



University of  
Zurich<sup>UZH</sup>

# Epidemiology

Data Analysis course 2016 – Project n° 3

Elena Graverini, Andreas Weiden

November 25, 2016

## 1 Science and medicine

Scientists have a tendency to describe various types of natural and artificial processes in terms of mathematical models. This approach is useful when looking at the future: it allows to make predictions and foresee the range of possible outcomes of a system, given its initial conditions.

Mathematical models can also be useful when dealing with contagious diseases. In this project, you will use your abilities in data modelling and analysis in order to help a community devise a plan against the epidemic propagation of a disease.

A dangerous and highly contagious viral disease, similar to the smallpox, has been observed in a few towns. Once an individual has been infected, the first symptoms appear very quickly, and the health of the individual degenerates rapidly, leading to his or her death. The file

`/disk/puma/da/projects/epidemiology_progress.dat`

contains medical records from 88 known cases. Each entry corresponds to the number of hours the individual survived after experiencing the first symptoms of the disease.

Even less data are available concerning the contagiousness of the illness. Old archives from the past have been explored and information has been extracted that can be used to model the present epidemic. The file

`/disk/puma/da/projects/epidemiology_contagion.dat`

has two columns. The first column reports the percentage of healthy people that incurred the disease after having been in contact, for one day, with the number of infected individuals quoted in the second column.

## 2 Tasks

You are asked to review these findings and use them to design a software simulating the progress of the epidemic.

- Extract from the provided data the mean expected survival time of an infected individual. You can use an unbinned Maximum Likelihood fit. Compute the uncertainty on this result.
- Extract from provided data the probability that an individual gets infected, as a function of the number of already infected people around him or her. You can use an unbinned Maximum Likelihood fit. Compute the uncertainty.
- Now consider a small village of 100 inhabitants. Keep your model simple: assume that each individual is in daily contact with all the others. Assume that initially one single inhabitant gets infected. Develop a simulation showing the evolution in time of:
  - the number of still healthy inhabitants
  - the number of infected inhabitants
  - the number of deceased inhabitants.
- Use your simulation to answer important questions such as:
  - after how many days does the epidemic stop (i.e. after how many days do people stop getting infected)?
  - how many people will survive the epidemic?
  - how much time do authorities have in order to put protection measures in place (e.g. quarantine, vaccination campaigns etc.) before 30% of the population is either infected or dead?
- Estimate the uncertainties on your predictions, for example by varying the values of the parameters that you extracted from the given data within their uncertainties. Use your simulation to determine how the previous result changes when accounting for these uncertainties.