Mathematical Modeling of Epidemics

Daniel Paliura

3/23/2021

Contents

Introduction

This document was created in addition to 'Control Theory' subject second laboratory work 'Mathematical Modeling of Epidemics'. This document on GitHub is here It contains the research for a dependence of input parameters on the problem solution. One of the well-known models for modeling epidemics is SIRD (Susceptible, Infected, Recovered, Dead). All parameters will be set randomly in recommended intervals that are respective to the COVID-19 pandemic. Also, the SIRD model can be complicated with some special values.

I chose the model with testing, so the system is described with additional value J - Justified (I guess, it could mean something other). Justified - people who tested with a positive result, in other words, people with proved infection. And so I - Infected and latent people.

This work refers to the control theory a little because the model contains control parameters. It can be used to make research of pandemic change and also to try different controls or strategies.

Description of chosen SIRD model

Model is based on differential equations system, that contains each variable change in time. SIRD with testing is described by the following system:

$$S(t) + I(t) + J(t) + R(t) + D(t) = N$$

$$\frac{dS}{dt} = -Sp(I) - Su(t)$$

$$\frac{dI}{dt} = Sp(I) - \alpha I - \beta I - Iv(t)$$

$$\frac{dJ}{dt} = Iv(t) - aJ - bJ$$
$$\frac{dR}{dt} = \alpha I + aJ + Su(t)$$
$$\frac{dD}{dt} = \beta I + bJ$$

with initial state $I(0) > 0, S(0) = N - I(0)$, where
N - number of people;
S, I, J, R, D - number of Susceptible, Infected (latent), Infected (tested), Recovered and Dead people at moment t;
$p(t)dt = p(I(t))dt$ - probability to infect while time dt;
$u(t) \in [0, 1]$ - vaccinated part of people. Control variable, by default $u(t) \equiv 0$;
$v(t) \in [0, 1]$ - tested part of people at the moment t. Control variable;
α - coefficient of recovery speed for latent infected;
β - mortality coefficient for latent infected;
a - coefficient of recovery speed for tested infected;
b - mortality coefficient for tested infected.

Let r is mean density of contacts for one person, c - probability of infection in case of contact with infected. So $\frac{I}{N}$ - probability to meet infected. $\frac{rI}{N}dt$ - number of contacts with infected while time dt. $1-c$ - probability not to infect while contact with infected. $(1-c)^{\frac{rI}{N}dt}$ - probability not to infect while time dt. So probability to infect while time dt is:

$$1 - (1-c)^{\frac{rI}{N}dt} = 1 - \exp(\frac{r\log(1-c)I}{N}dt) \approx 1 - (1 + \frac{r\log(1-c)I}{N}dt) = -\frac{r\log(1-c)I}{N}dt$$

So $p(t) = p(I(t)) = -\frac{rI\log(1-c)}{N}$.

To create a model of epidemic flow, there is a need to solve differential equations system with a numeric method. To do so I used R package `deSolve`, function `ode(y, times, func, parms, ...)`. Also, got values are decimals, which is not correct to represent people, so I rounded S, I, floored R, D and got $J = N - (S + I + R + D)$.

As model parameters will be taken:

- N - number of people
- I_0 - number of infected people at moment $t = 0$, $0 < I_0 < N$
- r - mean number of contacts per day for a single person
- c - probability of infection for single contact with infected
- α - probability to recover while one day for a single latent infected
- β - probability to die while one day for a single latent infected
- a - probability to recover while one day for a single tested infected
- b - probability to die while one day for a single tested infected

Research of influence of model parameters

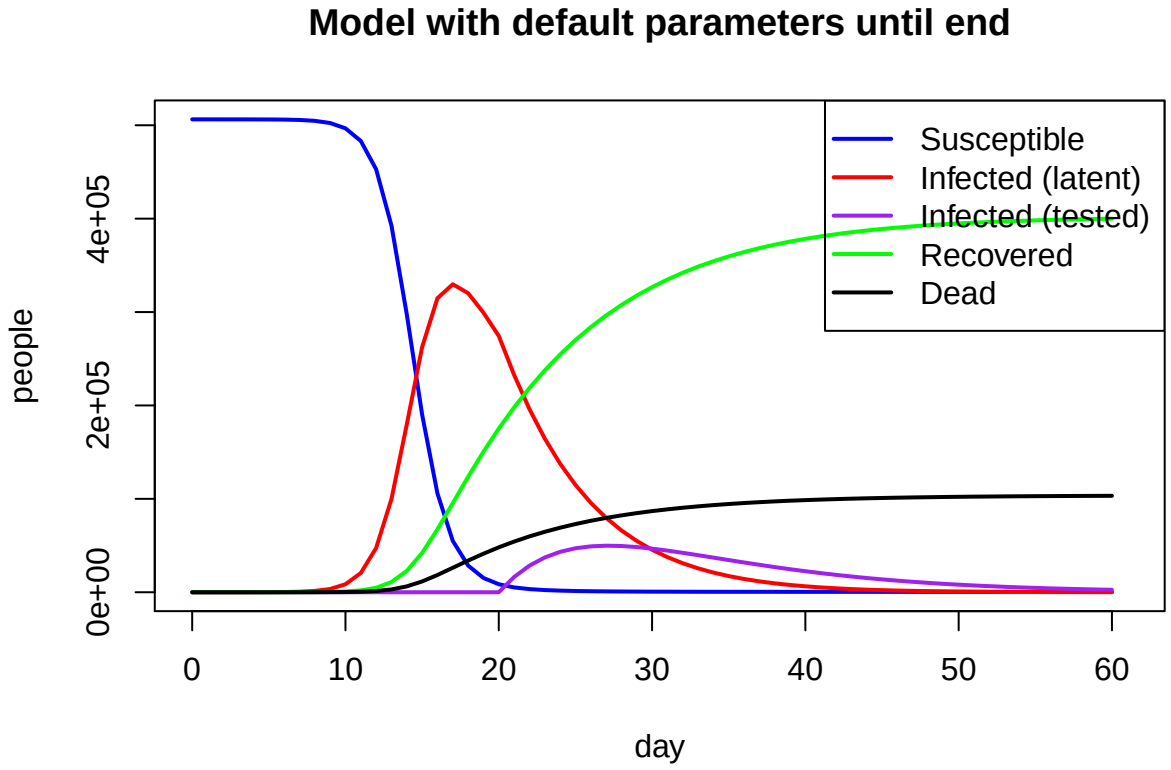
To do research I will variate each parameter with fixed others. I will set uniform random values in recommended intervals as default:
$N \in [10^3, 10^6], r \in [0.001, 50], c \in [0.4, 0.9]$
$\alpha \in [0.05, 0.1], \beta \in [0.01, 0.1]$
$a \in [\alpha, 0.1], b \in [0.01, \beta]$.
Testing part function $v(t)$ will be as it was in real experience, something like the cumulative function of normal distribution (first 20 days no tests, then testing speed increasing fast and again decreasing).

```r
set.seed(32)

input <- list(
    N = round(runif(1, 1e3, 1e6)),
    I0 = 1,
    r = 10^runif(1, log(0.001, 10), log(50,10)),
    c = runif(1, 0.4, 0.9),
    alpha = (runif(1, 0.05, 0.1) -> alpha),
    beta = (runif(1, 0.01, 0.1) -> beta),
    a = runif(1, alpha, 0.1),
    b = runif(1, 0.01, beta)
)

input <- lapply(input, round, 4)

print(input)
```

```
## $N
## [1] 506335
##
## $I0
## [1] 1
##
## $r
## [1] 0.6237
##
## $c
## [1] 0.8044
##
## $alpha
## [1] 0.0864
##
## $beta
## [1] 0.0237
##
## $a
## [1] 0.0994
##
## $b
## [1] 0.0203
```

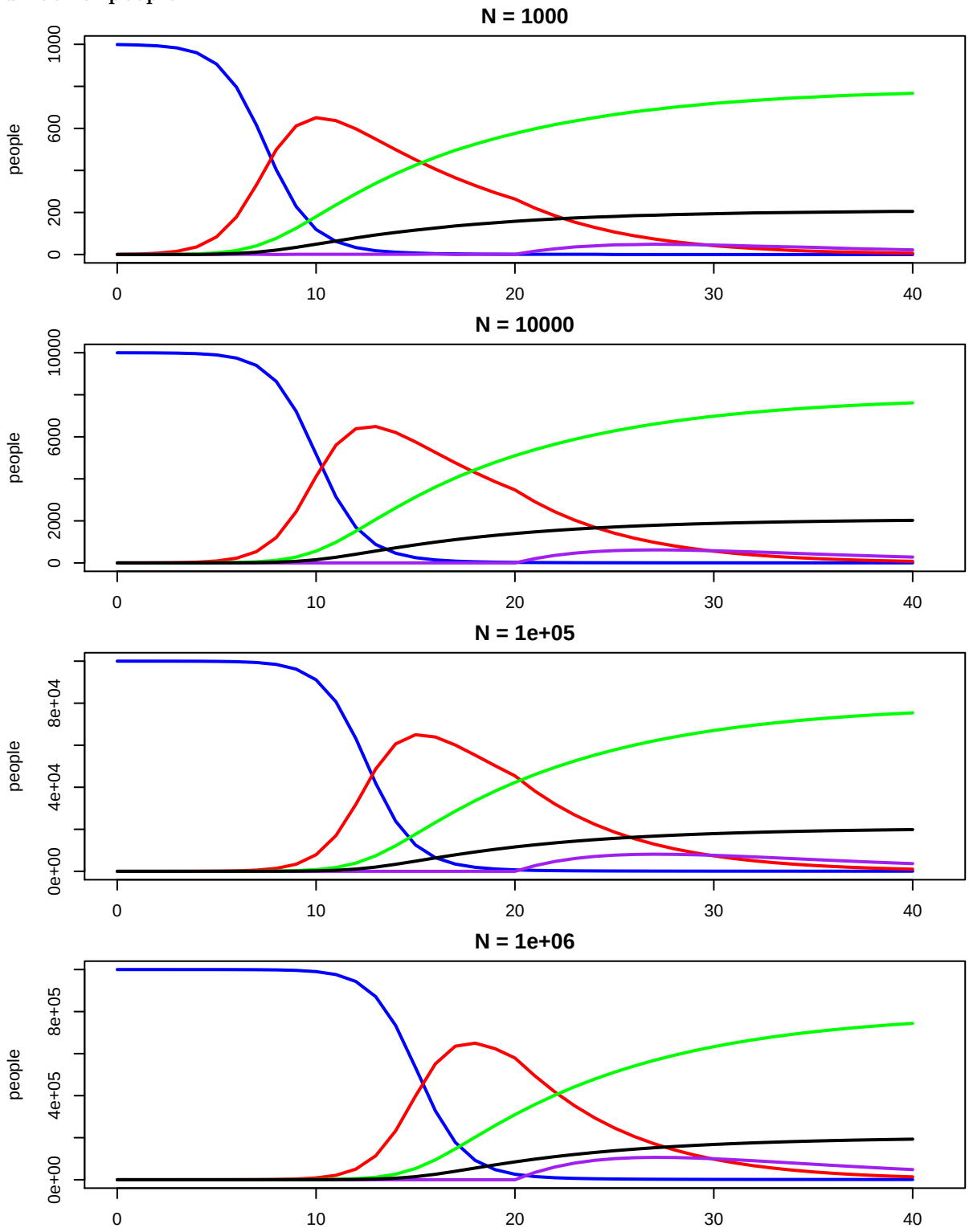
```r
input$v = function(t) ifelse(test = t>20,
                             yes = 0.7*pnorm(t, 150, 100),
                             no = 0)
```

The built model is represented on the plot:



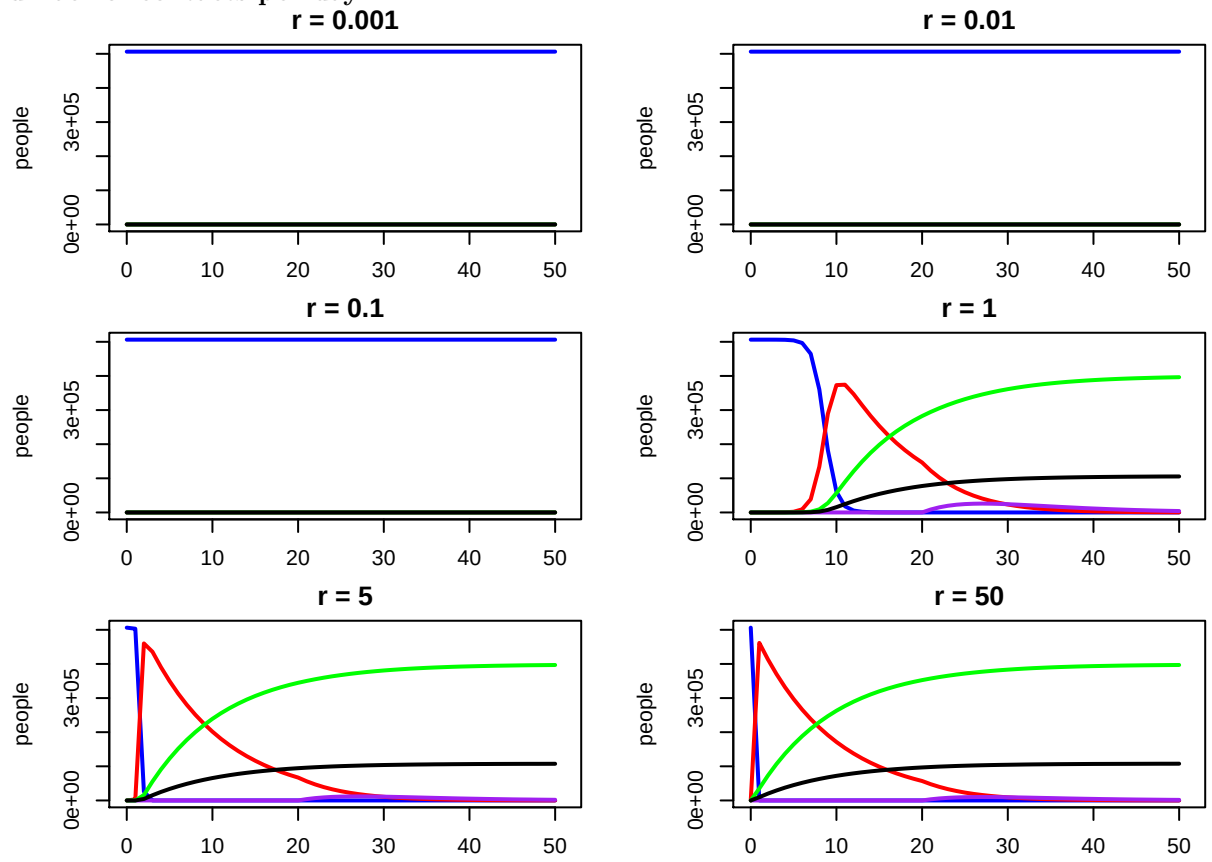
Looks controversial...

Number of people



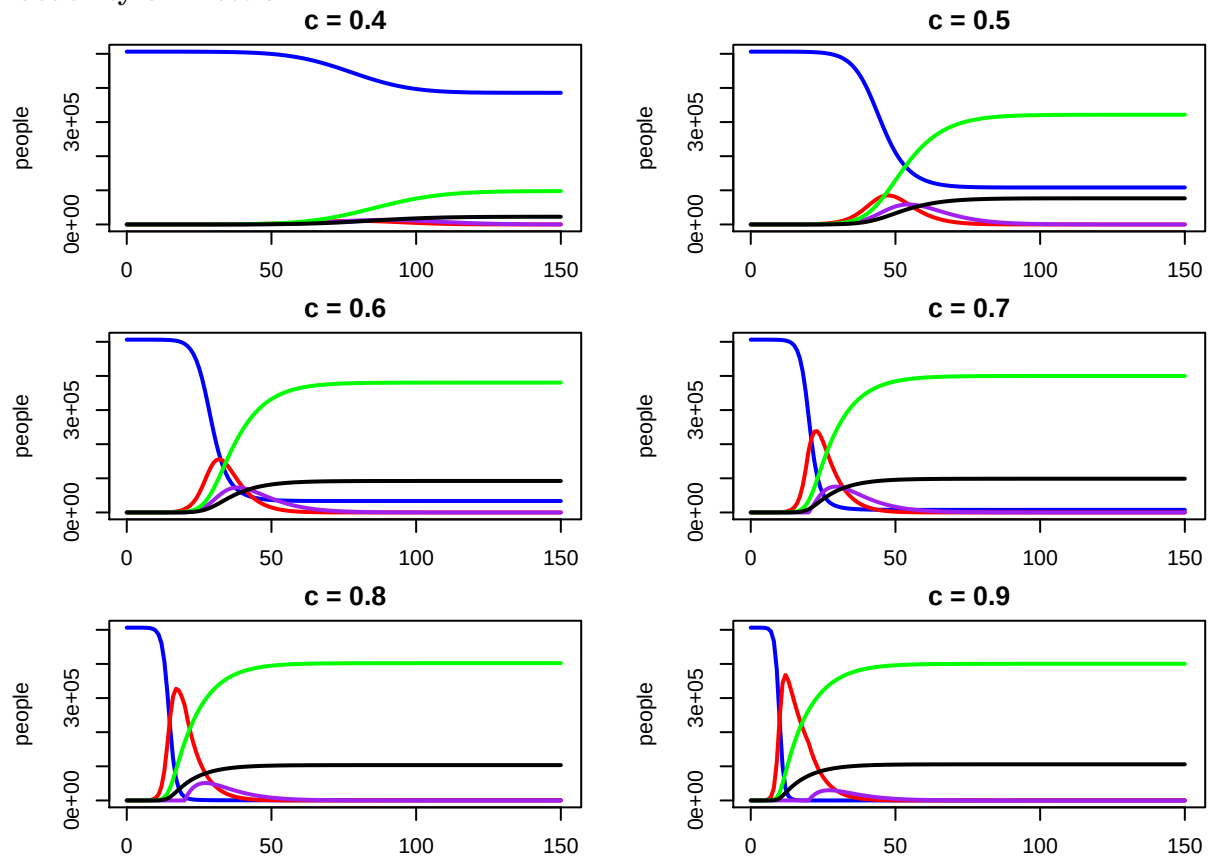
All plots are almost similar, so N affects weakly on epidemic flow. There is a small shift: I tops are at $t \approx 10$ for $N = 1e3$ and $t \approx 17$ for $N = 1e6$, so the difference is scanty considering N change.
I conclude: N **affects the speed of infection spread not significantly**.

Number of contacts per day



Those plots show weak sides of the model. First of all: we see no infection spread for $r \leq 0.1$, it might be because the single infected person died before infecting someone other. This situation is very controversial. Also, a lack of states is observed, because, for example, COVID-19 has a long incubation term, while people can infect others but are not described as infected. And also, I think intervals for parameters are given badly because the infection spreads very fast in each example and also people begin to die from the first day of infection. In my opinion, it isn't correct to say about probabilities to recover or die (α, β, a, b), when those probabilities are used in determined the model. I think the word 'probability' is not compatible with determined models. OK, also, we see that **increase of the r parameter leads to faster infection spread**. But limit distribution of Recovered and Dead almost the same, because it's very few tested infected (testing starts too late). For $r = 50$ there is the situation that almost all people are infected on the second day. **Contacts intensity highly affects on begin of infection spread**.

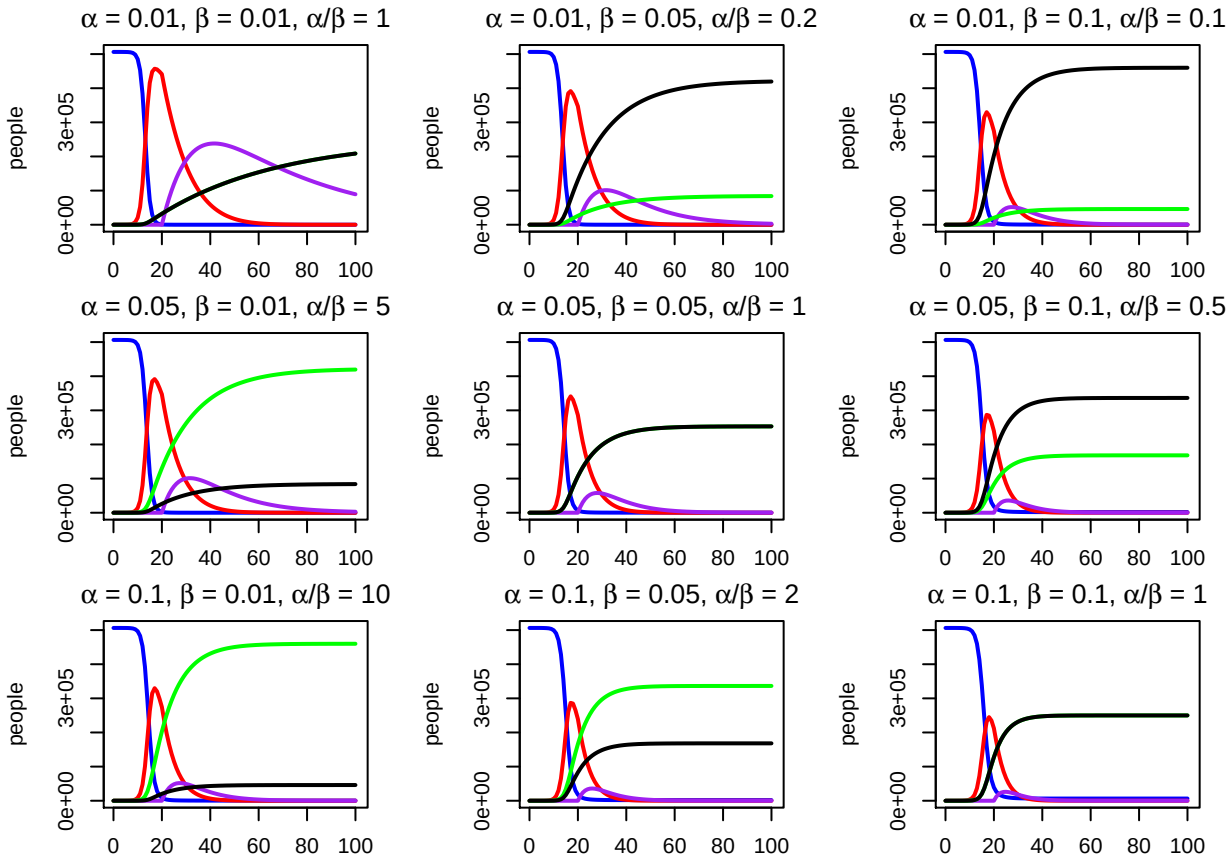
Probability of infection



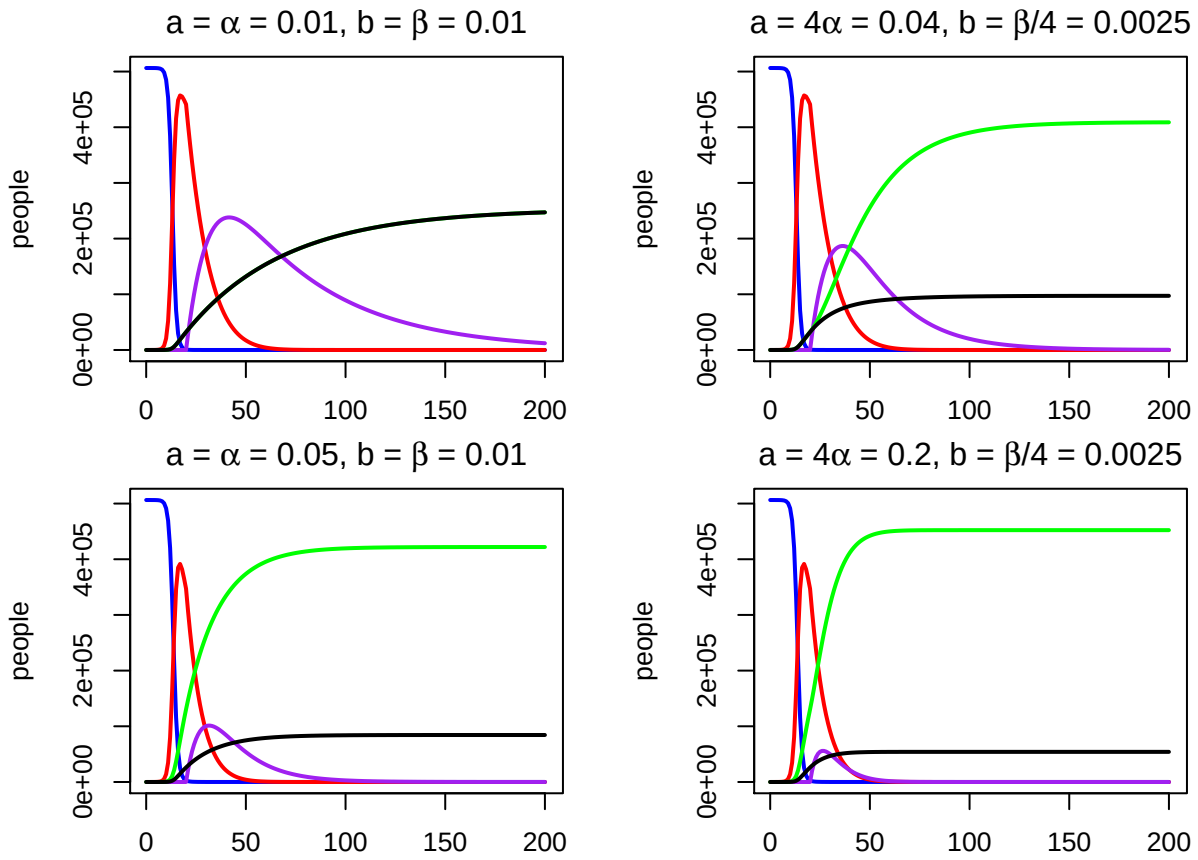
Very interesting. There is such an interesting situation: if the probability is small, then there are non-zero rest if Susceptible people. I guess it's because of testing: testing occupies latent Infected people before all Susceptible become Infected, and it can be proved: on plots with $c \geq 0.7$ the testing begins too late - when almost all Susceptibles become Infected. In other cases, testing begins when Infection spread is not such rapid and does not occupy almost all people. So according to these models, I conclude: **if testing exists and it is begun on time, it's available to stop infection spread until all people infected in case infection probability is low**. But there is an issue: it is not realistic, just demonstrative. And who knows whether it's better to recover all people, or to stop infection. Probably we can choose the strategy depending on this value. If the infection is low contagious, then it should be isolated, because it's available, like for papilloma or AIDS. But I mentioned sexually transmitted viruses. What about airborne droplets? I think there are no ways to stop it because we have to make people not breathe. For COVID-19 we have what we have, it has occupied the whole Earth.

Recovery and death probabilities

I will compare equal sets of α-s and β-s to get a more understandable picture. I will also set $a = \alpha, b = \beta$.



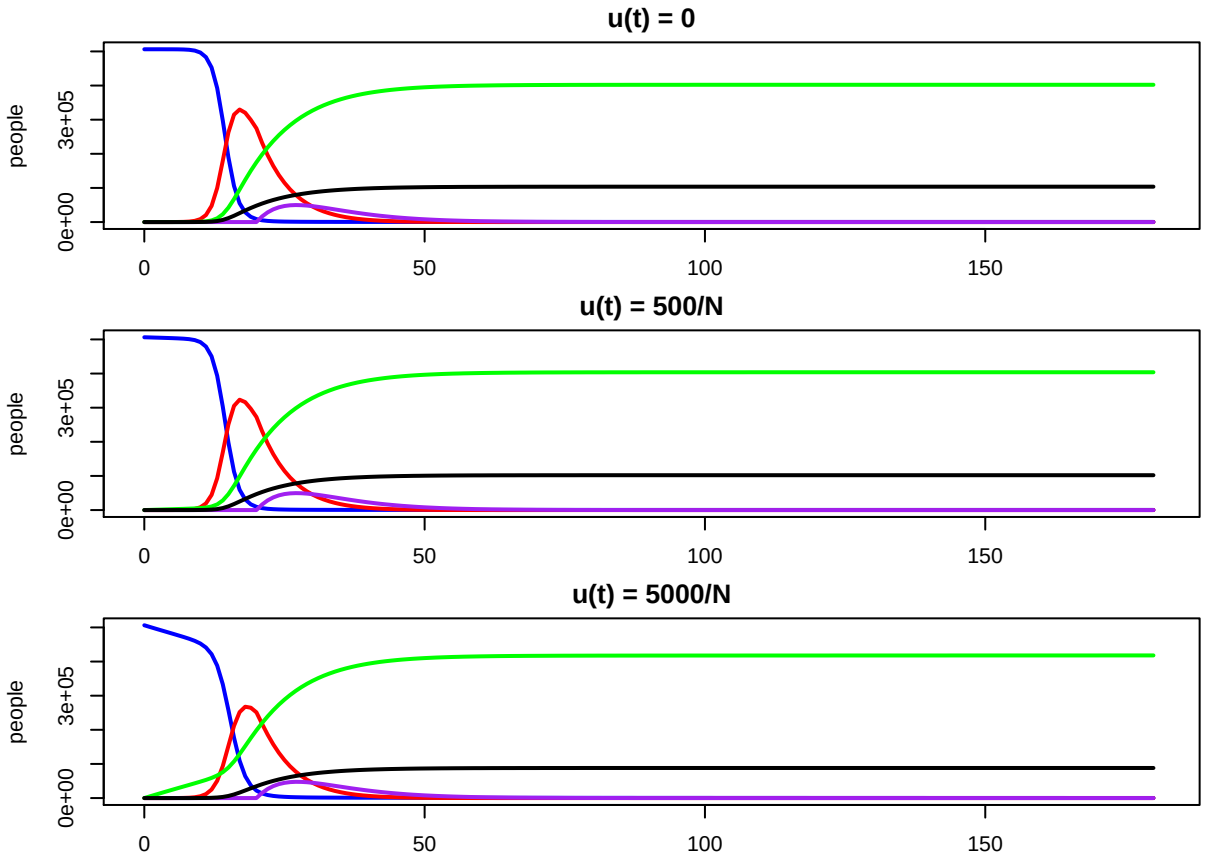
All plots look similar, but there are tiny differences. Let's look at the diagonal set of plots. We see equal distributions of Recovered and Dead but different forms of lines. And on the last plot, there is very little rest of Susceptible. Distributions of Recovered and Dead obey to $\frac{\alpha}{\beta}$ rate but, I noticed that height of Infected and hence Tested tops are different and they are less when $\alpha + \beta$ greater and, also, lines becomes straight earlier. I conclude: **if testing works nice and a, b are better than α, β, then there is a sense to provide more testing (but we have to remember that hospitals are not dimensionless)**.
To prove this conclusion I shall demonstrate this. So let's see plots, where $\alpha \in \{0.01, 0.05\}, \beta = 0.01$ and $a \in \{\alpha, 4\alpha\}, b \in \{\frac{\beta}{4}, \beta\}$,



The first pair of plots is very demonstrative about how can testing affect death volume. The second pair of plots example shows a simple idea: **Testing is more sensible as sooner it begins**.

Vaccination control research

Let try next vaccination controls: $u(t) = 500/N$ and $u(t) = 5000/N$ (500 and 5000 persons per day).




```
## u(t) = 0:

##     S      I      J     R      D
## T   396    0      1     402242 103696
## max 506334 329733 49836 402242 103696

## u(t) = 500/N:

##     S      I      J     R      D
## T   376    0      1     403814 102144
## max 506334 323313 49642 403814 102144

## u(t) = 5000/N:

##     S      I      J     R      D
## T   235    0      1     417944 88155
## max 506334 267614 47572 417944 88155
```

There is a valuable difference only in the last case, but is it available to make 5000 vaccinations per day in a half-million city? It's about 1% per day. I guess it's about maximum because it can be a lack of staff to make such many vaccinations per day. Only the last example shows notable changes in the plot. Numbers differ for each case, but not such significant for first and second cases as for second and third ones. I believe that human life is priceless and it is only the reason to vaccinate. Also, I think it makes sense to vaccinate on epidemic's very beginning to provide a maximum of Recovered people and hence a minimum of Infected to prevent many deaths.

My opinion

I think this method is good for the fast inaccurate modeling, but not sufficient to make some good realistic model, because this method gives a determined model. I think such a process like an epidemic can't be determined, it has to be stochastic. I guess it would be better to make something much complicated (even complex to formulate) but stochastic. And use the Monte-Carlo method. I tried to imagine what I would do if I would make my own model. Probably, it would be nice to keep numbers of Healthy (not infected), Infected (can be latent and tested), Recovered and Immunized (aren't equal), and Dead too. But the model has to be complicated with such parameters as:

- Incubation period mean duration
- Malady mean duration
- Quarantine properties (less contacts and probability of infection)
- Mean severity at the moment
- Severity deviation
- Social activity of people at weekdays and weekends
- Mean recovery times
- Probabilities to acquire immunity after recovery
- And others...

I imagine it like several capacities where can be people. Like in the Markov chain, each person is in some group (Healthy/Infected/Dead...) but it's hard to process each person with the Markov chain. So I propose to move random amounts of people from each state (except Dead and Immunized) like in the Markov chain but with consideration of mentioned parameters. For example, high severity brings high death or higher mean recovery times decreases the probability to infect again.