Mathematical Statistics
Stockholm University
Research Report **2014:3**,
http://www.math.su.se

Introduction to statistical inference for infectious diseases

Tom Britton and Federica Giardina

November 12, 2014

Abstract

In this paper, we first introduce the general stochastic epidemic model for the spread of infectious diseases. Then we give methods for inferring model parameters such as the basic reproduction number R_0 and vaccination coverage v_c assuming different types of data from an outbreak such as final outbreak details and temporal data or observations from an ongoing outbreak. Both individual heterogeneities and heterogeneous mixing are discussed. We also provide an overview of statistical methods to perform parameter estimation for other stochastic epidemic models. In the last section we describe the problem of early outbreak detection in infectious disease surveillance and statistical models used for this purpose.

Keywords: Stochastic epidemic models, basic reproduction numbers, vaccination coverage, MCMC, infectious disease surveillance, outbreak detection.

1 Introduction

Infectious disease models aim at understanding the underlying mechanisms that influence the spread of diseases and predicting disease transmission. Modelling has been increasingly used to evaluate the potential impact of different control measures and to guide public health policy decisions.

Deterministic models for infectious diseases in humans and animals have a vast literature, e.g. Anderson and May (1991); Keeling and Rohani (2008). Although these models can sometimes be sufficient to model the mean behaviour of the underlying stochastic system and guide towards parameter estimates, they do not allow the quantification of the uncertainty associated to model parameters estimates (Becker, 1989). Stochastic models (Andersson and Britton, 2000; Britton, 2004; Diekmann et al., 2013), can be used to infer relevant epidemic parameters and provide estimates of their variability.

Infectious disease data are commonly collected by surveillance systems at certain space and time resolutions. The main objectives of surveillance systems are early outbreak detection and the study of spatio-temporal patterns. Early outbreak detection commonly relies on statistical algorithms and regression models for (multivariate) time series of counts accounting for both time and space variations.