The Value and Pricing of Information

Haifeng Xu

Assistant Professor in Computer Science

University of Chicago

Outline

- Economic Foundations for Value of Information

(15 min break)

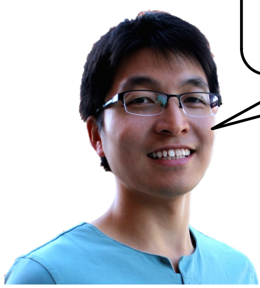
- Optimal Pricing of Information

- Summary and Open Problems

The Value of Distilled Data (i.e., Information)?

Anyone attending Haifeng's talk gets $100 – if correctly guess his coin toss

I tossed, and learnt the side



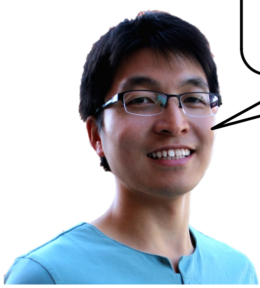
Question: How much is my information worth to James?

- Value of my information = \$100 - \$50 = \$50
  - Without my information, he gets \$50 via an arbitrary guess
  - With my information, he gets \$100
- Understanding its economic value is crucial for pricing information
  - Here, can sell my information to James at price \$49.9

Value of Info Depends to How It Is Used

Anyone attending Haifeng's talk gets $100 – if correctly guess his coin toss

I tossed, and learnt the side



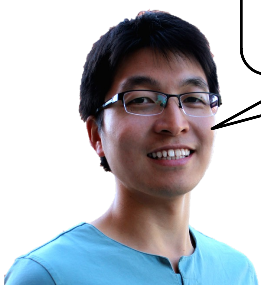
Question: How much is my information worth to James?

➢ Value of my information = \$100 - \$50 = \$50

Value of Info Depends to How It Is Used

Anyone attending Haifeng's talk gets ~~\$100~~ $2t$ if correctly guess his coin toss

I tossed, and learnt the side



Question: How much is my information worth to James?

➢ Value of my information = ~~\$100 - \$50 = \$50~~ t

Value of Info Depends to How It Is Used

Coin state ω

Prior belief $q(\omega)$

		1/2 tail	1/2 head
action	0	$2t$	0
	1	0	$2t$

$u(a,\omega)$

Value of Info Depends to How It Is Used

Coin state ω

Prior belief $q(\omega)$

		1/2 tail	1/2 head
action	0	$2t$	0
	1	0	$2t$

$u(a, \omega)$

- Without information, decision maker (DM) gets $\max_a [\mathbb{E}_{\omega \sim q} u(a, \omega)] = t$
- With my (full) information, DM gets $\mathbb{E}_{\omega \sim q} \left[\max_a u(a, \omega) \right] = 2t$

Value of (full) info = $\mathbb{E}_{\omega \sim q} \left[\max_a u(a, \omega) \right] - \max_a [\mathbb{E}_{\omega \sim q} u(a, \omega)]$

The Decision Problem Matters

Coin state ω

Prior belief $q(\omega)$

		1/2 tail	1/2 head
action	0	t	0
	1	0	$3t$

$u(a,\omega)$

Value of (full) info = $\mathbb{E}_{\omega \sim q}\left[\max_a u(a,\omega)\right] - \max_a\left[\mathbb{E}_{\omega\sim q} u(a,\omega)\right]$

$= \frac{1}{2}(t + 3t) - \max\left\{\frac{t}{2}, \frac{3t}{2}\right\}$

$= t/2$

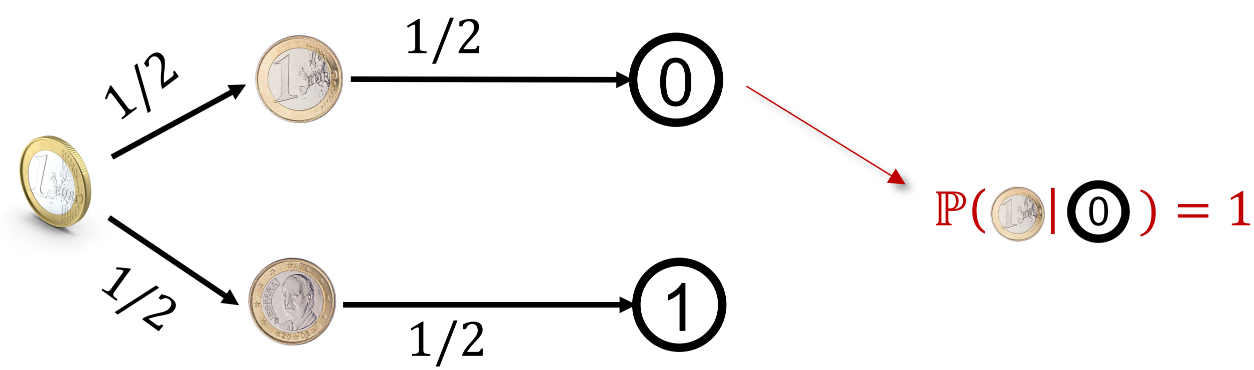
Moreover, Accuracy of Information Matters

Coin state ω

Prior belief $q(\omega)$

action	1/2 tail	1/2 head
0	t	0
1	0	$3t$

$u(a, \omega)$



Full information revelation

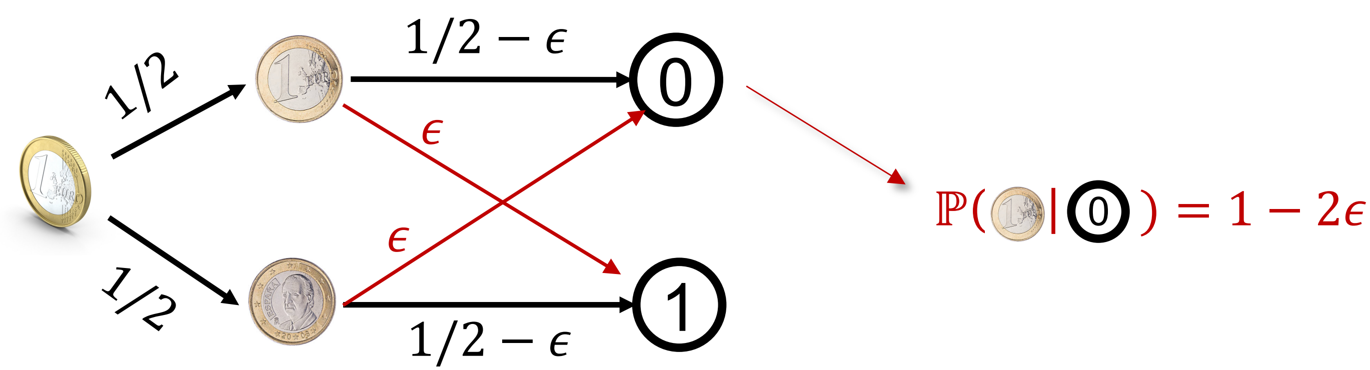
Moreover, Accuracy of Information Matters

Coin state ω

Prior belief $q(\omega)$

action		tail (1/2)	head (1/2)
	0	t	0
	1	0	$3t$

$u(a, \omega)$



Noisy information revelation

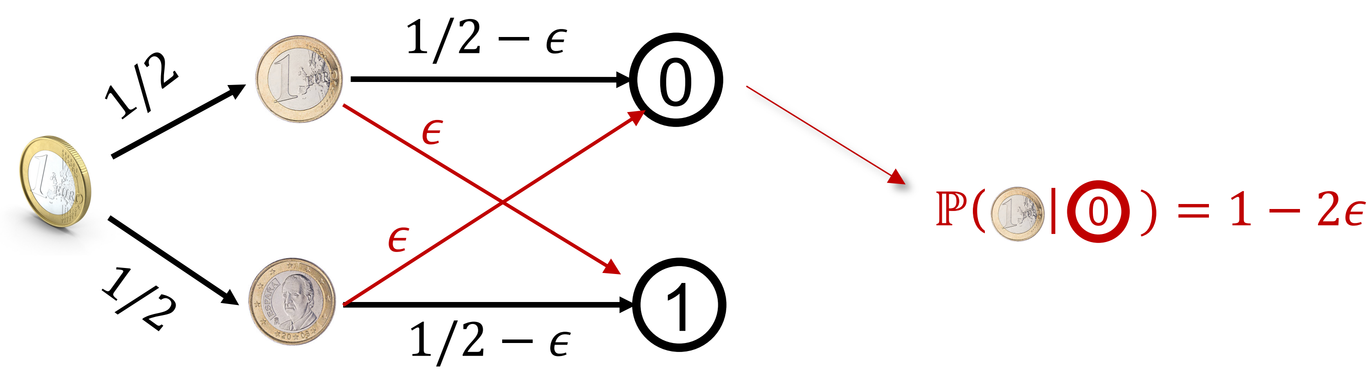
Moreover, Accuracy of Information Matters

Posterior belief $p(\omega)$, given ⓪

Prior belief $q(\omega)$

		tail	head
Posterior belief $p(\omega)$		$1 - 2\epsilon$	2ϵ
Prior belief $q(\omega)$		1/2	1/2
action	0	t	0
action	1	0	$3t$

$u(a, \omega)$



Noisy information revelation

Moreover, Accuracy of Information Matters

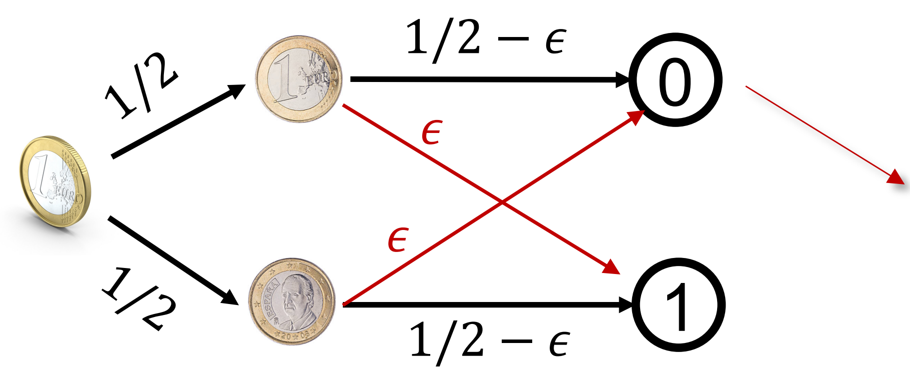
Posterior belief $p(\omega)$, given ⓪

Prior belief $q(\omega)$

		$1-2\epsilon$	2ϵ
		1/2	1/2
		tail	head
action	0	t	0
	1	0	$3t$

$u(a,\omega)$

In reality, can think of ⓪ as noisy prediction of state ω (e.g., stock trend, purchase prob)



$\mathbb{P}($  $|$ ⓪ $) = 1 - 2\epsilon$

Noisy information revelation

Moreover, Accuracy of Information Matters

Posterior belief $p(\omega)$, given 0: $1-2\epsilon$ (tail), 2ϵ (head)

Prior belief $q(\omega)$: 1/2 (tail), 1/2 (head)

action	tail	head
0	t	0
1	0	$3t$

$u(a,\omega)$

In reality, can think of 0 as noisy prediction of state ω (e.g., stock trend, purchase prob)

Predict stock trend ←  Financial data

Predict purchase prob ←  Consumer data

The Value of Knowing A Noisy Signal

Posterior belief $p(\omega)$, given ⓪

Prior belief $q(\omega)$

		$1-2\epsilon$	2ϵ
		1/2	1/2
		tail	head
action	0	t	0
	1	0	$3t$

$u(a, \omega)$

action 0 → $(1-2\epsilon)t$

action 1 → $6\epsilon t$

Question: What is the value of this noisy signal ⓪ ?

- Without knowing this signal, DM takes action 1
- With this signal ⓪, DM takes action 0 (assuming ϵ very small)
- However, true distribution is the posterior p regardless

Value of knowing ⓪ = $\mathbb{E}_{\omega\sim p}[u(0,\omega)] - \mathbb{E}_{\omega\sim p} u(1,\omega)$

More Generally: Value of Knowing a Signal

Definition (FK'19). Consider an arbitrary decision making problem $u(a, \omega)$, suppose a signal updates the DM's belief about state ω from $q \in \Delta(\Omega)$ to $p \in \Delta(\omega)$, the value of this signal is defined as

$$D^u(p; q) = \mathbb{E}_{\omega \sim p}[u(a^*(p), \omega)] - \mathbb{E}_{\omega \sim p} u(a^*(q), \omega)$$

Payoff is evaluated based on updated/refined uncertainty distribution

Alexander Frankel and Emir Kamenica, *Quantifying Information and Uncertainty*, American Economic Review 2019.

More Generally: Value of Knowing a Signal

Definition (FK'19). Consider an arbitrary decision making problem $u(a, \omega)$, suppose a signal updates the DM's belief about state ω from $q \in \Delta(\Omega)$ to $p \in \Delta(\omega)$, the value of this signal is defined as

$$D^u(p; q) = \mathbb{E}_{\omega \sim p}[u(a^*(p), \omega)] - \mathbb{E}_{\omega \sim p} u(a^*(q), \omega)$$

Value is generated from
more informed decisions

Alexander Frankel and Emir Kamenica, *Quantifying Information and Uncertainty*, American Economic Review 2019.

More Generally: Value of Knowing a Signal

Definition (FK'19). Consider an arbitrary decision making problem $u(a, \omega)$, suppose a signal updates the DM's belief about state ω from $q \in \Delta(\Omega)$ to $p \in \Delta(\omega)$, the value of this signal is defined as

$$D^u(p; q) = \mathbb{E}_{\omega\sim p}[u(a^*(p), \omega)] - \mathbb{E}_{\omega\sim p} u(a^*(q), \omega)$$

Example 1.

- $a \in A = \Delta(\Omega)$ → action is to pick a distribution over states
- $u(a, \omega) = \log a_\omega$
- Which action $a \in \Delta(\Omega)$ maximizes expected utility $\mathbb{E}_{\omega\sim p}[u(a, \omega)]$?

$$a^*(p) = ?$$

Alexander Frankel and Emir Kamenica, *Quantifying Information and Uncertainty*, American Economic Review 2019.

More Generally: Value of Knowing a Signal

Definition (FK'19). Consider an arbitrary decision making problem $u(a, \omega)$, suppose a signal updates the DM's belief about state ω from $q \in \Delta(\Omega)$ to $p \in \Delta(\omega)$, the value of this signal is defined as

$$D^u(p; q) = \mathbb{E}_{\omega\sim p}[u(a^*(p), \omega)] - \mathbb{E}_{\omega\sim p}u(a^*(q), \omega)$$

Example 1.

- $a \in A = \Delta(\Omega)$ → action is to pick a distribution over states
- $u(a, \omega) = \log a_\omega$
- Which action $a \in \Delta(\Omega)$ maximizes expected utility $\mathbb{E}_{\omega\sim p}[u(a, \omega)]$?

$$a^*(p) = p$$

$$D^u(p; q) = ?$$

Alexander Frankel and Emir Kamenica, *Quantifying Information and Uncertainty*, American Economic Review 2019.

More Generally: Value of Knowing a Signal

Definition (FK'19). Consider an arbitrary decision making problem $u(a, \omega)$, suppose a signal updates the DM's belief about state ω from $q \in \Delta(\Omega)$ to $p \in \Delta(\omega)$, the value of this signal is defined as

$$D^u(p; q) = \mathbb{E}_{\omega\sim p}[u(a^*(p), \omega)] - \mathbb{E}_{\omega\sim p} u(a^*(q), \omega)$$

Example 1.

- $a \in A = \Delta(\Omega)$ → action is to pick a distribution over states
- $u(a, \omega) = \log a_\omega$
- Which action $a \in \Delta(\Omega)$ maximizes expected utility $\mathbb{E}_{\omega\sim p}[u(a, \omega)]$?

$$a^*(p) = p$$

$$D^u(p; q) = \sum_\omega p_\omega \log \frac{p_\omega}{q_\omega}$$

KL-divergence

Alexander Frankel and Emir Kamenica, *Quantifying Information and Uncertainty*, American Economic Review 2019.

More Generally: Value of Knowing a Signal

Definition (FK'19). Consider an arbitrary decision making problem $u(a, \omega)$, suppose a signal updates the DM's belief about state ω from $q \in \Delta(\Omega)$ to $p \in \Delta(\omega)$, the value of this signal is defined as

$$D^u(p;q) = \mathbb{E}_{\omega\sim p}[u(a^*(p), \omega)] - \mathbb{E}_{\omega\sim p} u(a^*(q), \omega)$$

Example 2.

- $a \in A = \Delta(\Omega)$ → action is to pick a distribution over states
- $u(a, \omega) = -||a - e_\omega||^2$

$$a^*(p) = p$$

$$D^u(p;q) = ||p - q||^2$$ Squared distance

Alexander Frankel and Emir Kamenica, *Quantifying Information and Uncertainty*, American Economic Review 2019.

More Generally: Value of Knowing a Signal

Definition (FK'19). Consider an arbitrary decision making problem $u(a, \omega)$, suppose a signal updates the DM's belief about state ω from $q \in \Delta(\Omega)$ to $p \in \Delta(\omega)$, the value of this signal is defined as

$$D^u(p; q) = \mathbb{E}_{\omega \sim p}[u(a^*(p), \omega)] - \mathbb{E}_{\omega \sim p} u(a^*(q), \omega)$$

Some obvious properties

- ✓ Non-negativity: $D^u(p; q) \geq 0$
- ✓ Null information has no value: $D^u(q; q) = 0$
- ✓ Order-invariant: if DM receives signal σ_1, σ_2, the order of receiving them does not affect final expected total value

Alexander Frankel and Emir Kamenica, *Quantifying Information and Uncertainty*, American Economic Review 2019.

What kind of $D(p;q)$ is decision-theoretically grounded?

- ✓ Non-negativity: $D^u(p;q) \geq 0$
- ✓ Null information has no value: $D^u(q;q) = 0$
- ✓ Order-invariant: if DM receives signal σ_1, σ_2, the order of receiving them does not affect final expected total value

Alexander Frankel and Emir Kamenica, *Quantifying Information and Uncertainty*, American Economic Review 2019.

What kind of $D(p;q)$ is decision-theoretically grounded?

In this case, we say $D(p;q)$ is a valid measure for value of information

Theorem 1 (FK'19). Consider any $D(p;q)$ function.
There exists a decision problem $u(a,\omega)$ such that

$$D(p;q) = \mathbb{E}_{\omega\sim p}[u(a^*(p),\omega)] - \mathbb{E}_{\omega\sim p}u(a^*(q),\omega)$$

if and only if $D(p;q)$ satisfies

- ✓ Non-negativity: $D^u(p;q) \geq 0$
- ✓ Null information has no value: $D^u(q;q) = 0$
- ✓ Order-invariant: if DM receives signal σ_1, σ_2, the order of receiving them does not affect final expected total value

Alexander Frankel and Emir Kamenica, *Quantifying Information and Uncertainty*, American Economic Review 2019.

"Equivalence" between Value of Information and Concavity

Theorem 2 (FK'19).
1. For any concave function H, its Bregman divergence is a valid measure for value of information.

2. Conversely, for any valid measure $D(p;q)$ for value of information,
$$H(q) = \sum_\omega q^\omega D(e_\omega, q)$$
is a concave function whose Bregman divergence is $D(p;q)$.

{measures for the value of information}

= {Bregman divergences of concave functions}

Alexander Frankel and Emir Kamenica, *Quantifying Information and Uncertainty*, American Economic Review 2019.

“Equivalence” between Value of Information and Concavity

Theorem 2 (FK’19).
1. For any concave function H, its Bregman divergence is a valid measure for value of information.

2. Conversely, for any valid measure $D(p;q)$ for value of information,

$$H(q) = \sum_{\omega} q^{\omega} D(e_{\omega}, q)$$

is a concave function whose Bregman divergence is $D(p;q)$.

Why useful?

- Many functions – even natural ones like l_2 distance $||p - q||$ – are not valid measures
- In fact, any metric is not valid, since metric cannot be a Bregman divergence
- There are efficient ways to tell whether a $D(p,q)$ is valid

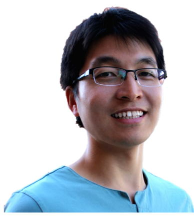
Alexander Frankel and Emir Kamenica, *Quantifying Information and Uncertainty*, American Economic Review 2019.

So far: How to measure the value of information

Next: How to price information based on its economic value

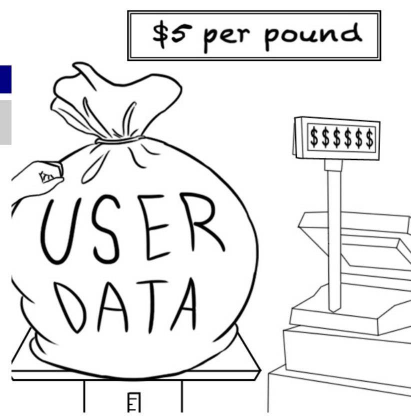
15 mins break

AAAI 2023 Tutorial: Economics of Data and ML


Haifeng Xu (Chicago)


Shuran Zheng (CMU)


James Zou (Stanford)



2/8/2023