FOR STATISTICIANS AND HEALTH AUTHORITIES

The MATHEMATICAL THEORY *of* EPIDEMICS

N. T. J. BAILEY, M.A.

CHARLES GRIFFIN AND COMPANY LIMITED

THE MATHEMATICAL THEORY OF EPIDEMICS

NORMAN T. J. BAILEY, M.A.

Reader in Biometry, University of Oxford;
Formerly Statistician to the Medical School,
University of Cambridge

LONDON
CHARLES GRIFFIN & COMPANY LIMITED

Printed in Great Britain by John Wright & Sons Ltd., The Stonebridge Press, Bristol

PREFACE

It is just about fifty years ago that the mathematical theory of epidemics, in the modern sense of the phrase, was first started by the work of William Hamer and Ronald Ross. Considerable progress has since been made, and this has been accelerated in recent years by the availability of new mathematical methods of handling random processes. A great number of interesting and valuable results are widely scattered in the literature, but there appears to be no single text-book giving a systematic treatment of the whole field. The present volume attempts to meet this need. Some selection of the existing material has been inevitable, but the bibliography has been made as complete as possible so far as predominantly mathematical references are concerned.

This book is primarily intended for those who wish to learn more about the use of mathematical and statistical methods in understanding the mechanisms underlying the spread of infectious diseases. It is, however, by no means addressed exclusively to professional mathematicians and statisticians, although some sections rely fairly heavily on rather specialized techniques. The theories discussed should be of general interest to all those who are involved in any form of biometrical investigation. Moreover, many of the results obtained may well have considerable relevance to the work of general practitioners, epidemiologists and Medical Officers of Health, who are all concerned more with the practical implications of the theory than with its mode of derivation. Those who do not want to follow the mathematical arguments in detail should be able to learn enough for their purpose from the general discussions of the basic models used and of the consequences that flow from them.

The mathematical theory of epidemics appears at present to be developing fairly rapidly. If it is to continue in the future as a useful branch of applied mathematics then there must be adequate co-operation between mathematician and epidemiologist. The theory is only likely to have valuable applications in so far as it is developed in the context of a proper understanding of the epidemiological realities. Many of the models used in this book are of necessity over-simplified: nevertheless, useful results are already available. Future refinements should enable much further progress to be made.

I have greatly enjoyed writing this book on a subject which has fascinated me ever since Dr. A. M. McFarlan first drew my attention to it at Cambridge in 1948. It is a pleasure therefore to acknowledge my

indebtedness to Dr. McFarlan for stimulating my interest and for many useful subsequent discussions. I have also derived great benefit from continued contacts with Dr. R. E. Hope Simpson, of the Cirencester Public Health Laboratory, who has not only introduced me to many practical aspects of field epidemiology but has also very kindly allowed me to make use of much of his own excellent but unpublished data.

Dr. Hope Simpson also very kindly consented to read and criticize the first draft of Chapters 1, 2, 3 and 9. A similar service was rendered by Professor M. S. Bartlett for Chapter 8; by Mr. D. G. Kendall for Chapters 4 and 5; and by my colleague, Mr. A. M. Walker, who studied and commented on the whole manuscript. I am most grateful for helpful and constructive criticism by all these gentlemen, though they are not, of course, responsible for any errors that remain. I should, in addition, like to thank Professor Bartlett and Mr. D. G. Kendall for allowing me to see several of their own papers in draft prior to publication.

I am indebted to the editors of *Biometrika* and of the *Journal of the Royal Statistical Society*, Series B, for allowing me to draw freely on my own papers in these journals. In particular, the *Biometrika* Trustees have kindly permitted the reprinting of Figs. 4.1, 4.2, 5.1, 5.2 and 5.3, which originally appeared in two papers of mine in *Biometrika*; and, with the agreement of Dr. P. Whittle, the use of Fig. 5.4, which is a redrawn version of a diagram he exhibited in the same journal. I should also like to thank the University of California Press for permission to make considerable use of two excellent papers by Professor Bartlett and Mr. D. G. Kendall in Vol. 4 of the *Third Berkeley Symposium on Mathematical Statistics and Probability*. The Appendix Tables were computed on EDSAC and are published here for the first time by permission of the Director of the University Mathematical Laboratory, Cambridge.

Finally, I should like to take this opportunity of thanking Mrs. Jill Esnouf for drawing some of the diagrams, Mrs. Tamara Hazlewood for her assistance with many of the computations, and last but not least my secretary, Mrs. Daphne Russen, for the rapid production of an extremely neat and legible typescript.

NORMAN T. J. BAILEY

Oxford,
August 1957

CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

The fearful toll of human life and happiness exacted through the ages by widespread disease and pestilence affords a spectacle that is both fascinating and repellent. A recital of the astronomical numbers of casualties suffered in this way by the human race is almost stupefying in its effect, and makes the consequences of all past wars seem almost trivial in comparison. Thus in Europe in the 14th century there were some 25 million deaths out of a population of perhaps 100 million from the Black Death alone. In 1520 the Aztecs lost about half their population of $3\frac{1}{2}$ million from smallpox. The downfall of their empire in 1521 was due more to smallpox than to Cortes. It has been estimated that Russia suffered about 25 million cases of typhus in the years from 1918 to 1921 with a death-rate of approximately 10 per cent. In the world pandemic of influenza in 1919 the total number of deaths is thought to have been in the region of 20 million over twelve months. Examples such as these could be multiplied *ad nauseam*.

Although modern medicine can now do much to alleviate or cure many infectious diseases once they have been contracted, by far the most spectacular results have been in the field of prevention. Malaria, for example, of which there are still said to be about 100 million attacks per year in India, has been eradicated from many areas of the world where it was previously endemic by the relatively simple procedure of draining swamps and marshes. The elimination of poverty and hunger, and the provision of adequate social and public health measures such as quarantine, isolation of infectious cases, provision of clean water supplies, proper disposal of sewage, vaccination and inoculation, etc. have provided the main contributions to the fight against disease. Though in the more advanced communities many diseases once rampant have virtually disappeared, except perhaps for merely sporadic cases, others like influenza, poliomyelitis, infective hepatitis and the

common cold, not only continue to defy prevention but still lack specific cures. All methods of study are therefore welcome, whether clinical, biological, ecological or mathematical.

It is, of course, with the last of these approaches that this book is primarily concerned, though it should be emphasized straight away that such methods are not entirely independent of other disciplines. Thus clinical questions of diagnosis, prognosis and efficacy of treatment often depend on the statistical interpretation of appropriate data. Advice to patients and their immediate contacts is considerably influenced by current views as to when a patient is infectious and by what is known about variations in the incubation period. The cogency of many epidemiological arguments about, for instance, the possibility of virulence changing with time, or even the existence of infectiousness itself, may well depend on whether the effects apparently observed could in fact be due merely to non-significant chance fluctuations. Again, we may be interested in developing mathematical models because of the light they shed on some aspect of the biological mechanism at work, such as the life-cycle of the parasite involved. Alternatively, we can use these models to study the large-scale population phenomena of immediate relevance to any social and public health measures that might be advocated or undertaken. In particular, we want to know more about the transmission and spread of infectious disease, about trying to predict the course of an epidemic, and about the recognition of threshold densities of population which must be surpassed before a flare-up is likely.

In all these matters mathematical and statistical investigations have an important part to play. They originated in the rudimentary medical statistics of Graunt and Petty, who studied the London Bills of Mortality in the 17th century. Progress was slow, but the great sanitary revolution of the mid-19th century, followed by the rise of bacteriology in its second half, had by 1900 created the conditions required for further theoretical developments of the kind treated in this book. The growing availability of mortality statistics served to bring into sharper focus the problems facing public health authorities, while the new discoveries of bacteriology suggested suitable models for more exact mathematical investigation. Since the turn of the century there has been continued and accelerated progress in this field. Although it may never be possible to attain the kind of fine-drawn elaboration

of theory now available in physics, the advancement already achieved by mathematical investigation in biological subjects such as evolution and genetics is extremely encouraging.

Most, but not quite all, of the earlier work on epidemics from about 1900 to 1930 was essentially deterministic in character, that is, it did not take into account the probability aspects of the processes studied. These aspects are of considerable importance in epidemic theory, even with large populations. The deficiency was beginning to be made good from about 1930 onwards by the idea of using chains of binomial distributions to represent successive crops of new cases. As a result of new developments in the 1940's in the handling of stochastic processes further progress became possible. In recent years new ways have been found of dealing with the very variable phenomena that occur in practice. With small groups like individual families it is possible to obtain fairly homogeneous data from large numbers of such units. Analysis of epidemic patterns observed on the basis of specific mathematical models then permits both numerical estimation of parameters of epidemiological importance, such as chance of infection or length of incubation period, and statistical tests of how reliably hypothesis and observation agree. With large groups like whole communities the problems are more difficult, especially if we try to take account of spatial as well as temporal patterns of events. Nevertheless, appreciable success has already been achieved. The newer stochastic versions give much more satisfactory explanations than the older deterministic models of such observed phenomena as *undamped* epidemic waves in recurrent outbreaks, and critical community sizes for the existence of fade-out effects.

The purpose of this book is to give a fairly full account of the mathematical theory of epidemics as it stands at present. Although purely mathematical points will be dealt with if they are essential to the main argument, or if they seem likely to be of use in future developments, the main emphasis will be on the biometrical and epidemiological aspects of the theory. Whereas some investigations are pursued in order to gain insight into the general character of epidemic processes, others are concerned with statistical methods of analysing specific kinds of observational data.

Those readers who are primarily mathematicians or statisticians should find little difficulty anywhere. At the same time it is hoped that much of the book will also be of interest and value to

mathematically inclined biologists, epidemiologists and medical research workers. Since these may not all wish to study the theory in detail, an attempt has been made in each chapter both to explain the general methods of investigation adopted and to make clear the practical consequences that follow. Where possible worked examples are provided to facilitate applications by the reader to fresh data. This is particularly important where statistical estimation of parameters is involved and goodness-of-fit tests are to be performed. It is common nowadays for scientific workers who are not primarily mathematicians to acquire mastery of specific techniques like maximum-likelihood scoring. Since this latter procedure is very frequently required in analysing actual data, full details are given of the appropriate scores and information functions wherever they are needed.

Chapter 2 gives a historical sketch of the development of the mathematical theory of epidemics, but some readers may prefer to omit this on a first reading, returning to it when a fuller acquaintance with the whole field has been obtained. Except for those who already have some knowledge of the subject, the remaining chapters should probably be taken in the order in which they stand. Chapter 3 deals in relatively non-technical language with epidemiological principles in so far as they are required for a proper appreciation of the mathematical theories developed in this book. Standard text-books should be consulted for specialized accounts. Deterministic models of the 'continuous-infection' type are introduced in Chapter 4, and their stochastic counterparts are treated in Chapter 5. The chain-binomial type of probability model is discussed in the next chapter, and this is followed in Chapter 7 by an extension which attempts to provide statistical estimates of latent and infectious periods. Chapter 8 then returns to population problems and uses the models of Chapters 4 and 5 to study recurrent epidemics and endemic states. The problem of detecting infectiousness is dealt with in Chapter 9. Finally, in Chapter 10 there is a short survey of the results so far obtained in mathematical epidemic theory, together with some remarks on the prospects for further research.

It seems quite probable to the writer that considerable advances in mathematical epidemiology will be made in the not very distant future. If these expectations are realized, the results obtained are likely to be of great importance for the prevention of infectious

disease. At one end of the scale we may look for a greater understanding of the detailed processes involved in the transmission of infection; while at the other end further knowledge of large-scale phenomena will have an important bearing on problems facing public health authorities, in so far as it will facilitate prediction of the epidemiological consequences of proposed administrative measures.

Such applications could be of immense value to the community if wisely employed. However, it is not perhaps entirely inappropriate here to point out what is no doubt already realized by certain government departments. Progress in understanding the nature of epidemic processes will not only assist the prevention of infectious disease, but will also increase the power and scope of the deliberately organized outbreaks contemplated by specialists in bacteriological warfare. Although decisions to prosecute such researches must rest primarily with the community as a whole, or at least with its elected representatives, it is the duty of responsible scientists working in special fields to point out to the general public the risks, if any, inherent in their activities.

CHAPTER 2

HISTORICAL SKETCH OF MATHEMATICAL EPIDEMIOLOGY

In this chapter it is proposed to give a short historical account of the development of mathematical theories of the spread of epidemic diseases. It is hoped that this will enable the detailed mathematical discussions appearing later to be seen in perspective, and that a broad view will facilitate the choice of problems for further research. Some readers may prefer to go straight on to the mathematics and return to this chapter later. There is no reason why this should not be done, though as already mentioned the remaining chapters should probably be taken in the order in which they appear, except perhaps by those readers who are to some extent already acquainted with the subject.

2.1 The beginnings

Recorded accounts of epidemic outbreaks and speculation as to possible causes go back at least as far as the ancient Greeks, e.g. the *Epidemics* of Hippocrates (459–377 B.C.), but genuine progress in epidemiology was hardly forthcoming until the 19th century. The spectacular rise of bacteriological science in the second half of that century, due to the researches of Pasteur (1822–95) and Koch (1843–1910), was perhaps the outstanding feature of the commencement of modern scientific achievement in this field. Nevertheless, some progress had already been made on a less fundamental level in the statistical appraisal of records showing the incidence and locality of known cases of disease. Indeed, men like John Graunt (1620–74) and William Petty (1623–87) had in the 17th century paid considerable attention to the London Bills of Mortality. Their work may be taken to mark the beginning of vital and medical statistics and the understanding of large-scale phenomena connected with disease and mortality, but the time was not yet ripe for anything approaching a connected theory of epidemics. In the first place, the requisite mathematical techniques

were themselves only then in process of development, and in the second place there were no sufficiently precise hypotheses about the spread of disease suitable for expression in mathematical terms. Although a good start was being made in the field of physics, particularly mechanics and astronomy, nearly 200 years passed before any real progress was achieved in the biological sphere. It is true that as early as 1546 Fracastorius had postulated a living principle of contagion, which could be spread from person to person. However, it was not until an unmistakable physical basis for the cause of infectious disease had been established in the second half of the 19th century that the stage was set for the development of adequate mathematical theories of large-scale phenomena, as opposed to purely empirical descriptions.

Even before the fundamental advances in the new science of bacteriology, extremely valuable work was going on in the field of what we now call social medicine. By studying the temporal and spatial pattern of cholera cases, John Snow showed in 1855 that the disease was being spread by the contamination of water supplies. In particular there was the celebrated affair of the Broad Street Pump. Later, in 1873, William Budd established a similar manner of spread for typhoid. Parallel to these detailed investigations were the broader studies of statistical returns made by William Farr (1840), who hoped to discover empirical laws underlying the waxing and waning of epidemic outbreaks. These attempts and their later developments are described in the next section.

2.2 Curve fitting and prediction

Apart from the highly successful *ad hoc* studies made by men like Snow and Budd, we have the more deliberate investigation of pooled statistical returns by Farr. His work was very much in the spirit of Graunt and Petty, but was mathematically more sophisticated. In the *Second Report of the Registrar-General of England and Wales*, Farr (1840) effectively fitted a normal curve to smoothed quarterly data on deaths from small-pox, assuming the constancy of 'second ratios' of successive pairs of frequencies. Later in 1866 (letter to London *Daily News*, 17th February, quoted by Brownlee, 1915b) he attempted to use a similar method, based this time on the constancy of third ratios, to predict the course of an outbreak of rinderpest amongst cattle. The curve was fitted to four rising

successive monthly totals and extrapolated values used for prediction. Although observed and predicted curves were both bell-shaped, agreement in detail was not very good. Similar curve-fitting methods used by Evans (1875) on the small-pox outbreak of 1871–2 also met with little real success.

More intensive studies of the same type were later undertaken by Brownlee (1906), who fitted various Pearson curves to epidemic data on many diseases occurring at different times and places. Further investigations of this type were reported in a series of papers (Brownlee, 1909 to 1918). These were all largely of an empirical nature, although to some extent related to the current ideas of Hamer (1906), mentioned in the next section, which involved the use of a specific *a priori* model. Such methods, if successful, would be extremely useful to public health authorities, but they have now been largely abandoned because of their intrinsic inaccuracy. All the same, we may still hope that the development of alternative lines of investigation will eventually permit some kind of predictions to be made, even if these should be more or less vague probability statements.

2.3 Deterministic models

By the end of the 19th century the general mechanism of epidemic spread, as revealed by bacteriological research, and the long familiarity with epidemiological data together made possible developments of a new kind. Hamer (1906) considered that the course of an epidemic must depend *inter alia* on the number of susceptibles and the contact-rate between susceptibles and infectious individuals. The simple mathematical assumptions used by Hamer are basic to all subsequent deterministic theories, and indeed appear in probability versions as well, in suitably modified form. Moreover, by using these ideas in a simple way, Hamer could deduce the existence of periodic recurrences. This was taken up again later by Soper (1929), as mentioned below.

In the meantime Ross (1911 and later) was working with a more developed mathematical model taking into account a whole set of basic parameters describing various aspects of the transmission of malaria. It is important to notice that although Ross employed the idea of chance or probability in formulating his basic equations, these were actually still deterministic in character. This means that, for such a model, the future state of the epidemic

process can be determined precisely when we are given the initial numbers of susceptibles and infectious individuals, together with the attack-, recovery-, birth- and death-rates. For the first time it was possible to use a well-organized mathematical theory as a genuine research tool in epidemiology. Deductions from the theory, possibly unforeseen and unexpected, could now be tested out in practice.

More elaborate mathematical studies of the same general type were later undertaken by Kermack and McKendrick (1927 to 1939). A greater degree of generality was introduced, including variable rates of infection, recovery, etc. These authors also considered the problem of endemic disease and related their findings (see also McKendrick, 1940) to experimental mouse epidemics. The most outstanding result obtained was, however, the celebrated threshold theorem, according to which the introduction of infectious cases into a community of susceptibles would not give rise to an epidemic outbreak if the density of susceptibles were below a certain critical value. If, on the other hand, the critical value were exceeded then there would be an epidemic of magnitude sufficient to reduce the density of susceptibles as far below the threshold as it originally was above.

Further deterministic work specifically associated with measles was carried out by Soper (1929). Although his basic relationship was written as a difference equation, it was in essence very similar to the differential equations of other writers. The most important result here was the discovery that the basic assumptions entailed, so far as recurrent epidemics were concerned, a *damped* train of harmonic waves. Published data on measles, however, while exhibiting a marked oscillation in incidence from year to year, show no tendency to damping. Soper believed, wrongly, that allowance for an incubation period would remove the damping. It is the essential failure of such deterministic models to square with the facts that has led to their abandonment in many quarters and consequent replacement by corresponding probability, or stochastic, representations.

2.4 Stochastic models

As epidemiological data became more extensive and on occasion dealt with much smaller groups than those relevant to returns for large areas, the elements of chance and variation became ever

more prominent. This was specially evident when small family or household groups were contemplated. The need for some kind of probability model was becoming increasingly necessary.

McKendrick (1926) was apparently the first to publish a genuinely stochastic treatment of an epidemic process. Whereas in deterministic models one takes the actual *number* of new cases in a short interval of time to be proportional to the numbers of both susceptibles and infectious cases, as well as to the length of the interval, McKendrick assumed that the *probability* of one new case in a short interval was proportional to the same quantity. This is a 'continuous-infection' model and entails an individual being himself infectious from the instant he receives infection until the moment he dies, recovers or is isolated. In the paper quoted, examples were given of probability distributions for the total number of cases in a household when infection was introduced from outside. This brilliant pioneering effort did not unfortunately attract much attention, and similar models were not again investigated until twenty years later. No doubt the absence of satisfactory methods of handling such models had much to do with this lapse, and it is curious to reflect that McKendrick himself subsequently embarked, with Kermack, on the series of deterministic investigations already mentioned.

An alternative probability treatment by Greenwood (1931) appearing five years later did, however, establish itself. Moreover, similar work was independently in progress in the United States, where in 1928 Lowell J. Reed and Wade Hampton Frost were already using the same kind of ideas in lectures and discussions (see Wilson and Burke, 1942; Abbey, 1952). The new model assumed that the period of infectiousness was comparatively short and that the latent and incubation periods could be regarded as approximately constant. Starting with a single case in a closed group (or several simultaneously infectious cases), new cases would then occur in a series of stages or generations. We should, under suitable conditions, expect the cases occurring at any stage to have a binomial distribution depending on the numbers of susceptibles and infectious individuals present at the previous stage. We should thus have a chain of binomial distributions.

Greenwood's treatment is fully stochastic in the sense that once the probability element has been introduced, *via* the chance of contact adequate for an infectious person to transmit disease to a