

Information Cascade & rational herding

근래, 인터넷 괴담이라는 말이 펍 떠돌고 있습니다. 근거가 부족한 내용의 말이 인터넷을 타고 걸잡을 수 없이 퍼지고 있는 현상을 말하죠.

근거가 부족한 면이 있는 것 같습니다. 하지만, 그렇다고 이런 말들이 꼭 틀리다라고 할 수도 없습니다. 아직 맞다, 틀리다 정확한 판단이 서지 않는 것이죠. 특별히 해당 정보에 접근할 수 있는 사람이던가, 아니면 전문가인 분들은 자신들의 판단에 더 확신을 갖을 수 있을 것입니다. 그렇지만, 한번 바람이 불면 이런 의견들이 묻혀 버리곤 하죠.

인류 역사상 이런 현상들은 수 없이 반복되었습니다. 개인들이 판단을 하려고 할 때 자신이 갖고 있는 데이터가 불충분하기에 다른 사람들이 어떻게 하였는가를 보고 이를 따라 하는 그런 행동 말입니다. 예전 네덜란드에서 있었던 튜립사건(http://en.wikipedia.org/wiki/Tulip_mania) 부터 수많은 “묻지마” 사건들 중 많은 것이 이 부류에 속할 것 같습니다.

위키, http://en.wikipedia.org/wiki/Information_cascade, 에서 information cascade를 “뒤에 있는 사람들이 자신들의 개인적인 판단에 의지하지 않고, 앞에 있던 사람들이 이미 내린 의사결정/행동을 참고하여 앞의 사람들의 의사결정/행동을 따라하는 것” 을 말합니다. 핵심은 이런 무리/집단적 행동이 반드시 비합리적이지 않다는 것입니다. 또 모든 사람들이 개인적인 판단에 근거해 의사결정을 하는 것보다 더 많은 사람들에게 “정확한” 의사결정을 제공해 주기도 합니다. 나쁘다고만 할 수 없죠. 어찌 보면, 우리 인류가 진화해 오면서 터득한 좋은 학습효과이기도 합니다.

이제 우리의 관심은 인터넷 시대에 정보의 확산속도가 엄청나게 빨라지고, 정말로 다양한 온갖 말들이 떠 돌아다닐 때 이런 information cascade가 어떻게 발생하고 그 이로움과 해악은 어떻게 나타나는지에 있습니다. 상식적으로 보아도 information cascade의 발생, 속도, 규모, 파급효과가 예전보다 더 커질 것 같지 않습니까? 본격적으로 공부하고 싶으면 <http://www.info-cascades.info/> 에 가면 많은 자료 있습니다. 우리도 이것을 활용합니다.

Information Cascading : 약간의 수학 배경 지식

몇 가지 기본 확률 공식 :

- $P(A|B) = \frac{P(AB)}{P(B)}$: B가 발생한 가운데, A도 생길 (조건부) 확률

- 위의 공식에 따라서,

$P(AB) = P(A|B) \cdot P(B)$ 가 된다. 즉, A와 B가 동시에 발생할 확률은 B가 발생할 확률에 $P(A|B)$ 조건부 확률을 곱한 값이 된다.

또한, $P(AB) = P(BA)$ 이므로 $P(BA) = P(B|A) \cdot P(A)$ 를 적용하면

$P(AB) = P(B|A) \cdot P(A)$ 또한 성립한다.

- $P(A + B) = P(A) + P(B) - P(AB)$: A 또는 B 상황이 발생할 확률

- * 만약 A와 B가 동시에 발생한 확률, $P(AB)$ 가 0 이라면

$P(A + B) = P(A) + P(B)$ 가 된다

Total Probability

If $B = [B_1, B_2, \dots, B_n]$ is a finite or countably infinite partition of a probability space and each set B_n is measurable, then for any event A we have

$$P(A) = P(A|B_1) \cdot P(B_1) + P(A|B_2) \cdot P(B_2) + \dots + P(A|B_n) \cdot P(B_n)$$

• Independent Events

Two events are independent if an occurrence of one has no effect on the probability of other. Then,

$$P(AB) = P(A) \cdot P(B)$$

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

Bayes' rule

앞서 우리는

$$P(A|B) = \frac{P(AB)}{P(B)} \quad \text{을 보았다. 또한, } P(B|A)P(A) = P(AB) = P(A|B)P(B) \text{ 이다.}$$

따라서,

$$P(A_n|B) = \frac{P(B|A_n) \cdot P(A_n)}{P(B)} \quad : \text{ Bayes rule. } P(A_n) \text{를 } a \text{ priori, } P(A_n|B) \text{를 } a \text{ posteriori}$$

확률이라 부른다.

$$= \frac{P(B|A_n) \cdot P(A_n)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3) + \dots\dots\dots P(B|A_n)P(A_n)}$$

* Bayes rule 은 conditional probability $P(A|B)$, 즉 B 를 이미 했을 때 (또는, 가정하고) A를 구하려 할 때, 이를 직접 구하는데 필요한 정보는 없고 대신 $P(A)$, $P(B)$ 와 $P(B|A)$ 를 알 때 $P(A|B)$ 를 구할 수 있다는 얘기이다. 즉, 어떤 상황 (A)이 이미 발생하였다고 전제/가정하고 상황(B)의 발생 확률($P(B|A)$)을 구하고자 할 때, 전제와 실제 알고자 하는 상황의 순서가 바뀐 ($P(A|B)$)를 알면 구할 수 있다는 말이다. 흔히, $P(B|A)$ 를 알고자 할 때, $P(A|B)$ 는 직관적으로 또는 쉽게 구할 수 있는 값일 경우가 많다.

Bayes' 활용 예

서울에는 3 가지 색깔의 택시가 영업한다고 하자. 흰색 택시가 전체의 50%, 검은색이 30%, 빨강색이 20% 이다.

어느날 강남역에서 택시관련 뺑소니 사고가 났는데 목격자의 말이 뺑소니 택시가 빨강색이었다고 한다. 그런데, 이 때 날이 어둑어둑해서 색깔 구분이 조금 어려웠다. 이럴 때, 맞게 색깔 구분을 할 수 있는 비율이 색깔에 관계없이 80% 라 한다.

이럴 경우, 이 뺑소니 사고 목격자의 진술이 맞을 확률은 얼마나 될까?

→

우선, 문제를 파악해 보자. 진술이 맞다는 말을 확률로 표현해 보자. 목격자는 빨강색이라 진술을 먼저 했다. 그러니 진술이 맞다는 것은 목격자의 말이 빨강색이었으니, 실제 택시의 색깔도 빨강색이어야 함을 의미한다. 즉, 실제 택시색깔과 목격자의 진술 관련 확률은 서로 다르다.

흰색을 **W**, 검은색을 **B**, 빨강을 **R**로 표시하고, 실제 택시색깔 표시를 **T**로, 목격자의 얘기를 **S**로 표현하면, 목격자의 진술이 맞다는 얘기는,

$P(T=R | S=R)$ 가 된다. 우리가 이미 아는 것을 표현해 보면, $P(T=W)=0.5$, $P(T=B)=0.3$, $P(T=R)=0.2$ 이다. 진술의 정확성을 표현해보면, $P(S=R | T=R) = 0.8$ 이다. 다른 흰색, 검은색에 대해서도 같이 0.8이 된다. 이제 $P(T=R | S=R)$ 에 Bayes theorem 을 적용하면,

$P(T=R | S=R) = P(S=R | T=R) \cdot P(T=R) / P(S=R)$. 이는 다시

$$= \frac{P(S=R | T=R) \cdot P(T=R)}{P(S=R | T=R)P(T=R) + P(S=R | T=W)P(T=W) + P(S=R | T=B)P(T=B)}$$

$$\frac{P(S=R | T=R) \cdot P(T=R)}{P(S=R | T=R)P(T=R) + P(S=R | T=W)P(T=W) + P(S=R | T=B)P(T=B)}$$

에 값을 적용하면,

$$\frac{0.8 \times 0.2}{0.8 \times 0.2 + 0.2 \times 0.5 + 0.2 \times 0.3} = 0.16 / 0.32 = 0.5$$

즉, 목격자의 진술이 맞을 확률이 **50%** 이다.

✓ 어떻게 된 것이지? 처음에 사람이 정확히 색깔을 맞출 확률이 80%라 했다. 그러니, 직관적으로 생각하면 사고를 목격한 사람이 정확히 색깔을 보았을 확률이 80%라 하기 쉽다. 그런데, 앞에서 본 것과 같이 50%의 확률이 나왔다.

이런 직관과 사실과의 차이는 정확하지 않은 표현에서 나온다. 앞서 정확성이 80%라 함은 정확히는 한 색깔을 보고 그 색깔을 정확히 말하는 확률이 80%라는 말이다.

우리가 알고자 했던 것은, 한 사람이 색깔을 보았다고 하는데 그 색깔이 정말로 그 사람이 보았다고 하는 색깔이 맞는가 하는 것이다. 왜냐하면, 그 사람은 다른 색깔을 보고도 그 색깔을 보았다고 20%의 확률로 말 할 수 있기 때문이다.

Information Cascades ;

출처 : 1. A theory of fads, fashion, custom, and cultural change as Information Cascades,
[Sushil Bikhchandani](#), [David Hirshleifer](#), [Ivo Welch](#) , 1992
2. Easley & Jon Kleinberg, 코넬대학, 2007

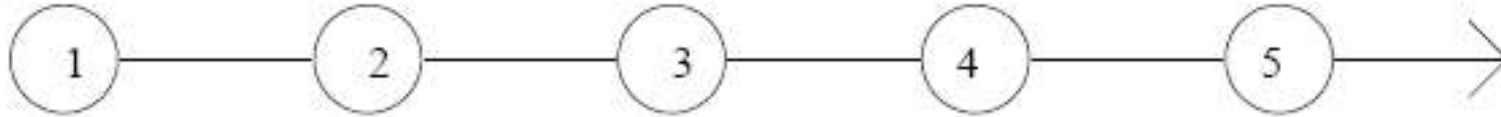


Figure 1: Information Network

“In many social settings individuals have an opportunity to observe the actions of others before they have to act. In these settings it may be rational for an individual to imitate the choices of others even if the individual's own information suggests an alternative choice. In this case **herding** or an **information cascade** is said to have occurred.”

* www.wikipedia.org 나 구글에서 “**information cascade**” thread 를 따라 가 보면 게임이론과 많은 경제학적, 사회적 관련 글들을 접할 수 있다.

Information Cascades and Rational Herding

출처 : <http://www.info-cascades.info/>

Definition: An *information cascade* is a situation in which every subsequent actor, based on the observations of others, makes the same choice independent of his/her private signal.

Erroneous Mass Behavior: In an *information cascades* everyone is individually acting rationally. Still, even if all participants as a collective have overwhelming information in favor of the correct action, each and every participant may take the wrong action. The probability that everyone is taking the wrong action is less than 50%, but it is easy to construct examples in which everyone is wrong with 30-40% probability.

Fragility: A little bit of public information (or an unusual signal) can overturn a long-standing *informational cascades*. That is, even though a million people may have chosen one action, seemingly little information can induce the next million people to choose the opposite action. Fragility is an integral component of the *Informational cascades* theory!

Information Cascade 의 한 예:

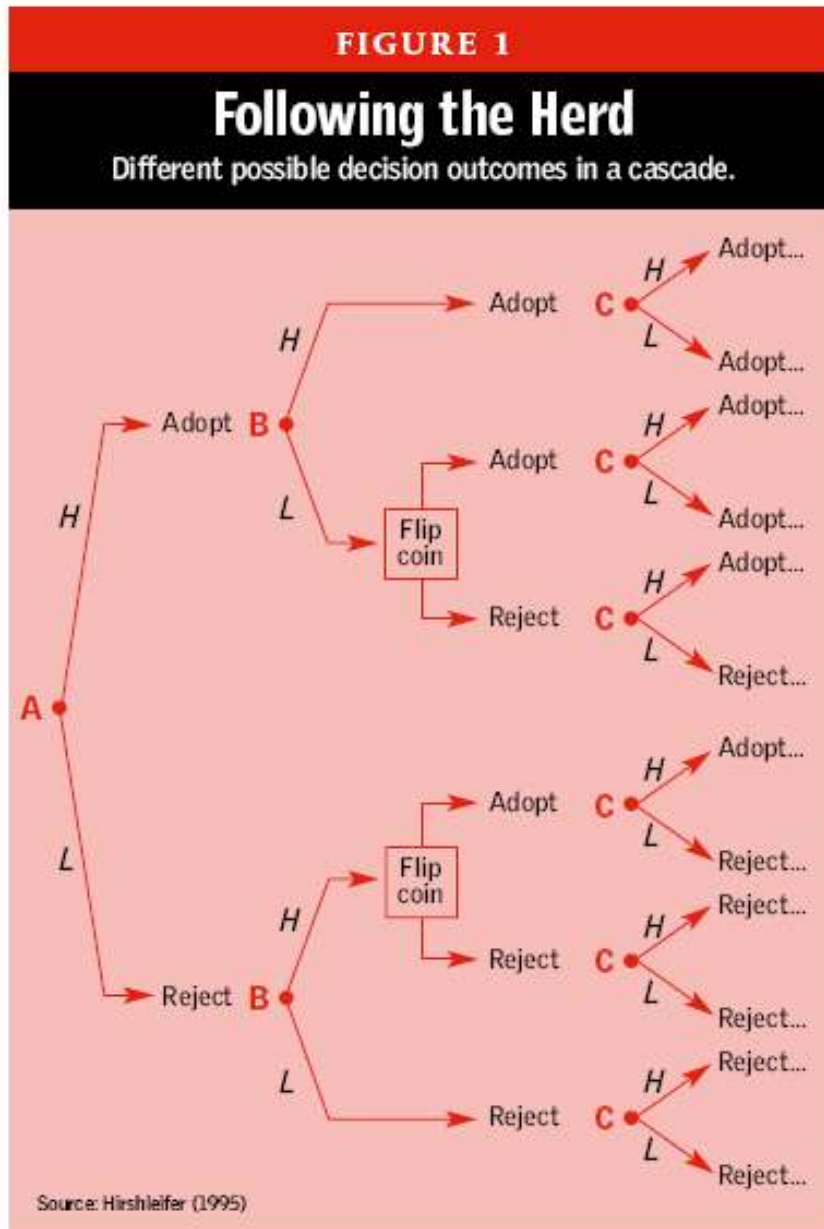
사람들은 다른 사람들의 선택을 보고 그들을 따르려는 경향이 있다. 특히, 자신이 생각하는 자신의 판단 근거가 약할 때 많이 나타나는데 이를 특별히 비합리적이라 말할 수는 없다. 이렇게 남들을 따라 행동하는 것이 합리적일 수 있을 때 이를 rational herding (합리적 군중심리) 또는 information cascade 이라 한다.

예를 들어, 사람들은 어떤 식당에서 먹을지 고를 때 어느 식당이 더 손님이 많은지 보고 많은 곳으로 찾아가는 것은 합리적인 행동이다. 특히나, 처음 간 고장에서 비슷한 식당들 중에 사람들이 많은 곳을 택하는 것은 이미 다른 사람들이 그 주위에 있는 여러 식당 중에서 먹어보고 이곳이 더 좋은 곳으로 선택했기 때문이라고 생각한 합리적인 결정이기 때문이다. 하지만 이런 경우 내가 그 식당에서 먹고, 그 식당은 더 손님들이 많아지고, 또 다른 나와 같은 사람이 나와 같이 판단해서 그 식당으로 오고, 이렇게 되면 처음에 시작이 어떻게 되었는지는 모르지만 모든 사람들이 그 식당으로만 집중되게 되는 결과가 생길 수 있다.

원지는 정확히 모르겠지만, 남들이 하니 “뭔가 이유가 있으니 그럴겠지” 하고 이를 따라 하는 것이 도처에 있다. 예를 들어, 주식시장, 패션, 유행, 정치판 등에서 다양한 형태로 일어난다.

직관적인 Information Cascade 현상 :

- 대부분의 개별 사람들이 결정을 내리기에 충분한 데이터가 없다고 하자. 즉, 대부분의 개인들이 문제의 판단에 대해 자신이 없다고 가정하자.
- 사람들은 결정을 내릴 때 자기 자신 스스로가 느끼는 신호/판단/주관 (Private Signal) 와 앞서 다른 사람들의 행동으로부터 유추한 신호 (Public signal) 를 근거로 어떤 결정을 내린다.
- 개별 사람들이 자신의 Private signal 에 대한 큰 믿음이 없으니 다른 사람들로 부터 얻은 Public signal 에 큰 영향을 받는다. 여기서, 자신의 private signal과 public signal이 같은 크기로 개인의 의사 결정에 동등한 영향을 미친다고 가정하자.
- 따라서 가령 내가 A라 생각해도 앞에 사람 2명이 B이면 나는 B라 결정한다. Fig-1 참조. 즉, 개인의 주관이 확실하지 않을 때 그 개인은 다른 사람들의 행동을 참고하여 나의 주관에 관계없이 다른 사람들의 결정을 따르게 되고, 그러면 그 뒤에 다른 사람들도 같은 결정에 따르게 되어 향후 뒤의 모든 사람들도 자신들의 주관에 관계없이 집단의 결정을 따르게 되는 것이 information cascade이다.
- Information cascade가 나타나면 더 이상 개인들이 지녔던 private signal (주관) 이 변수가 되지 못한다.
- 따라서 information cascade가 나타나는 이유는 개인들이 자신의 private signal에 대한 확신이 없는 가운데, 다른 사람들의 private signal은 보지 못하고 다른 사람들의 결정만 보고 판단을 하면서 발생한다.



Up(correct) cascade : 맞는 cascade.

Down(incorrect) cascade : 틀린 cascade

-개인들이 지닌 private signal이 약 60% 정확성을 갖을 경우에도 down cascade가 1/3 정도 생김.

- information cascade의 유지는 취약함. 왜냐하면 information cascade에 동참한 개인들도 자신들이 같은 결정을 내린 이유가 정보의 부족때문이라는 것을 알기 때문에 더 새로운 정보가 나오면 information cascade는 중단될 수 있음. 즉, information cascade가 incorrect cascade일 때, 어떤 맞는 얘기가 나오기 시작해서 information cascade에 동참한 사람들에게 “너 왜 그렇게 행동했어” 라고 물어보자. 그러면 그들은 “다른 사람들이 그렇게 하기에 그것이 맞는 줄 알았어” 대답할 가능성이 많다. 이렇게 되면 information cascade는 힘을 잃는다.

논리적인 Information Cascade 분석

Information Cascade가 재미있는 현상이다 보니 이를 보다 더 정밀하고, 논리적으로 분석하려는 시도가 경제학, 사회학 분야 전문가들에 의해서 진행되었다.

자료, <http://www.info-cascades.info/>, 에 의하면 information cascade가 본격적으로 다루어진 것이 다음과 같다고 한다:

- [Abhijit V. Banerjee](#). August 1992. A Simple Model of Herd Behavior. *Quarterly Journal of Economics* 107:3, 797-818
- [Sushil Bikhchandani](#), [David Hirshleifer](#), [Ivo Welch](#). October 1992. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy* 100:5, 992-1026
- [Ivo Welch](#). June 1992. Sequential Sales, Learning, and Cascades. *Journal of Finance* 47:2, 695-732

■ 원래 우리의 관심은 다양한 의견/정보/데이터가 나오고 엄청나게 쏟아지고, 쉽게 이런 말/글들에 접근할 수 있고, 전파속도가 무척 빠르고, 광범위하게 다양한 주제로 소셜네트워크가 형성되어 있고 또 형성될 수 있는 웹시대에서 information cascade를 살펴보고자 하는 것이다. 특히, 웹구조가 어떻게 information cascade에 영향을 미치고, 또 어떻게 information cascade를 예측, 발견할 수 있을까를 알아보고 싶다. But, 이 주제가 간단하지가 않아 “[Sushil Bikhchandani](#), [David Hirshleifer](#), [Ivo Welch](#). Summer 1998. Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades” 를 기본으로 information cascade를 조금만 분석적으로 보는 것으로 한다.

저 음식점으로 들어가는(Accept) 것이 좋은 결정일까, 아닐까?

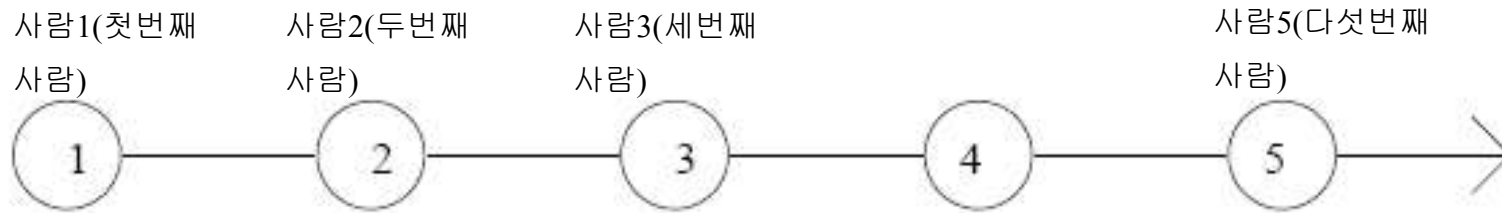


Figure 1: Information Network

- p : Accept 하는 것이 좋을 (Good state) 확률, (또는, 나의 결정이 옳을 확률)

$$P(G) = p, \quad 1 > p > 0$$

- $(1-p)$: Accept 하는 것이 나쁠 (Bad state) 확률 (또는, 나의 결정이 틀린 확률)

$$P(B) = 1 - p$$

* 개인들은 $P(G)$ 가 어떤 값일지 알고 싶다. $P(G)$ 를 알면 들어갈까, 아니면 말까를 판단할 수 있기 때문이다. 가령 $P(G)$ 가 0.1 이라는 것을 만약 안다면, 그 음식점 말고 다른 대안이 있을 때 그곳으로 들어갈 사람은 없을 것이다. 그런데, 우리들의 갖고 있는 이런 곤란한 문제에 도움을 주는 것이 있으니, 우리 앞에 있던 사람들이 어떻게 행동을 했는가를 보고 우리의 의사결정에 참고할 수 있다는 것이다. 이렇게 우리들의 의사결정에 참고할 수 있는 정보를 신호(signal) 이라 한다. 신호는 private signal과 public signal로 나뉘는데, private signal은 각자 사람들이 개인적으로 느끼는/판단하는 그 어떤 개인 고유의 판단 잣대이다. 주관이라고 해도 좋겠고, 개인적 직관, 믿음이라고 해도 좋겠다. 기본적으로 이런 private signal도 개인의 합리적인 사고에 따른다고 생각한다.

Public signal은 개인이 다른 사람들의 행동/의사결정을 보고 받게 되는 signal이다. 앞의 한 사람은 하나의 public signal을 뒤에 있는 사람에게 제공한다.

따라서 사람 1을 제외한 사람2, 3, 4 등의 뒤에 있는 사람들은 개인 고유의 private signal외에 앞에 있는 사람들의 행동들에 따른 public signal을 참고하여 의사결정/행동을 한다.

이 신호(signal 또는 s)는 L 또는 H 값을 갖는다. $s=H$ 이면 accept 하는 것이 좋다는 신호이고, $s=L$ 이면 나쁘다는 신호이다. 첫번째 사람은 자기 자신의 private signal 밖에 없었으니, 자신만을 믿고 의사결정을 할 것이지만, 두번째 사람2 부터는 앞의 사람들의 행동도 참고하여 자신들의 의사결정을 할 수 있다.

가령 세번째 사람3은 3개의 신호를 받게 되는데, s_1 은 사람1로부터의 public signal, s_2 는 사람2로부터의 public signal이고, s_3 은 자신의 private signal이다. 따라서 public signal은 앞의 사람들이 이미 한 행동/의사결정에 따른 것으로, 사람1이 accept 하였다면 $s_1=H$ 이고, 사람2가 accept 하지 않았으면 $s_2=L$ 이다. Signal의 H와 L의 값을 1과 -1로 나타내는 경우도 있다. 이제, 이렇게 다른 사람들의 행동을 참고한 public signal을 포함한 사람들의 의사결정의 정확도가 어떤 지를 보고, 어떻게 되면 information cascade가 시작되는지 살펴보자. 몇 가지 확률적 가정이 더 필요합니다 ;

- $P(H|G) = q$, 즉 Accept 하는 것이 좋은 결과를 가져온다는 가정하에, 신호가 H일 확률을 q 로 나타낸다. 즉 q 는 음식점이 실제로 좋고, 내가 보내는 신호도 H일 확률이다. 무슨 얘기인가? 음식점이 좋는데 신호는 L일 수도 있다는 것인가? 그렇다. 앞서 Bayes rule 얘기에서 목격자가 흰택시를 보았는데, 증언도 흰택시라고 할 확률이 100%가 아닐 수 있다는 것을 참고하면 이해된다. 그러면, $P(L|G) = 1-q$ 가 되어 실제로는 좋는데 신호는 나쁘게 나올 확률이 $(1-q)$ 이다. 물론 이 때 q 가 0.5가 넘을 가능성이 많겠다.

- $P(L|B) = q$, 즉, 음식점이 실제로 나쁘고 신호도 나쁘게 나올 확률이 q 이다.

- $P(H|G) = q$ 와 $P(L|B) = q$ 를 보아, 신호가 맞을 확률이 q 라고 생각하며 되겠다.  마이윙 www.miwing.com, 이현봉

첫 번째 사람의 의사결정



Figure 1: Information Network

첫 번째 사람 1은 자신의 **private signal** 외에 다른 판단근거가 없다. 왜냐하면, 사람1이 첫 번째 고객이기 때문이다. 이 때 실제로 $P(G)$ 값은 있겠지만, 그것은 아직 신만 알고 있다. 사람 1은 아직 식당에서 음식을 먹어보지 않았기 때문이다. 다만 1은 자신의 느낌으로 판단할 뿐이다. 우리의 관심은 사람1의 **private signal** $s_1=H$ 신호를 받았는데, 실제로 식당이 **Good** 일 확률이 얼마나 될 것인가 하는 것이다 ($P(G|H)$). 왜냐하면, 손님이 들어갈까 아니면 말까 하는 판단은 그가 받는 신호/ s 에 근거하기 때문이다. 그리고, 사람 2는 자신고유의 **private signal** s_2 외에 1에서부터의 **public signal** s_1 도 받고.... 이런 식으로 진행된다.

■ 주의 : 사람1이 의사결정/행동할 때 s_1 은 자신의 **private signal**이고, 사람1이 의사결정한 후에 사람2가 의사결정할 때 참고하는 s_1 은 사람2에게 보이는 사람1의 **public signal**로 앞서의 **private signal**과 비교해 그 값이 다를 수 있다. 즉, 한 사람이 의사결정을 한 후에 외부에서 보이는 **signal**은 원래 그 사람이 가졌던 **private** 신호/생각이 다를 수 있다는 것이다. 외부에서는 개인들의 속마음이 보이지 않고, 행동만 보인다.

$$\begin{aligned}
 P(G|H) &= \frac{P(G)P(H|G)}{P(H)} \\
 &= \frac{P(G)P(H|G)}{P(G)P(H|G) + P(B)P(H|B)} \\
 &= \frac{pq}{pq + (1-p)(1-q)} \\
 &> p.
 \end{aligned}$$

첫 번째 사람의 의사결정 - 계속

- 앞의 식에서 $P(G|H) > p$ 을 주목하자.

p 는 식당에 들어가는 것이 좋을 확률이다 ($P(G)$). 그렇지만, 아직 식당에 들어가지 않은 손님에게는 $P(G)=p$ 값은 실제로 보이지를 않는다. 이 때 p 값은 누구도 모른다. 하지만, 우리에게 느낌이나 참고할 만한 데이터가 있는데 이것이 signal이다. 첫 번째 사람에게서는 오직 자신의 private signal s_1 만 있다. $P(G|H)$ 는 사람1의 $s_1=H$ 일 경우에 식당에 들어가는 것이 올바름(좋음) 확률을 나타낸다.

$P(G|H) > p$, 즉 H 신호를 받았을 때, 식당에 들어가는 것이 좋을 확률이 p 보다 크다는 말은, 최소한 H 신호가 있으면 식당에 들어가는 것이 좋을 확률이 H 신호가 없을 때의 p 값보다 어쨌든 더 크다는 말이다. P 가 어떤 값인 지는 모르지만, 내가 accept 할까를 의사결정을 함에 있어 H 신호가 있다면 최소한 H 신호도 없이 그냥 들어갈 때보다는 accept 한 것이 좋을 확률이 더 커진다는 의미이다. (당연한 얘기인가?)

가령 $p=0.4$ 라면 식당에 들어가는 것이 좋은 생각일 확률이 40%라는 의미이다. 그런데, 손님은 아직 이런 사실을 모른다. 알고 있다면 들어가지 않을 것이다. 그런데, H 신호를 받으면 식당에 들어가는 것이 좋을 확률이 40%를 넘는다는 얘기이다. 즉, 손님의 판단이 옳을 확률이 올라간다는 의미이다.

즉, 사람1은 어차피 p 값을 모르니, H 신호를 받으면 식당에 들어가고, L 신호를 받으면 들어가지 않는 것이 현명하다.

■ 사람 2 의 경우를 보자.

사람2는 자신 고유 private 신호 s_2 와 앞서 사람 1의 행동을 볼 수 있다. 사람1의 행동은 식당에 들어갔는가 ($s_1=H$) 또는 들어가지 않았는가 ($s_1=L$) 의 신호 형태로 사람2에게 영향을 끼친다. 즉, 사람2는 s_1 과 s_2 에 의지해 자신의 행동을 결정한다.

이때 $s_1=s_2=H$ 이면 사람2는 식당에 들어가고, $s_1=s_2=L$ 이면 사람2는 식당에 들어가지 않을 것이다.

그런데, 만약 $s_1 \neq s_2$ 라면 어떨까?

$$\begin{aligned} P(G|L, H) &= \frac{P(G)P(L, H|G)}{P(L, H)} \\ &= \frac{P(G)P(L, H|G)}{P(G)P(L, H|G) + P(B)P(L, H|B)} \\ &= \frac{p(1-q)q}{p(1-q)q + (1-p)q(1-q)} \\ &= p. \end{aligned}$$

■ 사람 2 의 경우를 보자 ($P(G | L, H) = p$)

앞의 s_1 과 s_2 신호가 서로 달라 $P(G | L, H) = p$ 라는 말은 이럴 경우 사람2 는 식당에 들어가는 것이 좋을까, 아닐까 하는 의사결정에 신호들이 도움이 되지 않는다는 말이다. 신호들이 상쇄되어 신호가 아예 없는 경우 $P(G)=p$, 즉 하나님만이 알고 있는 값 p 인 상황에 처한 것이다. 이 경우 사람2가 어떻게 해야 할까? 주사위를 던지던지, 자신 고유의 판단에 따르던지, 아니면 다른 사람이 한 행동에 따르던지 할 것이다.

이럴 경우, 어떤 글에서는 주사위를 던진다고 가정하고, 또 다른 곳에서는 자신의 private 신호, 즉 s_2 의 값에 의존한다고 가정한다. 즉 $s_2=H$ 이면 accept하고, L 이면 reject한다고 가정하기도 한다.

■ 사람 3 의 경우를 보자.

사람3은 s_1 과 s_2, s_3 신호를 받는다. 이 경우에는 tie가 있을 수 없다.

사람 3의 의사결정 과정을 자세히 보자.

1. $s_1 = s_2 = H$ 일 때

앞서 두 사람이 다 accept하여 $s_1=s_2=H$ 신호를 사람3이 받았다고 하자. 이 경우 사람3은 자신의 private 신호/생각이 L 이어서 $s_3=L$ 라 하여도 사람3은 앞서 2 사람의 판단을 모방하여 자신도 accept 한다. 이래서, 사람3은 앞으로 외부에는 $s_3=H$ 신호를 주게 되어 4번째 사람부터는 $s_3=H$ 로 보인다. 즉, 자신이 원래 가졌던 private 생각인 L 은 사라지고 더 이상 외부에 전해지지 않는다. 4번째 사람도 accept 할 것이고, 그 이후 모든 사람들도 accept 할 것이다. 3번째 사람을 시작으로 information cascade가 발생했다.

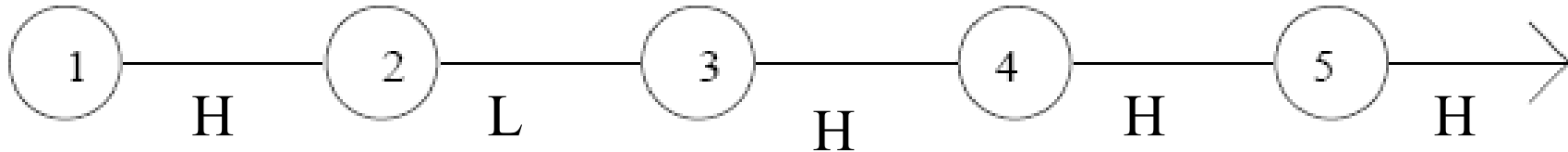
2. $s_1 = s_2 = L$ 일 때

앞서 두 사람이 다 reject하여 $s_1=s_2=L$ 신호를 사람3이 받았다고 하자. 이 경우 사람3은 자신 스스로의 생각이 s_3 에 관계없이 자신도 reject한다. 사람3은 다른 사람들에게 $s_3=L$ 을 나타낸다. 사람3부터 information cascade가 발생했다.

3. $s_1 \neq s_2$ 일 때

앞서 두 사람으로부터의 public 신호가 서로 상쇄된다. 사람3은 이 경우 자신의 s_3 신호에 따른다.

Cascade의 시작



사람5는 자신의 private 신호 s_5 에 관계없이 식당에 들어가고 public 신호 H를 내보낸다. 5 이후의 모든 사람들은 다 식당에 들어가게 된다.

“Let a_t be the number of Acceptances in the first t choices; then $r_t = t - a_t$ is the number of Rejections in the first t choices. Consider the first individual who observes two more Acceptances than Rejections, $a_t - r_t = 2$. No matter what signal he receives the number of high signals is more than the number of low signals (and he knows this) so he Accepts. Then the next individual to choose learns nothing from the choice of individual $t + 1$ and so he too Accepts regardless of his own signal. At this point an Accept cascade occurs as all subsequent individuals also Accept.

Similarly, the first individual who observes two more Rejections than Acceptances will Reject no matter what signal he receives. Then a Reject cascade occurs.”

Information Cascade의 음미

Information cascade를 다른 말로 rational herding, 합리적인 군중심리 또는 떼짓기라고 말하기도 합니다. 개개인으로 볼 때에 합리적일 수 있고, 사회 전체적으로도 많은 information cascade가 발생할 경우들을 평균하면 합리적인 선택일 경우가 많습니다. 결과가 반드시 바쁘다는 것이 아닙니다.

언뜻 생각하면 information cascade가 데이터가 부족한 가운데서 발생할 소지가 많기에 정보가 넘쳐 나는 인터넷 시대에서 생길 가능성이 전보다 작아지지 않을까 여겨집니다. 그렇지만, 요즘과 같이 웹정보화 시대에서 information cascade는 전보다 더 문제소지가 많을 수도 있습니다. 인터넷에 돌아다니는 얘기들을 다 신뢰할 수 없기에, 워낙 다양한 얘기들이 난무해 오히려 어떤 것이 맞는 것일까 혼동이 되고, 충분히 생각할 여유도 없고, 그러니 자칫하면 선동자의 꾀임에 휘말릴 수도 있겠죠. 특히, 앞에서 보았듯이 information cascade 발생에는 처음 소수의 사람들의 의사결정이 매우 중요한 역할을 합니다. 서로가 반신반의하는 가운데, 처음 몇 사람들의 목소리와 행동이 뒤에 오는 다수의 사람들의 의사결정에 큰 영향을 미치죠.

이렇게 보면 지금과 같은 정보화시대에서 더욱 중요한 것이 소수 사람들의 말입니다. 그것이 information cascade를 만들어, 다수의 사람들을 좋은 방향으로 인도하던가 또는 아니던가, 또는 information cascade 중에도 소수의 (정확한) 의견이 나와 (false) information cascade 를 중지시키던가.

Information cascade는 거대 담론, 학문적 흥미거리가 아닙니다. 싸이에서의 댓글, 인터넷 쇼핑몰에서의 사용자평, 댓글들이 모두 information cascade를 염두에 둔 기능들입니다. 여론 형성을 통해 목적하는 이익을 얻으려는 시도들이죠.