ENTROPY FOR DTMC SIS EPIDEMIC MODEL

TAYEBE WAEZIZADEH* AND FARZAD FATEHI
DEPARTMENT OF PURE MATHEMATICS, FACULTY OF MATHEMATICS
AND COMPUTER AND MAHANI MATHEMATICAL RESEARCH CENTER,
SHAHID BAHONAR UNIVERSITY OF KERMAN, KERMAN, IRAN
DEPARTMENT OF MATHEMATICS, SCHOOL OF MATHEMATICAL AND
PHYSICAL SCIENCES, UNIVERSITY OF SUSSEX, BRIGHTON, UK
E-MAILS: WAEZIZADEH@UK.AC.IR & F.FATEHI@SUSSEX.AC.UK

(Received: 17 January 2017, Accepted: 3 February 2017)

ABSTRACT. In this paper at first, a history of mathematical models is given. Next, some basic information about random variables, stochastic processes and Markov chains is introduced. As follows, the entropy for a discrete time Markov process is mentioned. After that, the entropy for SIS stochastic models is computed, and it is proved that an epidemic will be disappeared after a long time.

AMS Classification: 92D30, 60G20, 94A17, 37A50. Keywords: Epidemic Model, Entropy, Markov chain, Stochastic process.

1. Introduction

The first mathematical epidemic model was introduced by Bernouli in 1760. The basic continuous epidemic model was studied by Kermak and Mckendrick in 1927 [2, 8]. In next years many deterministic models for different infectious diseases were established [2, 8]. The aim was to investigate the dynamical behavior of the model and determine how many infected will get if the epidemic takes off [2].

However, deterministic models are not applicable for studying a community with a small number of susceptible or infected individuals. In this case, it is better to use stochastic epidemic models [5, 6, 3].

In this paper, we consider SIS stochastic epidemic model. In this model, susceptible

^{*} CORRESPONDING AUTHOR

SPECIAL ISSUE FOR SELECTED PAPERS OF CONFERENCE ON DYNALMICAL SYSTEMS AND GEOMETRIC THEORIES, 11-12 DECEMBER 2016, MAHANI MATHEMATICAL RESEARCH CENTER, SHAHID BAHONAR UNIVERSITY OF KERMAN JOURNAL OF MAHANI MATHEMATICAL RESEARCH CENTER VOL. 5, NUMBERS 1-2 (2016) 59-67.

 $^{@\}mathit{MAHANI}\ \mathit{MATHEMATICAL}\ \mathit{RESEARCH}\ \mathit{CENTER}$

individuals (S) can get infected (I) and after a while, an infected individual recovers and becomes susceptible immediately. We assume that the birth, death, immigration or emigration rate during the study period are equal, i.e. the population is closed. Moreover, we consider there is no vertical transmission of the disease, that is the disease is not passed from the mother to her offspring.

One of the most important characterizations that one can attach to a random variable and a stochastic process is its entropy. The entropy was appeared in physics in 1865. It was extended in information theory as the measure of the uncertainty in a random variable.

In recent years, the entropy has been developed for random transformations [4, 7]. The entropy of stochastic epidemic models was introduced in [1]. In this paper, we compute the entropy of SIS DTMC model which is mentioned in [5].

In the next section, random processes and homogeneous Markov chains are reviewed. And some of their properties is investigated. In section 3, the entropy of a discrete random variable and its meaning in the discrete time stochastic processes has been investigated. In section 4, SIS stochastic model is introduced and the entropy of the model is computed. In the last section, by using numerical simulation provided in an example, it is shown that entropy changes with β change.

2. Basic notions

Let (Ω, β, μ) be a probability space and (S, A) be a measurable space which is called the state space. An S-valued random variable is a measurable function from Ω to S. An S-valued stochastic process is a collection of S-valued random variables on Ω , indexed by a totally ordered set T. $t \in T$ shows time. That is a stochastic process X is a collection $\{X_t : t \in T\}$, where X_t is an S- valued random variable on Ω .

A stochastic process called continuous time if $T=\mathbb{R}$ or $T=[0,\infty)$, and discrete time if $T=\mathbb{Z}$ or $T=\mathbb{N}$. It is called a one-sided stochastic process if $T=\mathbb{N}$ [4]. In this paper, since the population is constant, the stochastic process is discrete time and finite state, where $S=\{0,1,2,...,N\}$ and N is the population size. The joint probability of the discrete random variables $X_0,X_1,...,X_n$ is defined as follows:

$$\mu(\{\omega \in \Omega: X_0(\omega) = x_0,...,X_n(\omega) = x_n\}) = \text{Prob}\{X_0 = x_0,X_1 = x_1,...,X_n = x_n\} = P(x_0,x_1,...,x_n).$$

Definition 2.1. A stochastic process $X = \{X_n\}_{n \in \mathbb{N}_0}$ is called a Markov process or Markov chain if $Prob\{X_n = x_n \mid X_{n-1} = x_{n-1}, ..., X_0 = x_0\} = Prob\{X_n = x_n \mid X_{n-1} = x_{n-1}\}, \ n \geq 1, \ where \ x_0, x_1, ..., x_n \in S = \{0, 1, 2, ..., N\}.$

The probability mass function associated with the random variable X_n which is denoted by $\{p_i(n)\}_{i=0}^N$, where

$$p_i(n) = \text{Prob}\{X_n = i\}.$$

Definition 2.2. The one-step transition probability or only the transition probability is noted as $p_{ji}(n)$, is defined as $p_{ji}(n) = Prob\{X_{n+1} = j \mid X_n = i\}$.

If the transition probabilities $p_{ji}(n)$ in a Markov chain do not depend on time n, then the Markov process X is called time homogeneous or homogeneous. Let $X = \{X_n\}_{n \in \mathbb{N}_0}$ be a homogeneous Markov chain, the matrix $P = (p_{ji})$ is called the transition matrix of X.

Definition 2.3. Let $X = \{X_n\}_{n \in \mathbb{N}_0}$ be a homogeneous Markov chain, the n-step transition probability which is denoted by $p_{ii}^{(n)}$, is defined by

(2.1)
$$p_{ji}^{(n)} = Prob\{X_n = j \mid X_0 = i\}.$$

The n-step transition matrix is denoted as $P^{(n)} = (p_{ji}^{(n)}), P^{(0)} := I_n \text{ and } P^{(1)} = P.$

Based on the Chapman-Kolmogorov equation [4], $P^{(n)} = P^{(n-s)}P^{(s)}$. Since $P^{(1)} = P$ then $P^{(n)} = P^n$, for all $n \in \mathbb{N}$.

Let p(n) denotes the vector of probability mass function associated with X_n , that is $p(n) = (p_0(n), ..., p_N(N))^T$, and $\sum_{i=0}^{N} p_i(n) = 1$, then the vector of probability associated with X_{n+1} can be founded by multiplying the transition matrix P by p(n), that is

$$(2.2) p(n+1) = Pp(n).$$

In general, $p(n) = P^n p(0)$, so we have $p(X_n = i) = p_i(n) = \sum_{i=0}^{N} p_{ii}^{(n)} p_j(0)$.

3. The entropy of a discrete random variable

Definition 3.1. Let X be a random variable with probability space (Ω, β, μ) and finite state S. The entropy of X is defined by

(3.1)
$$H(X) = -\sum_{x \in S} p(x) log p(x),$$

where
$$p(x) = \mu(\{\omega \in \Omega : X(\omega) = x\})$$
.

Let X, Y be two random variables on a probability space (Ω, β, μ) and state space S. The joint entropy of X and Y is defined as

(3.2)
$$H(X,Y) = -\sum_{x \in S_1} \sum_{y \in S_2} p(x,y) \log p(x,y),$$

where p(x, y) = Prob(X = x, Y = y).

The conditional entropy of Y given X is defined as follows,

(3.3)
$$H(Y \mid X) = -\sum_{x \in S_1} \sum_{y \in S_2} p(x, y) log p(y \mid x).$$

Where $p(y \mid x) = \frac{p(x,y)}{p(x)}$. It's clear that $H(X,Y) = H(X) + H(Y \mid X)$ [4, 7]. The (joint) entropy of the random variable vector $X_0^{n-1} = \{X_0, ..., X_{n-1}\}$ is defined

$$H(X_0,X_1,...,X_{n-1}) = -\sum_{x_0,x_1,...,x_{n-1} \in S} p(x_0,x_1,...,x_{n-1}) log p(x_0,x_1,...,x_{n-1}),$$

where

$$p(x_0, x_1, ..., x_{n-1}) = \text{prob}(X_0 = x_0, ..., X_{n-1} = x_{n-1}).$$

Theorem 3.2. Let the random variables $X_0, ..., X_{n-1}$ be given, then $i)p(X_0, ..., X_{n-1}) = \prod_{i=0}^{n-1} p(X_i \mid X_{i-1}, ..., X_0).$ $ii)H(X_0, ..., X_{n-1}) = \sum_{i=0}^{n-1} H(X_i \mid X_{i-1}, ..., X_0),$ where $p(X_0 \mid X_{-1}) := p(X_0)$ and $H(X_0 \mid X_{-1}) := H(X_0).$ iii) If the stochastic process is a homogeneous Markov process, then $H(X_i | X_{i-1}, ..., X_0) = H(X_i | X_{i-1}).$ iv) If the stochastic process is a homogeneous process, then $H(X_0, X_1, ..., X_n) = H(X_0) + H(X_1 | X_0) + ... + H(X_n | X_{n-1}).$ Proof. See [4, 1].

3.1. The entropy of a discrete time stochastic process.

Definition 3.3. The entropy of a stochastic process $X = \{X_n\}_{n=0}^{\infty}$ on a probability space (Ω, β, μ) with finite state S is defined by

(3.4)
$$h(X) = \lim_{n \to \infty} \frac{1}{n} H(X_0, ..., X_{n-1}),$$

provided that the limit exists [4].

The term $h(X_0, ..., X_{n-1}) = \frac{1}{n} H(X_0, ..., X_{n-1})$ is called the entropy of order n of X. $h(X_0^{n-1}) = h(X_0, ..., X_{n-1})$ is the average uncertainty about n consecutive outcomes of the random experiment modeled by X. On the other hand, if $X = \{X_n\}_{n=0}^{\infty}$ is a discrete- time stochastic process, then after n-time intervals Δt there are n random variables as outcomes. $h(X_0^{n-1})$ measures the average uncertainty and determines how close this model is to reality. The more $h(X_0^{n-1})$ is smaller, the more our model $(X = \{X_n\}_{n=0}^{\infty})$ is closer to the real.

As follows we introduce SIS DTMC (discrete time Markov chain) epidemic model and compute its entropy.

4. SIS DTMC, EPIDEMIC MODEL

Let S(t) and I(t) be discrete random variables for the number of susceptible and infected individuals at time t, respectively. Let $t \in \{0, \Delta t, 2\Delta t, ...\}$ and N be the population size which is constant. $S(t), I(t) \in \{0, 1, ..., N\}$ and S(t) + I(t) = N, so S(t) = N - I(t) [5, 2]. The stochastic process $\{I(t)\}_{t=0}^{\infty}$ has an associated probability function

$$(4.1) p_i(t) = prob\{I(t) = i\},$$

for i = 0, 1, ..., N and $t = 0, \Delta t, 2\Delta t, ...,$ that $\sum_{i=0}^{N} p_i(t) = 1$.

Let $p(t) = (p_0(t), ..., p_N(t))^T$ denotes the probability vector associated with I(t). For $t = n\Delta t$, I(t) and $p_i(n\Delta t)$ are denoted by I_n and $p_i(n)$, respectively.

Lemma 4.1. The sequence $I = \{I_n\}_{n=0}^{\infty}$ is a Markov chain.

Proof. We know that I_n is the random variable of infected individuals at time $t = n\Delta t$. Since the number of infected people at time $t = (n+1)\Delta t$ only dependes on the number of infected people at time $t = n\Delta t$. So it is clear that $Prob(I_{n+1})$ $I_n, I_{n-1}, ..., I_0$ = Prob $(I_{n+1} | I_n)$ so $I = \{I_n\}_{n=1}^{\infty}$ is a Markov chain.

The probability of a transition from state $I_n = i$ to state $I_{n+1} = j$, $i \longrightarrow j$, at time Δt , is denoted as

$$p_{ii}(n, \Delta t) = Prob\{I_{n+1} = j \mid I_n = i\}.$$

The transition probability does not depend on n, so the process is time homogeneous. The time step Δt is chosen sufficiently small such that the number of infected individuals changes by at most one during the interval Δt , on the other hand either there is at most a new infection, a birth, a death, or a recovery during the time interval Δt .

The probability of a new infection, $i \to i+1$, is $\beta i(N-i)\Delta t$, where β is the infection transition rate and the probability of a death or a recovery, $i \to i-1$, is $(b+\gamma)i\Delta t$, where b and γ are death and recovery rates, respectively. Moreover, Δt is chosen sufficiently small such that $\beta i(N-i)\Delta t + (b+\gamma)i\Delta t < 1$ for all $0 \le i \le N$. Finally, the probability of no change in states is $1 - [\beta i(N-i)\Delta t + (b+\gamma)i\Delta t]$. So we have

$$p_{ji}(\Delta t) = \begin{cases} \beta i(N-i)\Delta t & j=i+1\\ (b+\gamma)i\Delta t & j=i-1\\ 1-[\beta i(N-i)\Delta t+(b+\gamma)i\Delta t] & j=i\\ 0 & otherwise \end{cases}$$

$$(4.2) \quad p_i(n+1) = p_{i-1}b(i-1)\Delta t + p_{i+1}(n)d(i+1)\Delta t + p_i(n)[1-b(i)-d(i)]\Delta t.$$

Denote the transition matrix as $P(\Delta t)$. It is a $(N+1)\times(N+1)$ matrix which is given by

where $A_i = 1 - [b(i) + d(i)]\Delta t$ for i = 1, ..., N - 1.

According to the 4.2, it is obvious that

$$p(n+1) = P(\Delta t)p(n) = P(\Delta t)^{n+1}p(0),$$

 $p(n+1)-1 \leftarrow P(x),$ and $p_{ji}^{(n)}=((P(\Delta t))^n)_{ji}$. From now on, we will be using P instead of $P(\Delta t)$. We consider that there is i_0 infected at the initial time (t=0), so $p(0)=(0,...,\underbrace{1}_{i_0^{th}place},0,...,0)$

and $p(n) = P^n p(0)$, then

$$p_i(n) = ((p^n)_{i_0}, (p^n)_{i_1}, ..., (p^n)_{i_N})p(0)^T = (P^n)_{ii_0} = P_{ii_0}^{(n)}$$

4.1. The entropy of SIS epidemic model. The entropy of a stochastic process $X = \{X_n\}_{n=0}^{\infty}$ on a probability space (Ω, β, μ) is defined by the following formula

(4.3)
$$h(X) = \lim_{n \to \infty} \frac{1}{n} H(X_0, ..., X_{n-1}),$$

provided that the limit exists [4].

Definition 4.2. If $X = \{X_n\}_{n=0}^{\infty}$ is a homogeneous Markov chain, the matrix H(P) is defined by

(4.4)
$$H(P) = (-p_{ij}logp_{ij})_{(N+1)\times(N+1)},$$

and 0log 0 := 0.

Remark 4.3. Let $X = \{X_n\}_{n=0}^{\infty}$ be a homogeneous Markov chain, then $H(X_n|X_{n-1}) = -\sum_{j=0}^{N} \sum_{i=0}^{N} Prob(X_{n-1} = j, X_n = i)logProb(X_n = i|X_{n-1} = j) = -\sum_{j=0}^{N} \sum_{i=0}^{N} Prob(X_{n-1} = j)Prob(X_n = i|X_{n-1} = j)logp_{ij} = -\sum_{i=0}^{N} \sum_{j=0}^{N} Prob(X_{n-1} = j)p_{ji}logp_{ji}.$

Theorem 4.4. Let $I = \{I_n\}_{n=0}^{\infty}$ be the stochastic process of the SIS epidemic model with constant population and $I_0 = k$, then $h(I_0, ..., I_n)$ is the summation of k^{th} column entries of matrix $H_n = H(p)(\frac{0+I+P+...+P^{n-1}}{n+1})$.

Proof. Using part (iv) in theorem 4.2, $h(I_0,...,I_n) = \frac{1}{n+1} H(I_0,...,I_n) = \frac{H(I_0) + H(I_1|I_0) + ... + H(I_n|I_{n-1})}{n+1}. \text{ Since } I_0 = k, \text{ clearly } H(I_0) = 0, \text{ so } h(I_0,...,I_n) = \frac{H(I_1|I_0) + ... + H(I_n|I_{n-1})}{n+1}.$ Also by the above remark, $H(I_n|I_{n-1}) = \sum_j \sum_i \operatorname{Prob}(I_{n-1} = j) p_{ij} \log \frac{1}{p_{ij}}.$ By the

assumption $I_0 = k$, we have $Prob(I_{n-1} = j) = p_{jk}^{(n-1)}$, so

$$\begin{split} h(I_0,...,I_n) &= \sum_j \sum_i \frac{Prob(I_0 = j) + ... + Prob(I_{n-1} = j)}{n+1} p_{ij} log \frac{1}{p_{ij}} \\ &= \sum_{j=0}^N \sum_{i=0}^N \frac{p_{jk}^{(0)} + ... + p_{jk}^{(n-1)}}{n+1} p_{ij} log \frac{1}{p_{ij}} \\ &= \sum_{j=0}^N \sum_{i=0}^N \frac{p_{jk}^{(0)} + ... + p_{jk}^{(n-1)}}{n+1} (H(p))_{ij} \\ &= \sum_{j=0}^N \sum_{i=0}^N \frac{I_{jk} + P_{jk} + ... + (P^{n-1})_{jk}}{n+1} (H(p))_{ij} \\ &= \sum_{i=0}^N \frac{(H(p))_{ik} + (H(p)P)_{ik} + ... + (H(p)P^{n-1})_{ik}}{n+1} \\ &= \sum_{i=0}^N (\frac{H(p) + H(p)P + ... + H(p)P^{n-1}}{n+1})_{ik}. \end{split}$$

So $h(I_0,...,I_n)$ is the summation of k^{th} column entries of the matrix

$$\frac{H(p) + H(p)P + \ldots + H(p)P^{n-1}}{n+1} = H(p)\frac{0 + I + P + \ldots + P^{n-1}}{n+1}.$$

The matrix $H_n = H(p)(\frac{0+I+P+\ldots+P^{n-1}}{n+1})$ is called the entropy matrix of I. In the SIS DTMC epidemic model, the state i=0 is recurrent and the states i = 1, 2, ..., N are transient. We know for any state j and any transient state i, $\lim_{n\to\infty} P_{ij}^n = 0$ [1]. Therefore,

In the next theorem, we prove that an epidemic with constant population which is modelled by SIS, will disappear after a long time.

Theorem 4.5. h(I) = 0.

Proof. By the pervious assumptions, since the limit of P^n as $n \to \infty$ exists, then according to Cesaro's mean theorem, if $\lim_{n\to\infty}a_n=a$ and $b_n=\frac{\sum_{k=0}^n a_k}{n+1}$, then $\lim_{n\to\infty}b_n=a$. Therefore, we have $\lim_{n\to\infty}H(P)(\frac{0+I+\ldots+P^{n-1}}{n+1})=\lim_{n\to\infty}H(P)P^n$

$$\lim_{n \to \infty} H(P)(\frac{0+I+\dots+P^{n-1}}{n+1}) = \lim_{n \to \infty} H(P)P^n$$

We can say as time increases, this model becomes closer to the real. This model predicts that after a long time, the disease will disappear, so this prediction is true, because h(I) = 0.

In the next section, we compute $h(I_0^n)$ which shows the uncertainty about n consecutive outcomes of a random experiment.

5. Numerical simulation

The entropy is used to compute the uncertainty in dynamical systems. In this paper, SIS DTMC epidemic model is introduced. As follows, in the next example the entropy at different time steps n for different values of β is computed. Figure 1 shows how the entropy of the model is changed with respect to the time.

Example 5.1. Consider the population size N=10, $\Delta t=0.01$, b=0.25, $\gamma=0.25$, I(0)=2 and s(0)=8. The resulted table for different values of β is

| $n \setminus \beta$ | 0.01 | 0.02 | 0.03 | 0.04 |
|---------------------|--------|--------|--------|--------|
| 5 | 0.0296 | 0.0380 | 0.0475 | 0.058 |
| 10 | 0.0167 | 0.0226 | 0.0299 | 0.0391 |
| 15 | 0.0113 | 0.0155 | 0.0210 | 0.0286 |
| 20 | 0.0085 | 0.0117 | 0.0160 | 0.0228 |
| 20 | 0.0068 | 0.0155 | 0.0130 | 0.0208 |
| 30 | 0.0057 | 0.0094 | 0.0110 | 0.0249 |
| 40 | 0.0043 | 0.0059 | 0.008 | 0.0111 |
| 50 | 0.0034 | 0.0047 | 0.0064 | 0.0088 |
| 60 | 0.0028 | 0.0039 | 0.0054 | 0.0074 |
| 100 | 0.0017 | 0.0023 | 0.0032 | 0.0044 |
| 200 | 0.0008 | 0.0012 | 0.0016 | 0.0022 |

Fig.1 indicates that by increasing β , the entropy generally will increase as well.

Entropy with different transition rates

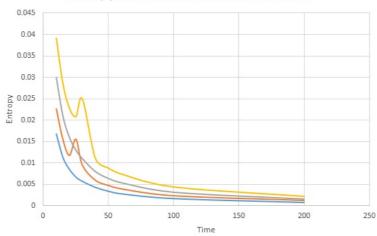


FIGURE 1. blue: $\beta=0.01,$ orange: $\beta=0.02,$ violet: $\beta=0.03$ and yellow: $\beta=0.04.$

By theorem 4.5, the entropy is zero after a long time. As one can consider, the entropy has allow amount for all values of time. Figure 1 also shows that by increasing β , the entropy will increase. Thus, the smaller β is the more accurate the prediction will be.

REFERENCES

- [1] Farzad Fatehi, Tayebe Waezizadeh, Entropy for the stochastic mathematical models of epidemics, Preprint.
- [2] Fred Brauer, Peter Van den Driessche, Jiahong Wu J, Mathematical Epidemiology, Springer-Verlag, Berlin Heidelberg, 2008.
- [3] Glenn Lahodny Jr E, Linda Allen, Probability of a disease outbreak in stochastic multi-patch epidemic models, Bulltein of Mathematical Biology, Number 75, 1157-1180 (2013).
- [4] Joze Mario Amigo, Permutation Complexity in Dynamical Systems-Ordinal Paterns, Permutation Entropy and All that, Springer-Verlag, Berlin, 2010.
- [5] Linda Allen, An Introduction to Stochastic Processes with Applications to Biology, 2th edition, Chapman Hall/CRC Press, Boca Raton, FL, 2010.
- [6] Linda Allen, Edward Allen, Comparison of three different stochastic population models with regard to persistence time, Theoretical Population Biology, Number 64, 439-449 (2003).
- [7] Peter Walter, An Introduction to Ergodic Theory, Springer-Verlag, New York, 2000.
- [8] Zhien Ma, Jia Li, Dynamical Modeling and Analysis of Epidemics, World Scientific publishing Co. Pte. Ltd, 2009.