Concave Value of Information in Multiple Decision Problems

Olivier Gossner and Penélope Hernández and Abraham Neyman

Olivier.Gossner@enpc.fr

CERAS, Fédération Jourdan

Universidad d'Alicante

Hebrew University, Jerusalem

General motivation

General motivation

- How are investments into information acquisition taken? Trade off between costs and benefits.

General motivation

- How are investments into information acquisition taken? Trade off between costs and benefits.
- How does this value evolve with the resources invested?

General motivation

- How are investments into information acquisition taken? Trade off between costs and benefits.
- How does this value evolve with the resources invested?
- Increasing, or decreasing returns to scale?

Plan of the talk

Plan of the talk

- Recall a result on nonconcavity of the value of information.

Plan of the talk

- Recall a result on nonconcavity of the value of information.
- Recall a model with concave value of information.

Plan of the talk

- Recall a result on nonconcavity of the value of information.
- Recall a model with concave value of information.
- Study the value of information in a “canonical” model of information acquisition.

Information acquisition before decision

Information acquisition before decision

- Decision maker has to decide $x \in X$.

Information acquisition before decision

- Decision maker has to decide $x \in X$.
- Payoff $g(x, k)$ also depends on some random state of nature k, of known prior distribution.

Information acquisition before decision

- Decision maker has to decide $x \in X$.
- Payoff $g(x, k)$ also depends on some random state of nature k, of known prior distribution.
- A statistical experiment μ gives a probability over signals S for each value of k.

Information acquisition before decision

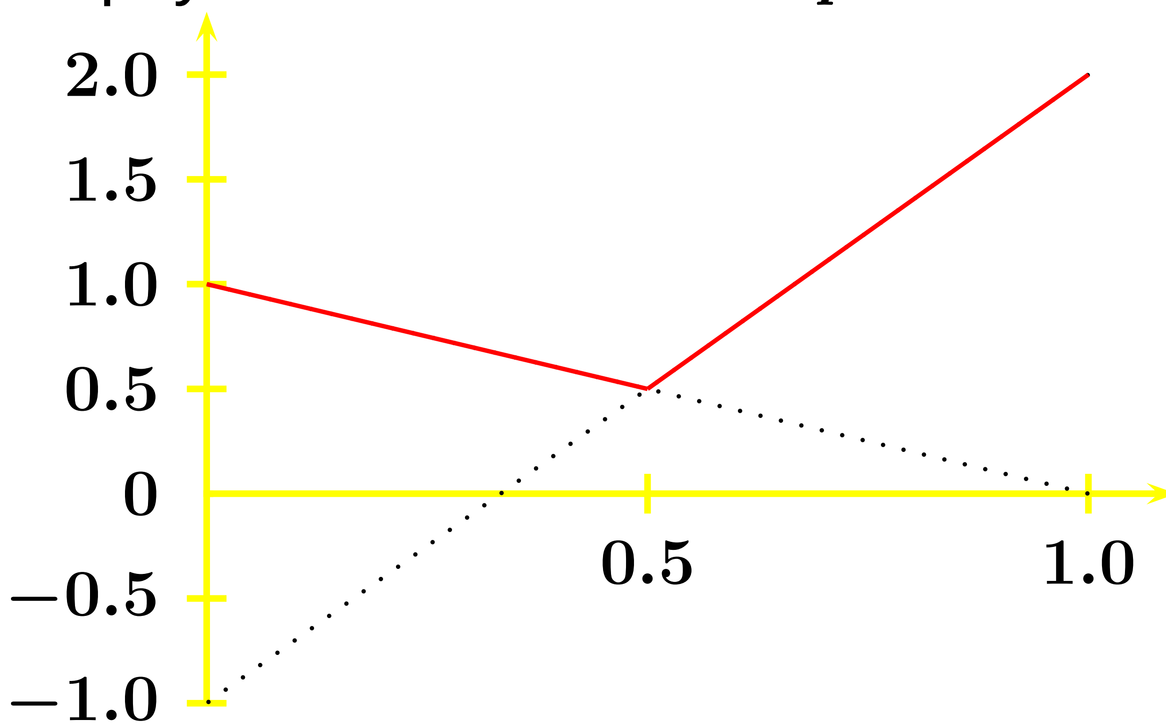
- Decision maker has to decide $x \in X$.
- Payoff $g(x, k)$ also depends on some random state of nature k, of known prior distribution.
- A statistical experiment μ gives a probability over signals S for each value of k.
- The decision maker can choose among statistical experiments prior to decision of x, at a cost that depends on the experiment chosen.

Example of decision problem

	k_1	k_2
x_1	2	-1
x_2	0	1

Prior is $P(k_1) = \frac{1}{3}$.

Expected payoffs as a function of $p =$ *belief in* k_1.



Example of statistical experiments

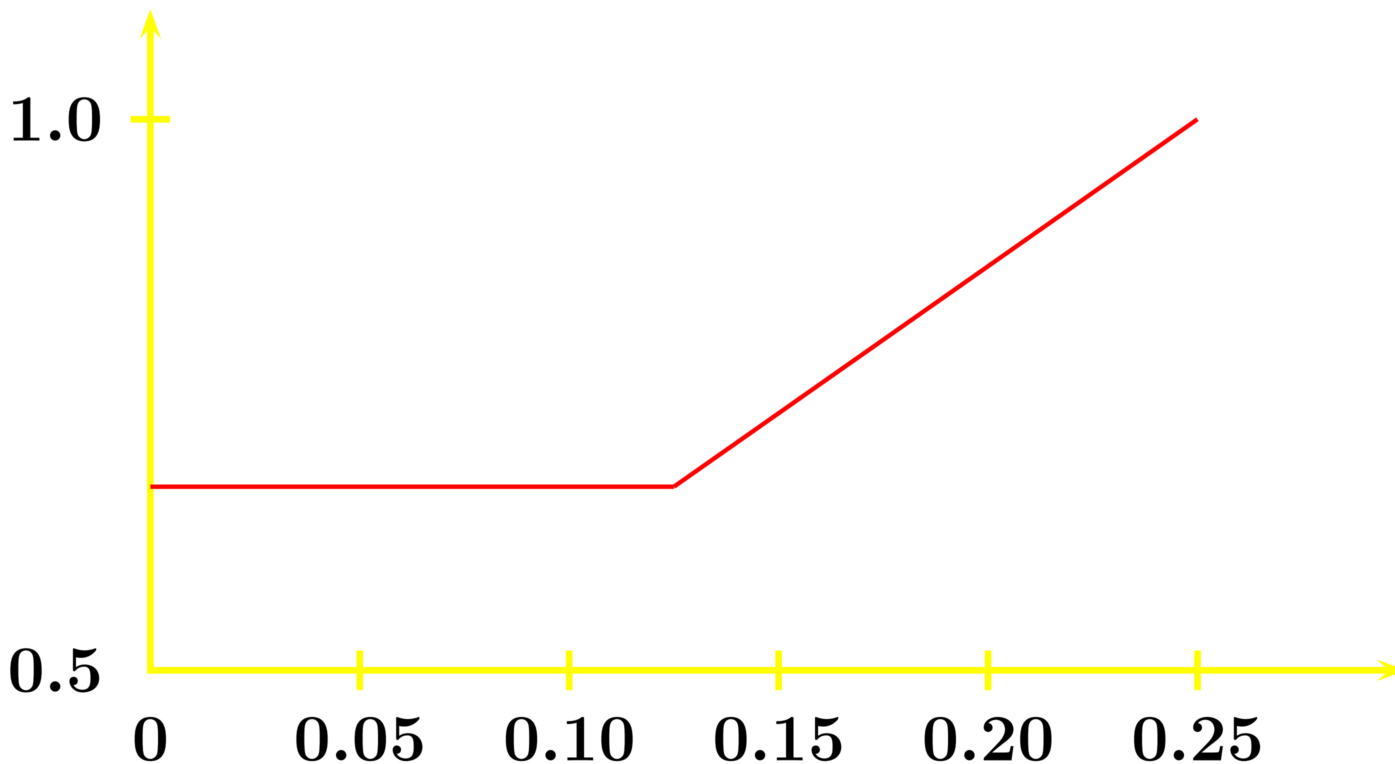
	k_1	k_2
s_1	$\frac{1}{2} + 2\alpha$	$\frac{1}{2} - \alpha$
s_2	$\frac{1}{2} - 2\alpha$	$\frac{1}{2} + \alpha$

$$0 \leq \alpha \leq \frac{1}{4}$$

$$P(s_1) = \tfrac{1}{2},\ P(k_1|s_1) = \tfrac{1}{3} + \tfrac{4}{3}\alpha$$
$$P(s_2) = \tfrac{1}{2},\ P(k_1|s_2) = \tfrac{1}{3} - \tfrac{4}{3}\alpha$$

"Splitting" around original beliefs.

Expected payoff as a function of α



Radner and Stiglitz (1984)

Radner and Stiglitz (1984)

Family of statistical experiments $(\mu_\alpha)_{\alpha>0}$.

Radner and Stiglitz (1984)

Family of statistical experiments $(\mu_\alpha)_{\alpha>0}$.

- $\alpha = 0$ corresponds to no information.

Radner and Stiglitz (1984)

Family of statistical experiments $(\mu_\alpha)_{\alpha>0}$.

- $\alpha = 0$ corresponds to no information.
- Differentiability assumption of μ_α at $\alpha = 0$.

Radner and Stiglitz (1984)

Family of statistical experiments $(\mu_\alpha)_{\alpha>0}$.

- $\alpha = 0$ corresponds to no information.
- Differentiability assumption of μ_α at $\alpha = 0$.

How does the payoff the D.M. can obtain evolve with α?

Non-concavity of value of information

Non-concavity of value of information

- The marginal value of α at $\alpha = 0$ is zero.

Non-concavity of value of information

- The marginal value of α at $\alpha = 0$ is zero.
- If marginal price is positive, there is no interest in buying small amounts of information.

Non-concavity of value of information

- The marginal value of α at $\alpha = 0$ is zero.
- If marginal price is positive, there is no interest in buying small amounts of information.

Value of information is not concave

Comments on the non-concavity

Comments on the non-concavity

- Negative implications:

Comments on the non-concavity

- Negative implications:
  - Demand of information is discontinuous with price.

Comments on the non-concavity

- Negative implications:
  - Demand of information is discontinuous with price.
  - Difficulties in models of information acquisition, in particular strategic experimentation.

Comments on the non-concavity

- Negative implications:
  - Demand of information is discontinuous with price.
  - Difficulties in models of information acquisition, in particular strategic experimentation.
  - No existence of competitive or rational expectation equilibria.

Comments on the non-concavity

- Negative implications:
  - Demand of information is discontinuous with price.
  - Difficulties in models of information acquisition, in particular strategic experimentation.
  - No existence of competitive or rational expectation equilibria.
- Zero marginal value of “small” information not very intuitive.

Comments on the non-concavity

- Negative implications:
  - Demand of information is discontinuous with price.
  - Difficulties in models of information acquisition, in particular strategic experimentation.
  - No existence of competitive or rational expectation equilibria.
- Zero marginal value of “small” information not very intuitive.
- Paradox: If you never agree to pay “for a little more” information, how can you ever agree to pay at all?

Comments on the non-concavity

- Negative implications:
  - Demand of information is discontinuous with price.
  - Difficulties in models of information acquisition, in particular strategic experimentation.
  - No existence of competitive or rational expectation equilibria.
- Zero marginal value of “small” information not very intuitive.
- Paradox: If you never agree to pay “for a little more” information, how can you ever agree to pay at all?
- Little choice in the model about “what” information to obtain.

Arrow (1971, 1985)

Arrow (1971, 1985)

- Investment model

Arrow (1971, 1985)

- Investment model
- $\log$ utilities functions

Arrow (1971, 1985)

- Investment model
- $\log$ utilities functions
- Prior to investment, decision maker can buy a signal s.

Arrow (1971, 1985)

- Investment model
- $\log$ utilities functions
- Prior to investment, decision maker can buy a signal s.
- The price of a signal depends on the *Mutual Information* between s and k

Arrow (1971, 1985)

- Investment model
- $\log$ utilities functions
- Prior to investment, decision maker can buy a signal s.
- The price of a signal depends on the *Mutual Information* between s and k

Obtained value of information is concave

Summary

Summary

Radner and Stiglitz show that for a class of models, value of information is not concave.

Summary

Radner and Stiglitz show that for a class of models, value of information is not concave.

Arrow suggests the use of mutual information, and obtains concave value of information.

Summary

Radner and Stiglitz show that for a class of models, value of information is not concave.

Arrow suggests the use of mutual information, and obtains concave value of information.

Radner and Stiglitz make an assumption, apparently innocuous but not met in Arrow's model.

Open questions

Open questions

- Are Radner and Stiglitz's assumptions "generally" satisfied?

Open questions

- Are Radner and Stiglitz's assumptions "generally" satisfied?
- In Arrow's model:

Open questions

- Are Radner and Stiglitz's assumptions "generally" satisfied?
- In Arrow's model:
  - Is there some fundamental justification for the use of mutual information?

Open questions

- Are Radner and Stiglitz's assumptions “generally” satisfied?
- In Arrow's model:
  - Is there some fundamental justification for the use of mutual information?
  - Can the concavity result be extended to other utility functions?

Model: nature and decisions

Model: nature and decisions

There is a stream of states of nature in K

$$k_\infty = (k_1, \ldots, k_t, \ldots)$$

drawn i.i.d. according to $\nu \in \Delta(K)$.

Model: nature and decisions

There is a stream of states of nature in K

$$k_\infty = (k_1, \ldots, k_t, \ldots)$$

drawn i.i.d. according to $\nu \in \Delta(K)$.

At stage t, the decision maker chooses $x_t \in X$, and gets a payoff

$$g(x_t, k_t)$$

Model: nature and decisions

There is a stream of states of nature in K

$$k_\infty = (k_1, \ldots, k_t, \ldots)$$

drawn i.i.d. according to $\nu \in \Delta(K)$.

At stage t, the decision maker chooses $x_t \in X$, and gets a payoff

$$g(x_t, k_t)$$

Payoffs are discounted, λ discount factor.

Model: information acquisition

Model: information acquisition

D.M. can ask basic questions at any stage to a research group, or to an expert, about states of nature

$$\boldsymbol{Q_0 = \{q_0\colon K^{\mathbb{N}} \to A_0\}}$$

$\boldsymbol{A_0}$ set of basic answers $\boldsymbol{A_0 = \{0, 1\}}$.

Model: information acquisition

D.M. can ask basic questions at any stage to a research group, or to an expert, about states of nature

$$Q_0 = \{q_0 \colon K^{\mathbb{N}} \to A_0\}$$

A_0 set of basic answers $A_0 = \{0, 1\}$.

Many questions can be asked $Q = \cup_{n \le N} Q_0^n$, N "large".

Model: information acquisition

D.M. can ask basic questions at any stage to a research group, or to an expert, about states of nature

$$Q_0 = \{q_0\colon K^{\mathbb{N}} \to A_0\}$$

A_0 set of basic answers $A_0 = \{0, 1\}$.

Many questions can be asked $Q = \cup_{n \leq N} Q_0^n$, N "large".

For $q \in Q$, $q(k_\infty) \in A = \cup_{n \leq N} A_0^n$ is the list of answers.

Model: information acquisition

D.M. can ask basic questions at any stage to a research group, or to an expert, about states of nature

$$Q_0 = \{q_0 \colon K^{\mathbb{N}} \to A_0\}$$

A_0 set of basic answers $A_0 = \{0, 1\}$.

Many questions can be asked $Q = \cup_{n \leq N} Q_0^n$, N “large”.

For $q \in Q$, $q(k_\infty) \in A = \cup_{n \leq N} A_0^n$ is the list of answers.

The cost for asking n questions is $c(n)$, $c \geq 0$, $c(0) = 0$.

Model: information acquisition

D.M. can ask basic questions at any stage to a research group, or to an expert, about states of nature

$$Q_0 = \{q_0 \colon K^{\mathbb{N}} \to A_0\}$$

A_0 set of basic answers $A_0 = \{0, 1\}$.

Many questions can be asked $Q = \cup_{n \leq N} Q_0^n$, N "large".

For $q \in Q$, $q(k_\infty) \in A = \cup_{n \leq N} A_0^n$ is the list of answers.

The cost for asking n questions is $c(n)$, $c \geq 0$, $c(0) = 0$.

Linear cost $c(n) = p.n$, is a particular case.

Model: information acquisition

D.M. can ask basic questions at any stage to a research group, or to an expert, about states of nature

$$Q_0 = \{q_0 \colon K^{\mathbb{N}} \to A_0\}$$

A_0 set of basic answers $A_0 = \{0, 1\}$.

Many questions can be asked $Q = \cup_{n \leq N} Q_0^n$, N "large".

For $q \in Q$, $q(k_\infty) \in A = \cup_{n \leq N} A_0^n$ is the list of answers.

The cost for asking n questions is $c(n)$, $c \geq 0$, $c(0) = 0$.

Linear cost $c(n) = p.n$, is a particular case.

We examine perfect answers first, "noisy" answers later.

Strategies

Strategies

Strategy at stage t is

$$\sigma_t = (\sigma_{t,q}, \sigma_{t,x}) : \{K \times A\}^{t-1} \to Q \times X$$

Strategies

Strategy at stage t is

$$\sigma_t = (\sigma_{t,q}, \sigma_{t,x}) : \{K \times A\}^{t-1} \to Q \times X$$

A strategy, together with k_∞, induces a sequence of decisions $(x_t)_t$, questions $(q_t)_t$, and answers $(a_t)_t$

Strategies

Strategy at stage t is

$$\sigma_t = (\sigma_{t,q}, \sigma_{t,x}) : \{K \times A\}^{t-1} \to Q \times X$$

A strategy, together with k_∞, induces a sequence of decisions $(x_t)_t$, questions $(q_t)_t$, and answers $(a_t)_t$ where

$$\begin{cases} q_{t+1} &= \sigma_{t+1,q}((k_1, a_1), \ldots, (k_t, a_t)) \\ x_{t+1} &= \sigma_{t+1,x}((k_1, a_1), \ldots, (k_t, a_t)) \\ a_{t+1} &= q_{t+1}(k_\infty) \end{cases}$$

Payoffs

Payoffs

Let $l(q_t)$ be the number of questions at stage t.

Payoffs

Let $l(q_t)$ be the number of questions at stage t.
The total discounted payoff for the D.M. is

$$V_\lambda(\sigma) = \mathbf{E} \sum_t \lambda^t (1 - \lambda)(g(x_t, k_t) - c(l(q_t)))$$

Payoffs

Let $l(q_t)$ be the number of questions at stage t.
The total discounted payoff for the D.M. is

$$V_\lambda(\sigma) = \mathbf{E} \sum_t \lambda^t (1-\lambda)(g(x_t, k_t) - c(l(q_t)))$$

The "decision" payoff is

$$U_\lambda(\sigma) = \mathbf{E} \sum_t \lambda^t (1-\lambda) g(x_t, k_t)$$

Payoffs

Let $l(q_t)$ be the number of questions at stage t.
The total discounted payoff for the D.M. is

$$V_\lambda(\sigma) = \mathbf{E}\sum_t \lambda^t(1-\lambda)(g(x_t,k_t) - c(l(q_t)))$$

The "decision" payoff is

$$U_\lambda(\sigma) = \mathbf{E}\sum_t \lambda^t(1-\lambda)g(x_t,k_t)$$

The information cost is

$$\mathbf{E}\sum_t \lambda^t(1-\lambda)c(l(q_t))$$

Payoffs

Payoffs

Uninformed payoff is

$$\underline{v} = \max_x \mathrm{E}_\nu g(x, k)$$

Payoffs

Uninformed payoff is

$$\underline{\boldsymbol{v}} = \max_{\boldsymbol{x}} \mathrm{E}_{\boldsymbol{\nu}} \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{k})$$

Fully informed payoff

$$\overline{\boldsymbol{v}} = \mathrm{E}_{\boldsymbol{\nu}} \max_{\boldsymbol{x}} \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{k})$$

Payoffs

Uninformed payoff is

$$\underline{v} = \max_x \mathrm{E}_\nu g(x, k)$$

Fully informed payoff

$$\overline{v} = \mathrm{E}_\nu \max_x g(x, k)$$

We assume

$$\underline{v} \neq \overline{v}$$

Payoffs

Uninformed payoff is

$$\underline{\boldsymbol{v}} = \max_{\boldsymbol{x}} \mathrm{E}_{\boldsymbol{\nu}} \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{k})$$

Fully informed payoff

$$\overline{\boldsymbol{v}} = \mathrm{E}_{\boldsymbol{\nu}} \max_{\boldsymbol{x}} \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{k})$$

We assume

$$\underline{\boldsymbol{v}} \neq \overline{\boldsymbol{v}}$$

In general

$$\underline{\boldsymbol{v}} \leq \max_{\boldsymbol{\sigma}} \boldsymbol{V}_{\boldsymbol{\lambda}}(\boldsymbol{\sigma}) \leq \overline{\boldsymbol{v}}$$

Study of values and optimal strategies

Study of values and optimal strategies

1. We define the information acquisition rate of a strategy.

Study of values and optimal strategies

1. We define the information acquisition rate of a strategy.
2. We study the best payoff that can be obtained for a fixed information acquisition rate, as a function of this rate.

Study of values and optimal strategies

1. We define the information acquisition rate of a strategy.

2. We study the best payoff that can be obtained for a fixed information acquisition rate, as a function of this rate.

3. We then derive the optimal information acquisition rate.

Acquisition rate of a strategy

Acquisition rate of a strategy

The *acquisition rate* of a strategy σ is

$$\mathbf{E} \sum_t \lambda^t (1 - \lambda) l(q_t)$$

Acquisition rate of a strategy

The *acquisition rate* of a strategy σ is

$$\mathbf{E} \sum_t \lambda^t (1 - \lambda) l(q_t)$$

Σ_r is the set of strategies of acquisition rate r.

Value of information

Value of information

The best payoff with acquisition rate r is

$$v_{\lambda}(r) = \max_{\sigma \in \Sigma_r} U_{\lambda}(\sigma)$$

Value of information

The best payoff with acquisition rate r is

$$v_{\lambda}(r) = \max_{\sigma \in \Sigma_r} U_{\lambda}(\sigma)$$

What can we say about $v_{\lambda}(r)$?

Value of information

The best payoff with acquisition rate r is

$$v_\lambda(r) = \max_{\sigma \in \Sigma_r} U_\lambda(\sigma)$$

What can we say about $v_\lambda(r)$?

Or about $\lim_{\lambda \to 1} v_\lambda(r)$?

Measure of information

Measure of information

X, Y pair of discrete random variables

Measure of information

$\boldsymbol{X, Y}$ pair of discrete random variables

$$\boldsymbol{H(X) = -\sum_x P(x) \log P(x), \log = \log_2, 0 \log 0 = 0}$$

Measure of information

X, Y pair of discrete random variables

$$H(X) = -\sum_x P(x) \log P(x), \log = \log_2, 0 \log 0 = 0$$

$$H(X|Y) = H(X, Y) - H(Y)$$

Mutual Information

Mutual Information

$$H(X) - H(X|Y) = H(Y) - H(Y|X)$$

Mutual Information

$$H(X) - H(X|Y) = H(Y) - H(Y|X)$$

This is the information that one of the r.vs yields about the other one.

Mutual Information

$$\boldsymbol{H}(\boldsymbol{X}) - \boldsymbol{H}(\boldsymbol{X}|\boldsymbol{Y}) = \boldsymbol{H}(\boldsymbol{Y}) - \boldsymbol{H}(\boldsymbol{Y}|\boldsymbol{X})$$

This is the information that one of the r.vs yields about the other one.

It is the *mutual information* $\boldsymbol{I}(\boldsymbol{X}, \boldsymbol{Y})$ between $\boldsymbol{X}$ and $\boldsymbol{Y}$.

Main result

Main result

$$v(r) = \lim_{\lambda \to 1} v_\lambda(r) = \max \mathrm{E} g(x, k)$$

Main result

$$v(r) = \lim_{\lambda \to 1} v_\lambda(r) = \max \mathrm{E} g(x, k)$$

where the maximum is taken over pairs of r.v.'s (x, k) s.t. $k \sim \nu$ and $I(x, k) \le r$.

Main result

$$v(r) = \lim_{\lambda \to 1} v_\lambda(r) = \max \mathrm{E} g(x, k)$$

where the maximum is taken over pairs of r.v.'s (x, k) s.t. $k \sim \nu$ and $I(x, k) \leq r$.

The limit value is thus the value of a one-shot decision problem in which price of information is given by mutual information.

Example

In the decision problem

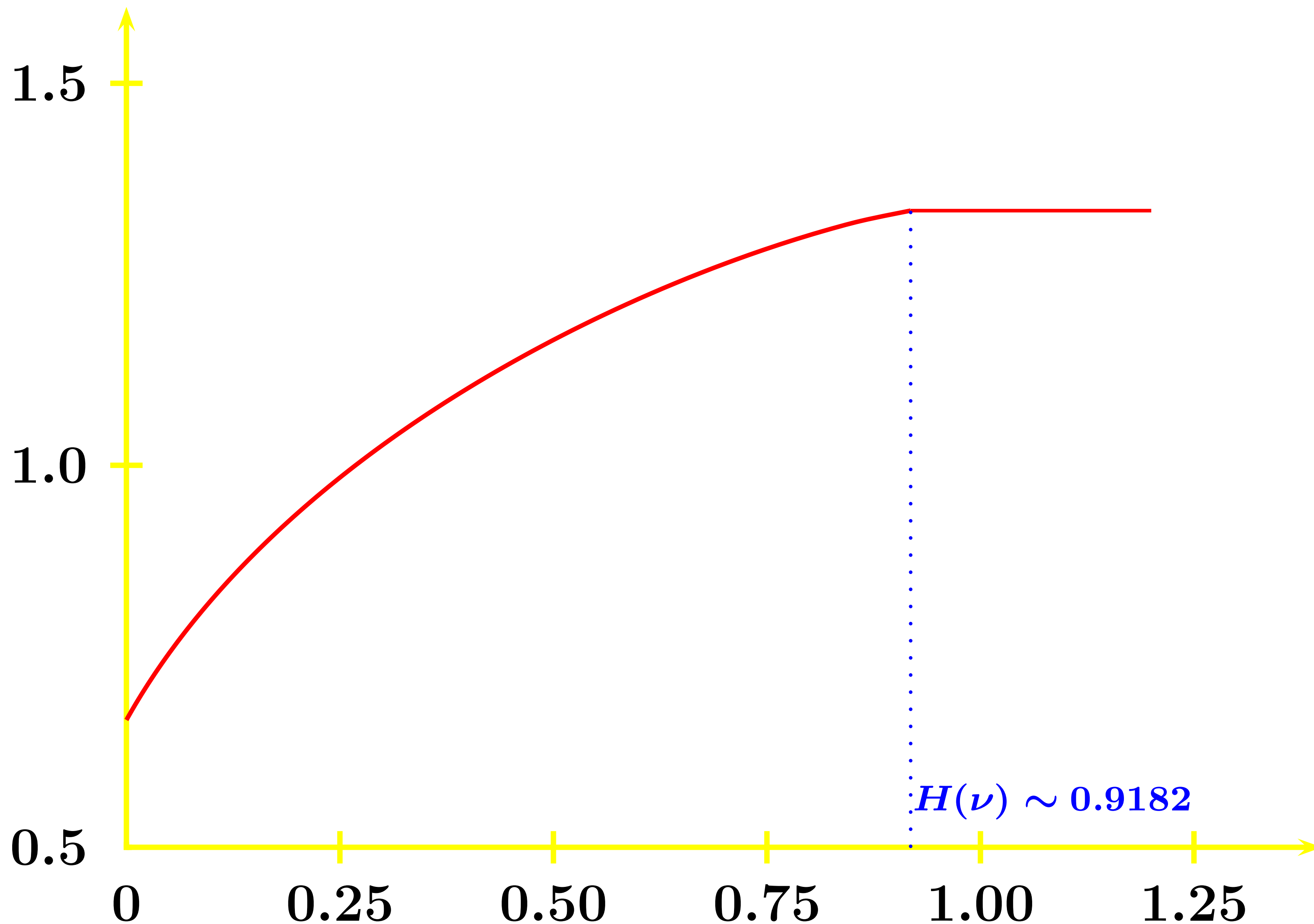
	k_1	k_2
x_1	2	-1
x_2	0	1

experiments that maximize expected payoff and minimize information costs correspond to “splittings”

$$\begin{array}{ll} P(s_1) = \lambda & P(k_1|s_1) = x \\ P(s_2) = 1-\lambda & P(k_1|s_2) = 1-x \end{array}$$

$$\lambda x + (1-\lambda)(1-x) = \frac{1}{3}$$

The derived graph of $v(r)$



Properties of $v(r)$

Properties of $v(r)$

1. $v(r)$ is a non-decreasing, concave function of r.

Properties of $v(r)$

1. $v(r)$ is a non-decreasing, concave function of r.
2. $r_0 = H(\nu)$ is such that $v(r) = \overline{v}$ iff $r \geq r_0$.

Properties of $v(r)$

1. $v(r)$ is a non-decreasing, concave function of r.
2. $r_0 = H(\nu)$ is such that $v(r) = \overline{v}$ iff $r \geq r_0$.
3. $v(r) = \underline{v}$ iff $r = 0$.

Properties of $v(r)$

1. $v(r)$ is a non-decreasing, concave function of r.
2. $r_0 = H(\nu)$ is such that $v(r) = \overline{v}$ iff $r \geq r_0$.
3. $v(r) = \underline{v}$ iff $r = 0$.
4. v is increasing on $[0, r_0]$.

Properties of $v(r)$

1. $v(r)$ is a non-decreasing, concave function of r.
2. $r_0 = H(\nu)$ is such that $v(r) = \overline{v}$ iff $r \geq r_0$.
3. $v(r) = \underline{v}$ iff $r = 0$.
4. v is increasing on $[0, r_0]$.
5. The left derivative of v at r_0 is 0.

Noisy answers: model

Noisy answers: model

The set of basic answers is $\boldsymbol{A_0}$, a basic question is a map

$$\boldsymbol{q_0\colon K^{\mathbb{N}} \longrightarrow A_0}$$

Noise is given by a probability transition $\boldsymbol{N}$ from $\boldsymbol{A_0}$ to $\boldsymbol{B_0}$.

If $\boldsymbol{A_0}$ is the (true) answer, D.M. observes $\boldsymbol{b_0 \in B_0}$ with probability $\boldsymbol{N(a_0)(b_0)}$. Noise is independent between questions.

Noisy answers: model

The set of basic answers is $\boldsymbol{A_0}$, a basic question is a map

$$\boldsymbol{q_0} \colon \boldsymbol{K}^{\mathbb{N}} \longrightarrow \boldsymbol{A_0}$$

Noise is given by a probability transition $\boldsymbol{N}$ from $\boldsymbol{A_0}$ to $\boldsymbol{B_0}$.

If $\boldsymbol{A_0}$ is the (true) answer, D.M. observes $\boldsymbol{b_0} \in \boldsymbol{B_0}$ with probability $\boldsymbol{N(a_0)(b_0)}$. Noise is independent between questions.

Sampling, where the same basic question is asked an arbitrary number of times, is a particular case.

Noisy answers: value of information

Noisy answers: value of information

Let

$$C = \max_{a_0} I(a_0, b_0)$$

C may be larger or smaller than 1.

Noisy answers: value of information

Let

$$C = \max_{a_0} I(a_0, b_0)$$

C may be larger or smaller than 1.

Main Result

$$v(r) = \lim_{\lambda \to 1} v_\lambda(r) = \max \mathrm{E} g(x, k)$$

where $k \sim \nu$ and $I(x, k) \leq Cr$.

Extensions

Extensions

Results are unchanged if

Extensions

Results are unchanged if

- questions may depend on finite histories only

Extensions

Results are unchanged if

- questions may depend on finite histories only
- only a signal on current state of nature is observed, or even no signal.

Extensions

Results are unchanged if

- questions may depend on finite histories only
- only a signal on current state of nature is observed, or even no signal.

Results extend to more general processes, e.g. Markov, with or withour observation of past states of nature.

Demand for information

Demand for information

1. Demand for information is continuous with price of information.

Demand for information

1. Demand for information is continuous with price of information.

2. Demand decreases with price (law of demand).

Demand for information

1. Demand for information is continuous with price of information.

2. Demand decreases with price (law of demand).

3. Demand never reaches r_0, which means that full information never obtains.