

ARTICLE

HIGHLIGHTS

This article provides a brief account of fuzzy set theory applied to explain epidemiological problems

ABSTRACT

The mathematical study of epidemics is essential for a better perception of their evolution and control potential. The typical deterministic mathematical models do not provide subjective modeling of these phenomena. The application of fuzzy set theory in order to model epidemiological problems and overcome this issue is suggested as an efficient predictive procedure for the epidemiology of infectious diseases. Although this is a recent research field, authors objective is to provide a review of the state of the art of fuzzy logic theory implementation in epidemiology. Findings suggest that fuzzy sets implementation in epidemiology is a really promising field of research and specify the future stages of fuzzy sets application in epidemiology.



An initiative by

PRECISION-MEDICINE.IN

www.precision-medicine.in

KEY WORDS

Fuzzy set theory, epidemic, epidemiology, modeling

Received 29 March 2016; Revised 21 April 2016; Accepted 29 April 2016; Published 8 May 2016

Antonia Plerou^{1*}, Elena Vlamou², Basil Papadopoulos³

¹Bioinformatics and Human Electrophysiology Laboratory, Department of Informatics Ionian University, Corfu, GREECE

²Department of Civil Engineering Democritus University of Thrace Xanthi, GREECE

³Department of Civil Engineering, Democritus University of Thrace, Xanthi, GREECE

*Corresponding author: tplerou@ionio.gr, papadob@civil.duth.gr



INTRODUCTION

The basic aspect of medicine is intervention. However, in order to intervene, one must predict the natural course of a system in the absence of intervention and to predict the evolution of these systems after the proposed intervention. Mathematical epidemiology is very recent and it is blending mathematics and medicine science. Traditional expert systems based on binary logic have been used successfully as diagnostic decision tools through the use of probability theory and information sequential processing are responsible these method limitations in several medical applications. Despite fuzzy logic potential in dealing with uncertainties, limited applications of fuzzy logic concepts in epidemiological problems have been obtained so far. The development of numerous mathematical methods and concepts lead to practical applications of fuzzy logic.

Epidemics in numbers

Epidemics of infectious diseases afflict millions of people worldwide. Namely, over than 40 million people throughout the world are infected with human immunodeficiency virus (HIV) in accordance with to the World Health Organization, statistics in 2001. Additionally, more than 300 million people were infected with sexually transmitted diseases other than HIV every year worldwide each year in accordance with World Health Organization in 1996. Almost 8 million people get sick and 2 million people are dying because of tuberculosis every year according to World Health Organization data in 2000. The optimal way to allocate prevention funds, in the case that the resources are unlimited, is to afford money in order to reduce the disease transmission. However, in the case that resources are constrained the challenge is to minimize the growth of an epidemic allocating properly the limited prevention funds (Brandeau, Zaric, & Richter, 2003).

Disease diagnosis

Disease diagnosis is related to several levels of imprecision and uncertainty, notably in epidemiological studies. A single disease may occur quite differently in different patients and with different disease status. Furthermore, a single symptom may suggest several diseases, and the occurrence of different diseases in a single patient may unsettle the typically expected symptom pattern. This is possible to cause a formidable amount of fuzziness and vagueness in the interpretation of effect measures of covariates of interest. A typical general aspect of the disease is that health and disease are opposed and that they are dual and contradictory attributes. Innovative fuzzy logic methods consider health and disease as supplementary situations. Uncertainty in epidemiology is not limited to random variations. Therefore, alternative approaches from other domains of fuzzy logic are used, like linguistic models, fuzzy decision making, fuzzy clustering, possibility distributions, to deal with some of the epidemiology and public health problems (Massad, Ortega, Struchiner, & Burattini, 2003).

Fuzzy logic

Fuzzy logic is a superset of conventional (Boolean) logic, developed in order to access the concept of partial truth (Massad et al., 2003) Fuzzy logic is related to an inference morphology that allows approximate human reasoning abilities to be applied to knowledge-based systems. The fuzzy logic is an approach to computing based on “degrees of truth” instead of the classic “true or false” (1 or 0) Boolean logic. Natural language is hard to be assigned into the absolute terms of 0 and 1. Fuzzy logic comprises 0 and 1 as extreme cases of truth but also contains the various states of truth in between so that (Kawatra, 2006) Fuzzy logic is applied to several fields like control theory and artificial intelligence. Namely, a fuzzy system advantage is 1. The capacity to represent inherent uncertainties of the human knowledge with linguistic variables; 2. A simple interaction with the expert of the domain with the engineering designer of the system; 3. An easy interpretation of the results, because of the natural rules representation; 4. An easy extension of the base of knowledge by adding new rules; 5. Robustness in a relation of the possible disturbances in the system (Vieira, Dias, & Mota, 2004).

Fuzzy inference systems

Fuzzy inference systems (FIS) are extensively used for process simulation or control. They can be designed either from expert knowledge or from data. The fuzzy inference mechanism consists of three stages. In the first stage, the values of the numerical inputs are mapped with the use of a function with accordance to a degree of compatibility of the respective fuzzy sets and this operation is named fuzzification. In the second stage, the fuzzy system processes the rules in accordance with the firing strengths of the inputs. In the third stage, the subsequent fuzzy values are altered again into numerical values. This operation is named defuzzification. Essentially, this procedure enables the use of fuzzy categories in the representation of words and abstracts ideas of the humans in the description of the decision-making procedure (Viharos & Kis, 2014).



Mathematical modeling

Mathematical models are always subject to inaccuracies related to the nature of the variables and parameters involved. In these models, the estimation of the parameters is usually based on statistical methods (Regina Siqueira Ortega, Cesar Sallum, & Massad, 2000). Mathematical modeling of infectious diseases is a tool to investigate the mechanisms for the occurrence and spread of diseases and foresee the future direction in order to control an epidemic. The earliest mathematical epidemic model is by Daniel Bernoulli. The results showed that the universal vaccination could raise the life expectancy. McKendrick and Kermack in 1927 formulated a simple deterministic model and successfully predict the behavior of outbreaks in many recorded epidemics using mathematical epidemic modeling (Anderson, 1991).

METHODS

Stochastic epidemic models

The epidemic process has random nature. Stochastic models are used for estimating the probabilistic quantities for the outcome events, like the probability distribution of extinction time or of final outbreak size and the associated mean. A stochastic model has a random variable and is a method in order to estimate probability distributions of possible outcomes. Stochastic models depend on the chance variations in risk of exposure, disease, and several illness dynamics.

The SI Model

Analysis of epidemic models is a vital research topic. The simplest classical model to describe the directly transmitted diseases with the interaction between susceptible and infected individuals is the SI model without neither vital dynamics (the rates of birth and mortality are not included), nor immunity, nor additional disease fatality rate. The model can be represented by the diagram to follow:

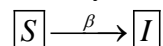


Figure 1: SI model diagram where the flow between the susceptible and infective compartments are explicit.

The classical normalized differential equations which describe such dynamics are given by:

$$\frac{dS}{dt} = -\beta SI \quad - \quad \frac{dI}{dt} = \beta SI$$

where $S + I = 1$, S is the proportion of impressionable individuals, I is the proportion of infected individuals at each instant and is the transmission coefficient of the disease. The basic assumption in this formula is that the population is homogeneous. Namely, every infected individual transmits the disease with the same chance, given by the real number β . So, the number of infected individuals at any instant t is given by:

$$I = \frac{I_0 e^{\beta t}}{S_0 + I_0 e^{\beta t}}$$

where S_0 and I_0 is the initial conditions. Both concepts of impressionable and infectious are unclear in the sense that there are different degrees in susceptibility and infectivity among the individuals of the population. Such differences can arise, for example, when we consider the population's distinct habits and customs, different degrees of resistance, etc. Thus, more realistic models consider different degrees of impressionability and/or infectivity of the individuals. The parameter β (presenting the chance that in one contact between an impressionable and an infected individual the transmission of the disease occurs) is considered as a fuzzy number (Massad et al., 2003).

The SI fuzzy model

The population heterogeneity is assumed to be given by the parasite load of infected individuals. Thus, the higher the parasite load, the higher is the possibility of disease transmission. Namely, it is hypothesized that $\beta = \beta(v)$ measures the chance of a transmission to occur in a meeting between a susceptible and an infected individual with an amount of pathogens v . So, several values of β are more likely comparing to others and that change β into a membership function of a fuzzy number. To obtain the membership function β it is considered that while a number of pathogens in an individual is rather low, the possibility of transmission is insignificant. Consequently, there is a minimum amount of pathogens v_{min} needed to cause disease transmission. Furthermore, for a specific amount of pathogens v_M , the chance of disease transmission is maximum and equal to 1. Finally, it is assumed that the individual's amount of pathogens is always limited by v_{max} for every disease. For the fuzzy subset, the following membership function is defined:



$$\beta(v) = \begin{cases} 0, & \text{if } v < v_{\min} \\ \frac{v - v_{\min}}{v_M - v_{\min}}, & \text{if } v_M < v < v_{\max} \\ 1, & \text{if } v_{\min} \leq v \leq v_M \end{cases}$$

where v_{\min} represents the minimum amount of pathogens needed for disease transmission to occur. This value is perceived as the one that provides the susceptibility of a particular population. Actually, as higher the v_{\min} value, the higher the amount of pathogens needed for transmission to occur. Namely, the populations have a low susceptibility to the disease (Massad et al., 2003) (Barros, Leite, & Bassanezi, 2003).

The SIS model

The simplest model to describe a disease in which the individual recovers but does not develop any kind of immunity, i.e. becomes susceptible again, is presented.

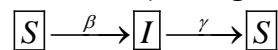


Figure 2: SIS model diagram showing the flow between susceptible and infective compartments.

It is considered that the flow of an individual from S class to I class occurs at a rate β depending only on the contact of a susceptible with an infected individual and that the individual recovers at a rate γ , returning to the susceptible condition. The dynamical system is described by the differential equations system that follows:

$$\frac{dS}{dt} = -\beta SI + \gamma I \quad \frac{dI}{dt} = \beta SI - \gamma I$$

where $S + I = 1$, S is the proportion of impressionable individuals, I is the proportion of infected individuals, β is the contact rate and γ the recovering rate. Consequently, γ^{-1} is the average period of infectiousness (Massad et al., 2003).

The SIS fuzzy model

As in the SI fuzzy model, $\beta = \beta(v)$ and the individuals' recovery rate (γ) is also a function of the parasite load v is assumed. The higher the parasite load, the longer it will take to recover from infection. Consequently, γ should be a decreasing function of v :

$$\gamma(v) = \frac{(\gamma_0 - 1)}{v_{\max}} v + 1$$

where $\gamma_0 > 0$ is the lowest recovery rate. So:

$$\frac{dI}{dt} = \beta I \left[\left(1 - \frac{\beta}{\gamma} \right) - 1 \right]$$

Thus, the equilibrium solution I^* is obtained

$$I^* = \frac{\beta - \gamma}{\beta + \left[\frac{\beta - \gamma}{I_0 - \beta} \right] e^{-(\beta - \gamma)t}}$$

And this occurs when $\beta \geq \gamma$. From the hypothesis of our model, β and γ depend on the parasite load v . In this way, the number of infected people, at each instant of time, is given by

$$I(v, t) = \frac{\beta(v) - \gamma(v)}{\beta(v) + \left[\frac{\beta(v) - \gamma(v)}{I_0} - \beta(v) \right] e^{-[\beta(v) - \gamma(v)]t}}$$

So, as in the SI fuzzy model, the average number of infected individuals is provided from FEV $[I(v, t)]$ or $E [I(v, t)]$. Thus, the objective is to evaluate the stability of the disease. To study the temporal evolution of the number of infected people, i.e., whether the number of infected increases indefinitely or not, the stability of the equilibrium points should be evaluated.



Making $\frac{dS}{dt} = 0$ and $\frac{dI}{dt} = 0$ the equilibrium points $P_1 = (1, 0)$ and $P_2 = \left(\frac{\gamma}{\beta}, 1 - \frac{\gamma}{\beta}\right)$ for the system are reserved.

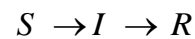
Deterministic epidemic models

The transition rate from one class to the other one is featured by derivative mathematically. Assuming that the population size is differentiable with respect to time, in the limiting of a large population, the time evolution of behavior of each subgroup can be approximated by the deterministic dynamics.

SIR Model

In 1927 Kermack and McKendrick created a model in which they considered a fixed population of three compartments only: susceptible; $S(t)$ infected, $I(t)$; and removed, $R(t)$. The compartments used for this model comprises of three classes: $S(t)$ is representing the number of individuals not yet infected with the disease at time t , or those susceptible to the disease. Additionally, $I(t)$ denotes the number of individuals who have been infected with the disease and is possible to spread the disease to those in the susceptible category. Finally, $R(t)$ is the compartment used for individuals who have been infected but detached of the disease, either due to immunization or to passing away. These individuals unable to be infected again or to transmit the infection to others.

The flow of this model may be considered as follows:



With the use of a fixed population, $N = S(t) + I(t) + R(t)$, Kermack and McKendrick result in the following equations:

$$\frac{dS}{dt} = -\frac{\beta SI}{N} \quad \frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I \quad \frac{dR}{dt} = \gamma I$$

for the standard stochastic SIR epidemic model.

Similarly, with the deterministic common epidemic model a closed homogeneous uniformly mixing community is assumed and n denote the size of the community. The $S(t)$, $I(t)$ and $R(t)$ denote the number of susceptible, infectives and recovered at time t , respectively, and suppose that at time $t = 0$ these numbers are given by $S(0) = n - m$, $I(0) = m$ and $R(0) = 0$. The dynamics of the model are defined as follows: Infectious individuals have "close contact" with other individuals rarely in time at continuous rate λ , and each such contact is with a randomly selected individual, all contacts of different infective being defined to be mutually independent. The term "close contact" denotes a contact close enough to result in infection whether the other individual is susceptible, otherwise, the contact has no effect. Any susceptible receiving such a contact is instantly infected and begin spreading the disease in accordance with the equal rules. Infected individuals continue to be infectious for a random time I (the infectious period) and thereafter they pause of being infectious, recover and become immune to the disease. The infectious periods are defined to be independent and identically distributed and the have distribution F_I and mean $E(I) = 1/\gamma$. The epidemic starts at time $t = 0$. As the epidemic evolves, in accordance with the above-mentioned rules, it is possible new individuals to get infected and finally recover, until the first time T when there is no infective in the population. Then no more individuals could get infected and that implies that the epidemic ends.

The final state of the epidemic is described by the ultimate number $R(T)$ infected (note that $I(T) = 0$, so $S(T) = n - R(T)$ make up the rest of the community). The final number of infected $R(T)$ consist of m who were initially infected plus Z , who were infected during the outbreak (Britton, 2010).

Fuzzy decision making in epidemiology

Making decisions is essential in public health where decisions frequently are relevant for millions of people. In the field of vaccination strategy design, decision making concerning the target population for the immunization program, the proportion of impressionable to be vaccinated, the optimal age to immunize children and the nature of the strategy, e.g. selective or indiscriminate, are cases of the variables to be optimized, subject to a set of constraints.

The objective of decision making is the study of the way decisions are made in order to become more effective and essential. Models of human decision making, in general, comprise the aggregation or constraint criteria. In the case that this cannot be modeled crisply, a decision could be defined as the intersection of fuzzy sets representing either objectives or constraints. The degree of membership of an



object in the intersection of two fuzzy sets, namely, the “fuzzy set decision” is determined with the minimum or the product operator procedure. Decision-making under risk conditions has been modeled by probabilistic and game theories, nevertheless fuzzy decision theories challenge is to deal with the imprecision of human preferences.

Fuzzy probabilities of epidemic events

The majority of significant application of fuzzy logic in epidemiology are related with the field of linguistic models. These methods are additionally used for a large amount of uncertainties characteristic of epidemic problems and to solve problems that probability theory fails to solve efficiently. Fuzzy logic and probability theory, although analogous in particular in perspectives, were designed to deal different tasks. The values of probability measure have been classically defined as a number between 0 and 1 that preserves the additive property. Fuzzy probability, in turn, is a generalization of interval probability in which the probability value is bounded by a fuzzy set.

CONCLUSION

Fuzzy logic concepts are applied to population biology with an emphasis on epidemiological problems like causal studies, epidemic models, and designing of vaccination strategies. The objective of the present paper is to provide a review on the present state of fuzzy logic applications in epidemiology. Alike to several biomedical fields of fuzzy logic applications, the key advantage of fuzzy methods as applied to epidemiology is the simulation of human cognitive processes. Biomedical applications of fuzzy logic methods in spite of their potential in problem-solving haven't received adequate attention and, therefore, additional research need to be done in the field of fuzzy epidemic control (Massad et al., 2003).

CONFLICT OF INTEREST

The author declares no competing interests.

ACKNOWLEDGEMENT

None.

FINANCIAL DISCLOSURE

None

REFERENCES

- [1] Anderson, R. (1991). The Kermack-McKendrick epidemic threshold theorem. *Bulletin of Mathematical Biology*, 53(1-2), 3–32. [http://doi.org/10.1016/S0092-8240\(05\)80039-4](http://doi.org/10.1016/S0092-8240(05)80039-4)
- [2] Barros, L. C. D., Leite, M. B. F., & Bassanezi, R. . (2003). The SI epidemiological models with a fuzzy transmission parameter. *Computers & Mathematics with Applications*, 45(10-11), 1619–1628. [http://doi.org/10.1016/S0898-1221\(03\)00141-X](http://doi.org/10.1016/S0898-1221(03)00141-X)
- [3] Brandeau, M. L., Zaric, G. S., & Richter, A. (2003). Resource allocation for control of infectious diseases in multiple independent populations: beyond cost-effectiveness analysis. *Journal of Health Economics*, 22(4), 575–98. [http://doi.org/10.1016/S0167-6296\(03\)00043-2](http://doi.org/10.1016/S0167-6296(03)00043-2)
- [4] Britton, T. (2010). Stochastic epidemic models: a survey. *Mathematical Biosciences*, 225(1), 24–35. <http://doi.org/10.1016/j.mbs.2010.01.006>
- [5] Costa, A. J. L., Massad, E., Ortega, N. R. S., & Araújo, A. de Q.-C. (2004). Perception of disability in a public health perspective: a model based on fuzzy logic. *International Journal of Medical Informatics*, 73(7-8), 647–56. <http://doi.org/10.1016/j.ijmedinf.2004.04.010>
- [6] Kawatra, S. K. (2006). *Advances in Communitation*. SME.
- [7] Li, T.-J., & Zhang, W.-X. (2008). Rough fuzzy approximations on two universes of discourse. *Information Sciences*, 178(3), 892–906. <http://doi.org/10.1016/j.ins.2007.09.006>
- [8] Massad, E., Ortega, N. R. S., de Barros, L. C., & Struchiner, C. J. (2008). *Fuzzy Logic in Action: Applications in Epidemiology and Beyond* (Vol. 232). Berlin, Heidelberg: Springer Berlin Heidelberg. <http://doi.org/10.1007/978-3-540-69094-8>
- [9] Massad, E., Ortega, N. R. S., Struchiner, C. J., & Burattini, M. N. (2003). Fuzzy epidemics. *Artificial Intelligence in Medicine*, 29(3), 241–259. [http://doi.org/10.1016/S0933-3657\(02\)00070-2](http://doi.org/10.1016/S0933-3657(02)00070-2)
- [10] Nedjah, N., & Mourelle, L. de M. (2005). *Fuzzy Systems Engineering: Theory and Practice*. Springer Science & Business Media.
- [11] Regina Siqueira Ortega, N., Cesar Sallum, P., & Massad, E. (2000). Fuzzy dynamical systems in epidemic modelling. *Kybernetes*, 29(2), 201–218. <http://doi.org/10.1108/03684920010312768>
- [12] Vieira, J., Dias, F., & Mota, A. (2004). Neuro-fuzzy systems: a survey. In *5th WSEAS NNA International Conference on Neural Networks and Applications*, Udine, Italia.
- [13] Viharos, Z. J., & Kis, K. B. (2014). *Fuzzy Systems and their Applications in Technical Diagnostics*. In *Workshop on Technical Diagnostics Advanced measurement tools in technical diagnostics for systems' reliability and safety*. Warsaw, Poland.

