성과 가환함수에 관한 강연≫은 1924년에 출판되었다.

1925년 65세의 나이에 볼테라는 훗날 사위가 된 동물학자 움베르토 단코나(D'Ancona)의 연구에 관심을 갖게 되었다. 이는 1905년부터 1923년까지 아드리아해의 세 항구인 트리에스테, 피우메(현재 크로아티아의 리예카), 베네치아에서 어획된 연골어류(상어, 가오리 등)의 비율에 대한 것이다. 단코나는 어획량이 줄어든 제1차 세계대전 기간 동안 이 어류의 비율이 증가했다는 사실을 발견했다(표 13.1).

연골 어류는 작은 물고기의 포식자이기 때문에 어획 노력이 감소하면 포식자 종에 유리한 것으로 보였다. 로트카의 연구에 대해 몰랐던 볼테라는 동일한 모델을 사용하여 이 현상을 설명했다.

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

그림 13.3: 1900년 케임브리지 대학교에서 명예 박사 학위를 받은 볼테라(1860-1940).

표 13.1: 제1차 세계대전 전후 트리에스테, 피우메, 베네치아의 어획량 중 연골어류의 비율.

| 연도 | 1910 | 1911 | 1912 | 1913 | 1914 | 1915 | 1916 |
|---|---|---|---|---|---|---|---|
| 트리에스테 | 5.7 | 8.8 | 9.5 | 15.7 | 14.6 | 7.6 | 16.2 |
| 피우메 | - | - | - | - | 11.9 | 21.4 | 22.1 |
| 베니스 | 21.8 | - | - | - | - | - | - |

| 연도 | 1917 | 1918 | 1919 | 1920 | 1921 | 1922 | 1923 |
|---|---|---|---|---|---|---|---|
| 트리에스테 | 15.4 | - | 19.9 | 15.8 | 13.3 | 10.7 | 10.2 |
| 피우메 | 21.2 | 36.4 | 27.3 | 16.0 | 15.9 | 14.8 | 10.7 |
| 베니스 | - | - | 30.9 | 25.3 | 25.9 | 25.8 | 26.6 |

여기서 $x(t)$는 피식자 수를, $y(t)$는 포식자의 수를 나타낸다. 그는 로트카와 마찬가지로 이 시스템이 초기 조건 $(x_0, y_0)$에 의존하는 주기 $T$의 주기성을 가지고 진동할 수 있다는 것을 알아냈다. 그는 다음의 사실도 발견했다.

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

이를 한 주기 $T$에 걸쳐 적분하면 ($x(0) = x(T)$와 $y(0) = y(T)$가 되도록) 다음과 같은 결과를 얻는다.

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

따라서 한 주기 동안 피식자 수와 포식자 수의 평균은 초기 조건에 무관하다. 또한, 어획 노력이 감소하면 피식자의 성장률 $a$는 증가하고 포식자의 사망률 $c$는 감소한다. 따라서 $x(t)$의 평균은 감소하고 $y(t)$의 평균은 증가하여 포식자의 비율이 증가한다. 이것이 바로 아드리아해의 어업 통계에서 관찰된 내용이다.

볼테라는 1926년 이탈리아어로 논문을 처음 발표했다. 몇 달 후 네이처에 영문 요약본이 게재되었다. 로트카는 볼테라와 다른 과학자들에게 포식자-피식자 시스템에 대한 자신의 연구가 선행했음을 알렸다. 그러나 그의 1920년 논문과 1925년 저서가 항상 언급되지는 않았다. 당시 로트카는 이미 보험회사에서 일하고 있었기 때문에 그의 연구는 인간 인구학에 집중되어 있었다. 볼테라는 10년 동안 포식자-피식자 시스템의 변형에 대한 연구를 계속했다. 그는 1928-1929년 파리에 새로 설립된 앙리 푸앵카레 연구소에서 연강을 했다. 이 강의의 노트는 1931년 ≪삶을 위한 투쟁의 수학적 이론에 관한 강연≫이라는 제목으로 출판되었다. 1935년 볼테라는 움베르토 단코나와 공동으로 ≪수학적 관점에서 본 생물학적 연관성≫에 관한 또 다른 책을 출판했다.

포식자-피식자 모델이 어업 데이터를 정확하게 설명하는 것처럼 보이지만, 생태학에서 단순화된 모델의 현실성에 관한 논쟁은 이제 막 시작되었으며 여전히 과학적 논쟁의 대상이 되고 있다. 오늘날 포식자-피식자 모델은 로트카-볼테라 모델이라고도 불리며 생태학에서 가장 흔히 인용되는 모델 중 하나이다.

1931년 볼테라는 무솔리니(Mussolini)에게 충성을 바치기를 거부했다. 그는 로마 대학에서 교수직을 잃었고 그가 가장 유명한 회원 중 한 명이었던 이탈리아 과학 아카데미에서 배제되었다. 그 후 그는 주로 이탈리아 이외의 지역에 머물며 유럽을 여행하고 강연을 했다. 그는 J. 페레스와 함께 ≪일반 함수론≫(1936)을, B. 호스틴스키(B. Hostinský)와 함께 ≪무한소 선형 연산≫(1938)을 출간했다. 그는 1940년 로마에서 사망했다.

# 참고문헌


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, p. 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)

14장

# 피셔와 자연 선택 (1922)

1922년 영국의 수학 생물학자 로널드 피셔는 인구 유전학에 관한 매우 영향력 있는 논문을 발표했다. 이 장에서는 자연 선택을 포함한 하디-바인베르크 모델의 변형에 초점을 맞춘 이 논문의 한 부분만 살펴본다. 피셔는 이형접합형이 유리하면 두 대립유전자가 공존할 수 있음을 보여주었다. 두 동형접합형 중 하나가 유리하면 다른 대립유전자는 사라진다. 중요한 문제는 일부 유전자가 여러 대립유전자를 가질 수 있는 메커니즘을 설명하는 것이다.

로널드 에일머 피셔(Fisher)는 1890년 런던에서 6남매 중 막내로 태어났다. 그의 아버지는 경매업자였으나 나중에 파산했다. 피셔는 1909년부터 1913년까지 케임브리지 대학교의 곤빌 앤드 카이우스 칼리지에서 수학과 물리학을 공부했다. 당시 유전학은 빠르게 발전하고 있었다. 1911년부터 피셔는 골턴(Galton)이 창시한 우생학회에 참여했다. 그는 골턴과 멘델(Mendel)의 연구와 관련된 통계적 문제에 파고들기 시작했다. 대학 공부를 마친 후 그는 여름 동안 캐나다의 한 농장에서 일한 후 런던시의 상업 및 일반 투자 회사에서 일했다. 극심한 근시 때문에 1차 세계 대전에 자원했지만 참전할 수 없었다. 그는 이 기간 동안 고등학교에서 학생들을 가르쳤다. 여가 시간에는 농장을 돌보며 연구를 계속했다. 그는 상관 계수와 멘델 유전학을 결합하여 중요한 새로운 결과를 얻었다.

1919년 그는 농업에 중점을 둔 로담스테드 실험 연구소에서 통계학자로 일하기 시작했다.

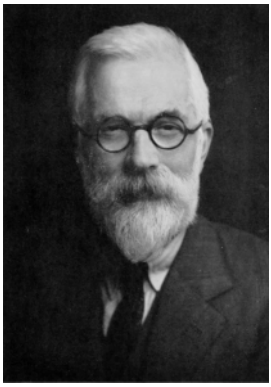

그림 14.1:
피셔 (1890–1962)

1922년 피셔는 ≪우점 비율에 대하여≫라는 제목의 논문을 발표했다. 여러 중요한 새로운 아이디어와 함께, 피셔는 멘델의 법칙과 다윈이 진화론에서 강조한 자연 선택의 개념을 결합한 수학적 모델을 다루었다. 피셔는 하디와 같이 두 개의 대립유전자 $A$와 $a$와 무작위 교배 가설이 있는 상황을 고려했다. 그러나 그는 유전자형 $AA$, $Aa$, $aa$를 가진 개체들이 성인이 되기 전에 서로 다른 사망률을 가진다고 가정하여 자연 선택을 모방하였다. $n$세대 성인의 세 가지 유전자형 빈도를 $p_n$, $2q_n$, $r_n$이라고 가정하면 $n+1$세대에 이러한 유전자형을 가진 신생아는 각각 $(p_n+q_n)^2$, $2(p_n+q_n)(q_n+r_n)$ 및 $(q_n+r_n)^2$가 된다. 출생부터 성인이 될 때까지의 각 생존 확률을 $u$, $v$, $w$라고 한다. 그러면 $n+1$ 세대에 속한 성인 개체의 유전자형 빈도 $p_{n+1}$, $2q_{n+1}$, $r_{n+1}$는 다음과 같다.

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

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

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

여기서

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

이다. $p_n + 2q_n + r_n = 1$ 였으므로, $u = v = w$ 일 때(즉, 자연 선택이 없을 때), (14.1)-(14.3)의 계는 하디가 고려한 계 (11.1)-(11.3)로 줄어든다.

$x_n = p_n + q_n$을 $n$ 세대의 성인 개체 중 대립유전자 $A$의 빈도라고 두자. 그러면 $q_n + r_n = 1 - x_n$은 대립유전자 $a$의 빈도이다. (14.1)과 (14.2)를 더하면 다음과 같다.

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

이 방정식을 다음과 같이 다시 적을 수 있다.

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

세대에 걸쳐 $x_n$ 빈도가 일정하게 유지되는 정상 상태가 항상 두 가지 이상 존재한다. 하나는 $x = 0$ (집단이 전부 동형접합형 $aa$로 구성)이고 다른 하나는 $x = 1$ (집단이 전부 동형접합형 $AA$로 구성) 이다.

방정식(14.4)을 사용하면 동형접합형 $AA$가 다른 두 유전자형보다 생존 확률이 더 높으면 ($u > v$, 그리고 $u > w$) 대립유전자 $a$가 집단에서 점차 사라질 것이라는 것을 알 수 있다. 두 대립유전자가 공존한다는 것을 알고 있다면 이 경우는 자연에서 흔하지 않을 것이다. 그러나 이형접합 $Aa$가 동형접합 $AA$ 및 $aa$에 비해 선택에서 우위를 갖는다면($v > u$, 그리고 $v > w$) 세 가지 유전자형이 집단에 공존할 수 있다. 이것이 가장 흔한 경우이며 농부들이 발견한 잡종의 '생존력'을 설명할 수 있다.

실제로 $x_n$이 1에 가까울 때 $x_{n+1} - x_n \approx (1-x_n)(u-v)/u$ 이므로 $u > v$ 일 때 정상 상태 $x = 1$은 안정적이다. 인구는 이 정상 상태로 향하게 된다. 정상 상태 $x = 1$은 $u < v$ 일 때 불안정하며, 이 경우 세 번째

정상 상태

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

가 존재하며, $0 < x^* < 1$ 이다. 또한 이 상태가 안정적이라는 것을 보여줄 수 있다. 정상 상태 $x^*$ 는 세 가지 유전자형의 혼재에 해당한다.

따라서 멘델의 법칙과 자연 선택 가설(여기서는 세 가지 유전자형에 대한 서로 다른 생존 확률)을 결합하면 유전자형의 공존 또는 소멸이라는 두 가지 상황을 설명할 수 있다. 피셔의 뒤를 이어 이 모델은 홀데인(J. B. S. Haldane) (17장 참조)과 시월 라이트(Wright) (19장 참조)에 의해서도 개발되었다.

20장에 대비해, $A$ 가 완전히 우성이고 동형접합형 $aa$ 가 다른 두 유전자형에 비해 불리하여 $u : v : w$ 가 $1 : 1 : 1-\varepsilon$ 의 비율을 가질 경우 ($\varepsilon \ll 1$) 방정식 (14.4)이 다음이 됨을 알아두자.

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

만약 이형접합형 $Aa$ 의 생존율이 두 동형접합형의 생존율의 중간 지점에 있어 $u : v : w$ 가 $1 : 1-\varepsilon/2 : 1-\varepsilon$ 의 비율이 된다면

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

이다 (여기서, $\varepsilon \ll 1$).

피셔는 로담스테드에서 농작물 수확량과 기상 조건에 관한 장기적 데이터를 분석했다. 그러나 그는 통계적 방법론에도 큰 공헌을 했다. 1925년에는 ≪연구원을 위한 통계적 방법≫이라는 제목의 책을 출간하여 큰 성공을 거두었고 여러 번 재출간되었다. 그는 1929년 왕립 학회 회원이 되었다. 1930년 피셔는 인구 유전학의 역사에 이정표가 될 ≪자연 선택의 유전학적 이론≫을 출간했다. 1933년에는 골턴 연구소 칼 피어슨의 뒤를 이어 유니버시티 칼리지 런던의 우생학 교수로 임명되었다. 1943년에는 퍼넷(R. C. Punnett)의 후임으로 케임브리지 대학 유전학과 학과장으로 자리를 옮겼다 (11장 참조). 그는 또한 ≪실험의 설계≫(1935), ≪근친

교배 이론≫(1949), ≪통계적 방법과 과학적 추론≫(1956) 등 여러 권의 책을 출판했다. 1952년 기사 작위를 받은 그는 1959년 은퇴 후 호주에 정착하여 1962년 애들레이드에서 사망했다. 그의 업적의 다른 부분은 20장에서 다시 다룬다.

## 참고문헌

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

15장

# 율과 진화 (1924)

1924년 영국의 통계학자 율은 종은 작은 돌연변이에 의해 새로운 종을 생성하고 속은 큰 돌연변이에 의해 새로운 속을 생성할 수 있다는 진화 모델을 연구했다. 그의 목적은 속 내의 종 수 분포를 설명하는 것이었는데, 대부분의 속은 한 종만 포함하고 있고 소수의 속이 많은 종을 포함하고 있다. 율이 그의 모델에 도입한 확률적 '탄생 과정'은 계통수 및 기타 여러 분야의 연구에서 여전히 기본적인 도구로 사용되고 있다.

조지 우드니 율(Yule)은 1871년 스코틀랜드에서 태어났으며, 그의 아버지는 인도 주재 영국 행정부에서 고위직을 지냈다. 16세 때 율은 엔지니어가 되기 위해 유니버시티 칼리지 런던에서 공부하기 시작했다. 1892년 그는 방향을 바꾸어 몇 년 전에 전자파의 존재를 입증한 물리학자 하인리히 헤르츠(Hertz)의 지도 아래 본에서 1년간 연구를 수행했다. 율이 영국으로 돌아왔을 때 칼 피어슨(Pearson)은 그에게 유니버시티 칼리지의 응용수학 조교수 자리를 제안했다. 율은 피어슨의 뒤를 이어 통계학에 집중하기 시작했다. 1911년 그는 ≪통계학 입문≫을 출간했으며, 이 책은 14번이나 재판인쇄 되었다. 이듬해 그는 케임브리지 대학으로 자리를 옮겼다. 그의 연구는 통계의 이론적 측면뿐만 아니라 농업과 질병 역학에 대한 응용도 다루었다. 그는 1922년 왕립 학회 회원이 되었다.

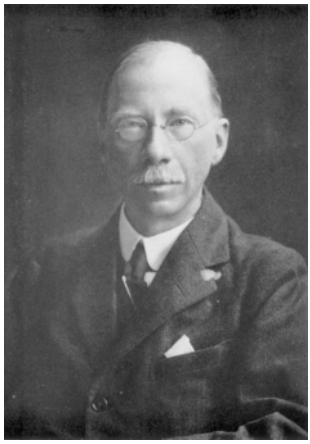

그림 15.1: 율(1871-1951)

1924년 율은 ≪윌리스 박사의 결과에 기반한 진화의 수학적 이론≫이라는 제목의 논문을 발표했다. 윌리스(Willis)는 1922년에 ≪연령과 지역, 종의 지리적 분포와 기원에 관한 연구≫라는 제목의 책을 출판한 왕립 학회 동료였다. 그는 식물과 동물의 분류에서 여러 속 중의 종의 분포를 연구했다. 그가 수집한 데이터에 따르면 대부분의 속은 한 종만 포함하고, 적은 수의 속은 몇 가지 종을 포함하고 있으며, 극소수의 속이 많은 종을 포함하고 있었다. 표 15.1는 뱀에 관한 데이터, 도마뱀에 관한 데이터, 딱정벌레의 두 과(Chrysomelidae와 Cerambycinae)에 관한 데이터를 보여준다. 당시 알려진 1,580종의 도마뱀은 259개의 속으로 분류되었는데, 그 중 105속은 1종만을, 44속은 2종만을, 23속은 3종만을 포함하는 식이었고, 100종 이상이 포함된 속은 2속이었다. 다른 동식물 과의 경우에도 포함 종 수별 속의 분포는 매우 유사한 형태를 보였다. 율은 윌리스에게 이 데이터의 그래프를 로그 스케일로 그려보라고 제안했다. 그 결과는 놀라웠다 (그림 15.2). $n$ 종을 포함하는 속의 수 $Q_n$의 로그는 $\log(n)$에 따라 선형적으로 감소하는 경향을 보였다. 즉, 어떤 상수 $\alpha > 0$와 $\beta > 0$에 대하여 $Q_n \approx \alpha n^{-\beta}$의 '멱함수 법칙'의 분포를 따르는 것이다. 1924년 논문에서 율은 이러한 통계적 분포를 설명할 수 있는 진화의 수학적 모델을 찾았다.

이를 위해 그는 먼저 한 속 내의 종 수 증가에 대한 연속시간 확률모형

표 15.1: 윌리스가 수집한 데이터.

| 종 수 | 속 수 | | | |
|---|---|---|---|---|
| | Chrysomelidae | Cerambycinae | 뱀 | 도마뱀 |
| 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 |
| 합계 | 627 | 1,024 | 293 | 259 |

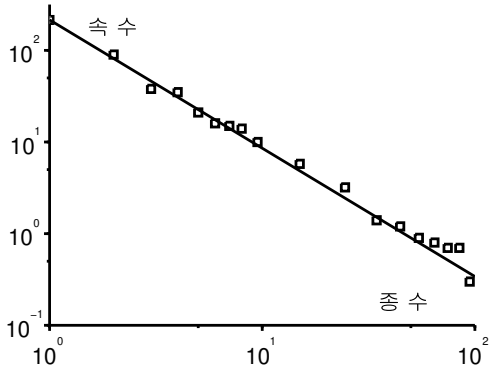


그림 15.2: 속이 포함하는 종의 수에 따른 속 수의 그래프를 상용 로그 스케일로 나타낸 것. 딱정벌레의 한 분류군(Chrysomelidae)에 대한 데이터. $n$(종의 수)이 클 때 변동을 부드럽게 하기 위해 표 15.1에서와 같이 $n$ 값의 범위에 대해 속의 수를 세었다. 따라서 특정한 $n$ 값에 대한 평균 속 수는 1보다 작을 수 있다.

을 구상했다(그림 15.3).[1] 그는 $t = 0$ 시점에 단 하나의 종으로 시작하여, 진화의 시간 척도에서 '짧은' 시간 간격 $dt$ 동안 한 종이 돌연변이를 통해 같은 속의 새로운 종을 탄생시킬 확률은 $r\,dt$ 와 같다고 가정했다 (여기서 $r > 0$).

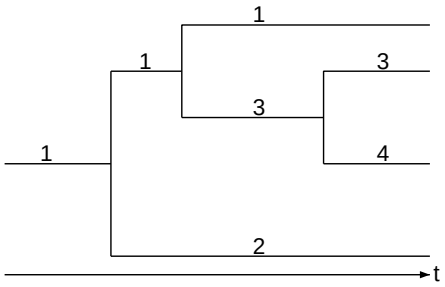


그림 15.3: 한 속 내의 종 수 진화에 대한 시뮬레이션. 종 1은 종 2와 3을 생성한다. 종 3은 종 4를 생성한다.

시간 $t$ 에 $n$ 개의 종이 존재할 확률을 $p_n(t)$ 라고 한다($n$은 정수이지만 $t$는 실수이다). 율은 $p_n(t + dt)$ 를 계산하기 위해 다음의 몇 가지 경우를

[1] 맥켄드릭(16장 참조)은 1914년에 발표한 논문에서 인구 역학에서 이러한 모델을 이미 연구하기 시작했다.

고려했다.

- 시간 $t$ 에 $n-1$ 종이 있다면, 각 종은 $t$ 와 $t+dt$ 사이에 하나의 새로운 종을 생성할 확률 $r\,dt$ 를 가지며, $dt \to 0$ 의 극한에 대해서는 시간 $t+dt$ 에 $(n-1)\,r\,dt$ 의 확률로 $n$ 종이 존재할 것이다.
- 시간 $t$ 에 $n$ 개의 종이 있다면, 시간 $t+dt$ 에는 $n\,r\,dt$ 의 확률로 $n+1$ 개의 종이 존재할 것이다.

따라서 $p_n(t)$ 는 다음과 같은 연립 미분 방정식으로 주어진다.

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

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

(모든 $n \geqslant 2$ 에 대해). 첫 번째 방정식에서 $p_1(0) = 1$ 이므로 $p_1(t) = e^{-rt}$ 를 얻는다. 초기 조건 $p_n(0) = 0$ 을 만족하는 두 번째 방정식의 해는 다음과 같다는 것을 보일 수 있다.

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

이는 모든 $n \geqslant 2$ 에 대해 성립한다 (그림 15.4). 따라서 어떤 고정된 시간 $t$ 에서 확률 $(p_n(t))_{n\geqslant 1}$ 의 분포는 두 연속 항 사이의 비율이 $1-e^{-rt}$ 와 같은 '기하학적' 분포이다.

실제로 방정식 (15.2)는 다음의 식과 같고

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

이를 이용해 $p_2(t)$, $p_3(t)$...를 연속적으로 계산할 수 있다. $p_2(t) = e^{-rt}\,(1-e^{-rt})$ 와 $p_3(t) = e^{-rt}\,(1-e^{-rt})^2$ 를 통해 일반 해에 대한 공식 (15.3)을 유추할 수 있다. 마지막으로 이 공식이 방정식 (15.4)의 해임을 확인할 수 있다.

율은 또한 식 (15.3)으로부터 종의 수에 대한 기대값이 시간이 지남에

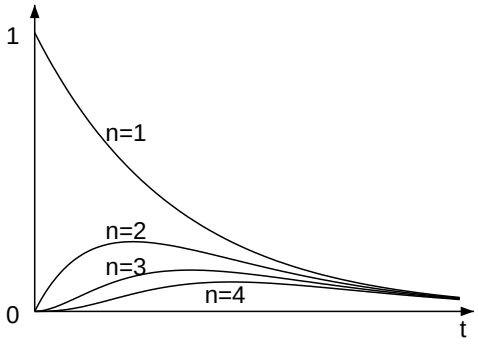


그림 15.4: 시간 $t$에 같은 속에 $n$종이 존재할 확률 $p_n(t)$. $(1 \leqslant n \leqslant 4)$

따라 다음과 같이 기하급수적으로 증가한다는 사실을 추론했다.

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

실제로 모든 $|x| < 1$에 대해

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

이므로 다음이 성립한다.

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

특히 $T$가 $e^{rT} = 2$로 정의되는 배가시간이라면, $t = T$ 시점의 종 수에 대한 확률 분포 $(p_n(t))_{n\geqslant 1}$는 비율 $1/2$를 갖는 기하 분포이다.

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

시간 $t = kT$에서는 비율 $1 - 1/2^k$를 갖는 기하 분포이며 $p_1(kT) = 1/2^k$

이다.

이어서 율은 같은 속에 속하는 종 수의 증가와 함께 더 큰 돌연변이로 인해 새로운 속이 만들어지는 유사한 과정을 생각했다. $s\,dt$ 를 짧은 시간 간격 $dt$ 동안 기존 속이 새로운 속을 생성할 확률이라고 한다. 이전과 마찬가지로, $t=0$ 시점에 속이 하나만 있다고 가정하면 $t$ 시점의 속 수의 기대값은 $e^{st}$ 이다. 시간 $t$ 에서 단위 시간당 생성된 속의 평균 개수는 미분인 $s\,e^{st}$ 이다. $t\to+\infty$ 의 극한에서, 시간 $t$ 에서 $x$ 와 $x+dx$ 시간 단위 사이에 존재했던 속의 평균 개수는 $s\,e^{s(t-x)}\,dx$ 이다.[2] 시간 $t$ 에서 무작위로 선택된 속이 $x$ 와 $x+dx$ 시간 단위 사이에 존재했을 확률은 $s\,e^{-sx}\,dx$ 이다.

시간 $t$ 에서 무작위로 선택된 속이 $x$ 와 $x+dx$ 시간 단위 사이에 존재했다면, 이 속이 $n$ 종을 가질 확률은 공식 (15.3)에 따라 $e^{-rx}\,(1-e^{-rx})^{n-1}$ 와 같다 (모든 $n\geqslant 1$ 에 대해). 따라서 시간 $t$ 에서 무작위로 선택된 속이 $n$ 종을 가질 확률 $q_n$ 은 다음과 같다.

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

여기서 $u=r/s$ 라고 두자. 간단한 계산을 통해 $q_1=1/(1+u)$ 와

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

를 보일 수 있다 (모든 $n\geqslant 2$ 에 대해).

실제로, $(1-e^{-rx})^{n-1}=(1-e^{-rx})^{n-2}\,(1-e^{-rx})$ 이다. 따라서

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

이다. 부분 적분을 통해 다음을 얻을 수 있다.

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

[2] 율은 $e^{st}$ 의 배가 시간에 비해 $t$ 가 크지 않은 경우도 고려했다. 계산은 조금 더 복잡하지만 결과는 크게 다르지 않다.

공식 (15.5)는 확률의 수열 $(q_n)_{n\geqslant 1}$이 감소하고 있음을 보여준다. 따라서 $n=1$일 때 최대값을 갖는다. 즉, 대부분의 속은 단 하나의 종만 포함한다. 이것은 정확히 데이터에서 관찰된 것이다. 또한, $q_n/q_{n-1}\to 1$이므로 $n$이 무한대로 향할 때 $q_n$의 0을 향한 감소세는 상대적으로 느리다. 이는 일부 속이 많은 종을 포함하는 이유를 설명할 수 있다. 더 정확하게는, 율은 $\log q_n$이 $\log(n)$에 대해 선형적으로 감소한다는 것을 보여주었다.

오일러의 감마 함수 $\Gamma(z)=\int_0^{+\infty} t^{z-1}e^{-t}\,dt$를 도입한다. $n$이 정수일 때 $\Gamma(n+1)=n!=n\times(n-1)\times\cdots\times 2\times 1$이고, $\Gamma(z+1)=z\Gamma(z)$이다. 따라서 (15.5)는 다음의 형태를 취한다.

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

하지만 스털링 근사에 따르면 $\log\Gamma(n)\approx n\log n-n-\frac{1}{2}\log n+\text{constant}$이다. 마찬가지로, $\log\Gamma(n+1+1/u)\approx n\log n-n+(\frac{1}{u}+\frac{1}{2})\log n+\text{constant}$이다. 따라서 $\log q_n\approx -(1+\frac{1}{u})\log n+\text{constant}$이다.

예를 들어 도마뱀의 경우를 생각해 보자. 매개변수 $u$는 한 종만 포함하는 속의 비율 $q_1=1/(1+u)$로부터 추정할 수 있다. 표 15.1에 따르면, $q_1=105/259$이므로 $u\approx 1.467$이다. 이제 이론적 확률 $q_n$를 구할 수 있고 $n$종을 포함하는 속의 예상 수 $Q_n$ 또한 $q_n$에 총 속의 수인 259를 곱하여 구할 수 있다 (표 15.2). 율은 수백만 년에 걸친 진화를 통해 종들이 겪은 대격변을 고려하지 않은 모델의 단순성을 고려할 때 관측과 계산의 일치도가 비교적 양호하다는 사실을 발견했다.[3]

1931년 이후 율은 케임브리지 대학에서 점진적으로 은퇴했다. 그는 책 저자를 식별하기 위해 문장 길이의 통계적 분포에 관심을 갖게되었다. 그는 특히 존 그란트가 출판한 책에 이 방법을 적용했는데(2장 참조), 아마 윌리엄 페티(Petty)에게서 영감을 받은 것으로 보인다. 1944년에는 ≪문학 어휘의 통계적 연구≫라는 책을 출판했다. 그는 1951년에 사망했다.

[3] 100종 이상의 종을 포함하는 속의 수에 대해, 율은 $t$가 $e^{st}$의 배가 시간에 비해 크지 않다고 가정했을 때 표 15.2보다 더 좋은 결과를 얻었다.

표 15.2: 도마뱀(259속 1,580종 분류)에 대한 데이터와 이론의 수치 비교.

| 속당 종 수 | 관찰된 속 수 | 계산된 속 수 |
|---:|---:|---:|
| 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 |
| 합계 | 259 | 259 |

오늘날에도 율의 모델은 '계통 나무'(종의 계통 나무)를 분석하는 데 사용된다. 그림 15.3과 유사한 이 나무는 분자 생물학에서 나온 새로운 데이터 덕분에 더 잘 알려져 있다. 그러나 방정식 (15.1)-(15.2)에서 정의된 확률 과정의 적용은 진화론에만 국한되지 않는다. 이 과정은 미시적 수준(예: 박테리아 군집 모델링)에서 거시적 수준(전염병의 시작 모델링)에 이르기까지 인구 역학에서 많은 모델의 기본 구성 요소이다. 이를 '순수 출생 과정' 또는 '율 과정'이라고 한다. 간단한 변형은 작은 시간 간격 $dt$ 동안 사망할 확률 $m\,dt$를 포함하며, 이 '출생 및 사망 과정'에 대한 시간 $t$의 인구의 기대값은 $e^{(r-m)t}$ 이다. 확률 분포(15.5)를 율 분포라고 부르기도 한다. 꼬리가 멱법칙을 만족하는 분포는 과학의 다양한 분야에서 많은

관심을 받고 있다. 멱함수 분포를 활용한 무작위 네트워크상의 전염병 전파 연구가 한 가지 예이다.

## 참고문헌

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

16장

# 전염병 모델링의 맥켄드릭과 커맥 (1926–1927)

1926년 맥켄드릭은 확률적 전염병 모델을 연구하여 전염병이 특정 최종 규모에 도달할 확률을 계산하는 방법을 발견했다. 또한 연속 시간 상에서 연령 구조로 구성된 인구를 설명하는 편미분 방정식도 발견했다. 1927년 커맥과 맥켄드릭은 결정론적 전염병 모델을 연구하여 인구 밀도에 대한 특정 임계값을 강조하는 최종 유행 규모에 대한 방정식을 얻었다. 대규모 전염병은 이 임계값 위에서는 발생할 수 있지만 그 아래에서는 발생하지 않는다. 이러한 연구는 현대 역학에서 여전히 많이 사용되고 있다.

앤더슨 그레이 맥켄드릭(McKendrick)은 1876년 에든버러에서 다섯 자녀 중 막내로 태어났다. 그는 글래스고 대학교에서 생리학 교수였던 아버지 밑에서 의학을 공부했다. 1900년 그는 인도 의무성에 합류했다. 인도에 가기 전에는 로널드 로스(Ross)와 함께 시에라리온에서 말라리아 퇴치 임무를 수행했다. 그 후 그는 수단에서 18개월 동안 군 복무를 했다. 인도에 도착한 그는 벵골의 한 교도소의 의사로 임명되어 이질 퇴치를 위해 노력하였다. 1905년에는 인도 북부에 있는 카사울리에 새로 생긴 중앙의학연구소에 합류했다. 그는 광견병을 연구하는 동시에 수학도

공부했다. 1920년 열대성 질병에 감염된 그는 에든버러로 돌아와 왕립 의과대학 연구소의 책임자가 되었다.

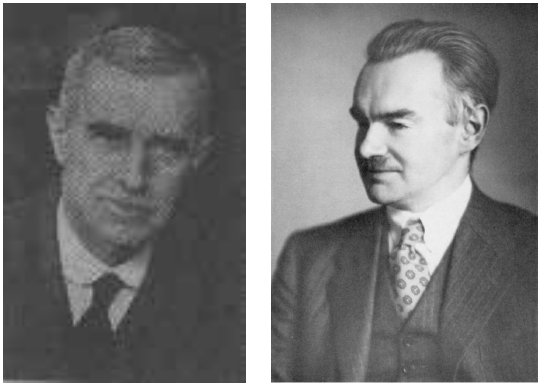

그림 16.1: 맥켄드릭(1876–1943)과 커맥(1898–1970)

1926년 맥켄드릭은 몇 가지 새로운 아이디어를 담은 ≪의학 문제에 대한 수학의 응용≫에 관한 논문을 발표했다. 그는 특히 감염과 회복의 확률적 측면을 고려한 연속시간 전염병 수리 모델을 소개했다.

처음에는 감염자가 한 명뿐인 $N$ 명의 인구가 있다고 가정해 보겠다. 사람들은 감염 가능 상태 $S$, 감염된 상태 $I$, 회복된 상태 $R$의 세 가지 상태를 단계적으로 거칠 수 있다 (그림 16.2).[1]

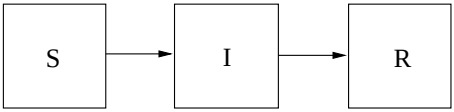


그림 16.2: 가능한 상: 감수성 있음(S), 감염됨(I), 회복됨(R).

$p_{i,r}(t)$ 는 시간 $t$에서 인구 집단이 상태 $I$에 있는 사람을 정확히 $i$명, 상태 $R$에 있는 사람을 정확히 $r$명 포함할 확률로, 여기서 $i$와 $r$은 $1 \leqslant i+r \leqslant N$의 정수이다. 이 경우 집단을 상태 $(i,r)$에 있다고 한다. 감염에 감수성이 있는 사람의 수는 $s = N-i-r$ 이다. 말라리아에 대한 로스의

[1] 다니엘 베르누이의 모델(4장 참조)에는 상태 $S$와 $R$이 포함되었지만 $I$는 포함되지 않았는데, 감염 기간이 평균 수명에 비해 훨씬 짧기 때문이다.

연구(12장 참조)를 따라 맥켄드릭은 작은 시간 간격 $dt$ 동안 한 번의 새로운 감염이 발생할 확률은 $a\,s\,i\,dt$ (즉, 감수성 있는 사람의 수와 감염된 사람의 수에 모두 비례)와 같다고 가정했다. 회복이 한 번 일어날 확률은 $b\,i\,dt$와 같다. $a$와 $b$는 모두 양의 매개변수이다. $p_{i,r}(t+dt)$를 계산하려면 몇 가지 경우를 구분해야 한다.

- 시간 $t$에서 인구가 $(i-1,r)$ 상태에 있고, $t$와 $t+dt$ 사이에 한 번의 새로운 감염이 일어나 인구가 $(i,r)$ 상태로 이동하는 경우. 이 사건의 확률은 $a\,s\,(i-1)\,dt$이며, $s=N-(i-1)-r$이다.
- 시간 $t$에서 인구는 $(i,r)$ 상태에 있고, $t$와 $t+dt$ 사이에 한 번의 새로운 감염으로 인구가 $(i+1,r)$ 상태로 이동하는 경우. 이 사건의 확률은 $a\,s\,i\,dt$이고 $s=N-i-r$이다.
- 시간 $t$에서 인구가 $(i+1,r-1)$ 상태에 있고, $t$와 $t+dt$ 사이에 한 번의 회복이 일어나 인구가 $(i,r)$ 상태로 이동하는 경우. 이 사건의 확률은 $b\,(i+1)\,dt$이다.
- 시간 $t$에서 인구가 $(i,r)$ 상태에 있고, $t$와 $t+dt$ 사이에 한 번의 회복이 일어나 인구가 $(i-1,r+1)$ 상태로 이동하는 경우. 이 사건의 확률은 $b\,i\,dt$이다.

따라서 맥켄드릭은 $1\leqslant i+r\leqslant N$에 대해 다음의 방정식을 얻었다.

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

우변의 첫 번째 항과 세 번째 항은 각각 $i=0$일 때와 $r=0$일 때 정의되지 않는다. 초기 조건은 $p_{1,0}(0)=1$을 제외한 모든 $(i,r)$에 대해 $p_{i,r}(0)=0$이다.

맥켄드릭은 이 모델을 사용하여 $n$명의 감염자를 발생시키고 전염병이 끝날 확률인 $t\to+\infty$일 때 $p_{0,n}(t)$의 극한을 계산할 수 있었다. 사실 계 (16.1)을 풀 필요는 없다. 감염자 수가 $i$이고 회복자 수가 $r$일 때 작은 시간 간격 $dt$ 동안 새로운 감염이 발생할 확률이 $a\,(N-i-r)\,i\,dt$이고 새

로운 회복이 발생할 확률은 $b\,i\,dt$ 라는 것을 알면 충분하다. 따라서 모든 $i \geqslant 1$ 에 대해 상태 $(i,r)$ 에서 상태 $(i+1,r)$ 로, 또는 상태 $(i-1,r+1)$ 로의 전이 확률(마르코프 연쇄 이론의 용어에 따라)은 각각 다음과 같다 (그림 16.3).

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

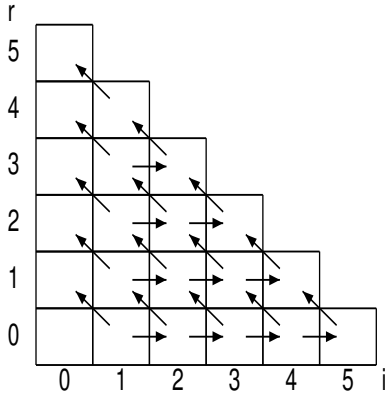


그림 16.3: $N = 5$ 인 인구의 가능한 상태(가로축은 $i$, 세로축은 $r$)와 감염(가로 화살표) 또는 회복(다른 화살표)으로 인해 발생 가능한 전이를 보여주는 도식식.

$q_{i,r}$ 를 전염병 유행 기간 동안 인구가 $(i,r)$ 상태를 거쳐갈 확률이라고 하면, $t = 0$ 일 때 $i = 1$ 이고 $r = 0$ 이므로 $q_{1,0} = 1$ 이 된다. 다른 상태는 감염 또는 회복 후에 도달한다.

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

$i = 0$ 또는 $i = 1$ 일 때 우변의 첫 번째 항은 존재하지 않는다. 두 번째 항은 $r = 0$ 일 때 없다. 이 공식으로부터 먼저 $(q_{i,0})_{2\leqslant i\leqslant N}$ 를, 그 다음 $(q_{i,1})_{0\leqslant i\leqslant N-1}$, 그리고 $(q_{i,2})_{0\leqslant i\leqslant N-2}$ 등을 계산할 수 있다. 전염병이 최종적으로 $n$ 명의 사람을 감염시킬 확률은 $q_{0,n}$ 이다. 1926년에는 이러한 계산

이 상당히 지루했다. 그래서 맥켄드릭은 한 가족처럼 아주 작은 인구에 관한 예로 제한했다. $N=5$ 명이고 $b/a=2$ 인 경우, 그는 표 16.1을 얻었다. 가장 높은 확률은 가족 중 한 사람만 감염된 경우 또는 가족 전체가 감염된 경우에 해당된다.

표 16.1: $b/a=2$ 일 때 전염병이 5인 가족의 $n$ 명을 감염시킬 확률.

| $n$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $q_{0,n}$ | 0.33 | 0.11 | 0.09 | 0.13 | 0.34 |

1926년에 발표된 동일한 논문에는 시간을 연속 변수로 간주할 때 인구 문제를 새롭게 공식화한 내용도 포함되어 있다. 무한히 작은 $dx$의 경우, $P(x,t)\,dx$를 시간 $t$에서 $x$와 $x+dx$ 사이의 연령을 가진 인구라고 하자. $m(x)$는 연령 $x$의 사망률이라고 한다. 그러면 $h$가 무한히 작은 경우 $P(x+h,t+h)\approx P(x,t)-m(x)\,P(x,t)\,h$ 이다. 함수 $P(x,t)$의 편도함수

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

를 생각해 보자.

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

임을 이용하여 맥켄드릭은 다음의 편미분 방정식을 얻는다.

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

이러한 방정식은 인구의 연령(25장 참조) 또는 역학의 감염 이후 시간과 같이 연속적인 변수로 구성된 인구 문제에서 자연스럽게 등장한다.

윌리엄 오길비 커맥(Kermack)은 1898년 스코틀랜드의 작은 마을에서 태어났다. 그는 애버딘 대학교에서 공부하고 옥스퍼드의 산업 실험

실에서 유기 화학 분야 연구를 시작했다. 1921년, 커맥은 에든버러 왕립 의과대학 연구소의 화학 부서를 담당하게 되었다. 1924년 에든버러 실험실에서 발생한 폭발 사고로 완전히 실명했지만 동료와 학생들의 도움으로 화학 연구를 계속했다. 또한 커맥은 전염병의 수학적 모델링을 위해 맥켄드릭과 협력하기 시작했다. 1927년부터 두 사람은 결정론적 전염병 모델을 연구한 ≪전염병의 수학적 이론에 대한 공헌≫ 시리즈를 함께 발표했다. 충분히 큰 규모의 $N$명의 인구를 가정하자. 1926년 논문에서와 같이 사람들이 감염에 감수성이 있거나 감염되었거나 회복된 상태 중 하나라고 가정한다. 질병이 치명적이라면 세 번째 상태는 사망을 의미한다. 세 가지 상태에 있는 사람의 수를 각각 $S(t)$, $I(t)$ 및 $R(t)$라고 하자. 모델(단순화된 형태의)은 다음 세 개의 미분 방정식으로 이뤄진 계다.

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

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

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

따라서 단위 시간당 새로운 감염자 수는 1926년의 확률 모델에서와 마찬가지로 감수성자의 수와 감염된 사람의 수 둘 다에 비례한다. 전염병이 시작되는 $t = 0$ 시점에는 $S(0) = N - I_0$, $I(0) = I_0$ 및 $R(0) = 0$로 몇 명의 사람이 감염되었다고 가정한다 $(0 < I_0 < N)$.

계 (16.2)-(16.4)는 닫힌 형태의 해는 갖지 않지만, 다음의 몇 가지 속성을 만족함을 증명할 수 있다.

- 총 인구 $S(t) + I(t) + R(t)$는 일정하게 유지되며 $N$과 같다.
- $S(t)$, $I(t)$ 및 $R(t)$는 음수가 아님을 유지한다 (각각은 인구이므로 음수가 아니어야 한다).
- $t \to +\infty$이면 $S(t)$는 극한 $S_\infty > 0$를 향해 감소하고, $I(t)$는 0에 가까워지며, $R(t)$는 극한 $R_\infty < N$를 향해 증가한다.

- 나아가, 극한 $S_\infty$ 은 다음의 음의 방정식

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

로부터 추정할 수 있고, 최종 유행 규모 $R_\infty = N - S_\infty$ 도 계산된다.

실제로 $\frac{d}{dt}(S+I+R) = 0$ 임을 알 수 있다. 따라서 $S(t)+I(t)+R(t) = S(0)+I(0)+R(0) = N$ 이다. 방정식 (16.2) 및 (16.3)는 다음과 같이 표현할 수 있다.

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

한쪽에서는 $S(t) = S(0)\,e^{-a\int_0^t I(\tau)\,d\tau} > 0$ 이고, 다른 쪽에서는 $I(t) = I(0)\,e^{a\int_0^t S(\tau)\,d\tau - bt} > 0$ 이 된다. 방정식 (16.2) 및 (16.4)는 함수 $S(t)$ 가 감소하고 함수 $R(t)$ 가 증가함을 보여준다(특히, $R(t) \geqslant 0$). $S(t) \geqslant 0$ 이고 $R(t) \leqslant N$ 이므로, 함수 $S(t)$ 와 $R(t)$ 는 $t \to +\infty$ 일 때 극한을 갖는다. $I(t) = N - S(t) - R(t)$ 이므로 $I(t)$ 도 $t \to +\infty$ 일 때 극한을 가지며, 이는 방정식 (16.4)을 적분하면 알 수 있듯이 0이 될 수밖에 없다. 방정식 (16.2)는 또한 $-\frac{d}{dt}[\log S] = aI$ 임을 보여준다. 이를 $t = 0$과 $t = +\infty$ 사이에서 적분하면 $\log S(0) - \log S_\infty = a\int_0^{+\infty} I(t)\,dt$ 를 얻는다. 방정식 (16.3)은 $\frac{dI}{dt} = -\frac{dS}{dt} - bI$ 로 표현할 수 있다. $t = 0$과 $t = +\infty$ 사이에서 적분하면 $I(0) = S(0) - S_\infty - b\int_0^\infty I(t)\,dt$ 를 얻는는다. 두 결과를 결합하면 공식 (16.5)을 구할 수 있으며, 이는 $S_\infty > 0$ 임을 보여준다.

초기 감염자 수 $I_0$ 가 인구 규모 $N$ 에 비해 적은 경우 (한 도시에서 전염병이 시작될 때 흔히 발생하는 경우), 식 (16.5)은 $S_\infty = N - R_\infty$ 를 사용하여 다음과 같이 다시 적을 수 있다.

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

여기에서 $\mathscr{R}_0$는 정의에 따라

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

이다. 방정식 (16.6)은 $\mathscr{R}_0 > 1$인 경우에만 양의 해를 갖는다. 따라서 커맥과 맥켄드릭은 다음의 결론에 도달한다. $\mathscr{R}_0 > 1$일 경우 전염병은 인구의 상당 수를 감염시킨다. 인구가 임계값 $N^* = b/a$ 보다 작으면 전염병이 발생할 수 없다.

인구의 크기 $N$가 이 임계값 ($N = N^* + \varepsilon$) 바로 위일 때에는 작은 규모의 전염병이 발생한다. 식 (16.6)로부터 $R_\infty \approx 2\varepsilon$ 이고, 따라서 $S_\infty \approx N^* - \varepsilon$ 이다.

실제로 $-\log(1-x) \approx x + \frac{x^2}{2}$의 근사를 사용하면 방정식 (16.6)은 다음과 같이 된다.

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

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

로스의 말라리아 모델 (12장)에서와 마찬가지로 $\mathscr{R}_0 > 1$일 조건은 간단하게 해석할 수 있다. $aN$은 전염병이 시작될 때 감염자 한 명이 단위 시간당 감염시키는 사람의 수이고 $1/b$는 평균 전염 기간이므로 $\mathscr{R}_0 = aN/b$는 전염병 전파 초기 한 명의 감염자로 인해 발생하는 평균 이차 감염자 수다.

치명적인 질병의 경우, $R(t)$는 전염병이 시작된 이후 누적 사망자 수이고 $dR/dt$는 단위 시간당 사망자 수다. 커맥과 맥켄드릭은 그들의 수리모델에서 $dR/dt$ 함수의 그래프가 전염병 곡선에서 예상할 수 있는 종 모양을 하고 있다는 사실을 발견했다 (그림 16.4).

$dR/dt$을 그리기 위해 (16.2)를 (16.4)로 나누어 $dS/dR = -aS/b$를 구한다. 따라서 $S(t) = S(0)\exp(-aR(t)/b)$가 된다. 이를 방정식 (16.4)에 대입하고 $S(t) + I(t) + R(t) = N$을 사용하면 다음의 방정식을 구할

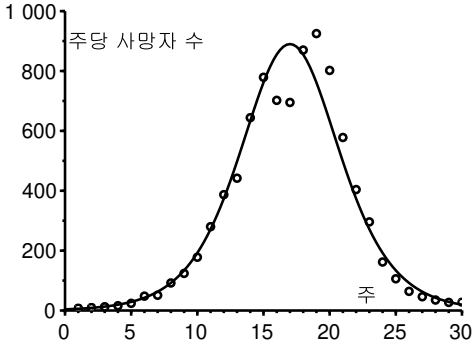


그림 16.4: 1905-1906년 봄베이에서 흑사병 유행시 주당 사망자 수에 대한 데이터와 시간에 대한 함수 $dR/dt$의 곡선.

수 있다.

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

이는 여전히 양함수 형태로 풀 수 없지만, 전체 유행 기간 동안 $\frac{a}{b}R(t)$가 작게 유지되면 $\exp(-u) \approx 1 - u + u^2/2$의 근사를 이용해

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

를 구할 수 있다. 이것은 소위 리카티 방정식으로, 양수 $R_+$와 음수 $R_-$의 두 개의 상수 해를 가지며, 이 해들은 (16.8) 우변의 $R$에 대한 2차 다항식의 근이다. $\tilde{R}(t)$를 (16.8)의 엄밀해라 하고 $Q(t) = \tilde{R}(t) - R_+$로 두자. 그러면 $Q(t)$는 다니엘 베르누이와 베르휼스트가 발견한 것과 유사한 베르누이 미분 방정식을 만족한다((4.5) 및 (6.1) 참조). 따라서 공식 (6.2)를 직접 적용하여 $Q(t)$를 구할 수 있다. 간단하지만 지루한 계산을 통해 $dQ/dt$는 $\alpha/\cosh^2(\beta t - \gamma)$의 형태라는 것을 알 수 있다 (여기서 $\alpha, \beta$ 및 $\gamma$는 모델의 매개변수에 복잡하게 의존하는 상수). $dR/dt \approx d\tilde{R}/dt = dQ/dt$이므로, 커맥과 맥켄드릭은 데이터에 맞는 $(\alpha, \beta, \gamma)$를 선택할 수 있다. 물론 최신 컴퓨터와 소프트웨어는

이러한 근사를 거치지 않고도 미분 방정식 (16.7)의 수치적 해를 쉽게 구할 수 있다.

이렇게 구한 $dR/dt$ 곡선은 1905년 12월부터 1906년 7월까지 봄베이에서 유행한 전염병의 주당 사망자 수 데이터와 잘 맞았다 (그림 16.4).

커맥과 맥켄드릭은 전염성 $a(x)$가 감염 이후 시간 $x$에 의존하고 회복률 $b(x)$도 $x$에 의존하는 보다 일반적인 모델도 고려했다. 최종 유행 규모에 대한 방정식(초기 감염자 수가 적을 때)은 여전히 (16.6)와 같지만

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

이다. 매개변수 $\mathscr{R}_0$는 전염병이 시작될 때 감염자 한 명으로 인한 평균 2차 감염자 수라는 점에서 앞의 경우와 동일한 해석을 갖는다. (16.9)와 인구에 대한 로트카(Lotka)의 공식(10.2) 의 $\mathscr{R}_0$는 유사하다. 나이는 감염 이후 시간으로, 생존율은 여전히 감염되어 있을 확률 $e^{-\int_0^x b(y)\,dy}$로, 출생률은 접촉률 $Na(x)$로 대체된다.

커맥과 맥켄드릭은 1930년대에 전염병에 대한 다수의 다른 수학적 모델을 개발했다. 이러한 모델은 오늘날 감염병 역학에서 사용되는 복잡한 모델 대부분의 뼈대가 된다. 매개변수 $\mathscr{R}_0$의 정의는 1990년 디크만(Diekmann), 히스터빅(Heesterbeek)과 메츠(Metz)에 의해 일반화 되었고, 모델 분석에서 여전히 핵심적인 역할을 하고 있다.

맥켄드릭은 1941년 은퇴하고 1943년 사망했다. 1930년부터 1933년 사이 커맥은 에든버러 대학 수학과의 윌리엄 맥크리(McCrea), 에드먼드 휘태커(Whittaker)와 함께 수리물리학에 관한 몇 편의 논문을 공동 집필했다. 1930년대와 1940년대에 커맥의 화학자 팀은 항말라리아 활성이 있는 새로운 분자를 합성하려고 시도했지만 큰 성과는 없었다. 1938년 커맥은 필립 에글턴(Eggleton)과 함께 기초 생화학에 관한 인기 도서 '우리를 이루고 있는 것들"을 공동 집필했다. 1944년 왕립 학회 회원으로 선출되었고 1949년 애버딘 대학교의 생화학과의 학과장으로 취임했다. 이후 자연과학대학 학장을 역임했다. 1968년에 은퇴한 뒤 1970년 사망했다.

# 참고문헌

1. Advisory Committee appointed by the Secretary of State for India: 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, p. 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

17장

# 할데인과 돌연변이 (1927)

1922년 논문의 다른 섹션에서 피셔는 주어진 확률 분포로 무작위 수의 자손에게 전달될 수 있는 돌연변이 유전자에 대한 문제를 고려했다. 이 문제는 형식적으로는 성의 소멸 문제와 같지만 유전적 맥락에서는 달랐다. 피셔는 확률 분포가 푸아송 분포이고 돌연변이 유전자가 선택적 우위를 갖지 않는다면 돌연변이 유전자가 개체군에서 매우 천천히 사라질 수 있음을 보여주었다. 1927년 영국의 생물학자 할데인은 이 모델에 대한 연구를 더욱 발전시켜 돌연변이에 유리한 유전자가 스스로 유지될 확률이 선택적 우위의 두 배라는 것을 보여주었다. 그는 또한 멸종 문제를 자세하게 다루었다.

존 버든 샌더슨 할데인(Haldane)은 1892년 옥스퍼드에서 태어났으며, 그의 아버지는 대학의 생리학 교수였다. 할데인은 이튼 칼리지에서 공부하고 1911년 이후 옥스퍼드 대학교의 뉴 칼리지에서 공부했다. 첫해에 수학에 집중한 그는 인문학으로 전향하였다. 그는 제1차 세계대전으로 인해 학업을 중단하였고, 그 기간 동안 프랑스와 이라크에서 복무하였다. 부상을 입은 그는 인도에 군사 교관으로 파견되었다. 1915년 그는 전쟁 전에 시작한 쥐를 대상으로 한 유전자 실험에 관한 첫 번째 논문을 발표하였다. 1919년에는 뉴 칼리지의 펠로우가 되어 생리학을 가르치며 아버지처럼 호흡을 연구하였다. 1923년에는 케임브리지 대학의 F. G.

홉킨스 생화학 연구실에 합류하여 효소의 동역학에 집중하였다.[1] 또한 공상과학 소설 ≪다이달로스 또는 과학과 미래≫(1923)와 에세이 ≪칼리니쿠스, 화학전 방어≫(1925)를 발표하였다. 1924년부터 1934년까지 ≪자연선택과 인공선택의 수학적 이론≫이라는 제목으로 10편의 논문을 연재하였다.

그림 17.1:
할데인(1892–1964)

1927년에 발표된 이 시리즈의 다섯 번째 논문에서 할데인은 피셔(Fisher)가 1922년에 연구했던 또 다른 유전 모델, 즉 돌연변이에 초점을 맞춘 모델을 재검토하였다. 피셔는 돌연변이 유전자가 집단에 침입하거나 사라질 확률을 연구하였다. 이 문제는 공식적으로는 비에나메(Bienaymé), 골턴(Galton), 왓슨(Watson)이 성의 소멸에 관해 연구한 것과 동일하였다. 피셔는 골턴의 1889년 저서 ≪자연유전≫의 부록에 실린 골턴과 왓슨의 글을 읽었을 수도 있지만 이러한 저작에 대해 언급하지 않았다. 9장에서와 같이, 어떤 유전자가 1세대에 $k$ 자손에게 전달될 확률 $(k \geqslant 0)$을 $p_k$ 라고 할때, 피셔는 $k$에 대한 상한을 고정하지 않는 생성 함수를 고려하였다. 이 합은 무한한 수의 항을 포함할 수 있다.

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

그는 0세대에 돌연변이 유전자를 가진 한 개체에서 시작하여, 이 유전

[1] 1929년 비타민 연구로 노벨 생리의학상을 수상한 프레드릭 고울랜드 홉킨스는 1929년 노벨 생리의학상을 수상하였다.

자가 $k$ 개체에 존재할 확률은 1세대의 경우 $f_1(x) = f(x)$, 2세대의 경우 $f_2(x) = f(f(x))$, 3세대의 경우 $f_3(x) = f(f(f(x)))$ 에서 $x^k$ 항의 계수라는 것을 깨달았다. 이러한 방법으로 다음 방정식이 성립한다는 것을 알 수 있다.

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

이 방정식은 왓슨이 도출한 $f_n(x) = f_{n-1}(f(x))$ 방정식보다 훨씬 더 실용적이다. 특히, (17.1)에서 알 수 있듯이 $n$ 세대 내의 멸종 확률 $x_n = f_n(0)$ 은 비에나메가 제시했던 반복식 $x_n = f(x_{n-1})$ 을 만족한다.

예를 들어 피셔는 돌연변이 유전자를 가진 식물이 $N$ 개의 씨앗을 생산할 수 있고, 각 씨앗이 생존하여 새로운 식물을 생산할 확률 $q$ 를 갖는 경우를 고려하였다. 돌연변이 유전자를 가진 $k$ 개의 자손을 얻을 확률 $p_k$ 는 다음과 같은 이항식으로 생각해 볼 수 있다.

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

$0 \leqslant k \leqslant N$ 이고 $k > N$ 의 경우 $p_k = 0$ 이다. 그러면 생성 함수는 $f(x) = (1-q+qx)^N$ 이다. 새로운 식물을 생산하기 위해 살아남는 씨앗의 평균 수를 $\mathscr{R}_0 = Nq$ 라고 할 때, $N$ 이 크고 $q$ 가 작으면 다음과 같다.

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

확률 분포 $(p_k)$ 는 $e^{-\mathscr{R}_0}(\mathscr{R}_0)^k/k!$ 를 따르고, 이를 푸아송 분포라고 한다. 그런 다음 피셔는 $x_0 = 0$, $x_n \approx e^{\mathscr{R}_0(x_{n-1}-1)}$ 와 수치 $N = 80$ 및 $q = 1/80$ 을 사용하여 $n$ 세대 내의 멸종 확률을 계산하였다. 이 경우 $\mathscr{R}_0 = Nq = 1$ 이 된다. 계산을 통해 $x_{100} \approx 0.98$ 이고, 선택적 이점이 없는 돌연변이 유전자 $(\mathscr{R}_0 = 1)$ 는 매우 느리게 사라진다는 것을 알 수 있다. 100세대가 지난 후에도 해당 유전자가 집단에 존재할 확률은 2% 이다. 1922년 피셔는 이 모델에 대한 연구를 더 이상 진행하지 않았다.

피셔의 연구를 이어받은 할데인은 1927년 논문에서 $p_0 > 0$ 인 확률 분포 $(p_k)$ 의 경우, 돌연변이 유전자를 가진 자손의 평균 수 $\mathscr{R}_0$ 가 1보다

매우 클 때, 즉 돌연변이 유전자가 선택적 우위를 가질 때 $(0,1]$ 구간에서 방정식 $x=f(x)$ 가 정확히 두 근을 갖는다는 사실을 처음 발견하였다. 또한, 소멸 확률 $x_\infty(n\to+\infty$ 일때 $x_n$ 의 한계값 $)$ 는 $x=f(x)$ 의 두 근 중 가장 작은 근이다. 이 유전자는 집단에 정착할 확률이 0이 아니다. 비에나메와 쿠르노와 달리 할데인은 이 결론에 대한 증명을 제시하였다.

실제로 $[0,1]$ 간격의 모든 $x$에 대해 $f'(x)\geqslant 0$ 및 $f''(x)\geqslant 0$이 된다. 즉, 함수 $f(x)$는 증가하며 볼록이다. $f(0)=p_0>0$ 및 $f'(1)=\mathscr{R}_0=p_1+2p_2+3p_3+\cdots>1$ 가정은 $f(x)=x$ 방정식이 $(0,1]$ 구간에서 $x=1$과 $x^*$의 정확히 두 가지 해를 가지며, $0<x^*<1$이라는 것을 암시한다. 할데인은 1883년 가브리엘 코니그스의 논문을 언급했는데, 이 논문은 $x_n=f(x_{n-1})$ 이고 $x_n\to x_\infty$ 라면 $x_\infty=f(x_\infty)$ 와 $|f'(x_\infty)|\leqslant 1$ 이라는 것을 보여줬다. $f'(1)>1$ 일 때, 유일한 가능성은 $x_\infty=x^*$ 이다.

$f(x)=e^{\mathscr{R}_0(x-1)}$ 이고 $\mathscr{R}_0$ 이 1보다 약간 큰 푸아송 분포의 경우, 소멸 확률$(x_\infty)$은 1에 매우 근접한다. 방정식 $f(x_\infty)=x_\infty$ 를 보면 알 수 있다.

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

이것은 $1-x_\infty\approx 2(\mathscr{R}_0-1)$과 같다. 할데인은 돌연변이 유전자가 멸종하지 않을 확률은 선택적 우위 $\mathscr{R}_0-1$의 두 배라는 결론을 내렸다. 피셔는 할데인을 인용하지 않고 1930년 저서에서 돌연변이 유전자가 멸종하지 않을 확률이 2% 인 $\mathscr{R}_0=1.01$ 인 경우를 예로 들었다.

할데인은 1932년 왕립 학회 회원이 되었다. 캠브리지를 떠나 런던 유니버시티 칼리지의 유전학 교수가 되었고, 이후에는 생체 인식학 교수가 되었다. 당시 그는 돌연변이율 추정, 염색체 유전자 지도 등 인간 유전학에 특히 관심이 많았다. 1927년 줄리언 헉슬리(Huxley)와 함께 쓴 ≪동물 생물학≫, 1930년 ≪효소≫, 1932년 ≪진화의 원인≫, 1954년 ≪유전학의 생화학≫ 등의 과학 저서 외에도 언론에 과학 관련 기사(예를 들어 생명의 기원) 및 에세이(1932년 ≪인간의 불평등≫, 1935년 ≪한 생물학자의 철학≫, 1938년 ≪마르크스주의 철학과 과학≫, 1938년 ≪유

전과 정치≫, 1947년 ≪과학의 진보≫)를 발표하였다. 내전 중 스페인을 여러 차례 방문한 후, 그는 자국에 공습에 대비한 피난처를 건설하도록 설득하려고 노력하였다. 제2차 세계대전 중에는 잠수함의 호흡 문제를 연구하였다. 1942년부터 공산당 당원이었던 그는 1950년 리센코의 영향으로 소련에서 멘델 유전학을 공식적으로 거부하는 바람에 사임하였다. 1957년 인도에 정착한 그는 처음에는 캘커타의 인도 통계연구소에서, 나중에는 부바네스와르에서 연구를 계속하였다. 인도 시민이 된 그는 1964년에 사망하였다.

## 참고문헌

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)

18장

# 멸종 문제에 대한 얼랑과 스테펜센의 견해 (1929–1933)

1929년 덴마크의 전화 기술자 얼랑은 성의 소멸 문제를 다시 한 번 고려하였다. 그의 동료 통계학자 스테펜센은 이 문제에 대한 완전한 해결책을 찾아냈다. 그는 특히 각 세대의 자손 수에 대한 기대치가 기하급수적으로 증가하여 확률론적 인구 모델과 결정론적 인구 모델 사이의 격차를 해소한다는 것을 보여주었다.

아그너 크라우프 얼랑(Erlang)은 1878년 덴마크 론보르그에서 태어났다. 그의 아버지는 교사였다. 1896년부터 1901년까지 얼랑은 코펜하겐 대학에서 수학, 물리학, 화학을 공부하였다. 그 후 고등학교에서 수년간 교편을 잡으며 수학, 특히 확률 이론에 관심을 가졌다. 그는 코펜하겐 전화 회사의 수석 엔지니어이자 아마추어 수학자였던 옌센(Jensen)을 만나 1908년 회사의 새로운 연구소에 합류하도록 설득하였다. 얼랑은 전화 통화 관리에 확률 이론을 적용하는 논문을 발표하기 시작하였다. 1917년 그는 대기 시간에 대한 공식을 발견했고, 이 공식은 전 세계 전화 회사에서 빠르게 사용되었다. 덴마크어로 처음 출판된 그의 논문은 이후 여러 다른 언어로 번역되었다.

1929년 얼랑은 비에나메(Bienaymé), 골턴(Galton), 왓슨(Watson)이

그림 18.1: 얼랑 (1878–1929) 과 스테펜센 (1873–1961).

가족 이름의 맥락에서, 피셔(Fisher)와 할데인(Haldane)이 돌연변이 유전자의 맥락에서 연구했던 것과 동일한 멸종 문제에 관심을 갖게 되었다. 이전 연구자들과 마찬가지로 그는 이미 출판된 연구 결과를 알지 못하였다. 한 개체가 $k$와 같은 수의 자손을 가질 확률 $p_k$에서, 얼랑은 그 가계가 $n$ 세대 내에 멸종할 확률 $x_n$에 대해 다음과 같은 등식이 성립한다는 것을 알아냈다.

$$x_n = p_0 + p_1 x_{n-1} + p_2 {x_{n-1}}^2 + \cdots = f(x_{n-1}), \quad x_0 = 0.$$

$x_\infty$를 $n \to +\infty$일때 $x_n$의 극한이라고 할때 이것이 최종 소멸 확률이다. 얼랑은 $x_\infty$가 $x_\infty = f(x_\infty)$ 방정식의 해라는 것을 알아냈다. 그는 $x = 1$이 항상 해이며, 자손의 평균 수 $\mathscr{R}_0 = f'(1)$가 1보다 클 때 0과 1 사이에 다른 해가 존재한다는 것을 깨달았다. 그러나 그는 이 두 해 중 어떤 것이 옳은지 알아낼 수 없었던 것 같다. 갤턴과 마찬가지로 그는 1929년 이 문제를 덴마크의 수학 저널인 ≪Matematisk Tidsskrift≫에 제출하였다.

> "문제 15. 한 개인이 $k$의 자녀를 가질 확률이 $p_k$이고, 여기서 $p_0 + p_1 + p_2 + \cdots = 1$ 일 때, 그의 가족이 멸종할 확률을 구하여라."

안타깝게도 얼랑은 같은 해인 1929년 51세의 나이로 세상을 떠났다. 사

실 그는 자식 없이 죽었다.[1]

코펜하겐 대학의 보험수학 교수인 요한 프레데릭 스테펜센(Steffensen)은 얼랑의 질문을 받아들였다. 그는 1930년 같은 덴마크 저널에 멸종 확률 $x_\infty$는 항상 닫힌 구간 $[0,1]$에서 방정식 $x = f(x)$의 가장 작은 근이라는 해법을 발표했는데, 이는 이미 비에나메와 할데인이 알고 있던 사실이었다. 스테펜센의 증명은 현대 교과서에서 찾아볼 수 있는 증명이다.

실제로, 우리는 소멸 확률 $x_\infty$가 닫힌 구간 $[0,1]$에서 $x = f(x)$의 해임을 확인하였다. $x^*$를 이러한 해의 최소값이라고 가정한다. 정의에 따르면, $x^* \leqslant x_\infty$이다. 스테펜센은 먼저 $x^* = f(x^*) \geqslant p_0 = x_1$이라는 것을 알아냈다. 귀납적으로 $x^* \geqslant x_n$이라고 가정할 때 $x^* = f(x^*) \geqslant f(x_n) = x_{n+1}$ 함수가 증가하기 때문에 $f(x)$가 증가한다. 따라서 모든 $n$에 대해 $x^* \geqslant x_n$이다. 극한을 취하면 $x^* \geqslant x_\infty$가 된다. 따라서 $x_\infty = x^*$이다.

스테펜센은 자손의 평균 수 $\mathscr{R}_0 = f'(1)$가 1보다 작거나 같을 때 왜 $x = 1$이 $x = f(x)$의 유일한 근인지(그림 18.2a)와 $\mathscr{R}_0 > 1$의 경우 왜 $x = 1$과 다른 근이 하나만 있는지에 대해서도 보다 공식적인 설명을 하였다(그림 18.2b). 수학적으로 $\mathscr{R}_0 = f'(1)$은 $x = 1$에서 함수 $f(x)$의 기울기임을 알 수 있다.

그는 $x = f(x)$의 모든 근에 대해 다음과 같은 사실을 발견하였다.

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

$x \neq 1$이라고 가정하고 $1 - x$로 나누면 다음과 같이 된다.

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

[1] 그의 업적을 기리기 위해 1946년 국제 전화 자문 위원회는 전화 트래픽의 강도를 측정하는 단위를 '얼랑'이라고 부르기로 결정하였다. '얼랑'은 에릭슨이라는 회사에서 프로그래밍 언어에 붙인 이름이기도 하다.

$x$가 0에서 1로 증가하면 방정식 (18.1)의 오른쪽은 $1-p_0$에서 $\mathscr{R}_0 = f'(1)$로 증가한다. 만약 $\mathscr{R}_0 < 1$이면, 방정식 (18.1)은 해가 없다. 만약 $\mathscr{R}_0 \geqslant 1$이고 $p_1 = 1$인 경우를 제외하면, 방정식 (18.1)의 오른쪽은 $x$의 순증가 함수이다. 그렇지 않으면 $p_k \neq 0$이 되는 $k \geqslant 2$가 없을 것이고, $\mathscr{R}_0$은 $p_1 < 1$과 같을 것이다. 결론적으로, (18.1)는 $\mathscr{R}_0 \geqslant 1$일 때 $[0,1]$ 구간에서 오직 하나의 해가 존재한다.

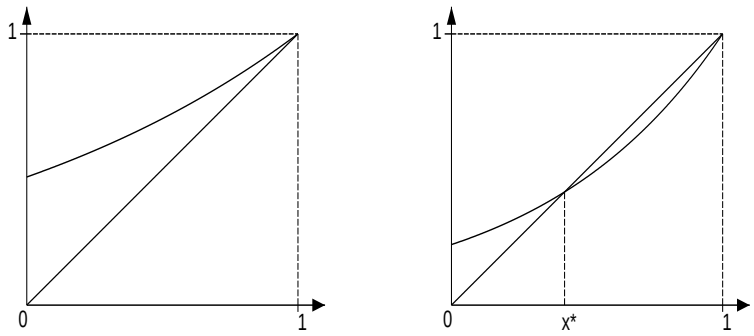


그림 18.2: 17장의 예제에서 $y = x$ 및 $y = f(x)$ 함수의 그래프를 나타낸다. $f(x) = e^{\mathscr{R}_0(x-1)}$, $\mathscr{R}_0 = 0.75 < 1$(왼쪽) 또는 $\mathscr{R}_0 = 1.5 > 1$(오른쪽).

덴마크 보험계리학회와 덴마크 수학학회 회장이기도 했던 스테펜센은 1930년 런던 대학교에 초빙되었다. 그의 영국 동료인 W. P. 엘더튼(Elderton)은 그에게 골턴과 왓슨의 연구에 대해 이야기하였다. 1933년 스테펜센은 1931년 강연을 했던 앙리 푸앵카레 연구소 연보에 새로운 논문을 발표하였다. 그는 자신의 논문 결과를 덴마크어로 요약하여 왓슨의 결과와 비교하였다. 그는 또한 $n$ 세대의 자손 수에 대한 수학적 기대치가 $(\mathscr{R}_0)^n$과 같다는 것을 보여주었다.

0세대의 한 개체에서 시작하여 $n$번째 세대에 $k$ 자손이 있을 확률을

$p_{k,n}$이라 하고, $n$번째 세대에 대한 생성 함수를 $f_n(x)$라 하자.

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

1930년 논문에서 스테펜센은 이전 연구자들과 마찬가지로 $f_1(x) = f(x)$이며 다음 방정식이 성립한다는 사실을 알아냈다.

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

$M_n$을 $n$번째 세대의 자손 수에 대한 기대치라 하면 다음과 같은 식을 유도할 수 있다.

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

공식(18.2)을 이용하면 $f_n'(x) = f'(f_{n-1}(x)) \times f_{n-1}'(x)$을 얻을 수 있다. 따라서 $M_n = f_n'(1) = f'(f_{n-1}(1)) \times f_{n-1}'(1) = f'(1) \times M_{n-1} = \mathscr{R}_0 \times M_{n-1}$. $M_1 = f_1'(1) = f'(1) = \mathscr{R}_0$이므로, 모든 $n$에 대해 $M_n = (\mathscr{R}_0)^n$이다.

따라서 예상되는 자손의 수는 $\mathscr{R}_0$이 1보다 크거나 작은지에 따라 기하학적으로 증가하거나 감소한다. 예상되는 자손의 수는 오일러(Euler), 맬서스(Malthus) 등이 고려한 인구 증가의 결정론적 모델에서와 같다. 그러나 $\mathscr{R}_0 > 1$일 때에도 해당 가계가 멸종할 확률은 0이 아닌 $x_\infty$이다. 이러한 가능성은 결정론적 모델에서는 발생하지 않는다.

스테펜센과 그의 전임자들이 연구한 확률과정은 여전히 많은 현실적인 인구 역학 모형의 기본 요소이다. 이 문제는 20장에서 마지막으로 한 번 더 언급할 것이다. 스테펜센은 1943년까지 코펜하겐 대학의 교수로 재직하다가 1961년에 사망하였다.

# 참고문헌

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

19장

# 라이트 및 무작위 유전적 드리프트 (1931)

1931년 미국의 생물학자 시월 라이트는 인구유전학 분야의 확률론적 모델 연구를 하였다. 이 연구에서 하디-바인베르크 법칙과 동일한 가정을 기반으로 하지만 인구가 무한하다고 가정하면 유전자형의 빈도는 더 이상 일정하지 않아진다. 두 대립 유전자 중 하나는 실제로 사라지지만 아주 오랜 시간이 지난 후에 사라질 수도 있다. 이 모델의 해석은 라이트와 피셔 사이에 여전히 논쟁의 대상이었는데, 후자는 자연 선택이 무작위성보다 진화에 더 중요한 역할을 한다고 추정하였다.

시월 라이트(Wright)는 1889년 매사추세츠에서 태어났다. 그는 아버지가 경제학을 가르치던 일리노이의 작은 대학에서 학부 과정을 마쳤다. 일리노이 대학교 어바나 캠퍼스에서 생물학 석사 학위를 받고 콜드스프링 하버 연구소에서 여름학교를 마친 라이트는 하버드 대학교에서 기니피그의 털색 유전과 관련하여 박사 학위를 받았다. 1915년부터 1925년까지 워싱턴에 있는 미국 농무부 축산국에서 기니피그를 이용한 근친교배 실험을 계속하였다. 그는 이러한 실험을 분석하기 위해 '경로계수법'을 개발하였다. 그 후 시카고 대학교 동물학과에 합류하였다.

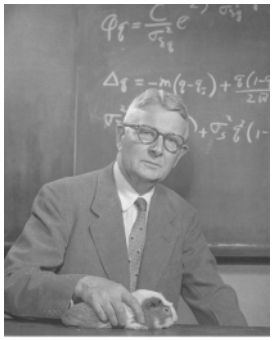

그림 19.1:
라이트 (1889–1988)

1922년 피셔(Fisher)의 인구 유전학 논문(14장 참조)에 영향을 받은 라이트는 1925년에 ≪멘델 인구의 진화≫라는 장문의 논문을 썼고, 1931년에 마침내 출판하였다. 그는 특히 피셔의 1930년 저서 ≪자연 선택의 유전적 이론≫에도 암시적으로 등장한 수학적 모델을 연구하였다. 하디-바인베르크 법칙에서와 마찬가지로, 이 모델은 하나의 유전자좌에 대해 가능한 대립유전자 $A$와 $a$가 두 개뿐인 경우를 고려하지만, 개체군이 무한히 크다고 가정하지 않는다. 요점은 이 가정을 제거하면 모집단의 유전적 구성에 어떤 영향을 미치는지 확인하는 것이다. 따라서 $N$을 모든 세대에서 동일하다고 가정한 총 개체 수라고 하자. 각 개인은 두 개의 대립 유전자를 가지고 있다. 따라서 각 세대의 인구에는 총 $2N$ 개의 대립 유전자가 존재한다. 이 모델은 또한 짝짓기가 무작위로 일어난다고 가정한다. n번째 세대에서 $A$ 유형의 대립유전자 수가 $i$이고 $a$ 유형의 대립유전자 수가 $2N-i$라면, $n+1$ 세대의 개체 중에서 무작위로 선택된 대립유전자가 $A$가 될 확률은 $\frac{i}{2N}$ 이거나 $a$가 될 확률은 $1-\frac{i}{2N}$ 이다. 따라서 $n+1$ 세대의 $A$ 대립유전자 수가 $j$가 될 확률은 다음과 같다

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

여기에서 $\binom{2N}{j}$는 이항 계수이며 $\frac{(2N)!}{j!(2N-j)!}$와 같다. $X_n$은 세대 $n$에 있는 $A$ 대립 유전자의 수로, 무작위 변수이다(그림 19.2). $X_n = i$라는 것을 알면 $X_{n+1}$의 기대값이 $i$와 같다는 것을 증명할 수 있는데, 이는 대립 유전자

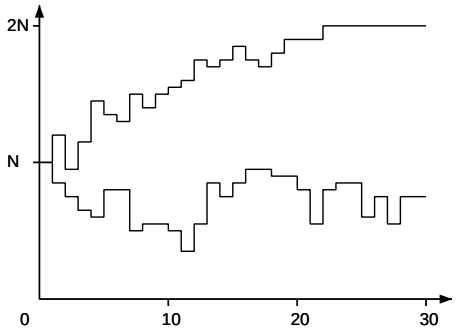


그림 19.2: $N = 20$ 및 $X_0 = 10$ 일 때 30세대 동안 $A$ 대립 유전자의 수 $X_n$의 변화를 보여주는 두 가지 시뮬레이션.

$A$의 빈도가 세대를 거쳐 일정하게 유지되는 하디-바인베르크 법칙을 연상시킨다.

실제로 다음과 같은 '생성 함수'를 고려해 보자.

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

$X_n = i$라고 가정할 때 $X_{n+1}$의 기대값은 다음과 같다.

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

그러나 이 모델에서는 $0 < i < 2N$인 초기 조건 $X_0 = i$에서 시작하여 특정 세대 후에 우연히 $X_n = 0$이라는 이벤트가 발생할 수 있다. 이 경우 모든 대립 유전자는 $a$ 유형이 되고 $X_n$은 모든 미래 세대에서 0과 동일하게 유지된다. 특정 세대 후 $X_n = 2N$인 경우 대립유전자 $A$에 대해서도 동일하게 유지된다. 요약하면, 하디-바인베르크 모델에서와 같이 인구가 무한히 많다고 가정하면 두 대립 유전자의 빈도가 일정하게 유지되므로 두 대립 유전자는 사라질 수 없다. 피셔-라이트 모델에서와 같이 인구의 유한한 크기를 고려하면 두 대립 유전자의 빈도가 변동하고 대립 유전자 중 하나가 사라질 수 있다.

$X_0 = i$ 에서 시작하면, 모집단이 $X = 0$ 상태로 고정될 확률 $Q_i$ 를 쉽게 계산할 수 있다. 실제로 $Q_i$ 는 '경계 조건'을 만족해야 한다.

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

게다가,

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

$p_{i,j} Q_j$ 는 $X_0 = i$ 에서 시작하여 $X_1 = j$ 를 통과하는 상태 $X = 0$ 에 고정될 확률임을 알 수 있다. $\sum_{j=0}^{2N} p_{i,j} = 1$ 이므로, (19.2)를 사용하면 $Q_i = 1 - \frac{i}{2N}$ 가 시스템 (19.3)-(19.4)의 해임을 알 수 있다. 따라서 $N$ 크기의 모집단에서 $A$ 유형의 대립유전자 수가 $i$ 인 상황에서 시작하여 시스템이 대립유전자 $a$ 만을 포함하는 모집단으로 진화할 확률은 $1 - \frac{i}{2N}$ 와 같다. 마찬가지로, 대립 유전자 $A$ 만 포함하는 집단으로 진화할 확률은 $\frac{i}{2N}$ 와 같다.

라이트는 두 극단 상태 중 하나에서 고정되기까지 경과하는 세대 수가 $2N$ 세대 정도라는 것을 보여주었다 (그림 19.3). 수백만 명의 개인으로 구성된 집단에서는 이 시간이 너무 길어 하디-바인베르크 법칙에서처럼 대립 유전자의 빈도가 거의 일정한 것으로 간주될 수 있다.

실제로, 0 세대의 모집단에 $A$ 유형의 대립유전자가 $i_0$ 개 있다고 가정한다. $u_i^{(n)}$ 를 $n$ 세대의 모집단에 유형 $A$ 의 대립유전자가 $i$ 개 존재할 확률이다.

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

우리는 이미 모든 $0 < i < 2N$ 에 대해 $n \to +\infty$ 일 때를 다음을 보였다.

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

라이트는 모든 $i = 1, \ldots, 2N-1$ 에 대해 $u_i^{(n)} = v$ 라면 모든 $1 < j < 2N$

에 대해 ($p_{0,j} = p_{2N,j} = 0$ 이므로) 다음과 같다는 것을 알아냈다.

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

$N$ 이 충분히 클 때,

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

적분의 값은 부분별 연속 적분을 통해 얻었다. (19.5)과 (19.6)를 결합하면 $0 < j < 2N$ 에 대해 최종적으로 다음과 같은 값을 유도할 수 있다.

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

따라서 모든 $0 < j < 2N$ 에 대한 확률 $u_j^{(n)}$ 은 세대당 약 $1/2N$ 의 비율로 감소한다. 이 속도는 $N$ 이 큰 경우 매우 느리다. 예를 들어 $N$ 이 수백만 단위인 경우에는 거의 감소하지 않는다.

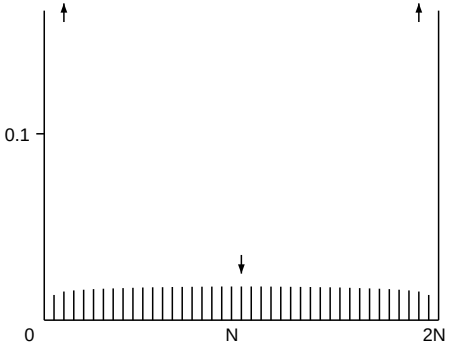


그림 19.3: $N = 20$ 이고 $X_0 = 10$ 인 경우 30 세대 후 모집단(가로축의 $i = 0, \ldots, 2N$)에 $i$ 대립유전자 $A$ 가 있을 확률이다.

1922 년에 피셔는 이미 이 고정률($\frac{1}{2N}$)을 추정하려고 시도했지만 요소

2를 놓쳤다. 어쨌든 두 과학자는 번식 개체군의 일반적인 크기 $N$에 대해 의견이 일치하지 않았다. 진화론에서 라이트의 연구는 작은 개체군에서 무작위적인 유전적 이동이 종의 기원에 대한 메커니즘이 될 수 있다고 제안하였다. 종의 분류를 연구하는 생물학자들은 실제로 종이나 아종 간의 차이가 자연 선택의 관점에서 명백하게 설명되지 않는 경우가 많다는 사실을 발견하였다. 1940년대와 1950년대에 피셔와 그의 동료인 E. B. 포드(Ford)는 자연 선택에 비해 무작위적인 유전적 이동은 무시할 수 있는 수준이라고 생각하여 이러한 생각에 강력히 반대하였다. 이들은 특히 옥스퍼드 근처에서 고립된 나방(파낙시아 도미눌라) 개체군에서 특정 유전자에 대한 세 가지 유전자형(일반 동형 접합체, 이형 접합체, 희귀 동형 접합체)을 눈으로 구분할 수 있는 유전자 주파수의 변동을 연구한 것을 언급하였다. 자연 선택과 무작위 표류의 각각의 영향에 대한 또 다른 유명한 논쟁은 '세파에아' 속의 달팽이에 초점을 맞췄다. 이제 보다 현실적인 진화 모델에는 무작위 표류, 선택, 돌연변이, 이동, 비무작위 짝짓기 등이 결합되어 있다. 무작위 표류의 역할은 나중에 일본 과학자 기무라(木村) 모토오가 '중립적 분자 진화 이론'을 통해 다시 강조하였다. 또 다른 발전은 1982년 존 킹맨(Kingman)이 소개한 유합 이론의 발전으로, 유전자의 조상을 시간을 거슬러 올라가 하나의 공통 조상이 있는 지점까지 추적하는 이론이다.

라이트는 1934년 국립과학아카데미 회원이 되었다. 그는 테오도시우스 돕잔스키(Dobzhansky)와 함께 데스밸리 지역의 자연 개체군 파리(초파리 슈도옵스큐라)의 유전학에 대해 수년간 연구하였다. 1955년 시카고 대학교에서 은퇴했지만 위스콘신-매디슨 대학교에서 5년 더 교수로 재직하였다. 1968년부터 1978년까지 ≪진화와 개체군의 유전학≫에 대한 자신의 연구를 요약한 4권짜리 논문을 발표하였다. 1984년 발잔상을 수상했으며 1988년 98세의 나이로 사망하였다.

# 참고문헌

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 (1983)
4. Malécot, G.: Sur un problème de probabilités en chaîne que pose la génétique. C. R. Acad. Sci. Paris 219, 379–381 (1944)
5. Provine, W.B.: Sewall Wright and Evolutionary Biology. Chicago (1989)
6. Wright, S.: Evolution in Mendelian populations. Genetics 16, 97–159 (1931) www.esp.org
7. Wright, S.: Evolution and the Genetics of Populations, Vol. 2. Chicago (1969)

20장

# 유전자의 확산 (1937)

1937년 로널드 피셔와 세 명의 러시아 수학자 콜모고로프, 페트로프스키, 피스쿠노프는 유리한 유전자의 지리적 확산에 대한 편미분 방정식을 연구하였다. 그들은 유전자 빈도가 유전자의 우세와 확산 계수에 따라 잘 정의된 속도로 이동하는 파동처럼 작용한다는 것을 보였다. 이들의 연구는 반응-확산 방정식 이론의 출발점이 되었다.

1937년에는 인구 역학에서 공간적 이질성 연구에 대한 새로운 접근법을 소개하는 두 편의 논문이 발표되었다. 피셔는 ≪우생학 연대기≫에 실린 ≪유리한 유전자의 진보의 물결≫이라는 제목의 첫 번째 논문의 저자였다. 그는 한 집단에서 유리한 유전자의 공간적 전파를 연구하였다. 그는 공간을 한 차원으로 축소한 후, 시간 $t$에서 점 $x$에 위치한 인구 중 유리한 유전자를 가진 인구의 비율을 $u(x,t)$라고 불렀다. 따라서 $0 \leqslant u(x,t) \leqslant 1$이다. 자연 선택을 포함하기 위해 그는 연속적인 시간 변수가 있는 방정식 (14.6)을 사용했다.

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

$a$는 양의 매개변수이다. 주어진 값 $x$에 대해, 우리는 베르휼스트(Verhulst)의 로지스틱 방정식(챕터 6 참조)을 알고 있으며, 해 $u(x,t)$는 $t \to$

$+\infty$ 일때 1에 가까워지는 경향이 있다. 또한 피셔(Fisher)는 유리한 유전자를 가진 점 $x$에 위치한 개체의 자손은 같은 지점에 머무르지 않고 $x$의 이웃에 무작위로 흩어진다고 가정하였다. 그는 물리학에 비유하여 $u(x,t)$ 방정식에 확산 항을 추가하여 다음과 같은 편미분 방정식을 유도해야 한다고 주장하였다.

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

선택 계수 $a$가 0이면 이는 푸리에(Fourier)가 열 이론에서 도입한 확산 방정식으로 축소되며, 훗날 피크(Fick)가 물리적 입자의 확산에 사용하였다. 1904년 로널드 로스(Ross)는 인구 역학에서 무작위 분산을 고려하기 시작하였다. 당시 그는 번식지로부터 거리가 멀어질수록 모기의 밀도가 어떻게 감소하는지 궁금해하고 있었다. 이 문제는 칼 피어슨(Pearson)과 레이리(Rayleigh) 경의 관심을 끌었다. 1937년까지 확산 방정식에 대한 과학 문헌은 특히 브라운 운동에 대한 아인슈타인(Einstein)의 연구 이후 상당히 증가하였다.

피셔는 세 가지 조건을 만족하는 $u(x,t) = U(x+vt)$ 형식의 방정식 (20.1)의 해가 존재한다는 것을 보였다.

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

하지만 $v \geqslant v^*$ 인 경우에만 해당된다.

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

이 해는 우성 유전자가 있는 정상 상태 $u=1$과 그러한 유전자가 없는 정상 상태 $u=0$을 연결한다. 이는 $x$의 값이 감소하는 방향으로 속도 $v$로 전파되는 파동을 나타낸다. 실제로 $u(x-vT,t+T) = u(x,t)$ 이고 시간 $t$에서 위치 $x$에 있던 파동의 일부가 시간 $t+T$에서 위치 $x-vT$로 이동한다.

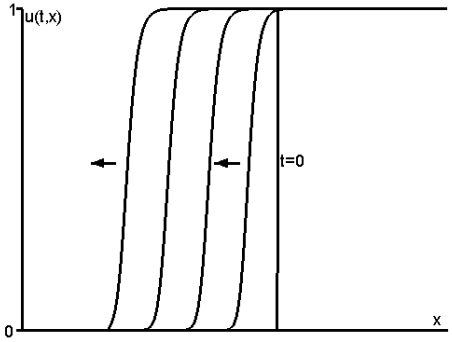


그림 20.1: 유리한 유전자의 오른쪽에서 왼쪽으로 전파되는 속도 $v^*$. $t=0$에서 유전자 빈도 $u(t,x)$는 단계 함수이다.

실제로 $z=x+vt$를 설정한 피셔는 $u(x,t)=U(z)$이면 다음과 같다는 것을 알아냈다.

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

$u$가 방정식 (20.1)의 해라면 다음과 같다.

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

$u\approx 0$에 가까울 때, 즉 '$z\to-\infty$' 일 때, 피셔는 $U(z)\to 0$과 $U'(z)\to 0$이 될 것으로 예상하였다. $z\to-\infty$일 때 $k$를 $U'(z)/U'(z)$의 극한이라고 부르면, 우리는 로피탈의 법칙으로부터 $U''(z)/U'(z)$도 $k$에 가까워진다는 것을 알 수 있다. 따라서

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

이다. 방정식 (20.2)를 $U(z)$로 나누고 $z\to-\infty$로 하면 2차 방정식 $Dk^2-vk+a=0$이 나오고 $k$는 실수여야 한다. 따라서 이 방정식의 판별자는 양수여야 한다. 즉, $v^2-4aD\geqslant 0$, 혹은 $v\geqslant 2\sqrt{aD}=v^*$

이다. 따라서 $v \geqslant v^*$는 $v$의 속도로 전파하는 파동이 존재하기 위한 필요 조건이다. 또한 아래에서 설명하는 것처럼 충분 조건이기도 하다.

피셔는 $x < 0$의 경우 $u(x,0) = 0$, $x \geqslant 0$의 경우 $u(x,0) = 1$과 같이 큰 종류의 초기 조건에 대해 정확히 $v^*$의 속도로 전파하는 파동만이 선택된다는 것을 알아냈다. 그림 20.1은 이 불연속 초기 조건이 $v^*$의 속도로 $x$가 감소하는 방향으로 점진적으로 전파되는 부드러운 파동이 되는 과정을 보여준다.

같은 해인 1937년, 피셔의 연구와는 별개로 안드레이 니콜라에비치 콜모고로프(Kolmogorov), 이반 게오르기예비치 페트로프스키(Petrovsky), 니콜라이 세메노비치 피스쿠노프(Piskunov)는 우성 유전자의 공간적 전파라는 동일한 문제를 연구하였다.

콜모고로프는 1903년 러시아 탐보프에서 태어났다. 모스크바 국립대학교에서 수학을 공부하는 동안 그는 삼각급수에 관한 중요한 연구를 수행하였다. 1929년 수학 및 역학 연구소의 연구원이 되었고 1931년 대학교 교수가 되었다. 그는 확률 과정과 미분 및 편미분 방정식과의 연관성에 대해 연구하였다. 1933년에는 확률 이론의 현대적 토대를 마련한 논문을 발표하였다. 그의 연구 관심 분야는 위상수학, 근사 이론, 마르코프 사슬, 브라운 운동, 생물학적 문제에 대한 응용 등이었다. 1935년에는 하디(Hardy), 피셔, 라이트(Wright)의 연구 결과를 논의하는 유전학에 관한 논문을 발표하였다. 1936년에는 로트트카-볼테라 시스템의 일반화에 관한 논문을 발표하였다.

페트로프스키는 1901년 세브스크에서 태어났다. 그는 또한 모스크바 국립 대학교에서 수학을 공부했으며 1933년 교수가 되었다. 그는 주로 편미분 방정식 이론과 실수 대수 곡선의 위상학을 연구했지만 일반 미분 방정식과 확률 이론에 관한 논문도 썼다. 1908년에 태어난 피스쿠노프는 모스크바 국립대학교에서 수학을 전공한 또 다른 수학자였다.

1930년대에 콜모고로프는 모스크바에서 인구 유전학의 선구자인 A. S. 세레브로프스키(Serebrovsky)와 접촉하였다. 당시 소련에서는 멘델 유전학이 '부르주아 사이비 과학'에 불과하다고 스탈린을 설득한 농학자 리센코(Lysenko)의 등장으로 멘델 유전학을 옹호하는 것이 점점 더

그림 20.2: 콜모고로프(1903–1987)와 페트로프스키(1901–1973)

위험해지고 있었다. 1937년 모스크바에서 열릴 예정이었던 제7차 국제 유전학 회의는 취소되었다. 많은 소련 유전학자들이 처형되거나 노동 수용소로 보내졌다.

콜모고로프, 페트로프스키, 피스쿠노프는 1937년 ≪모스크바 국립대학 회보≫에 실린 ≪물질의 양 증가에 따른 확산 방정식 연구와 생물학적 문제에 대한 적용≫이라는 논문에서 멘델유전학에 기반한 수학적 모델을 사용하였다. 그들의 모델은 다음과 같은 형태의 편미분 방정식이다.

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

$u(x,t)$는 다시 점 $x$와 시간 $t$에서 유리한 유전자의 빈도다. 함수 $f(u)$는 몇 가지 조건을 만족한다고 가정한다. $f(0) = f(1) = 0, 0 < u < 1$인 경우 $f(u) > 0, f'(0) > 0, 0 < u \leqslant 1$인 경우 $f'(u) < f'(0)$이다. 저자들은 피셔의 결과와 유사하지만 좀 더 엄격한 증명을 통해 초기 조건이 $0 \leqslant u(x,0) \leqslant 1$인 경우, 모든 $x < x_1$에 대해 $u(x,0) = 0$이고 모든 $x > x_2 \geqslant x_1$에 대해 $u(x,0) = 1$이라면 유전자는 $v^* = 2\sqrt{f'(0)D}$ 속도로 전파된다는 결과를 보여주었다.

$z = x + vt$인 $u(x,t) = U(z)$ 형태의 해를 구하면 방정식 (20.2)을 다음과 같이 일반화할 수 있다 $vU'(z) = f(U(z)) + DU''(z)$. 이 2차 미분

방정식은 1차 미분 방정식의 시스템으로 다시 쓸 수 있다.

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

$U(z)$는 두 가지 조건을 만족해야 한다. $z \to -\infty$일 경우 $U(z) \to 0$, $z \to +\infty$일 경우 $U(z) \to 1$. 시스템 (20.4)의 정상 상태 $(U = 0, p = 0)$ 근처에서 우리는 $f(U) \approx f'(0)U$를 가진다. 따라서 (20.4)은 선형 시스템으로 근사화할 수 있다.

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

$U(z) = U_0 e^{kz}$ 및 $p(z) = p_0 e^{kz}$ 형태의 지수 해를 구하면 피셔의 논문에서와 같이 특성 방정식 $Dk^2 - vk + f'(0) = 0$이 나온다. 여기서도 $k$는 실수여야 한다 (그렇지 않으면 $u$가 진동하여 음의 값이 된다). 따라서 $v \geqslant 2\sqrt{f'(0)D} = v^*$가 된다. 그러면 $k$의 두 근은 실수이고 양수이다. $v > v^*$이면 두 근이 서로 다르며, 정상 상태 $(U = 0, p = 0)$는 불안정한 노드다. $v = v^*$이면 그림과 같이 두 근이 동일하고 $(U = 0, p = 0)$는 불안정한 퇴행 노드이다. 마찬가지로, 정상 상태 $(U = 1, p = 0)$에 가까운 시스템(20.4)의 선형화는 다음과 같은 시스템을 유도한다.

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

특성 방정식은 $Dk^2 - vk + f'(1) = 0$이다. $f'(1) \leqslant 0$이므로 판별식은 양수다. 즉, $v^2 - 4Df'(1) \geqslant 0$ 이다. $f'(1) < 0$이면 부호가 반대인 두 개의 실근이 있고 $(U = 1, p = 0)$은 안장점이다. $f'(1) = 0$이면 한 근은 0이고 다른 근은 양수다 (그림 참조 20.3). 자세한 분석 결과, 그림의 특수한 경우에서와 같이 모든 $v \geqslant 2\sqrt{f'(0)D}$에 대해 두 정상 상태 $(U = 0, p = 0)$와 $(U = 1, p = 0)$를 연결하는 고유 적분 곡선이 존재한다는 것을 알 수 있다.

콜모고로프, 페트로프스키, 피스쿠노프는 아래와 같은 몇 가지 특성을 엄밀하게 증명하였다.

- 편미분 방정식(20.3)에는 초기 조건을 만족하는 고유한 해 $u(x,t)$가 있다.
- 이 해는 모든 $x$와 $t>0$에 대해 $0<u(x,t)\leqslant 1$이다.
- $u(x,t)$는 $t=0$에서 이미 증가 함수인 경우 $x$의 증가 함수로 남아 있다.
- $u(x,t)$는 $v^*$의 속도로 전파하는 파동 프로파일을 향해 수렴한다.

증명은 너무 길어서 여기서는 생략한다.

피셔가 사용한 함수 $f(u)=au(1-u)$는 $f'(0)=a$로 이 모든 조건을 만족한다는 것을 알 수 있다. 콜모고로프, 페트로프스키, 피스쿠노프는 방정식 (14.5)에서 영감을 얻어 동일한 조건을 만족하고 동일한 전파 속도를 제공하는 함수 $f(u)=au(1-u)^2$를 고려하였다.

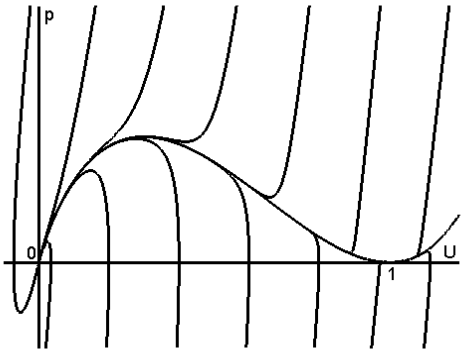


그림 20.3: 시스템 (20.5)의 일부 적분 곡선, 특히 $(U=1,p=0)$에서 $(U=0,p=0)$로 가는 유일한 곡선인 전파파의 모양을 나타내는 곡선을 보여주는 $(U,p)$ 다이어그램. 가정: $f(u)=au(1-u)^2$, $a=1, D=1$ 및 $v=v^*=2$.

피셔와 콜모고로프, 페트로프스키, 피스쿠노프의 논문은 유전학, 생태학, 역학 분야에서 지리적 확산을 고려한 많은 수학적 모델 구축의 출발점이 되었다. 이러한 모델을 '반응-확산 시스템'이라고 한다.

콜모고로프의 경우 1938년부터 비에나메(Bienaymé), 골턴(Galton),

왓슨(Watson), 피셔, 할데인(Haldane), 얼랑(Erlang), 스테펜센(Steffensen)이 고려한 이름의 소멸 문제를 연구했는데, 그는 이 모든 연구에 공통적으로 적용되는 확률과정을 '분지 과정'이라고 불렀다. 1939년 그는 소련 과학 아카데미 회원이 되었다. 이후 유체 역학의 난류 문제(1941년), 천체 역학과 관련된 동역학 시스템 이론(1953년), 정보 이론(1956년부터)에 중요한 공헌을 하였다. 또한 백과사전과 고등학교 및 대학교 교과서 집필에 기여했으며 실험 고등학교 설립을 도왔고 대중 과학 잡지를 편집하였다. 그는 1962년 발잔상, 1980년 울프상 등 많은 국제 상을 수상했으며 1987년 모스크바에서 사망하였다.

페트로프스키는 1940년 모스크바 국립대학교의 기계 및 수학 학부 학장이 되었다. 1951년부터 1973년 사망할 때까지 이 대학의 총장을 역임하였다. 1946년부터 소련 과학 아카데미의 정회원이었으며 1966년 모스크바에서 열린 국제 수학자 회의의 회장을 역임하였다. 또한 일반 미분 방정식, 편미분 방정식 및 적분 방정식에 관한 교과서를 저술하였다. 피스쿠노프는 육군 사관학교의 교수가 되었다. 미분 및 적분 미적분에 관한 그의 교과서는 많은 기술 대학에서 사용되었다. 그는 1977년에 사망하였다.

## 참고문헌


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. Méc. 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, p. 4–30. Gordon and Breach, Amsterdam (1996)
4. Pearson, K.: Mathematical Contributions to the Theory of Evolution,

XV, A Mathematical Theory of Random Migration. Dulau, London (1906) archive.org

5. Rosenfeld, B.A.: Reminiscences of Soviet Mathematicians. In: Zdravkovska, S., Duren, P.L. (eds.) Golden Years of Moscow Mathematics, 2nd edn., 75–100. Am. Math. Soc. (2007)
6. Shiryaev, A.N.: Selected Works of A. N. Kolmogorov, vol. 2. Kluwer (1992)
7. Shiryaev, A.N.: Andrei Nikolaevich Kolmogorov (April 25, 1903 to October 20, 1987). In: Kolmogorov in Perspective, 1–88. Am. Math. Soc. (2000)

21장

# 레슬리 매트릭스 (1945)

1945년 영국의 생태학자 P. H. 레슬리는 연령 구조로 구성된 설치류 집단에 대한 행렬 모델을 분석하여 로트카의 연구를 이산 시간 프레임워크에 적용하였다. 그는 성장률은 고유값에 해당하고 안정적인 연령 구조는 고유벡터에 해당한다는 점을 강조하였다. 또한 그는 갈색 쥐의 순 번식률 $\mathscr{R}_0$을 수치적으로 추정하였다.

패트릭 홀트 레슬리(Leslie)는 1900년 스코틀랜드 에든버러 근처에서 태어났다. 그는 옥스퍼드 대학교의 그리스도 교회 대학에서 공부하고 1921년 생리학 학사 학위를 취득하였다. 그러나 건강 문제로 인해 의학 공부를 마칠 수 없었다. 몇 년 동안 병리학과의 세균학 조교로 일한 후 통계학으로 전향하여 1935년 찰스 엘튼(Elton)이 설립한 새로운 연구 센터인 동물 개체수국에 합류하였다. 이 센터의 목적은 현장 연구와 실험실 실험을 통해 동물 개체군의 변동을 연구하는 것이었다. 캐나다 허드슨베이 회사의 기록 보관소를 이용한 토끼와 포식자인 스라소니의 주기 분석, 영국에서 붉은 다람쥐를 희생시킨 회색 다람쥐의 영토 확장에 대한 후속 연구, 옥스퍼드 인근의 들쥐에 대한 데이터 수집 등 대부분의 연구는 설치류에 대한 것이었다. 레슬리는 로트카(Lotka)가 인간 인구학을 위해 개발한 방법을 들쥐에 대한 데이터에 적용하였다. 2차 세계대전 중 이 센터의 연구는 사일로에서 쥐와 생쥐를 통제하는 방법에 초점을

맞추었다.

그림 21.1:
P. H. 레슬리 (1900–1972)

1945년 레슬리는 1901년 골턴(Galton), 피어슨(Pearson), 웰던(Weldon)이 창간한 저널인 ≪바이오메트리카≫에 그의 가장 유명한 논문을 발표하였다. 논문 제목은 ≪특정 인구 수학에서 행렬의 사용에 관하여≫였다. 레슬리는 동물 개체군, 예를 들어 쥐 개체군(사람일 수도 있지만)의 암컷 수 증가에 대한 모델을 고려하였다. 인구는 $K+1$ 연령 그룹으로 나뉘며, $P_{k,n}$ 은 시간 $n$ $(k=0,1,\ldots,K;\ n=0,1,\ldots)$에서 나이 $k$인 암컷의 수다. $f_k$를 나이 $k$의 출산율, 더 정확하게는 시간 $n$과 시간 $n+1$ 사이에 여성 한 명당 태어난 딸의 숫자이다. 그런 다음 $K$는 출산율이 0이 아닌 최대 연령이다 $(f_K>0)$. 나이 $k$인 동물이 적어도 나이 $k+1$까지 생존할 확률을 $s_k$라고 부른다. 그러면 모집단의 연령 구조는 다음 방정식으로 주어진다.

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

숫자 $f_k$는 양수이고, $s_k$는 $0<s_k<1$을 만족한다. 19세기 말과 20세기 초에 수학자들은 이러한 방정식 체계를 다음과 같은 축약된 형식으로

작성하였다.[1]

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

$P_n$ 는 열 벡터 $(P_{0,n}, \ldots, P_{K,n})$ 이고 $M$ 은 정사각형 행렬(즉, $K+1$ 행과 $K+1$ 열을 가진 숫자 테이블)이다.

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

레슬리는 시간의 함수로서 시스템(21.1)의 동작을 이해하기 위해 기하학적으로 증가 또는 감소하는 해 $P_n = r^n V$ 를 찾았다. 수 $r$과 벡터 $V$ 는 다음을 만족해야 한다.

$$MV = rV \,. \tag{21.2}$$

이 경우 $r$을 고유값, $V$ 를 행렬 $M$의 고유벡터라고 한다. 즉, 문제는 각 시간 단계에서 상수 $r$을 곱한 연령 분포 $V$ 를 구하는 것이다. 로트카의 용어에 따라 이러한 분포를 안정 분포라고 한다. 좀 더 일반적인 표기법으로 돌아가서 방정식 (21.2)을 다음과 같이 다시 작성할 수 있다.

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

이 방정식에서 다음과 같은 결론을 유도할 수 있다

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

이를 첫 번째 방정식에 대입하고 $V_0$ 로 단순화하고 $r^K$ 를 곱하면 다음과

[1] 즉, 모든 $k$에 대해 $P_{k,n+1} = M_{k,0} P_{0,n} + M_{k,1} P_{1,n} + \cdots + M_{k,K} P_{K,n}$ 이 된다.

같은 특성 방정식을 구할 수 있다.

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

이것은 차수 $K+1$의 변수 $r$의 다항식 방정식이다. 따라서 $K+1$개의 실근 또는 복소근 $r_1, \ldots, r_{K+1}$이 존재한다. 또한 레슬리는 (다항식에 대한 데카르트의 부호 규칙을 사용하여) 실수이자 양수인 근이 하나만 있다는 것을 알아냈다. 이를 $r_1$이라고 부른다.

레슬리는 또한 대부분의 생물학적으로 현실적인 조건(양의 행렬에 대한 페론과 프로베니우스의 이론을 사용하여 정확하게 만들 수 있음)에서 고유값 $r_1$은 다른 모든 실수 또는 복소 고유값의 모듈러스($r_2, \cdots, r_{K+1}$라고 함)보다 엄격하게 크다는 것을 제시하였다. 게다가 (21.3)의 모든 근은 일반적으로 서로 다르다. 각 고유값 $r_i$에 대해 연관된 고유 벡터를 찾을 수 있다. $Q$를 $K+1$ 크기의 정사각형 행렬로 $K+1$ 열에 각각 $r_1, \ldots, r_{K+1}$와 연관된 고유 벡터를 포함한다고 가정하면, $MQ = QD$이며, 여기서 $D$는 대각선 행렬 $[r_1, \ldots, r_{K+1}]$이다. 따라서 $M = QDQ^{-1}$ 및

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

$D^n$은 대각선 행렬 $[(r_1)^n, \ldots, (r_{K+1})^n]$이고 다음과 같은 관계식을 알 수 있다.

$$D^n / r_1^n \underset{n \to +\infty}{\longrightarrow} \mathscr{D} = [1, 0, \ldots, 0].$$

왜냐하면 $i \neq 1$에 대해 $r_1 > |r_i|$이기 때문이다. 따라서 $P_n/(r_1)^n$은 $Q\mathscr{D}Q^{-1}P_0$으로 수렴한다.

연령 구조 벡터 $P_n$의 각 구성 요소는 $(r_1)^n$처럼 증가하거나 감소한다. $r_1 > 1$이면 인구는 기하급수적으로 증가한다. $r_1 < 1$이면 기하급수적으로 감소한다. 방정식 (21.3)에서 다음 부등식이 성립하는 경우에만 $r_1 > 1$

조건이 참이라는 것을 쉽게 알 수 있다.

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

$s_0 s_1 \cdots s_{k-1}$은 최소 $k$ 나이까지 생존할 확률이다. 따라서 매개변수 $\mathscr{R}_0$은 한 여성이 평생 동안 낳은 딸의 평균 수이며 (10.2), (12.2) 및 (16.9) 공식과 유사하다. 현재의 모델은 로트카의 연구(10장 참조)와 오일러(Euler)의 연구(3장 참조)의 연령 의존적 출산을 포함한 일반화(3장 참조)의 일종의 이산 시간 형태다.

레슬리는 미국 동료가 갈색 쥐의 번식력과 생존 계수 $f_k$와 $s_k$에 대해 발표한 데이터를 사용하여 자신의 방법을 설명하였다. 합리적인 방식으로 데이터를 완성하기 위해 몇 가지 통계적 연산을 수행한 후 $\mathscr{R}_0 \approx 26$을 구하였다.

인구 역학 문제에 대한 레슬리의 행렬 공식화는 현재 많은 생물학자들이 사용하고 있다. 계산은 모든 행렬의 고유값과 고유 벡터를 계산할 수 있는 최신 컴퓨터와 과학 소프트웨어에 의해 크게 단순화되었다. 매개변수 $\mathscr{R}_0$과 성장률 $r_1$을 모두 쉽게 계산할 수 있다.

2차 세계대전 이후 레슬리는 자신의 방법을 사용하여 새, 딱정벌레 등 다른 동물 종의 성장률을 계산하였다. 그는 또한 확률론적 모델, 종 간의 경쟁 모델, 포획-재포획 데이터 분석에 관한 연구도 수행하였다. 그는 1967년에 은퇴하였다. 같은 해 찰스 엘튼도 은퇴하면서 동물 개체수국은 독립 연구 센터로서 더 이상 존재하지 않게 되었고 옥스퍼드 대학교 동물학과의 일부가 되었다. 레슬리는 1972년에 사망하였다.

## 참고문헌


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)

22장

# 퍼콜레이션 및 전염병 (1957)

1957년 해머슬리와 브로드벤트는 인접한 두 노드가 주어진 확률로 연결되는 무한 정사각형 네트워크에서 '유체'의 전파를 고려하였다. 가능한 예로 과수원에서 전염병이 전파되는 것을 설명하였다. 그들은 대규모 전염병이 발생하지 않는 임계 확률과 그 이상에서 양수 확률로 대규모 전염병이 발생하는 임계 확률이 있음을 보여주었다. 그들의 논문은 침투 이론의 출발점이었다.

존 마이클 해머슬리(Hammersley)는 1920년 스코틀랜드에서 태어났으며, 그의 아버지는 철강을 수출하는 미국 회사에서 근무하였다. 그는 케임브리지 대학의 엠마누엘 칼리지에서 공부를 시작했지만 1940년에 군대에 입대해야 했다. 그는 포병용 계산기 개선에 종사하였다. 1948년 학업을 마친 후 옥스퍼드 대학에서 실험 설계 및 분석 그룹에서 조교로 일하였다. 1955년 옥스퍼드 인근 하웰에 있는 원자력 연구소에 입사하였다.

사이먼 랄프 브로드벤트(Broadbent)는 1928년에 태어났다. 그는 케임브리지에서 공학을, 옥스퍼드의 매그달렌 칼리지에서 수학을 공부했으며 (시를 쓰기도 했다), 런던의 임페리얼 칼리지에서 ≪균일 분산으로부터의 이탈에 대한 몇 가지 테스트≫로 통계학 박사 학위를 취득하였다. 박사 과정 중에 그는 영국 석탄 이용 연구 협회의 지원을 받아 석탄 생산

그림 22.1:
해머슬리 (1920–2004)

과 관련된 통계적 문제를 조사하였다.

1954년 런던의 왕립 통계학회에서 원자력 연구소가 후원하는 몬테카를로 방법에 관한 심포지엄이 열렸다. 1940년대에 로스 알라모스 연구소의 존 폰 노이만(Neumann), 스타니슬라프 울람(Ulam), 니콜라스 메트로폴리스(Metropolis)가 시작한 이 방법은 확률론적 컴퓨터 시뮬레이션을 사용하여 미지의 수학적 양을 추정하는 방법이다. 해머슬리는 런던 심포지엄에서 하웰의 동료 모튼(Morton)과 공동으로 준비한 논문을 발표하였다. 이 논문은 ≪왕립 통계학회지≫에도 게재되었다. 심포지엄 발표 후 이어진 토론에서 브로드벤트는 몬테카를로 방법을 사용하여 연구할 수 있는 흥미로운 문제를 언급하였다. 2차원 또는 3차원의 규칙적인 기공 네트워크가 주어지고 두 개의 인접한 기공이 확률 $p$로 연결되어 있다면, 기체가 이 기공 중 하나를 통해 유입될 경우 네트워크의 몇 퍼센트가 기체로 채워질 것인가에 대한 문제였다. 브로드벤트는 사실 탄광 노동자를 위한 방독면을 설계할 때, 특히 방독면이 작동하는 데 필요한 기공의 크기에 대해 생각하고 있었다.

해머슬리는 브로드벤트와 함께 이 방독면 문제를 연구하기 시작하였다. 그들은 이 문제가 아직 연구되지 않은 문제군, 즉 임의의 매체에서 '유체'(맥락에 따라 의미가 달라짐)의 결정론적 전파의 프로토타입에 불과하다는 것을 깨달았다. 해머슬리는 이를 커피 포트에서 일어나는 일과 비유하여 '퍼콜레이션'이라고 불렀다. 원자력 연구소에서 해머슬리는 당시 가장 강력한 컴퓨터를 사용하여 퍼콜레이션 문제에 대한 몬테카를

로 방법을 테스트 할 수 있었다.

1957년 브로드벤트와 해머슬리는 마침내 침투의 수학적 이론에 관한 첫 번째 논문을 발표하였다. 그들이 고려한 예 중 하나는 과수원에서 전염병이 전파되는 인구 역학 모델이다. 매우 큰 과수원의 나무는 정사각형 네트워크의 노드에 배치되어 있다고 가정한다. 각 나무에는 가장 가까운 이웃이 네 개 있다. 감염된 나무의 주어진 이웃도 감염될 확률 $p$가 있다. 문제는 많은 수의 나무가 감염될지 아니면 전염병이 국지적으로 유지될지 여부다. 이는 물론 확률 $p$에 따라 달라지며, 이는 다시 나무를 구분하는 거리, 즉 네트워크 메시의 폭과 관련이 있다.

브로드벤트와 해머슬리는 과수원이 무한하고 전체 면을 덮고 있으며 처음에 감염된 나무가 하나뿐인 경우를 살펴봤다. $f(p)$를 이 감염원으로부터 무한한 수의 나무가 감염될 확률이라고 하면, $f(p)$는 $f(0)=0$, $f(1)=1$로 $p$의 증가 함수일 것으로 예상된다. 주요 결과는 $p^*, 0<p^*<1$의 임계 확률이 존재한다는 것이었다.

- $p<p^*$이면 $f(p)=0$이므로 유한한 수의 트리만 감염된다.
- 만약 $p>p^*$라면, $f(p)>0$이고 무한한 수의 트리가 감염될 수 있다.

이 증명은 감염원으로부터 시작하여 평면에서 서로 다른 '자기 회피 산책'의 수를 추정하는 것이다. 이 보행은 어떤 나무도 두 번 이상 방문하지 않고 특정 수의 이웃 나무를 통과한다(각 나무에는 네 개의 이웃이 있다는 것을 기억하자). 방문한 각 나무에서 다음 나무로 감염이 $p$의 확률로 전염될 수 있으므로, $n$개의 보폭을 가진 자기 회피 보행은 $p^n$의 확률로 감염 경로가 된다. 이제 $q(j,n)$를 $n$개의 보폭을 가진 모든 자기 회피 보행 중에서 감염 경로인 보행이 정확히 $j$개 있을 확률이라고 하자. 감염된 나무가 무한히 많다면, 모든 정수 $n$에 대해 감염 경로인 n보의 자기 회피 보행이 하나 이상 존재한다. 따라서

$$0 \leqslant f(p) \leqslant \sum_{j=1}^{+\infty} q(j,n) \leqslant \sum_{j=1}^{+\infty} j\, q(j,n) \quad \forall n.$$

그러나 $\sum_{j=1}^{+\infty} j\,q(j,n)$는 감염 경로가 $n$ 걸음인 자기 회피 보행의 예상 횟수이다. 이 수치는 $p^n s(n)$과 같으며, 여기서 $s(n)$은 $n$보의 자기 회피 보행의 총 횟수다. 해머슬리는 논문에서 $s(n)$이 $n \to +\infty$에 따라 $e^{\kappa n}$처럼 증가한다는 것을 보여줄 수 있는데, 여기서 $\kappa$를 연결 상수라고 부른다. $p < e^{-\kappa}$이면 $p^n s(n)$는 $n \to +\infty$에 따라 0이 되고 $f(p) = 0$이 된다. 따라서 $p^* \geqslant e^{-\kappa} > 0$이다.

따라서 실제로는 전염병이 발생할 경우 나무가 너무 가깝지 않아 $p$를 $p^*$ 이하로 유지하는 것이 좋다. 그러나 나무가 가까울수록 헥타르당 생산량이 높아진다. 타협점을 찾아야 한다.

브로드벤트와 해머슬리가 밝혀낸 것처럼, 퍼콜레이션 프로세스에서 임계 확률의 존재와 분기 프로세스에서 임계값의 존재 사이에는 어느 정도 유사성이 있다(7 장 참조).

임계 확률 $p^*$를 수치적으로 추정해 볼 수 있다. 이를 위해 $p$의 값을 고정하고 $N$이 충분히 큰 크기 $N \times N$의 유한 정사각형 네트워크로 무한 네트워크를 근사화한다. 예를 들어 네트워크 중앙에 있는 나무가 감염되었다고 가정한다. 컴퓨터를 사용하면 어떤 나무가 다른 나무를 감염시킬 수 있는지 무작위로 선택할 수 있다. 그림 22.2과 그림 22.3는 그래프 이론에서와 같이 에지를 사용하여 무작위로 선택된 감염 경로를 보여준다. 그림 22.2에서 $p$는 $p^*$보다 작다. 그림 22.3에서 $p$는 $p^*$보다 크다. 감염된 나무, 즉 중앙의 감염된 나무에서 시작하는 경로로 도달할 수 있는 나무를 쉽게 확인할 수 있다. 그림에서 작은 검은색 사각형으로 표시되어 있다.

그런 다음 전염병이 적어도 $N \times N$ 크기의 네트워크의 경계에 도달했는지 확인할 수 있다. 만약 그렇다면, 그리고 $N$이 충분히 크다면, 감염된 나무의 수는 '거의 무한대'라고 생각할 수 있다. 이런 종류의 시뮬레이션을 여러 번 반복하면 감염된 나무의 수가 무한대일 확률 $f(p)$의 근사값을 구할 수 있다(이것이 몬테카를로 방법이다). 마지막으로, $p$를 0과 1 사이에서 변화시키면 임계값 $p^*$의 근사치를 구할 수 있는데, 이는 $p > p^*$일 때 $f(p) > 0$이 되는 가장 작은 값이다.

브로드벤트와 해머슬리의 논문에는 임계값 $p^*$의 존재에 대한 증명만 포함되어 있다. 그 후 몇 년 동안 해머슬리는 침투에 대한 수학적 이론을 계속 발전시켰고, 브로드벤트는 다른 주제로 눈을 돌렸다. 1970년대에

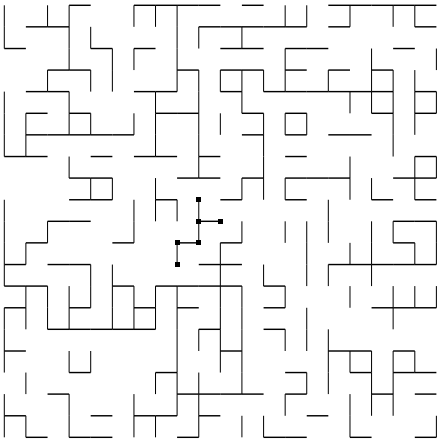

그림 22.2: $p = 0.4$를 사용한 퍼콜레이션.

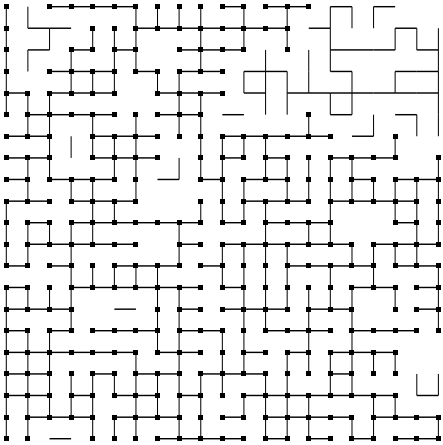

그림 22.3: $p = 0.55$를 사용한 퍼콜레이션.

컴퓨터가 개발되면서 위에서 설명한 시뮬레이션을 실행하는 것이 더 쉬워졌다(그림 22.4). 그 후 $p^* = 1/2$라고 추측하였다. 이 결과는 1980년 코넬 대학교의 해리 케스텐에 의해 마침내 증명되었다.

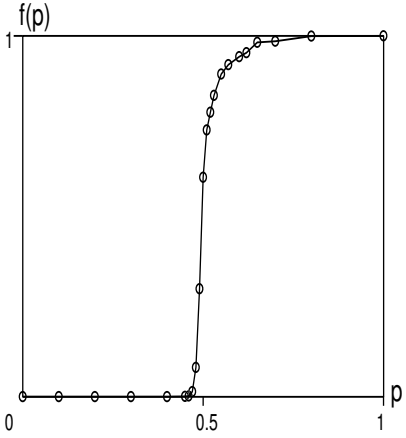


그림 22.4: 무한히 많은 나무가 감염될 확률 $f(p)$는 $p$의 함수다. 이 곡선은 $200 \times 200$ 네트워크에서 1,000회의 시뮬레이션을 실행하여 얻는다.

1959년부터 1969년까지 해머슬리는 옥스퍼드 대학교의 경제 및 통계 연구소에서 근무하였다. 그는 트리니티 칼리지의 펠로우가 되었다. 1964년에는 데이비드 핸스콤(Handscomb)과 공동으로 ≪몬테카를로 방법≫이라는 제목의 책을 출간하였다. 1976년 왕립학회 회원으로 선출되었다. 1987년에 은퇴했지만 옥스퍼드 산업 및 응용수학 센터를 계속 방문하였다. 그는 2004년에 사망하였다.

브로드벤트는 1957년 임페리얼 칼리지에서 박사 학위를 받았다. 그는 유나이티드 유리병 제조업체라는 산업 회사에 취직하였다. 산업계에서 10년을 일한 후 그는 과학적 독자층 연구를 수행하는 런던 프레스 익스체인지라는 뉴스 에이전시에서 일하기 시작하였다. 이 대행사는 1969년 미국 광고 회사인 레오 버넷(Leo Burnett)에 인수되었다. 브로드벤트는 광고의 효과를 측정하는 방법을 연구하여 ≪광고비 지출≫(1975), ≪광고 예산≫(1989), ≪책임 있는 광고≫(1997), ≪언제 광고해야 하는가≫(1999) 등 여러 권의 저서를 출간하였다. 1980년에는 광고 효과상 창설을

도왔다. 시카고에 있는 레오 버넷 본사에서 브랜드 경제학 이사로 수년간 근무하였다. 또한 자신의 컨설팅 회사인 BrandCon Limited를 운영하기도 하였다. 그는 2002년에 사망하였다.

# 참고문헌


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.campaign-live.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)

23장

# 게임 이론과 진화 (1973)

1973년 메이너드 스미스와 프라이스는 동물들이 종내 갈등에서 가장 위험한 무기를 사용하지 않는 이유를 분석하는 논문을 발표하였다. 이들의 모델은 게임 이론을 사용했으며, 이 수학적 이론을 진화 문제에 적용하기 시작한 모델 중 하나였다.

존 메이너드 스미스(Maynard Smith)는 1920년 런던에서 태어났다. 외과의사였던 아버지는 그가 여덟 살 때 돌아가셨다. 메이너드 스미스는 이튼 칼리지에서 공부한 후 케임브리지 대학교 트리니티 칼리지에서 공학을 전공하였다. 당시 그는 영국 공산당의 당원이었다. 1939년 전쟁이 발발하자 그는 군대에 지원하려 했지만 시력이 좋지 않아 거절당하였다. 그는 공학 공부를 마치고 몇 년 동안 군용 항공기 설계에 종사하였다. 마침내 그는 생물학으로 전향하여 런던의 유니버시티 칼리지에서 할데인(Haldane)을 지도교수로 모시고 유전학을 공부하기로 결정하였다. 1952년 동물학 강사가 되었다. 1956년 헝가리에서 일어난 사건 이후 공산당을 떠났다. 그의 첫 번째 저서인 ≪진화론≫은 1958년에 출판되었다. 1965년 새로 설립된 서섹스 대학교의 생물학 교수가 되었다. 그 후 두 권의 다른 책을 출간하였다. ≪생물학의 수학적 아이디어≫(1968)와 ≪진화에 대하여≫(1972)가 그것이다.

조지 R. 프라이스(Price)는 1922년 미국에서 태어났다. 시카고 대학

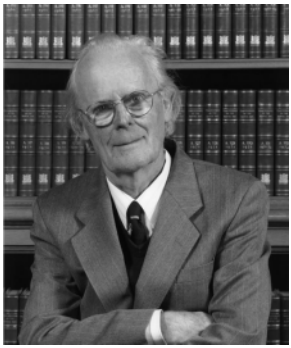

그림 23.1: 메이너드 스미스 (1920–2004) 과 프라이스 (1922–1975).

교에서 화학을 전공하고 원자폭탄 개발 프로젝트인 맨해튼 프로젝트에 참여한 후 1946년 박사 학위를 취득하였다. 1950년에는 미네소타 대학교에서 의학 부연구원이 되었다. 이후 여러 잡지의 독립 저널리스트로 활동하다가 IBM에서 다시 연구원으로 복귀하였다. 1967년 갑상선암 치료를 받은 후 영국에 정착하여 완전히 다른 분야인 진화 생물학을 연구하기 시작하였다. 그는 1968년부터 런던에 있는 유니버시티 칼리지의 골턴 연구소에서 근무하였다. 이 새로운 분야에 대한 그의 첫 번째 논문인 ≪선택과 공분산≫은 W. D. 해밀턴의 도움으로 1970년 ≪네이처≫ 호에 발표되었으며, 현재 프라이스 방정식으로 불리는 내용을 담고 있다.

프라이스는 이번에는 동물 갈등에 관한 또 다른 논문을 ≪네이처≫에 제출하였다. 하지만 이 저널에 적합한 형식이 아니었다. 그래서 심사위원이었던 메이너드 스미스는 더 짧은 버전을 준비할 것을 제안하였다. 프라이스는 다른 작업을 시작했고, 메이너드 스미스는 프라이스의 아이디어를 혼자서 발전시키기 시작하였다. 마침내 메이너드 스미스와 프라이스는 1973년 <네이처>에 ≪동물 갈등의 논리≫라는 제목의 공동 논문을 발표하였다. 이 논문은 진화 생물학에서 게임 이론을 사용하는 데 흥미로운 기여를 했다. 그 이전까지 게임 이론은 주로 경제학과 정치학, 특히 1944년 존 폰 노이만(Neumann)과 오스카 모겐슈타인(Morgenstern)의 저서 ≪게임과 경제 행동 이론≫ 이후 발전해 왔다. 메이너드 스미스와 프라이스의 출발점은 다음과 같은 질문이었다. 같은 종의 동물들 간의

충돌에서 어떻게 뿔, 발톱, 독 등 동물이 사용할 수 있는 '무기'가 살상에 거의 사용되지 않는가? 다윈의 생명 투쟁에 대한 생각에 따르면, 더 공격적인 동물은 더 많은 전투에서 승리하고 더 많은 수의 새끼를 낳아야 하며, 이는 '무기' 사용의 증가로 이어진다. 이 시기는 냉전 시대였기 때문에 이 주제는 정치적 색채도 띠고 있었다.

메이너드 스미스와 프라이스는 두 동물이 자원(예: 유리한 서식지의 영역)을 놓고 경쟁하는 일련의 게임을 상상했다. 메이너드 스미스가 1982년 저서 ≪진화와 게임 이론≫에서 사용한 단순화된 설명에서 각 동물은 '매 전략' 또는 '비둘기 전략' 중 하나를 채택한다. 앞으로 매와 비둘기로 간단히 지칭하겠지만, 이들은 같은 종의 동물이 채택하는 전략을 의미한다. $V>0$을 자원의 가치라고 하자. 즉, $\mathscr{R}_0$이 한 동물의 정상적인 평균 새끼 수라면 경쟁에서 승리한 동물은 평균적으로 $\mathscr{R}_0+V$의 새끼를 갖게 된다는 뜻이다.

매가 다른 매를 만나면 자원을 놓고 경쟁을 벌이게 되는데, 승자는 가치 $V$의 자원을 얻고 패자는 '비용' $C>0$를 부담하게 된다. 두 매는 각각 경쟁에서 이길 확률이 $1/2$이고 질 확률은 동일하다. 따라서 두 매의 싸움에서 기대되는 보상은 두 경쟁자의 경우 $\frac{1}{2}(V-C)$이다. 그러나 매 한 마리가 비둘기를 만나면 매는 자원 $V$를 얻고 비둘기는 싸우지 않고 탈출하며 비용은 0이다. 마지막으로, 두 마리의 비둘기가 만나면 한 마리는 자원 $V$를 얻고, 다른 한 마리는 싸우지 않고 0의 비용으로 탈출한다. 두 비둘기의 승리 확률은 각각 $1/2$이므로, 두 비둘기가 만났을 때 예상되는 보상은 $V/2$이다. 보상은 다음의 표23.1과 같이 요약할 수 있다.

표 23.1: 매와 비둘기 게임의 예상 결과.

| | 매를 상대로 | 비둘기를 상대로 |
|---|---|---|
| 매의 보상 | $\frac{1}{2}(V-C)$ | $V$ |
| 비둘기의 보상 | 0 | $V/2$ |

보다 일반적으로, 1과 2라는 두 가지 전략 중 하나를 채택할 수 있는 개인 간의 싸움을 예상 보상 행렬 $(G_{i,j})_{1\leqslant i,j\leqslant 2}$와 함께 상상해 볼 수 있다.

위의 예에서 매는 전략 1을 따르고 비둘기는 전략 2를 따른다.

$$G_{1,1} = \tfrac{1}{2}(V-C), \quad G_{1,2} = V, \quad G_{2,1} = 0, \quad G_{2,2} = V/2.$$

메이너드 스미스와 프라이스는 1973년에 이미 컴퓨터 시뮬레이션을 통해 두 가지 이상의 가능한 전략(매, 쥐, 가해자, 보복자, 시험자-보복자)을 테스트한 적이 있다.

이제 n세대에 매의 비율이 $x_n$ 이고 비둘기의 비율이 $1-x_n$ 인 같은 종의 큰 동물 집단을 생각해 보자. $n$ 번째 세대의 매는 평균적으로 다음과 같은 수의 새끼를 낳는다.

$$R_1(n) = \mathscr{R}_0 + x_n G_{1,1} + (1-x_n) G_{1,2}. \tag{23.1}$$

마찬가지로 비둘기의 평균 새끼 수는 다음과 같다.

$$R_2(n) = \mathscr{R}_0 + x_n G_{2,1} + (1-x_n) G_{2,2}. \tag{23.2}$$

따라서 전체 개체군의 평균 자손 수는 와 같다.

$$R(n) = x_n R_1(n) + (1-x_n) R_2(n).$$

성적 생식으로 인한 가능한 미묘함을 잊어 버리면 다음 세대의 매의 비율이 다음과 같다는 것을 알 수 있다.

$$x_{n+1} = x_n R_1(n)/R(n). \tag{23.3}$$

따라서 $R_1(n) > R(n)$ 이면 $x_{n+1} > x_n$ 이고, $R_1(n) < R(n)$ 이면 $x_{n+1} < x_n$ 이 된다. 가능한 정상 상태는 세 가지이다. $x=0$, $x=1$, 그리고

$$x^* = \frac{G_{1,2} - G_{2,2}}{G_{2,1} - G_{1,1} + G_{1,2} - G_{2,2}}$$

이다 ($0 < x^* < 1$ 인 경우에만). 매-비둘기 게임에서 $V < C$ 일 경우 $x^* = V/C < 1$ 이다.

실제로 $x=0$은 (23.3)의 명백한 정상상태다. 만약 $x\neq 0$이 또 다른 정상상태라면, $R_1 = R = xR_1 + (1-x)R_2$가 된다. 따라서 $x=1$ 또는 $R_1 = R_2$ 중 하나다. 후자는 $xG_{1,1} + (1-x)G_{1,2} = xG_{2,1} + (1-x)G_{2,2}$와 동일하며, 이로부터 정상 상태 $x^*$를 알 수 있다.

정상 상태 $x=1$은 전략 1을 따르는 개체가 100% 집단에 해당한다. 이 정상 상태는 전략 2를 따르는 소수의 개인이 침입할 수 없는 경우 안정적이다. (23.3)에서 이 조건은 1에 충분히 가까운 모든 $x_n$에 대해 $R_1(n) > R(n)$을 갖는 것과 같다는 것을 알 수 있다. $R(n) = x_n R_1(n) + (1-x_n)R_2(n)$이므로, 조건은 1에 충분히 가까운 모든 $x_n$에 대해 $R_1(n) > R_2(n)$이 된다. $R_1$과 $R_2$의 식 (23.1)-(23.2)를 살펴보면, 다음 두 조건 중 하나만 만족하면 $x=1$이 안정적이라는 결론에 도달할 수 있다.

- $G_{1,1} > G_{2,1}$.
- $G_{1,1} = G_{2,1}$ 및 $G_{1,2} > G_{2,2}$이다.

그렇다면 전략 1은 '진화적으로 안정적인 전략'이라고 할 수 있다. 매-비둘기 게임에서 $G_{1,2} > G_{2,2}$ 조건은 항상 참이다. 따라서 매 전략은 $G_{1,1} \geqslant G_{2,1}$, 즉 $V \geqslant C$일 경우에만 진화적으로 안정적이다.

정상 상태 $x=0$은 모든 개인이 전략 2를 따르는 집단에 해당한다. 이 상황은 인덱스 1과 2를 교환하면 앞의 상황과 대칭이 된다. 매-비둘기 게임에서는 $G_{1,2} = V > G_{2,2} = V/2$이므로 정상 상태 $x=0$은 항상 불안정하다. 비둘기 집단에 소수의 매를 도입하면 매의 점진적인 침입으로 이어질 수 있다.

마찬가지로, $0 < x^* < 1$인 경우 세 번째 정상 상태 $x^*$는 항상 안정적이라는 것을 보여줄 수 있다. 매-비둘기 게임에서 $x^* = V/C$는 매와 비둘기가 모두 있는 혼합 집단에 해당한다.

결론적으로 매-비둘기 게임에는 두 가지 경우가 있다. 만약 $V \geqslant C$, 즉 자원의 가치가 가능한 비용보다 크다면, $0 < x(0) < 1$의 초기 조건 $x(0)$에 상관없이 개체군은 매는 있지만 비둘기는 없는 안정 상태로 향한다. 따라서 매 전략은 진화적으로 안정적인 전략이다. 반대로 $V < C$인 경우, 개체군은 매의 비율 $x^*$와 비둘기의 비율 $1-x^*$가 혼합된 안정 상태로

변하는 경향이 있다. 따라서 이 모델은 공격적 행동이 덜한 개체가 $V < C$ 일 때 생존할 수 있는 이유를 설명할 수 있다. 또한 $x^* = V/C$ 공식은 패자의 비용 $C$가 높을수록 개체군에서 매의 비율 $x^*$가 작아진다는 것을 보여준다. 따라서 가장 위험한 '무기'를 가진 종은 무기를 종내 싸움에는 거의 사용하지 않으며, 경쟁 동물들이 서로에게 깊은 인상을 남기려고 노력하지만 부상을 초래할 수 있는 실제 싸움은 피하는 비공격적인 싸움의 의식을 선호한다.

1973년 메이너드 스미스와 프라이스의 논문은 진화적으로 안정적인 전략의 개념을 논의하고 주로 동물의 경쟁에 대한 컴퓨터 시뮬레이션을 사용하여 다양한 전략의 보상을 기록하였다. (23.3)와 같은 동역학 방정식을 사용하는 접근 방식은 다소 후에 테일러(Taylor)와 존커(Jonker)에 의해 개발되었다. 그 이후로 많은 저자들이 게임 이론의 아이디어를 진화 생물학의 문제에 적용하거나 반대로 게임 이론의 보다 고전적인 문제에 동적 진화 접근법을 적용하였다. 동물의 갈등과 관련된 문제 외에도 부모의 투자 문제나 성비(출생 시 수컷과 암컷의 비율) 문제를 예로 들 수 있는데, 후자는 1884년 칼 뒤싱(Düsing)이 이미 연구했고 1930년 로널드 피셔가 ≪자연 선택의 유전적 이론≫이라는 저서에서 연구하였다. 다른 모델들은 '죄수의 딜레마' 또는 '가위바위보' 게임의 역동적인 측면에 초점을 맞추고 있다. 또한 진화적으로 안정적인 전략의 개념이 게임 이론의 내쉬 균형 개념과 밀접한 관련이 있다는 것을 깨달았다.

확고한 무신론자였던 프라이스는 1970년 신비로운 경험을 한 후 기독교 신앙으로 개종하였다. 그는 1974년 "(자신이) 하고 있던 이론적 수학 유전학이 인간의 문제와 별로 관련이 없다"고 생각하여 연구를 포기했다. 그는 자신의 모든 재산을 노숙자들에게 기부하고 몇 달 후 자살하였다.

반면 메이너드 스미스는 이러한 생각을 계속 이어나가 1977년 왕립학회 회원으로 선출되었다. 그는 많은 책을 출간하였다. ≪생태학의 모델≫(1974), ≪성의 진화≫(1978), ≪진화와 게임 이론≫(1982), ≪생물학의 문제들≫(1986), ≪다윈은 옳았을까≫(1988), ≪진화 유전학≫(1989) 등이 있다. 또한 E. Szathmáry와 공동으로 ≪진화의 주요 전환≫(1995)과 ≪생명의 기원≫(1999)을 출간하였다. 1985년에 은퇴하였다. 1999년 스웨덴 왕립과학원으로부터 '진화 생물학의 개념적 발전에 대한 근

본적인 공헌'을 인정받아 생명과학 분야 크라푸어드 상을 수상하였다. 2003년에는 하퍼와 공동으로 ≪동물의 신호≫를 출판하였다. 2004년 서섹스에서 사망하였다.

## 참고문헌


1. Charlesworth, B., Harvey, P.: John Maynard Smith, 6 January 1920–19 April 2004. Biog. Mem. Fellows R. Soc. 51, 253–265 (2005)
2. Edwards, A.W.F.: Carl Düsing (1884) on the regulation of the sex-ratio. Theor. Pop. Biol. 58, 255–257 (2000)
3. Frank, S.A.: George Price's contributions to evolutionary genetics. J. Theor. Biol. 175, 373–388 (1995)
4. Harman, O.: The Price of Altruism. W. W. Norton, London (2010)
5. Maynard Smith, J., Price, G.R.: The logic of animal conflict. Nature 246, 15–18 (1973)
6. Maynard Smith, J.: Evolution and the Theory of Games. Cambridge (1982)
7. Schwartz, J.: Death of an altruist: Was the man who found the selfless gene too good for this world? Lingua Franca 10, 51–61 (2000) bio.kuleuven.be/ento/pdfs/schwartz2000.pdf
8. Sigmund, K.: John Maynard Smith and evolutionary game theory. Theor. Pop. Biol. 68, 7–10 (2005)
9. Taylor, P.D., Jonker, L.B.: Evolutionary stable strategies and game dynamics. Math. Biosci. 40, 145–156 (1978)
10. Von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press (1944) archive.org

24장

# 카오스적인 개체수 (1974)

1974년 호주의 물리학자에서 생태학자로 전향한 로버트 메이는 인구 역학 모델로서 이산 시간 로지스틱 방정식을 연구하였다. 그는 예상치 못한 분기점이 발생하고 점근적 행동이 혼란스러울 수 있다는 것을 발견하였다. 따라서 단순한 결정론적 모델로는 장기 예측이 불가능할 수 있다. 메이의 논문은 '카오스 이론'의 시발점이 된 논문 중 하나였다.

로버트 맥크레디 메이(May)는 1936년 호주에서 태어났다. 이론 물리학을 공부하고 1959년 시드니 대학교에서 박사 학위를 받은 후 하버드 대학교 응용수학과에서 2년을 보냈다. 호주로 돌아온 그는 이론 물리학 교수가 되었다. 1971년 프린스턴 고등 연구소를 방문하던 중 연구 주제를 바꿔 동물 개체군 역학에 집중하기 시작하였다. 1973년 그는 프린스턴의 동물학 교수가 되었다. 같은 해에 ≪모델 생태계의 안정성과 복잡성≫이라는 제목의 책을 출간하였다.

1974년 5월 ≪사이언스≫에 ≪세대가 겹치지 않는 생물학적 개체군: 안정점, 안정 주기, 카오스≫이라는 제목의 논문을 발표하여 인구 역학에서 매우 단순한 수학적 모델이 혼란스러운 방식으로 작동할 수 있음을 보여주었다.

이 문제의 기원을 이해하려면 약 10년 전으로 거슬러 올라가야 한다.

그림 24.1:
로버트 M. 메이 (1936–2020)

1963년 매사추세츠 공과대학(M.I.T.)에서 근무하던 미국 기상학자인 에드워드 로렌츠(Lorenz)는 컴퓨터로 수치 시뮬레이션을 하던 중 미분 방정식 세 개만 있는 단순화된 대기 모델이 매우 놀라운 방식으로 작동할 수 있다는 사실을 발견하였다. 초기 조건의 작은 변화가 시뮬레이션의 최종 결과와 기상 예보를 완전히 바꿀 수 있다는 것이다. 수학자 앙리 푸앵카레(Poincaré)는 태양계 행성의 움직임을 연구한 후 사실 컴퓨터 시대 훨씬 이전인 20세기 초에 이미 이러한 가능성을 생각하였다. 하지만 1970년대 초에는 소수의 연구자만이 이 이상한 속성을 더 자세히 살펴보기 시작하였다. 메릴랜드 대학교의 제임스 요크(Yorke)는 로렌츠의 연구에 대해 생각하던 중 이러한 맥락에서 '카오스'라는 용어를 도입하였다. 그가 제자 티엔이엔 리(Li)와 함께 쓴 ≪제3기는 카오스를 의미한다≫라는 제목의 논문은 1975년에 발표되었다. [1]

메이는 다음 모델에 집중하였다.

$$p_{n+1} = p_n + a\,p_n(1 - p_n/K). \tag{24.1}$$

$a$와 $K$는 양수 매개변수이며, $p_n$은 연도 $n$의 동물 개체군 크기를 나타낸다. 수용 능력 $K$에 비해 $p_n$이 작을 때, 동역학은 기하학적 성장 $p_{n+1} \approx (1+a)\,p_n$에 가까워진다. 전체 방정식은 베르휼스트(Verhulst)가 도입한 로지스틱 방정식의 일종의 이산 시간 유사 방정식이다 (6장 참조). 그러

[1] 놀랍게도 더 일반적인 결과는 1964년 O. M. 샤르코프스키에 의해 증명되었지만, 우크라이나 수학 저널에 발표된 그의 논문은 잘 알려지지 않았다.

나 후자와 달리 메이는 이산 시간 방정식이 훨씬 더 놀라운 동작을 할 수 있다는 것을 보여주었는데, 이는 간단한 휴대용 계산기로 덧셈과 곱셈을 하면 쉽게 관찰할 수 있다 (그림 24.2). 메이너드 스미스(Smith)는 1968년 저서인 ≪생물학의 수학적 아이디어≫에서 이미 방정식(24.1)을 고려하였다. 그러나 $a$에 대해 몇 가지 수치값을 시도해 보았음에도 불구하고 그는 특별한 무언가가 있다는 것을 깨닫지 못했다.

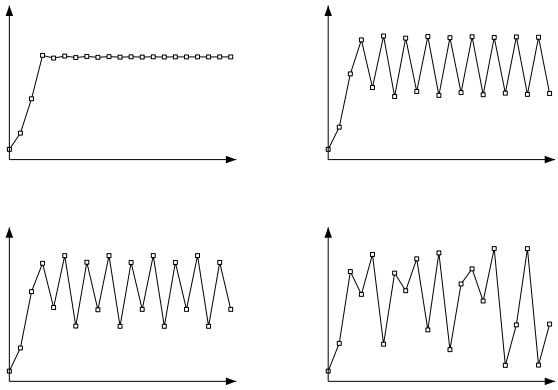

그림 24.2: 모든 그림에서 가로축에 $n$, 세로축에 $p_n$, 세로축에 $p_0 = K/10$이 나타나 있다. 선은 좌표 $(n, p_n)$로 점을 결합하여 얻었다. 왼쪽 위: $0 < a < 2$ (정상 상태). 오른쪽 상단: $2 < a \leqslant 2.449$ (주기 2의 주기). 왼쪽 하단: $2.450 \leqslant a \leqslant 2.544$ (주기 4의 주기). 오른쪽 아래: $2.570 \leqslant a \leqslant 3$ (카오스 상태일 수 있음).

1974년 로버트 메이의 논문에서와 유사한 그림 24.2는 $0 < a < 2$일 때 모집단 $p_n$이 정상 상태로 수렴하는 것을 보여준다. $2 < a \leqslant 2.449$(상한 2.449는 근사치)일 때, 모집단 $p_n$은 주기 2의 주기로 수렴하는 경향이 있다. $2.450 \leqslant a \leqslant 2.544$ 일 때, 모집단 $p_n$ 은 주기 4의 주기를 따른다. $2.545 \leqslant a \leqslant 2.564$일 때, $p_n$은 주기 8의 주기를 따르는 경향이 있다. $p_n$이 주기 $2^n$의 주기로 향하는 매개변수 $a$의 간격은 $n$이 증가함에 따라 작아지고 2.570을 초과하지 않는다. $a \geqslant 2.570$일 때, $p_n$은 '혼란스러운' 방식으로 동작할 수 있다.

1976년 5월에는 ≪매우 복잡한 역학을 가진 간단한 수학적 모델≫

이라는 제목으로 ≪네이처≫에 이 문제에 대한 리뷰를 썼다. 거기서 그는 자신의 결과뿐만 아니라 다른 연구자들의 결과도 수집하였다. 먼저 $x_n = \frac{a p_n}{K(1+a)}$ 와 $r = 1 + a (r > 1)$를 설정하면 방정식(24.1)을 보다 간단한 형태로 다시 쓸 수 있다는 것을 알 수 있다.

$$x_{n+1} = r x_n (1 - x_n) . \tag{24.2}$$

이 방정식이 인구 역학에서 의미를 가지려면 $x_n$이 모든 $n$에 대해 양수여야 한다. 따라서 초기 조건 $x_0$이 $0 \leqslant x_0 \leqslant 1$을 만족하고 $r \leqslant 4$를 만족한다고 가정한다. 후자의 조건은 (24.2)의 오른쪽이 0과 1 사이를 유지하도록 한다. 놀랍게도 1947년에 이미 스타니슬라프 울람(Ulam)과 존 폰 노이만(Neumann)이 난수 생성기로서 카오스의 경우 $r = 4$를 사용하였다. 함수

$$f(x) = r x (1 - x),$$

를 도입하면 방정식 (24.2)는 $x_{n+1} = f(x_n)$로 다시 쓸 수 있으며, 정상 상태는 $x = f(x)$의 해가 된다. 그래프로 보면 $y = f(x)$와 $y = x$ 곡선의 교집합이다 (그림 24.3). $x = 0$은 항상 정상 상태임을 알 수 있다. $r > 1$이므로 $x^* = r x^* (1 - x^*)$, 즉 $x^* = 1 - 1/r$인 또 다른 정상 상태 $x^* > 0$도 존재한다.

$r > 1$이므로 정상 상태 $x = 0$은 불안정하다. 실제로 $x_n$이 0에 가까울 때, 우리는 $x_{n+1} \approx r x_n$을 갖게 된다. 따라서 $x_n$은 0에서 멀어지는 경향이 있다. 정상 상태 $x^*$는 $1 < r < 3$일 때만 국소적으로 안정적이다.

> 실제로 $y_n = x_n - x^*$로 두면 (24.2)는 $y_{n+1} = (2 - r - r y_n) y_n$과 같다. $x_n$이 $x^*$에 가까우면 $y_n$은 0에 가까우며 $y_{n+1} \approx (2 - r) y_n$이 된다. 그러나 $y_{n+1} = k y_n$이면 $y_n = k^n y_0$이 된다. 따라서 $1 < k < 1$일 경우에만 $n \to \infty$일 때 $y_n \to 0$이 된다. 여기서 정상 상태 $x^*$는 $-1 < 2 - r < 1$, 즉 $1 < r < 3$일 경우에만 국소적으로 안정적이다.

$1 < r < 3$일 때, 모든 초기 조건 $0 < x_0 < 1$에 대해 수열 $x_n$이 실제로 $x^*$를 따르는 경향이 있음을 알 수 있다(그림 24.3a). 하지만 $3 < r \leqslant 4$가 되면 어떻게 될까? 이 질문에 답하기 위해 먼저 $x_{n+2} = f(x_{n+1}) = f(f(x_n))$

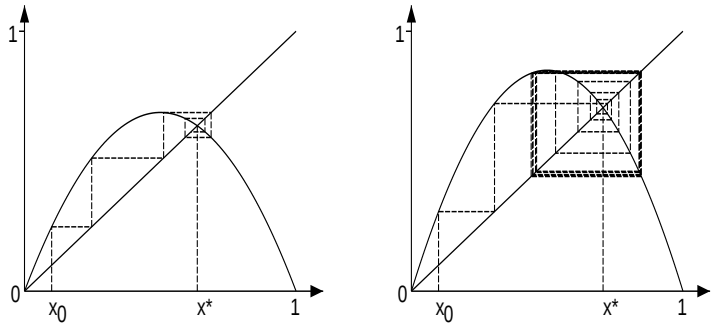


그림 24.3: 함수 $y=f(x)$, 직선 $y=x$, 정상 상태 $x^*$, 수열은 $x_{n+1}=f(x_n)$로 정의된다. 위쪽: $r=2.75$, 수열은 $x^*$를 향하는 경향을 보인다. 아래: $r=3.4$, 정상 상태 $x^*$는 불안정하고 수열은 주기 2의 주기로 향한다.

임에 주목하자. 함수

$$f_2(x)=f(f(x))=r^2x(1-x)\,(1-rx(1-x))$$

를 도입하하고 함수 $f_2(x)$의 고정점이라고 하는 방정식 $x=f_2(x)$의 해를 고려한다. 그래프상으로는 $y=f_2(x)$와 $y=x$ 곡선의 교집합이다 (그림 24.4).

$x=f(x)$이면 $x=f(f(x))=f_2(x)$이다. 따라서 $x=0$과 $x=x^*$도 함수 $f_2(x)$의 고정점이다. 그러나 $r>3$일 때, 함수 $f_2(x)$에는 $f(x_-)=x_+$와 $f(x_+)=x_-$를 만족하는 두 개의 다른 고정점 $x_-$와 $x_+$가 있다.

실제로 $f_2'(x)=f'(f(x))\,f'(x)$임을 알 수 있다. 따라서 $f_2'(x^*)=[f'(x^*)]^2$이다. 그러나 $f'(x)=r(1-2x)$이고 $x^*=1-1/r$이다. 따라서 $f'(x^*)=2-r$ 및 $f_2'(x^*)=(2-r)^2$이다. 따라서 $x=x^*$에서 함수 $f_2(x)$의 기울기는 $r>3$일 경우 $f_2'(x^*)>1$이 된다. 그러나 $f_2(0)=0$, $f_2'(0)=r^2>1$, $f_2(1)=0$이므로, 그림 24.4b에서 $0<x_-<x^*$ 및 $x^*<x_+<1$을 갖는 $x=f_2(x)$ 방정식의 다른 두 해 $x_-$와 $x_+$가 반드시 존재함을 알 수 있다. 동일한 결론에 도달하는 또 다른 방법은 $x=f_2(x)$ 방정식을 푸는

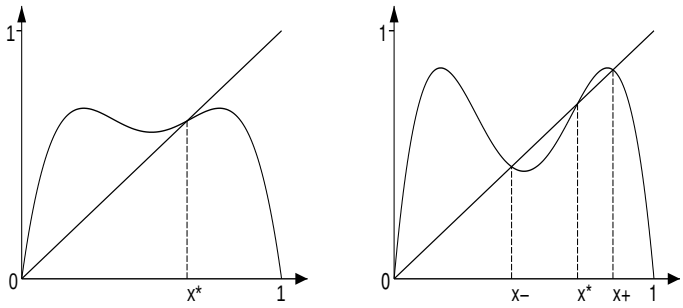


그림 24.4: 곡선 $y = f_2(x) = f(f(x))$와 $y = x$, 그리고 평형해 $x^*$이다. 상단: $r = 2.75$. 아래: $r = 3.4$ 및 방정식 $x = f_2(x)$의 다른 두 해 $x_-$ 및 $x_+$.

것인데, 이 방정식은 차수 4의 다항 방정식으로, 두 개의 자명해인 $x = 0$과 $x = x^*$가 있다. 서로 다른 두 해 $x_-$와 $x_+$는 다항식 방정식

$$x^2 - \frac{1+r}{r}x + \frac{1+r}{r^2} = 0 \tag{24.3}$$

의 근이다. 이 두 해는 판별자가 양수인 경우, 즉 $r > 3$인 경우 실수이다. $f_2(f(x_-)) = f(f(f(x_-))) = f(f_2(x_-)) = f(x_-)$이므로 $f(x_-)$점 역시 $f_2(x)$의 고정점이다. 그러나 $x_-$는 $f(x)$의 고정점이 아니기 때문에 $f(x_-) \neq x_-$이다. 그리고 $f(x_-) \neq x^*$, 그렇지 않으면 $x_- = f(f(x_-)) = f(x^*) = x^*$가 될 것이다. $f(x_-) \neq 0$이므로 $f(x_-) = x_+$라는 결론을 내릴 수 있다. 마찬가지로 $f(x_+) = x_-$이다.

따라서 $r > 3$의 경우 예를 들어 $x_0 = x_-$이면 $x_1 = x_+, x_2 = x_-, x_3 = x_+$임을 알 수 있다. 또한 거의 모든 초기 조건 $0 < x_0 < 1$에서 $x_n$ 수열은 주기 2의 주기인 $x_-, x_+, x_-, x_+$ 등으로 향하는 경향이 있음을 보여줄 수 있다 (그림 24.3b 및 24.4b 참조). 이 주기는 $r$이 임계값 $r_1 = 1 + \sqrt{6} \approx 3.449$ 이하이면 안정적으로 유지되며, 여기서 $f_2'(x_-) = -1$이다.

실제로 (24.3)를 사용하면 다음과 같은 결과를 얻을 수 있다.

$$\begin{aligned} f_2'(x_-) &= f'(f(x_-))\,f'(x_-) = f'(x_+)\,f'(x_-) \\ &= r^2(1-2x_+)(1-2x_-) = r^2(1-2(x_++x_-)+4x_+x_-) \\ &= r^2\left(1-2\frac{1+r}{r}+4\frac{1+r}{r^2}\right) = -r^2+2r+4. \end{aligned}$$

따라서 $r^2+2r+5=0$, 특히 $r=1+\sqrt{6}$이면 $f_2'(x_-)=-1$이 된다.

$r_1<r<r_2$의 경우 주기 4의 주기가 안정적이게 된다. 함수

$$f_4(x) = f_2(f_2(x)) = f(f(f(f(x))))$$

에는 4개의 새로운 고정점이 있다 (그림 24.5). $r_2<r<r_3$의 경우, 길이 8 등의 주기가 된다. 숫자 $r_n$은 $n\to+\infty$일 때 한계 $r_\infty\approx 3.570$에 도달하는 경향이 있다. $r_\infty<r\leqslant 4$일 때는 시스템이 혼란스러워질 수도 있다. 그림 24.6은 동역학의 복잡성에 대한 아이디어를 제공하는 분기 다이어그램을 보여준다. [2]

로버트 메이는 매우 단순한 동역학 시스템도 매우 복잡한 동작을 할 수 있다는 점을 강조하며 결론을 내렸다. 이는 $x_{n+1}=rx_n(1-x_n)$ 방정식에만 국한되지 않는다. 카오스로 이어지는 동일한 '주기 배가 캐스케이드'는 '범프' 모양의 함수 $f(x)$를 가진 다른 방정식에서도 나타난다. 예를 들어 인구 생물학에서 사용되는 다른 방정식인 $x_{n+1}=x_n\exp(r(1-x_n))$가 이에 해당한다.

이 연구는 인구 역학에 관한 많은 데이터 세트를 분석하기 어렵다고 해서 놀라지 말아야 한다는 것을 시사한다. 이 모델은 또한 결정론적 모델과 확률론적 모델의 구분이 이전에 생각했던 것만큼 명확하지 않다는 것을 보여준다. 단순한 결정론적 모델을 사용하더라도 매개변수가 카오스 시스템에 있으면 장기 예측이 불가능할 수 있다.

[2] 이 다이어그램은 $r$의 각 주어진 값에 대해 좌표 $(r,x_{200})$, $(r,x_{201})$, ⋯, $(r,x_{220})$를 가진 점들을 찍어서 얻었으며, 여기서 $x_{n+1}=f(x_n)$ 및 $x_0=0.1$이다. $x_n$이 정상 상태로 향하는 경우 다이어그램에는 한 점만 표시된다. $x_n$이 주기 2의 주기로 다가가면 두 점이 표시된다.

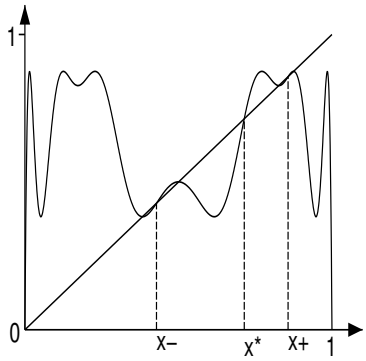


그림 24.5: $r = 3.5$ 일 때 곡선 $y = f_4(x)$ 와 선 $y = x$. $x^*$, $x_+$, $x_-$ 외에 4개의 다른 고정점이 있는데, 이 고정점은 구별하기 쉽지 않다.

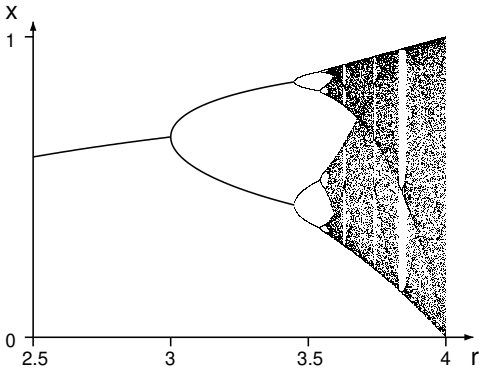


그림 24.6: 방정식의 분기 다이어그램(24.2).

1979년 메이는 왕립학회 회원으로 선출되었다. 1988년부터 1995년까지 옥스퍼드 대학교와 런던 임페리얼 칼리지에서 교수로 재직하였다. 1995년부터 2000년까지 영국 정부의 수석 과학 고문을 역임하였다. 1996년에는 '인구, 공동체 및 생태계의 역학에 대한 이론적 분석에 관한 선구적인 생태학 연구'로 크라푸어드 상을 수상하였다. 생태학에서 그는 역학 및 면역학으로 전환하여 두 권의 책을 출판하였다. ≪인간의 전염병≫(1991년, 로이 앤더슨과 공저)과 ≪바이러스 역학, 면역학 및 바이러스학의 수학적 기초≫(2000년, 마틴 노왁과 공저)가 그것이다. 후자의 책은 면역계 세포와 HIV(에이즈를 유발하는 바이러스) 사이의 상호작용을 일종의 피식자-포식자 시스템으로 분석한다 (13장 참조). 2000년부터 2005년까지 메이는 왕립학회 회장을 역임하였다. 그는 1996년에 기사 작위를 받았고 2001년에 종신 회원이되었다. 그는 2020년에 사망하였다.

# 참고문헌


1. Gleick, J.: Chaos, Making a New Science. Viking Penguin, New York (1987)
2. Levin, S.A.: Robert May receives Crafoord prize. Not. Amer. Math. Soc. 43, 977–978 (1996) ams.org
3. Li, T.Y., Yorke, J.A.: Period three implies chaos. Amer. Math. Monthly 82, 985–992 (1975)
4. Lorenz, E.N.: Deterministic nonperiodic flow. J. Atmosph. Sci. 20, 130–141 (1963) journals.ametsoc.org
5. May, R.M.: Biological populations with nonoverlapping generations: stable points, stable cycles and chaos. Science 186, 645–647 (1974)
6. May, R.M.: Simple mathematical models with very complicated dynamics. Nature 261, 459–467 (1976)
7. May, R.M., Oster, G.F.: Bifurcations and dynamic complexity in simple ecological models. Amer. Natur. 110, 573–599 (1976)
8. Maynard Smith, J.: Mathematical Ideas in Biology. Cambridge (1968)
9. Poincaré, H.: Science et Méthode. Flammarion, Paris (1908) gal-

lica.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

25장

# 중국의 한 자녀 정책 (1980)

1980년 공역 공학을 응용한 제어 이론 전문가였던 송지안과 그의 동료들은 중국의 출산율이 현재 수준을 유지한다면 21세기에 인구가 20억 명을 넘을 것이라는 계산을 내놓았다. 연령 구조화된 수학적 모델을 기반으로 한 이들의 결과는 중국 정부가 한 자녀 정책으로 전환하는 데 기여하였다.

송지안(宋健)은 1931년 중국 산둥성 룽청에서 태어났다.[1] 1950년대 소련에서 바우만 모스크바 공과대학교와 모스크바 국립대학교 수학 및 역학과에서 공부하였다. 그 후 중국으로 돌아와 중국과학원 수학연구소의 사이버네틱 연구실 책임자가 되었다. 그는 미사일 유도에 제어 이론을 적용하는 전문가였다. 그는 또한 나중에 항공 우주부로 이름이 변경된 제7기계 건설부에서 일하였다. 1978년 그는 제어 이론과 인구학 사이의 연관성에 집중하기 시작하였다.

인구 역학에 대한 송지안의 연구 맥락을 이해하려면 먼저 '통제 이론'이 무엇인지 알아야 한다. 제어 이론은 주어진 기준을 최적화하기 위해 시간이 지남에 따라 수정할 수 있는 일부 매개변수에 따라 행동이 달라지는 동적 시스템에 대한 연구이다. 이 이론은 특히 미국과 소련의 우주 프로그램과 관련하여 개발되었다. 실제로 엔지니어들은 인공위성을 지구

[1] 송은 그의 성씨이다.

그림 25.1: 송지안

궤도에 진입시키기 위해 우주왕복선의 궤도를 '제어'해야 했다. 그러나 응용 범위는 물리적 또는 공학적 문제에만 국한되지 않았다. 산아제한 정책도 수학적 의미에서 일종의 최적 제어 문제로 간주할 수 있다.

1972년 매사추세츠 공과대학(M.I.T.) 연구팀이 작성한 ≪성장의 한계 - 로마 클럽의 인류 곤경 프로젝트 보고서≫라는 제목의 에세이도 언급할 필요가 있다. 이 연구는 천연자원, 인구 규모, 오염을 고려한 세계 경제 성장에 대한 수학적 모델을 기반으로 하였다. 이 보고서는 재생 불가능한 자원의 고갈, 인구에 대한 식량 부족 또는 과도한 오염으로 인해 세계 경제가 재앙을 향해 가고 있다고 제안하였다. 자발적인 출산 제한은 제안된 해결책 중 하나였다. 요약하자면 맬서스의 논문을 현대적으로 재해석한 일종의 보고서였다. 이 보고서는 1970년대 서구에서 큰 반향을 불러일으켰다.

1949년 중화인민공화국이 건국된 이래 중국의 출산율은 재앙적인 '대약진 운동' 시기를 제외하고는 매우 높았다. 1970년대 중반 중국은 문화대혁명에서 서서히 회복하고 있었다. 가족 계획은 여성들에게 출산을 늦추고, 두 번의 연속 출산 간격을 늘리고, 자녀를 적게 낳도록 촉구하였다. 1976년 마오쩌둥(모택동) 사망 후 새로운 지도자로 등장한 덩샤오핑(등쇼평)은 1978년 농업, 공업, 과학 기술, 국방 등 '4대 현대화' 정책을 시작하였다. 당시 중국 인구의 규모와 증가는 이러한 현대화의 중요한 장애물로 인식되었다. 그때까지 군사용 애플리케이션을 연구하던 과학자들은 이 어려운 문제에 대한 해결책을 찾도록 독려받았다.

이러한 배경을 염두에 두고 송지안은 1978년 국제자동제어연맹 회의

에 참석하기 위해 헬싱키에 갔다. 그곳에서 그는 유럽의 일부 연구자들이 엄격한 산아제한이 결국 '성장의 한계' 보고서에서 발표한 재앙을 막을 수 있다는 생각으로 인구 문제에 통제 이론을 적용하려는 시도를 하고 있다는 사실을 알게 되었다. 중국으로 돌아온 그는 동료인 위징위안과 컴퓨터 전문가 리광위안을 포함한 소규모 팀을 꾸려 중국 인구 관련 데이터에 이러한 수학적 모델링을 적용했다. 당시에는 중국과 다른 국가 간의 과학적 소통이 거의 이루어지지 않았다. 연구팀은 로트카(Lotka)와 맥켄드릭(McKendrick)이 했던 것과 같은 방식으로 인구 연령 구조의 변화를 설명하는 방정식을 재개발하였다 (10 및 16 참조). 연속 시간 모델을 사용하여 다음을 정의한다.

- $P(x,t)$, 시간 $t$에 $x$세인 인구,
- $m(x)$, 연령 $x$에서의 사망률,
- $P_0(x), t=0$ 시점의 인구 연령 구조,
- $b(t), t$ 시점에 여성의 총 출산율, 즉 연령별 출산율이 $t$ 시점의 출산율과 동일하게 유지될 경우 여성이 평생 동안 낳을 평균 자녀 수,
- $f$, 여성 출생 비율,
- $h(x)$, 아이가 태어날 때 어머니의 나이에 대한 확률 분포 ($\int_0^{+\infty} h(x)\,dx = 1$).

이러한 표기법과 가설을 통해 연령 구조의 진화를 다음과 같은 편미분 방정식으로 모델링할 수 있다

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

초기 조건은 $P(x,0) = P_0(x)$이고 경계 조건은 다음과 같다

$$P(0,t) = b(t)\,f \int_0^{+\infty} h(x)\,P(x,t)\,dx\,.$$

$b(t)$는 제어할 매개변수다. 여성의 총 출산율이 일정하고 임계값

$$b^* = 1/\left[f\int_0^{+\infty} h(x)\,e^{-\int_0^x m(y)\,dy}\,dx\right]$$

를 초과하면 인구는 기하급수적으로 증가한다. 이 기준은 공식 (10.2)을 사용하여 로트카가 얻은 기준과 유사하다. 송지안 팀은 레슬리의 모델과 유사한 시간 이산형 모델도 고려하였다(21장 참조). $P_{k,n}$를 연도 $n$에 $k$세인 인구다. 이와 유사하게 $m_k$, $b_n$, $h_k$를 도입한다. 그럼 다음과 같은 식을 얻을 수 있다

$$P_{k+1,n+1} = (1-m_k)P_{k,n}\,, \quad P_{0,n+1} = b_n f\sum_{k\geqslant 0} h_k P_{k,n}\,.$$

표본 조사에서 사망률 $m_k$(그림 25.2)를 알고, 여성 출생 비율 $f \approx 0.487$, 산모의 연령 분포 $h_k$(그림 25.3), 1978년 인구의 연령 구조인 초기 조건 $P_{k,0}$(그림 25.4), 총 출산율 $b$(각 시뮬레이션 내내 일정하다고 가정)를 변경하여 송지안 팀은 1980년부터 2080년까지 100년의 시간 범위로 자국에 대한 인구 예측을 만들 수 있었다(그림 25.5). 수천 번의 덧셈과 곱셈이 필요했기 때문에(연도 $n$은 0~100년 사이, 연령 $k$는 0~90년 사이) 컴퓨터가 필요했다. 당시 중국에서는 군대에서 일하는 사람을 제외하고는 이러한 장비를 사용할 수 있는 사람이 거의 없었다. 미사일 유도 분야의 최고 전문가인 송지안도 그 중 한 명이었다.

이 예측에 따르면 중국이 1978년 출산율인 여성 1인당 2.3명($b^* = 2.19$로 추정되는 임계치 바로 위)을 유지하더라도 인구는 1980년 9억 8천만 명에서 2080년 21억 2천만 명으로 늘어날 것으로 예상하였다. 그러나 중국은 이미 농업에 사용할 수 있는 거의 모든 토지를 사용하고 있었다. 심지어 사막화와 도시화로 인해 이 땅의 일부를 잃는 경향도 있었다. 농업 생산량이 충분하지 않다면 어떻게 그런 인구를 먹여 살릴 수 있을까? 맬서스가 2세기 전에 고민했던 것과 같은 질문이다. 1975년 출산율이 $b = 3.0$일 때 2080년에는 인구가 42억 6천만 명에 달할 수도 있다. $b = 2.0$일 경우 인구는 2050년경에 최대 15억 3천만 명에 도달한 후 약간 감소하기 시작할 것이다. $b = 1.5$의 경우 2030년경에 최대 11억 7천만

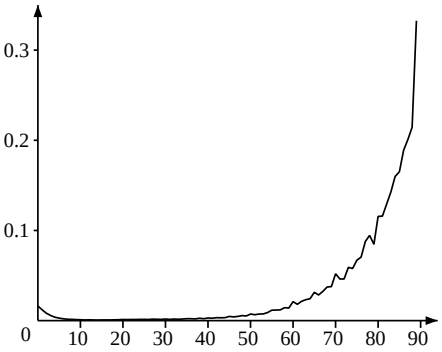


그림 25.2: 1978 년 연령에 따른 사망률 (연간).

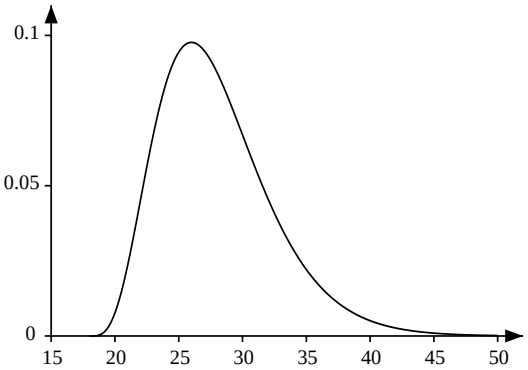


그림 25.3: 1978 년 연령에 따른 출산율 (연간).

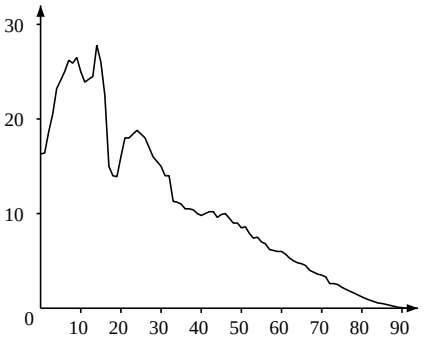


그림 25.4: 1978 년 연령 피라미드. 가로축: 연령. 세로축: 인구 (백만 명 단위).

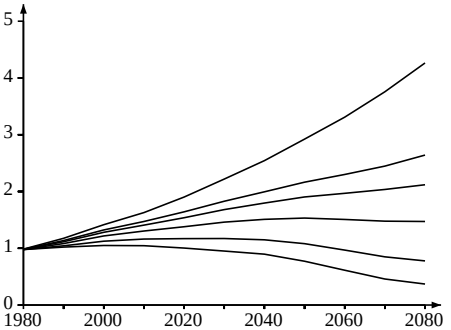


그림 25.5: 여성 1인당 평균 자녀 수에 대한 다양한 가설에 따른 인구 통계학적 예측 (십억 단위). 아래에서부터 $b = 1.0,\ 1.5,\ 2.0,\ 2.3,\ 2.5,\ 3.0$.

명에 도달할 것이다. $b = 1.0$일 경우, 최대치는 10억 5천만 명에 불과하며 2000년경에 도달할 것이다. 이 가정에 따르면 인구는 2025년에야 1978년 수준으로 돌아갈 것이다.

이 연구에서 가장 놀라운 부분은 수학적 인구 역학 역사상 비교할 수 없을 정도로 중요한 실제적인 결과였다. 실제로 리광위안은 1979년 12월 쓰촨성 청두에서 열린 인구 관련 심포지엄에서 연구팀의 시뮬레이션 결과를 발표했다.[2] 1980년 1월, 송지안, 위징위안, 리광위안은 이 연구 결과를 경제학에 초점을 맞춘 중국 학술지에 발표하여 한 자녀 정책을 옹호하였다. 또한 이들은 ≪중국 인구 개발 문제에 대한 정량적 연구에 관한 보고서≫라는 논문을 중국 최고 과학자 첸쉐센에게 보냈고, 첸쉐센은 이 논문을 출산 계획 관리국 책임자에게 추천서와 함께 전달하였다. 송지안 연구팀의 연구 결과는 대부분의 정치 지도자들에게 깊은 인상을 남겼다. 이들은 이미 마르크스의 저서(5장 참조) 에도 불구하고 산아제한의 필요성을 확신하고 있었지만, 그 통제 수준에 대해서는 여전히 망설이고 있었다. 1980년 2월, 국무원과 당 중앙위원회는 2000년까지 중국 인구 목표를 12억 명으로 정했다. 1980년 3월, 송지안 연구팀의 연구 결과가 ≪인민일보≫에 발표되었다. 4월에는 정치 지도자와 인구 전문가로 구성된 위원회가 인구 증가의 환경적, 경제적 결과를 조사한 결과 덩샤오핑이 제시한 2000년의 1인당 소득 목표를 달성하기 위해서는 한 자녀 정책이 필요하다는 결론을 내렸다. 이 정책은 같은 해 9월에 공식화되었고, 이를 설명하는 공개 서한이 ≪인민일보≫ 1면에 게재되었다.

1983년까지도 여전히 많은 무단 출산이 이루어졌다. 정부는 이미 두 자녀를 둔 부부의 경우 한 명은 불임 수술을 받고 금지된 임신은 중단하기로 결정하였다. 그러나 1984년부터 딸이 한 명뿐인 시골 부부는 두 번째 아이를 가질 수 있게 되었다. 한 자녀 정책은 2015년까지 계속되었다. 지난 몇 년 동안 몇 가지 개선안이 도입되었다. 부부의 남성과 여성이 모두 한 자녀 가정 출신 인 경우 두 자녀를 가질 수 있다. 한 명 이상의 자녀를 둔 부부에 대한 억압 조치는 가혹하였다. 공무원은 직장을 잃을 수 있고, 두 번째 자녀의 학교 교육을 위해 행정 서류를 받기 위해 값 비싼 벌금을 지불해야하는 등 가혹했다. 요약하자면, 수학적 모델링 역사에서

[2] 여기와 아래에는 Susan Greenhalgh의 자세한 설명이 요약되어 있다[1,2].

이처럼 사회적으로 큰 영향을 미친 사례는 찾기 어렵다. 물론 송지안과 그의 협력자들의 연구는 한 자녀 정책의 선택으로 이어진 일부 요소에 불과하였지만 중요한 역할을 하였다.

이전 장에서와 마찬가지로 수학적 모델링의 역할은 고민의 대상이 될 수 있다. 실제 상황에서 시작하여 모델을 구축한다. 이 모델은 수학적으로 분석하거나 컴퓨터로 시뮬레이션할 수 있다. 그러면 일부 매개변수가 변할 때 모델이 어떻게 작동하는지 이해할 수 있다. 그러나 수학은 모델이 실제 상황을 충실하게 묘사하는지 여부를 말해주지 않는다. 매우 중요한 몇 가지 측면이 무시되었을 수 있다. 예를 들어 일부 모델에는 중국 인구를 12억 명 미만으로 유지하는 것과 같은 목표가 포함되어 있다. 수학은 이 목표가 적절한지 말해주지 않는다.[3]

1980년 송지안은 중국 우주 프로그램의 '아버지'로 불리는 첸쉐센이 쓴 ≪엔지니어링 사이버네틱스≫라는 책의 새 판에 공동 저자로 참여하기도 했다. 그 후 항공우주부 차관 겸 수석 과학자-엔지니어(1981~1984), 중국 공산당 중앙위원회 위원(1982~2002), 국가과학기술위원회 위원장(1985~1998), 국무위원(1986~1998) 등 다양한 고위 정치 직책을 역임하였다. 그는 영어로 번역된 다른 두 권의 저서인 ≪중국의 인구 통제≫(1985년, 투안 치시엔, 유징위안 공저)와 ≪인구 시스템 통제≫(1988년, 유징위안 공저)도 출간하였다. 이 책들은 인구 역학에 적용되는 최적 제어 이론을 발전시켰다. 송지안은 1991년 중국 과학원, 1994년 중국 공학원에 선출되었으며, 1998년부터 2002년까지 중국 공학원 회장을 역임하였다.

## 참고문헌

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.

[3]2000년의 인구는 12억 6,400만 명으로 추산된다. 1인당 소득은 1980년과 2000년 사이에 약 $200에서 $1000로 증가했다. 동시에 성비는 주로 성 선택적 낙태로 인해 남아에게 극도로 편향되었다.

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)

26장

# 현대의 몇 가지 문제

이 장에서는 인구학의 인구 고령화, 역학의 신종 질병(에이즈, 사스, 매개체 매개 질병 등) 및 예방 접종 정책, 생태학의 어업 정책, 인구유전학의 유전자 변형 생물체의 확산 등 수학적 인구 역학에 관한 현대적 문제에 대한 간략한 개요를 제공한다. 이러한 문제를 모델링하기 위해 프랑스의 전문 기관이 언급되고 연구의 다양한 측면도 강조된다.

인구 통계학에서는 지난 수십 년 동안 인구 고령화라는 비교적 새로운 문제가 나타났다. 이 문제는 프랑스(그림 참조)뿐만 아니라 다른 많은 유럽 국가와 한국, 일본에서도 우려의 대상이 되고 있다. 연금 제도, 이민 정책 등 경제 및 사회적으로 중요한 영향을 미친다. 프랑스에서는 국립인구학연구소(INED)와 국립통계경제연구소(INSEE)에서 고령화 현상을 분석하기 위한 수학적 모델을 개발하고 있다. 인구 예측의 어려움 중 하나는 출산율이 10년 후도 예측할 수 없을 정도로 시간이 지남에 따라 크게 달라질 수 있다는 사실에 있다. 1968년 프랑스 인구에 대한 1985년 예측을 되돌아보면 이 점이 특히 두드러진다. 이 예측은 1970년대에 발생한 출산율 감소를 예상하지 못했다. 잘못된 것으로 판명된 수학적 모델에 기반한 모든 예측, 특히 언론에서 반향을 일으킨 예측을 검토하는 것은 흥미로울 것이다. 이것은 이 책의 한자녀 정책에 대한 장을 읽은 후

독자가 이미 의심스러워했을지도 모를 '진보'이라는 인상을 조정해줄 것이다. 후자의 주제와 관련하여 현재 새로운 문제는 향후 수십 년 동안 예상되는 급속한 고령화 현상을 피하기 위해 정책을 완화하는 방법이다. 이 역시 수학적 모델이 논쟁에 기여한다.

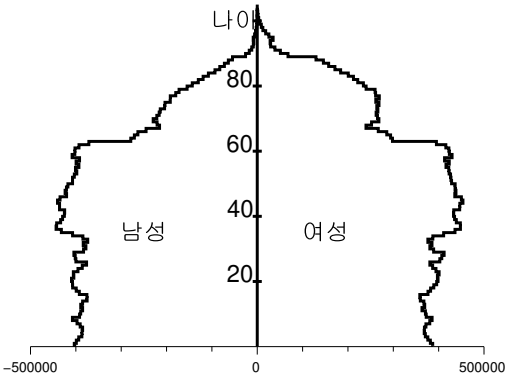


그림 26.1: 2010년 1월 1일 기준 프랑스 인구의 연령 피라미드. 출처: www.insee.fr.

역학 분야에서 지난 30년 동안 전 세계적으로 등장한 새로운 문제 중 에이즈 전염병의 발전이 특히 두드러진다. 일부 모델은 러시아, 인도 또는 중국과 같이 최근에 감염된 국가에서 전염병의 미래를 추측하려고 한다. 서유럽과 북미에서처럼 전염병이 둔화될지, 아니면 일부 사하라 사막 이남 국가에서처럼 인구의 상당한 비율에 도달할지 예측하기는 어렵다. 아프리카의 에볼라, 북미의 웨스트나일열, 사스(중증 급성 호흡기 증후군), 조류 독감, 치쿤구니야 또는 H1N1 인플루엔자와 같은 다른 신종 질병도 모두 수학적 모델을 통해 면밀히 조사되었지만 성공적이라고 보기는 어렵다.

사스의 경우 모델링의 어려움 중 하나는 전염병이 각 국가 내에서는 비교적 제한적으로 유지되지만 국가 간(홍콩과 중국, 싱가포르, 캐나다 등)으로 매우 빠르게 확산될 수 있다는 점이다. 새로운 유마다 나타나는 전염병 곡선의 무작위적인 특성도 무시할 수 없다. 16장과 22장에서 살펴본 것처럼 확률모델은 일반적으로 다루기가 더 어렵다.

2005~2006년 레위니옹 섬(인도양의 프랑스 해외 영토)에서 발생한

치쿤구니야열 유행에 대한 일부 수학적 모델은 말라리아에 대한 로스(Ross)의 모델에서 영감을 얻었다(12장 참조). 이 두 질병은 모기에 의해 전파된다. 고려해야 할 중요한 측면은 계절의 영향이다. 실제로 남쪽 겨울에는 모기 개체수가 감소하여 질병의 전파가 감소한다. 섬 인구의 극히 일부만을 담당하는 약 30명의 일반의들로 구성된 소규모 네트워크가 매주 보고하는 새로운 감염자 수를 보여주는 그림26.2에서 이를 확인할 수 있다. 네트워크는 2005년 9월과 10월 몇 주 동안 새로운 감염 사례를 발견하지 못했지만 질병의 전파는 여전히 계속되고 있었다. 전염병에 대한 수학적 모델은 국립보건의료연구소(INSERM)와 열대연구소(IRD)에서 개발되었다. 이러한 모델에도 불구하고 2005년 남부의 겨울이 끝나기 전, 불과 수천 명만 감염된 이 전염병이 사라지지 않을 것이라고는 아무도 예측할 수 없었다. 결국, 섬 인구의 거의 3분의 1에 해당하는 약 26만 6,000명이 감염되었다. 이는 전염병의 미래를 예측하는 것이 필요한 경우 매우 어려울 수 있으며 전염병 초기에 전염병이 경미한 전염병인지 주요 전염병인지 구별하기가 쉽지 않다는 것을 보여준다. 이는 일기 예보와 비슷한 점이 있다. 오늘날 일기예보는 해양과 대기의 복잡한 수학적 모델에 대한 거대 컴퓨터 시뮬레이션에 의존한다. 그럼에도 불구하고 며칠 이후의 예측은 신뢰할 수 없다.

좀 더 이론적인 관점에서 치쿤구니야 전염병은 환경에 계절적(예: 주기적) 변동이 있다고 가정하는 모델에서 기본 재생산 수 $\mathscr{R}_0$의 개념을 어떻게 적용해야 하는지에 대한 의문을 제기하였다. 이러한 적용은 그리 간단하지 않으며, 2009년 H1N1 인플루엔자 대유행과 같이 계절성의 영향을 받는 다른 전염병에 대해 매개변수 $\mathscr{R}_0$이 어떻게 사용되었는지에 대한 우려를 불러일으킨다.

모델러들이 분석하려고 시도한 또 다른 문제는 약물 내성(항생제, 항말라리아제 등)에 관한 문제이다. 백신 주사가 잠재적 위험을 수반할 때 비용과 편익의 균형을 맞추는 방법에 대한 다니엘 베르누이(Bernoulli)와 달랑베르(d'Alembert) 시대 이래로 역학에서 반복되는 질문은 여전히 논란의 대상이 되고 있으며, 위험에 대한 민감도가 변화함에 따라 앞으로도 계속될 수 있다. 따라서 1998년 프랑스 보건부는 B형 간염 백신이 극소수의 사례에서 심각한 합병증을 유발할 수 있다는 일부 제언에 따라

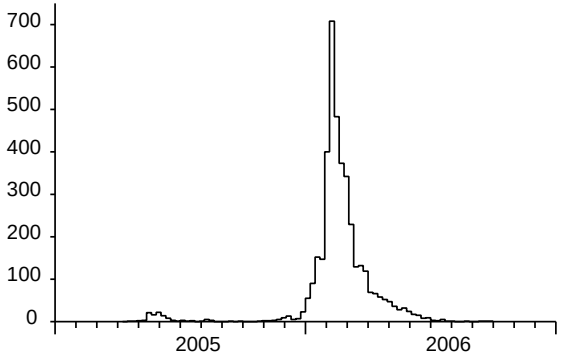


그림 26.2: 2005~2006년 레위니옹 섬에서 유행한 치쿤구니야열. 소규모 의사 네트워크에 의해 보고된 주당 신규 확진자 수를 시간의 함수로 나타낸 수치이다. 첫 번째 작은 정점은 2005년 5월에, 두 번째 큰 정점은 2006년 2월에 도달하였다. 이 수치의 숫자에 약 67을 곱해야 실제 전염병의 규모를 알 수 있다. 출처: www.invs.sante.fr.

B형 간염 바이러스 감염 후 사망 위험에 비해 그 위험성이 무시할 수 있는 수준이라고 하더라도 학교에서의 백신 접종 캠페인을 중단하였다.

생태학에서 어류 개체군 동역학에 대한 연구는 여전히 많은 문제를 제기하고 있다. 그럼에도 불구하고 어획 할당량 및 기타 제한에 대한 선택 위한 과학적 근거로 작용해야 한다. 비스카이 만의 멸치 남획과 지중해의 참다랑어 남획은 최근의 두 가지 예에 불과하다. 어족 자원의 추정치는 종종 신뢰할 수 없기 때문에 이러한 데이터를 사용하는 모델은 신중하게 고려해야 한다. 프랑스에서는 이러한 유형의 연구가 주로 해양개발연구소(IFREMER)에서 수행된다. 일부 수학적 모델은 국제포경위원회의 과거 결정에도 영향을 미쳤다.

집단 유전학에서 유전자 변형 생물체의 확산은 일부 연구자들이 피셔(Fisher)의 모델에서 영감을 얻은 '반응-확산' 모델을 사용하여 연구하려고 시도한 논란의 대상이기도 하다 ( 20장 참조). 이 분야는 국립농업과학원(INRAE)의 연구 분야이다.

좀 더 이론적인 측면에서 언급할 수 있는 것은 다음과 같다.

- 확산 방정식(20장 참조) 또는 연령 구조 방정식(16장 참조)과 같은 편미분 방정식에 대한 과학적 연구
- 전염병 확산을 모델링하는 무작위 네트워크에 대한 연구와 결정론적 근사에 대한 연구를 포함하는 공간적, 혹은 비공간적 확률론적 모델에 대한 과학적 연구 (16장 및 22장 참조)

이러한 유형의 연구는 주로 응용 수학자가 수행한다. 최근 몇 년 동안 프랑스 대학 및 기타 고등 교육 기관에 수리생물학 석사 과정이 여러 개 개설되었다.

다른 과학 분야와 마찬가지로 인구 역학에 대한 수학적 연구는 주로 다음을 통해 이루어졌다.

- 관련학회: 네덜란드 이론생물학회(1970년 이후), 수리생물학회(1973년), 프랑스 이론생물학회(1985년), 중국 수리생물학회(1985년), 일본 수리생물학회(1989년), 유럽 수리 및 이론생물학회(1991년), 라틴아메리카 수리생물학회(2002년), 한국 수리생물학회(2005년) 등
- 전문 저널: Bulletin of Mathematical Biology (1939년 이후), Mathematical Biosciences (1967), Journal of Mathematical Biology (1974), Mathematical Medicine and Biology (1984), Mathematical Population Studies (1988), Mathematical Biosciences and Engineering (2004), International Journal of Biomathematics (2008), Biomath (2012) 등
- 컨퍼런스: Annual Meeting of the Society for Mathematical Biology, Mathematical and Computational Population Dynamics, European Conference on Mathematical and Theoretical Biology 등

여기서는 수학과 인구 역학에 대한 수학의 응용 사이의 경계에 있다고 명시적으로 표현한 것들만 언급하였다.

그러나 각 특정 분야(인구학, 생태학, 인구 유전학, 역학 등)에 대해 어느 정도의 수학적 모델을 사용하는 과학 저널과 컨퍼런스가 있다.

끝으로, 관심 있는 독자는 월드 와이드 웹에서 제공되는 원문을 살펴보시기 바란다. 주소는 각 장의 마지막에 있는 참고 문헌에 나와 있다. 로널드 피셔가 멘델(Mendel)에 대해 쓴 구절은 다음과 같다.

> "과학의 역사는 교사들이 간접 자료를 사용하는 것과 이로 인해 과거의 위대한 발견들이 이루어진 환경과 그 당시의 지적 분위기가 잊혀짐으로써 큰 타격을 입었다. 연구의 초본에서는 항상 배울 점이 많고 때로는 예상치 못한 발견을 할 수도 있다."

# 참고문헌

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. Bennett, J.H.: Experiments in Plant Hybridisation. Oliver & Boyd, Edinburgh (1965)
3. Levin, S.A.: Mathematics and biology, the interface. www.bio.vu.nl/nvtb/

# 그림

- 6쪽. 런던 왕립학회가 소장하고 있는 토마스 머레이(Thomas Murray, 1687년경)가 그린 초상화. Chapman, S.: Edmond Halley, F.R.S. 1656–1742. Notes Rec. R. Soc. Lond. 12, 168–174 (1957) © The Royal Society.
- 13쪽. 바젤 쿤스트뮤지엄이 소장한 에마누엘 한트만(1753년)이 그린 초상화. Leonhard Euler 1707–1783, Beiträge zu Leben und Werk. Birkhäuser, Basel (1983)
- 20쪽. 베를린의 성 베드로 교회가 소장하고 있던 초상화. 1945년에 파괴된 것으로 추정. Reimer, K.F.: Johann Peter Süssmilch, seine Abstammung und Biographie. Arch. soz. Hyg. Demogr. 7, 20–28 (1932)
- 27쪽. 바젤 자연사 박물관이 소장하고 있는 요한 니클라우스 그로스(Johann Niclaus Grooth, 1750 1755년경)가 그린 초상화. Speiser, D.: Die Werke von Daniel Bernoulli, Band 2. Birkhäuser, Basel (1982)
- 35쪽. 모리스 쿠엔틴 드 라 투르가 그린 초상화(1753년), 파리 루브르 박물관 소장.
- 40쪽. 존 리넬이 그린 초상화(1833년), 영국 헤일리버리 대학 소장. Habakkuk, H.J.: Robert Malthus, F.R.S. (1766-1834). Notes Rec. R. Soc. Lond. 14, 99–108 (1959)
- 45쪽. 플라멩의 판화(1850년). Quetelet, A.: Pierre-François Verhulst. Annu. Acad. R. Sci. Lett. B.-Arts Belg. 16, 97–124 (1850)

- 52쪽. Heyde, C.C., Seneta, E.: I. J. Bienaymé, Statistical Theory Anticipated. Springer-Verlag, New York (1977) © Académie des sciences, Institut de France.
- 57쪽. Bateson, W.: Mendel's Principles of Heredity. Cambridge University Press (1913)
- 62쪽. Pearson, K.: The Life, Letters, and Labors of Francis Galton, vol. 1. Cambridge University Press (1914)
- 62쪽. 케임브리지 대학교 트리니티 칼리지 도서관에 있는 왓슨의 초상화. Kendall, D.G.: Branching processes since 1873. J. Lond. Math. Soc. 41, 385–406 (1966)
- 70쪽. 알프레드 J. 로트카 논문. 공공 정책 논문. 희귀 도서 및 특별 컬렉션 부서. © Princeton University Library.
- 76쪽. Titchmarsh, E. C.: Godfrey Harold Hardy 1877–1947. Obit. Not. Fellows R. Soc. 6, 446–461 (1949)
- 79쪽. Stern, C.: Wilhelm Weinberg. Genetics 47, 1–5 (1962)
- 82쪽. G.H.F.N.: Sir Ronald Ross 1857–1932. Obit. Not. Fellows R. Soc. 1, 108–115 (1933) © The Royal Society.
- 93쪽. Whittaker, E.T.: Vito Volterra 1860–1940. Obit. Not. Fellows R. Soc. 3, 690–729 (1941)
- 98쪽. Yates, F., Mather, K.: Ronald Aylmer Fisher, 1890–1962. Biog. Mem. Fellows R. Soc. 9, 91–120 (1963) © The Royal Society/Godfrey Argent Studio.
- 103쪽. Yates, F.: George Udny Yule. Obit. Not. Fellows R. Soc. 8, 308–323 (1952)
- 113쪽. Heyde, C.C., Seneta, E. (eds.): Statisticians of the Centuries. Springer, New York (2001)
- 124쪽. britannica.com/EBchecked/topic/252257/J-B-S-Haldane © Bassano and Vandyk Studios.
- 135쪽. Hill, W.G.: Sewall Wright, 21 December 1889-3 March 1988. Biog. Mem. Fellows R. Soc. 36, 568–579 (1990) © Llewellyn Studios, Chicago.

- 129쪽. Nybølle, H.C.: Agner Krarup Erlang f. 1. Januar 1878 - d. 3. Februar 1929. Mat. Tidsskr. B, 32–36 (1929)
- 145쪽. 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)
- 145쪽. I. G. Petrowsky Selected Works Part I. Gordon and Breach, Amsterdam (1996) © Taylor and Francis Books UK.
- 151쪽. 사진: 데니스 켐슨. Crowcroft, P.: Elton's Ecologists, a History of the Bureau of Animal Population. University of Chicago Press (1991)
- 156쪽. © Geoffrey Grimmett.
- 163쪽. 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.
- 170쪽. © Samuel Schlaefli / ETH Zürich.
- 180쪽. Selected works of J. Song. Science Press, Beijing (1999) © Song Jian.

이 책은 수학이 중요한 통찰을 준 유전학, 생태학, 역학 및 인구학과 밀접한 관련이 있는 이론 분야인 인구 역학의 역사를 추적합니다. 오일러와 맬서스에서 중국의 한 자녀 정책까지 인구의 기하급수적 성장, 멘델의 법칙과 성씨 소멸 문제에서 전염병 확산에 대한 퍼콜레이션 이론까지 확률적 모델의 발전, 결정론과 우연론이 얽혀 있는 혼돈계의 인구 등 중요한 주제들의 출현을 조사합니다.

최근 기계 번역의 발전으로 과학 문헌에서 단일 언어의 사실상 독점은 더 이상 정당화되지 않습니다. 대학 내 언어적 소외가 심화되는 상황을 역전시킬 수 있습니다. 세심하게 수정된 이 한국어 번역을 통해 우리는 이 새로운 길을 홍보하고 있습니다.

ISBN: 979-10-426-1860-5

15€

Nicolas Bacaër / 나경아 김광수 수학적 인구 역학의 짧은 역사