Chapter 2

Historical Overview, Recent Developments, Challenges and Course Perspective

Historical review of the development of stochastic models in population biology: deterministic models, branching processes, stochastic models in population genetics, molecular biology, epidemiology and ecology. Diffusion approximations, genealogy, particle systems and stepping stone models. Scientific challenges and the role of stochastic models.

2.1 Classical deterministic population dynamics

We begin our historical review with some basic models from demography, ecology and epidemiology.

The mathematical formulation of the growth of an age-structured population was developed by Euler (1760) ([231]). The (female) birth rate at time t $\{B(t)\}_{t\geq 0}$ satisfies the renewal equation

$$B(t) = \int_0^t B(t-s)(1-L(s))m(s)ds. \tag{2.1}$$

This leads to exponential growth $B(t) \sim e^{\alpha t}$ where the *Malthusian parameter* α is given by the *characteristic equation of demography* (Euler-Lotka equation)

$$1 = \int_0^\infty e^{-\alpha s}(1-L(s))m(s)ds, \tag{2.2}$$

where $m(s)$ is the average birth rate for an individual of age s and $L(s)$ is the cumulative distribution function of the lifetime of an individual.

The implications of exponential growth of the human population was the subject of the famous writings of Thomas Malthus - *Essay on the Principle of Population* - (1798) which had a major impact and which was one of the influences on Darwin.

2.1.1 The Logistic Equation

Verhulst (1838) ([586]) introduced the *logistic equation* which describes the more realistic situation in which resources are limited and the death rate increases as the resources are exhausted.

$$\frac{dx}{dt} = \alpha x(1 - \frac{x}{N}), \quad x(0) \geq 0, \quad \alpha \geq 0 \tag{2.3}$$

$$x(t) = \frac{Nx(0)e^{\alpha t}}{N + x(0)(e^{\alpha t} - 1)} \to N \text{ as } t \to \infty \tag{2.4}$$

Here $N < \infty$ is interpreted as the *carrying capacity* of the environment in which the population lived.

2.1.2 The Lotka-Volterra Equations for Competing Species

Equations to model the competition between species in ecology were proposed by (Lotka (1925) [415] and Volterra (1926) [585]):

$$\frac{dx_1}{dt} = r_1 x_1 \left(1 - \frac{x_1}{K_1} - a_{12}\frac{x_2}{K_1}\right) \tag{2.5}$$

$$\frac{dx_2}{dt} = r_2 x_2 \left(1 - \frac{x_2}{K_2} - a_{21}\frac{x_1}{K_2}\right) \tag{2.6}$$

Coexistence, that is, a *stable equilibrium* with both species present occurs if $\frac{1}{a_{21}} > \frac{K_1}{K_2} > a_{12}$.

Remark 2.1 *Gause (1934) proposed the* competitive-exclusion principle *that states two species cannot stably coexist if they occupy the same niche, for example, if* $a_{12} = a_{21}$.

2.1.3 The SIR Epidemic Model

A classical model for the progress of an epidemic due to Kermack and McKendrick (1927) [372] is given by the system of ode:

$$\frac{dS}{dt} = -\beta SI, \quad \frac{dI}{dt} = \beta SI - \gamma I, \quad \frac{dR}{dt} = \gamma I, \tag{2.7}$$

$$S(0) > 0,\ I(0) > 0,\ R(0) = 0. \tag{2.8}$$

Here S denotes the population of susceptible individuals, I the population of infectious individuals and R the population of removed individuals.

The *epidemiological threshold* quantity is defined by

$$R_0 = \frac{\beta S(0)}{\gamma} (\text{ reproductive ratio}) \tag{2.9}$$

If $R_0 < 1$, then the infected population never increases whereas if $R_0 > 1$ the epidemic "will spread".

2.1.4 Population models and dynamical systems

As suggested by these elementary examples, the modeling of interacting multitype populations leads to a rich area of dynamical systems and there exists an immense literature in this field.

For example the extension of the Lotka-Volterra equations to N interacting species is given by the system:

$$\frac{x_i(t)}{dt} = r_i x_i \left(1 - \sum_{j=1}^{N} \alpha_{ij} x_j\right), \quad i = 1, \ldots, N. \tag{2.10}$$

These multispecies dynamical systems can have very complex behavior including limit cycles or chaotic behaviour. In fact S. Smale (1976) [551] proved that for $N \geq 5$ these systems can exhibit any asymptotic behavior.

2.2 Small population effects

Deterministic models provide good approximations to the growth of large (non-critical) populations but for small populations and "nearly critical populations" it is essential to take account of their inherent discrete nature and randomness.

2.2.1 The Bienamyé-Galton-Watson Branching Process (BGW)

The importance of the fact that individuals produce a random number of offspring and the possibility exists that the population can become extinct led Bienaymé (1845) [33], and Galton-Watson (1874) [591] to introduce this probabilistic model.

The population size at generation n is denotes X_n. Starting with $X_1 = 1$, at each generation each individual gives rise to a random number of children as follows:

$$X_{n+1} = \sum_{k=1}^{X_n} \xi_k, \text{ where the } \quad \{\xi_k\}_{k=1,\ldots,X_n} \text{ are } \textbf{independent}.$$

Generating functions provide a basic tool for developing these processes. The generating function for the offspring distribution is given by:

$$f(s) = E[s^{\xi}] = \sum_{k=0}^{\infty} p_k s^k, \quad 0 \le s \le 1$$
$$f'(1) = m \quad = \quad \text{mean offspring size.}$$

The key relation is

$$E[s^{X_n}] = f_n(s), \text{ where } f_{n+1} = f[f_n(s)].$$

The **extinction probability**, is defined by $q := P(X_n = 0 \text{ for some } n < \infty)$.

Theorem 2.2 *(Steffensen (1930,1932)) If $m \le 1$ (critical, subcritical branching), the $q = 1$. If $m > 1$ (supercritical branching), the q is the unique nonnegative solution in $[0, 1)$*

(2.11) $s = f(s)$.

If $m < \infty$, then

$$\frac{X_n}{m^n} \text{ is a martingale .}$$

Propagation of initial randomness

Theorem 2.3 *(Hawkins and Ulam (1944), Yaglom (1947), Harris (1948), [292]) If $m > 1, \quad \sigma^2 = f''(1) + f'(1) - (f'(1))^2 < \infty$, then*

$$\frac{X_n}{m^n} \to W, \quad \textit{in } L^2 \textit{ and a.s. as } n \to \infty$$

and

$$EW = 1, \quad Var(W) = \frac{\sigma^2}{(m^2 - m)}, \; P(W = 0) = q.$$

2.2.2 Reed-Frost epidemic model

A probabilistic analogue of the SIR epidemic model known as the Reed-Frost model is given as follows. We consider an initial population of susceptible individuals $S_0 = N$ and one infected individual $I_0 = 1$.

$$S_{t+1} \sim \text{Bin}(S_t, (1-p)^{I_t}), \quad t \in \mathbb{N},$$

that is, in each time unit a susceptible individual has probability p of meeting each infected individual and one such contact results in infection. Individuals are infected during one time period so that $I_{t+1} = S_t - S_{t+1}$. If $1 - p = e^{-\lambda/N}$, then Von Bahr and Martin-Löf (1980) [448] showed that as $N \to \infty$ the critical threshold is $\lambda = 1$.

2.2.3 Multitype populations and the Wright-Fisher Model

The celebrated work of Mendel (1865) [455] on the inheritance of traits and its rediscovery around 1900 led to the development of the field of genetics. The modern theory of mathematical genetics was initiated in the work of Wright (1931), (1932) [609],[610] and Fisher (1930) [246]. They introduced a probabilistic model of finite population sampling that serves as a starting point for modern population genetics. This model deals with a population of individuals of different types. As a mathematical idealization they assume that the total population is constant in time and they focus on the changes in the relative proportions of the different types of individual. The key ingredients are:

- Fixed finite population size N
- Typespace (*alleles*)

  $$E_K = \{1, \ldots, K\}$$

- $X_n(i)$ is the number of individuals of type i, at generation n.

Let $\mathcal{N}(E_K)$ denote the counting measures on E. Then dynamics are defined by a Markov chain $X_n = (X_n(1), \ldots, X_n(K))$ with state space

$$\{(x_1, \ldots, x_K) \in \mathcal{N} : \sum_{i=1}^{K} x_i = N\}.$$

The intuitive idea leading to the transition mechanism for the *neutral model* is that first each individual in the nth generation produces a large number of potential offspring. Then in a second stage the population is pruned back (culling) so that the total population remains N (this can be thought of as an analogue of carrying capacity). Based on the *neutral assumption*, that is each of the individuals in produced in the first stage has equal probability of being selected, the $(n+1)$st generation consists of N individuals of types $\{1, \ldots, K\}$ obtained by

- *multinomial sampling* from the empirical distribution

$$\begin{aligned} &P(X_{n+1} = (y_1, \ldots, y_K) | X_n = (x_1, \ldots, x_K)) \\ &\quad = \frac{N!}{y_1! y_2! \ldots y_K!} \left(\frac{x_1}{N}\right)^{y_1} \cdots \left(\frac{x_K}{N}\right)^{y_K} \end{aligned}$$

An important feature of this process is the loss of information (diversity) leading to **fixation**, that is, the long time survival of exactly one type. To see this note that

$$p_n(i) \text{ is a martingale where } p_n(i) = \frac{X_n(i)}{N}$$

and

$$p_n(i) \to 0 \text{ or } 1 \text{ as } n \to \infty \text{ for each } i \text{ w.p.1.}$$

The dual perspective

If we choose k individuals at random from generation $n+1$ and look backwards in time to identify the parents in the nth generation, by an elementary conditional probability calculation, we see that the *each individual in generation* $n+1$ *picks its parents "at random".* This naturally leads to the notion of *identity by descent* introduced by Malécot (1941) [467], that is, two individuals are identical by descent if they have a common ancestor (and no mutations have occurred).

2.3 The Role of Stochastic Analysis

Basic developments in stochastic analysis:

The development of stochastic population modelling was made possible by the remarkable developments in stochastic analysis.

- Markov chains and processes (1906) [446], Kolmogorov (1931), [402]
- Brownian motion Wiener (1923) [607], Lévy (1948) [425]
- Ito stochastic calculus (1942), (1946), (1951) [329]
- Markov processes and their semigroup characterization
  Feller (1951) [242], Itô-McKean (1965).

  Given a Markov process $\{X(t)\}_{t\geq 0}$ with state space E (for example, compact metric space) and $f \in C(E)$ (bounded continuous functions on E) and $x \in E$, let

  $$T_t f(x) = E_x(f(X(t))$$