New model describing the dynamical behaviour of penetration rates

Tohru Tashiro, Hiroe Minagawa and Michiko Chiba

E-mail: tashiro@cosmos.phys.ocha.ac.jp
Department of Physics, Ochanomizu University, 2-1-1 Ohtsuka, Bunkyo, Tokyo 112-8610, Japan

Abstract. We propose a hierarchical logistic equation as a model to describe the dynamical behaviour of a penetration rate of a prevalent stuff. In this model, a memory, how many people who already possess it a person who does not process it yet met, is considered, which does not exist in the logistic model. As an application, we apply this model to iPod sales data, and find that this model can approximate the data much better than the logistic equation.

1. Introduction
How does fashion diffuse in our society? Fashion spreads out, although we, members of society, do not aim to do so: Most of us are not out to spread it, while some may have that aim. New stuffs that are somewhat of a curiosity at the beginning will become commonplace before we notice. Some of them may disappear from our life style. This phenomenon is essentially similar to various changes of phase in matter which we cannot imagine from an interaction between atoms or molecules. That is to say, the human being is “the social atom”[1].

We want to clarify the mechanism producing the occurrence of fashion. In order to deal with this as a scientific problem, quantitative data indicating the extent of diffusing are necessary. Here, we shall employ a penetration rate. This has the following universality: Generally, penetration rates increase slowly at the beginning and, then, the growth reaches its maximum. Finally, the rates become saturated. This change in time is called the S-shaped curve. The logistic function, solution of the logistic equation, has often been used to analyze the rates.

As is well known, the logistic equation was proposed as an equation describing a population growth with an upper limit by Verhulst [2]. Griliches made the first adoption of the logistic equation for the dynamical behaviour of innovation diffusion [3]. After that, the logistic function has been utilized for dynamical behaviour of penetration rates for various stuffs.

However, what kind of human communication results in such a dynamical behavior? It is not a self-evident question. In this work, we shall unveil it first of all. According to our study, it is clear that the logistic equation applying to penetration rates supposes the following human communication: those who do not have a prevalent stuff start to possess it shortly after they meet people already possessing it.

Hence, a new question arises: Are we influenced by others so easily?, which is the real start-line of this paper. Therefore, we constructed a brand new model supposing more natural human communication, which implies that we extend the logistic equation. Moreover, we adopt the
that is
\[ t \]
Here, we take the limits as \( \Delta t \) and one non-adopter can share a lattice. The third rule means that only one adopter probability at each step whose time interval is \( \Delta t \) number is not part with it. Such a group can be realized by considering people existing on lattices whose

3. Hierarchical logistic equation

It is clarified in the previous section that the dynamical behavior of the penetration rate depicted by the logistic function is based on imitation: People not possessing a prevalent stuff yet go buying it shortly after they meet people already possessing it. Here, most of us reach the same question: But are we really like that? In that rule, the definitive human psychology, when we imitate, is forgotten. That is, the memory, how many adopters we have met, is essential for us to start to possess the stuff.

In order to integrate this feature, hereby, we shall extend the rules of the group of random walker by the following way: We set the number of people starting to process the stuff after they meet \( \mu \) adopters at \( i \) step as \( Q_i^\mu \), in which we call \( \mu \) as remaining adopters number (RAN). Indeed, if a non-adopter, whose RAN is \( \mu \), meet one of adopters, his/her RAN becomes \( \mu - 1 \) at the next step. We do not alter other rules. Namely, we do not consider interactions between non-adopters despite the fact that the non-adopter gets more varied.

2. Penetration rate in an imitating group

In this section, we unveil the human communication yielding the penetration rate which can be described by the logistic function.

Let us consider a group composed of \( N \) people. For this group, we shall apply the following rules: i) At the beginning, some people have a stuff which will diffuse in this society. ii) If those who do not possess the stuff yet (non-adopters) meet people who already possessing it (adopters), they start to adopt it at once. iii) A non-adopter is not influenced by more than one adopter and an adopter can not influence more than one non-adopter at the same time. iv) Adopters do not part with it. Such a group can be realized by considering people existing on lattices whose number is \( n \times n \) \( (< N) \). We suppose that he/she moves to one of the next lattices with the same probability at each step whose time interval is \( \Delta t \). The third rule means that only one adopter and one non-adopter can share a lattice.

Here, we respectively set the number of adopters and non-adopters at \( i \)th step as \( P_i \) and \( Q_i \). Therefore, it is natural to consider that a probability to meet adopters or non-adopters is proportional to each number of them: The probability at \( i \)th step to meet adopters and non-adopters is \( P_i/n^2 \) and \( Q_i/n^2 \), respectively. Indeed, a probability to meet nobody at each step is \( (n^2 - P_i - Q_i)/n^2 = 1 - N/n^2 \). Therefore, \( (P_i/n^2) \times Q_i \) people of non-adopters become adopters at the next step, so that we can obtain the following recursion formulae:

\[
P_{i+1} = P_i + \frac{P_i}{n^2} Q_i, \quad Q_{i+1} = Q_i - \frac{P_i}{n^2} Q_i.
\]

We shall define the number of them at \( t \) as \( P(t) = P(i \cdot \Delta t) \equiv P_i \) and \( Q(t) = Q(i \cdot \Delta t) \equiv Q_i \). Here, we take the limits as \( \Delta t \to 0 \) and \( n \to \infty \) with \( n^2 \Delta t \) fixed. By setting the fixed value as \( a/N \) and using \( P(t) + Q(t) = N \), the following differential equation is derived:

\[
\frac{dP(t)}{dt} = \frac{a}{N} \{N - P(t)\} P(t),
\]
that is the exact logistic equation. The penetration rate \( p(t) \equiv P(t)/N \) satisfies the following logistic equation

\[
\frac{dp(t)}{dt} = a \{1 - p(t)\} p(t).
\]
If the maximum of RAN is \( m \), the recursion formulae exchange into
\[
P_{i+1} = P_i + \frac{P_i}{n^2} Q_1^i, \quad Q_1^i = Q_i^1 - \frac{P_i}{n^2} Q_1^1 + \frac{P_i}{n^2} Q_2^1, \quad \ldots, \quad Q_{i+1}^m = Q_i^m - \frac{P_i}{n^2} Q_i^m.
\]

By the previous continuation of time and space, we can obtain the following differential equations:
\[
\frac{dP(t)}{dt} = \frac{a}{N} Q_1(t) P(t), \quad \frac{dQ_1(t)}{dt} = -\frac{a}{N} P(t) Q_1(t) + \frac{a}{N} P(t) Q_2(t), \quad \ldots, \quad \frac{dQ_m(t)}{dt} = -\frac{a}{N} P(t) Q_m(t)
\]
where \( Q^m(t) = Q^m(i \cdot \Delta t) \equiv Q_i^m \).

We shall call this the hierarchical logistic equation. Indeed, \( P(t) + Q^1(t) + \ldots + Q^m(t) = N \) is always conserved.

We can solve Eq. (5) easily: The solution is \( Q^m(t) = Q^m(0) \exp \left[-a \int_0^t dt' P(t')\right] \). If \( Q^m(0) = 0 \), \( Q^m(t) \) is always 0. Then, the contribution of \( Q^m \) into the differential equation of \( Q^{m-1} \) disappears, and so \( Q^{m-1} \) can be calculated similarly. Therefore, if \( Q^m(0) = Q^{m-1}(0) = \ldots = Q^2(0) = 0, Q^m(t) = Q^{m-1}(t) = \ldots = Q^2(t) = 0 \), which means the normal logistic equation is recovered. Namely, the hierarchical logistic equation includes the normal one.

If we use ratios of adopters and non-adopters to the total number \( N \), the differential equations become
\[
\frac{dp(t)}{dt} = a q^1(t) p(t), \quad \frac{dq^1(t)}{dt} = -ap(t) q^1(t) + ap(t) q^2(t), \quad \ldots, \quad \frac{dq^m(t)}{dt} = -ap(t) q^m(t),
\]
where \( q^m(t) \equiv Q^m(t)/N \).

4. Fitting the iPod sales data

Now, let us apply the hierarchical logistic equation to fitting a real data. As this data, we shall employ the iPod, created and marketed by Apple Inc., sales which can be obtained from the official Website, http://www.apple.com/. We shall use the data from November on 2001 to May on 2006, which includes only one peak. We treat \( N \) as a fitting parameter, because the number of sales is not saturated and so we cannot obtain \( N \) from the data.

Setting November on 2001 as the origin of time, we construct the cumulative sales and, then, we fit the data with the hierarchical logistic equation. If we minimize the residual sum of squares or the sum of the absolute value of error (SAE) when fitting this data, the solution of the hierarchical logistic equation does not match the data with small values.

Therefore, we shall minimize the product of SAE and the sum of the absolute value of relative error (SARE). The results are show in Tab. 1. The parameters minimize the product. SARE diminishes with increase of \( m \): SARE with \( m = 4 \) reduces to nearly half that with \( m = 1 \). In other words, the average relative error for the hierarchical logistic model with \( m = 4 \), 5.3\%, is about as half as that for the logistic model, 9.7\%.

From the parameters with \( m = 4 \), we can find the following facts: the market size producing the first peak is about 66 million people; the ratio of the trend-conscious people, \( q^1(0) \), is about 33\%; the ratio of the cautious people, \( q^2(0) \), about 66\% and the ratio of the more cautious people, \( q^3(0) + q^4(0) \), is about 1\%.

Comparing the logistic and the hierarchical logistic model, we show the result with \( m = 1 \) and \( m = 4 \) in Fig. 1. The hierarchical logistic function with \( m = 4 \) matches the data which the logistic model cannot approximate.
Table 1. Parameters minimizing the product of SARE and SAE and SAREs by them. All values are rounded to a three-digit number, and so the sum of \( p(0) \) and \( p^u(0) \) is not equal to one.

|       | \( m = 1 \) (logistic) | \( m = 2 \) | \( m = 3 \) | \( m = 4 \) |
|-------|-------------------------|-------------|-------------|-------------|
| \( a \) | 1.43                    | 4.42        | 4.25        | 4.17        |
| \( N \) \([\times 10^6]\) | \( 1.28 \times 10^5 \) | 64.6        | 65.4        | 66.2        |
| \( p(0) \) | \( 9.26 \times 10^{-7} \) | 0.00207     | 0.00200     | 0.00189     |
| \( q^1(0) \) | 1.00                    | 0.288       | 0.311       | 0.325       |
| \( q^2(0) \) | –                       | 0.710       | 0.675       | 0.657       |
| \( q^3(0) \) | –                       | –           | 0.0125      | 0.00619     |
| \( q^4(0) \) | –                       | –           | –           | 0.00760     |
| SARE   | 1.84                    | 1.12        | 1.02        | 1.01        |

Figure 1. (color online) Cumulative iPod sales expresses by circles and fitting curves. The (purple) dashed and (light blue) full curves are solutions of the logistic and the hierarchical logistic equation \((m = 4)\) with parameters minimizing the product of SARE and SAE, respectively.

5. Concluding remarks
In this work, we have unveiled the following fact that the essential human communication within a group, where the dynamical behaviour of the penetration rate can be approximated by the logistic function, is imitation; non-adopters start to process a prevalent stuff shortly after meeting adopters. Indeed, this is not natural. Thereby, we have proposed the extended logistic equation, the hierarchical logistic equation, considering the memory of the number of adopters they met. In addition, we have applied this model to the change of iPod sales in time, and so the model has approximated the data much better than the logistic equation.

6. Acknowledgments
We would like to thank Tomo Tanaka and members of astrophysics laboratory at Ochanomizu University for extensive discussions.

References
[1] Buchanan M 2007 The Social Atom (New York: Bloomsbury Press)
[2] Verhulst P P 1838 Corr. Math. Phys 10 113
[3] Griliches Z 1957 Econometrica 25 501