

Disease spread as a second order (chemical) reaction

This notebook has been written in *Mathematica* by

Mark J. McCready
Professor of Chemical and Biomolecular Engineering
University of Notre Dame
Notre Dame IN 46556
USA

Mark.J.McCready.1@nd.edu
www.chemeprof.com

It is copyrighted to the extent allowed by whatever laws pertain to the World Wide Web and the Internet.

I would hope that as a professional courtesy, if you use it, this notice remain visible to other users.
There is no charge for copying and dissemination.

Version: 4/7/20

More recent versions of this notebook may be available at the web site:
<http://www.chemeprof.com/>

■ The motivating question, “Why does NYC have so many more cases than anywhere else.

To pass the disease, it takes 2 people! Thus the process is 2nd order.

Therefore if we want some measure of possible disease passing interactions, we need a “concentration” of people with the disease and a concentration of people who are susceptible to the disease.

Such concentrations seem to be (obviously) people/land-area. a.k.a “population density”

Since the interaction is second order, disease transmission rates will vary as the population density squared.

I don't know why the disease modellers choose “first order” in the “number” of people. When they do so, there is an “R0” that not only varies in different locations -- for apparently unexplained reasons -- but also is not constant in time!

(How can any prediction be done in this case?)

- **Hypothesis: If the spread rate is 2nd order in population density, maybe so are the numbers of cases**

Population densities of St. Joe, Marshall, Marion counties and NYC

Below, the disease spread is modelled as a 2nd order “chemical” reaction with $A + B \rightarrow 2 B$. The concentrations are people/land area

$A \rightarrow$ UNINFECTED PEOPLE
 $B \rightarrow$ INFECTED PEOPLE
 DISEASE TRANSMISSION IS 2ND ORDER
 $r = \beta C_A C_B$
 $A + B \rightarrow 2 B$
 $\therefore \frac{dA}{dt} = -\beta C_A C_B$
 $\frac{dB}{dt} = -\beta C_A C_B + 2\beta C_A C_B$
 $\frac{dB}{dt} = \beta C_A C_B$

COUNTY	POPULATION DENSITY	(POPULATION DENSITY) ²	COVID 19 CASES
ST. JOE	585 P/m ²	3.4×10^5	49
MARSHALL	104	1.1×10^4	3
MARION	2369	5.6×10^6	964
NYC	26,400	7×10^8	41,000

Cases and population density squared

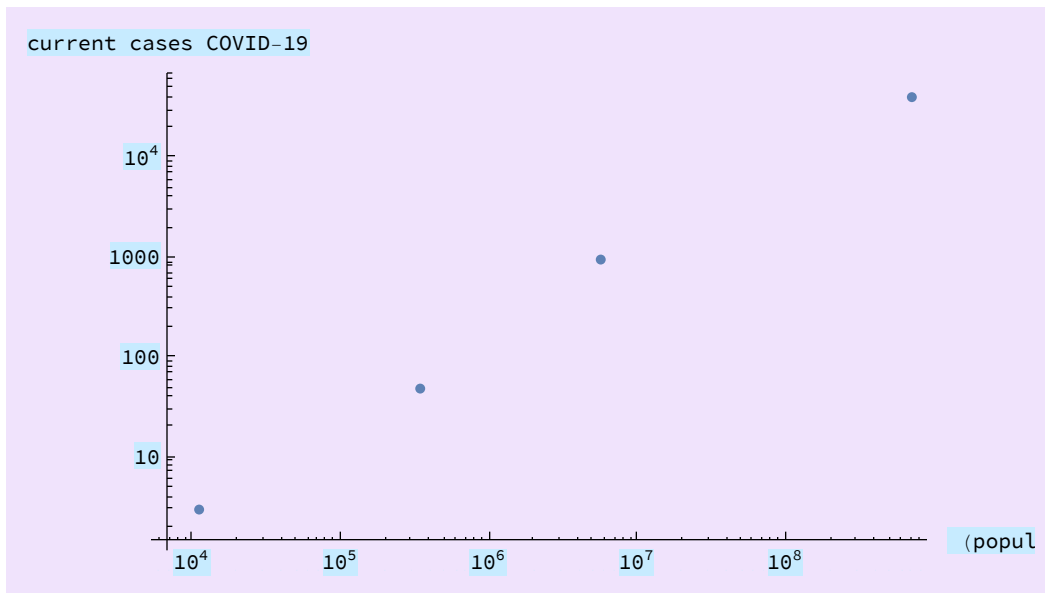
On March 31 for Marshall, St. Joseph, Marion counties in Indiana and New York city. Data for population density² in (people/mile²)² and numbers of reported COVID19 cases.

```
data = {{1.1 × 104, 3},
        {3.4 × 105, 49}, {5.6 × 106, 964}, {7 × 108, 41000}}
```

```
{{11000., 3}, {340000., 49}, {5.6 × 106, 964}, {700000000, 41000}}
```

Maybe just luck... but look at the correlation!

```
ListLogLogPlot[data,
  AxesLabel → {"(population/mi2)2", "current cases COVID-19"}]
```



equations for disease spread

1 person to 1 person, constant interaction rate, just a pair of equations for a 2nd order reaction, A + B → 2 B

$$\begin{aligned} D[ca[t], t] &= -\beta ca[t] \times cb[t], \\ D[cb[t], t] &= +\beta ca[t] \times cb[t], \quad ca[0] = ca0, \quad cb[0] = cb0 \end{aligned}$$

$$\begin{aligned} \{ca'[t] &= -\beta ca[t] \times cb[t], \\ cb'[t] &= \beta ca[t] \times cb[t], \quad ca[0] = ca0, \quad cb[0] = cb0\} \end{aligned}$$

An interaction coefficient. Could be different in different locations, but not likely to change over time naturally -- unlike R0!

$\beta = .01$

0.01



Some numbers for our St. Joseph county

$ca_0 = 585;$

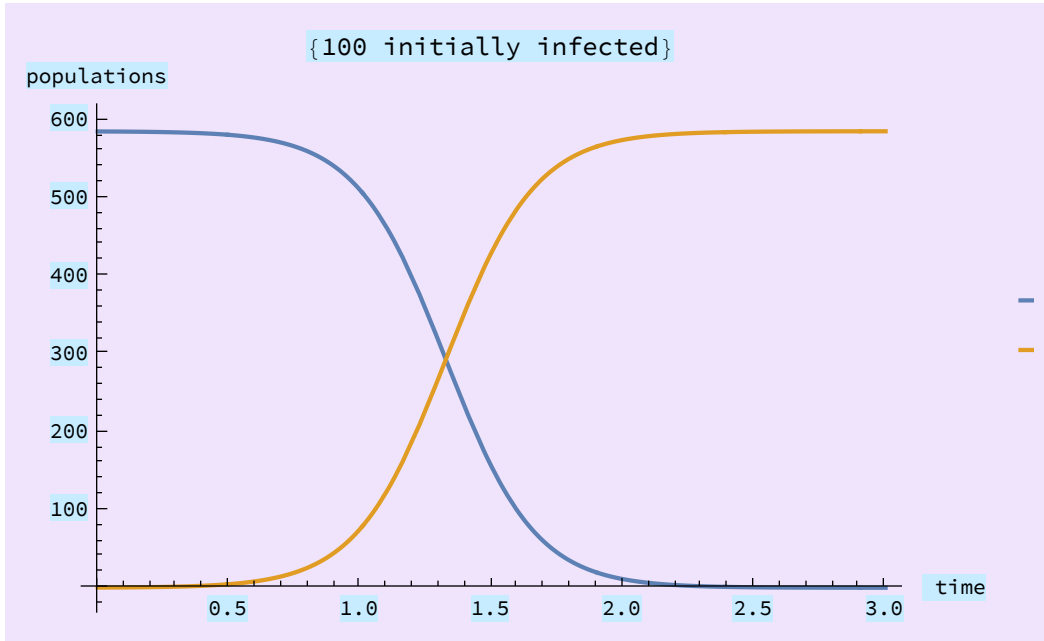
Supposed 100 initial infections some time in the past

$cb_0 = 100 / 400;$

```
ans = NDSolve[
  { D[ca[t], t] == -  $\beta$  ca[t]  $\times$  cb[t], D[cb[t], t] == +  $\beta$  ca[t]  $\times$  cb[t],
    ca[0] == ca0, cb[0] == cb0}, {ca[t], cb[t]}, {t, 10}]
```

```
{{ca[t] → InterpolatingFunction[ Domain: {{0., 10.}} Output: scalar][t],
  cb[t] → InterpolatingFunction[ Domain: {{0., 10.}} Output: scalar][t]}}
```

```
Plot[{ca[t] /. ans, cb[t] /. ans}, {t, 0, 3},
PlotLegends -> {"uninfected", "infected"},
AxesLabel -> {"time", "populations"},
PlotLabel -> {"100 initially infected"}]
```



```
ca0 = 585;
```

only 10 initial cases

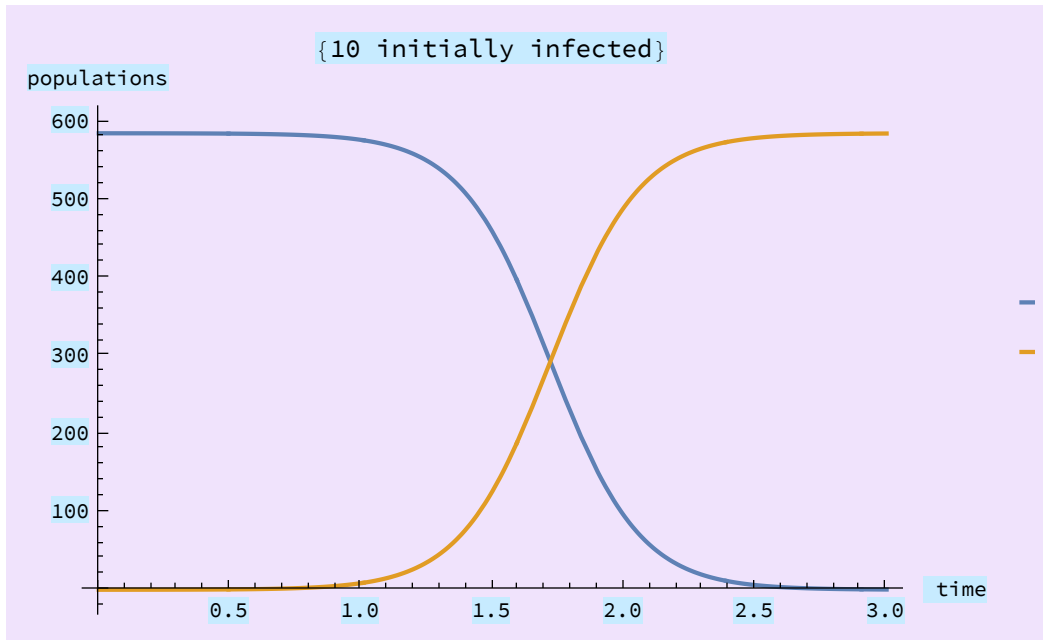
```
cb0 = 10 / 400;
```

```
ans = NDSolve[
  {D[ca[t], t] == -β ca[t] × cb[t], D[cb[t], t] == +β ca[t] × cb[t],
  ca[0] == ca0, cb[0] == cb0}, {ca[t], cb[t]}, {t, 10}]
```

$\{ \{ ca[t] \rightarrow \text{InterpolatingFunction} [\text{Domain: } \{ \{ 0., 10. \} \}] [t],$
 $cb[t] \rightarrow \text{InterpolatingFunction} [\text{Domain: } \{ \{ 0., 10. \} \}] [t] \} \}$

So it takes longer to get to the inflection point. This is an arbitrary criterion, but allows comparison between different initial conditions and population densities.

```
Plot[{ca[t] /. ans, cb[t] /. ans}, {t, 0, 3},
PlotLegends -> {"uninfected", "infected"},
AxesLabel -> {"time", "populations"},
PlotLabel -> {"10 initially infected"}]
```



```
ca0 = 585;
```

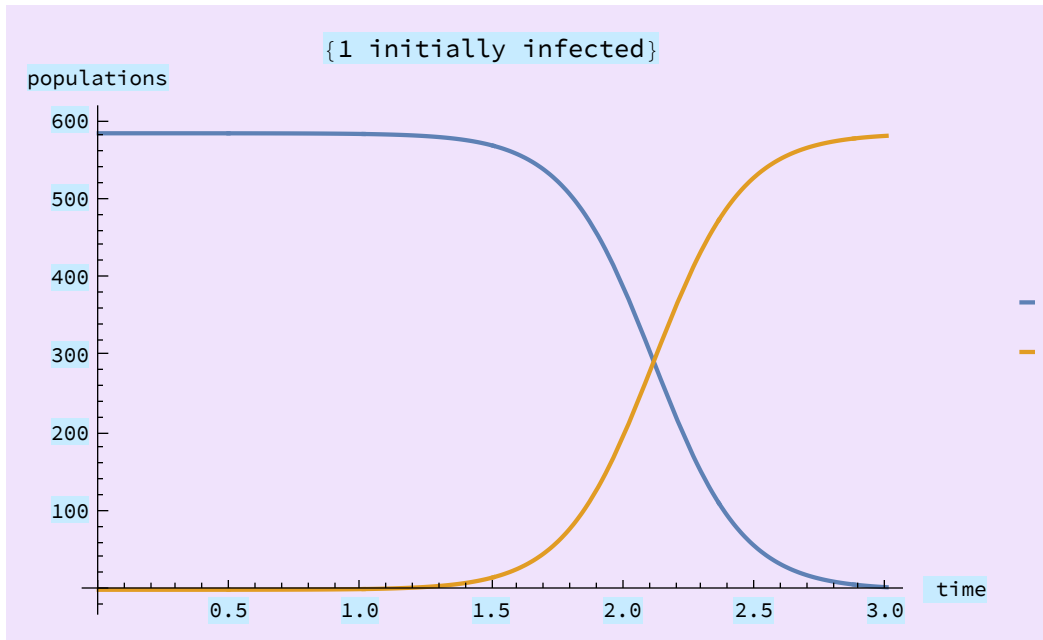
only 1 initial case

```
cb0 = 1 / 400;
```

```
ans = NDSolve[
  {D[ca[t], t] == -β ca[t] × cb[t], D[cb[t], t] == +β ca[t] × cb[t],
   ca[0] == ca0, cb[0] == cb0}, {ca[t], cb[t]}, {t, 10}]
```

```
{ {ca[t] -> InterpolatingFunction[ + [Domain: {{0., 10.}}
Output: scalar] [t],
cb[t] -> InterpolatingFunction[ + [Domain: {{0., 10.}}
Output: scalar] [t]] }
```

```
Plot[{ca[t] /. ans, cb[t] /. ans}, {t, 0, 3},
PlotLegends -> {"uninfected", "infected"},
AxesLabel -> {"time", "populations"},
PlotLabel -> {"1 initially infected"}]
```



So if initial concentration is $1/400$ time = 2.1, if the initial concentration was $100/400$ (that is 100 cases in the country), the time would have been 1.3. Even with this large uncertainty in initial cases, the time period until maximum “growth” would vary by less than a factor of 2.

So the growth in infected people will be exponential. So while this matches the standard expectation of an epidemic model, exponential growth, the process is really a “pseudo” first order reaction.

A value for the parameter β is not known, presumably it is determined by the mechanism of transmission, but there is no obvious reason for it to change in time... as opposed to R_0 which has to drop to stop the exponential growth!

■ NYC

The concentration of people in New York City.

ca0 = 26 400;

What if same initial concentration of cases in NYC as in St. Joseph country. Using our middle value above, $10/400$.

```

cb0 = 10 / 400;
ans = NDSolve[
  { D[ca[t], t] == - β ca[t] × cb[t], D[cb[t], t] == + β ca[t] × cb[t],
    ca[0] == ca0, cb[0] == cb0}, {ca[t], cb[t]}, {t, 10}]

```

```

{{ca[t] → InterpolatingFunction[ + [t],
  cb[t] → InterpolatingFunction[ + [t]]}

```

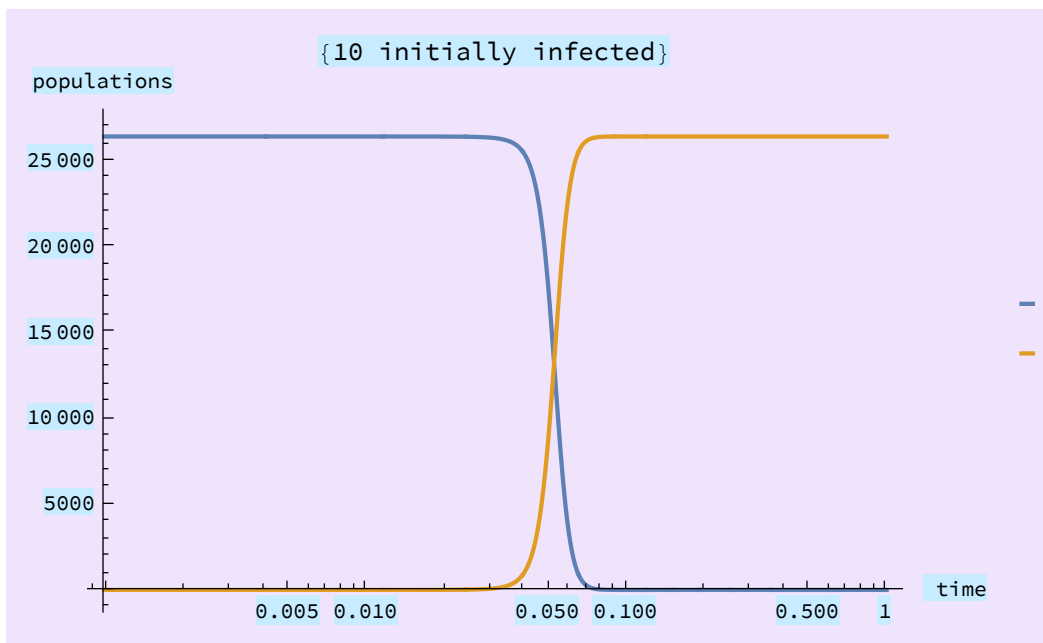
Domain: {{0., 10.}}
Output: scalar

Domain: {{0., 10.}}
Output: scalar

```

LogLinearPlot[{ca[t] /. ans, cb[t] /. ans},
  {t, 0, 1}, PlotLegends → {"uninfected", "infected"},
  AxesLabel → {" time", "populations"},
  PlotLabel → {"10 initially infected"}]

```



We see that the time to inflection is ~ 0.052 , compared to 1.75 for St. Joseph County Indiana. This is about $1/30$ the time interval for St. Joseph county.

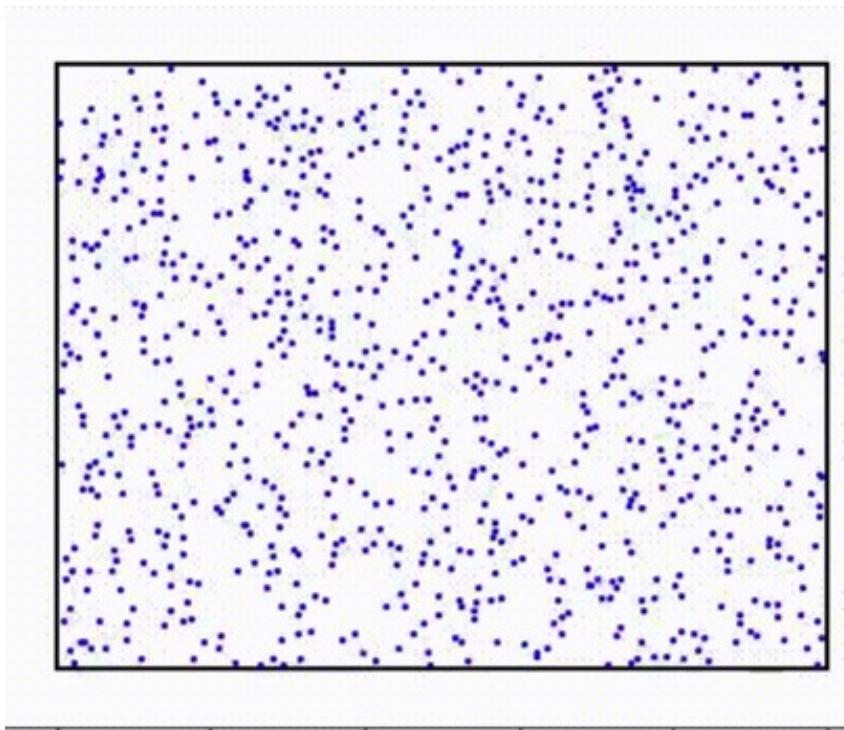
So it is not correct to say that one region is 2 weeks behind another!
What ever is happening to NYC is 30 times faster than us... At least!

Something that was a couple of weeks there would *never* happen here!

■ Differences in local geography.

We might expect that the local *fluctuations* in population density are different depending on location.

Sure, maybe. But I paste in a snapshot of an ideal gas simulation showing that the atoms “appear” to be in many clusters even though there are no attractive forces. Similarly, in a bar in NYC or in a small restaurant in Plymouth Indiana, there are relative clusters of people. Thus most of the transmission in either city would be in closely-packed, probably indoor, locations.



Suppose that people get better

$$\beta = .01$$

$$0.01$$

Define k as a “rate” coefficient for people getting over the virus and presumed to not be susceptible to immediate re-infection.

$$k = 1;$$

Some numbers for our county

```
ca0 = 585;
```

10 initial infections

```
cb0 = 10 / 400;
```

```
ansk2 = NDSolve[{ D[ca[t], t] == -  $\beta$  ca[t]  $\times$  cb[t],  
  D[cb[t], t] == +  $\beta$  ca[t]  $\times$  cb[t] - k cb[t], ca[0] == ca0,  
  cb[0] == cb0}, {ca[t], cb[t]}, {t, 10}]
```

```
anskp1 = NDSolve[{ D[ca[t], t] == -  $\beta$  ca[t]  $\times$  cb[t],  
  D[cb[t], t] == +  $\beta$  ca[t]  $\times$  cb[t] - k cb[t], ca[0] == ca0,  
  cb[0] == cb0}, {ca[t], cb[t]}, {t, 10}]
```

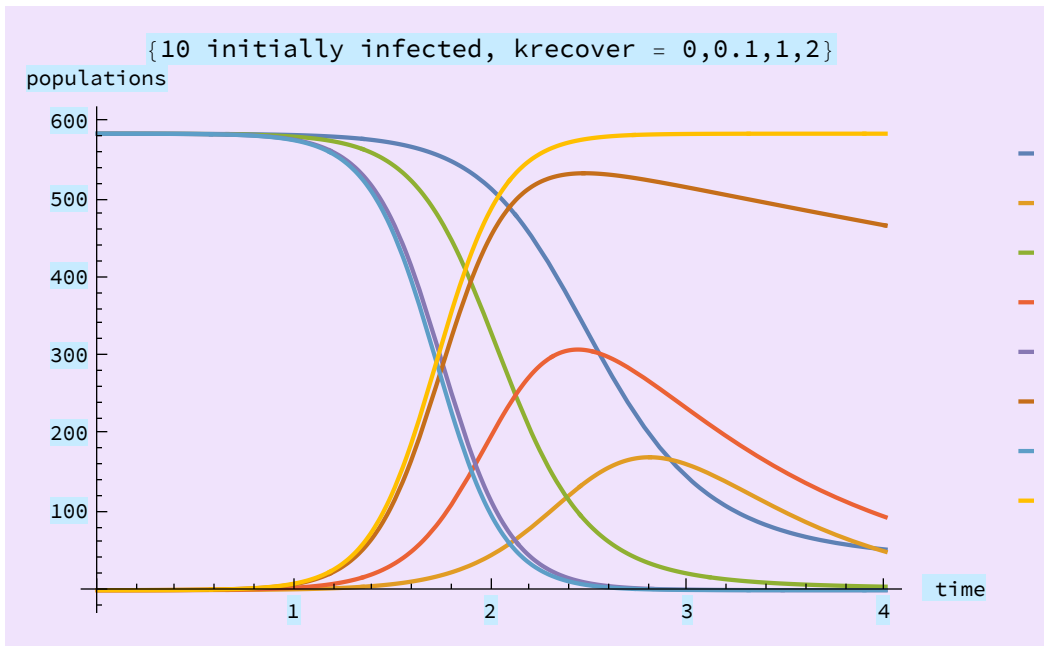
```
ansk1 = NDSolve[{ D[ca[t], t] == -  $\beta$  ca[t]  $\times$  cb[t],  
  D[cb[t], t] == +  $\beta$  ca[t]  $\times$  cb[t] - k cb[t], ca[0] == ca0,  
  cb[0] == cb0}, {ca[t], cb[t]}, {t, 10}]
```

```
ansk0 = NDSolve[{ D[ca[t], t] == -  $\beta$  ca[t]  $\times$  cb[t],  
  D[cb[t], t] == +  $\beta$  ca[t]  $\times$  cb[t] - k cb[t], ca[0] == ca0,  
  cb[0] == cb0}, {ca[t], cb[t]}, {t, 10}]
```

```

Plot[{ca[t] /. ans2, cb[t] /. ans2, ca[t] /. ans1, cb[t] /. ans1,
      ca[t] /. ansk1, cb[t] /. ansk1, ca[t] /. ansk0, cb[t] /. ansk0},
     {t, 0, 4}, PlotLegends → {"uninfected, k=2", "infected, k=2",
                               "uninfected, k=1", "infected, k=1", "uninfected, k=0.1",
                               "infected, k=1", "uninfected, k=0", "infected, k=0"},
     AxesLabel → {" time", "populations"},
     PlotLabel → {"10 initially infected, krecover = 0,0.1,1,2"}]

```



We see that if some people recover, the cross-over point moves to later in time. The total number infected at any time is also reduced -- but this is obvious.

Conclusions

1. The 2nd order model in population density gives a plausible reason for the very high numbers of infected people in NYC compared to elsewhere.
2. If this model is correct, regions that are much less densely populated have little worry about in terms of overwhelming any medical facilities. “Something” will happen to interfere with ever reaching a peak if the time to get there is ~year, not 1 few weeks!