

Tameez Latib

Problem 2

I, Tameez Latib, declare that this work is my own. I did this work honestly and can fully stand behind everything that I have written.

—
We will model the spread of a popular video/image/meme/etc. To do this, let us define some variables. Let there be 3 groups of people

The Infectious group, I , the people are sharing the cat video

The Exposed group, E , the people who have seen the cat video but are not sharing it

The Susceptible group, S , the people who have not seen the video

Let n be the average amount of people who see the video per link/share of the video. (Always positive)

Let p be the probability that any person who sees the video will be infected, and will share the video again. (Always positive)

To make the problem easier, let's assume a few things.

Suppose the image/video is shared on a website like imgur or youtube, so that it can reach any possible person (I.e., all people are 'connected' to each other, and any person can spread the video to any other person). Furthermore, due to the extremely large amount of people who use the internet, let's assume S is infinite- there is always a large number of people who have not seen the image/video. We are also going to assume that an infected person ONLY shares the video once. In other words, after an infected person shares a video, he/she leaves the Infectious group and enters the Exposed group. We will also assume that probabilities are certainties (for example, if the chance of landing heads in a coin flip is one half, then flipping a coin twice will yield one heads and one tails)

Let's model this in time-steps, t . We know n people see the video per Infectious person, and that p of those people become infectious. So np is the number of people who become infected by one infectious person, and likewise, $n(1 - p)$ is the number of people who become exposed by one infectious person. After one time step, the amount of infected people should be equal to the the number of people who are infectious times the number of people the average infectious person infects. In math terms, $I_t = I_{t-1} * np$. Note that we don't add the number of previously infected people to the number of total infected people at time t , as the previously infected people become part of the Exposed group. Furthermore, let I_0 be the amount of initially infected people. The amount of Exposed people is equal to the amount of people exposed by infectious individuals (who do not become infectious) plus the amount of previously exposed people plus the amount of people who are infectious that become exposed. Or in math terms, $E_t = I_{t-1} * n(1 - p) + E_{t-1} + I_{t-1}$ Furthermore, let E_0 be the amount of people who first saw the video who did not become exposed.

First, let's try to model this discretely. Expanding out I_t gives

$$I_t = np * I_{t-1} = np * np * I_{t-2} = \dots = np^k * I_{t-k} = np^t * I_0$$

E_t can also be simplified by making a physical observation; the number of exposed people is equal to the amount of people who have previously been

infected and exposed 2 time step or more earlier plus the amount of people who just became exposed plus the amount of initially exposed people plus the amount of initially infected people. We also note that the $I_t * n$ is the number of people that see the video. Furthermore, the number of people infected and exposed together is the number of people who see a video. Using this, we can simplify

$$\begin{aligned} E_t &= (\sum_{i=0}^{t-2} nI_i) + I_{t-1} * n(1-p) + E_0 + I_0 \text{ For } t > 1, \text{ with} \\ E_0 &= E_0 \\ E_1 &= E_0 + I_0 + I_0 * n(1-p) \end{aligned}$$

Where the summation is physically interpreted as the people that view the video except for the people who the currently infected people show the video to. The next term is the people who the currently infected people expose (but do not infect). We can also check this by expanding out our previous equation for E_t . We can also use our formula for I_t to get

$$E_t = (\sum_{i=0}^{t-2} n(np)^i * I_0) + I_{t-1} * n(1-p) + E_0 + I_0$$

Using some summation formulas,

$$E_t = (nI_0 * \sum_{i=0}^{t-2} (np)^i) + I_{t-1} * n(1-p) + E_0 + I_0$$

$$E_t = (nI_0 * \frac{1-(np)^{t-1}}{1-np}) + I_{t-1} * n(1-p) + E_0 + I_0$$

$$E_t = (nI_0 * \frac{1-(np)^{t-1}}{1-np}) + I_0 * (np)^{t-1} * n(1-p) + E_0 + I_0$$

From our equation for I_t , we see that it

if $np > 1$, $I_{t+1} > I_t$ For all t

if $np < 1$, $I_{t+1} < I_t$ For all t

if $np = 1$, $I_{t+1} = I_t$ For all t

Thus, we see that if $np < 1$, I_t goes to 0, and eventually E_t will approach an asymptote (physically, if there are no people to infect other people ($I = 0$), nobody will be exposed)

For $np = 1$, (or respectively $np > 1$) I_t remains the same (or respectively I_t increases), so E_t will continue to increase without bound.

Now, let's model this as a continuous system. To do this, let

$$I(t) = I_t$$

$$E(t) = E_t$$

Using our original equation for $I(t)$,

$$I(t+1) = I(t) * np$$

$$I(t+1) - I(t) = I(t) * np - I(t)$$

Let's utilize the approximation: $f(t + \Delta t) = f(t) + \Delta t * f'(t) + O$ Where O is the higher order terms. for this approximation, let $O = 0$, and $\Delta t = 1$. Then:

$$I'(t) = I(t) * np - I(t)$$

$$I'(t) = I(t)(np - 1)$$

The resulting ODE yields

$$I(t) = I_0 e^{(np-1)*t}$$

$E(t)$ can be found by using the above. Similar to the last result, we get

$$E(0) = E_0$$

$$\begin{aligned}
E(1) &= E_0 + I_0 + I_0 * n(1-p) \\
E_t &= (\sum_{i=0}^{t-2} nI(i)) + I(t-1) * n(1-p) + E_0 + I_0 \\
E_t &= nI_0(\sum_{i=0}^{t-2} e^{(np-1)*i}) + I(t-1) * n(1-p) + E_0 + I_0 \\
E_t &= nI_0(\sum_{i=0}^{t-2} e^{(np-1)^i}) + I_0 e^{(np-1)*(t-1)} * n(1-p) + E_0 + I_0 \\
E_t &= nI_0(\frac{1 - e^{(np-1)^{t-1}}}{1 - e^{(np-1)}}) + I_0 e^{(np-1)*(t-1)} * n(1-p) + E_0 + I_0
\end{aligned}$$

From our result

$$I(t) = I_0 e^{(np-1)*t}$$

We see that a similar result follows with the same reasoning. The models aren't too different, in fact $e^{(np-1)} \approx np$ by its taylor series, up to its 1st order term ($e^x = 1 + x + \dots \approx 1 + x$)

if $np > 1$, $I(t+1) > I(t)$ For all t

if $np < 1$, $I(t+1) < I(t)$ For all t

if $np = 1$, $I(t+1) = I(t)$ For all t

We see that if $np < 1$, $I(t)$ goes to 0, and eventually $E(t)$ will approach an asymptote (physically, if there are no people to infect other people ($I = 0$), nobody will be exposed)

Ultimately, we find that a video will go viral if $np \geq 1$. This makes sense, as if one infected person shares the link to a group of people, and at least one person from that group re-shares the video, the cycle can keep going on, increasing the number of exposed people without bound. Our equations $E(t)$ and E_t aren't too useful, as our models $I(t)$ and I_t , and understanding of the situation, tells us all we need. However, it doesn't hurt to have models for the exposed population. Our model also doesn't work very well for large values of np , as the susceptible population must be taken into account in that case (there is a finite amount of people). Furthermore, we assume that an exposed person cannot be a part of the people n who see the video (when S is infinite, the probability of this occurring is 0).