DIFFUSION PROCESSES IN GENETICS

WILLIAM FELLER
PRINCETON UNIVERSITY

1. Introduction

The theory of evolution provides examples of stochastic processes which have not yet been treated systematically. We find here many open problems whose solution promises new insights into the general theory.

There exists a huge literature on the mathematical theory of evolution and statistical genetics, but existing methods and results are due almost entirely to R. A. Fisher and Sewall Wright.[1] They have attacked individual problems with great ingenuity and an admirable resourcefulness, and had in some instances to discover for themselves isolated facts of the general theory of stochastic processes. However, as is only natural with such pioneer work, it is not easy to penetrate to the mathematical core of the arguments or to discover the explicit and implicit assumptions underlying the theory. In the following an attempt is made to formulate the basic mathematical problems and to discuss their connection with other stochastic processes. Such a systematic approach leads automatically to more general formulations which may be useful at least for a better understanding of the underlying assumptions.

We are concerned with mathematical models of population growth. Relatively small populations require discrete models, but for large populations it is possible to apply a continuous approximation, and this leads to processes of the diffusion type.

By way of introduction we start with the simple *branching process* which became popular in connection with its application to nuclear chain reactions, but which had been previously used by Galton in a discussion of the survival of family names, and by R. A. Fisher in his treatment of the survival of mutant genes. This branching process describes the simplest possible populations: the individuals are of like kind, and there is absolutely no interaction among them so that they are statistically independent. Thus, contrary to a widespread belief, the branching process does not represent an isolated type of stochastic process and is remarkable mainly because of its simplicity. For example, in the case of a growing population the process in its later stages converges to a diffusion process regulated by the Fokker-Planck equation describing the simplest growth (compare appendix II, and end of section 5).

Serious difficulties arise if one wishes to construct population models with in-

This research was done under an Office of Naval Research contract at Cornell University for developing probability theory.

[1] See references [5], [12], and [13]. It is difficult to give useful references to original papers, since these are mostly highly technical and inaccessible to nonspecialists.