

THE DISTRIBUTION OF THE TOTAL SIZE OF AN EPIDEMIC

H. E. DANIELS
UNIVERSITY OF BIRMINGHAM

1. Introduction

This paper examines in some detail the distribution of the total number of cases in an epidemic of the general stochastic type for a closed population. The assumed model is that of Bartlett [2] and McKendrick [11] which Bailey [1] used to study the stochastic analogue of the deterministic threshold theorem (Kermack and McKendrick [10], D. G. Kendall [9]). Bailey obtained recurrence relations from which the required probabilities were computed numerically. His calculations revealed a gradual transition from J -shaped distributions containing only small epidemics for population sizes below the threshold, to U -shaped distributions containing either large or small epidemics but practically no epidemics of intermediate size when the threshold is exceeded. There is also an interesting transitional form of distribution near the threshold value.

In an attempt to understand what motivates an epidemic to behave in this way, Whittle [13] and Kendall [9] constructed different models approximating to the one used by Bailey but easier to handle analytically. Both explained Bailey's results in terms of an initial birth and death process where extinction is certain in the first case and not certain in the second. This work is summarized, with additional references, in the book by Bailey [2]. In a paper presented at this Symposium, Gani [7] develops some recent work by Siskind [12] and himself [6] on a method of obtaining time dependent solutions of the epidemic equations. For the limiting case considered here he shows how the probabilities can be computed by successive multiplication of matrices.

My main object is to arrive at approximate formulae for the distribution of total epidemic size which are appropriate for large populations. The approach differs from that of most other investigations in that the backward equations of the process are used. (See, however, Bartlett [3].) I also find it convenient to work in terms of the number remaining uninfected, rather than the total number of new cases. The technique by which the approximations are obtained was previously used by me in an entirely different context (Daniels [4], [5]). As presented here it should not be regarded as rigorously establishing the approximations, though I have no doubt that the results are correct and numerical comparisons bear this out.