沈聖娥 教授指導

碩士學位 請求論文

Dynamics of Infectious Diseases: Epidemic Models and AIDS

감염질병의 역학 : 전염병 모델과 AIDS

2008

誠信女子大學校 教育大學院

教育學科 數學教育專攻

郭 다 니

Dynamics of Infectious Diseases: Epidemic Models and AIDS

감염질병의 역학 : 전염병 모델과 AIDS

沈聖娥 教授指導

이 論文을 碩士學位 論文으로 提出함

2008 年 2 月

誠信女子大學校 教育大學院
教育學科 數學教育專攻
郭 다 니

認 准 書

郭 다 니의 碩士學位 論文을 認准함

審査委員 _____印

審査委員 _____印

審査委員 _____ 印

誠信女子大學校 教育大學院

논 문 개 요

전 세계적으로 질병은 전쟁과 기아보다도 더 큰 사망원인이고, 새로운 질병의 출현과 기존의 질병의 재발현 현상들로 인해 여러 학문영역을 망라하는 학재간 연구의 필요성이 절실히 대두되고 있다.

이 연구 논문에서는 전염성 질병에 대한 수학적 모델을 세우고, 이를 이용하여 질병의 전염을 조절하고 궁극적으로는 질병을 박멸하는 방안을 제시하고자 한다.

수학적 모델의 실제에 사용될 수 있기 위해서는 실제에서 얻은 자료들이 반영되어야 한다.

하지만 얻을 수 있는 모든 자료들을 모델에 포함시키는 것이 아니라, 가장 중요하게 나타나는 요소들을 가능한 단순한 형태로 반영하게 된다.

이러한 모델들은 질병의 전염의 보이지 않는 기제를 설명하고, 질병의 전염을 조절하는 방법을 제시 할 수 있어야 한다.

목 차

I Introduction

Historically, the 14th century Black Death is the most famous epidemic. In Europe, which had a population of around 85 million at the time, about a third of the population died. One epidemic which has exercised classical scholars for a very long time is the Plague of Athens(430-428 BC) described in great detail by Thucydides including the symptoms and disease progression. He also gave some exact figures such as that 1050 of 4000 soldiers on an expedition died of the disease. The disease described so minutely by Thucydides, has been the source of numerous articles over some hundreds of years with cases being made for an incredible range of diseases such as bubonic plague, measles, Malta fever, smallpox, scarlet fever, typhus, typhoid fever and many others. One of the interesting aspects of Thucydides' account is that there is no mention of person-to-person contagion which we now accept so freely with diseases. It was only in the 19th century that it was beginning to be discussed. Evil exhalations from the earth, aerial miasmata and so on were generally accepted.

The study of epidemics with its long history has come up with an astonishing number and variety of models and explanations for the spread and cause of epidemic outbreaks. Since the end of World War II, public health strategy has focused on the elimination and control of organisms which cause disease. The advent of new antibiotics changed the whole ethos of disease control. Just over 20 years ago, in 1978, the United Nations signed the 'Health for All, 2000' accord which set the ambitious goal of the eradication of disease by the year 2000. AIDS at the time had not yet been discovered, or perhaps recognized is a better word, and in the year before, the last known case of smallpox had been treated. There was certainly cause for optimism albeit short lived. Scientists thought that microbes were biologically stationary targets and hence would not mutate in resistance to drugs and other biological influences.

This comforting image of unchanging microbes started to change shortly after this time with the emergence of microbes that could swim in a pool of bleach, grow on a bar of soap, and ignore doses of penicillin logarithmically larger than those effective in the 1950's([1]). The practical reality of bacterial

mutation is dramatically seen in New York City with tuberculosis, Control of the W-strain of the disease, which first appeared in the city in 1992, is resistant to every available drug and kills over half its victims, has already cost more than 1 billion. It was only 20 years ago that it was predicted that tuberculosis would be eradicated in the world by 2000.

Another aspect in the the current spread of disease is with the modern era of transportation allowing more than a million people a day to cross international borders, the threat of a major outbreak of exotic diseases is very real. The population explosion, especially in underdeveloped countries, is another factor in the microbes' favour. These played key roles in the proliferation of HIV(human immunodeficiency virus) in the 1980's. Recently the World Health Organization(WHO) estimated that over 30 million people worldwide are currently infected with HIV.

There are four main disease-causing microrganisms: viruses, bacteria, parasites and fungi. In this thesis, we describe a model for the population dynamics of disease agents. Such models have been commonly used to model the spread of viral, bacterial and parasitic infections but considerably less so

with fungal infections. We aim to exploit the models in the control, or ideally the eradication, of the disease or infection we are considering. The practical use of such models must rely heavily on the realism put into the models. As usual, this does not mean the inclusion of all possible effects, but rather the incorporation in the model mechanisms, in as simple a way as possible, of what appear to be the major components. However, even simple models should, and frequently do, pose important questions with regard to the underlying process and possible means of control of the disease or epidemic.

One such case study, which went through various hypothetical scenarios, is the model proposed by [2] for the 1973 cholera epidemic in the port city of Bari in southern Italy. An interesting early mathematical model, involving a nonlinear ordinary differential equation, by [3], considered the effect of cowpox inoculation on the spread of smallpox. The article has some interesting data on child mortality at the time. It is probably the first time that a mathematical model was used to assess the practical advantages of a vaccination control programme. The recent paper by [4] discusses vaccination protocols in connection with chickenpox and shingles and highlights certain dangers of

extensive vaccination. Among other things, they evaluate with their models the effects of different vaccination programmes. The classical theoretical papers on epidemic models by [5], [6], [7] have had a major influence in the development of mathematical models and are still relevant in a surprising number of epidemic situations.

In this thesis we consider the type of modes in that the total population divided into susceptible, infected and immune groups: other groupings are also possible, depending on the disease. And we only discuss deterministic models which are deficient in certain situations-eradication of a disease is one, since here the probability that the last few infected individuals will infect another susceptible is not deterministic. Nevertheless it is perhaps surprising how useful, and quantitatively predictive, deterministic models can often be; the examples in this thesis are only a very few examples where this has proven to be the case.

II Preliminaries

In this thesis we deal with a classical model with a constant population. When a small group of infected individuals is introduced into a large population, a basic problem is to describe the spread of the infection within the population as a function of time. This depends on many circumstances, including the actual disease involved, but as a first attempt at modelling directly transmitted disease we make some not unreasonable general assumptions.

Let us consider a disease which, after recovery, confers immunity which, if lethal, includes deaths: dead individuals are still counted. Suppose the disease is such that the population can be divided into three distinct classes: the susceptibles, S, who can catch the disease; the infectives, I, who have the disease and can transmit it; and the removed class, R, namely, those who have either had the disease, or are recovered, immune or isolated until

recovered. The progress of individuals is schematically represented by

$$S \to I \to R.$$

Such models are often called SIR models. The number of classes depends on the disease. SI models, for example, have only susceptible and infected classes while $SEIR$ models have a susceptible class, S, a class in which the disease is latent, E, an infectious class, I, and a recovered or dead class, R. The assumptions made about the transmission of the infection and incubation period are crucial in any model; these are reflected in the terms in the equations and the parameters.

With $S(t)$, $I(t)$ and $R(t)$ as the number of individuals in each class we assume here that:

(i) The gain in the infective class is at a rate proportional to the number of infectives and susceptibles, that is, rSI, where $r > 0$ is a constant parameter. The susceptibles are lost at the same rate.

(ii) The rate of removal of infectives to the removed class is proportional to the number of infectives, that is, aI where $a > 0$ is a constant; $\frac{1}{a}$ is a measure of the time spent in the infectious state.

(iii) The incubation period is short enough to be negligible; that is, a susceptible who contracts the disease is infective right away.

We now consider the various classes as uniformly mixed; that is, every pair of individuals has equal probability of coming into contact with one another. This is a major assumption and in many situations does not hold as in most sexually transmitted diseases(STD's). The model mechanism based on the above assumption is then

$$\frac{dS}{dt} = -rSI, \tag{1}$$

$$\frac{dI}{dt} = rSI - aI, \tag{2}$$

$$\frac{dR}{dt} = aI \tag{3}$$

where $r > 0$ is the infection rate and $a > 0$ the removal rate of infectives. This is the classic Kermack-MaKendrick(1972) model. We are only interested in

nonnegative solutions for S, I and R, because that these functions represent the populations of the corresponding classes. This is a basic model but, even so, we can make some highly relevant general comments about epidemics and, in fact, adequately describe some specific epidemics with such a model. The constant population size is built into the system (1)-(3) since, on adding the equations,

$$\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0 \qquad \Rightarrow \qquad S(t) + I(t) + R(t) = N, \tag{4}$$

where N is the total size of the population. Thus, S, T and R are all bounded above by N. The mathematical formulation of the epidemic problem is completed given initial conditions such as

$$S(0) = S_0 > 0, \qquad I(0) = I_0 > 0, \qquad R(0) = 0. \tag{5}$$

We study the mathematical model consisting of equations (1), (2), (3) and the initial condition (5) to make comparisons with actual data.

III Main Results

An important question in any epidemic situation is, given r, a, S and the initial number of infectives I, whether the infection will spread or not, and if it dose how it develops with time, and crucially when it will start to decline. In the Lemmas we show some analytic results from the mathematical model consisting of equations (1), (2), (3) and (5). In Theorem the function $R(t)$ of the removed population is obtained by solving the equations.

Lemma 1. *If* $S_0 > a/r$ *then* $I(t) > I_0$ *for some* $t > 0$, *while if* $S_0 < a/r$ *then* $I(t) < I_0$ *for all* $t > 0$.

Proof. From equation (2) we have that

$$\left[\frac{dI}{dt}\right]_{t=0} = rS_0I_0 - aI_0 = I_0(rS_0 - a). \tag{6}$$

Thus $\left[\frac{dI}{dt}\right]_{t=0} > 0$ if $S_0 > \rho$, and $\left[\frac{dI}{dt}\right]_{t=0} < 0$ if $S_0 < \rho$, where $\rho = \frac{a}{r}$.

Since we have that $\frac{dS}{dt} \leq 0$ from (1), if $S_0 < \frac{a}{r}$ then $S \leq S_0$ for all $t \geq 0$.

Therefore, in the case that $S_0 < \rho$ it is derived that

$$\frac{dI}{dt} = I(rS - a) \leq 0 \qquad \text{for all } t \geq 0, \tag{7}$$

and thus $I_0 > I(t) \to 0$ as $t \to \infty$ and so the infection dies out; that is, no epidemic can occur.

On the other hand if $S_0 > \frac{a}{r}$ then $\frac{dI}{dt} > 0$ for small $t > 0$. Hence $I(t)$ initially increases and thus $I(t) > I_0$ for some $t > 0$. In this case we have an epidemic. $\square$

Lemma 2. *The (I,S) phase plane trajectory of the epidemic system* (1)-(3) *with the initial condition* (5) *is given by the equation*

$$I + S - \rho \ln S = constant = I_0 + S_0 - \rho \ln S_0. \tag{8}$$

Proof. From (1) and (2)

$$\frac{dI}{dS} = -\frac{(rS - a)I}{rSI} = -1 + \frac{\rho}{S}, \qquad \rho = \frac{a}{r}, \qquad (I \neq 0).$$

The singularities all lie on the $I = 0$ axis because that the second fraction above would have zero denominator there.

Integrating the last equation gives

$$I(S) = -S + \rho \ln S + C.$$

Since $I(S_0) = I_0$, we have that

$$I_0 = -S_0 + \rho \ln S_0 + C,$$

and so

$$C = I_0 + S_0 - \rho \ln S_0.$$

Hence the (I,S) phase plane trajectory of the epidemic system (1)-(3) with the initial condition (5) is given by the equation

$$I + S - \rho \ln S = C = I_0 + S_0 - \rho \ln S_0.$$

□

The phase trajectories are sketched in Figure 1. We note that with (5), all initial values S_0 and I_0 satisfy $I_0 + S_0 = N$ since $R(0) = 0$ and so $0 \leq S + I < N$ for all $t > 0$. As Figure 1 illustrates we have a threshold phenomenon. If $S_0 > S_c = \frac{a}{r}$ there is an epidemic while if $S_0 < S_c$ there is

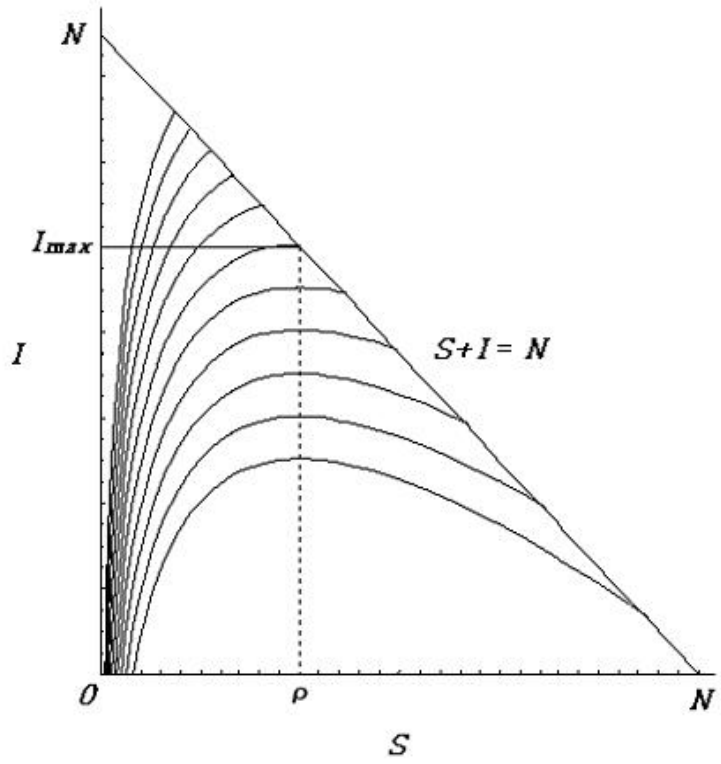

Figure 1: Phase trajectories in the susceptibles (S)-infectives (I) phase plane for the SIR model epidemic system (1)-(3). The curves are determined by the initial condition $I(0) = I_0$ and $S(0) = S_0$.

not. The critical parameter $\rho = \frac{a}{r}$ is sometimes called the *relative removal rate* and its reciprocal $\sigma(= \frac{a}{r})$ the *infection's contact rate*. We write

$$R_0 = \frac{rS_0}{a},$$

where R_0 is the *basic reproduction rate* of the infection, that is, the number of secondary infections produced by one primary infection in a wholly susceptible population. If $R_0 > 1$, clearly an epidemic ensues.

Lemma 3. *The maximum of I, I_{max}, and the total number of susceptibles who can catch the disease in the course of the epidemic, I_{total} are as in the following.*

$$\begin{aligned} I_{max} &= N - \rho + \rho \ln(\tfrac{\rho}{S_0}), \\ I_{total} &= I_0 + S_0 - S(\infty), \end{aligned}$$

where $S(\infty)$ is the positive root $0 < z < \rho$ of the transcendental equation

$$S_0 \exp\left[-\frac{N-z}{\rho}\right] = z.$$

Proof. From (7) the maximum I, I_{max}, occurs at $S = \rho$ where $\frac{dI}{dt} = 0$. From

(8), with $S = \rho$,

$$\begin{aligned} I_{max} &= \rho \ln \rho - \rho + I_0 + S_0 - \rho \ln S_0 \\ &= I_0 + (S_0 - \rho) + \rho \ln(\tfrac{\rho}{S_0}) \\ &= N - \rho + \rho \ln(\tfrac{\rho}{S_0}). \end{aligned}$$

Since the axis $I = 0$ is a line of singularities, on all trajectories $I \to 0$ as $t \to \infty$. From (1), S decreases since $\frac{dS}{dt} < 0$ for $S \neq 0$, $I \neq 0$. From (1) and (3),

$$\frac{dS}{dR} = -\frac{S}{\rho} \quad \Rightarrow \quad S = S_0\, e^{-\frac{R}{\rho}} \;\geq\; S_0\, e^{-\frac{N}{\rho}} > 0 \quad \Rightarrow \quad 0 < S(\infty) \leq N. \qquad (9)$$

In fact from Figure 1, $0 < S(\infty) < \rho$. Since $I(\infty) = 0$, (4) implies that $R(\infty) = N - S(\infty)$. Thus, from (9),

$$S(\infty) = S_0 \exp\left[-\frac{R(\infty)}{\rho}\right] \;=\; S_0 \exp\left[-\frac{N - S(\infty)}{\rho}\right]$$

and so $S(\infty)$ is the positive root $0 < z < \rho$ of the transcendental equation

$$S_0 \exp\left[-\frac{N - z}{\rho}\right] = z. \qquad (10)$$

We then get the total number of susceptibles who catch the disease in the course of the epidemic as

$$I_{total} = I_0 + S_0 - S(\infty), \qquad (11)$$

where $S(\infty)$ is the positive solution z of (10). $\square$

An important implication of the result in Lemma 3, namely, that $I(t) \to 0$ and $S(t) \to S(\infty) > 0$, is that the disease dies out from a lack infectives and not from a lack of susceptibles.

In most epidemics it is difficult to determine how many new infectives there are each day since only those that are removed, for medical aid or whatever, can be counted. Public Health records generally give the number of removed per day, week or month. So, to apply the model to actual epidemic situations, in general we need to know the number removed per unit time, namely, $\frac{dR}{dt}$, as a function of time.

From (10), (4) and (3) we get an equation for R alone; namely,

$$\frac{dR}{dt} = aI = a(N - R - S) = a(N - R - S_0 e^{-\frac{R}{\rho}}), \qquad R(0) = 0, \qquad (12)$$

which can only be solved analytically in a parametric way: the solution in this form however is not particularly convenient. Of course, if we know a, r,

S_0 and N it is a simple matter to compute the solution numerically. Usually we do not know all the parameters and so we have to carry out a best fit procedure assuming, of course, the epidemic is reasonably described by such a model. In practice, however, it is often the case that if the epidemic is not large, $\frac{R}{\rho}$ is small - at least $\frac{R}{\rho} < 1$. Following [5] we can then approximate (12) by

$$\frac{dR}{dt} = a[N - S_0 + (\frac{S_0}{\rho} - 1)R - \frac{S_0 R^2}{2\rho^2}], \qquad R(0) = 0 \tag{13}$$

In Theorem 4 under the condition that $\frac{R}{\rho} < 1$ we solve the equations (13) and find an approximation of the solution $R(t)$ for the equations (1), (2) and (3).

Theorem 4. *If $\frac{R}{\rho} < 1$, then the function $R(t)$ below is an approximation of the solution $R(t)$ of the system of equations* (1), (2) *and* (3).

$$R(t) = \frac{\rho(S_0 - \rho)}{S_0} + d \tanh\left[\frac{dS_0(at + C)}{2\rho^2}\right] \tag{14}$$

where $d = \left[\rho^2\left(\frac{(S_0-\rho)^2}{S_0^2} + \frac{2(N-S_0)}{S_0}\right)\right]^{\frac{1}{2}}$, *and* $C = \frac{2\rho^2}{dS_0}\tanh^{-1}\left[\frac{\rho(\rho-S_0)}{dS_0}\right]$.

Proof. We may rewrite equation (13) by separating the variables R and t as

in the following ;

$$\frac{dR}{N - S_0 + (\frac{S_0}{\rho} - 1)R - \frac{S_0 R^2}{2\rho^2}} \;=\; a\, dt,$$

and integrating both sides we have that

$$\int \frac{dR}{N - S_0 + (\frac{S_0}{\rho} - 1)R - \frac{S_0 R^2}{2\rho^2}} \;=\; \int a\, dt \;=\; at + C$$

$$= \; -\frac{2\rho^2}{S_0} \int \frac{dR}{R^2 - \frac{2\rho^2}{S_0}(\frac{S_0}{\rho} - 1)R - \frac{2\rho^2}{S_0}(N - S_0)}$$

$$= \; -\frac{2\rho^2}{S_0} \int \frac{dR}{R^2 - \frac{2\rho(S_0 - \rho)}{S_0} R - \frac{2\rho^2 (N - S_0)}{S_0}}$$

$$= \; -\frac{2\rho^2}{S_0} \int \frac{dR}{R - \frac{\rho(S_0 - \rho)}{S_0}^{\;2} - \rho^2 \;\; \frac{(S_0 - \rho)^2}{S_0^2} + \frac{2(N - S_0)}{S_0}} \;.$$

Now for simplicity of writing let us denote that $c \;=\; \frac{\rho(S_0 - \rho)}{S_0}$ and $d^2 \;=\; \rho^2 \left(\frac{(S_0 - \rho)^2}{S_0^2} + \frac{2(N - S_0)}{S_0} \right)$, and evaluate the integral ;

$$\int \frac{dR}{(R - c)^2 - d^2} \;=\; -\frac{1}{d} \tanh^{-1} \left[\frac{R - c}{d} \right]$$

Thus, we have that

$$at + C \;=\; \frac{2\rho^2}{dS_0} \tanh^{-1} \left[\frac{R - c}{d} \right],$$

and so

$$R(t) \;=\; c + d \tanh \left[\frac{dS_0(at + C)}{2\rho^2} \right] \;=\; \frac{\rho(S_0 - \rho)}{S_0} + d \tanh \left[\frac{dS_0(at + C)}{2\rho^2} \right].$$

Since $R(0) = 0$, we have that

$$0 \;=\; c + d\,\tanh\left[\frac{dS_0C}{2\rho^2}\right],$$

and thus

$$C \;=\; \frac{2\rho^2}{dS_0}\tanh^{-1}\left[-\frac{c}{d}\right] \;=\; \frac{2\rho^2}{dS_0}\tanh^{-1}\left[\frac{\rho(\rho - S_0)}{dS_0}\right].$$

□

From the result in Theorem 4 the removal rate is then given by

$$\frac{dR}{dt} \;=\; \frac{ad^2S_0}{2\rho^2}\,\mathrm{sech}^2\left[\frac{dS_0(at + C)}{2\rho^2}\right] \;=\; \frac{a\alpha^2\rho^2}{2S_0}\,\mathrm{sech}^2\left[\frac{\alpha at}{2} - \phi)\right], \qquad (15)$$

where $\alpha = \frac{dS_0}{\rho^2}$, and $\phi = -\frac{\alpha C}{2}$. Thus we obtained dR/dt in a form that involves only 3 parameters, namely, $\frac{a\alpha^2\phi^2}{2S_0}$, α and ϕ. With epidemics which are not large, it is this function of time which should fit to the public health records. On the other hand, if the disease is such that we know the actual number of the removed class then it is $R(t)$ in (14) we should use. If $\frac{R}{\rho}$ is not small, however, we must use the differential equation (12) to determine $R(t)$.

IV Practical Applications

We now present how the population model (1)-(3) is applied to two very different epidemic situations. In each situations we try to find out the parameters for the functions $S(t)$, $I(t)$, and $R(t)$.

Bombay plague Epidemic 1905-1906

This plague epidemic lasted for almost a year. Since most of the victims who got the disease died, the number removed per week, this is, $\frac{dR}{dt}$, was approximately equal to the number of deaths per week. On the basis that the epidemic was not severe(relative to the population size), [5] compared the actual data with (15) and determined the best fit for the three parameter which resulted in

$$\frac{dR}{dt} = 890\,\mathrm{sech}^2(0.2t - 3.4). \tag{16}$$

This is illustrated in Figure 2 together with the actual epidemic data.

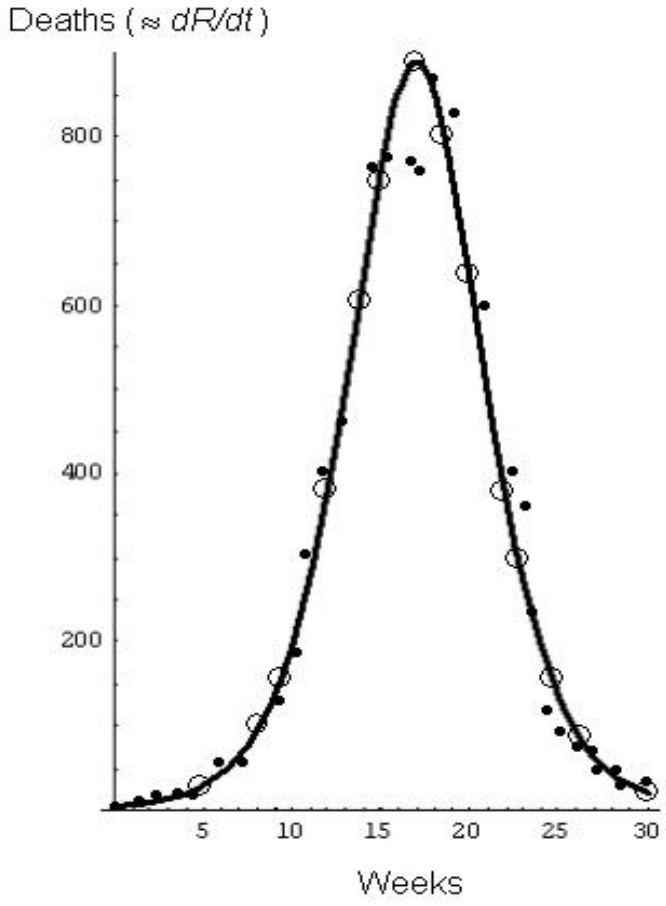

Figure 2: Bombay plague Epidemic 1905-1906. Comparison between the data (•) and theory (○) from equation (16).

Influenza Epidemic in an English Boarding School 1978

In 1978 in British medical journal, The Lancet, there was a report with detailed statistics of a flu epidemic in a boys' boarding school with a total of 763 boys. Of these 512 were confined to bed during the epidemic, which lasted from 22nd January to 4th February 1978. It seems that one infected boy initiated the epidemic. This situation has many of the requirements assumed in the above model derivation. Here, however, the epidemic was severe and the full system has to be used. Also, when a boy was infected he was put to bed and so we have $I(t)$ directly from the data. Since in this case we have no analytical solution for comparison with the data, a best fit numerical technique was used directly on the equation (1) - (3) for comparison of the data. Figure 3 illustrates the resulting time evolution for the infectives, $I(t)$, together with the epidemic statistics. The R-equation (3) is uncoupled; the solution for $R(t)$ is simply proportional to the area under the $I(t)$ curve.

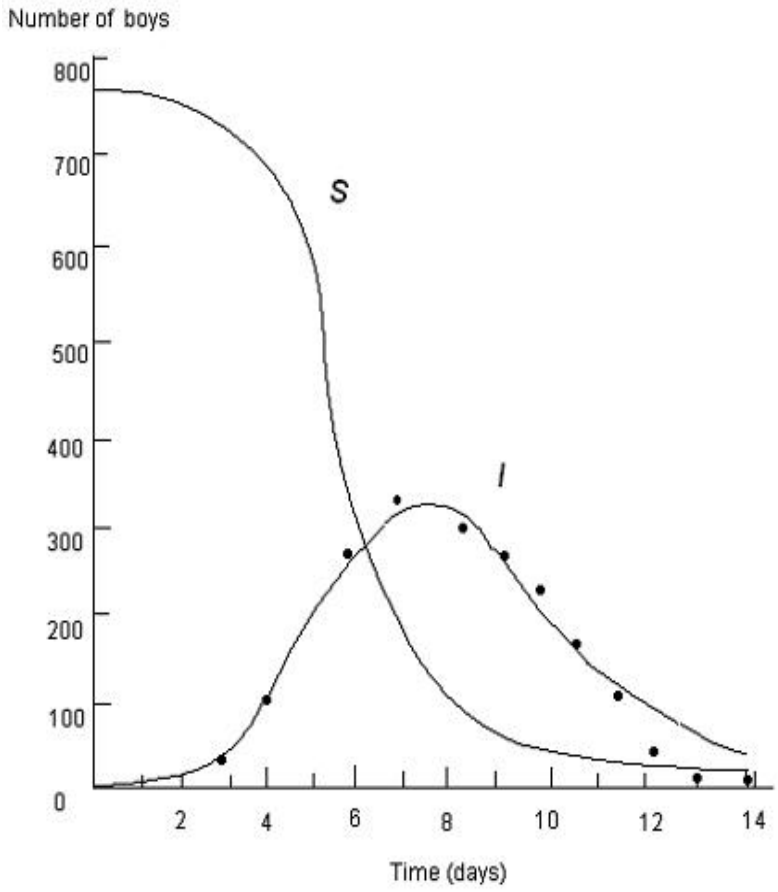

Figure 3: Influenza epidemic data (•) for a boy's boarding school as reported in the British medical journal, *The Lancet*, 4th March 1978.

References

[1] L.Garrett. The return of infectious disease. *Foreign Affairs* , 75:66-79,1996.

[2] V. Capasso and S.L. paveri-Fontana. A mathematical model for the 1973 cholera epidemic in the European Mediterranean region.*Rav. Epidem. et Sante Publ.*, 27:121-132, 1979.

[3] D.Bernoulli.Essai d'une nouvelle analyse de la mortailte causee par la petite verole, et des avantages de l'inoculation pour la prevenir. *Histoire de l'Acad. Roy. Sci. (Paris) avec Mem. des Math. et Phys. and Mem*, pages 1-45, 1760.

[4] M.C. Schuette and H.W. Hethcote. Modelling the effects of varicella vaccination programs on the incidence of chickenpox and shingles. *Bull.*

math. Biol, 61:1031-1064, 1999.

[5] W.O. Kermack and A.G. Mckendrick. Contributions to the mathematical theory of epidemics. *Proc. R. Soc. Lond.* A, 115:700-721, 1927.

[6] W.O. Kermack and A.G. Mckendrick. Contributions to the mathematical theory of epidemics. *Proc. R. Soc. Lond.* A, 138:55-83, 1932.

[7] W.O. Kermack and A.G. Mckendrick. Contributions to the mathematical theory of epidemics. *Proc. R. Soc. Lond.* A, 141:94-122, 1933.

ABSTRACT

Title of Paper

Kwak, Da Ni

Major in Mathematics Education

Graduate School of Education

Sungshin Women's University

Supervised by Seong-A Shim, Ph. D.

In the world epidemics is more the cause of death than the war or the hunger. Due to the appearance of a new epidemics and the revelation of the existing epidemics it became necessary to study epidemics over a large range of different fields of researches. In this thesis, we formulate mathematical a model of epidemics, and make use of this to analyze infectious diseases and ultimately build a plan to eradicate diseases. Mathematics models should reflect data of an actual situation, explain the invisible process of infections, and give a method to control the spread of diseases.