A Logistic Model for the Spread of Epidemics

Vajira Pathirana

University of Manitoba

A LOGISTIC MODEL FOR THE SPREAD OF EPIDEMICS

Submitted as a requirement for the course 6.337

by

Vajira Pathirana

(Student Number : 6747106)

on

April 5, 1999

1 What is an epidemic?

An epidemic is a widespread occurrence of an infectious disease in a community that may affect a significant portion of the population. Epidemics can show catastrophic effects on populations. Cholera, small-pox, leprosy, bubonic plague, AIDS are few reported epidemics. A customary approach to the control of contagious diseases is via medical intervention, either by preventive vaccination or by the use of antibiotics. But there is enough evidence that sociological control had been the customary method of combating epidemics in the history when the new medical achievements which we enjoy today were not available. But even today sociological control is the first step of prevention from an epidemic, especially for a new disease, despite the fact that they may be more painful.

2 Reported epidemics in the history

The history of the world shows major outbreaks of epidemics at different parts of the world at different times. These had led to severe destruction of human population and civilizations. The most common epidemics reported were bubonic plague, smallpox, cholera, and leprosy. The mortality rate had been so high for epidemics such as smallpox. The main method of control of such epidemics at those time had been sociological control (like isolation) until medical solutions were found. But it's not possible to find correct statistics or any sort of mathematical analysis done on such outbreaks of epidemics.

not be the same in the two cases. In such situations the population cannot be considered as uniformly mixing. So it is incorrect to consider a uniformly mixing population in the mathematical model.

In addition if the spreading of the disease is only significant in small groups of the population then it is difficult to build a model to represent the overall spread of the disease considering the total population. Although an overall model may be built considering the small groups the growth rate of the disease could be different in each area due to the physical parameters concerned.

Besides, the statistics available at an epidemic situation may not be complete due to the nature of the disease itself, and it would be difficult to calculate the parameters of the mathematical model which may lead to inaccurate predictions. Infective people may tend to pretend about the disease or susceptible population may start to keep away from the infected area, which makes it really difficult to collect accurate figures.

5 Description of the model

Consider different types of individual groups within a population in an epidemic situation. Four such groups can be identified :

S : The subgroup consisting of individuals who are uninfected, but susceptible to the disease. They are known as *susceptibles.*

A : The subgroup consisting of individuals who are infected, but not yet capable of spreading the disease. They are known as *affecteds.*

I : The subgroup consisting of individuals who are infected and spreading the disease. They are known as *infectives.*

U : The subgroup consisting of individuals who do not belong to any of the above categories. It consist of those who are not susceptible to the disease at all or those who have been infected, healed fully and removed from S. They are known as *unaffecteds.*

The number of individuals in each subgroup S,A,I and U at an arbitrary time t are denoted by *s(t), a(t), i(t)* and *u(t)* respectively.

If we consider that the epidemic started at time $t = t_o$ then the number of individuals in each subgroup S,A,I and U at time t_0 is given by s_0, a_0, i_0 and u_0 respectively.

Within a real epidemic situation we may observe all four categories of individuals depending on the disease. The duration in which an individual is a member of A or I varies from disease to disease and individual to individual. In addition according to the nature of the disease an individual leaving subgroup I may join either subgroup S or U.

The total population can be denoted by P. The number of individuals in P at an arbitrary time t is denoted by *p(t).* P is made up of above four categories.

$$p(t) = s(t) + a(t) + i(t) + u(t)$$

For an epidemic to develop there must be at least one infective in the population at the beginning. Therefore at time $t = 0$,

$$i(0) \geq 1$$

An epidemic does not exist anymore if there are no more individuals in the subgroups A and I. Therefore at a certain time $t = T$,

$$i(T) = 0 \; ; \qquad a(T) = 0$$

6 Assumptions made

To build a simple and realistic mathematical model several assumptions have to be made. It is assumed that an individual is capable of spreading the disease as soon as he gets infected. Then the subpopulation A does not have any members and $a(t) = 0$ for all t. Then the total population is given by

$$p(t) = s(t) + i(t) + u(t)$$

To make the model more simple it is assumed that all the individuals in P are susceptible to the disease. Therefore the group U can be incorporated into S. Then the subpopulation U does not have any members and $u(t) = 0$ for all t. Then number of individuals in P is given by

$$p(t) = s(t) + i(t)$$

It can be assumed that the time required to heal the disease is much longer than the time required for it to spread through the population. So the infected individuals are capable of spreading the disease to others throughout the entire period they are infected. Therefore we can assume that over a long period of time every possible contact between individuals of S and I can take place. This means that P is a uniformly mixing population.

Further we can assume that each individual has the same probability of getting infected independent of their age, living standards, sex etc. Therefore the population P can be assumed as homogeneous when the epidemic is considered.

With the assumptions made the number of individuals of the total population at time t is given by

$$p(t) = s(t) + i(t)$$

To develop a mathematical model we can consider $p(t)$ as a

1. a constant. If we assume that the disease exists only for relatively short periods and the population consists of moderate birth and death rates then the population is essentially a constant. Therefore

$$p(t) = P$$

where P is a constant

2. variable of time t. Taking competition also into consideration the population can be assumed to have a *logistic growth rate* given by

$$p(t) = \frac{C}{1 + Fe^{-kt}} \qquad k, C, F > 0$$

where C is the *carrying capacity* of the population, k the *initial relative growth rate* of the population and F a constant.

7 Mathematical model

If we consider that the spreading of the disease is due to contact, then the rate of growth of the infected population is proportional to the number of contacts between individuals of infected subpopulation I and individuals of susceptible subpopulation S. Assuming a uniformly mixing population, the number of contacts per unit time at time t is given by the product of the number of members in each group at time t . Therefore

$$\begin{aligned} \frac{di(t)}{dt} &\propto s(t).i(t) \\ &= m.s(t).i(t) \\ &= m.[p(t) - i(t)].i(t) \end{aligned}$$

where m is a positive constant known as *infection rate.* This might depend on the nature of the disease and other physical parameters involved in the situation. In the same way we can show that the rate of decay of susceptible population is given by,

$$\frac{ds(t)}{dt} = -m.[p(t) - s(t)].s(t)$$

7.1 Case 1 : $p(t)$ is constant

If the time interval involved is relatively small then the total population can be considered as a constant.

$$\begin{aligned} p(t) &= P \qquad (P > 0) \\ \frac{di(t)}{dt} &= m.[P - i(t)].i(t) \\ &= mP.[1 - \frac{i(t)}{P}].i(t) \end{aligned}$$

This gives a logistic curve with initial relative growth rate mP and carrying capacity P.

Figure1 gives the curve di/dt vs. i. $di/dt = 0$ at $i = 0$ and $i = P$. Out of the two, $i = 0$ is an unstable equilibrium point while $i = P$ is a stable equilibrium point. di/dt has a maximum and $d^2i/dt^2 = 0$ at $i = P/2$. Therefore $i(t)$ must have a point of inflection at $i = P/2$ if the initial population $i_0 < P/2$ and $i(t)$ will not have a point of inflection if $i_0 > P/2$.

$$\begin{aligned} \int \frac{di}{(P-i)i} &= \int mdt \\ ln \mid i \mid -ln \mid (P-i) \mid &= mPt + C_1 \end{aligned}$$

where C_1 is a constant. Since $i > 0$ and $P > i$

$$\frac{i}{(P-i)} = De^{mPt}$$

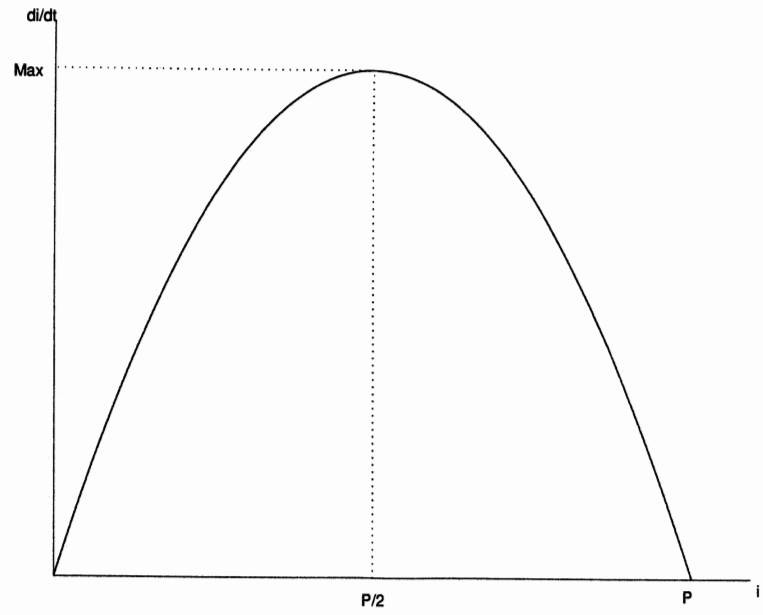

Figure 1: *Rate of change of infective population vs. number of infectives when $p(t)$ is constant*

$$i(t) = \frac{P}{1 + Ge^{-mPt}}$$

where $D = e^{C_1}$ and $G = \frac{1}{D}$.

The graph $i(t)$ vs. t has three types of variation depending on the initial population i_0 for same m and P as shown in *figure2.* The case $i_0 > P$ is not of interest. $i(t)$ will have a point of inflection at time $t = t^*$ if $i_0 < P/2$ and $i(t)$ will approach P for larger values of t. $i(t)$ will not have a point of inflection if $i_0 > P/2$. Since all epidemics initiate with a small number of infectives, the graph $i_0 < P/2$ is of more interest and is given by *figure 3.*

This represents a logistic curve. It can be observed that the number of infectives increase rapidly initially and then the growth rate starts to decrease after a time period t^*. At this time the infective population consists

actually has a very small growth rate initially, increases to a max & decreases to zero.

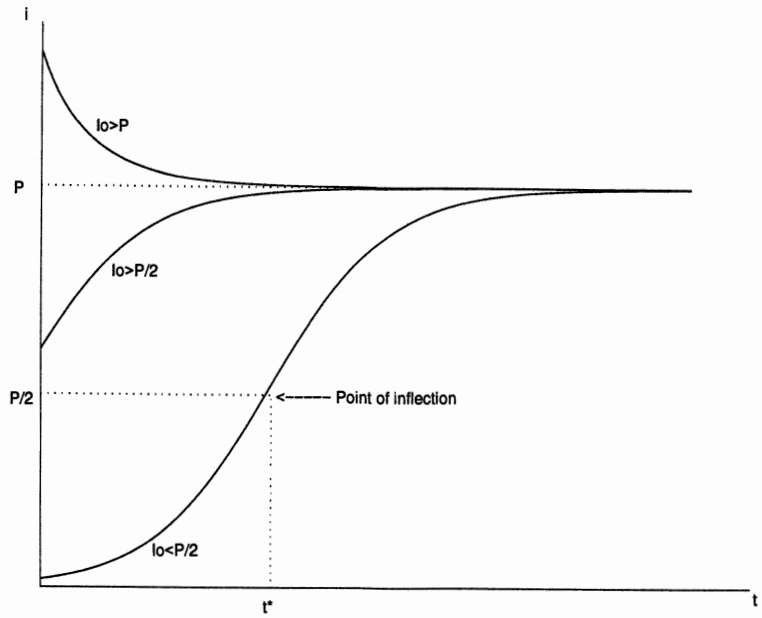

Figure 2: *Variation of infective population when p(t) is constant*

of half of the total population. This can be observed from *figure1* as well. The infective population has its maximum growth rate at this point. The number of infectives approaches P when time increases to infinity.

If the graph of di/dt vs. t is considered, it can be seen that at time t^*, the infective population grows most rapidly and then the rate of growth goes down and approaches zero as t increases to infinity. This is shown in *figure 4*.

$$\begin{aligned}\frac{di(t)}{dt} &= m.[P - i(t)].i(t) \\ &= m.[P - \frac{P}{1 + Fe^{-kPt}}].\frac{P}{1 + Fe^{-kPt}} \\ &= \frac{mP^2e^{-kPt}}{[1 + Fe^{-kPt}]^2}\end{aligned}$$

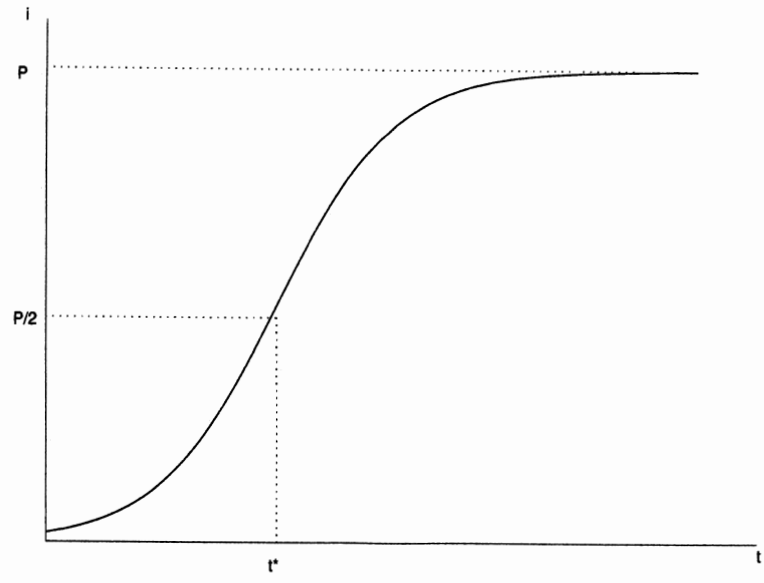

Figure 3: *Variation of infective population when p(t) is constant and* $i_0 < P/2$

The time t^* at which $i(t)$ shows the point of inflection is given by

$$t^* = -\frac{1}{kP} ln(\frac{1}{F})$$

At this time the rate of growth of the infective population shows a maximum.

7.2 Case 2 : Incorporating subpopulation U

In a more practical situation it can be seen that the number of individuals in the subpopulation U cannot be zero. Depending on the nature of the disease, there can be individuals in U who are permanent members of that group. For this the special case in which the individuals removed from I join U and does not return to I or S, is considered.

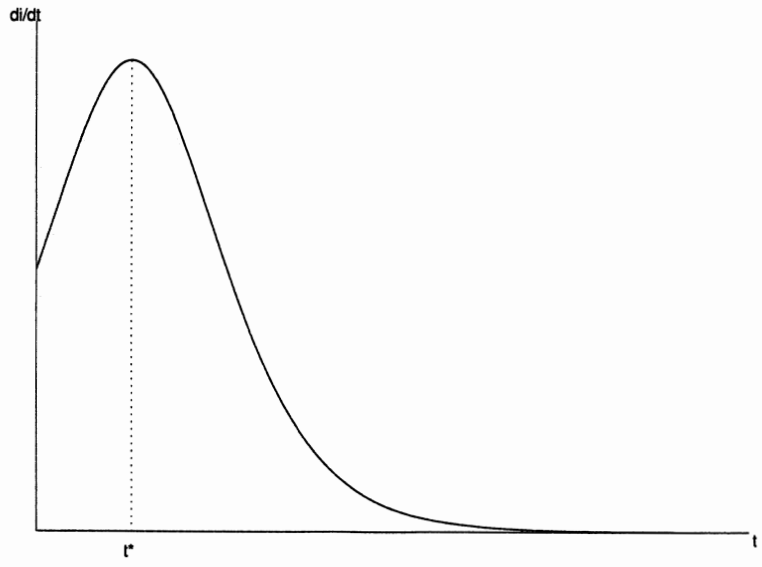

Figure 4: *Variation of rate of growth of infective population with time when p(t) is constant*

An epidemic may develop or a contagious disease may not enter the situation of an epidemic depending on the time taken for individuals to move between subpopulations S,I or U, the nature of the disease and the number of individuals in each group etc.

Now the total population at time t is given by,

$$p(t) = s(t) + i(t) + u(t)$$

If the total population is considered to be a constant, then the rate at which individuals leave I and join group U at time t can be assumed to be directly proportional to the number of infectives at that time. Therefore,

$$\frac{du}{dt} = r.i(t)$$

where $r(>0)$ is a constant known as the *removal rate* which depends on the nature of the disease and the physical parameters involved.

So the rate of change of total population $p(t)$ is given by the system of equations;

$$\begin{aligned}
\frac{ds(t)}{dt} &= -m.s(t).i(t) \\
\frac{di(t)}{dt} &= m.s(t).i(t) - r.i(t) \\
\frac{du(t)}{dt} &= r.i(t)
\end{aligned}$$

From the above system we can obtain the relationship

$$\begin{aligned}
\frac{ds(t)}{dt} &= -m.s(t).\frac{1}{r}\frac{du(t)}{dt} \\
lns(t) &= -\frac{m}{r}.u(t) + C_1
\end{aligned}$$

where C_1 is a constant. Applying the initial conditions

$$\begin{aligned}
C_1 &= lns_0 + \frac{m}{r}.u_0 \\
s(t) &= K_1 e^{-\frac{m}{r}u(t)}
\end{aligned}$$

where $K_1 = e^{C_1}$. This variation is shown on *figure5*. Therefore as expected in a real situation, the susceptible population decreases with the increase of unaffected population when the total population is constant.

Further,

$$\begin{aligned}
\frac{di(t)}{dt} &= m.s(t).i(t) - r.i(t) \\
&= m.i(t).[s(t) - \frac{r}{m}]
\end{aligned}$$

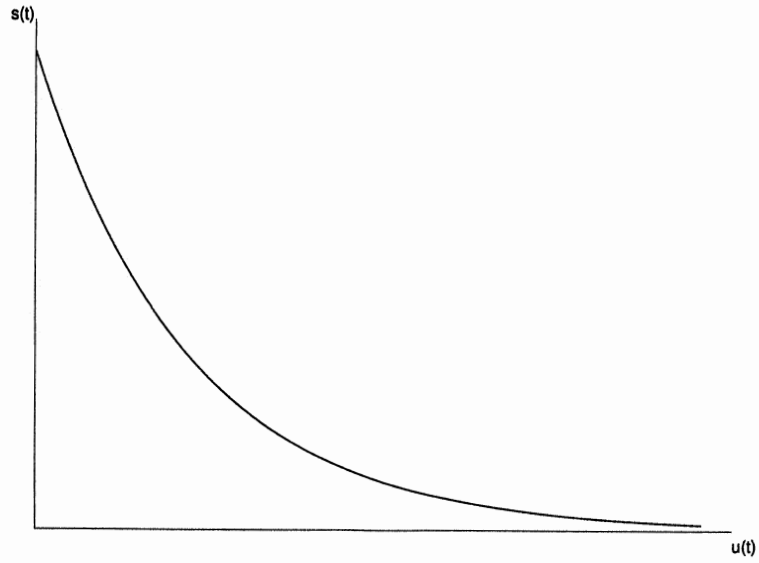

Figure 5: *Variation between susceptible and unaffected populations*

An epidemic may develop if di/dt is positive at a certain time t. Hence for a contagious disease to be developed to an epidemic $s(t) > r/m$, or the size of the susceptible population must be greater than the ratio between the removal rate and infection rate. This suggests that by improving r by means such as quarantining, the spreading of an epidemic can be hindered. In contrast, for a fixed ratio r/m, an epidemic will develop only if the size of the susceptible population increases above the threshold level r/m. As long as the number of susceptibles are below this level, a disease may prevail within a community without becoming an epidemic.

The infective population will increase with a positive growth rate until the size of the susceptible population decreases to its threshold value and then the growth rate will decrease with further decrease of susceptible population.

7.3 Case 3 : $p(t)$ is varying

If the time involved in spreading of the epidemic is significant, then the total population $p(t)$ cannot any longer be considered as a constant. Hence $p(t)$ has to be considered as a varying population. In this case, the total population is assumed to have a growth given by a logistic model.

$$p(t) = \frac{C}{1 + Fe^{-kt}}$$

where C and k are positive constants known as the carrying capacity and initial relative growth rate of the population respectively while F is a positive constant.

To develop a model lets assume that the total population $p(t)$ consists only of subpopulations I and S. Hence,

$$\begin{aligned} p(t) &= i(t) + s(t) \\ \frac{di}{dt} &= m.i(t).s(t) \\ &= m.i(t).[p(t) - i(t)] \\ &= m.i(t).[\frac{C}{1 + Fe^{-kt}} - i(t)] \\ \frac{di}{dt} - \frac{mC}{1 + Fe^{-kt}}.i &= -m.i^2 \end{aligned}$$

This is a Bernoulli's equation which could be solved by applying $z = i^{-1}$.

$$\begin{aligned} \frac{dz}{dt} + \frac{mC}{1 + Fe^{-kt}}.z &= m \\ \frac{d}{dt}[(F + e^{kt})^{\frac{mC}{k}}.z] &= m.[F + e^{kt}]^{\frac{mC}{k}} \\ (F + e^{kt})^{\frac{mC}{k}}.z &= m\int[F + e^{kt}]^{\frac{mC}{k}}dt + C_1 \\ i &= \frac{(F + e^{kt})^{\frac{mC}{k}}}{m\int[F + e^{kt}]^{\frac{mC}{k}}dt + C_1} \end{aligned}$$

where C_1 is a constant of integration. The integration part could be done by taking mC/k = n and expanding the series. But for this the ratio m/k has to be an integer.

$$\begin{aligned}
\int [F + e^{kt}]^{\frac{mC}{k}} dt &= \int [e^{kt} + F]^n dt \\
&= \int [e^{nkt} + \binom{n}{1} F e^{(n-1)kt} + \binom{n}{2} F^2 e^{(n-2)kt} + \cdots + \binom{n}{n} F^n] dt \\
&= \frac{e^{nkt}}{nk} + \binom{n}{1} F \frac{e^{(n-1)kt}}{(n-1)k} + \binom{n}{2} F^2 \frac{e^{(n-2)kt}}{(n-2)k} + \cdots + F^n t \\
&= \frac{e^{nkt}}{nk} [1 + \frac{n^2}{(n-1)} F e^{-kt} + \frac{n^2(n-1)}{2(n-2)} F^2 e^{-2kt} + \cdots + nkF^n t e^{-nkt}]
\end{aligned}$$

Thus,

$$\begin{aligned}
i &= \frac{e^{nkt}[1 + nFe^{-kt} + \frac{n(n-1)}{2} F^2 e^{-2kt} + \cdots + F^n e^{-nkt}]}{\frac{m.e^{nkt}}{nk}[1 + \frac{n^2}{(n-1)} F e^{-kt} + \frac{n^2(n-1)}{2(n-2)} F^2 e^{-2kt} + \cdots + nkF^n t e^{-nkt}] + C_1} \\
&= \frac{C.[1 + nFe^{-kt} + \frac{n(n-1)}{2} F^2 e^{-2kt} + \cdots + F^n e^{-nkt}]}{[1 + \frac{n^2}{(n-1)} F e^{-kt} + \frac{n^2(n-1)}{2(n-2)} F^2 e^{-2kt} + \cdots + nkF^n t e^{-nkt} + \frac{nkC_1}{m} e^{-nkt}]}
\end{aligned}$$

By neglecting the higher order terms of e^{-kt} from both top and bottom expressions of $i(t)$, we obtain :

$$i(t) = \frac{C[1 + nFe^{-kt} + \frac{n(n-1)}{2} F^2 e^{-2kt}]}{[1 + \frac{n^2}{(n-1)} F e^{-kt} + \frac{n^2(n-1)}{2(n-2)} F^2 e^{-2kt}]}$$

The variation of $i(t)$ with time is given by *figure6.*

It can be seen that the infective population initially has a very low growth rate and then it starts to grow rapidly until the number of infectives reaches the number of individuals in the total population. This situation is possible as the total population was considered to have only two subgroups - susceptibles and infectives. The infection rate m and growth rate of the population k

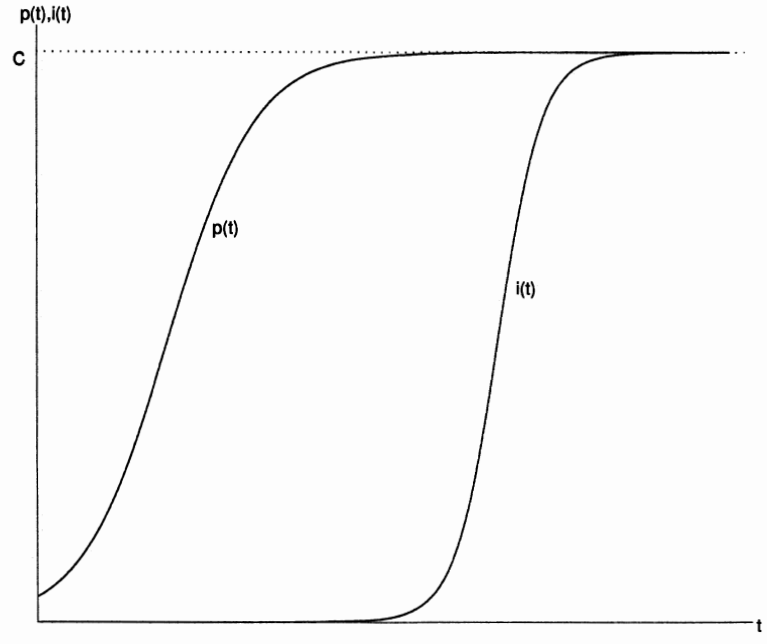

Figure 6: *The variation of infective population with time when the total population shows a logistic growth*

determines how fast the epidemic may spread over the total population. For this model to have a feasible solution mC/k must be an integer much greater than one. Hence $m > k/C$. This means that for an epidemic to spread the infection rate m must be much greater than the *apparent growth rate* k/C of the total population.

Considering the rate of growth of infective population and applying the solution of $i(t)$;

$$\frac{di(t)}{dt} = m.i(t).[\frac{C}{1+Fe^{-kt}} - i(t)]$$

$$= m.[\frac{C}{1+Fe^{-kt}} - \frac{C[1+nFe^{-kt}+\frac{n(n-1)}{2}F^2e^{-2kt}]}{[1+\frac{n^2}{(n-1)}Fe^{-kt}+\frac{n^2(n-1)}{2(n-2)}F^2e^{-2kt}]}].$$
$$\frac{C[1+nFe^{-kt}+\frac{n(n-1)}{2}F^2e^{-2kt}]}{[1+\frac{n^2}{(n-1)}Fe^{-kt}+\frac{n^2(n-1)}{2(n-2)}F^2e^{-2kt}]}$$

For an epidemic to start di/dt must be positive at some time t. If the variation of both $i(t)$ and di/dt are plotted against time t, it can be observed that the infective population starts to grow after a certain time when di/dt becomes positive. The graph of infective population has a point of inflection when di/dt has its maximum value. After that point the growth rate of $i(t)$ starts to decrease and $i(t)$ reaches the total population. This is shown in *figure 7.*

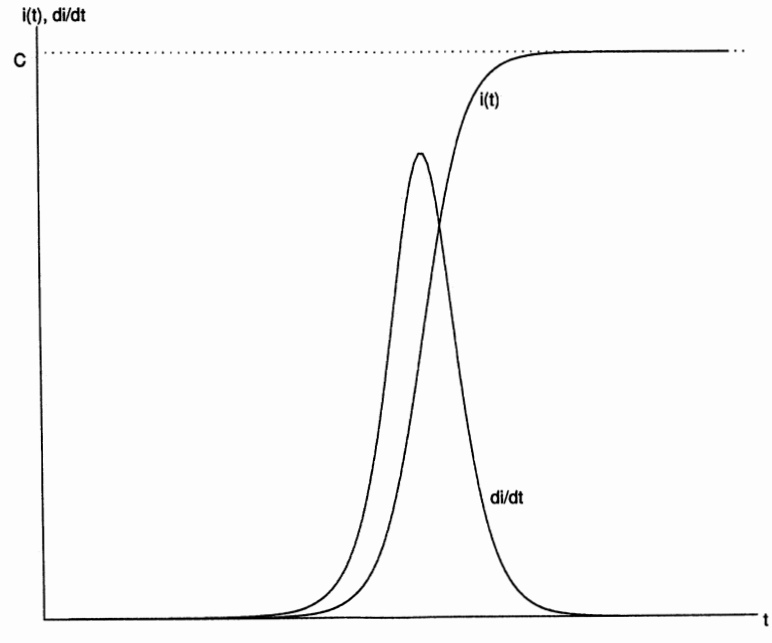

Figure 7: *The variation of $i(t)$ and di/dt with time when the total population shows a logistic growth*

8 The effectiveness of the model

When developing the model it was assumed that the population was uniformly mixing and homogeneous, which is not possible to have in the real world, especially during a period of an epidemic. So for a more realistic model the population has to be considered as nonhomogeneous and not uniformly mixing. In addition the diseases were assumed to have no incubation period so that an infective might spread the disease as soon as he gets affected with the disease. But most of the diseases have an incubation period.

In the first model, the population was assumed to be constant which is not a reasonable assumption for almost all of the diseases. The representation of total population with a logistic model as in case 3 would yield more accurate results, but the application may be restricted by the nature of the disease. The total population may not show a logistic growth if the time period involved is small.

The *encounter principle* used to develop the models (assumption that the growth rate of infectives is proportional to the product of the infectives and susceptibles) may not be applicable to all sorts of diseases. The spreading method and its rate may depend on the biological nature of the disease and other physical factors involved.

A more general improved model could be developed by incorporating the subgroup A also into the model so that most of the contagious diseases can be modeled. In addition the model must have provision for individuals of any subgroup to move from one group to another so that an individual leaving I may join S,U or A depending on the disease. As an improvement the growth of the total population must be replaced with a more suitable model.

9 A model of AIDS

AIDS [Acquired Immune Deficiency Syndrome] can be considered as the most reason outbreak of an epidemic in the world. Identified so far as incurable, AIDS has become the most concerned disease on which the largest number of medical research programs are carried out.

Although it is difficult to build a model of AIDS to depict the true nature of the disease, failure to do so might give a false impression on the severity of the epidemic. Such a model could be used to show the causes of the disease, the behavior of the epidemic and the possible steps that could be taken to intervene the spread of the disease.

9.1 Developing the model

To develop the model, lets assume that the population consists only of two groups of individuals ; *susceptibles* S and *infectives* I. At a particular time t, the number of susceptibles and infectives are given by $s(t)$ and $i(t)$ respectively. The two subpopulations *affecteds* A and *unaffecteds* do not have any members. Once an individual of S becomes a member of I, he cannot go back to S due to the nature of the disease. Therefore the total population $p(t)$ is given by,

$$p(t) = s(t) + i(t)$$

The total population is assumed to have a logistic growth given by :

$$p(t) = \frac{C}{1 + Fe^{-kt}}$$

where C, k are positive constants known as the carrying capacity and initial relative growth rate of the population respectively while F is a positive constant.

If we consider that the spreading of AIDS is due to contact between members of S and I, then the rate of growth of the infected population is proportional to the number of contacts between individuals of I and individuals of S. Assuming a uniformly mixing population, the number of contacts per unit time at time t is given by the product of the number of members in each group at time t. Therefore,

$$\begin{aligned}\frac{di(t)}{dt} &\propto s(t).i(t) \\ &= m.s(t).i(t) \\ &= m.[p(t) - i(t)].i(t) \\ &= m.i(t).[\frac{C}{1 + Fe^{-kt}} - i(t)]\end{aligned}$$

where m is a positive constant known as infection rate. But if the carrying capacity C of the total population is very much bigger than the number of AIDS patients concerned, then the total population may be assumed as a constant. If this is denotes as P ;

$$\frac{di(t)}{dt} = m.[P - i(t)].i(t)$$

This gives a logistic curve with initial relative growth rate mP and carrying capacity P.

$$i(t) = \frac{P}{1 + Ge^{-mPt}}$$

where G is a constant.

9.2 AIDS epidemic in North America

The statistics on AIDS cases in North America shows a logistic growth more similar to the model developed above. The figures from 1986 to 1995 were

The figures available for the number of AIDS infectives seem to be very low in the early years. This may cause an error in the model when the model parameters are calculated. Therefore it is difficult to find a model giving more accurate figures with the available data. The low numbers of AIDS infectives in the early years could be due to several reasons. Until early 1980's the existence of such a disease called AIDS was unknown. In addition there had been hardly any sensus done at those times to count the number of AIDS infectives.

So the model parameters were recalculated. If the data from 1984 to 1998 were used then a better model with more accurate results could be obtained. The parameter values obtained through least squares method are :

$$a = -9.775$$

$$b = 0.472$$

Therefore,

$$m = 4.9663 * 10^{-7}$$

$$G = 54.0126$$

The model with recalculated parameters is:

$$i(t) = \frac{950000}{1 + 54.0126.e^{-0.472(t-1984)}}$$

This is shown in *figure 9*. Figures used for calculations are given in *appendix2*

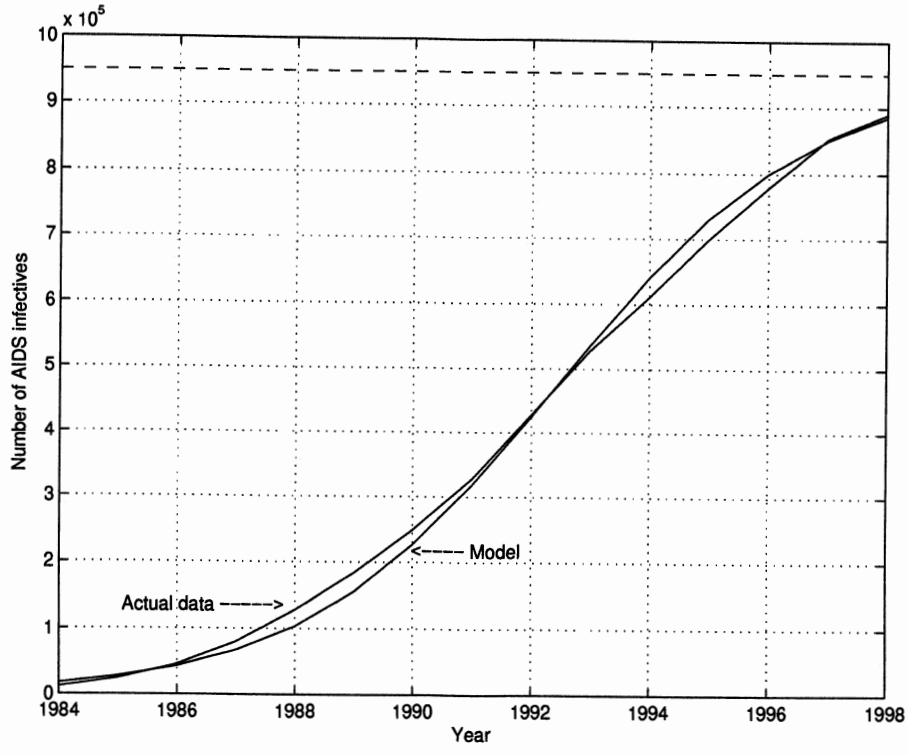

Figure 9: *The number of AIDS infectives and the the numbers given by the model (for the years 1984-1998)*

9.3 Drawbacks of the model

The model does not include the subpopulations A and U (affecteds and unaffecteds as defined on page 5). But according to the research done so far, some of the infectives develop the disease quickly while some become carriers of the disease without developing the physical manifestations of the disease for a long period of time. Therefore the subgroup A must also be included in the model.

Although it's hard to say that no one is immune to AIDS, the assump-

tion that all the individuals in the population are equally susceptible for the disease is not true. The sex patterns of individuals are not the same and the possibility an individual may involve in sex depends on his or her age. So it's unlikely that everyone will have the same probability of getting affected by AIDS. According to the surveys, the possibility of infection through homosexual behavior is higher than through heterosexual behavior. The transmission of the disease by nonsexual modes such as unsterilized needles, blood transfusion etc may not have the same growth rate as that due to sexual activities. Besides, the transmission of AIDS through nonsexual modes may not necessarily involve interaction between individuals which contradicts the assumption made when building the model.

10 Reference

1. An Introduction to Mathematical Modeling - *Edward A. Bender*
2. Empirical Model Building - *James R. Thompson*
3. Mathematical Modeling - *J.G.Andrews, R.R.Mclone*
4. Mathematics Applied to Deterministic Problems in the Natural Sciences - *C.C.Lin, L.A.Segel*
5. Mathematical Models and Applications - *Daniel P. Maki, Maynard Thompson*
6. Report by United Nations Joint Commission on AIDS
7. World Health Organization AIDS information service report
8. World Wide Web Virtual Library Information on AIDS