The effect of information constraints on decision-making and economic behaviour

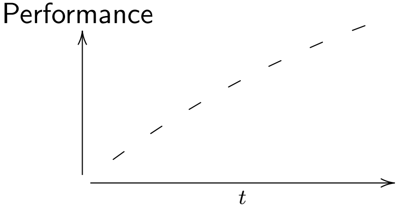
Roman V. Belavkin

Middlesex University

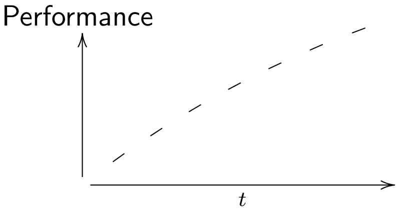
April 16, 2010

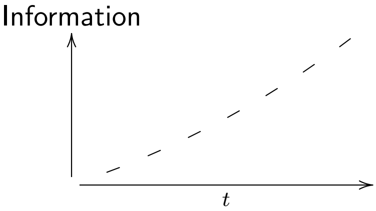
Learning Systems

Learning Systems

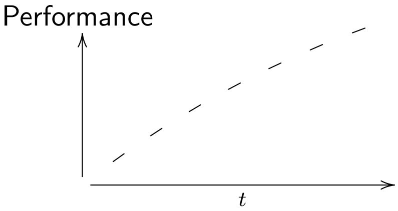


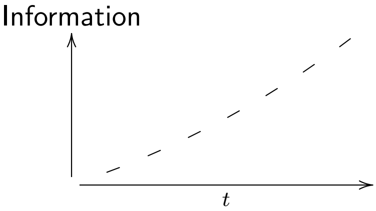
Learning Systems





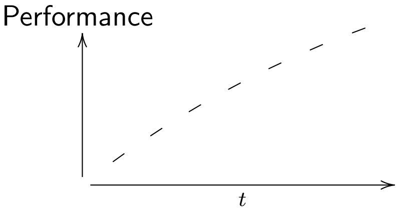
Learning Systems

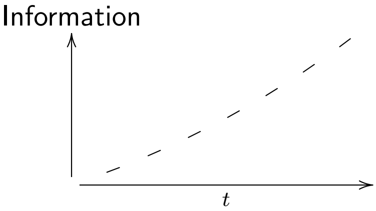




- Performance and information have orders, and the relation between them is monotonic.

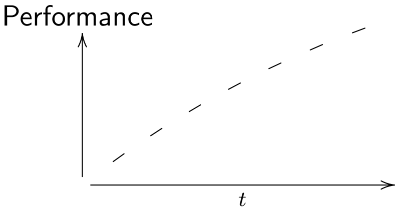
Learning Systems

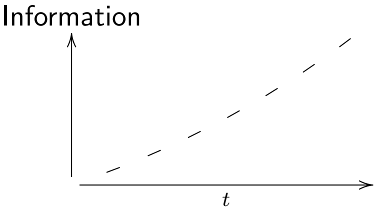




- Performance and information have orders, and the relation between them is monotonic.
- Complete partial orders, domain theory.

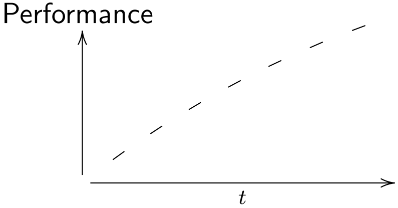
Learning Systems

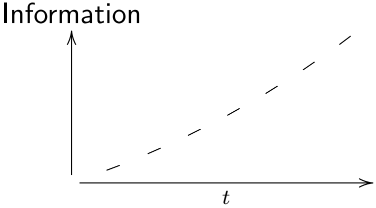




- Performance and information have orders, and the relation between them is monotonic.
- Complete partial orders, domain theory.
- Utility theory, information theory

Learning Systems





- Performance and information have orders, and the relation between them is monotonic.
- Complete partial orders, domain theory.
- Utility theory, information theory
- Allows for treating both deterministic and non-deterministic case:

$$x = f(\omega)\,, \qquad x = f(\omega) + \mathsf{rand}()$$

Expected Utility Theory

- $f : \Omega \to \mathbb{R}$ a utility function.

Expected Utility Theory

- $f : \Omega \to \mathbb{R}$ a utility function.
- $p : \mathfrak{R} \to [0, 1]$ a probability measure on $(\Omega, \mathfrak{R})$.

Expected Utility Theory

- $f : \Omega \to \mathbb{R}$ a utility function.
- $p : \mathfrak{R} \to [0, 1]$ a probability measure on $(\Omega, \mathfrak{R})$.
- The expected utility

$$\mathbb{E}_p\{x\} := \sum_{\omega \in \Omega} x(\omega)p(\omega)$$

Expected Utility Theory

- $f : \Omega \to \mathbb{R}$ a utility function.
- $p : \mathfrak{R} \to [0, 1]$ a probability measure on $(\Omega, \mathfrak{R})$.
- The expected utility

$$\mathbb{E}_p\{x\} := \int_\Omega x(\omega)\, dp(\omega)$$

- Choice under uncertainty

$$q \lesssim p \quad \iff \quad \mathbb{E}_q\{f\} \leq \mathbb{E}_p\{f\}$$

Expected Utility Theory

- $f : \Omega \to \mathbb{R}$ a utility function.
- $p : \mathfrak{R} \to [0, 1]$ a probability measure on $(\Omega, \mathfrak{R})$.
- The expected utility

$$\mathbb{E}_p\{x\} := \int_\Omega x(\omega)\, dp(\omega)$$

- Choice under uncertainty

$$q \lesssim p \quad \iff \quad \mathbb{E}_q\{f\} \leq \mathbb{E}_p\{f\}$$

Question (Why expected utility?)

1. $\mathbb{E}_y\{f\} = f(\omega)$ *if* $y(\Omega') = \delta_\omega(\Omega')$.
2. $x \lesssim y \quad \iff \quad \lambda x \lesssim \lambda y$, $\forall\, \lambda > 0$
3. $x \lesssim y \quad \iff \quad x + z \lesssim y + z$, $\forall\, z \in Y$.

St. Petersburg lottery

Due to Nicolas Bernoulli (1713)

- The lottery is played by tossing a fair coin repeatedly until the first head appears.

St. Petersburg lottery

Due to Nicolas Bernoulli (1713)

- The lottery is played by tossing a fair coin repeatedly until the first head appears.
- If the head appeared on nth toss, then you win $\pounds 2^n$.

St. Petersburg lottery

Due to Nicolas Bernoulli (1713)

- The lottery is played by tossing a fair coin repeatedly until the first head appears.
- If the head appeared on nth toss, then you win $£2^n$.
- For example

$$\begin{aligned} n &= 3 & £2^3 &= £8 \\ n &= 4 & £2^4 &= £16 \\ & \cdots & & \\ n &= 20 & £2^{20} &= £1,048,576 \end{aligned}$$

St. Petersburg lottery

Due to Nicolas Bernoulli (1713)

- The lottery is played by tossing a fair coin repeatedly until the first head appears.
- If the head appeared on nth toss, then you win $\pounds 2^n$.
- For example

$$\begin{aligned} n = 3 &\qquad \pounds 2^3 = \pounds 8 \\ n = 4 &\qquad \pounds 2^4 = \pounds 16 \\ &\cdots \\ n = 20 &\qquad \pounds 2^{20} = \pounds 1,048,576 \end{aligned}$$

- To enter the lottery, you must pay a fee of $\pounds X$

St. Petersburg lottery

Due to Nicolas Bernoulli (1713)

- The lottery is played by tossing a fair coin repeatedly until the first head appears.
- If the head appeared on nth toss, then you win $\pounds 2^n$.
- For example

$$\begin{aligned} n &= 3 & \pounds 2^3 &= \pounds 8 \\ n &= 4 & \pounds 2^4 &= \pounds 16 \\ & & \dots & \\ n &= 20 & \pounds 2^{20} &= \pounds 1,048,576 \end{aligned}$$

- To enter the lottery, you must pay a fee of $\pounds X$
- How much is $\pounds X$?

Why is it a paradox?

- We don't want to pay more than expect to win:

$$X \leq \mathbb{E}\{\text{win}\}$$

Why is it a paradox?

- We don't want to pay more than expect to win:

$$X \leq \mathbb{E}\{\text{win}\}$$

- How large is $\mathbb{E}_p\{\text{win}\}$?

Why is it a paradox?

- We don't want to pay more than expect to win:

$$X \leq \mathbb{E}\{\text{win}\}$$

- How large is $\mathbb{E}_p\{\text{win}\}$?
- Let $p(n)$ be the probability of $n \in \mathbb{N}$

head	1	2	3	4	$\cdots$	n	$\cdots$
win	£2	£4	£8	£16	$\cdots$	2^n	$\cdots$
$p(n)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\cdots$	$\frac{1}{2^n}$	$\cdots$

Why is it a paradox?

- We don't want to pay more than expect to win:

$$X \leq \mathbb{E}\{\text{win}\}$$

- How large is $\mathbb{E}_p\{\text{win}\}$?
- Let $p(n)$ be the probability of $n \in \mathbb{N}$

head	1	2	3	4	$\cdots$	n	$\cdots$
win	£2	£4	£8	£16	$\cdots$	2^n	$\cdots$
$p(n)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\cdots$	$\frac{1}{2^n}$	$\cdots$

- It is easy to see that

$$\mathbb{E}_p\{\text{win}\} = 2 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} + 8 \cdot \frac{1}{8} + 16 \cdot \frac{1}{16} + \cdots = \sum_{n=1}^{\infty} \frac{2^n}{2^n} = \infty$$

Why is it a paradox?

- We don't want to pay more than expect to win:

$$X \leq \mathbb{E}\{\text{win}\}$$

- How large is $\mathbb{E}_p\{\text{win}\}$?
- Let $p(n)$ be the probability of $n \in \mathbb{N}$

head	1	2	3	4	$\cdots$	n	$\cdots$
win	£2	£4	£8	£16	$\cdots$	2^n	$\cdots$
$p(n)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\cdots$	$\frac{1}{2^n}$	$\cdots$

- It is easy to see that

$$\mathbb{E}_p\{\text{win}\} = 2 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} + 8 \cdot \frac{1}{8} + 16 \cdot \frac{1}{16} + \cdots = \sum_{n=1}^{\infty} \frac{2^n}{2^n} = \infty$$

- One cannot buy what is not for sale.

Classical solutions

- Daniel Bernoulli (1738) proposed $f(n) = \log_2 2^n = n$.

Classical solutions

- Daniel Bernoulli (1738) proposed $f(n) = \log_2 2^n = n$.
- Note that for any $f(n)$ we can introduce a lottery $p(n) \propto f^{-1}(n)$:

$$\mathbb{E}_p\{f\} \propto \sum_{n=1}^{\infty} \frac{f(n)}{f(n)} = \infty$$

Classical solutions

- Daniel Bernoulli (1738) proposed $f(n) = \log_2 2^n = n$.
- Note that for any $f(n)$ we can introduce a lottery $p(n) \propto f^{-1}(n)$:

$$\mathbb{E}_p\{f\} \propto \sum_{n=1}^{\infty} \frac{f(n)}{f(n)} = \infty$$

- Some suggest to use only f such that

$$\|f\|_\infty := \sup |f(\omega)| < \infty$$

Northern Rock lottery

Due to unknown author (2008)

- You can borrow a mortgage of any amount $\pounds X$

Northern Rock lottery

Due to unknown author (2008)

- You can borrow a mortgage of any amount $£X$
- The amount you repay is decided by tossing a fair coin repeatedly until the first head appears.

Northern Rock lottery

Due to unknown author (2008)

- You can borrow a mortgage of any amount $£X$
- The amount you repay is decided by tossing a fair coin repeatedly until the first head appears.
- If the head appeared on nth toss, then you repay $£2^n$.

Northern Rock lottery

Due to unknown author (2008)

- You can borrow a mortgage of any amount $£X$
- The amount you repay is decided by tossing a fair coin repeatedly until the first head appears.
- If the head appeared on nth toss, then you repay $£2^n$.
- For example

$$\begin{array}{rl} n = 3 & £2^3 = £8 \\ n = 4 & £2^4 = £16 \\ & \dots \\ n = 20 & £2^{20} = £1,048,576 \end{array}$$

Northern Rock lottery

Due to unknown author (2008)

- You can borrow a mortgage of any amount $\pounds X$
- The amount you repay is decided by tossing a fair coin repeatedly until the first head appears.
- If the head appeared on nth toss, then you repay $\pounds 2^n$.
- For example

$$\begin{array}{rl} n = 3 & \pounds 2^3 = \pounds 8 \\ n = 4 & \pounds 2^4 = \pounds 16 \\ & \dots \\ n = 20 & \pounds 2^{20} = \pounds 1,048,576 \end{array}$$

- How much would you borrow? ($\pounds X =$?)

The Allais (1953) paradox

Consider two lotteries:

A : $p(\pounds 300) = \frac{1}{3}$ (and $p(\pounds 0) = \frac{2}{3}$)

B : $p(\pounds 100) = 1$

The Allais (1953) paradox

Consider two lotteries:

$$\begin{aligned} \text{A} &: p(\pounds 300) = \tfrac{1}{3} \qquad \text{(and } p(\pounds 0) = \tfrac{2}{3}\text{)} \\ \text{B} &: p(\pounds 100) = 1 \end{aligned}$$

- Most of the people seem to prefer $A \lesssim B$

The Allais (1953) paradox

Consider two lotteries:

$$\text{A} \; : \; p(\pounds 300) = \tfrac{1}{3} \qquad (\text{and } p(\pounds 0) = \tfrac{2}{3})$$
$$\text{B} \; : \; p(\pounds 100) = 1$$

- Most of the people seem to prefer $A \lesssim B$
- Note that

$$\begin{aligned} \mathbb{E}_A\{x\} &= 300 \cdot \frac{1}{3} + 100 \cdot 0 + 0 \cdot \frac{2}{3} = 100 \\ \mathbb{E}_B\{x\} &= 300 \cdot 0 + 100 \cdot 1 + 0 \cdot 0 = 100 \end{aligned}$$

The Allais (1953) paradox

Consider two lotteries:

A : $p(£300) = \frac{1}{3}$ (and $p(£0) = \frac{2}{3}$)

B : $p(£100) = 1$

- Most of the people seem to prefer $A \lesssim B$
- Note that

$$\mathbb{E}_A\{x\} = 300 \cdot \frac{1}{3} + 100 \cdot 0 + 0 \cdot \frac{2}{3} = 100$$
$$\mathbb{E}_B\{x\} = 300 \cdot 0 + 100 \cdot 1 + 0 \cdot 0 = 100$$

Remark

Safety is preferred (i.e. risk averse).

The Allais (1953) paradox (2)

Consider two lotteries:

$$\text{C} : p(-£300) = \tfrac{1}{3} \qquad (\text{and } p(£0) = \tfrac{2}{3})$$
$$\text{D} : p(-£100) = 1$$

The Allais (1953) paradox (2)

Consider two lotteries:

C : $p(-\pounds 300) = \frac{1}{3}$ (and $p(\pounds 0) = \frac{2}{3}$)

D : $p(-\pounds 100) = 1$

- Most of the people seem to prefer $C \gtrsim D$

The Allais (1953) paradox (2)

Consider two lotteries:

$$\begin{aligned} C &: p(-£300) = \tfrac{1}{3} \qquad (\text{and } p(£0) = \tfrac{2}{3}) \\ D &: p(-£100) = 1 \end{aligned}$$

- Most of the people seem to prefer $C \gtrsim D$
- Note that

$$\begin{aligned} \mathbb{E}_C\{x\} &= -300 \cdot \frac{1}{3} - 100 \cdot 0 - 0 \cdot \frac{2}{3} = -100 \\ \mathbb{E}_D\{x\} &= -300 \cdot 0 - 100 \cdot 1 - 0 \cdot 0 = -100 \end{aligned}$$

The Allais (1953) paradox (2)

Consider two lotteries:

$$\text{C} : p(-\pounds 300) = \tfrac{1}{3} \qquad (\text{and } p(\pounds 0) = \tfrac{2}{3})$$
$$\text{D} : p(-\pounds 100) = 1$$

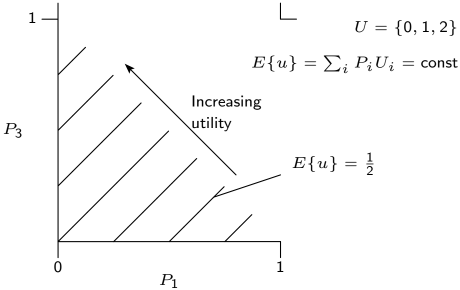
- Most of the people seem to prefer $C \gtrsim D$
- Note that

$$\begin{aligned} \mathbb{E}_C\{x\} &= -300 \cdot \frac{1}{3} - 100 \cdot 0 - 0 \cdot \frac{2}{3} = -100 \\ \mathbb{E}_D\{x\} &= -300 \cdot 0 - 100 \cdot 1 - 0 \cdot 0 = -100 \end{aligned}$$

Remark

Risk is preferred (i.e. risk taking).

Why is it a paradox?



Remark

Any linear functional (e.g. $\mathbb{E}_p\{x\}$) has parallel level sets. If people use expected utility to make choices, then they are either risk-averse or risk-taking, but not both.

Prospect theory

Due to Tversky and Kahneman (1981)

- It was proposed that the utility is convex, when the choice is among gains, and concave when the choice is among losses.

Prospect theory

Due to Tversky and Kahneman (1981)

- It was proposed that the utility is convex, when the choice is among gains, and concave when the choice is among losses.
- This would make the choice risk averse for gains and risk taking for losses.

Prospect theory

Due to Tversky and Kahneman (1981)

- It was proposed that the utility is convex, when the choice is among gains, and concave when the choice is among losses.
- This would make the choice risk averse for gains and risk taking for losses.

Remark

This theory is not normative (i.e. it is not derived using rational approach).

The Ellsberg (1961) paradox

Consider two lotteries:

$$\text{A} : p(£100) = \tfrac{1}{2} \qquad (\text{and } p(£0) = \tfrac{1}{2})$$

$$\text{B} : p(£100) = \text{unknown}$$

The Ellsberg (1961) paradox

Consider two lotteries:

$$\text{A} : p(£100) = \tfrac{1}{2} \qquad \text{(and } p(£0) = \tfrac{1}{2}\text{)}$$
$$\text{B} : p(£100) = \text{unknown}$$

- Most of the people seem to prefer $A \gtrsim B$

The Ellsberg (1961) paradox

Consider two lotteries:

$$\begin{aligned} \text{A} &: p(\pounds 100) = \tfrac{1}{2} \qquad \text{(and } p(\pounds 0) = \tfrac{1}{2}\text{)} \\ \text{B} &: p(\pounds 100) = \text{unknown} \end{aligned}$$

- Most of the people seem to prefer $A \gtrsim B$
- Note that

$$\begin{aligned} \mathbb{E}_A\{x\} &= 100 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = 50 \\ \mathbb{E}_B\{x\} &= \int_0^1 (100 \cdot p + 0 \cdot (1 - p))\, dp = 50 \end{aligned}$$

The Ellsberg (1961) paradox

Consider two lotteries:

$$\begin{aligned} A &: p(\pounds 100) = \tfrac{1}{2} \qquad (\text{and } p(\pounds 0) = \tfrac{1}{2}) \\ B &: p(\pounds 100) = \text{unknown} \end{aligned}$$

- Most of the people seem to prefer $A \gtrsim B$
- Note that

$$\begin{aligned} \mathbb{E}_A\{x\} &= 100 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = 50 \\ \mathbb{E}_B\{x\} &= \int_0^1 (100 \cdot p + 0 \cdot (1-p))\, dp = 50 \end{aligned}$$

Remark

More information is preferred.

Extreme Value Problems

Unconditional extremum

- Maximise $f(y)$:

$$\sup f(y)$$

Extreme Value Problems

Unconditional extremum

- Maximise $f(y)$:

$$\sup f(y)$$

- Necessary condition $\partial f(\bar{y}) \ni 0$.

Extreme Value Problems

Unconditional extremum

- Maximise $f(y)$:

$$\sup f(y)$$

- Necessary condition $\partial f(\bar{y}) \ni 0$.
- Sufficient, if f is concave.

Extreme Value Problems

Unconditional extremum

- Maximise $f(y)$:

$$\sup f(y)$$

- Necessary condition $\partial f(\bar{y}) \ni 0$.
- Sufficient, if f is concave.

Conditional extremum

- Maximise $f(y)$ subject to $g(y) \leq \lambda$:

$$\overline{f}(\lambda) := \sup\{f(y) : g(y) \leq \lambda\}$$

.

Extreme Value Problems

Unconditional extremum

- Maximise $f(y)$:
$$\sup f(y)$$
- Necessary condition $\partial f(\bar{y}) \ni 0$.
- Sufficient, if f is concave.

Conditional extremum

- Maximise $f(y)$ subject to $g(y) \leq \lambda$:
$$\overline{f}(\lambda) := \sup\{f(y) : g(y) \leq \lambda\}$$
.
- Necessary condition $\partial f(\bar{y}) - \alpha \partial g(y) \ni 0$.

Extreme Value Problems

Unconditional extremum

- Maximise $f(y)$:

$$\sup f(y)$$

- Necessary condition $\partial f(\bar{y}) \ni 0$.
- Sufficient, if f is concave.

Conditional extremum

- Maximise $f(y)$ subject to $g(y) \leq \lambda$:

$$\overline{f}(\lambda) := \sup\{f(y) : g(y) \leq \lambda\}$$

.

- Necessary condition $\partial f(\bar{y}) - \alpha \partial g(y) \ni 0$.
- Sufficient if $K(y, \alpha) := f(y) + \alpha[\lambda - g(y)]$ is concave.

Representation in Paired Spaces

- $x \in X$, $y \in Y$, $\langle \cdot, \cdot \rangle : X \times Y \to \mathbb{R}$

$$\langle x, y \rangle := \int_\Omega x(\omega)\, dy(\omega)$$

Representation in Paired Spaces

- $x \in X$, $y \in Y$, $\langle \cdot, \cdot \rangle : X \times Y \to \mathbb{R}$

$$\langle x, y \rangle := \int_\Omega x(\omega)\, dy(\omega)$$

- Separation:

$$\langle x, y \rangle = 0\,, \quad \forall\, x \in X \quad \Rightarrow \quad y = 0 \in Y$$
$$\langle x, y \rangle = 0\,, \quad \forall\, y \in Y \quad \Rightarrow \quad x = 0 \in X$$

Representation in Paired Spaces

- $x \in X$, $y \in Y$, $\langle \cdot, \cdot \rangle : X \times Y \to \mathbb{R}$

$$\langle x, y \rangle := \int_\Omega x(\omega)\, dy(\omega)$$

- Separation:

$$\langle x, y \rangle = 0\,, \quad \forall\, x \in X \quad \Rightarrow \quad y = 0 \in Y$$
$$\langle x, y \rangle = 0\,, \quad \forall\, y \in Y \quad \Rightarrow \quad x = 0 \in X$$

- Can equip X and Y with $\|x\|_\infty$ and $\|y\|_1$.

Representation in Paired Spaces

- $x \in X$, $y \in Y$, $\langle \cdot, \cdot \rangle : X \times Y \to \mathbb{R}$

$$\langle x, y \rangle := \int_\Omega x(\omega)\, dy(\omega)$$

- Separation:

$$\begin{aligned} \langle x, y \rangle = 0\,, \quad \forall\, x \in X \quad &\Rightarrow \quad y = 0 \in Y \\ \langle x, y \rangle = 0\,, \quad \forall\, y \in Y \quad &\Rightarrow \quad x = 0 \in X \end{aligned}$$

- Can equip X and Y with $\|x\|_\infty$ and $\|y\|_1$.
- Statistical manifold:

$$\mathcal{M} := \{y \in Y : y \geq 0\,,\ \|y\|_1 = 1\}$$

Representation in Paired Spaces

- $x \in X$, $y \in Y$, $\langle \cdot, \cdot \rangle : X \times Y \to \mathbb{R}$

$$\langle x, y \rangle := \int_\Omega x(\omega)\, dy(\omega)$$

- Separation:

$$\langle x, y \rangle = 0\,, \quad \forall\, x \in X \quad \Rightarrow \quad y = 0 \in Y$$
$$\langle x, y \rangle = 0\,, \quad \forall\, y \in Y \quad \Rightarrow \quad x = 0 \in X$$

- Can equip X and Y with $\|x\|_\infty$ and $\|y\|_1$.
- Statistical manifold:

$$\mathcal{M} := \{y \in Y : y \geq 0\,,\ \|y\|_1 = 1\}$$

- Expected value

$$\mathbb{E}_p\{x\} = \langle x, y \rangle|_{\mathcal{M}}$$

Information

Definition (Information resource)

a closed functional $F : Y \to \mathbb{R} \cup \{\infty\}$ with $\inf F = F(z)$.

Information

Definition (Information resource)

a closed functional $F : Y \to \mathbb{R} \cup \{\infty\}$ with $\inf F = F(z)$.

Example (Relative Information (Belavkin, 2010b))

- For $z > 0$, let

$$F(y) := \begin{cases} \left\langle \ln \frac{y}{z}, y \right\rangle - \langle 1, y - z \rangle \,, & \text{if } y > 0 \\ \langle 1, z \rangle \,, & \text{if } y = 0 \\ \infty \,, & \text{if } y < 0 \end{cases}$$

Information

Definition (Information resource)

a closed functional $F : Y \to \mathbb{R} \cup \{\infty\}$ with $\inf F = F(z)$.

Example (Relative Information (Belavkin, 2010b))

- For $z > 0$, let

$$F(y) := \begin{cases} \left\langle \ln \frac{y}{z}, y \right\rangle - \langle 1, y - z \rangle \,, & \text{if } y > 0 \\ \langle 1, z \rangle \,, & \text{if } y = 0 \\ \infty \,, & \text{if } y < 0 \end{cases}$$

- $\partial F(y) = \ln \frac{y}{z} = x \quad \iff \quad y = e^x \, z = \partial F^*(x)$

Information

Definition (Information resource)

a closed functional $F : Y \to \mathbb{R} \cup \{\infty\}$ with $\inf F = F(z)$.

Example (Relative Information (Belavkin, 2010b))

- For $z > 0$, let

$$F(y) := \begin{cases} \left\langle \ln \frac{y}{z}, y \right\rangle - \langle 1, y - z \rangle \,, & \text{if } y > 0 \\ \langle 1, z \rangle \,, & \text{if } y = 0 \\ \infty \,, & \text{if } y < 0 \end{cases}$$

- $\partial F(y) = \ln \frac{y}{z} = x \quad \iff \quad y = e^x \, z = \partial F^*(x)$
- The dual $F^* : X \to \mathbb{R} \cup \{\infty\}$ is

$$F^*(x) := \langle 1, e^x \, z \rangle$$

Utility of Information

- If $x \in X$ is utility, then the value of event y relative to z is

$$\langle x, y - z \rangle = \mathbb{E}_y\{x\} - \mathbb{E}_z\{x\}$$

Utility of Information

- If $x \in X$ is utility, then the value of event y relative to z is

$$\langle x, y - z \rangle = \mathbb{E}_y\{x\} - \mathbb{E}_z\{x\}$$

Definition (Utility of information)

$$U_x(I) := \sup\{\langle x, y \rangle : F(y) \leq I\}$$

Utility of Information

- If $x \in X$ is utility, then the value of event y relative to z is

$$\langle x, y - z \rangle = \mathbb{E}_y\{x\} - \mathbb{E}_z\{x\}$$

Definition (Utility of information)

$$U_x(I) := \sup\{\langle x, y \rangle : F(y) \leq I\}$$

- Stratonovich (1965) defined $U_x(I)$ for Shannon information.

Utility of Information

- If $x \in X$ is utility, then the value of event y relative to z is

$$\langle x, y - z \rangle = \mathbb{E}_y\{x\} - \mathbb{E}_z\{x\}$$

Definition (Utility of information)

$$U_x(I) := \sup\{\langle x, y \rangle : F(y) \leq I\}$$

- Stratonovich (1965) defined $U_x(I)$ for Shannon information.
- Related functions

$$\begin{aligned} -U_{-x}(I) &:= \inf\{\langle x, y \rangle : F(y) \leq I\} \\ I_x(U) &:= \inf\{F(y) : U_0 \leq U \leq \langle x, y \rangle\} \\ I_x(U) &:= \inf\{F(y) : \langle x, y \rangle \leq U < U_0\} \end{aligned}$$

Information Bounded Utility

Definition (Information Bounded Utility)

A function $f : \Omega \to \mathbb{R}$ that admits a solution to the utility of information problem $U_f(I)$ for $I \in (\inf F, \sup F)$

Information Bounded Utility

Definition (Information Bounded Utility)

A function $f : \Omega \to \mathbb{R}$ that admits a solution to the utility of information problem $U_f(I)$ for $I \in (\inf F, \sup F)$

Theorem

A solution to $U_f(I)$ and $I_f(U)$ exists if and only if set $\{x : F_q^*(x) \leq I^*\}$ absorbs function f:

$$\exists\, \beta^{-1} > 0 : F_q^*(\beta f) < \infty$$

Information Bounded Utility

Definition (Information Bounded Utility)

A function $f : \Omega \to \mathbb{R}$ that admits a solution to the utility of information problem $U_f(I)$ for $I \in (\inf F, \sup F)$

Theorem

A solution to $U_f(I)$ and $I_f(U)$ exists if and only if set $\{x : F_q^*(x) \leq I^*\}$ absorbs function f:

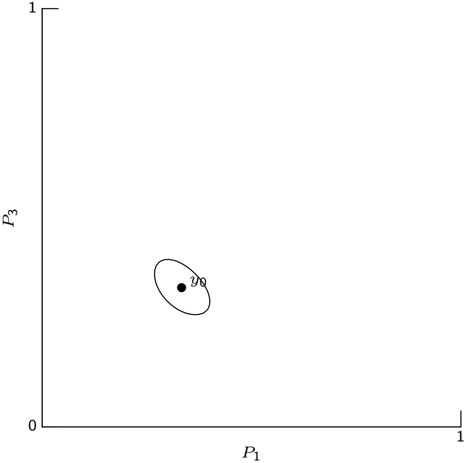
$$\exists\, \beta^{-1} > 0 : F_q^*(\beta f) < \infty$$

Remark (Separation of information)

For all $I \in (\inf F, \sup F)$ there exist β_1^{-1}, $\beta_2^{-1} > 0$:

$$F_q(\partial F_q^*(\beta_1 f)) < I < F_q(\partial F^*(\beta_2 f))$$

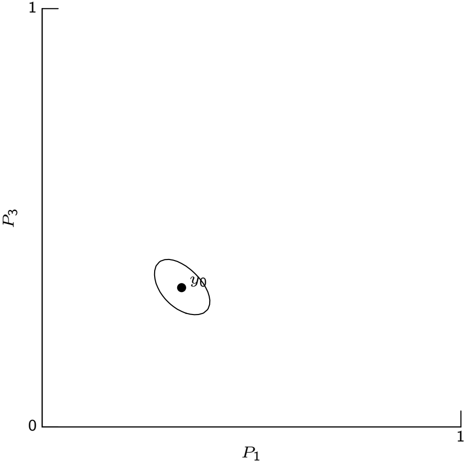
Information Topology (Belavkin, 2010a)



- The topology is defined using an information resource:

$$F : L \to \mathbb{R} \cup \{\infty\}\,, \quad \inf F = F(y_0)$$

Information Topology (Belavkin, 2010a)



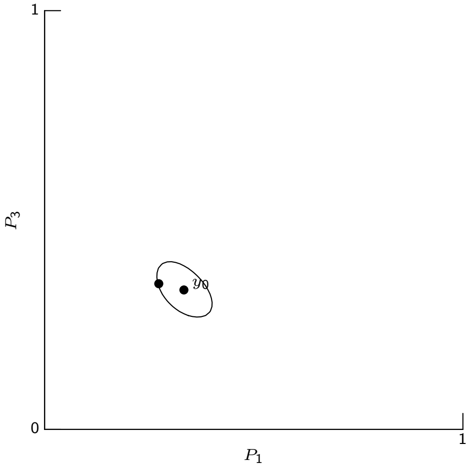
- The topology is defined using an information resource:

$$F : L \to \mathbb{R} \cup \{\infty\}\,, \quad \inf F = F(y_0)$$

- Neighbourhoods of y_0:

$$C := \{y : F(y) \leq I\}$$

Information Topology (Belavkin, 2010a)



- The topology is defined using an information resource:

$$F : L \to \mathbb{R} \cup \{\infty\}\,, \quad \inf F = F(y_0)$$

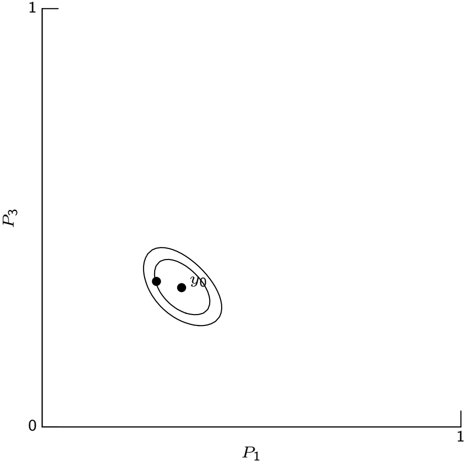
- Neighbourhoods of y_0:

$$C := \{y : F(y) \leq I\}$$

- The topology of information bounded functions:

$$C^\circ := \{x : F^*(x) \leq I^*\}$$

Information Topology (Belavkin, 2010a)



- The topology is defined using an information resource:

$$F : L \to \mathbb{R} \cup \{\infty\}\,, \quad \inf F = F(y_0)$$

- Neighbourhoods of y_0:

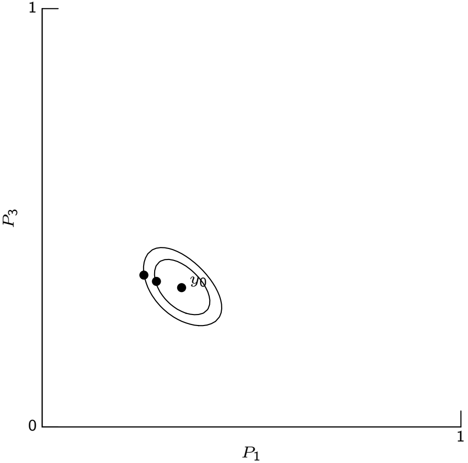
$$C := \{y : F(y) \leq I\}$$

- The topology of information bounded functions:

$$C^\circ := \{x : F^*(x) \leq I^*\}$$

- Generally, $f \in C^\circ$ does not imply $-f \in C^\circ$.

Information Topology (Belavkin, 2010a)



- The topology is defined using an information resource:

$$F : L \to \mathbb{R} \cup \{\infty\}\,, \quad \inf F = F(y_0)$$

- Neighbourhoods of y_0:

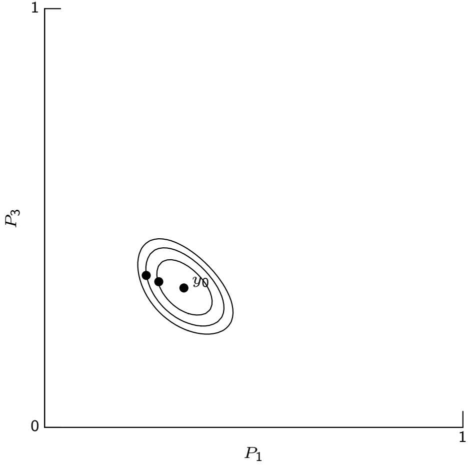
$$C := \{y : F(y) \leq I\}$$

- The topology of information bounded functions:

$$C^\circ := \{x : F^*(x) \leq I^*\}$$

- Generally, $f \in C^\circ$ does not imply $-f \in C^\circ$.

Information Topology (Belavkin, 2010a)



- The topology is defined using an information resource:

$$F : L \to \mathbb{R} \cup \{\infty\}\,, \quad \inf F = F(y_0)$$

- Neighbourhoods of y_0:

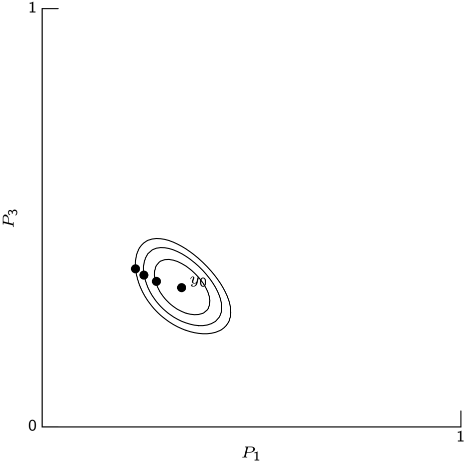
$$C := \{y : F(y) \leq I\}$$

- The topology of information bounded functions:

$$C^\circ := \{x : F^*(x) \leq I^*\}$$

- Generally, $f \in C^\circ$ does not imply $-f \in C^\circ$.

Information Topology (Belavkin, 2010a)



- The topology is defined using an information resource:

$$F : L \to \mathbb{R} \cup \{\infty\}\,, \quad \inf F = F(y_0)$$

- Neighbourhoods of y_0:

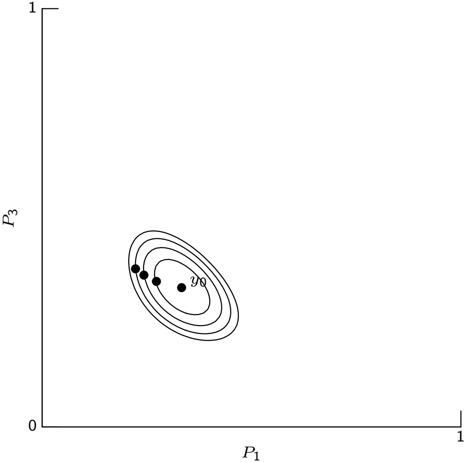
$$C := \{y : F(y) \leq I\}$$

- The topology of information bounded functions:

$$C^{\circ} := \{x : F^*(x) \leq I^*\}$$

- Generally, $f \in C^{\circ}$ does not imply $-f \in C^{\circ}$.

Information Topology (Belavkin, 2010a)



- The topology is defined using an information resource:

$$F : L \to \mathbb{R} \cup \{\infty\}\,, \quad \inf F = F(y_0)$$

- Neighbourhoods of y_0:

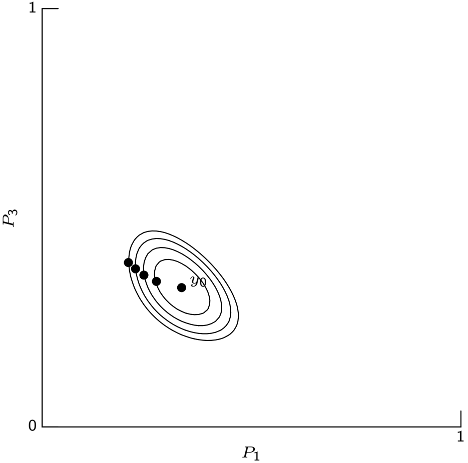
$$C := \{y : F(y) \leq I\}$$

- The topology of information bounded functions:

$$C^\circ := \{x : F^*(x) \leq I^*\}$$

- Generally, $f \in C^\circ$ does not imply $-f \in C^\circ$.

Information Topology (Belavkin, 2010a)



- The topology is defined using an information resource:

$$F : L \to \mathbb{R} \cup \{\infty\} , \quad \inf F = F(y_0)$$

- Neighbourhoods of y_0:

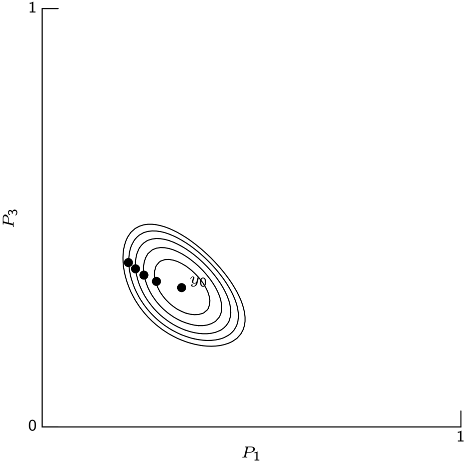
$$C := \{y : F(y) \leq I\}$$

- The topology of information bounded functions:

$$C^\circ := \{x : F^*(x) \leq I^*\}$$

- Generally, $f \in C^\circ$ does not imply $-f \in C^\circ$.

Information Topology (Belavkin, 2010a)



- The topology is defined using an information resource:

$$F : L \to \mathbb{R} \cup \{\infty\}\,, \quad \inf F = F(y_0)$$

- Neighbourhoods of y_0:

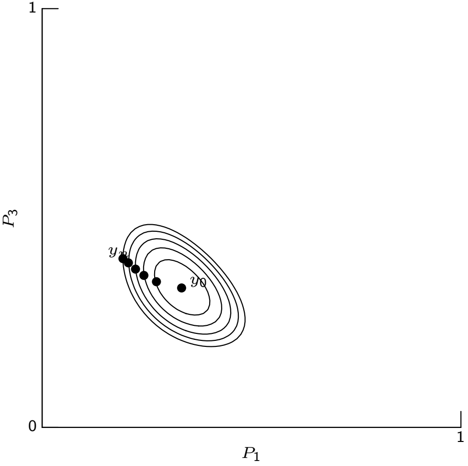
$$C := \{y : F(y) \leq I\}$$

- The topology of information bounded functions:

$$C^\circ := \{x : F^*(x) \leq I^*\}$$

- Generally, $f \in C^\circ$ does not imply $-f \in C^\circ$.

Information Topology (Belavkin, 2010a)



- The topology is defined using an information resource:

$$F : L \to \mathbb{R} \cup \{\infty\}\,, \quad \inf F = F(y_0)$$

- Neighbourhoods of y_0:

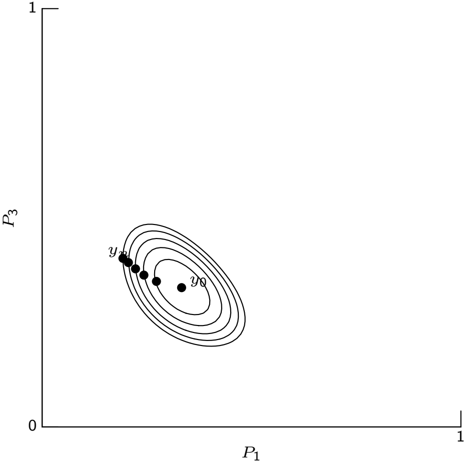
$$C := \{y : F(y) \leq I\}$$

- The topology of information bounded functions:

$$C^\circ := \{x : F^*(x) \leq I^*\}$$

- Generally, $f \in C^\circ$ does not imply $-f \in C^\circ$.

Information Topology (Belavkin, 2010a)



- The topology is defined using an information resource:

$$F : L \to \mathbb{R} \cup \{\infty\}\,, \quad \inf F = F(y_0)$$

- Neighbourhoods of y_0:

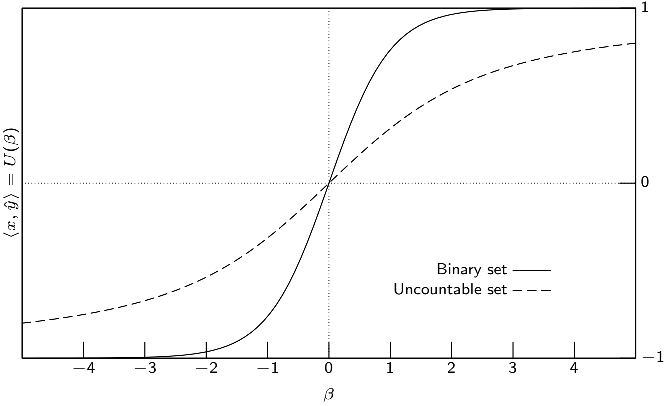
$$C := \{y : F(y) \leq I\}$$

- The topology of information bounded functions:

$$C^\circ := \{x : F^*(x) \leq I^*\}$$

- Generally, $f \in C^\circ$ does not imply $-f \in C^\circ$.

Parametrisation by the Expected Utility

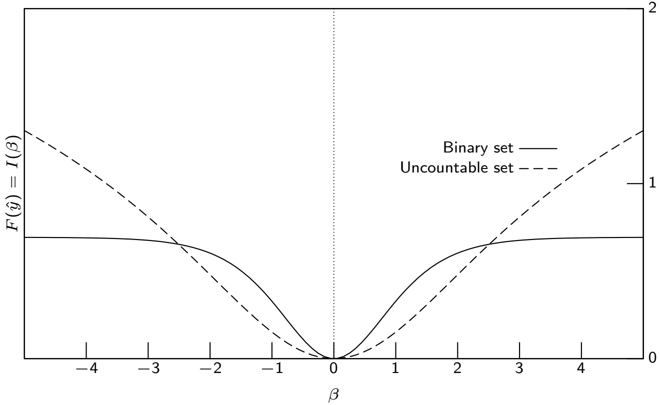


Let $F(y)$ be negative entropy (i.e. $F(y)$ is minimised at $y_0(\omega) = \text{const}$)

$$x : \Omega \to \{c - d, c + d\} \qquad U(\beta) = \Psi'(\beta) = c + d \tanh(\beta\, d)$$

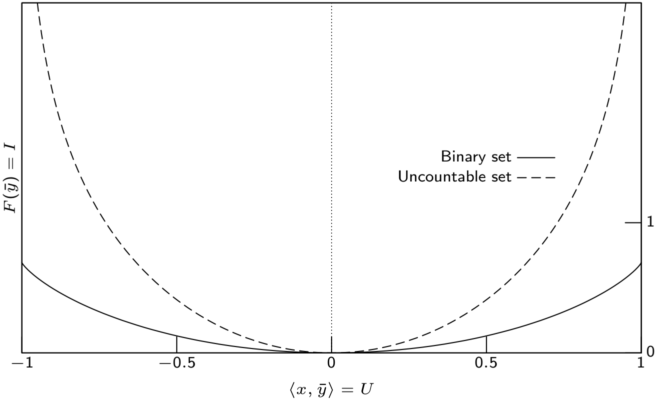
$$x : \Omega \to [c - d, c + d] \qquad U(\beta) = \Psi'(\beta) = c + d \coth(\beta\, d) - \beta^{-1}$$

Parametrisation by Information



$$I = \Phi'(\beta^{-1})\,, \qquad I = \beta\,\Psi'(\beta) - \Psi(\beta)$$

Parametric Dependency



A Solution of the SPB Paradox

- $f : \mathbb{N} \to \mathbb{R}$ is information bounded iff for some $\beta^{-1} > 0$:

$$F^*(\beta f) = \sum_{n=1}^{\infty} q(n) e^{\beta f(n)} < \infty$$

A Solution of the SPB Paradox

- $f : \mathbb{N} \to \mathbb{R}$ is information bounded iff for some $\beta^{-1} > 0$:

$$F^*(\beta f) = \sum_{n=1}^{\infty} q(n) e^{\beta f(n)} < \infty$$

- Using $\sum_{n=1}^{\infty} r^n < \infty$

$$\exists\, \beta^{-1} > 0 : \quad \beta f(n) < -\ln z(n)\,, \quad \forall\, n \in \mathbb{N}$$

A Solution of the SPB Paradox

- $f : \mathbb{N} \to \mathbb{R}$ is information bounded iff for some $\beta^{-1} > 0$:

$$F^*(\beta f) = \sum_{n=1}^{\infty} q(n) e^{\beta f(n)} < \infty$$

- Using $\sum_{n=1}^{\infty} r^n < \infty$

$$\exists\, \beta^{-1} > 0 : \quad \beta f(n) < -\ln z(n)\,, \quad \forall\, n \in \mathbb{N}$$

- where $z(n) = q(n) \sum q(n)$.

A Solution of the SPB Paradox

- $f : \mathbb{N} \to \mathbb{R}$ is information bounded iff for some $\beta^{-1} > 0$:

$$F^*(\beta f) = \sum_{n=1}^{\infty} q(n) e^{\beta f(n)} < \infty$$

- Using $\sum_{n=1}^{\infty} r^n < \infty$

$$\exists\, \beta^{-1} > 0 : \quad \beta f(n) < -\ln z(n)\,, \quad \forall\, n \in \mathbb{N}$$

- where $z(n) = q(n) \sum q(n)$.
- Let $q(n) = (e-1)e^{-n}$ (i.e. 2^{-n}).

A Solution of the SPB Paradox

- $f : \mathbb{N} \to \mathbb{R}$ is information bounded iff for some $\beta^{-1} > 0$:

$$F^*(\beta f) = (e - 1) \sum_{n=1}^{\infty} e^{\beta f(n) - n} < \infty$$

- Using $\sum_{n=1}^{\infty} r^n < \infty$

$$\exists\, \beta^{-1} > 0 : \quad \beta f(n) < -\ln z(n) \,, \quad \forall\, n \in \mathbb{N}$$

- where $z(n) = q(n) \sum q(n)$.
- Let $q(n) = (e - 1)e^{-n}$ (i.e. 2^{-n}).

A Solution of the SPB Paradox

- $f : \mathbb{N} \to \mathbb{R}$ is information bounded iff for some $\beta^{-1} > 0$:

$$F^*(\beta f) = (e-1)\sum_{n=1}^{\infty} e^{\beta f(n) - n} < \infty$$

- Using $\sum_{n=1}^{\infty} r^n < \infty$

$$\exists\, \beta^{-1} > 0 : \quad \beta f(n) < n\,, \quad \forall\, n \in \mathbb{N}$$

- where $z(n) = q(n) \sum q(n)$.
- Let $q(n) = (e-1)e^{-n}$ (i.e. 2^{-n}).

A Solution of the SPB Paradox

- $f : \mathbb{N} \to \mathbb{R}$ is information bounded iff for some $\beta^{-1} > 0$:

$$F^*(\beta f) = (e-1)\sum_{n=1}^{\infty} e^{\beta f(n) - n} < \infty$$

- Using $\sum_{n=1}^{\infty} r^n < \infty$

$$\exists\, \beta^{-1} > 0 : \quad \beta f(n) < n\,, \quad \forall\, n \in \mathbb{N}$$

- where $z(n) = q(n) \sum q(n)$.
- Let $q(n) = (e-1)e^{-n}$ (i.e. 2^{-n}).
- For $f(n) = n$, we have $\beta < 1$.

A Solution of the SPB Paradox

- Using $U = \Psi'_f(\beta)$ obtain

$$U = \frac{1}{1 - e^{\beta - 1}}\,, \qquad U_0 = \frac{1}{1 - e^{-1}}$$

A Solution of the SPB Paradox

- Using $U = \Psi'_f(\beta)$ obtain

$$U = \frac{1}{1 - e^{\beta - 1}}\,, \qquad U_0 = \frac{1}{1 - e^{-1}}$$

- The inverse of function $U(\beta)$ is $\beta = 1 + \ln(1 - U^{-1})$.

A Solution of the SPB Paradox

- Using $U = \Psi'_f(\beta)$ obtain

$$U = \frac{1}{1 - e^{\beta-1}}\,, \qquad U_0 = \frac{1}{1 - e^{-1}}$$

- The inverse of function $U(\beta)$ is $\beta = 1 + \ln(1 - U^{-1})$.
- Using $I = \beta(\ln\Psi(\beta))' - \ln\Psi(\beta)$:

$$I_f(U) = (1 + \ln(1 - U^{-1}))U - \ln(e - 1) - \ln(U - 1)$$

A Solution of the SPB Paradox

- Using $U = \Psi'_f(\beta)$ obtain

$$U = \frac{1}{1 - e^{\beta - 1}}\,, \qquad U_0 = \frac{1}{1 - e^{-1}}$$

- The inverse of function $U(\beta)$ is $\beta = 1 + \ln(1 - U^{-1})$.
- Using $I = \beta(\ln\Psi(\beta))' - \ln\Psi(\beta)$:

$$I_f(U) = (1 + \ln(1 - U^{-1}))U - \ln(e - 1) - \ln(U - 1)$$

- Change e to 2 (ln to $\log_2$).

A Solution of the SPB Paradox

- Using $U = \Psi'_f(\beta)$ obtain

$$U = \frac{1}{1 - 2^{\beta-1}}\,, \qquad U_0 = 2$$

- The inverse of function $U(\beta)$ is $\beta = 1 + \ln(1 - U^{-1})$.
- Using $I = \beta(\ln\Psi(\beta))' - \ln\Psi(\beta)$:

$$I_f(U) = (1 + \log_2(1 - U^{-1}))U - \log_2(U - 1)$$

- Change e to 2 (ln to $\log_2$).

A Solution of the SPB Paradox

- Using $U = \Psi'_f(\beta)$ obtain

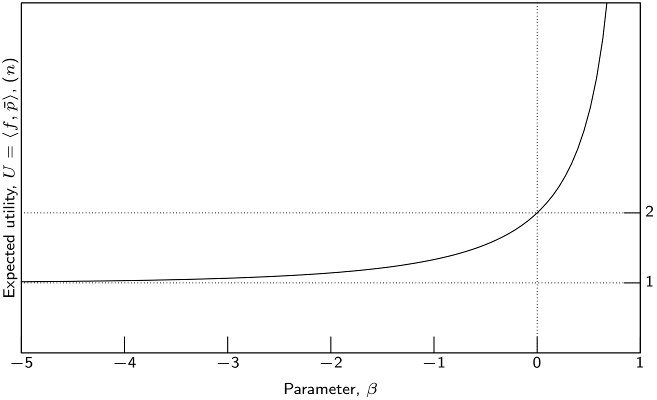
$$U = \frac{1}{1 - 2^{\beta - 1}}\,, \qquad U_0 = 2$$

- The inverse of function $U(\beta)$ is $\beta = 1 + \ln(1 - U^{-1})$.
- Using $I = \beta(\ln \Psi(\beta))' - \ln \Psi(\beta)$:

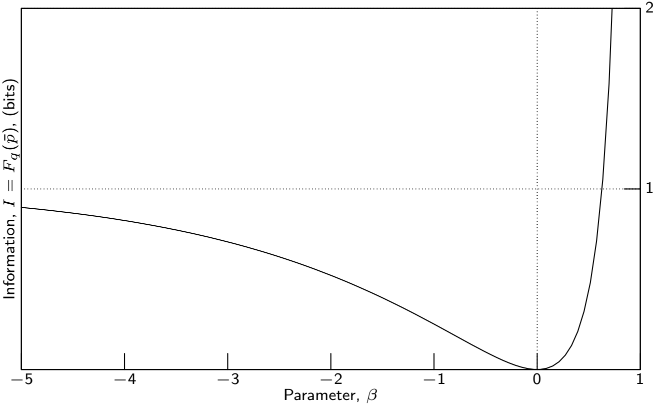
$$I_f(U) = (1 + \log_2(1 - U^{-1}))U - \log_2(U - 1)$$

- Change e to 2 (ln to $\log_2$).
- For the information amount of 0 bits, the optimal entrance fee is $c \leq U_0 = 2$.

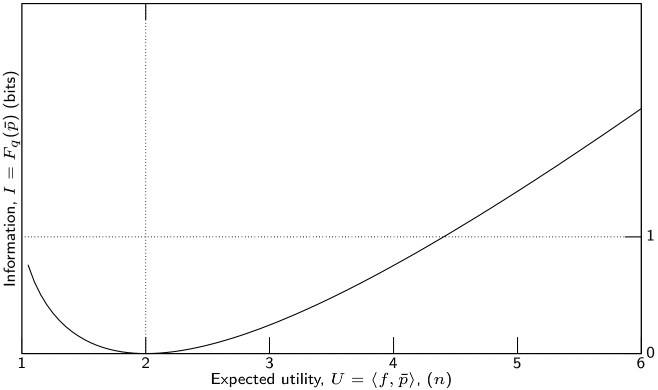
SPB Lottery


Expected utility, $U = \langle f, \bar{p} \rangle$, (n)
Parameter, β
−5
−4
−3
−2
−1
0
1
2
1

SPB Lottery



SPB Lottery


Information, $I = F_q(\bar{p})$ (bits)
1
0
1
2
3
4
5
6
Expected utility, $U = \langle f, \bar{p} \rangle$, (n)

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'École americaine. *Econometrica*, *21*, 503–546.

Belavkin, R. V. (2010a). Information trajectory of optimal learning. In M. J. Hirsch, P. M. Pardalos, & R. Murphey (Eds.), *Dynamics of information systems: Theory and applications* (Vol. 40). Springer.

Belavkin, R. V. (2010b). Utility and value of information in cognitive science, biology and quantum theory. In L. Accardi, W. Freudenberg, & M. Ohya (Eds.), *Quantum Bio-Informatics III* (Vol. 26). World Scientific.

Ellsberg, D. (1961, November). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics*, *75*(4), 643–669.

Stratonovich, R. L. (1965). On value of information. *Izvestiya of USSR Academy of Sciences, Technical Cybernetics*, *5*, 3–12. (In Russian)

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453–458.