Statistical inference for epidemic models approximated by diffusion and Gaussian processes

Romain GUY[1,2]
Joint work with C. Laredo[1,2] and E. Vergu[1]

January 30th 2013

[1] *UR 341 Mathématiques et Informatique Appliquées, INRA, Jouy-en-Josas, France*
[2] *UMR 7599 Laboratoire de Probabilités et Modèles aléatoires, Université Paris Diderot, Paris, France*

Abstract

Estimation of key parameters of epidemic models is still a challenging problem despite the progress made last years through computer intensive algorithms. From an applied perspective, it is classical to use compartmental SIR-like models (Susceptible-Infectious-Removed), where each individual is, at a given time, in one of these three mutually exclusive states. These models can be described in various mathematical frameworks and, amongst them, a natural representation is a multidimensional continuous-time Markov jump process (Z_t). However, epidemic data are often partially observed and temporaly aggregated and the tractability in large populations of such processes is difficult. In this context, diffusion processes allow shedding new light on inference problems of epidemic data.

In this talk, we extend the results of [ETH05] to show the interest of developping a statistical framework fitting epidemic models based on diffusion or Gaussian processes. Indeed, the normalization by the population size N of the continuous time Markov jump process (Z_t) leads, as $N \to \infty$, to the solution of an ODE system. Recentering (Z_t) yields a multidimensional Gaussian process. Another approximation, based on infinitesimal generators, leads to a diffusion process with small diffusion coefficient $(\frac{1}{\sqrt{N}})$, close to the previous Gaussian process. We present the results of [GUY12] on discretely observed diffusion processes with small diffusion coefficient, which yields good results for epidemics. We study minimum contrast estimators derived from the likelihood of the Gaussian process. We obtain consistent and asymptotically normal estimators for the parameters in two different asymptotics: fixed number n of observations and $N \to \infty$; $n \to \infty$ and $N \to \infty$ simultaneously. Then, we propose a correction term to the previous contrast process which yields better estimates non asymptotically ([GUY13]). Finally, we use simulations of several epidemics (SIR, $SEIR$, $SIRS$ models) with different characteristics (R_0,N,n) to compare the performances of the estimators.

References

[ETH05] Ethier SN, Kurtz TG. (2005) Markov processes: characterization and convergence. *Wiley Series in Probability and Statistics, Wiley.*

[GUY12] Guy R., Laredo C., Vergu E. (2012) Parametric inference for discretely observed multidimensionnal diffusions with small diffusion coefficient. *ArXiv: 1206.0916*

[GUY13] Guy R., Laredo C., Vergu E. (2013) Statistical inference for epidemic models approximated by diffusion processes. *To be submitted*