Information Cascade in a Sequential Voting Experiment

S. Mori[a] and **M. Hisakado**[b]

[a]Department of Physics, Kitasato University
1-15-1 Kitasato, Sagamihara 252-0373, Japan
mori@sci.kitasato-u.ac.jp

[b]Standard and Poor's
1-6-5 Marunouchi, Chiyoda-ku, Tokyo 100-0005, Japan
masato_hisakado@standardandpoors.com

Keywords
voting model, phase transition, information cascade, herder

In this paper, we study how people gather information by observing others' choices and what happens as the result. We perform an simple experiment of sequential voting and we control the transmission of the information about others' choices among people. We observe an information cascade transition in the distribution of the percentage x of persons who answer correctly. We clarify how people refer to others' choices and introduce a voting model with two parameters λ and z. The model describes the transition quantitatively well.
We perform an experiment of sequential voting. We prepared 200 questions with two choices - knowledge and no knowledge. 31 persons answered these questions sequentially and their order is denoted as $t \in \{1, 2, 3, \cdots, T = 31\}$ First, they answered without any information about the others' answers, i.e., their answers were based on their knowledge. Persons who know the answers can select the correct answers. Persons who do not know the answers can select the correct answers with a probability of $\frac{1}{2}$. We denote the ratio of the latter as p. The expectation value of x is $\mathrm{E}(x) = (1-p) + \frac{1}{2}p$. Next, the person who answered the questions could see the previous r persons' answers [1]. Persons who do not know the answers refer to this information. In the experiment, we increase r from 0 to ∞ in the set $r \in \{0, 1, 2, 3, 5, 7, 9, \infty\}$ and control the information transmission. People need to answer eight times for each question at maximum. We denote the number of the person who choose the correct answer after t as $N_1^r(t)$.
Fig.1 shows the distribution $P(x)$ of $x = \frac{N_1^r}{T}$ for $r = 0$ and $r = \infty$. The mean value of x is about 0.6 and we can estimated $p = 2 \cdot (1 - \mathrm{E}(x))$ as $p \simeq 0.8$. About 80% of the people are ignorant persons. They can refer to the previous answers and changes their answers. The resulting distribution of x is also shown for $r = \infty$. We see two peaks in the distribution. One peak is at $x = 1$, which means that they succeed in choosing the correct answers by referring to the previous answers. The other peak is at $x = 0.2$, which means that all ignorant person fail in choosing the correct answers and only the person who knows the answer can choose the correct ones.
In order to study the mechanism of the change in $P(x)$ from $r = 0$ to $r = \infty$, we study how the ignorant person's answer depends on the previous answers. We study the relation between the number n_1^r of the persons who answer correctly in the previous r persons and the probability p_1 that the ignorant person choose the correct answer. If the person answers without any information, the probability is 0.5. For the $r = 1$ case, as n_1^r/r varies from 0 to 1, p_1 increases slightly. For $r = 9$ and $r \geq 10$, the answers are greatly affected by the previous answers and p_1 increases much. The increases reflect how the previous answers

[1]If the order t of the person is less than $r + 1$, he can see all the previous $t - 1$ persons' answers.

affect the persons choice. In general, we can describe the relation by

$$p_1(r, n_1^r) = (1-p) + p \times 0.5 \left(\tanh(\lambda(\frac{n_1^r - \frac{r}{2}}{r+z}) + 1 \right). \tag{1}$$

The parameters $\lambda \in \{-\infty, \infty\}$ and $z > 0$ describe how their answers depend on the previous answers. If λ is positive and large, the ignorant person trusts the previous answers much. In the $\lambda \rightarrow \infty$ limit, he follows the majority of the previous r answers. We call him digital herder [2]. If $\lambda \leq 2$, the model reduces to the case of analog herder [1]. z means a weight of one vote. If z is large, the weight is small and the answer is not so much affected by the previous one answer. The maximum likelihood estimate of λ and z are $\lambda = 4.15$ and $z = 11.21$. The 95% confidence interval for them is $3.57 \leq \lambda \leq 4.92$ and $8.72 \leq z \leq 14.64$. In the figure 2, we show the fitted results.
We introduce a general voting model where people vote with the conditional probability eq.(1). The probability $\mathrm{P}(n,t)$ that the correct answer gets n answers up to t obeys the following master equation

$$P(n, t+1) = (1 - p_1(r, n))P(n, t) + p_1(r, n-1))P(n-1, t) \tag{2}$$

with the initial condition $P(0,0) = 1$. The model shows a phase transition for the appropriate choice of the parameters r, p, λ, z in the limit $t \rightarrow \infty$. By using the mean field approximation, we can study the phase diagram of the model. In the limit $r \rightarrow \infty$, the critical value of λ is 2. For $\lambda > 2.0$ and $p \geq p_c(\lambda)(< 1.0)$, the system is in the two peak phase [2]. If $\lambda \leq 2.0$ or $p < p_c(\lambda)$, the system is in the one peak phase. The order parameter of the phase transition is the probability that all ignorant person choose the wrong answers. Fig 3 shows the probability $\alpha = \mathrm{Prob}(N_1^\infty < \frac{T}{2})$ as the function of p. For the experimental data, the value of p is the maximum likelihood value and estimated as $p_{MLE} = 2(1 - \frac{N_1^0}{T})$.
Our voting model describes the information cascade transition quantitatively well. In the talk, we explain in detail how to estimate p_c based on the experimental data.

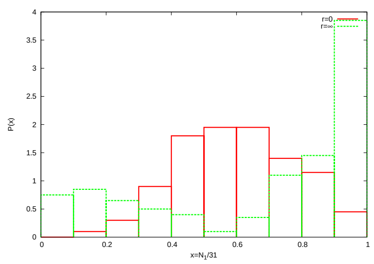

Figure 1: $\mathrm{P}(x)$ vs $x = \frac{N_1}{31}$ for $r = 0$(red) and $r = \infty$(green).

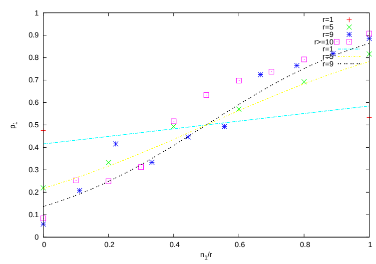

Figure 2: n_1^r/r vs $p_1(r, n_1^r)$ for $r = 1, 5, 9$ and $r \geq 10$.

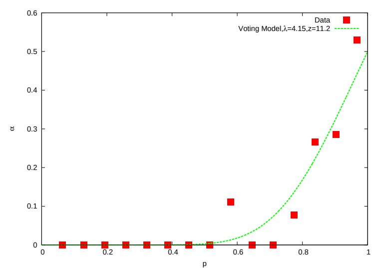

Figure 3: $\alpha = \mathrm{Prob}(N_1^\infty < \frac{T}{2})$ vs p. $T = 31$.

References
[1] M. Hisakado and S. Mori, *Phase transition and information cascade in a voting model*, J.Phys.A,Math.Theor.43(2010)315207.
[2] M. Hisakado and S. Mori, *Digital herders and phase transition in a voting model*, J.Phys.A,Math.Theor.44(2011)275204.