Martina Favero
Department of Mathematics
Stockholm University
Contact: martina.favero@math.su.se

MT5020: The mathematics and statistics of infectious disease outbreaks

Written Exam, June 2, 2026

Allowed to bring to exam: Pocket calculator (not phone). Summary or formula sheet is **not allowed**.

Exam points: Each exercise gives a prescribed maximum points, with an overall maximum sum of 35 points. How these affect final grade is explained on course page.

Your solutions: (that you turn in) must be clearly written and answers need to be motivated (except Exercise 1).

Final size equation: If the reproduction number equals R_0, then the final fraction τ^* of infected is the largest solution to the equation $1 - \tau = e^{-R_0\tau}$ (for a homogeneously mixing community with no prior immunity).

Problem 1 [5P]

Multiple choice exercise. Just answer i), ii), iii) or nothing on each of the 5 subexercises. No motivation or computation! Each correct answer gives 1 point and each wrong answer gives -1 point (so better not to answer if you don't know). A negative total sum gives 0 points on the exercise.

(a) Which one of the following statements is *false*?

i) If $R_0 \leq 1$, there will be no major outbreak.

ii) If $R_0 > 1$, there will be a major outbreak.

iii) If $R_0 > 1$, there could be a major outbreak.

(b) When is the deterministic SIR model more appropriate?

i) At the very beginning of an outbreak, when only a few cases have occurred.

ii) In the middle of a large epidemic, when many individuals are infected.

iii) At the very end of an epidemic, when case counts are declining towards zero.

(c) When is a branching-process approximation suitable?

i) At the very beginning of an outbreak, when only a few cases have occurred.

ii) In the middle of a large epidemic, when many individuals are infected.

iii) At the peak of a large epidemic, when the number of cases starts to decrease.

(d) Which of the following is the time between symptom onset in a primary case and symptom onset in a secondary case?

i) The generation interval

ii) The serial interval

iii) The incubation period

(e) Which of the following do you need to compute R_0 from the Euler-Lotka equation?

i) The exponential growth rate and the infectious period

ii) The exponential growth rate and the generation time distribution

iii) The final fraction infected and the generation time distribution

Problem 2 [10P]

Consider a highly infectious disease with a rather high vaccination coverage (e.g. measles) in a homogeneously mixing community. Suppose a community fraction $v = 0.9$ is vaccinated with a vaccine having 100% efficacy and suppose that we know $R_0 = 15$.

(a) What is the reproduction number R_v? Is the vaccination coverage v sufficient to prevent an outbreak? **[2P]**

(b) What is the critical vaccination coverage v_c required to surely prevent an outbreak (i.e. the minimal vaccination coverage to obtain herd immunity)? **[2P]**

(c) Suppose an outbreak occurs and a fraction τ^* among those who are not vaccinated becomes infected. Provide a final size equation for τ^* taking into account the prior immunity given by the vaccine (a fraction $v = 0.9$ is vaccinated prior to the outbreak). **[3P]**

(d) Suppose that prior to an outbreak various preventive measures are introduced to reduce the reproduction number by a fraction p (and a fraction $v = 0.9$ is vaccinated). How large must p be to surely prevent an outbreak? **[3P]**

Problem 3 [10P]

Consider the Reed-Frost model (in each generation, each infectious individual can infect each of the susceptible individuals independently with probability p and then recovers). Suppose that in a household of 4 individuals, one individual is initially infected outside of the household, the other 3 are susceptible.

(a) Compute the probability of the transmission chain $1 \to 2 \to 0$, that is, the probability that the externally infected individual infects two other individuals (first generation) who in turn infect nobody else. **[2P]**

(b) List all possible transmission chains. **[2P]**

(c) Compute the probability that all individuals except one become infected, i.e the final size Z is equal to 2. **[2P]**

(d) Compute the distribution of the final size Z (i.e. $\mathbb{P}(Z = z)$ for all possible values z of Z). **[2P]**

Hint: probabilities sum up to one, so you can skip the explicit calculation for one possible value of Z (choose wisely).

(e) Suppose $p = 1/2$. How many can we expect to become infected on average? **[2P]**

Problem 4 [10P]

The deterministic SIR epidemic with demography is defined as follows:

$$\begin{aligned} s'(t) &= \mu - \beta s(t)i(t) - \mu s(t) \\ i'(t) &= \beta s(t)i(t) - \gamma i(t) - \mu i(t) \\ r'(t) &= \gamma i(t) - \mu r(t) \end{aligned}$$

The parameter μ is the (constant) birth rate of susceptibles, and all individuals die at rate μ (so no increased death rate from the disease). Please note that $s'(t) + i'(t) + r'(t) = 0$ so the population size remains constant.

Suppose we want to modify this model to the situation where immunity wanes at constant rate, meaning that all recovered individuals become susceptible again at rate ν. This model is called the SIRS epidemic with demography.

a) Define this model and make sure that $s'(t) + i'(t) + r'(t) = 0$ still holds. **[2P]**

b) Compute R_0 for the model. R_0 equals the average number of infections by an infectious individual before recovering (or dying) in a fully susceptible community. **[3P]**

c) Compute the endemic levels, i.e. the values $\tilde{s}, \tilde{i}, \tilde{r}$ for which all derivates are 0 (assuming $R_0 > 1$). **[3P]**

d) Compute R_0 and the endemic levels if: the infectious period has mean $1/\gamma = 1$ week, the immune period has mean $1/\nu = 2$ years, and life length has mean $1/\mu = 80$ years, and $\beta = 2/\text{week}$ (make sure to convert to one single time unit). **[2P]**