Extended SIR Model

We consider an extension of the SIR model from epidemics. We will use the model proposed in the paper of Ferretti et al. Consider a population of size N, we assume N is fixed, so there is no migration. At day t we say $S(t)$ is the number of susceptible people, $I_k(t)$ is the number of people that were infected on day $t-k$ for $k = 1, 2, ..., K$, where K is the estimated duration of infectiousness. $H(t)$ is the number of hospitalised patients on day t, with hospitalised we mean patients on the IC. $D(t)$ and $R(t)$ denote the number of deceased and recovered patients on day t respectively.

In the model proposed in the paper, $Y(t, \tau, \tau')$ is introduced as the number of individuals at time t who where infected at a time $t-\tau$ by individuals who where in turn infected at a time $t-\tau'$. We want to discretize this and therefore introduce $Y^*(t, \tau)$, as the number of people that are infected in time interval $[t, t+1)$ by people that where infected in time interval $[t-\tau, t-\tau+1)$ (or in discrete terms, the number of people that where infected on day t by someone infected on day $t-\tau$). This gives us $Y^*(t,\tau) = \int_\tau^{\tau+1} \int_t^{t+1} Y(t', 0, \tau') dt' d\tau'$.
Now we will consider the following model.

$$
\begin{aligned}
S(t+1) &= S(t) - \sum_{k=1}^{K} Y^*(t,k) \\
I_1(t+1) &= \sum_{k=1}^{K} Y^*(t,k) \\
I_k(t+1) &= I_{k-1}(t) - \gamma_{k-1,H} I_{k-1}(t) \qquad k = 2, 3, ..., K \\
H(t+1) &= H(t) + \sum_{k=1}^{K} \gamma_{k,H} I_k(t) - (\gamma_{H,D} + \gamma_{H,R}) H(t) \\
R(t+1) &= R(t) + (1 - \gamma_{K,H}) I_K(t) + \gamma_{H,R} H(t) \\
D(t+1) &= D(t) + \gamma_{H,D} H(t)
\end{aligned}
$$

Here $\gamma_{k,H}$ is the rate of patients that were infected k days ago and have to go to the intensive care. $\gamma_{H,D}$ and $\gamma_{H,R}$ are the rates at which hospitalized patients become deceased and recovered respectively.

We will use results from the Science paper. One of the assumptions that is made is that for $I(t)$, the total number of infected patients at time t, it holds that there is some r such that $\frac{d}{dt} I(t) = e^{-rt}$. This assumption does not necessarily hold when we plug it into an SIR-model, as Sander Borst pointed out.

Let ϵ_I be the efficacy of case isolation, and ϵ_T the contact tracing plus quarantine efficacy. Then we define

$$
f(\tau', \hat{\tau}) = \beta(\hat{\tau})(1 - \epsilon_I s(\hat{\tau})) \left(1 - \epsilon_T + \epsilon_T \frac{1 - s(\tau')}{1 - s(\tau' - \hat{\tau}))} \right).
$$

Where $\beta(\tau)$ is the mean rate at which an individual infects others at time τ, and $s(\tau)$ is the probability of having started showing symptoms after being infected for time τ. Then, by the Science article, we have

$$
Y(t, 0, \tau) = \int_\tau^\infty f(\tau', \tau) Y(t, \tau, \tau') d\tau'.
$$

Sander Borst used this to derive the following approximation for $Y^*(t, \tau)$.

$$
\begin{aligned}
Y^*(t,\tau) &= \int_\tau^{\tau+1} \int_t^{t+1} Y(\hat{t}, 0, \hat{\tau}) d\hat{t} d\hat{\tau} \\
&= \int_\tau^{\tau+1} \int_t^{t+1} \int_{\hat{\tau}}^\infty f(\tau', \hat{\tau}) Y(\hat{t}, t\hat{a}u, \tau') d\tau' d\hat{t} d\hat{\tau} \\
&= \int_\tau^{\tau+1} \int_t^{t+1} \int_{\hat{\tau}}^\infty f(\tau', \hat{\tau}) Y(\hat{t} - \hat{\tau}, 0, \tau' - \hat{\tau}) d\tau' d\hat{t} d\hat{\tau} \\
&= \int_\tau^\infty \int_t^{t+1} \int_\tau^{\min(\tau', \tau+1)} f(\tau', \hat{\tau}) Y(\hat{t} - \hat{\tau}, 0, \tau' - \hat{\tau}) d\tau' d\hat{t} d\hat{\tau} \\
&\approx \int_\tau^\infty \int_t^{t+1} \int_\tau^{\min(\tau', \tau+1)} f(\tau', \hat{\tau}) Y(\lceil \hat{t} - \hat{\tau} \rceil, 0, \lceil \tau' - \hat{\tau} \rceil) d\tau' d\hat{t} d\hat{\tau} \\
&= \sum_{\tau'=\tau}^\infty Y^*(t - \tau, \tau' - \tau) \int_{\tau'}^{\tau'+1} \int_\tau^{\min(\hat{\tau}', \tau+1)} f(\hat{\tau}', \hat{\tau}) d\hat{\tau} d\hat{\tau}'
\end{aligned}
$$

Taking into account the probability of someone being susceptible to the virus we get the following recursion for $Y*$.

$$Y^*(t,\tau) \approx \frac{S(t)}{N} \sum_{\tau'=\tau}^{\infty} Y^*(t-\tau, \tau'-\tau) \int_{\tau'}^{\tau'+1} \int_{\tau}^{\min(\hat{\tau}', \tau+1)} f(\hat{\tau}', \hat{\tau}) d\hat{\tau} d\hat{\tau}'$$

Some notions: We assume that an infected patient remains infectious for K days, so in the we can replace (and, for implementation, should replace) the infinite sum by a sum summing up to K.
In order to use this recursion we should have values for $Y^*(t,\tau)$ where $-K \leq t < 0$.
The implementation takes the longest time on calculating the integrals, the integrals only have to be calculated once for running the model, but if we want to change values for efficacy etc they have to be calculated again.

The article we referred to can be found here:
`https://science.sciencemag.org/content/sci/early/2020/04/09/science.abb6936.full.pdf` With supplementary materials:
`https://science.sciencemag.org/content/sci/suppl/2020/03/30/science.abb6936.DC1/abb6936_Ferretti_SM.pdf`

Parameter Estimation

In order to implement the model, we should estimate the parameters and initial values of the model.

We use the Science article to get estimates on $\beta(\tau)$, $s(\tau)$ and K. In the results of the article we see that on average the infectiousness is very small after 13 days, so we set $K = 13$. We have $\beta(\tau) = R_0 \cdot w(\tau)$, where R_0 is estimated to be 2, and $w(\tau)$, the probability density function of the generation time, estimated as Weibull distributed with shape parameter 2.826027 and scale parameter 5.665302. $s(\tau)$ is estimated as being lognormal distributed with meanlog 1.644 and sdlog 0.363.

We assume that patients will go to the hospital only after their period of being infected for K days. For the values of $\gamma_{k,H}$, this means $\gamma_{k,H} = 0$ for $k < K$. In the article `https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf` some estimates are given for the percentage of symptomatic cases that require critical care per age-group. Note that this is the percentage of symptomatic cases, so this does not fit to the model where we also have asymptomatic cases. Also, this table is made more than a month ago, so they are not based on the most recent data.
Using this and the age distribution in the Netherlands from `https://opendata.cbs.nl/statline/?dl=308BE#/CBS/nl/dataset/7461bev/table` we can get the following estimates.

$\gamma_{K,H}$	$\gamma_{H,D}$	$\gamma_{H,R}$
0.026445154	0.003718003	0.096281997

These parameters are also based on an average time on the IC of 10 days.

Now we need to find initial values. There is an estimate that 3% of the people has already recovered from the virus. So we set $R(0) = 0.03N$. It is very hard to estimate $I_k(0)$, something we could do is look at the number of new deceased patients in the previous K days, and use these to estimate the number of infected people the days before that. In this case we set $t = 0$ at a moment in the past. $H(0)$ and $D(0)$ can be obtained by data of the RIVM. $S(0)$ is the remaining population size.

We should also estimate $Y^*(t,\tau)$ for $-K \leq t < 0$. We will estimate this as $Y^*(t,\tau) = \int_{K-\tau}^{K-\tau+1} \beta\tau' d\tau' I_{-t}(0)$.