Methods in Transportation Econometrics and Statistics (Master)

Winter semester 2020/21, Solutions to Tutorial No. 1a

Solution to Problem 1a.1: SI model

(a) Just insert the proposed solution into the differential equation

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta I(1-I): \tag{1}$$

– left-hand side with the ratio rule $(u/v)' = (u'v - v'u)/v^2$ for differentiating a ratio u/v:

$$\begin{aligned}\text{lhs} = \frac{\mathrm{d}I}{\mathrm{d}t} &= \frac{\beta e^{\beta t}\left(1+e^{\beta t}\right) - e^{2\beta t}}{\left(1+e^{\beta t}\right)^2} \\ &= \frac{\beta e^{\beta t}}{\left(1+e^{\beta t}\right)^2}\end{aligned}$$

– Inserting the ansatz into the right-hand side:

$$\begin{aligned}\text{rhs} = \beta I(1-I) = \beta \frac{e^{\beta t}}{1+e^{\beta t}}\left(1 - \frac{e^{\beta t}}{1+e^{\beta t}}\right) &= \beta \frac{e^{\beta t}\left(1+e^{\beta t} - e^{\beta t}\right)}{\left(1+e^{\beta t}\right)^2} \\ &= \frac{\beta e^{\beta t}}{\left(1+e^{\beta t}\right)^2}\end{aligned}$$

so we have lhs=rhs, so the equation is valid and the proposed function is, in fact, a solution to the SI model (1).

– Initial condition

$$I(0) = \frac{e^0}{1+e^0} = 1/2$$

(b) With $I \ll 1$ and $1 - I \approx 1$, the SI model (1) becomes

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta I(1-I) \approx \beta I$$

with the class of solutions (just insert!) $I(t) = I_0 e^{\beta t}$.

Solution to Problem 1a.2: SIR model

(a) Just take the balance of all three equations resulting in

$$\frac{\mathrm{d}S}{\mathrm{d}t} + \frac{\mathrm{d}I}{\mathrm{d}t} + \frac{\mathrm{d}R}{\mathrm{d}t} = 0 \quad \Rightarrow \quad S + I + R = \text{const.}$$

You can set the constant to any value but only if you set it =1, the quantities S, I, and R have the meaning of fractions adding up to 1.

(b) The first two equations are *linked* in both directions that means the S equation depends on I and the I equation on S, so they need to be treated simultaneously. However, the third equation is just unidirectionally *chained* to the I equation, i.e., there is no backlink and the dynamics of S and I does not depend on the "dropouts" R. So, R can be calculated afterwards once $S(t)$ and $I(t)$ is known

(c) At the beginning of the infection spread, we have $I \ll 1$, $S = 1 - I \approx 1$ and $R = 0$. Then, the middle equation for the dynamics of $I(t)$ becomes uncoupled by virtue of $S \approx 1$ and can be considered on its own:

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta I 1 - \gamma I = (\beta - \gamma) I$$

This is the model for unlimited growth with the growth rate $\beta - \gamma$

(d) This can be shown by inserting $I = 0$ and setting all time derivatives to zero in the model (**??**). We obtain three trivial identities 0=0 showing that any combination $I = 0$, and $S = \text{const.}$, $R = \text{const.}$ is a valid solution. Since, in order to obtain the meaning of fractions for the dynamic variables we set $S + I + R = 1$, we have $S(t) = S_0 = \text{const.}$, $I = 0$ and $R(t) = R_0 = 1 - S_0$. Since there is no time dependency, this is a *steady-state* solution.

(e) Here, again, the SIR equation for $\frac{\mathrm{d}I}{\mathrm{d}t}$ gives the right information. Inserting the ansatz $S(t) = S_0 - \tilde{S}(t)$, $I(t) = \tilde{I}(t)$, and $R(t) = R_0 + \tilde{R}(t)$ into this equation and neglecting the product $\tilde{I}\tilde{S}$ since it is much smaller than the other terms, we get

$$\frac{\mathrm{d}\tilde{I}}{\mathrm{d}t} = \beta S_0 \tilde{I} - \gamma \tilde{I}.$$

This is again a model/equation for unlimited growth with the solution

$$\tilde{I}(t) = \tilde{I}(0) e^{(\beta S_0 - \gamma)t}$$

So,

- for the growthrate $\lambda = \beta S_0 - \gamma < 0$, i.e., sufficiently low fraction of susceptible persons, the new infection will die out,
- for $\lambda = \beta S_0 - \gamma < 0$ it will trigger a new infection wave

With $R_0 = \beta/\gamma$ and the fact that β relates to the dynamics of new infections while γ is just a constant for a given virus, we eliminate $\beta = R_0\gamma$ in favour of the "true" constant γ. Hence, we obtain following criterion for the onset of a new wave:

$$\lambda = (R_0 S_0 - 1)\gamma > 0$$

Notice that, for $R_0 < 1$, this condition is never satisfied which is consistent: In this case, every infected person infects less than one other person, even if everybody is susceptible.

Solution to Problem 1a.3: Tests: sensitivity and specificity

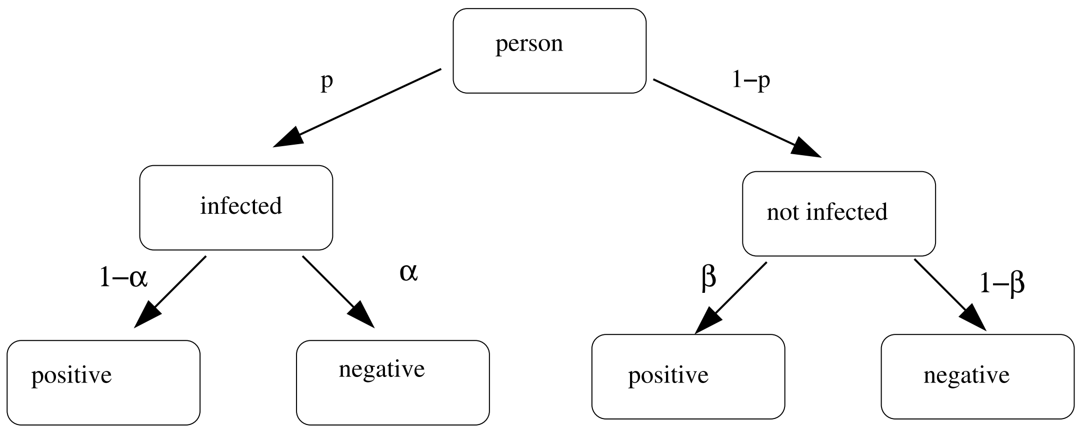
(a) Since $1 - \alpha$ is the conditional probability that an infected person is tested positive, α is the detection error that an infected person is tested negative. Under the null hypothesis

$$H_0: \quad \text{the person is infected}$$

α is the error of the first kind $P(H_0 \text{ rejected}|H_0)$

The error β is the *false positive* rate or, under the above null hypothesis, an error of the second kind "H_0 is not true but not rejected"

(b) Tree diagram:



(c) The expected fraction of positive tests is given by the two "positive" exits of the tree diagram:

$$\begin{aligned} P(\text{pos}) &= P(\text{infected})P(\text{pos}|\text{infected}) + P(\text{not infected})P(\text{pos}|\text{not infected}) \\ &= p(1-\alpha) + (1-p)\beta \\ &= 1.97\,\% \end{aligned}$$

The conditional probability $P(\text{infected}|\text{pos})$ can be calculated using Bayes's theorem or considering the two trajectories through the tree with "positive" exits:

$$P(\text{infected}|\text{pos}) = \frac{P(\text{pos}|\text{infected})P(\text{infected})}{P(\text{pos})} = \frac{p(1-\alpha)}{p(1-\alpha)+(1-p)\beta} = 49.7\,\%$$

The conditional probability $P(\text{not infected}|\text{negative})$ is determined similarly by analysing the trajectories with "negative" exits:

$$P(\text{not infected}|\text{neg}) = \frac{(1-p)(1-\beta)}{(1-p)(1-\beta)+p\alpha} = 99.98\,\%$$

Although only one percent is infected and the test has a false-positive error of only 1 %, nearly 50 % of all persons actually tested positive are false positives! This is due to the low population incidence $p = 1\,\%$. This low incidence also leads to the result that, after a negative test, you know pretty certainly (0.02 % error) that you are not infected.

Notice that this is *not* true for persons with symptoms because, then, the a-priori rate is much higher than the population incidence $p = 1\,\%$