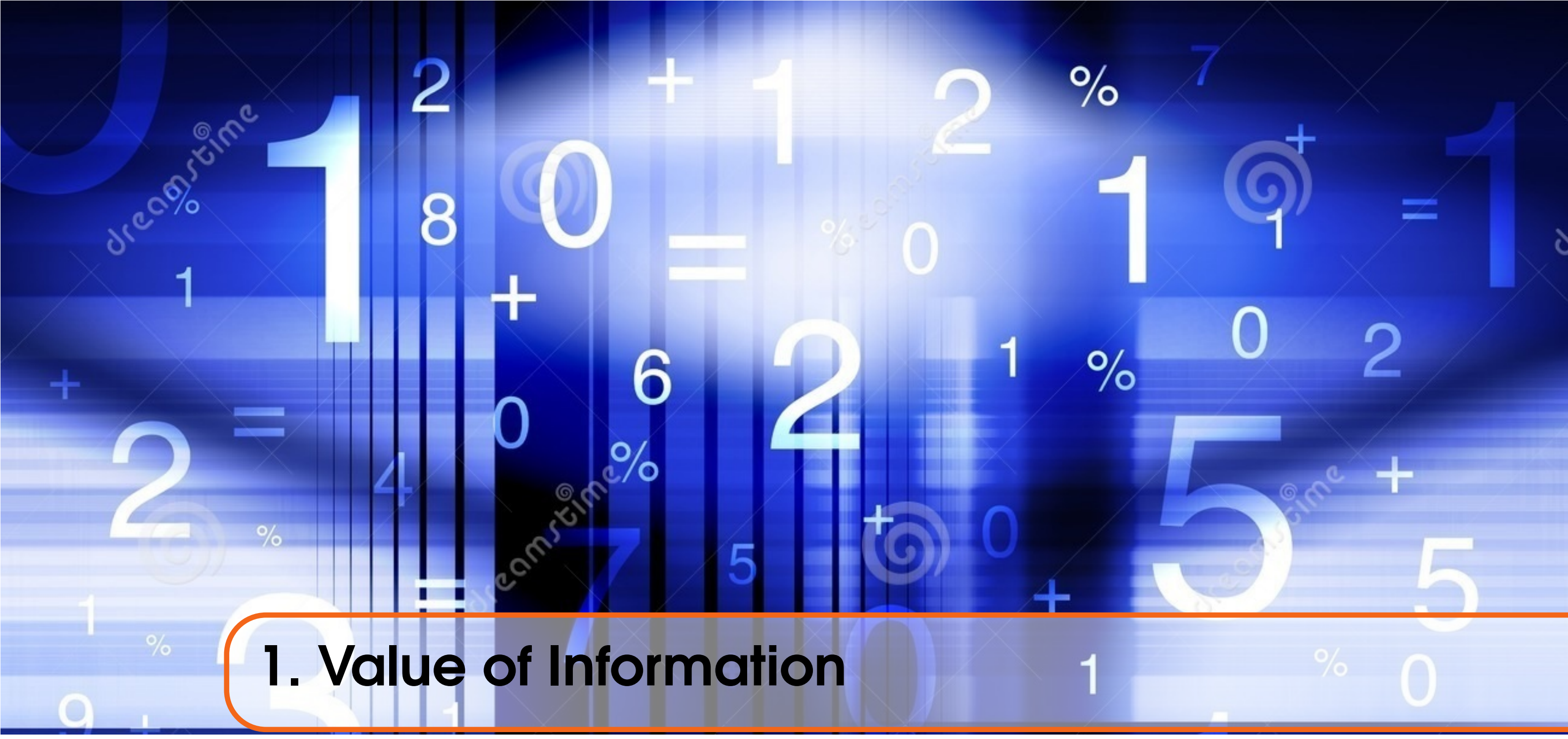
ECON 514: Information and Incentives

Lecture Notes

Jasmine Hao

Índice general

Value of Information



1. Value of Information

Given one experiment $P:[p_{ij}]$, where denotes probability of observing s_i given signal x_j

The matrix of P looks like $\begin{bmatrix} x_1 & x_2 & \dots & x_n(Priors) \\ p_{11} & p_{21} & \dots & p_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1N_1} & p_{2N_1} & \dots & p_{nN_1} \end{bmatrix}$ $Q:[q_{ik}]$, can artificially expand number of signals y_k. e.g, y_{k_j} could be uninformative signals.

Theorem 1.0.1 We say that P more informative than $Q \Leftrightarrow B(P,A) \supset B(Q,A)$.

Intuition: Every point achievable in Q is achievable in P.

Definition 1.0.1 — Sufficiency. $P \succ Q, \exists \quad M_{N_1 \times N_2} \quad s.t. \quad Q = PM. \, q_{ik} = \sum_{j=1}^{N_1} p_{ij} m_{jk} \quad \forall k \in \{1, \dots, N_2\}$

Can generate Q from P.

R Can we say the rank of P is higher than Q Not only, it has to be a Markov matrix.

Theorem 1.0.2 — Blackwell. $P \succ Q \Rightarrow P \supset Q$

Demostración. $P \succ Q \Rightarrow P \supset Q$ Suffice if any point in $B(P,A)$ under decision function f is in $B(P, QA)$.
Consider

$$h(x_j) = \sum m_{jk} f(y_k)$$

. Then

$$
\begin{aligned}
v_i(f)(\in B(Q,A)) &= \sum_1^{N_2} q_{ik} f(y_k) \\
&= \sum_1^{N_2}\sum_1^{N_1} p_{ij} m_{jk} f(y_k) \\
&= \sum_1^{N_1} p_{ij} h(x_j) \\
&= v_i(h)(\in B(P,A))
\end{aligned}
$$

$h(x_j)$ is also a decision function, just a linear transformation of $f(y_k)$.
$P \succ Q \Leftarrow P \supset Q$ If $P \supset Q$, suppose that $Q \neq PM \quad \forall \quad M_{N_1 \times N_2}$.
$\mathscr{K} \equiv \{K = PM \quad | \quad M_{N_1 \times N_2}\}$. Claim: $Q \not\subset \mathscr{K}$.
$\exists \quad U$ s.t. $tr(KU) < tr(QU)$ for all $K \in \mathscr{K}$, contradiction.

R U is a decision rule.

R Add up the trace is the sum of expected payoff (??).

The more "spread out"the posteriors are, the more of DM can adjust decision accordingly. $U_{N_2 \times N_1}, tr(QU) > tr(PMU) \forall m$. implies exist a strategy, such that the payoff under Q is higher than under PM. contradict the fact that $B(P,A) \supset B(Q,A)$. ■

Proposition 1.0.3 P is more informative than $Q \iff \sum_{j=1}^{N_1} \mu_j g(\pi_P) \geq \sum_{j=1}^{N_2} \mu_j g(\pi_Q) \forall g$

In Bayesian learning, uniform prior is very important.

■ **Example 1.0.3.1** State: $n \in \{1,2,3\}$
Signal $X = \{x_1, x_2, x_3\}$
Experiment 1: $\pi_{ji} = \frac{1}{3}, \forall i,j$
Experiment 2: x_1 generate distribution c, x_2, x_3 generate distribution b instead.
Experiment 3: x_1 generate distribution c, x_2 generate distribution d, x_3 generate e. Posterior distribution under experiment 2.

$$
\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \\ 2/3 & 1/6 & 1/6 \end{bmatrix}
$$

Posterior distribution under experiment 3.

$$
\begin{bmatrix} 1/3 & 1/2 & 1/6 \\ 0 & 1/3 & 2/3 \\ 2/3 & 1/6 & 1/6 \end{bmatrix}
$$

The experiment is generated by mean-preserving spread.

Proposition 1.0.4 G is a mean-preserving spread of F if $\exists \tilde{x}, \tilde{y}, \tilde{z}$ with $E[\tilde{z}|\tilde{y}]$. $\tilde{x}$ G, $\tilde{y}$ F.

Definition 1.0.2 G riskier than F. $\iff \int u(x)dG(x) \leq \int u(x)dF(x)$ for all concave function u

R [Mean preserving spread] Claim: DM like more volatile experiment.
Question: Why would anyone love risk?
It's the uncertainty in different space, the volatility in signal space is preferred.

Conclusion: $E3 \succ E2 \succ E1$ if care all three states.

R If only care about outcome 1,3, not 2, experiment 3 no better than 2.
In this case, the first experiment not helpful. If not mean-preserving spread,then not **Blackwell comparable**.

Definition 1.0.3 Y is a garbling of X if $Pr(y_k|s_i) = \sum_j Pr(y_k|x_j)Pr(x_j|s_i)$

Proposition 1.0.5 P is more informative than Q $\iff$ Q are garbling of signal from P

How to garble experiment 3 to get experiment 2 ?
Be careful of signals vs. posteriors.
Garbling signal is a bad thing, garbling posterior is a good thing.
$PM = Q$. First garbling the signal.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

Recover garbling posterior from garbling of signal. Suppose the prior is $\alpha, 1-\alpha$.

$$Q->P\begin{cases} 1/2 & with 1/2 \\ -1/2 & with 1/2 \end{cases}$$

Example 1.0.5.1 $S = \{S_1, S_2\}$. P, Q two experiment. $PM = Q$.

$$\begin{bmatrix} \gamma_1 & 1-\gamma_1 \\ \gamma_2 & 1-\gamma_2 \end{bmatrix} \times \begin{bmatrix} \alpha & 1-\alpha \\ \beta & 1-\beta \end{bmatrix} = \begin{bmatrix} \gamma_1' & 1-\gamma_1' \\ \gamma_2' & 1-\gamma_2' \end{bmatrix}$$

When we solve for $M = \begin{bmatrix} \alpha & 1-\alpha \\ \beta & 1-\beta \end{bmatrix}$, we can say $P \succ Q$.
How to work for posterior, $\pi_{11} = \frac{\gamma_1}{\gamma_1+\gamma_2}$, true state 1, conditional signal 1.
$\pi_{21} = \frac{1-\gamma_1}{2-\gamma_1-\gamma_2}$.

Claim: if P more informative than Q, it must be $\pi_{11}' \in (\pi_{11}, \pi_{12})$.The Q posterior lies in between the P posteriors, $\Rightarrow$ able to construct mean-preserving spread.Similarly, check for all. Actually , can use either methods.

If take a risk-averse and risk-neutral, who is willing to pay more for the information, recall the value of information is $\sum_{j=1}^{N_1} \mu_j g(\pi_j) \geq g(\phi)$.

Example 1.0.5.2 Suppose DM has binary choice $d \in \{0,1\}$
payoff is $\{10,-5\}$.
$s_1 = (0{,}75, 0{,}25)$ and $s_2 = (0{,}25, 0{,}75)$
risk-neutral $\frac{20}{2} + \frac{5}{2} > 10$, always invest without information. With information invest if first posterior. value of information $0{,}75 * (20) * 0{,}5 + (10) * 0{,}5 - 12{,}5 = \frac{105}{8}$.
Use $C(x) = \frac{x^{1-r}}{1-r}$ to represent the risk-adverseness.
If γ is extremely high. can form $CE(\gamma)$ - certainty equivalent.
Call $w(CE(\tilde{x})) = EU(\tilde{x})$ the certainty equivalent of the best outcome.

R The certainty equivalent is decreasing with risk aversion.

Analysis of value of information

- There has to be a certain point of risk aversion, above which the information value is 0. The certainty equivalent of taking the bet without the experiment is lower than 10, agent won't take the experiment, VOI decrease with concavity.
- In the case where the agent choose $d = 1$.

$$
\begin{aligned}
VOI &= \frac{1}{2}(\frac{3}{4}u(20) + \frac{1}{4}u(5)) + \frac{1}{2}u(10) - (\frac{1}{2}u(20) + \frac{1}{2}u(5) \\
&= \frac{1}{2}(\frac{3}{4}u(20) + \frac{1}{4}u(5) - u(10)) - \left(\frac{1}{2}u(20) + \frac{1}{2}u(5) - u(10)\right)
\end{aligned}
$$

With the area where the agent will choose $d = 1$. Both term decrease with concavity , not sure how this perform.

There are two opposing effect, one reduce risk, anti risk and ex-post risk.

R Blackwell comparison is a partial order. Note it might not able to compare all experiments. It require every DM to prefer P to Q.

■ **Example 1.0.5.3 — Location experiment.** $x|s = s + \varepsilon^i$, $E(\varepsilon^i) = 0$.

- $X \sim U[s-\frac{1}{2}, s+\frac{1}{2}], Y \sim U[s-1, s+1]$.
- $\varepsilon^P, \varepsilon^Q$ both normal, $var(\varepsilon^P) < var(\varepsilon^Q)$

Question: Small variance = more informative? ■

Note : $U[-\frac{1}{2}, \frac{1}{2}] + binary(-\frac{1}{2}, \frac{1}{2}, 0{,}5) = U[-1, 1]$. Then can garble X to get Y.

Corollary 1.0.6 For example 1, Note that if $X \sim U[s-\eta, s+\eta], 0 < \eta < 1$ and $Y \sim U[s-1, s+1]$ are not comparable except for $\eta = 1/k, \exists k \in \mathbb{N}$ (cf. Lehmann(1988)) .
For example 2, unless two are normally distributed, otherwise not comparable. (since normal add up to normal ,can garble P(find noise) to achieve another distribution Q).

Summary 1.0.7 Blackwell: more spread out posteriors= more informative Problem: can be too restrictive
What if outside expected utility framework? (cf. Kreps and Porteus (1978))

- Nonlinear in current and future expected utility
- Value of information can be negative? ?
- $U(c_0, \tilde{c}_1) = u_0(c_0) + u_1(v^{-1}Ev(\tilde{c}_1))$
- Prefers early(later) resolution of uncertainty if u is less(more) concave than v

■ **Example 1.0.7.1** $U(c_0, \tilde{c}_1) = u_0(c_0) + u_1(v^{-1}Ev(\tilde{c}_1))$ where $(5, \tilde{c}_1), (5, c_1)$ ■

Read: .^ nonconcavity in the value of information", Radner and Stiglitz (1984) the end