The Study of Spread of Rumors in Society

Study in Deterministic Mathematical Modeling

Submitted to:

Professor Berry

By:

Danielle Southern
670 0180

April 7, 1997

<u>The Study of the Spread of Rumors in a Population</u>

Infectious diseases spread through a population by process of social diffusion. Much like communicable diseases, fads and technological innovations also spread through a community by the same process. There are a variety of models to explain and study the spread, severity and cause of infectious diseases. The study of these diseases or epidemics has been an on going study in Applied Mathematics. One of the first models ever used to assess the advantages of the vaccination control of the smallpox epidemic involves a nonlinear differential equation by Bernoulli written in 1760, and some of the theoretical papers by Kermack and McKendrick from 1927 on epidemic models have had major influence in the development of mathematical models in the fields of both biology and epidemiology. Epidemics are modeled in order to determine which features of a community can be controlled, thus reducing the risk of spread. This paper will focus on simple models which incorporate Kermack and McKendrick's theories as well as general aspests of epidemiological modelling of disease transmission and the spread of epidemics over time. We will look at a type of model in which the total population is approximately constant and can be subdivided into three distinct groups, namely, those susceptible to the disease, those infected or suffering from the disease and those who are removed (by isolation or death). Such models are refered to as SIR models, as they can be denoted schematically as:

$$S \rightarrow I \rightarrow R$$

An epidemic is the appearance of an infectious disease attacking a whole community or population within a certain time period. A rumor can be studied much like an epidemic. A rumor, appears in a population and spreads (attacks) rapidly through the whole population. This spread can also be studied over a period of time. We consider a small group of villages knowing a rumor and insert them into a large population of equally susceptible villages. We want to know what occurs as time evolves and if the rumor dies out rapidly or spreads similar to an epidemic? Once studying these questions, we will ask: How many villages will ultimately learn of the rumor?

Formulation of the Model:

We will consider a population of villages which are small and isolated. The only form of communication between these villages is through a telephone system by which only one pair of villages may be in contact with each other at a time. Once a village has heard the rumor, the entire village knows of it immeadiately. (ie. they are infected immeadiately).

Assumptions:

First we must assume that:
i) The rumor is transmitted between a village that has heard the rumor and a village that has not heard it (ie. once a village has heard a rumor they only make calls and do not receive anymore).
ii) The rumor is known to the whole village immeadiately upon hearing it. In other words the incubation period is short enough to be negligible.
iii) All villages have equal probability of hearing the rumor.
iv) The population being studied is fixed in size and no one is entering the population by either birth or migration. We also do not take into account the deaths.

First we consider a trivial model in which we take the population as effectively infinite in size and having, initially, all villages in the population susceptible to hearing the rumor, with the exception of a small fraction of villages that already know the rumor. These two distinct groups (suceptible and infective) change with time. Susceptible villages become infective upon hearing the rumor. We let time be an independent variable, t and I(t) be the number of villages that have heard the rumor at time t and let:

> B = (constant) the average number of contacts with susceptible individuals which lead to a new village hearing the rumor/time/village already familiar with the rumor. Also let $B \geq 0$.

To deduce the number of villages learning of the rumor at time $t+\Delta t$ in terms of the number of villages already familiar with the rumor at time t, we take the sum of the number of infectives at time t and the number of new infectives in the time interval from $t+\Delta t$:

$$I(t+\Delta t) = I(t) + B * I(t) * \Delta t \qquad (1)$$

Note: There is an inconsistency here since $I(t+\Delta t)$ should not be expected to be an integer even if $I(t)$ is. We will notice later that this model is inexact so we interpret $I(t)$ to the nearest integer. Rearranging (1) we get:

$$\frac{I(t+\Delta t) - I(t)}{\Delta t} = B * I(t) \qquad (2)$$

Taking the limit of the left hand side of (2) as $\Delta t \to 0$ is simply the definition of a derivative. Taking the limit then gives:

$$\frac{d\,I(t)}{dt} = B * I(t) \qquad (3)$$

<u>Analysis of Model:</u>

Now (3) is a simple, separable ordinary differential equation. Solving with the initial condition, $I(0)=I_0$ gives:

$$\frac{1}{I(t)}\,d\,I(t) = B\,dt \qquad (4)$$

Then taking the integral of both sides of (4) gives:

$$\ln| I(t) | = B * t + C$$

Note: since $I(t) \geq 0$, $\ln| I(t) | = \ln(I(t))$.

$$\ln(I(t)) = B * t + C$$

Then taking the inverse log of both sides gives:

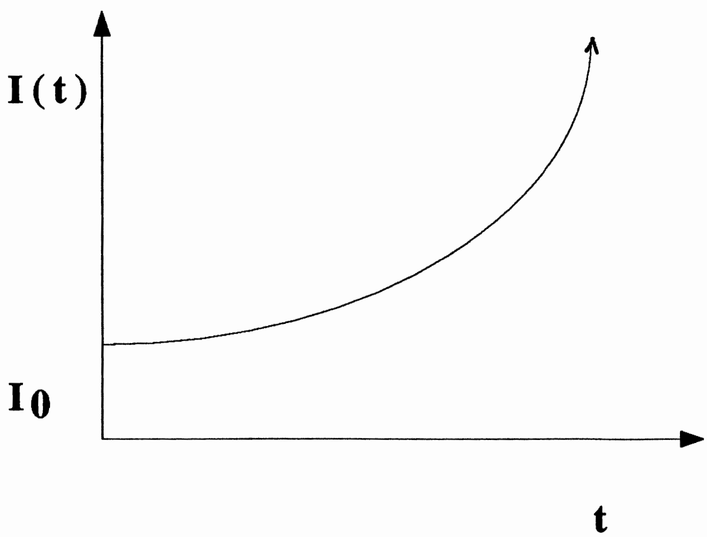
$$I(t) = D * Exp(B * t) \qquad \text{where } D = Exp(C) > 0$$

Now, with the initial condition:

$$I(0) = D * 1$$

$$=> \qquad I(t) = I_0 * Exp(B * t) \qquad (5)$$

Graphing this curve of I(t) versus t gives an exponentially increasing curve since $B \geq 0$. We note that as $t \to \infty$, I_0 grows with unlimited growth.



In order to make up for this problem we can assume that rather than having an infinite population, we have a finite number of susceptible villages. Modifying the above model gives a new model in which every village is either susceptible to the rumor or already familiar with the rumor.

<u>Modification of Model:</u>

We define the new variables as:

S(t) = the average number of villages susceptible to the rumor at time t.

$$ß = \frac{B}{S(t)}$$ (no longer a constant, varies with the number of susceptible villages.)

The new model is now:

$$\frac{d\,I(t)}{dt} = ß * S(t) * I(t) \qquad (6)$$

<u>Analysis of New Model:</u>

Since the total population is finite in size, we'll say that the total population is:

$$N = I(t) + S(t)$$

$$\Rightarrow \qquad S(t) = N - I(t)$$

Now (5) becomes:

$$\frac{d\,I(t)}{dt} = ß * (N - I(t)) * I(t) \qquad (7)$$

exponentially logistic.

Separating and integrating both sides gives:

$$\frac{1}{N}\int \frac{1}{I(t) * (N - I(t))}\, d\,I(t) = \int ß\, dt$$

missing factor!

Using partial fractions on the left hand side and solving gives:

$$\int \frac{1}{I(t)}\, d\,I(t) + \int \frac{1}{(N - I(t))}\, d\,I(t) = ß * t + C$$

$$\Rightarrow \qquad \ln|\, I(t) \,| - \ln|\, N - I(t) \,| = ß * t + C$$

$$\Rightarrow \qquad \ln\left(\frac{I(t)}{N - I(t)}\right) = ß * t + C$$

Taking the inverse log of both sides gives:

$$\frac{I(t)}{N - I(t)} = D * Exp(ß * t) \qquad \text{where } D = Exp(C) > 0$$

$$\Rightarrow \qquad I(t) = D * N * Exp(ß * t) - D * I(t) * Exp(ß * t)$$

$$I(t) = \frac{D * N * Exp(ß * t)}{(1 + D * Exp(ß * t)}$$

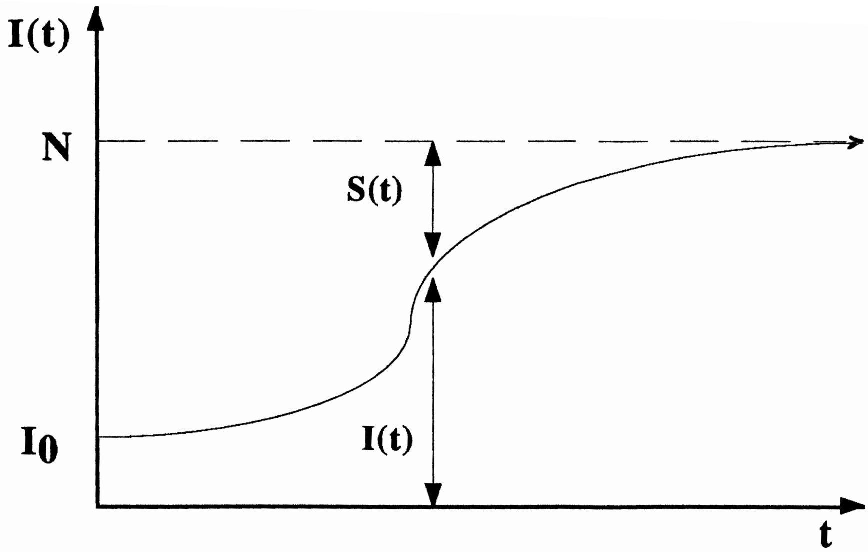
$$I(t) = \frac{N}{(1 + F * Exp(-ß * t))} \qquad \text{where } F = 1 / D > 0$$

Solving for F with the initial condition $I(0) = I_0$ gives:

$$I(t) = \frac{N * I_0}{(I_0 + (N - I_0) * Exp(-ß * t))} \qquad (8)$$

Note we can see here that as $t \rightarrow \infty$ the denominator approaches I_0 and so I(t) approaches N. Also if we divide N from the numerator and denominator and then take $N \rightarrow \infty$ we can see that we will obtain the first model (5).

We can thus conclude here that with a small number of villages familiar with the rumor, the rumor spreads through the villages exponentially until only a small number of villages susceptible to hearing the rumor are left. The rate of spread then decreases and ends when all the villages have heard the rumor. We can see this on the following page in the graph of I(t) versus t.



This model is significantly improved in comparison to the first model, however, this model implies that all villages hear the rumor and are forever familiar with it as well as forever interested in spreading the rumor. It also assumes that the whole population learns of the rumor. Realistically some villages can avoid the rumor altogether (ie. they don't answer the phone). A more realistic approach to the problem must then take into account the villages that either forget the rumor or loose interest in spreading to the remaining susceptable villages.

good

<u>Modification of second model:</u>

In this modified model we will define the new variables:

> $R(t)$ = the number of removed villages at time t. (ie. villages that are no longer interested in spreading or have forgotten the rumor)
> ∂ = average rate of removal of villages learning the rumor/unit time/village already familiar with the rumor. Also note $\partial \geq 0$.

The gain in villages learning the rumor is at a rate proportional to both the number of villages familiar and those still susceptible to the rumor, ß SI. The susceptible villages are lost at the same rate, namely -ß SI. The rate of removal of villages familiar with the rumor to those removed is proportional to the number of villages familiar with the rumor, ∂I.

The modified model now becomes:

$$\frac{d\,S(t)}{dt} = -ß * S(t) * I(t) \tag{9}$$

$$\frac{d\,I(t)}{dt} = ß * S(t) * I(t) - \partial * I(t) \tag{10}$$

$$\frac{d\,R(t)}{dt} = \partial * I(t) \tag{11}$$

Analysis of Modified Model:

All villages in the population are either susceptible, familiar with the rumor, or no longer interested in spreading the rumor. So the total population is:

$$N = S(t) * I(t) * R(t)$$

Assume initial conditions:

$$S(0) = S_0 = N - I_0 > 0$$

$$I(0) = I_0\,,\quad 0 < I_0 << N$$

$$R(0) = 0$$

We want to look at whether the rumor will spread and if so, how it does with time. We know that $\partial > 0$ and $I(t) \geq 0$, so we can deduce that dR/dt is always positive, which makes R a non-decreasing function of t. We also know that ß > 0 and both I(t) and S(t) are positive so $dS/dt \leq 0$ for all t, making S a non-increasing function. This makes sense as the number of susceptibles must decrease as an epidemic spreads.

$$\frac{d\,I(t)}{dt} = I(t)\,[\,ß * S(t) - \partial]$$

Setting $f = \partial/ß$:

$$d\,I(t) = ß * I(t)\,[\,S(t) - f]$$

9

We know $I(t) \geq 0$, so

$$d\,I(t) > 0 \qquad \text{when } S(t) > f$$

$$d\,I(t) < 0 \qquad \text{when } S(t) < f$$

There will be no epidemic if $S < f$. The critical value $S = f$ for which $dI/dt = 0$ is what epidemiologists term a threshold value. The initial number of susceptible villages must exceed the threshold value for the rumor to spread. This is an important result of Kermack-McKendrick's Threshold Theorem. The theorem states that an infection determines a threshold size for the susceptible population, above which an epidemic propagates.[1]

Eliminating I(t) in (9) and (11) gives:

$$\frac{d\,S(t)}{dt} = -\frac{S(t)}{f} * \frac{d\,R(t)}{dt}$$

Separating and multiplying through by dt then gives:

$$\frac{d\,S(t)}{S(t)} = -\frac{d\,R(t)}{f}$$

$$\int \frac{d\,S(t)}{S(t)} = -\int \frac{d\,R(t)}{f}$$

$$\Rightarrow \qquad \ln |\,S(t)\,| = -\frac{R(t)}{f} + C$$

[1] Hoppensteadt, *Mathematics in Medicine and the Life Sciences*, 1992

$$S(t) = D * Exp(-R(t)/f) \qquad \text{where } D = S_0 \qquad (12)$$

We know that $R(t) \leq N$ for all t, so $Exp(-R/f) \geq Exp(-N/f)$ giving:

$$S(t) \geq S_0 \, Exp(-N/f)$$

$$\Rightarrow \qquad S(\infty) = \lim_{t \to \infty} S(t) > 0$$

We can deduce from this that there will always be some number $S(\infty)$ of villages that avoid learning the rumor. The rumor will eventually die out, but not due to the loss of susceptible villages.

Now, studying the relationship between the infectives and susceptibles involves dividing (2) by (1):

$$\frac{dI}{dS} = \frac{ßSI - \partial I}{-ßSI}$$

$$\frac{dI}{dS} = -1 + \frac{f}{S}$$

Separating and integrating gives:

$$\int dI = \int \left(-1 + \frac{f}{S}\right) dS$$

$$I(S) = -S + f \ln|S| + C$$

Setting t = 0, we can solve for C:

$$I_0 = -S_0 + f \ln|S| + C$$

$$\Rightarrow \qquad C = I_0 + S_0 - f \ln|S_0|$$

$$I(S) = I_0 + S_0 - S + f \ln\left(\frac{S}{S_0}\right) \tag{13}$$

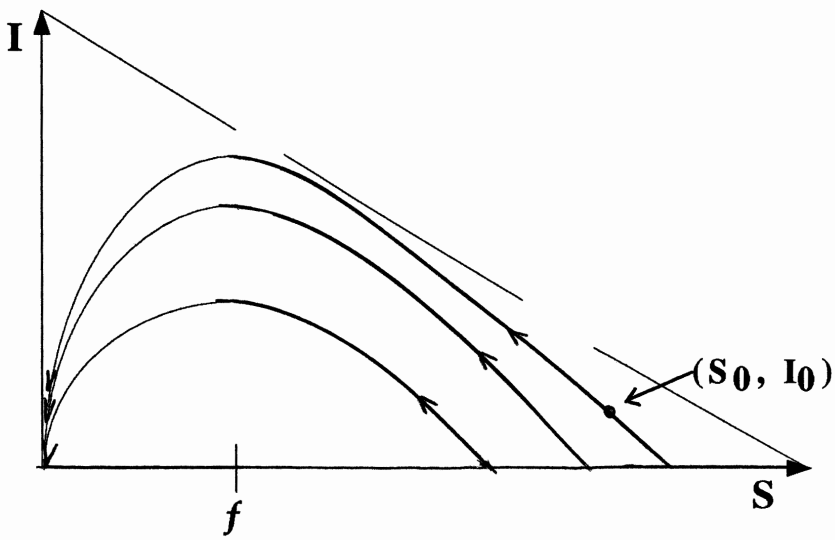
$N = I_0 + S_0$ so $I(S)$ can be re-written as:

$$I(S) = N - S + f \ln\left(\frac{S}{S_0}\right)$$

It can be seen that $I(0) = -\infty$, $I(S_0) = I_0 > 0$ and $I(S(\infty)) = 0$. Taking the second derivative of I gives:

$$I(S)'' = -\frac{f}{S^2} < 0$$

Since the second derivative is negative the curve will be concave down with a maximum occuring at $dI/dt = 0$, or $S = f$. Graphing I versus S looks like the following:



As t travels from t° to ∞ $(S(t), I(t))$ travels along (13) and $S(t)$ decreases monotonically with time. If $S < f$, $I(t)$ decreases to zero and $S(t) \rightarrow S(\infty)$. So if I_0 is inserted into a population with S_0 susceptibles and $S_0 < f$, the rumor will die out rapidly. If on the otherhand, $S_0 > f$, $I(t)$ will increase as $S(t)$ decreases to f and achieves maximum at $S = f$. We can thus conclude that the rumor spreads if and only if $S_0 > f$ where

$f = \partial / ß$ is the threshold value. We can also conclude that the rumor spreads only with the lack of villages knowing the rumor and not with the lack of susceptible villages.

We can never know for sure the number of villages that have learned the rumor since the only recognizable villages are those who seek attention/aid (ie. they ask for verification of the rumor). We can get an idea of the number however, by studying dR/dt rather than dI/dt. Since the relative removal rate varies with the community, it is dR/dt which determines whether a rumor can spread in one population and not another. For example, if the density of the susceptible villages is high and the removal rate ∂ of villages having learnt the rumor is low (ie. population is full of gossips), then a rumor is likely to spread through a community.

Using $N = S(t) + I(t) + R(t)$ in (11) gives:

$$d\,R(t) = \partial\,[N - R(t) - S(t)]$$

then replacing S(t) by (12), we have:

$$d\,R(t) = \partial\,[N - R(t) - S_0 * \mathrm{Exp}(-R(t)/f)] \qquad (14)$$

The methods for solving this differential equation are rather complicated, however we can compute a numerical solution. By expanding the exponential term we get:

$$\mathrm{Exp}(-R(t)/f) = 1 - \frac{R(t)}{f} + \frac{[R(t)]^2}{2!\,f^2} - \frac{R(t)]^3}{3!\,f^3} + \dots\,]$$

If we truncate this upto the cubic, the resulting integration becomes difficult so we will truncate upto the quadratic term to obtain an approximate solution. So:

$$\frac{d\,R(t)}{dt} = \partial^*[I_0 + S_0 - R(t) - S_0\{1 - \frac{R(t)}{f} + \frac{R(t)^2}{2!f^2}\}]$$

$$\frac{d\,R(t)}{dt} = \partial^*[I_0 + \{\frac{S_0}{f} - 1\}R(t) - S_0 * \frac{R(t)^2}{2!f^2}] \qquad (15)$$

Completing the square and then separating, making a hyperbolic trigonometric substitution and integrating will then give:

$$\frac{d\,R(t)}{dt} = -\frac{S_0\,\partial}{2^*f^2}{}^*\,[R^2 - \frac{2^*f^2}{S_0}\{\frac{S_0}{f} - 1\}^*R - \frac{2^*f^2{}^*I_0}{S_0}]$$

$$\frac{d\,R(t)}{dt} = -\frac{S_0\,\partial}{2^*f^2}{}^*\,[(R - \frac{f^2}{S_0}\{\frac{S_0}{f} - 1\})^2 - \frac{2^*f^2{}^*I_0}{S_0} - \frac{f^4}{S_0^2}\{\frac{S_0}{f} - 1\}]$$

$$\frac{d\,R(t)}{dt} = -\frac{S_0\,\partial}{2^*f^2}{}^*\,[(R - \frac{f^2}{S_0}\{\frac{S_0}{f} - 1\})^2 - \{\frac{f^2}{S_0}((\frac{S_0}{f} - 1)^2 + \frac{2^*S_0{}^*I_0}{f^2})^{(1/2)}\}^2]$$

Setting:

$$A = ((\frac{S_0}{f} - 1)^2 + \frac{2^*S_0{}^*I_0}{f^2})^{(1/2)}$$

$$B = \frac{f^2{}^*A}{S_0}$$

$$\Omega = \frac{f^2}{S_0}{}^*(\frac{S_0}{f} - 1)$$

$$\frac{d\,R(t)}{dt} = -\frac{S_0\,\partial}{2^*f^2}{}^*\,[(R - \Omega)^2 - (B)^2] \qquad (16)$$

$$\frac{dR}{\{R - \Omega\}^2 - B^2} = - \frac{S_0 \partial}{2f^2} dt$$

let $R - \Omega = B \tanh(y)$ (17)

$$\Rightarrow \quad dR = B \{\text{sech}(y)\}^2 dy$$

$$\Rightarrow \quad \frac{B \{\text{sech}(y)\}^2}{B^2 (\{\tanh(y)\}^2 - 1)} dy = - \frac{S_0 \partial}{2f^2} dt$$

$$- \frac{dy}{B} = - \frac{S_0 \partial}{2 f^2} dt$$

$$- \int dy = - \frac{B \int S_0 \partial}{2 f^2} t$$

$$y = \frac{B S_0 \partial}{2 f^2} t - B{*}C$$ where C = constant of integration

Rearranging (16) we have:

$$R = \Omega + B \tanh(y)$$

$$\Rightarrow \quad R = \Omega + B \tanh\left(\frac{B S_0 \partial}{2f^2} t - B{*}C \right) \qquad (18)$$

At time t = 0, R(0) = 0 so (17) becomes:

$$0 = \Omega - B \tanh(B{*}C)$$

$$\Rightarrow \quad \frac{\Omega}{B} = \tanh(B{*}C)$$

$$B^*C = \tanh^{-1}(\Omega/B) = ø \tag{19}$$

$$=> \qquad R(t) = \frac{f^2}{S_0}\left[\left(\frac{S_0}{f} - 1\right) + A\tanh(ø)\right]$$

where:

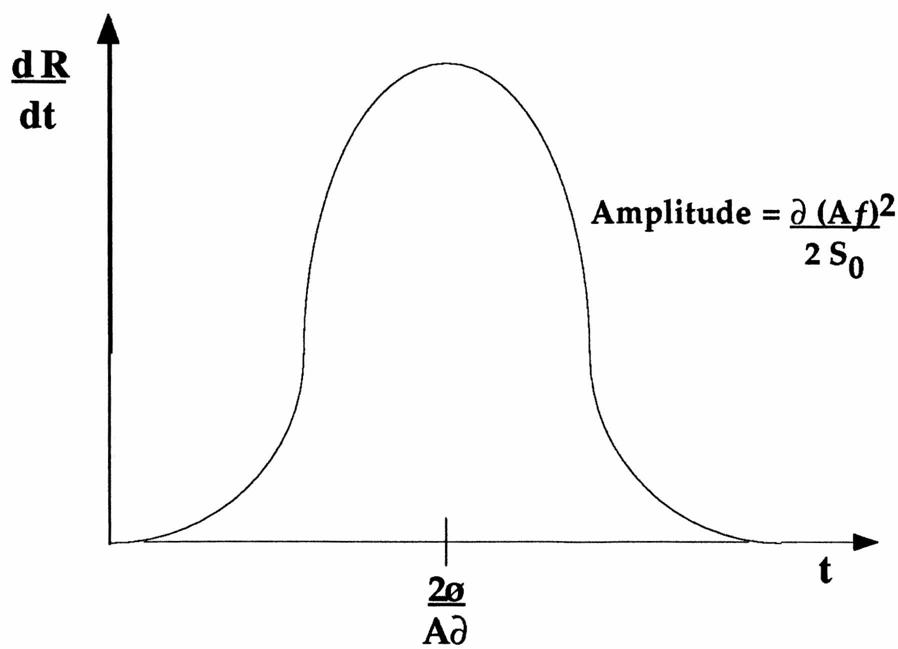
$$B = \left\{\frac{f^2 A}{S_0}\right\}^2$$

$$A = \left\{\left(\frac{S_0}{f} - 1\right)^2 + \frac{2 S_0 I_0}{f^2}\right\}^{(1/2)}$$

The epidemic curve is now:

$$\frac{dR}{dt} = \partial\,\frac{(Af)^2}{2 S_0}\left\{\mathrm{sech}\left(\frac{A\partial}{2} t - ø\right)\right\}^2 \tag{20}$$

Which when graphed looks like the following:



Consider now the total size of the population. Looking at the limit of R(t) as t -> ∞, we have:

$$\lim_{t \to \infty} R(t) = R(\infty) = \frac{f^2}{S_0} \left[\frac{S_0}{f} - 1 + A \right]$$

We assumed earlier that I_0 was very small, so that:

$$\frac{2 S_0 I_0}{f^2} << \frac{S_0}{f} - 1 \qquad \text{where } S_0 > f$$

it follows then that:

$$A \approx \frac{S_0}{f} - 1$$

And $R(\infty)$ be further written as:

$$R(\infty) = 2 f \left(1 - \frac{f}{S_0} \right)$$

Since we deduced earlier that $S_0 > f$ must be a condition for an epidemic outbreak, we will assume: $S_0 = f + E$. Where $0 < E << f$. Now:

$$R(\infty) \approx 2f \left(\frac{E}{f + E} \right) \approx 2 E$$

So a population with, initially, $S_0 = f + E$ susceptibles, is reduced to $S(\infty) = f - E$ susceptibles left (the value is as far below the threshold, f, as it originally was above). Kermack and McKendrick compared equation (20) with data from the 1905 Plague in Bombay. They found that (20) was a good approximation to the data and thus a reliable and accurate model to explain the spread of infectious diseases or rumors as in this case.

This model however can still be modified. If the time period over which the rumor spreads is not of short duration, birth and death terms should be included in the susceptible term (9). Natural deaths should also be taken into consideration in the infective term (10) and the removed term (11). Many diseases also have latent periods. For example when a village hears of a rumor, the receiver passes the rumor on in their village before they spread it further (ie. the rumor takes time to actually hit the village once they've heard it). This factor could be added as a delay factor and we can introduce a new class, say L(t) in which the susceptible remains for a period of time before becoming infective. Age is also an important factor in diseases since some are more prone to becoming infected with a disease than others depending on their age. We can say for this example that some villages are more prone to hear of a rumor depending on their age as well. For example a village that is fairly young may not have interest in spreading a rumor an older village communicates to a younger village. In general the model does not incorporate the variation in behaviour of villages. In other words it does not take into account age as well as sexual and racial differentiation. All these factors should be taken into account and the model should be further modified and analysed. The final model deduced is however sufficient for our assumptions and seems to be an adequate model to describe the spread of rumors in a population.

good

Bibliography

1. Bailey, N.T.J. Mathematical Theory of Epidemics; Charles Griffin and Company Ltd., London, 1957.

2. Bailey, N.T.J. The Mathematical Theory of Infectious Diseases and its Applications; Griffin, London, 1975.

3. Braun, Martin, Differential Equations and its Applications (Applied Mathematical Sciences, v.15); Springer-Verlag, 1975.

4. Coles, C.W., Eason, G. and Gettinby, G. Mathematics and Statistics for the Bio-Sciences; Ellis Horwood Ltd., 1980, p.267-268.

5. Edelstein-Keshet, L. Mathematical Models in Biology, Random House, New York, 1988, p.242-269.

6. Frauenthal, J.C. Mathematical Modelling in Epidemiology; Springer-Verlag, New York, 1980.

7. Goel, N.S. and Richter-Dyn, N. Stochastic Models in Biology; Academic Press, New York, 1974.

8. Hoppensteadt, F.C and Peskin, C.S. Mathematics in Medicine and the Life Sciences; Springer-Verlag, New York, Inc., 1992, p.67-79.

9. Jones, D.S. and Sleeman, B.D. Differential Equations and Mathematical Biology; George Allen and Unwin Ltd., UK, 1983.

10. Murray, J.D. Mathematical Biology; Springer-Verlag, Berlin Heidelberg, 1989, p.610-696.

11. Murthy, D.N.P, Page, N.W. and Rodin, E.Y. Mathematical Modelling: A Tool for Problem Solving in Engineering, Physical, Biological and Social Sciences; Pergamon Press, 1990, p.237-258.

12. Olinick, Michael, An Introduction to Mathematical Models in the Social and Life Sciences; Addison-Wesley Publishing Company, Inc., Phillipines, 1978.

13. Waltman, P. Lecture Notes in Biomathematics: Deterministic Threshold Models in the Theory of Epidemics; Springer-Verlag, New York, 1974.

14. Wickwise, K. Mathematical Models for the Control of Pests and Infectious Diseases: A Survey of Theoretical Population Biology; 1977, p.182-258.

15. Epidemiolgy: Proceedings of a Sims Conference on Epidemiology, Alta, Utah, July 8-12, 1974; Society for Industrial and Applied Mathematics, U.S.A., 1975.

16. Lecture Notes in Biomathematics: Mathematics in Biology and Medicine, Proceedings, Bari 1983; Springer-Verlag, Berlin Heidelberg, 1985, p.106-185

17. Lecture Notes in Biomathematics: Mathematical Models in Medicine, Workshop, Mainz, March 1976; Springer-Verlag, Berlin Heidelberg, 1976.

18. Mathematical Models in Biology and Medicine: Proceedings of IF18-TC4, Working Conference on Mathematical Models in Biology and Medicine, Varna, 6-11, Sept, 1972; North-Holland Publishing Company, Amsterdam London, p.1-7.