Exercises for "het Project 2007"
Martin Bootsma
25-09-2007

Exercise 1

(a) Draw the compartments for the SIR-model in a closed population and convince yourself that under the assumption of exponentially distributed waiting times and homogenenous mixing, this results in the following set of differential equations:

$$\frac{d}{dt}S(t) = -\lambda(t)S(t)$$

$$\frac{d}{dt}I(t) = \lambda(t)S(t) - \alpha I(t)$$

$$\frac{d}{dt}R(t) = \alpha I(t)$$

with $\lambda(t) = \beta I(t)$.

(b) Determine the basic reproduction number R_0 and the real-time exponential growth rate r on the basis of this model.

(c) Show that $\frac{\alpha}{\beta}\log S - S - I$ is a conserved quantity. (Hint: Consider $\frac{dI}{dS}$ and seperate variables.)

(d) Derive an equation for the final size of the epidemic. (Hint, what is $S(-\infty)$ and $S(\infty)$?)

(e) Analyze this final size equation graphically. Both for $R_0 > 1$ and $R_0 < 1$.

(f) What will change in the box diagram and in the differential equations if birth and death is added? and how do R_0 and r change?

(g) Calculate the equilibria for this system including birth and death. Are all equilibria biologically relevant?

(h) Can you say something about stability of the equilibra?

Exercise 2: The Reed-Frost model

We consider the spread of an infectious disease, e.g., common cold, in a population of N individuals where N is small. We assume an SIR-structure and we look at the spread in discrete time, more specifically, on a generation basis. This is a realistic assumption if the latency period is much longer than the infectious period. If this is the case, new infections will occur in generations. We consider a Markovian model, i.e., the event probabilities in a generation only depend on the state of the epidemic in the previous generation. Let X_j and Y_j be the number of susceptibles and infectives respectively at time (or generation) j. The number of recovered individuals at time j is given by $N - X_j - Y_j$ and all infectious individuals will be recovered on etimestep later. The conditional probabilities for the chain-binomial Reed-Frost model are given by:

$$P(Y_{j+1} = y_{j+1}|X_0 = x_0, Y_0 = y_0, \ldots, X_j = x_j, Y_j = y_j) =$$
$$P(Y_{j+1} = y_{j+1}|X_j = x_j, Y_j = y_j) = \binom{x_j}{y_{j+1}}(1 - q^{y_j})^{y_{j+1}}(q^{y_j})^{x_j - y_{j+1}}$$

(a) What is the biological meaning of q and explain the biological meaning of this formula.

(b) What happens if $y_j = 0$?

(c) Define $Z = \sum_{j=1}^{\infty} Y_j$. What is the meaning of Z? Calculate $P(Z = 0|X_0 = 10, Y_0 = 1)$, $P(Z = 1|X_0 = 10, Y_0 = 1)$, $P(Z = 2|X_0 = 10, Y_0 = 1)$, and $P(Z = 3|X_0 = 10, Y_0 = 1)$. Note that this model is not suitable when Z becomes large.

Exercise 3 Fitting a model to data

(a) Consider the following data for a measles epidemic (data from Aycock). Assume that a Reed-Frost type model is suitable to analyse this data set.

t	0	1	2	3	4	5
X_t	117	108	86	25	12	12
Y_t	1	9	22	61	13	0

(a) Calculate how likely each transition from time t to time $t+1$ is as a function of q according to the model.

(b) Calculate at each time step which value of q gives the maximum likelihood

(c) Assume q is constant throughout all time intervals. Determine numerically which value of q yields the maximum likelihood.

(d) Optional: How well do the data fit the model?

Exercise 4 Heterogeneity

Consider a population divided into two subpopulations, which are indicated the the labels 1 and 2. We look at a sexually transmitted disease, with sexual intercourse as a contact at which the disease will spread if one of the partners has the disease. Assume that population 1 consists of promiscuous individuals who have $c_1 = 100$ contacts per year, while individuals of population 2 have $c_2 = 10$ contacts per year. Assume there are $N_1 = 10^5$ individuals of type 1 and $N_2 = 10^7$ individuals of type 2. Assume that a fraction $\frac{c_2 N_2}{c_1 N_1 + c_2 N_2}$ of the contacts of a 1-individual are with 2-individuals and a fraction $\frac{c_1 N_1}{c_1 N_1 + c_2 N_2}$ of the contacts of a 2-individual are with 1-individuals. Assume the infectious period is exactly 1 year. What is R_0?

(a) Calculate the mean number of contacts per year in this population.

(b) How many infections does an infectious 1-individual generate among 1- and among 2-individuals if all of his contacts are susceptible? And how many infections does a 2-individual generate among both groups?

(c) Calculate R_0.

(d) What is the effect of heterogeneity? Can you relate this to a 'real' disease?

(e) Make a graph (numerically) how the fraction of S, I and R-individuals is changing during the epidemic when at time $t = 0$ $S_1 = 9999$, $I_1 = 1$, $R_1 = 0$, $S_2 = 10^7$, $I_2 = R_2 = 0$. Do this for 1-individuals, 2-individuals and for the population as a whole.