Probability distributions, stochastic processes and infectious disease models

Ping Yan
Centre for Infectious Diseases Prevention and Control
Public Health Agency of Canada
Surveillance and Risk Assessment Division
Ottawa, Canada

Abstract: This talk addresses two distinct phases of a typical infectious outbreak: the initial phase when the depletion of the susceptible population is ignorable; and the final phase when a proportion of the susceptibles escapes from the infection (final size). Based on personal experience working in a public health agency, examples from some epidemiology studies will be given to demonstrate how mathematical theories are applied in public health.

For the initial phase, there are two approximations to describe the spread of an infectious disease: (i) the branching processes; (ii) the exponential growth curve. The former is based on generation concept characterized by a mean parameter: the reproduction number (Ro); whereas the latter is based on trend over calendar time characterized by an exponential growth rate as its parameter. It is common to use the empirically-observed early exponential growth rate to estimate Ro for different infections in different settings using formulae derived when both the latent and infectious periods follow exponential distributions. There has been numerous recent literature involving mathematical models related to SARS, pandemic influenza and other infectious diseases in some very high impact journals (e.g. Nature, Science, etc.) by just doing this. This talk will first show that all these formulae are special cases when the latent and infectious periods are gamma-distributed where a closed-form solution for Ro was established by Anderson, D. and Watson, R. (Biometrika, 1980). This talk will use continuous time branching processes to establish a more general result that takes that published in Anderson, D. and Watson, R. (1980) as its special case. The implication of this general result has two folds: (1) in general, it is *not* possible to deduce the reproductive number from the exponential growth rate and the mean latent and infectious periods alone; one needs to know the distributions of the latent and infectious periods, as well as the (perhaps time-dependent) disease transmission rate; (2) the following three aspects: (i) the intensity of infectious contacts as a counting process; (ii) the distribution of the latent period during which the infected individual does not transmit; (iii) the distribution of the infectious period, independently shape the relationship between the growth rate and the reproduction number.

For the final phase, this talk will address the final size along the line of a central limit theorem by von Bahr and Martin-Lof (1980) and its generalities. Unlike the exponential growth rate during the initial phase, there is a general monotone relationship between asymptotic mean final size and Ro. Under some fairly general conditions, this relationship is invariant if there exists a latent period, an arbitrarily distributed infectious period and with any number of infectious stages and/or a stage during which infectives are isolated. This generality allows one to modify and use the final size equation to formulate Operations Research framework where the objective is to minimize a function of the final size (e.g. hospitality, absenteeism, deaths, etc.) with constraints such as limited resources.

(If time allows, this talk may also address related topics on the "degree distributions" when the transmission of diseases is viewed as a realization of random graphs, and show how the properties of the infectious period distribution shapes these degree distributions.)

Acknowledgement: This talk is built upon my studies of early work by Professors Peter Jagers and Andrew Barbour on branching processes and on modelling the transmission of diseases using stochastic processes.