Description

The following system of differential equations describes the evolution in time of the number of susceptible (S), infectious (I) and recovered (R) individuals in the SIR compartmental model. The additional parameter m allows for deviations of the epidemic from the deterministic solutions. We assume a constant population: $S(t) + I(t) + R(t) \equiv N$.

$$\begin{cases} \frac{dS}{dt} = \mu(m-1)I - \frac{\beta}{N}IS \\ \frac{dI}{dt} = -\mu(2m-1)I + \frac{\beta}{N}IS \\ \frac{dR}{dt} = \mu m I \end{cases} \tag{1}$$

Parameters:

- $\mu = 0.2$ transmission rate (known);
- $N = 10^3$ population size (known);
- $\beta > 0$ transmission rate (to be estimated);
- $m \geq 0$ trajectory parameter (to be estimated).
- Initial conditions at $t = 0$ (known): $I(0) = 10$, $R(0) = 0$.

You are given the data `obs_I.csv`, containing the number of infectious individuals during the epidemic wave, every $\Delta t = 1$. The data follow a typical, bell-shaped epidemic wave: infections grow in the initial phase, then peak, then decrease until the wave subsides.

Tasks

Your overall task is to estimate β, m from the data. The data go from $t = 0$ to $T = 60$, i.e., they report infections during 61 time steps. In reality, you might want to be able to estimate the epidemic parameters before the epidemic wave is over. Therefore, you will estimate the parameters at varying $T_{obs} \leq T$. At fixed T_{obs}, you assume you have access to data only from $t = 0$ to $t = T_{obs} - 1$.

To infer the parameter, you should use two approaches

1. one approach of your choice, that you consider to be appropriate,
2. neural simulation-based inference, which can be performed using this library: https://sbi.readthedocs.io/en/latest/index.html (you can choose any of the implemented algorithms).

You should briefly compare the performance of the two approaches.

Specific tasks

1. Estimate the parameters in the initial epidemic phase ($T_{\text{obs}} \approx 15$);
2. Estimate the parameters towards the end of the epidemic $T_{\text{obs}} \approx T$;

3. Can you obtain unique and well-identified estimates of β and m using only data from the initial phase of the epidemic ($T_{\mathrm{obs}} \approx 15$)? If not, explain why.

Requested output

You can present your results in a brief presentation. Please detail how you decided to approach the problem, your thought processes, the choices you made and their justifications. Highlighting your problem-solving skills is more important that getting an accurate value of the parameters.