Econ 520 - Second Half

Section 1: Comparisons of Experiments

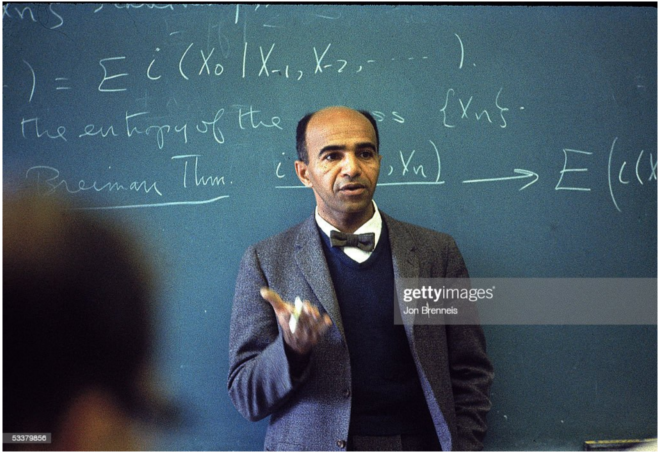
Ernesto Rivera Mora

Yale University

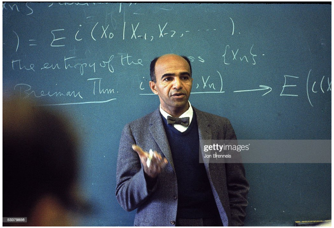
Winter 2023

Table of Contents

David Blackwell



David Blackwell



- 7th african american to receive a Ph.D in mathematics
  - Illinois at Urbana-Champaign
- 1st african american professor at Berkeley
- 1st african american inducted into the National Academy of Sciences

- **States:** Θ (finite)
- **Decision Problem** (for Θ) denoted $\mathrm{DP} = (A, u, p)$
  - Compact action set: A
  - Continuous utility Function: $u : A \times \Theta \to \mathbb{R}$
  - Support Prior: $p \in \Delta(\Theta)$
  - Objective:

$$\max_{a \in A} \sum_{\theta} p(\theta) \cdot u(a, \theta)$$

Experiments

- **Experiment:** $E = (S, \pi)$
  - Set of signals: S (finite)
  - Conditional distribution mapping $\pi : \Theta \to \Delta(S)$
- **Strategy** for E: mapping $\alpha : S \to \Delta(A)$
  - Σ_E: Set of Strategies for E
- **Expected utility** for DP given E:

$$V(E, \mathrm{DP}) = \max_{\alpha \in \Sigma_E} \sum_{\theta, a, s} u(a, \theta) \cdot \alpha(a \mid s) \cdot \pi(s \mid \theta) \cdot p(\theta)$$

Garblings

Definition

An experiment $E' = (S', \pi')$ is a **garbling** of another experiment $E = (S, \pi)$ if there is a mapping $\gamma : S \to \Delta(S')$ such that

$$\pi'(s' \mid \theta) = \sum_{s \in S} \gamma(s' \mid s) \cdot \pi(s \mid \theta)$$

- E' "adds" noise to E
- Defines partial order on the set of experiments
  - **Remark:** Not all pairs of experiments are ranked

Garblings

Definition

An experiment $E' = (S', \pi')$ is a **garbling** of another experiment $E = (S, \pi)$ if there is a mapping $\gamma : S \to \Delta(S')$ such that

$$\pi'(s' \mid \theta) = \sum_{s \in S} \gamma(s' \mid s) \cdot \pi(s \mid \theta)$$

- E' "adds" noise to E
- Defines partial order on the set of experiments
  - **Remark:** Not all pairs of experiments are ranked

More Information

Definition

E is **more informative** than E' if for any decision problem DP,

$$V(E, \mathrm{DP}) \geq V(E', \mathrm{DP})$$

- Any Bayesian agent prefers to observe E rather than E'
- This ranking defines **partial order** on the set of experiments
  - Remark: Not all pairs of experiments are ranked

More Information

Definition

E is **more informative** than E' if for any decision problem DP,

$$V(E, \mathrm{DP}) \geq V(E', \mathrm{DP})$$

- Any Bayesian agent prefers to observe E rather than E'
- This ranking defines **partial order** on the set of experiments
  - **Remark:** Not all pairs of experiments are ranked

More Information

Definition

E is more informative than E' if for any decision problem DP,

$$V(E, \mathrm{DP}) \geq V(E', \mathrm{DP})$$

- Any Bayesian agent prefers to observe E rather than E'
- This ranking defines **partial order** on the set of experiments
  - Remark: Not all pairs of experiments are ranked

Blackwell's Equivalence Result

Theorem

The following statements are equivalent:

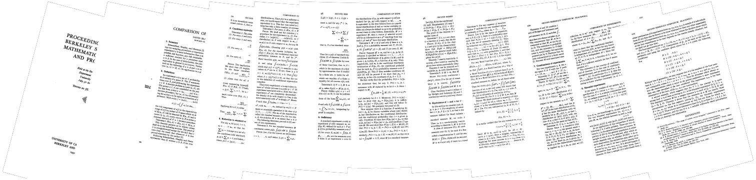
1. E' *is a* ***garbling*** *of* E.
2. E *is* ***more informative*** *than* E'.

Blackwell's Informativeness Theorem: Proof

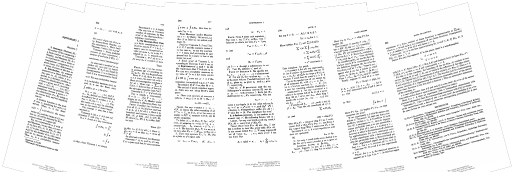
Proof?

Original Proof

- Blackwell 1951



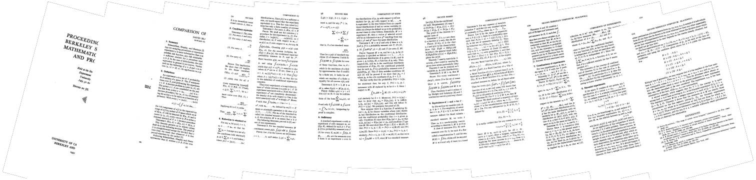
- Blackwell 1953



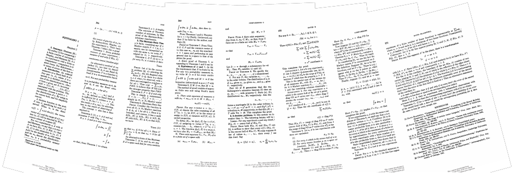
- Blackwell and Girschick (1954): 6 Equivalent conditions

Original Proof

- Blackwell 1951



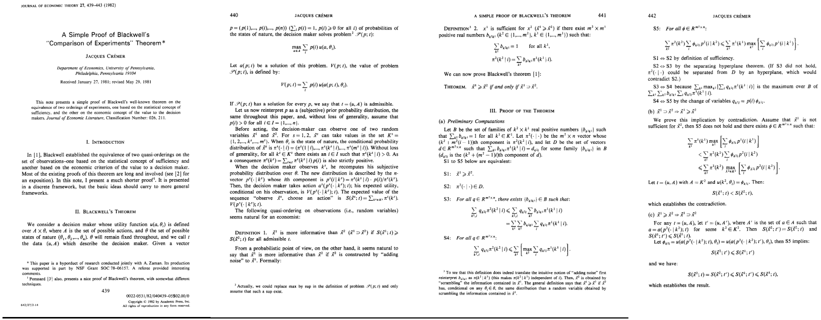
- Blackwell 1953



- Blackwell and Girschick (1954): 6 Equivalent conditions

Simpler Proof

Jacques Cremer (1981)



JOURNAL OF ECONOMIC THEORY 27, 439–443 (1982)

A Simple Proof of Blackwell's "Comparison of Experiments" Theorem*

JACQUES CRÉMER

Department of Economics, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received January 27, 1981; revised May 29, 1981

This note presents a simple proof of Blackwell's well-known theorem on the equivalence of two orderings of experiments, one based on the statistical concept of sufficiency, and the other on the economic concept of the value to the decision makers. *Journal of Economic Literature*, Classification Number: 026, 211.

I. INTRODUCTION

In [1], Blackwell established the equivalence of two quasi-orderings on the set of observations–one based on the statistical concept of sufficiency and another based on the economic criterion of the value to a decision maker. Most of the existing proofs of this theorem are long and involved (see [2] for an exposition). In this note, I present a much shorter proof[1]. It is presented in a discrete framework, but the basic ideas should carry to more general frameworks.

II. BLACKWELL'S THEOREM

We consider a decision maker whose utility function $u(a, \theta_i)$ is defined over $A \times \theta$, where A is the set of possible actions, and θ the set of possible states of nature $(\theta_1, \theta_2, ..., \theta_n)$. θ will remain fixed throughout, and we call t the data (u, A) which describe the decision maker. Given a vector

* This paper is a byporduct of research conducted jointly with A. Zaman. Its production was supported in part by NSF Grant SOC 78–06157. A referee provided interesting comments.

[1] Ponssard [3] also, presents a nice proof of Blackwell's theorem, with somewhat different techniques.

0022-0531/82/040439–05$02.00/0

642/27/2-14

$p = (p(1), ..., p(i), ..., p(n))$ ($\sum_i p(i) = 1$, $p(i) \geqslant 0$ for all i) of probabilities of the states of nature, the decision maker solves problem[2] $\mathscr{P}(p; t)$:

$$\max_{a \in A} \sum_i p(i)\, u(a, \theta_i).$$

Let $a(p; t)$ be a solution of this problem. $V(p; t)$, the value of problem $\mathscr{P}(p; t)$, is defined by:

$$V(p; t) = \sum_i p(i)\, u(a(p; t), \theta_i).$$

If $\mathscr{P}(p; t)$ has a solution for every p, we say that $t = (u, A)$ is admissible.

Let us now reinterpret p as a (subjective) prior probability distribution, the same throughout this paper, and, without loss of generality, assume that $p(i) > 0$ for all $i \in I = \{1, ..., n\}$.

Before acting, the decision-maker can observe one of two random variables $\tilde{x}^1$ and $\tilde{x}^2$. For $s = 1, 2$, $\tilde{x}^s$ can take values in the set $K^s = \{1, 2, ..., k^s, ..., m^s\}$. When θ_i is the state of nature, the conditional probability distribution of $\tilde{x}^s$ is $\pi^s(\cdot \mid i) = (\pi^s(1 \mid i), ..., \pi^s(k^s \mid i), ..., \pi^s(m^s \mid i))$. Without loss of generality, for all $k^s \in K^s$ there exists an $i \in I$ such that $\pi^s(k^s \mid i) > 0$. As a consequence $\pi^s(k^s) = \sum_{i \in I} \pi^s(k^s \mid i)\, p(i)$ is also strictly positive.

When the decision maker observes k^s, he recomputes his subjective probability distribution over θ. The new distribution is described by the n-vector $p^s(\cdot \mid k^s)$ whose ith component is $p^s(i \mid k^s) = \pi^s(k^s \mid i) \cdot p(i)/\pi^s(k^s)$. Then, the decision maker takes action $a^s(p^s(\cdot \mid k^s); t)$; his expected utility, conditional on his observation, is $V(p^s(\cdot \mid k^s); t)$. The expected value of the sequence "observe $\tilde{x}^s$, choose an action" is $S(\tilde{x}^s; t) = \sum_{k^s \in K^s} \pi^s(k^s) \cdot V(p^s(\cdot \mid k^s); t)$.

The following quasi-ordering on observations (i.e., random variables) seems natural for an economist:

DEFINITION 1. $\tilde{x}^1$ is more informative than $\tilde{x}^2$ ($\tilde{x}^1 \supset \tilde{x}^2$) if $S(\tilde{x}^1; t) \geqslant S(\tilde{x}^2; t)$ for all admissible t.

From a probabilistic point of view, on the other hand, it seems natural to say that $\tilde{x}^1$ is more informative than $\tilde{x}^2$ if $\tilde{x}^2$ is constructed by "adding noise" to $\tilde{x}^1$. Formally:

[2] Actually, we could replace max by sup in the definition of problem $\mathscr{P}(p; t)$ and only assume that such a sup exist.

DEFINITION[3] 2. x^1 is sufficient for x^2 ($\tilde{x}^1 \geqslant \tilde{x}^2$) if there exist $m^2 \times m^1$ positive real numbers $b_{k^2k^1}$ ($k^2 \in \{1, ..., m^2\}$, $k^1 \in \{1, ..., m^1\}$) such that:

$$\sum_{k^1} b_{k^2k^1} = 1 \qquad \text{for all } k^1,$$

$$\pi^2(k^2 \mid i) = \sum_{k^1} b_{k^2k^1} \pi^1(k^1 \mid i).$$

We can now prove Blackwell's theorem [1]:

THEOREM. *$\tilde{x}^1 \geqslant \tilde{x}^2$ if and only if $\tilde{x}^1 \supset \tilde{x}^2$.*

III. PROOF OF THE THEOREM

(a) *Preliminary Computations*

Let B be the set of families of $k^2 \times k^1$ real positive numbers $\{b_{k^2k^1}\}$ such that $\sum_{k^2} b_{k^2k^1} = 1$ for all $k^1 \in K^1$. Let $\pi^2(\cdot \mid \cdot)$ be the $m^2 \times n$ vector whose $(k^2 + m^2(i-1))$th component is $\pi^2(k^2 \mid i)$, and let D be the set of vectors $d \in R^{m^2 \times n}$ such that $\sum_{k^1} b_{k^2k^1} \pi^1(k^1 \mid i) = d_{k^2i}$ for some family $\{b_{k^2k^1}\}$ in B (d_{k^2i} is the $(k^2 + (m^2 - 1)i)$th component of d).

S1 to S5 below are equivalent:

S1: $\tilde{x}^1 \geqslant \tilde{x}^2$.

S2: $\pi^2(\cdot \mid \cdot) \in D$.

S3: *For all $q \in R^{m^2 \times n}$, there exists $\{b_{k^2k^1}\} \in B$ such that:*

$$\sum_{k^2, i} q_{k^2i} \pi^2(k^2 \mid i) \leqslant \sum_{k^2, i} q_{k^2i} \sum_{k^1} b_{k^2k^1} \pi^1(k^1 \mid i) = \sum_{k^1} \sum_{k^2} b_{k^2k^1} \sum_i q_{k^2i} \pi^1(k^1 \mid i).$$

S4: *For all $q \in R^{m^2 \times n}$:*

$$\sum_{k^2, i} q_{k^2i} \pi^2(k^2 \mid i) \leqslant \sum_{k^1} \left[\max_{k^2} \sum_i q_{k^2i} \pi^1(k^1 \mid i) \right].$$

[3] To see that this definition does indeed translate the intuitive notion of "adding noise" first reinterpret $b_{k^2k^1}$ as $\pi(k^2 \mid k^1)$ (this makes $\pi(k^2 \mid k^1)$ independent of i). Then, $\tilde{x}^2$ is obtained by "scrambling" the information contained in $\tilde{x}^1$. The general definition says that $\tilde{x}^1 \geqslant \tilde{x}^2$ if $\tilde{x}^2$ has, conditional on any $\theta_i \in \theta$, the same distribution than a random variable obtained by scrambling the information contained in $\tilde{x}^1$.

S5: *For all $\phi \in R^{m^2 \times n}$:*

$$\sum_{k^2} \pi^2(k^2) \sum_i \phi_{k^2i} p^2(i \mid k^2) \leqslant \sum_{k^1} \pi^1(k^1) \max_{k^2} \left[\sum_i \phi_{k^2i} p^1(i \mid k^1) \right].$$

S1 $\Leftrightarrow$ S2 by definition of sufficiency.

S2 $\Leftrightarrow$ S3 by the separating hyperplane theorem. (If S3 did not hold, $\pi^2(\cdot \mid \cdot)$ could be separated from D by an hyperplane, which would contradict S2.)

S3 $\Leftrightarrow$ S4 because $\sum_{k^1} \max_{k^2} [\sum_i q_{k^2i} \pi^1(k^1 \mid i)]$ is the maximum over B of $\sum_{k^1} \sum_{k^2} b_{k^2k^1} \sum_i q_{k^2i} \pi^1(k^1 \mid i)$.

S4 $\Leftrightarrow$ S5 by the change of variables $q_{k^2i} = p(i)\, \phi_{k^2i}$.

(b) $\tilde{x}^1 \supset \tilde{x}^2 \Rightarrow \tilde{x}^1 \geqslant \tilde{x}^2$

We prove this implication by contradiction. Assume that $\tilde{x}^1$ is not sufficient for $\tilde{x}^2$, then S5 does not hold and there exists $\phi \in R^{m^2 \times n}$ such that:

$$\begin{aligned} &\sum_{k^1} \pi^1(k^1) \max_{k^2} \left[\sum_i \phi_{k^2i} p^1(i \mid k^1) \right] \\ &\quad < \sum_{k^2} \pi^2(k^2) \sum_i \phi_{k^2i} p^2(i \mid k^2) \\ &\quad \leqslant \sum_{k^2} \pi^2(k^2) \max_{\lambda^2 \in K^2} \left[\sum_i \phi_{\lambda^2 i} p^2(i \mid k^2) \right]. \end{aligned}$$

Let $t = (u, A)$ with $A = K^2$ and $u(k^2, \theta_i) = \phi_{k^2i}$. Then:

$$S(\tilde{x}^1; t) < S(\tilde{x}^2; t),$$

which establishes the contradiction.

(c) $\tilde{x}^1 \geqslant \tilde{x}^2 \Rightarrow \tilde{x}^1 \supset \tilde{x}^2$

For any $t = (u, A)$, let $t' = (u, A')$, where A' is the set of $a \in A$ such that $a = a(p^2(\cdot \mid k^2); t)$ for some $k^2 \in K^2$. Then $S(\tilde{x}^2; t') = S(\tilde{x}^2; t)$ and $S(\tilde{x}^1; t') \leqslant S(\tilde{x}^1; t)$.

Let $\phi_{k^2i} = u(a(p^2(\cdot \mid k^2); t), \theta_i) = u(a(p^2(\cdot \mid k^2); t'), \theta_i)$, then S5 implies:

$$S(\tilde{x}^2; t') \leqslant S(\tilde{x}^1; t')$$

and we have:

$$S(\tilde{x}^2; t) = S(\tilde{x}^2; t') \leqslant S(\tilde{x}^1; t') \leqslant S(\tilde{x}^1; t),$$

which establishes the result.

Jacques Cremer (1981)

S1: $\tilde{x}^1 \geqslant \tilde{x}^2$.

S2: $\pi^2(\cdot \mid \cdot) \in D$.

S3: *For all* $q \in R^{m^2 \times n}$, *there exists* $\{b_{k^2 k^1}\} \in B$ *such that:*

$$\sum_{k^2, i} q_{k^2 i} \pi^2(k^2 \mid i) \leqslant \sum_{k^2, i} q_{k^2 i} \sum_{k^2} b_{k^2 k^1} \pi^1(k^1 \mid i)$$
$$= \sum_{k^1} \sum_{k^2} b_{k^2 k^1} \sum_{i} q_{k^2 i} \pi^1(k^1 \mid i).$$

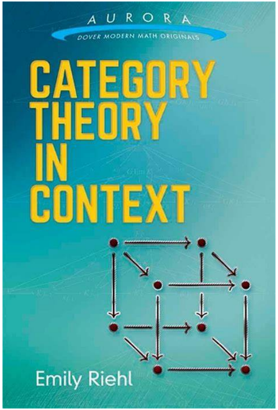
S4: *For all* $q \in R^{m^2 \times n}$:

$$\sum_{k^2, i} q_{k^2 i} \pi^2(k^2 \mid i) \leqslant \sum_{k^1} \left[\max_{k^2} \sum_{i} q_{k^2 i} \pi^1(k^1 \mid i) \right].$$

S5: *For all* $\phi \in R^{m^2 \times n}$:

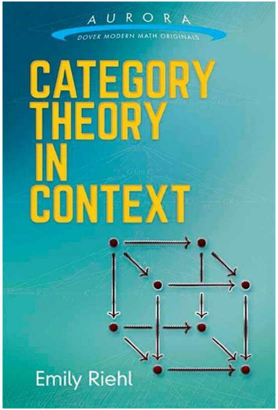
$$\sum_{k^2} \pi^2(k^2) \sum_{i} \phi_{k^2 i} p^2(i \mid k^2) \leqslant \sum_{k^1} \pi^1(k^1) \max_{k^2} \left[\sum_{i} \phi_{k^2 i} p^1(i \mid k^1) \right].$$

Proof of Blackwell theorem using **Category Theory**



- Studies *relations* between objects
- We will use only chapter 1
- Applications in **Economic Theory**, Neuroscience, Physics, Linguistics, etc.

Proof of Blackwell theorem using **Category Theory**

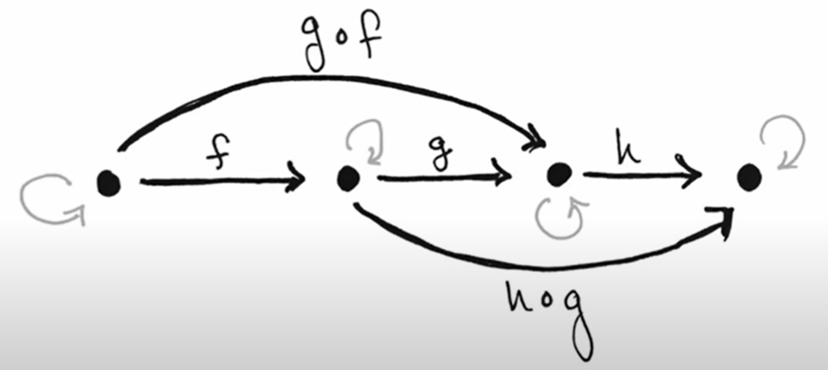


- Studies *relations* between objects
- We will use only chapter 1
- Applications in **Economic Theory**, Neuroscience, Physics, Linguistics, etc.

Set Theory vs Category Theory

- **Set Theory:** Interested in different kinds of objects
- **Category Theory:** Interested in relations between objects

Composition of Morphisms



Category - Definition

- **Objects:** $X, Y, Z, \ldots$
  - Eg.: numbers, sets, vectors, hyper-planes, states, signals, actions
- **Morphisms** $f, g, h, \ldots$
  - "Generalization" of a mapping
  - Notation: $f : X \dashrightarrow Y$ where "domain" and "codomain" are the objects
  - Each object X has a **identity morphism** $\mathrm{Id}_X : X \dashrightarrow X$
  - For each pair $f : X \dashrightarrow Y, g : Y \dashrightarrow Z$, there is a **composition** $g \circ f : X \dashrightarrow Z$
- **Category** is a set of Objects and Morphisms that satisfy two **Axioms**

Category - Definition

- **Objects:** $X, Y, Z, \ldots$
    - Eg.: numbers, sets, vectors, hyper-planes, states, signals, actions
- **Morphisms** $f, g, h, \ldots$
    - "Generalization" of a mapping
    - Notation: $f : X \dashrightarrow Y$ where "domain" and "codomain" are the objects
    - Each object X has a **identity morphism** $\mathrm{Id}_X : X \dashrightarrow X$
    - For each pair $f : X \dashrightarrow Y, g : Y \dashrightarrow Z$, there is a **composition** $g \circ f : X \dashrightarrow Z$
- Category is a set of Objects and Morphisms that satisfy two **Axioms**

Category - Definition

- **Objects:** $X, Y, Z, \ldots$
  - Eg.: numbers, sets, vectors, hyper-planes, states, signals, actions
- **Morphisms** $f, g, h, \ldots$
  - "Generalization" of a mapping
  - Notation: $f : X \dashrightarrow Y$ where "domain" and "codomain" are the objects
  - Each object X has a **identity morphism** $\mathrm{Id}_X : X \dashrightarrow X$
  - For each pair $f : X \dashrightarrow Y, g : Y \dashrightarrow Z$, there is a **composition** $g \circ f : X \dashrightarrow Z$
- **Category** is a set of Objects and Morphisms that satisfy two **Axioms**

Axioms

1. **Axiom 1:**

   For each $f : X \dashrightarrow Y$, $\quad \mathrm{Id}_Y \circ f = f \circ \mathrm{Id}_X = f$

2. **Axiom 2:**

   For each $f : X \dashrightarrow Y$, $g : Y \dashrightarrow Z$, $h : Z \dashrightarrow W$,

   $$f \circ (g \circ h) = (f \circ g) \circ h$$

Axioms

1. **Axiom 1:**

   For each $f : X \dashrightarrow Y$, $\quad \mathrm{Id}_Y \circ f = f \circ \mathrm{Id}_X = f$

2. **Axiom 2:**

   For each $f : X \dashrightarrow Y$, $g : Y \dashrightarrow Z$, $h : Z \dashrightarrow W$,

   $$f \circ (g \circ h) = (f \circ g) \circ h$$

Rermark: Morphisms vs Mappings

- $f : X \dashrightarrow Y$ is not necessarily a function from X to Y
- $f : X \dashrightarrow Y$ “represents" a function $f : Obj(X) \to Obj(Y)$
  - $Obj(X)$ is a mathematical object that depends on X (can be X)
  - $Obj(Y)$ is a mathematical object that depends on Y (can be Y)
- $f : X \dashrightarrow Y$ is different from $f : X \to Y$

Examples

1. Category of sets and functions
2. Category of topological spaces and continuous functions
3. Category of measurable spaces and measurable functions
4. Category of natural numbers and matrices:
   - An object is a number $n \in \mathbb{N}$
   - A morphism $f : n \dashrightarrow m$ is a linear function from $\mathbb{R}^n$ to $\mathbb{R}^m$
     - Represented by a matrix of size $m \times n$
     - Identity of n : I_n
     - Composition: Matrix multiplication

Examples

1. Category of sets and functions
2. Category of topological spaces and continuous functions
3. Category of measurable spaces and measurable functions
4. Category of natural numbers and matrices:
   - An object is a number $n \in \mathbb{N}$
   - A morphism $f : n \dashrightarrow m$ is a linear function from $\mathbb{R}^n$ to $\mathbb{R}^m$
     - Represented by a matrix of size $m \times n$
     - Identity of n : I_n
     - Composition: Matrix multiplication

Examples

1. Category of sets and functions
2. Category of topological spaces and continuous functions
3. Category of measurable spaces and measurable functions
4. Category of natural numbers and matrices:
    - An object is a number $n \in \mathbb{N}$
    - A morphism $f : n \dashrightarrow m$ is a linear function from $\mathbb{R}^n$ to $\mathbb{R}^m$
        - Represented by a matrix of size $m \times n$
        - Identity of n : I_n
        - Composition: Matrix multiplication

Category of Stochastic Mappings

- **Objects:** Θ, A, S, S' (finite sets)
- A **morphism** $f : X \dashrightarrow Y$ is a function $f : X \to \Delta Y$
  - **Identity:** $\mathrm{Id}_X : X \dashrightarrow X$ Identity mapping $\mathrm{Id}_X : X \to \Delta(X)$
    - $\mathrm{Id}_X(x \mid x') = 1$ iff $x = x'$
  - **Composition:** $f : X \dashrightarrow Y$, $g : Y \dashrightarrow Z$, define $g \circ f : X \dashrightarrow Z$ as
    - $g \circ f(z \mid x) = \sum_y g(z \mid y) \cdot f(y \mid x)$

Category of Stochastic Mappings

- **Objects:** Θ, A, S, S' (finite sets)
- A **morphism** $f : X \dashrightarrow Y$ is a function $f : X \to \Delta Y$
  - **Identity:** $\mathrm{Id}_X : X \dashrightarrow X$ Identity mapping $\mathrm{Id}_X : X \to \Delta(X)$
    - $\mathrm{Id}_X(x \mid x') = 1$ iff $x = x'$
  - **Composition:** $f : X \dashrightarrow Y$, $g : Y \dashrightarrow Z$, define $g \circ f : X \dashrightarrow Z$ as
    - $g \circ f(z \mid x) = \sum_y g(z \mid y) \cdot f(y \mid x)$

Category of Stochastic Mappings

- **Objects:** Θ, A, S, S' (finite sets)
- A **morphism** $f : X \dashrightarrow Y$ is a function $f : X \to \Delta Y$
  - **Identity:** $\mathrm{Id}_X : X \dashrightarrow X$ Identity mapping $\mathrm{Id}_X : X \to \Delta(X)$
    - $\mathrm{Id}_X(x \mid x') = 1$ iff $x = x'$
  - **Composition:** $f : X \dashrightarrow Y$, $g : Y \dashrightarrow Z$, define $g \circ f : X \dashrightarrow Z$ as
    - $g \circ f(z \mid x) = \sum_y g(z \mid y) \cdot f(y \mid x)$

Category of Stochastic Mappings

- **Objects:** Θ, A, S, S' (finite sets)
- A **morphism** $f : X \dashrightarrow Y$ is a function $f : X \to \Delta Y$
  - **Identity:** $\mathrm{Id}_X : X \dashrightarrow X$ Identity mapping $\mathrm{Id}_X : X \to \Delta(X)$
    - $\mathrm{Id}_X(x \mid x') = 1$ iff $x = x'$
  - **Composition:** $f : X \dashrightarrow Y$, $g : Y \dashrightarrow Z$, define $g \circ f : X \dashrightarrow Z$ as
    - $g \circ f(z \mid x) = \sum_y g(z \mid y) \cdot f(y \mid x)$

Axioms

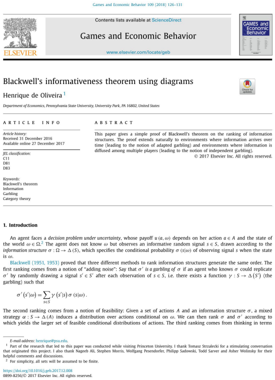
1. For each $f : X \dashrightarrow Y$,
   - $\mathrm{Id}_Y \circ f = f \circ \mathrm{Id}_X = f$
2. For each $f : X \dashrightarrow Y$, $g : Y \dashrightarrow Z$, $h : Z \dashrightarrow W$,
   - $f \circ (g \circ h)(w \mid x) = \sum_y \sum_z h(w \mid z) \cdot g(z \mid y) \cdot f(y \mid x)$
   - $(f \circ g) \circ h(w \mid x) = \sum_y \sum_z h(w \mid z) \cdot g(z \mid y) \cdot f(y \mid x)$

Commuting Diagrams

$$
\begin{tikzcd}
X \arrow[r, dashed, "f"'] \arrow[d, dashed, "f'"] & Y \arrow[d, dashed, "g"'] \\
Y' \arrow[r, dashed, "g'"] & Z
\end{tikzcd}
$$

- Say the *"diagram commutes"* if $g \circ f = g' \circ f'$
- Any two paths with same beginning and end define the same morphism

Henrique de Oliveira, 2018


Games and Economic Behavior 109 (2018) 126–131

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Blackwell's informativeness theorem using diagrams

Henrique de Oliveira [1]

Department of Economics, Pennsylvania State University, University Park, PA 16802, United States

ARTICLE INFO

Article history:
Received 31 December 2016
Available online 27 December 2017

JEL classification:
C11
D81
D83

Keywords:
Blackwell's theorem
Information
Garbling
Category theory

ABSTRACT

This paper gives a simple proof of Blackwell's theorem on the ranking of information structures. The proof extends naturally to environments where information arrives over time (leading to the notion of adapted garbling) and environments where information is diffused among multiple players (leading to the notion of independent garbling).

1. Introduction

An agent faces a *decision problem under uncertainty*, whose payoff $u(a, \omega)$ depends on her action $a \in A$ and the state of the world $\omega \in \Omega$.[2] The agent does not know ω but observes an informative random signal $s \in S$, drawn according to the *information structure* $\sigma : \Omega \to \Delta(S)$, which specifies the conditional probability $\sigma(s|\omega)$ of observing signal s when the state is ω.

Blackwell (1951, 1953) proved that three different methods to rank information structures generate the same order. The first ranking comes from a notion of "adding noise": Say that σ' *is a garbling of* σ if an agent who knows σ could replicate σ' by randomly drawing a signal $s' \in S'$ after each observation of $s \in S$, i.e. there exists a function $\gamma : S \to \Delta(S')$ (the garbling) such that

$$\sigma'(s'|\omega) = \sum_{s \in S} \gamma(s'|s)\, \sigma(s|\omega).$$

The second ranking comes from a notion of feasibility: Given a set of actions A and an information structure σ, a mixed strategy $\alpha : S \to \Delta(A)$ induces a distribution over actions conditional on ω. We can then rank σ and σ' according to which yields the larger set of feasible conditional distributions of actions. The third ranking comes from thinking in terms

E-mail address: henrique@psu.edu.

[1] Part of the research that led to this paper was conducted while visiting Princeton University. I thank Tomasz Strzalecki for a stimulating conversation that originated this project. I also thank Nageeb Ali, Stephen Morris, Wolfgang Pesendorfer, Philipp Sadowski, Todd Sarver and Asher Wolinsky for their helpful comments and discussions.

[2] For simplicity, all sets will be assumed to be finite.

https://doi.org/10.1016/j.geb.2017.12.008

Observation: Strategies and experiments belong to the same category: The category of stochastic mappings.

Advantage: Uses "commuting diagrams" to provide an elegant proof

$$\begin{tikzcd} \Theta \arrow[r, dashed, "\pi"] \arrow[d, dashed, "\pi'"'] & S \arrow[dl, dashed, "\gamma"] \\ S' & \end{tikzcd}$$

Remark

$E' = (S', \pi')$ is a **garbling** of $E = (\pi, S)$ if there is some stochastic mapping $\gamma : S \to \Delta(S')$ so the diagram commutes.

Conditional Distribution of Actions

$$
\begin{tikzcd}
\Theta \arrow[r, dashed, "\pi"'] \arrow[d, dashed, "\pi'"] & S \arrow[d, dashed, "\alpha"'] \\
S' \arrow[r, dashed, "\alpha'"] & A
\end{tikzcd}
$$

- For each $E = (S, \pi)$ and A, define

$$\Lambda(E, A) := \{\alpha \circ \pi \mid \alpha : S \dashrightarrow A\}$$

- When does $\Lambda(E', A) \subseteq \Lambda(E, A)$ holds ?
  - Holds if for each $\alpha' : S' \dashrightarrow A$ there exists some $\alpha : S \dashrightarrow A$ so that $\alpha \circ \pi = \alpha' \circ \pi'$ (the diagram commutes)

Blackwell's Equivalence Theorem

Theorem

The following statements are equivalent:

1. E' *is a garbling of* E.
2. *For any action set* A, $\Lambda(E', A) \subseteq \Lambda(E, A)$.
3. E *is more informative than* E': $V(E, \mathrm{DP}) \geq V(E', \mathrm{DP})$ *for each* DP

Blackwell's Equivalence Theorem

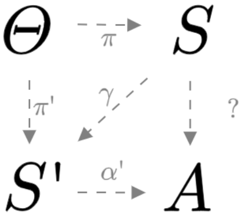
Theorem

The following statements are equivalent:

1. *Statement in terms of **information***
2. *Statement in terms of **possibilities***
3. *Statement in terms of **payoffs***

Proof (1) Implies (2)

To show: There is some γ so that $\gamma \circ \pi = \pi'$ $\Rightarrow$ $\Lambda(E', A) \subseteq \Lambda(E, A)$



Fix α', wish to find α that commutes the diagram

Proof (1) Implies (2)

To show: There is some γ so that $\gamma \circ \pi = \pi'$ $\quad\Rightarrow\quad$ $\Lambda(E', A) \subseteq \Lambda(E, A)$

$$\begin{tikzcd}
\Theta \arrow[r, dashed, "\pi"] \arrow[d, dashed, "\pi'"] & S \arrow[dl, dashed, "\gamma"'] \arrow[d, dashed, "\gamma \circ \alpha'"] \\
S' \arrow[r, dashed, "\alpha'"] & A
\end{tikzcd}$$

Define $\alpha = \alpha' \circ \gamma$,

Proof (2) Implies (1)

To show: $\Lambda(E', A) \subseteq \Lambda(E, A) \quad \Rightarrow \quad$ There is some γ so that $\gamma \circ \pi = \pi'$

$$\begin{tikzcd} \Theta \arrow[r, dashed, "\pi"'] \arrow[d, dashed, "\pi'"] & S \\ S' & \end{tikzcd}$$

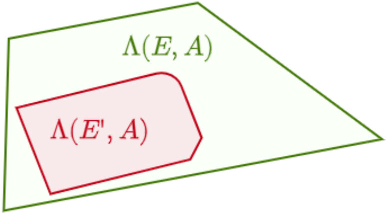
Proof (2) Implies (1)

To show: $\Lambda(E', A) \subseteq \Lambda(E, A) \quad \Rightarrow \quad$ There is some γ so that $\gamma \circ \pi = \pi'$

$$\begin{tikzcd}
\Theta \arrow[r, dashed, "\pi"'] \arrow[d, dashed, "\pi'"] & S \arrow[d, dashed, "\alpha"] \\
S' \arrow[r, dashed, "1_{S'}"] & S'
\end{tikzcd}$$

- Set $A = S'$ and take $\alpha' = \mathrm{Id}_{S'}$
- By (2), there is some α that commutes the diagram
  - Notice: $\pi' = \pi' \circ \mathrm{Id}_{S'} = \alpha \circ \pi$

Proof (2) Implies (3)

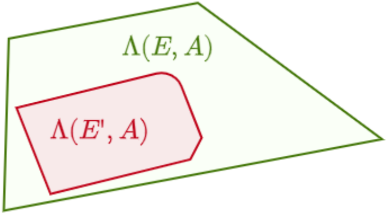


$$V(E, \mathrm{DP}) = \max_{\alpha \in \Sigma_E} \sum_{\theta, a, s} u(a, \theta) \cdot \alpha(a \mid s) \cdot \pi(s \mid \theta) \cdot p(\theta)$$

$$= \max_{\lambda \in \Lambda(E, A)} \sum_{\theta, a} u(a, \theta) \cdot \lambda(a \mid \theta) \cdot p(\theta)$$

Thus, $\quad V(E, \mathrm{DP}) \geq V(E', \mathrm{DP})$

Proof (2) Implies (3)

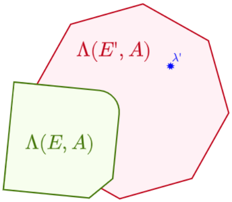


$$V(E, \mathrm{DP}) = \max_{\alpha \in \Sigma_E} \sum_{\theta, a, s} u(a, \theta) \cdot \alpha(a \mid s) \cdot \pi(s \mid \theta) \cdot p(\theta)$$

$$= \max_{\lambda \in \Lambda(E, A)} \sum_{\theta, a} u(a, \theta) \cdot \lambda(a \mid \theta) \cdot p(\theta)$$

Thus, $\quad V(E, \mathrm{DP}) \geq V(E', \mathrm{DP})$

Proof (3) Implies (2)

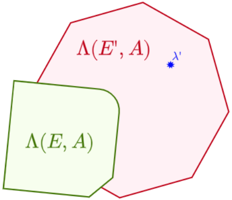


- Assume exists $\lambda' \in \Lambda(E', A) \backslash \Lambda(E, A)$
- $\Lambda(E, A) \subset \mathbb{R}^{\Theta \times A}$ is compact and convex
- There is a hyperplane $u \in \mathbb{R}^{\Theta \times A}$ such that, for each $\lambda \in \Lambda(E, A)$,

$$\sum_{\theta, a} u(\theta, a) \cdot \lambda(a \mid \theta) < \sum_{\theta, a} u(\theta, a) \cdot \lambda'(a \mid \theta)$$

- Then $V(E, \mathrm{DP}) < V(E', \mathrm{DP})$ for $\mathrm{DP} = (A, u, p)$ with **uniform** p

Proof (3) Implies (2)

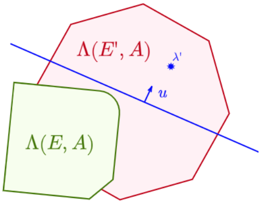


- Assume exists $\lambda' \in \Lambda(E', A) \backslash \Lambda(E, A)$
- $\Lambda(E, A) \subset \mathbb{R}^{\Theta \times A}$ is compact and convex
- There is a hyperplane $u \in \mathbb{R}^{\Theta \times A}$ such that, for each $\lambda \in \Lambda(E, A)$,

$$\sum_{\theta, a} u(\theta, a) \cdot \lambda(a \mid \theta) < \sum_{\theta, a} u(\theta, a) \cdot \lambda'(a \mid \theta)$$

- Then $V(E, \mathrm{DP}) < V(E', \mathrm{DP})$ for $\mathrm{DP} = (A, u, p)$ with **uniform** p

Proof (3) Implies (2)

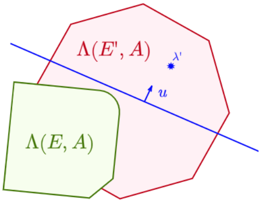


- Assume exists $\lambda' \in \Lambda(E', A) \backslash \Lambda(E, A)$
- $\Lambda(E, A) \subset \mathbb{R}^{\Theta \times A}$ is compact and convex
- There is a hyperplane $u \in \mathbb{R}^{\Theta \times A}$ such that, for each $\lambda \in \Lambda(E, A)$,

$$\sum_{\theta, a} u(\theta, a) \cdot \lambda(a \mid \theta) < \sum_{\theta, a} u(\theta, a) \cdot \lambda'(a \mid \theta)$$

- Then $V(E, \mathrm{DP}) < V(E', \mathrm{DP})$ for $\mathrm{DP} = (A, u, p)$ with **uniform** p

Proof (3) Implies (2)



- Assume exists $\lambda' \in \Lambda(E', A) \backslash \Lambda(E, A)$
- $\Lambda(E, A) \subset \mathbb{R}^{\Theta \times A}$ is compact and convex
- There is a hyperplane $u \in \mathbb{R}^{\Theta \times A}$ such that, for each $\lambda \in \Lambda(E, A)$,

$$\sum_{\theta, a} u(\theta, a) \cdot \lambda(a \mid \theta) < \sum_{\theta, a} u(\theta, a) \cdot \lambda'(a \mid \theta)$$

- Then $V(E, \mathrm{DP}) < V(E', \mathrm{DP})$ for $\mathrm{DP} = (A, u, p)$ with **uniform** p

Blackwell Theorem:

Applications

Application

- **Buyer** has valuation v of object
  - v is distributed uniformly in $[0, 1]$
- **Seller** does not observe v
  - Seller observes some experiment (E, π) of v
  - Provides a take it or leave it price p to Buyer
  - Buyer buys iff $v \geq p$
- **To Show:** Seller's expected revenue is in $[\frac{1}{4}, \frac{1}{2}]$

Application

- **Buyer** has valuation v of object
  - v is distributed uniformly in $[0, 1]$
- **Seller** does not observe v
  - Seller observes some experiment (E, π) of v
  - Provides a take it or leave it price p to Buyer
  - Buyer buys iff $v \geq p$
- **To Show:** Seller's expected revenue is in $[\frac{1}{4}, \frac{1}{2}]$

Application

- **Buyer** has valuation v of object
  - v is distributed uniformly in $[0, 1]$
- **Seller** does not observe v
  - Seller observes some experiment (E, π) of v
  - Provides a take it or leave it price p to Buyer
  - Buyer buys iff $v \geq p$
- **To Show:** Seller's expected revenue is in $[\frac{1}{4}, \frac{1}{2}]$

Best Experiment vs. Worst Experiment

- Seller faces a decision problem:
  - Unknown state: Buyer's valuation v
  - Objective: Choosing the right price p
- **Best experiment:** Fully Revealing θ
  - Any other experiment is a garbling of a fully revealing experiment.
- **Worst experiment:** Observe nothing
  - Observing nothing is a garbling of any experiment
- **Blackwell:** Payoffs bounded by the best and the worst experiments

Best Experiment vs. Worst Experiment

- Seller faces a decision problem:
  - Unknown state: Buyer's valuation v
  - Objective: Choosing the right price p
- **Best experiment:** Fully Revealing θ
  - Any other experiment is a garbling of a fully revealing experiment.
- **Worst experiment:** Observe nothing
  - Observing nothing is a garbling of any experiment
- **Blackwell:** Payoffs bounded by the best and the worst experiments

Best Experiment vs. Worst Experiment

- Seller faces a decision problem:
  - Unknown state: Buyer's valuation v
  - Objective: Choosing the right price p
- **Best experiment:** Fully Revealing θ
  - Any other experiment is a garbling of a fully revealing experiment.
- **Worst experiment:** Observe nothing
  - Observing nothing is a garbling of any experiment
- **Blackwell:** Payoffs bounded by the best and the worst experiments

Best Experiment vs. Worst Experiment

- Seller faces a decision problem:
  - Unknown state: Buyer's valuation v
  - Objective: Choosing the right price p
- **Best experiment:** Fully Revealing θ
  - Any other experiment is a garbling of a fully revealing experiment.
- **Worst experiment:** Observe nothing
  - Observing nothing is a garbling of any experiment
- **Blackwell:** Payoffs bounded by the best and the worst experiments

Best Experiment

- **Best Experiment:** Observing buyer's valuation
  - Optimal price: $p = v$.
  - Expected revenue: $\mathbb{E}[v] = \frac{1}{2}$.
- **Blackwell's Theorem:** If a seller can observe experiment (E, π), the expected revenue will be lower than $\frac{1}{2}$.

Best Experiment

- **Best Experiment:** Observing buyer's valuation
  - Optimal price: $p = v$.
  - Expected revenue: $\mathbb{E}[v] = \frac{1}{2}$.
- **Blackwell's Theorem:** If a seller can observe experiment (E, π), the expected revenue will be lower than $\frac{1}{2}$.

Worst Experiment

- **Worst Experiment:** Observing nothing
  - Optimal price: $p = \frac{1}{2}$.
  - Probability of sale: $\frac{1}{2}$.
  - Expected revenue: $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.
- **Blackwell's Theorem:** If a seller can observe experiment (E, π), the expected revenue will be higher than $\frac{1}{4}$.

Worst Experiment

- **Worst Experiment:** Observing nothing
  - Optimal price: $p = \frac{1}{2}$.
  - Probability of sale: $\frac{1}{2}$.
  - Expected revenue: $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.
- **Blackwell's Theorem:** If a seller can observe experiment (E, π), the expected revenue will be higher than $\frac{1}{4}$.

Exercise: Comparison of Gaussian Signals

Two signals: $X \sim N(\mu, \sigma^2)$ and $Y \sim N(\mu, 0.25 \cdot \sigma^2)$

Case 1 Known variable: σ^2 - Unknown state $\theta = \mu$

Case 2 Known variable: μ - Unknown state $\theta = \sigma^2$

Case 3 Unknown state $\theta = (\mu, \sigma^2)$

When are signals comparable? Which signal is more informative?

Application: Comparison of Gaussian Signals

Two signals: $X \sim N(\mu, \sigma^2)$ and $Y \sim N(\mu, 0.25 \cdot \sigma^2)$

Case 1 Known variable σ^2 - Unknown state $\theta = \mu$

- X is less informative than Y:

  X has same distribution as $Y + Z$ with $Z \sim N(0, 0.75 \cdot \sigma^2)$

Case 2 Known variable μ - Unknown state $\theta = \sigma^2$

- X and Y are equally informative:
  - $2Y - \mu$ has same distribution as X

Case 3 Unknown state $\theta = (\mu, \sigma)$

- X and Y are not comparable (Hansen and Torgersen, 1954)

Application: Comparison of Gaussian Signals

Two signals: $X \sim N(\mu, \sigma^2)$ and $Y \sim N(\mu, 0.25 \cdot \sigma^2)$

Case 1 Known variable σ^2 - Unknown state $\theta = \mu$

- X is less informative than Y:

  X has same distribution as $Y + Z$ with $Z \sim N(0, 0.75 \cdot \sigma^2)$

Case 2 Known variable μ - Unknown state $\theta = \sigma^2$

- X and Y are equally informative:
  - $2Y - \mu$ has same distribution as X

Case 3 Unknown state $\theta = (\mu, \sigma)$

- X and Y are not comparable (Hansen and Torgersen, 1954)

Application: Comparison of Gaussian Signals

Two signals: $X \sim N(\mu, \sigma^2)$ and $Y \sim N(\mu, 0.25 \cdot \sigma^2)$

Case 1 Known variable σ^2 - Unknown state $\theta = \mu$

- X is less informative than Y:

  X has same distribution as $Y + Z$ with $Z \sim N(0, 0.75 \cdot \sigma^2)$

Case 2 Known variable μ - Unknown state $\theta = \sigma^2$

- X and Y are equally informative:
  - $2Y - \mu$ has same distribution as X

Case 3 Unknown state $\theta = (\mu, \sigma)$

- X and Y are not comparable (Hansen and Torgersen, 1954)

Blackwell's Theorem:

Other Useful Equivalent Statements

Blackwell's Equivalence Theorem:

Theorem

The following statements are equivalent:

1. E' *is a garbling of* E.
2. *For any action set* A, $\Lambda(E', A) \subseteq \Lambda(E, A)$.
3. E *is more informative than* E': $V(E, \mathrm{DP}) \geq V(E', \mathrm{DP})$ *for each* DP

Blackwell's Equivalence Theorem:

Theorem

Equivalence between:

1. *Statement in terms of **information***
2. *Statement in terms of **possibilities***
3. *Statement in terms of **payoffs***

Blackwell's Equivalence Theorem:

Theorem

Equivalence between:

1. *Statement in terms of* ***information***
2. *Statement in terms of* ***possibilities***
3. *Statement in terms of* ***payoffs***
4. ***New:*** *Statement in terms of* ***posterior beliefs***
5. ***New:*** *Statement in terms of* ***convex optimization***

Remark

- Fix $E = \{S, \pi\}$ and $E' = \{S', \pi'\}$
- The garbling ranking is silent how signals s and s' are correlated
  - Garbling are in terms of distributions $\pi : \Theta \to \Delta(S)$ and $\pi' : \Theta \to \Delta(S')$
  - Not about linking a realization s from a realization s'
- If $\mathcal{P}$ is a probability space that captures the realisation of (θ, s, s')
  - then ranking E and E' does not pin down the distribution (θ, s, s')

Rick Durrett - Probability: Theory and Examples

Law of Iterated Expectations (Theorem 4.1.13)

Fix a probability space $\mathcal{P} = (\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{F}_2 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}$ be σ-algebras. If the random variable $\mathbf{X}$ satisfies $\mathbb{E}[|\mathbf{X}|] < \infty$, then:

$$\mathbb{E}\,[\,\mathbb{E}[\,\mathbf{X} \mid \mathcal{F}_1] \mid \mathcal{F}_2] = \mathbb{E}\,[\,\mathbb{E}[\,\mathbf{X} \mid \mathcal{F}_2] \mid \mathcal{F}_1] = \mathbb{E}[\mathbf{X} \mid \mathcal{F}_1].$$

- If $\mathcal{F}_2 = \mathcal{F}$ and $\mathcal{F}_1 = \mathcal{F}(\mathbf{Y})$ then:

$$\mathbb{E}[\,\mathbb{E}[\mathbf{X} \mid \mathbf{Y}]\,] = \mathbb{E}[\mathbf{X}]$$

- If $\mathcal{F}_2 = \mathcal{F}(\mathbf{Y}_1)$ and $\mathcal{F}_1 = \mathcal{F}(\mathbf{Y}_1, \mathbf{Y}_2)$ then:

$$\mathbb{E}[\,\mathbb{E}[\mathbf{X} \mid \mathbf{Y}_1, \mathbf{Y}_2] \mid \mathbf{Y}_1\,] = \mathbb{E}[\mathbf{X} \mid \mathbf{Y}_1]$$

Distributions of Beliefs

- We will be interested in objects in the space $\Delta(\Delta(\Theta))$
  - an $F \in \Delta(\Delta(\Theta))$ is called a **distribution of beliefs**
- **Example:**
  - Fix $\Theta = \{\theta_1, \theta_2, \theta_3\}$
  - Fix beliefs $p_1 = (0.6, 0.2, 0.2)$ and $p_2 = (0.1, 0.2, 0.7)$ in $\Delta(\Theta)$
  - $F(p_1) = F(p_2) = \frac{1}{2}$ is a distribution of beliefs

Mean Preserving Spreads

Definition

A distribution of beliefs $F \in \Delta(\Delta(\Theta))$ is a **mean-preserving spread (MPS)** of $\widetilde{F} \in \Delta(\Delta(\Theta))$ if there is some probability space $\mathcal{P}$ with $\Delta(\Theta)$-valued random variables $\boldsymbol{Z}, \tilde{\boldsymbol{Z}}$ such that

1. $\boldsymbol{Z} \sim F, \quad \widetilde{\boldsymbol{Z}} \sim \widetilde{F}$
2. $\mathbb{E}(\boldsymbol{Z}|\widetilde{\boldsymbol{Z}}) = \widetilde{\boldsymbol{Z}}$

- By the law of iterated expectations: $\mathbb{E}(\boldsymbol{Z}) = \mathbb{E}(\widetilde{\boldsymbol{Z}})$
- Exercise: Show (2) is equivalent to the existence of a random variable $\boldsymbol{W}$ such that: $\boldsymbol{Z} = \widetilde{\boldsymbol{Z}} + \boldsymbol{W}$ and $\mathbb{E}[\boldsymbol{W} \mid \widetilde{\boldsymbol{Z}}] = 0$
- Exercise: Show (2) implies $\text{supp}(\widetilde{\boldsymbol{Z}}) \subset conv(\text{supp}(\boldsymbol{Z}))$

Mean Preserving Spreads

Definition

A distribution of beliefs $F \in \Delta(\Delta(\Theta))$ is a **mean-preserving spread (MPS)** of $\widetilde{F} \in \Delta(\Delta(\Theta))$ if there is some probability space $\mathcal{P}$ with $\Delta(\Theta)$-valued random variables $\boldsymbol{Z}, \tilde{\boldsymbol{Z}}$ such that

1. $\boldsymbol{Z} \sim F, \quad \widetilde{\boldsymbol{Z}} \sim \widetilde{F}$
2. $\mathbb{E}(\boldsymbol{Z}|\widetilde{\boldsymbol{Z}}) = \widetilde{\boldsymbol{Z}}$

- By the law of iterated expectations: $\mathbb{E}(\boldsymbol{Z}) = \mathbb{E}(\widetilde{\boldsymbol{Z}})$
- **Exercise:** Show (2) is equivalent to the existence of a random variable $\boldsymbol{W}$ such that: $\boldsymbol{Z} = \widetilde{\boldsymbol{Z}} + \boldsymbol{W}$ and $\mathbb{E}[\boldsymbol{W} \mid \widetilde{\boldsymbol{Z}}] = 0$
- **Exercise:** Show (2) implies $\text{supp}(\widetilde{\boldsymbol{Z}}) \subset conv(\text{supp}(\boldsymbol{Z}))$

Mean Preserving Spreads

Definition

A distribution of beliefs $F \in \Delta(\Delta(\Theta))$ is a **mean-preserving spread (MPS)** of $\widetilde{F} \in \Delta(\Delta(\Theta))$ if there is some probability space $\mathcal{P}$ with $\Delta(\Theta)$-valued random variables $\boldsymbol{Z}, \tilde{\boldsymbol{Z}}$ such that

1. $\boldsymbol{Z} \sim F, \quad \widetilde{\boldsymbol{Z}} \sim \widetilde{F}$
2. $\mathbb{E}(\boldsymbol{Z}|\widetilde{\boldsymbol{Z}}) = \widetilde{\boldsymbol{Z}}$

- By the law of iterated expectations: $\mathbb{E}(\boldsymbol{Z}) = \mathbb{E}(\widetilde{\boldsymbol{Z}})$
- **Exercise:** Show **(2)** is equivalent to the existence of a random variable $\boldsymbol{W}$ such that: $\boldsymbol{Z} = \widetilde{\boldsymbol{Z}} + \boldsymbol{W}$ and $\mathbb{E}[\boldsymbol{W} \mid \widetilde{\boldsymbol{Z}}] = 0$
- **Exercise:** Show **(2)** implies $\text{supp}(\widetilde{\boldsymbol{Z}}) \subset conv(\text{supp}(\boldsymbol{Z}))$

Mean Preserving Spreads

Definition

A distribution of beliefs $F \in \Delta(\Delta(\Theta))$ is a **mean-preserving spread (MPS)** of $\widetilde{F} \in \Delta(\Delta(\Theta))$ if there is some probability space $\mathcal{P}$ with $\Delta(\Theta)$-valued random variables $\boldsymbol{Z}, \tilde{\boldsymbol{Z}}$ such that

1. $\boldsymbol{Z} \sim F, \quad \widetilde{\boldsymbol{Z}} \sim \widetilde{F}$
2. $\mathbb{E}(\boldsymbol{Z}|\widetilde{\boldsymbol{Z}}) = \widetilde{\boldsymbol{Z}}$

- By the law of iterated expectations: $\mathbb{E}(\boldsymbol{Z}) = \mathbb{E}(\widetilde{\boldsymbol{Z}})$
- **Exercise:** Show **(2)** is equivalent to the existence of a random variable $\boldsymbol{W}$ such that: $\boldsymbol{Z} = \widetilde{\boldsymbol{Z}} + \boldsymbol{W}$ and $\mathbb{E}[\boldsymbol{W} \mid \widetilde{\boldsymbol{Z}}] = 0$
- **Exercise:** Show **(2)** implies $\text{supp}(\widetilde{\boldsymbol{Z}}) \subset conv(\text{supp}(\boldsymbol{Z}))$

Definition

A distribution of beliefs $F \in \Delta(\Delta(\Theta))$ **dominates** $\widetilde{F} \in \Delta(\Delta(\Theta))$ **in the convex order** if for every continuous convex function $h : \Delta(\Theta) \to \mathbb{R}$,

$$\int_{\Delta(\Theta)} h(p)\, dF(p) \geq \int_{\Delta(\Theta)} h(p)\, d\widetilde{F}(p).$$

Blackwell's Equivalence Theorem

Theorem

The following statements are equivalent:

1. E' *is a garbling of* E.
2. *For any action set* A, $\Lambda(E', A) \subseteq \Lambda(E, A)$.
3. E *is more informative than* E': $V(E, \mathrm{DP}) \geq V(E', \mathrm{DP})$ *for each* DP.
4. ***New:*** *For any prior* $p \in \Delta(\Theta)$, *if* F *and* F' *are the distributions of posterior beliefs induced by* E *and* E', *then* F *is a MPS of* F'.
5. ***New:*** *For any prior* $p \in \Delta(\Theta)$, *if* F *and* F' *are the distributions of posterior beliefs induced by* E *and* E', *then* F *dominates* F' *in the convex order.*

Proof

- Follow Annie Liang notes
- We have shown: $(1) \iff (2) \iff (3)$
- Suffices to show:
  - $(1) \Rightarrow (4)$,
  - $(4) \Rightarrow (5)$, and
  - $(5) \Rightarrow (1)$

Proof: $(1) \Rightarrow (4)$

- $E = (S, \pi)$ is a garbling of $E' = (S', \pi')$ (with $\gamma : S \to \Delta(S')$).
- Fix $p \in \Delta(\Theta)$ and define probability space $\mathcal{P}$ on $\Omega = \Theta \times S \times S'$ with
  - $\mathbb{P}[\theta, s, s'] = p(\theta) \cdot \pi(s \mid \theta) \cdot \gamma(s' \mid s)$
  - $\boldsymbol{\Theta}$: random realization of state θ
  - $\boldsymbol{S}$ and $\boldsymbol{S'}$: random realizations of signals s and s'
  - $\boldsymbol{Z}$ and $\boldsymbol{Z'}$: random realizations of posterior beliefs of s and s'

$$\boldsymbol{Z}_\theta := \mathbb{E}[\mathbb{1}_\theta[\boldsymbol{\Theta}] \mid \boldsymbol{S}] \qquad \text{and} \qquad \boldsymbol{Z}'_\theta := \mathbb{E}[\mathbb{1}_\theta[\boldsymbol{\Theta}] \mid \boldsymbol{S}']$$

- To show: $\mathbb{E}[\boldsymbol{Z}_\theta \mid \boldsymbol{Z}'] = \boldsymbol{Z}'_\theta$ for each $\theta \in \Theta$

Proof: $(1) \Rightarrow (4)$

In the probability space $\mathcal{P}$: $\boldsymbol{\Theta}$ and $\boldsymbol{S'}$ are correlated only trough $\boldsymbol{S}$

$$\Theta \dashrightarrow S \dashrightarrow S'$$

$$\mathbb{P}[\theta, s, s'] = p(\theta) \cdot \pi(s \mid \theta) \cdot \gamma(s' \mid s)$$

- By construction, for each $(\theta, s, s') \in \Theta \times S \times S'$:

$$\mathbb{P}[\boldsymbol{\Theta} = \theta, \boldsymbol{S} = s' \mid \boldsymbol{S} = s] = \mathbb{P}[\boldsymbol{\Theta} = \theta \mid \boldsymbol{S} = s] \cdot \mathbb{P}[\boldsymbol{\Theta} = \theta \mid \boldsymbol{S} = s]$$

- Hence, $\boldsymbol{\Theta}$ and $\boldsymbol{S'}$ are conditionally independent given $\boldsymbol{S}$:

Proof: $(1) \Rightarrow (4)$

By conditinal independence of $\mathbf{\Theta}$ and $\boldsymbol{S}'$ conditional on $\boldsymbol{S}$:

$$\boldsymbol{Z}_\theta = \mathbb{E}[\mathbb{1}_\theta[\mathbf{\Theta}] \mid \boldsymbol{S}] = \mathbb{E}[\mathbb{1}_\theta[\mathbf{\Theta}] \mid \boldsymbol{S}, \boldsymbol{S}']$$

Then,

$$\begin{aligned} \mathbb{E}[\boldsymbol{Z}_\theta|\boldsymbol{S}'] &= \mathbb{E}[\ \mathbb{E}[\mathbb{1}_\theta[\Theta] \mid \boldsymbol{S}, \boldsymbol{S}'] \mid \boldsymbol{S}'] \\ &= \mathbb{E}[\ \mathbb{1}_\theta[\Theta] \mid \boldsymbol{S}'] && \text{by law of I.E.} \\ &= \boldsymbol{Z}'_\theta \end{aligned}$$

Thus,

$$\begin{aligned} \mathbb{E}[\boldsymbol{Z}_\theta \mid \boldsymbol{Z}'] &= \mathbb{E}[\ \mathbb{E}[\boldsymbol{Z}_\theta \mid \boldsymbol{S}', \boldsymbol{Z}'] \mid \boldsymbol{Z}'] && \text{by law of I.E.} \\ &= \mathbb{E}[\ \mathbb{E}[\boldsymbol{Z}_\theta \mid \boldsymbol{S}'] \mid \boldsymbol{Z}'] && \text{since } \boldsymbol{Z}' \text{ is a function of } \boldsymbol{S}' \\ &= \mathbb{E}[\ \boldsymbol{Z}'_\theta \mid \boldsymbol{Z}'] && \text{by the eq. above} \\ &= \boldsymbol{Z}'_\theta \end{aligned}$$

Proof: $(1) \Rightarrow (4)$

By conditinal independence of $\boldsymbol{\Theta}$ and $\boldsymbol{S}'$ conditional on $\boldsymbol{S}$:

$$\boldsymbol{Z}_\theta = \mathbb{E}[\mathbb{1}_\theta[\boldsymbol{\Theta}] \mid \boldsymbol{S}] = \mathbb{E}[\mathbb{1}_\theta[\boldsymbol{\Theta}] \mid \boldsymbol{S}, \boldsymbol{S}']$$

Then,

$$\begin{aligned} \mathbb{E}[\boldsymbol{Z}_\theta | \boldsymbol{S}'] &= \mathbb{E}[\ \mathbb{E}[\mathbb{1}_\theta[\boldsymbol{\Theta}] \mid \boldsymbol{S}, \boldsymbol{S}'] \mid \boldsymbol{S}'] \\ &= \mathbb{E}[\ \mathbb{1}_\theta[\boldsymbol{\Theta}] \mid \boldsymbol{S}'] && \text{by law of I.E.} \\ &= \boldsymbol{Z}'_\theta \end{aligned}$$

Thus,

$$\begin{aligned} \mathbb{E}[Z_\theta \mid Z'] &= \mathbb{E}[\ \mathbb{E}[Z_\theta \mid S', Z'] \mid Z'] && \text{by law of I.E.} \\ &= \mathbb{E}[\ \mathbb{E}[Z_\theta \mid S'] \mid Z'] && \text{since } Z' \text{ is a function of } S' \\ &= \mathbb{E}[\ Z'_\theta \mid Z'] && \text{by the eq. above} \\ &= Z'_\theta \end{aligned}$$

Proof: $(1) \Rightarrow (4)$

By conditinal independence of $\boldsymbol{\Theta}$ and $\boldsymbol{S}'$ conditional on $\boldsymbol{S}$:

$$\boldsymbol{Z}_\theta = \mathbb{E}[\mathbb{1}_\theta[\boldsymbol{\Theta}] \mid \boldsymbol{S}] = \mathbb{E}[\mathbb{1}_\theta[\boldsymbol{\Theta}] \mid \boldsymbol{S}, \boldsymbol{S}']$$

Then,

$$\begin{aligned}
\mathbb{E}[\boldsymbol{Z}_\theta | \boldsymbol{S}'] &= \mathbb{E}[\ \mathbb{E}[\mathbb{1}_\theta[\boldsymbol{\Theta}] \mid \boldsymbol{S}, \boldsymbol{S}'] \mid \boldsymbol{S}'] \\
&= \mathbb{E}[\ \mathbb{1}_\theta[\boldsymbol{\Theta}] \mid \boldsymbol{S}'] && \text{by law of I.E.} \\
&= \boldsymbol{Z}'_\theta
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathbb{E}[\boldsymbol{Z}_\theta \mid \boldsymbol{Z}'] &= \mathbb{E}[\ \mathbb{E}[\boldsymbol{Z}_\theta \mid \boldsymbol{S}', \boldsymbol{Z}'] \mid \boldsymbol{Z}'] && \text{by law of I.E.} \\
&= \mathbb{E}[\ \mathbb{E}[\boldsymbol{Z}_\theta \mid \boldsymbol{S}'] \mid \boldsymbol{Z}'] && \text{since } \boldsymbol{Z}' \text{ is a function of } \boldsymbol{S}' \\
&= \mathbb{E}[\ \boldsymbol{Z}'_\theta \mid \boldsymbol{Z}'] && \text{by the eq. above} \\
&= \boldsymbol{Z}'_\theta
\end{aligned}$$

Proof: $(4) \Rightarrow (5)$

Assume $\boldsymbol{Z}$ and $\boldsymbol{Z'}$ satisfy $\mathbb{E}[\boldsymbol{Z} \mid \boldsymbol{Z'}] = \boldsymbol{Z'}$. Then, if h is **convex** and **continuous**,

$$\begin{aligned}
\int_{\Delta(\Theta)} h(p)\, dF(p) &= \mathbb{E}[h(\boldsymbol{Z})] \\
&= \mathbb{E}[\, \mathbb{E}[h(\boldsymbol{Z}) \mid \boldsymbol{Z'}]] && \text{by law of I.E.} \\
&\geq \mathbb{E}\,[h(\mathbb{E}[\boldsymbol{Z} \mid \boldsymbol{Z'}])] && \text{by Jensen's inequality} \\
&= \mathbb{E}[h(\boldsymbol{Z'})] && \text{by assumption} \\
&= \int_{\Delta(\Theta)} h(p)\, dF'(p)
\end{aligned}$$

Proof: $(5) \Rightarrow (2)$

- Fix a decision problem $\mathrm{DP} = (A, u, p)$.
- Write $h(q) = \max_{a \in A} \sum_{\theta \in \Theta} q(\theta) \cdot u(a, \theta)$ for the max payoff under $q \in \Delta(\Theta)$.
  - Pointwise max of linear functions.
  - Continuous and convex.
- If F and $\widetilde{F}$ are induced posterior beliefs of E and $\widetilde{E}$, then

$$V(E, \mathrm{DP}) = \int_{\Delta(\Theta)} h(q) dF(q) \geq \int_{\Delta(\Theta)} h(q) d\widetilde{F}(q) = V(\widetilde{E}, \mathrm{DP})$$

Final Remark

- Conditions **(1)-(5)** are hard to verify when $|\Theta| \geq 3$
- Condition **(4)** (F is a MPS of $\widetilde{F}$) is easy to verify when $\Theta = \{0, 1\}$
- Abuse of notation:
  - F, $\widetilde{F}$ are the cdf of the posterior $\mathbb{P}(\boldsymbol{\Theta} = 1 \mid \boldsymbol{S})$, $\mathbb{P}(\boldsymbol{\Theta} = 1 \mid \widetilde{\boldsymbol{S}})$

Lemma

Assume F and $\widetilde{F}$ are distributions on $[0, 1]$. F is **mean preserving spread** of F' if and only if

- $\mathbb{E}_F[\boldsymbol{Z}] = \mathbb{E}_{\widetilde{F}}[\boldsymbol{Z}]$ (Same expected posterior)
- $\displaystyle\int_0^z F(t)dt \geq \int_0^z \widetilde{F}(t)dt$ for each $z \in [0, 1]$ (second-order dominance)

Blackwell's Theorem: Extensions

Extensions

1. **Dynamic decision problems?**
2. **Multiple players?**

Extension I : Dynamic Decision Problems

- What if the decision maker receives information and takes decisions over time?
- What is the right notion of *"dynamic"* garbling?
- Does Blackwell's Theorem hold?
  - What does it exactly state?

Extension I : Dynamic Decision Problems

- Focus on two periods of time: $t = 1, 2$
- **Dynamic experiment:** $E = (S_1, S_2, \pi)$ where $\pi : \Theta \to \Delta(S_1 \times S_2)$
  - $\pi : \Theta \to \Delta(S_1 \times S_2)$
  - S_t: Signal at period t
- **Dynamic decision problem:** $\mathrm{DP} = (A_1, A_2, u, p)$
  - A_t action taken at period t
  - Utility $u : \Theta \times A_1 \times A_2 \to \mathbb{R}$
  - Prior of the state: $p \in \Delta(\Theta)$
- At each period t agent observes $s_t \in S_t$ and chooses action $a_t \in A_t$

Strategies as Adapted Mappings

- **Strategy** (α_1, α_2): $\alpha_1 : S_1 \to \Delta(A_1)$ and $\alpha_2 : S_1 \times S_2 \times A_1 \to \Delta(A_2)$
  - A pair of (α_1, α_2) is not a stochastic mapping.
  - How do we represent (α_1, α_2) as a stochastic mapping?
- $\alpha : S_1 \times S_2 \to \Delta(A_1 \times A_2)$ is **Adapted mapping** if for each $(s_1, s_2) \in S_1 \times S_2$ and $(a_1, a_2) \in A_1 \times A_2$, $\alpha(a_1 \mid s_1, s_2) := \sum_{a_2} \alpha(a_1, a_2 \mid s_1, s_2)$ is constant in s_2.
  - The marginal distribution on A_1 depends only on S_1
  - S_2 does not affect A_1
  - Induces a stochastic mapping $\alpha_1 : S_1 \dashrightarrow A_1$

Strategies as Adapted Mappings

- **Strategy** (α_1, α_2): $\alpha_1 : S_1 \to \Delta(A_1)$ and $\alpha_2 : S_1 \times S_2 \times A_1 \to \Delta(A_2)$
  - A pair of (α_1, α_2) is not a stochastic mapping.
  - How do we represent (α_1, α_2) as a stochastic mapping?
- $\alpha : S_1 \times S_2 \to \Delta(A_1 \times A_2)$ is **Adapted mapping** if for each $(s_1, s_2) \in S_1 \times S_2$ and $(a_1, a_2) \in A_1 \times A_2$, $\alpha(a_1 \mid s_1, s_2) := \sum_{a_2} \alpha(a_1, a_2 \mid s_1, s_2)$ is constant in s_2.
  - The marginal distribution on A_1 depends only on S_1
  - S_2 does not affect A_1
  - Induces a stochastic mapping $\alpha_1 : S_1 \dashrightarrow A_1$

Strategies as Adapted Mappings

- **Strategy** (α_1, α_2): $\alpha_1 : S_1 \to \Delta(A_1)$ and $\alpha_2 : S_1 \times S_2 \times A_1 \to \Delta(A_2)$
  - A pair of (α_1, α_2) is not a stochastic mapping.
  - How do we represent (α_1, α_2) as a stochastic mapping?
- $\alpha : S_1 \times S_2 \to \Delta(A_1 \times A_2)$ is **Adapted mapping** if for each $(s_1, s_2) \in S_1 \times S_2$ and $(a_1, a_2) \in A_1 \times A_2$, $\alpha(a_1 \mid s_1, s_2) := \sum_{a_2} \alpha(a_1, a_2 \mid s_1, s_2)$ is constant in s_2.
  - The marginal distribution on A_1 depends only on S_1
  - S_2 does not affect A_1
  - Induces a stochastic mapping $\alpha_1 : S_1 \dashrightarrow A_1$

Strategies as Adapted Mappings

Lemma

1. Every strategy (α_1, α_2) induces a unique adapted mapping α.
2. Every adapted mapping α induces a unique strategy (α_1, α_2).

To show (1):

$$\alpha(a_1, a_2 \mid s_1, s_2) := \alpha_1(a_1 \mid s_1) \cdot \alpha_2(a_2 \mid s_1, s_2, a_1)$$

To show (2):

α_1 is the induced marginal of α

$$\alpha_2(a_2 \mid s_1, s_2, a_1) := \frac{\alpha(a_1, a_2 \mid s_1, s_2)}{\alpha(a_1 \mid s_1)}$$

Strategies as Adapted Mappings

Lemma

1. Every strategy (α_1, α_2) induces a unique adapted mapping α.
2. Every adapted mapping α induces a unique strategy (α_1, α_2).

To show **(1)**:

$$\alpha(a_1, a_2 \mid s_1, s_2) := \alpha_1(a_1 \mid s_1) \cdot \alpha_2(a_2 \mid s_1, s_2, a_1)$$

To show **(2)**:

α_1 is the induced marginal of α

$$\alpha_2(a_2 \mid s_1, s_2, a_1) := \frac{\alpha(a_1, a_2 | s_1, s_2)}{\alpha(a_1 | s_1)}$$

Strategies as Adapted Mappings

Lemma

1. Every strategy (α_1, α_2) induces a unique adapted mapping α.
2. Every adapted mapping α induces a unique strategy (α_1, α_2).

To show **(1)**:

$$\alpha(a_1, a_2 \mid s_1, s_2) := \alpha_1(a_1 \mid s_1) \cdot \alpha_2(a_2 \mid s_1, s_2, a_1)$$

To show **(2)**:

α_1 is the induced marginal of α

$$\alpha_2(a_2 \mid s_1, s_2, a_1) := \tfrac{\alpha(a_1, a_2 \mid s_1, s_2)}{\alpha(a_1 \mid s_1)}$$

Adapted Garblings as Commuting Diagrams

$$\begin{tikzcd} \Theta \arrow[r, dashed, "\pi"] & S_1 \times S_2 \arrow[r, dashed, "Proj"] \arrow[d, dashed, "\alpha"] & S_1 \arrow[d, dashed, "\alpha_1"] \\ & A_1 \times A_2 \arrow[r, dashed, "Proj"] & A_1 \end{tikzcd}$$

α is **adapted** if there exists α_1 that make the diagram commute

Adapted Garblings

Definition

an experiment $E' = (S'_1, S'_2, \pi)$ is an **adapted garbling** of $E = (S_1, S_2, \pi)$ if there is some $\gamma : (S_1 \times S_2) \to \Delta(S'_1 \times S'_2)$ such that

1. $\pi' = \gamma \circ \pi$
2. $\gamma(s'_1 \mid s_1, s_2) := \sum_{s'_2} \gamma(s'_1, s'_2 \mid s_1, s_2)$ is constant in s_2

- "s_2 has no effect on s'_1"
- Notice: (2) induces $\gamma_1 : S_1 \to \Delta(S'_1)$

Adapted Garblings

Definition

an experiment $E' = (S'_1, S'_2, \pi)$ is an adapted garbling of $E = (S_1, S_2, \pi)$ if there is some $\gamma : (S_1 \times S_2) \to \Delta(S'_1 \times S'_2)$ such that

1. $\pi' = \gamma \circ \pi$
2. $\gamma(s'_1 \mid s_1, s_2) := \sum_{s'_2} \gamma(s'_1, s'_2 \mid s_1, s_2)$ is constant in s_2

- "s_2 has no effect on s'_1"
- Notice: (2) induces $\gamma_1 : S_1 \to \Delta(S'_1)$

Conditional Distribution of Actions

$$
\begin{tikzcd}
 & \Theta \arrow[r, dashed, "\pi"] \arrow[d, dashed, "\pi'"] & S_1 \times S_2 \arrow[d, dashed, "\alpha"] \\
 & S'_1 \times S'_2 \arrow[r, dashed, "\alpha'"] & A_1 \times A_2
\end{tikzcd}
$$

- For each $E = (S_1, S_2, \pi)$ and A_1, A_2, define

$$\Lambda(E, A_1, A_2) := \{\alpha \circ \pi \mid \alpha : S_1 \times S_2 \dashrightarrow A_1 \times A_2 \text{ is } \textbf{adapted}\}$$

- When does $\Lambda(E', A_1, A_2) \subseteq \Lambda(E, A_1, A_2)$ holds ?
  - Holds if for each **adapted** $\alpha' : S'_1 \times S'_2 \dashrightarrow A_1 \times A_2$ there exists some **adapted** $\alpha : S_1 \times S_2 \dashrightarrow A_1 \times A_2$ so that $\alpha \circ \pi = \alpha' \circ \pi'$

Informativeness Characterization in Dynamic Settings

Theorem

The following statements are equivalent:

1. E' *is an adapted garbling of* E.
2. *For any action sets* A_1, A_2, $\Lambda(E', A_1, A_2) \subseteq \Lambda(E, A_1, A_2)$.
3. E *is more informative than* E' : $V(E, \mathrm{DP}) \geq V(E', \mathrm{DP})$

- What about (4) (MPS) and (5) (convex optimization) ?
- Proof:
  - (2) $\iff$ (3) analogous argument as before
  - Suffices to show (1) $\iff$ (2)

Informativeness Characterization in Dynamic Settings

Theorem

The following statements are equivalent:

1. E' *is an adapted garbling of* E.
2. *For any action sets* A_1, A_2, $\Lambda(E', A_1, A_2) \subseteq \Lambda(E, A_1, A_2)$.
3. E *is more informative than* E' : $V(E, \mathrm{DP}) \geq V(E', \mathrm{DP})$

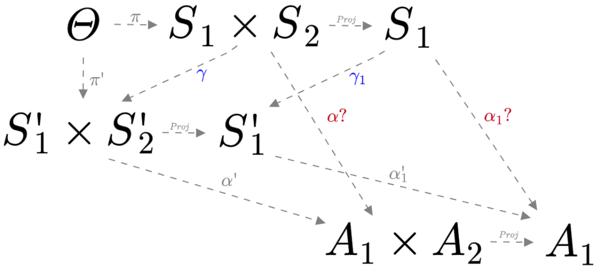
- What about **(4)** (MPS) and **(5)** (convex optimization) ?
- Proof:
  - (2) $\iff$ (3) analogous argument as before
  - Suffices to show (1) $\iff$ (2)

$$
\begin{tikzcd}
\Theta \arrow[r, dashed, "\pi"] \arrow[d, dashed, "\pi'"] & S_1 \times S_2 \arrow[r, dashed, "Proj"] & S_1 & \\
S_1' \times S_2' \arrow[r, dashed, "Proj"] \arrow[drr, dashed, "\alpha'"] & S_1' \arrow[drr, dashed, "\alpha_1'"] & & \\
 & & A_1 \times A_2 \arrow[r, dashed, "Proj"] & A_1
\end{tikzcd}
$$

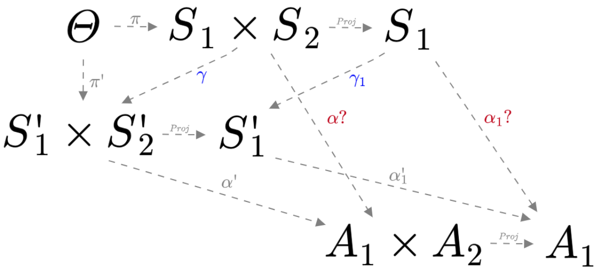
- **Hypothesis:** E' is a garbling of E
- **Goal:** $\Lambda(E, A_1, A_2) \subseteq \Lambda(E', A_1, A_2)$

$$\begin{tikzcd}
\Theta \arrow[r, dashed, "\pi"] \arrow[d, dashed, "\pi'"] & S_1 \times S_2 \arrow[r, dashed, "Proj"] \arrow[ddr, dashed, "\alpha?"] & S_1 \arrow[ddr, dashed, "\alpha_1?"] & \\
S'_1 \times S'_2 \arrow[r, dashed, "Proj"] \arrow[drr, dashed, "\alpha'"] & S'_1 \arrow[drr, dashed, "\alpha'_1"] & & \\
 & & A_1 \times A_2 \arrow[r, dashed, "Proj"] & A_1
\end{tikzcd}$$

- **Hypothesis:** E' is a garbling of E
- **Goal:** Find α, α_1 that make the diagram to commute



- **Hypothesis:** Exist γ, γ_1 that make the diagram to commute
- **Goal:** Find α, α_1 that make the diagram to commute



- **Hypothesis:** Exist γ, γ_1 that make the diagram to commute
- **Solution:** Take $\alpha = \alpha' \circ \gamma$ and $\alpha_1 = \alpha'_1 \circ \gamma_1$

Part II of Proof: (2) Implies (1)

$$\begin{tikzcd}
& \Theta \arrow[r, dashed, "\pi"] \arrow[d, dashed, "\pi'"] & S_1 \times S_2 \arrow[r, dashed, "Proj"] & S_1 \\
S'_1 \times S'_2 \arrow[r, dashed, "Proj"] & S'_1 & &
\end{tikzcd}$$

- **Hypothesis:** $\Lambda(E', A_1, A_2) \subseteq \Lambda(E, A_1, A_2)$ for each A_1, A_2
- **Goal:** show E' is a garbling of E

$$\begin{tikzcd}
\Theta \arrow[r, dashed, "\pi"] \arrow[d, dashed, "\pi'"] & S_1 \times S_2 \arrow[r, dashed, "Proj"] & S_1 & \\
S'_1 \times S'_2 \arrow[r, dashed, "Proj"] \arrow[drr, dashed, "Id"] & S'_1 \arrow[drr, dashed, "Id"] & & \\
& & S'_1 \times S'_2 \arrow[r, dashed, "Proj"] & S'_1
\end{tikzcd}$$

- **Hypothesis:** $\Lambda(E', A_1 = S_1, A_2 = S_2) \subseteq \Lambda(E, A_1 = S_1, A_2 = S_2)$
- **Goal:** show E' is a garbling of E

Part II of Proof: (2) Implies (1)

$$\begin{tikzcd}
\Theta \arrow[r, dashed, "\pi"] \arrow[d, dashed, "\pi'"] & S_1 \times S_2 \arrow[r, dashed, "Proj"] \arrow[ddr, dashed, "\alpha"] & S_1 \arrow[ddr, dashed, "\alpha_1"] & \\
S'_1 \times S'_2 \arrow[r, dashed, "Proj"] \arrow[drr, dashed, "Id"] & S'_1 \arrow[drr, dashed, "Id"] & & \\
& & S'_1 \times S'_2 \arrow[r, dashed, "Proj"] & S'_1
\end{tikzcd}$$

- **Hypothesis:** there are some α, α_1 that make diagram commute
- **Goal:** Show E' is an adaptive garbling of E

$$\begin{tikzcd}
\Theta \arrow[r, dashed, "\pi"] \arrow[d, dashed, "\pi'"] & S_1 \times S_2 \arrow[r, dashed, "Proj"] \arrow[ddr, dashed, "\alpha"] & S_1 \arrow[ddr, dashed, "\alpha_1"] & \\
S'_1 \times S'_2 \arrow[r, dashed, "Proj"] \arrow[drr, dashed, "Id"] & S'_1 \arrow[drr, dashed, "Id"] & & \\
& & S'_1 \times S'_2 \arrow[r, dashed, "Proj"] & S'_1
\end{tikzcd}$$

- **Hypothesis:** there are some α, α_1 that make diagram commute
- **Solution:** Take $\gamma = \alpha$, $\gamma_1 = \alpha_1$

Extension II : Multiple Agents

- Suppose that there are n agents.
- Suppose that each agent observes a **private** signal of a given experiment.
- What is the right notion of garbling?
- Does Blackwell's theorem holds for games?
  - Some parts hold
  - Some parts do not

Extension II : Multiple Agents

- Suppose that there are n agents.
- Suppose that each agent observes a **private** signal of a given experiment.
- What is the right notion of garbling?
- Does Blackwell's theorem holds for games?
  - Some parts hold
  - Some parts do not

Extension II : Multiple Agents

- Suppose that there are n agents.
- Suppose that each agent observes a **private** signal of a given experiment.
- What is the right notion of garbling?
- Does Blackwell's theorem holds for games?
  - Some parts hold
  - Some parts do not

Extension II : Multiple Agents

- Focus on 2 agents
- **Experiment:** $E = (S_1, S_2, \pi)$ where $\pi : \Theta \to \Delta(S_1 \times S_2)$
  - Agent i observes S_i
- **Game:** $G = (A_1, A_2, u_1, u_2, p)$ where each $u_i : \Theta \times A_1 \times A_2 \to \mathbb{R}$
  1. i's strategy: $\alpha_i : S_i \to \Delta(A_i)$
  2. A profile $\alpha = (\alpha_1, \alpha_2)$ induces $\alpha : (S_1 \times S_2) \dashrightarrow (A_1 \times A_2)$
     - Induces $\alpha \circ \pi : \Theta \dashrightarrow (A_1 \times A_2)$
- What is the right notion of garbling?

Extension II : Multiple Agents

- Focus on 2 agents
- **Experiment:** $E = (S_1, S_2, \pi)$ where $\pi : \Theta \to \Delta(S_1 \times S_2)$
  - Agent i observes S_i
- **Game:** $G = (A_1, A_2, u_1, u_2, p)$ where each $u_i : \Theta \times A_1 \times A_2 \to \mathbb{R}$
  1. i's strategy: $\alpha_i : S_i \to \Delta(A_i)$
  2. A profile $\alpha = (\alpha_1, \alpha_2)$ induces $\alpha : (S_1 \times S_2) \dashrightarrow (A_1 \times A_2)$
     - Induces $\alpha \circ \pi : \Theta \dashrightarrow (A_1 \times A_2)$
- What is the right notion of garbling?

Extension II : Multiple Agents

- Focus on 2 agents
- **Experiment:** $E = (S_1, S_2, \pi)$ where $\pi : \Theta \to \Delta(S_1 \times S_2)$
  - Agent i observes S_i
- **Game:** $G = (A_1, A_2, u_1, u_2, p)$ where each $u_i : \Theta \times A_1 \times A_2 \to \mathbb{R}$
  1. i's strategy: $\alpha_i : S_i \to \Delta(A_i)$
  2. A profile $\alpha = (\alpha_1, \alpha_2)$ induces $\alpha : (S_1 \times S_2) \dashrightarrow (A_1 \times A_2)$
     - Induces $\alpha \circ \pi : \Theta \dashrightarrow (A_1 \times A_2)$
- What is the right notion of garbling?

Independent Garblings

Definition

Say $E' = (S_1', S_2', \pi')$ is an **independent garbling** of $E = (S_1, S_2, \pi)$ if there is some $\gamma : (S_1 \times S_2) \dashrightarrow (S_1' \times S_2')$ such that

- $\pi' = \gamma \circ \pi$
- $\gamma(s_1', s_2' \mid s_1, s_2) = \gamma_1(s_1' \mid s_1) \cdot \gamma_2(s_2' \mid s_2)$ for some mappings γ_1, γ_2.

Informativeness Characterization for Games

Theorem

The following statements are equivalent:

1. E' *is an independent garbling of* E.
2. *For any action sets* A_1, A_2, $\Lambda(E, A_1, A_2) \subseteq \Lambda(E', A_1', A_2')$.

- Silent about payoffs
- Proof:
  - Verify that the set of independent garblings forms a category
  - Use same diagrams as before

Informativeness Characterization for Games

Theorem

The following statements are equivalent:

1. E' *is an independent garbling of* E.
2. *For any action sets* A_1, A_2, $\Lambda(E, A_1, A_2) \subseteq \Lambda(E', A_1', A_2')$.

- Silent about payoffs
- Proof:
  - Verify that the set of independent garblings forms a category
  - Use same diagrams as before

Informativeness Characterization for Games

Theorem

The following statements are equivalent:

1. E' *is an independent garbling of* E.
2. *For any action sets* A_1, A_2, $\Lambda(E, A_1, A_2) \subseteq \Lambda(E', A_1', A_2')$.

- Silent about payoffs
- Proof:
  - Verify that the set of independent garblings forms a category
  - Use same diagrams as before

Blackwell's theorem states that "more information" is always weakly better

- True for **individual decision problems**
- False for multiple agents playing a **Game**
  - More information can make each player strictly worse off
  - Example?

Remark

Blackwell's theorem states that "more information" is always weakly better

- True for **individual decision problems**
- False for multiple agents playing a **Game**
  - More information can make each player strictly worse off
  - Example?

Example: More Information Hurts Players

State $\theta = 1$ or $\theta = 0$ with equal probability

	D_0	C	D_1
D_0	(0,0)	(2,-1)	(-9,-9)
C	(-1,2)	(1,1)	(-9,-9)
D_1	(-9,-9)	(-9,-9)	(-11,-11)

$\theta = 0$

	D_0	C	D_1
D_0	(-11,-11)	(-9,-9)	(-9,-9)
C	(-9,-9)	(1,1)	(-1,2)
D_1	(-9,-9)	(2,-1)	(0,0)

$\theta = 1$

Example: More Information Hurts Players

State $\theta = 1$ or $\theta = 0$ with equal probability

	D_0	C	D_1
D_0	(0,0)	(2,-1)	(-9,-9)
C	(-1,2)	(1,1)	(-9,-9)
D_1	(-9,-9)	(-9,-9)	(-11,-11)

$\theta = 0$

	D_0	C	D_1
D_0	(-11,-11)	(-9,-9)	(-9,-9)
C	(-9,-9)	(1,1)	(-1,2)
D_1	(-9,-9)	(2,-1)	(0,0)

$\theta = 1$

- Prisoner's dilemma with a Bomb!
- If agents know θ: Only equilibrium payoffs are $(0, 0)$

What if agents do not know θ?

Average payoffs for **common prior** $(\frac{1}{2}, \frac{1}{2})$:

	D_0	C	D_1
D_0	(-5.5, -5.5)	(-3.5, -5)	(-9, -9)
C	(-5, -3.5)	(1, 1)	(-5, -3.5)
D_1	(-9, -9)	(-3.5, -5)	(-5.5, -5.5)

- If agents do not know where the bomb is only eq. payoff is $(1, 1)$
- If agents know where the bomb is only eq. payoff is $(0, 0)$

What if agents do not know θ?

Average payoffs for **common prior** $(\frac{1}{2}, \frac{1}{2})$:

	D_0	C	D_1
D_0	(-5.5, -5.5)	(-3.5, -5)	(-9, -9)
C	(-5, -3.5)	(1, 1)	(-5, -3.5)
D_1	(-9, -9)	(-3.5, -5)	(-5.5, -5.5)

- If agents do not know where the bomb is only eq. payoff is $(1, 1)$
- If agents know where the bomb is only eq. payoff is $(0, 0)$

What if agents do not know θ?

Average payoffs for **common prior** $(\frac{1}{2}, \frac{1}{2})$:

	D_0	C	D_1
D_0	(-5.5, -5.5)	(-3.5, -5)	(-9, -9)
C	(-5, -3.5)	(1, 1)	(-5, -3.5)
D_1	(-9, -9)	(-3.5, -5)	(-5.5, -5.5)

- If agents do not know where the bomb is only eq. payoff is $(1, 1)$
- If agents know where the bomb is only eq. payoff is $(0, 0)$

Alternative Rankings of Experiments

Other Rankings of Information

- **Stronger than Blackwell**
  - Strong Blackwell order (Brooks, Frankel, Kamenica, 2022)
  - Sufficiency order
  - Strong martingale order
- **Weaker than Blackwell**
  - Blackwell' ranking on dichotomies
  - Lehmann's ranking (Lehmann, 1988)

Strong Blackwell Order

Brooks, Frankel, Kamenica (2022)

Strong Blackwell Order

- Blackwell's Theorem assumes absence of complementary information.
  - Assume X is more informative than Y in Blackwell's order
  - In addition the agent observes signal Z
  - Is (X, Z) more informative than (Y, Z)?
    - Not necessarily
    - **Important:** complementarity and substitutitability of experiments
- Can we find an order that is robust to complementary information?
  - Yes! Brooks, Frankel, Kamenica (2023)

Strong Blackwell Order

- Blackwell's Theorem assumes absence of complementary information.
  - Assume X is more informative than Y in Blackwell's order
  - In addition the agent observes signal Z
  - Is (X, Z) more informative than (Y, Z)?
    - Not necessarily
    - **Important:** complementarity and substitutitability of experiments
- Can we find an order that is robust to complementary information?
  - Yes! Brooks, Frankel, Kamenica (2023)

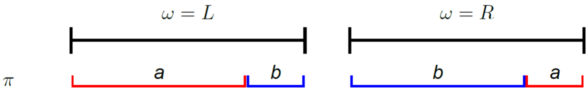
Strong Blackwell Order

- Blackwell's Theorem assumes absence of complementary information.
    - Assume X is more informative than Y in Blackwell's order
    - In addition the agent observes signal Z
    - Is (X, Z) more informative than (Y, Z)?
        - Not necessarily
        - **Important:** complementarity and substitutitability of experiments
- Can we find an order that is robust to complementary information?
    - Yes! Brooks, Frankel, Kamenica (2023)

Signals vs. Experiments

State space: Ω

- **Experiment:** (S, π) where $\pi : \Omega \to \Delta(S)$
- **Signal:** π is a finite partition of $\Omega \times [0, 1]$

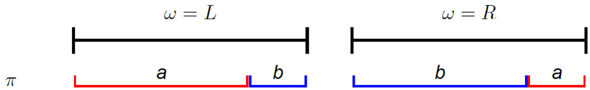


- Induces an stochastic mapping $\Omega \to \Delta(\{a, b\})$
- Captures correlation between experiments

Signals vs. Experiments

State space: Ω

- **Experiment:** (S, π) where $\pi : \Omega \to \Delta(S)$
- **Signal:** π is a finite partition of $\Omega \times [0, 1]$

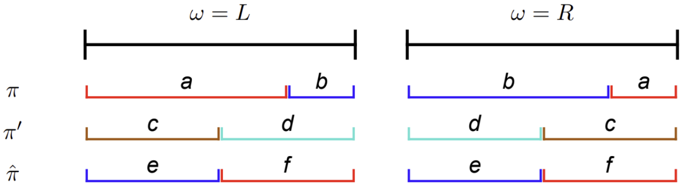


  - Induces an stochastic mapping $\Omega \to \Delta(\{a, b\})$
  - Captures correlation between experiments

Signals vs. Experiments

State space: Ω

- **Experiment:** (S, π) where $\pi : \Omega \to \Delta(S)$
- **Signal:** π is a finite partition of $\Omega \times [0, 1]$



- Induces an stochastic mapping $\Omega \to \Delta(\{a, b\})$
- Captures correlation between experiments

Blackwell vs. Strong Blackwell

Definition

Say π **Blackwell dominates** π' (denoted $\pi \succsim_B \pi'$) if for each decision problem DP,

$$V(\pi, \mathrm{DP}) \geq V(\pi', \mathrm{DP})$$

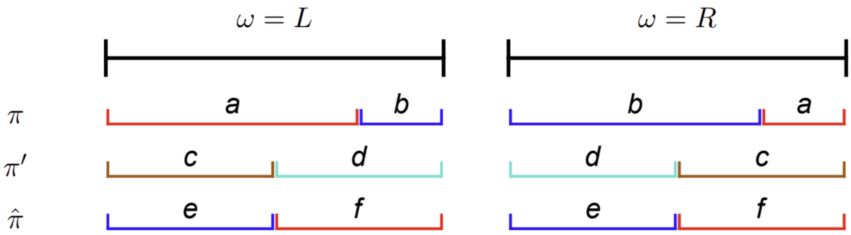
The **meet** of two signals $\pi \vee \hat{\pi}$: Partition created by observing π and $\hat{\pi}$

Definition

Say π **strongly Blackwell dominates** π' (denoted $\pi \succsim_{SB} \pi'$)) if for each signal $\hat{\pi}$,

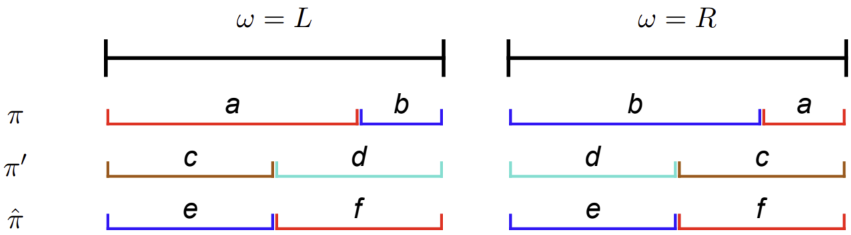
$$\pi \vee \hat{\pi} \succsim_B \pi' \vee \hat{\pi}$$

Blackwell vs. Strong Blackwell



$\pi \succsim_B \pi'$ does not imply $\pi \succsim_{SB} \pi'$

Blackwell vs. Strong Blackwell



$\pi \succsim_B \pi'$ does not imply $\pi \succsim_{SB} \pi'$

Reveal or Refines

Definition

Say π **reveals-or-refines** π' if each $s \in \pi$ either

- **Reveals the state**
- **Reveals the other signal π'**

Reveal or Refines

Definition

Say π **reveals-or-refines** π' if each $s \in \pi$ either

- **Reveals the state:** $\mathbb{P}(s \mid \omega) > 0$ for at most one ω
- **Reveals the other signal π':** $s \subseteq s'$ for some $s' \in \pi'$

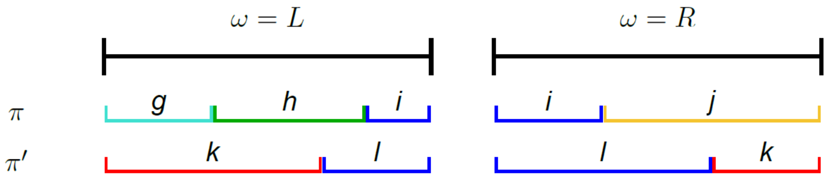
Reveal or Refines

Definition

Say π **reveals-or-refines** π' if each $s \in \pi$ either

- **Reveals the state:** $\mathbb{P}(s \mid \omega) > 0$ for at most one ω
- **Reveals the other signal π':** $s \subseteq s'$ for some $s' \in \pi'$

Example:



Characterization of Strong Blackwell Order

Theorem

Signal π **reveals-or-refines** π' if and only if $\pi \succsim_{SB} \pi'$

$\Rightarrow$ Easy: Check signal by signal

- If $s \in \pi$ reveals ω, then there is nothing else to reveal
- If $s \in \pi$ reveals signal of π', then $\pi' \vee \hat{\pi}$ cannot do better than $\pi \vee \hat{\pi}$

$\Leftarrow$ By contrapositive:

- If some $s \in \pi$ does not reveal ω or π', then construct $\hat{\pi}$ that makes π' is more attractive than π

Characterization of Strong Blackwell Order

Theorem

Signal π **reveals-or-refines** π' if and only if $\pi \succsim_{SB} \pi'$

$\Rightarrow$ Easy: Check signal by signal

- If $s \in \pi$ reveals ω, then there is nothing else to reveal
- If $s \in \pi$ reveals signal of π', then $\pi' \vee \hat{\pi}$ cannot do better than $\pi \vee \hat{\pi}$

$\Leftarrow$ By contrapositive:

- If some $s \in \pi$ does not reveal ω or π', then construct $\hat{\pi}$ that makes π' is more attractive than π

Characterization of Strong Blackwell Order

Theorem

Signal π **reveals-or-refines** π' if and only if $\pi \succsim_{SB} \pi'$

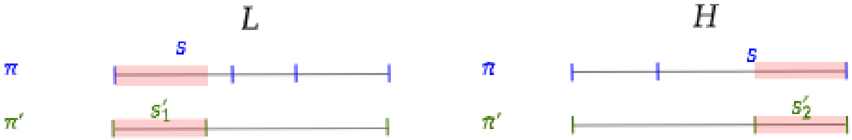
$\Rightarrow$ Easy: Check signal by signal

- If $s \in \pi$ reveals ω, then there is nothing else to reveal
- If $s \in \pi$ reveals signal of π', then $\pi' \vee \hat{\pi}$ cannot do better than $\pi \vee \hat{\pi}$

$\Leftarrow$ By contrapositive:

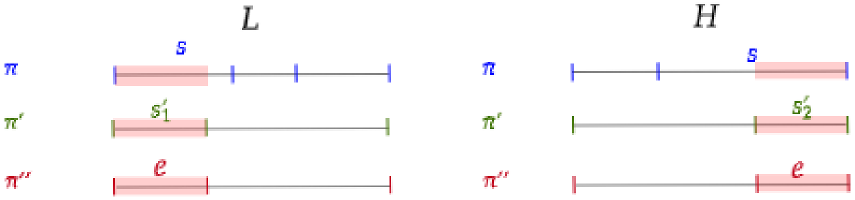
- If some $s \in \pi$ does not reveal ω or π', then construct $\hat{\pi}$ that makes π' is more attractive than π

Proof ($\Leftarrow$)

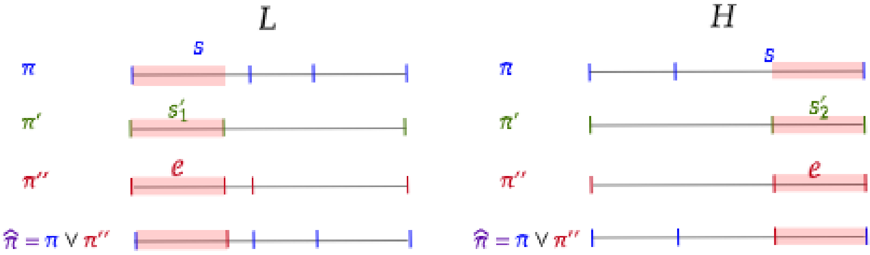


- π is not informative given $\hat{\pi}$ $\Rightarrow$ $\pi \vee \hat{\pi} \sim \hat{\pi}$
- π' is (very) informative given $\hat{\pi}$ $\Rightarrow$ $\pi' \vee \hat{\pi} \succ \hat{\pi}$

Proof ($\Leftarrow$)



- π is not informative given $\hat{\pi}$ $\quad\Rightarrow\quad$ $\pi \vee \hat{\pi} \sim \hat{\pi}$
- π' is (very) informative given $\hat{\pi}$ $\quad\Rightarrow\quad$ $\pi' \vee \hat{\pi} \succ \hat{\pi}$



- π is not informative given $\hat{\pi}$ $\quad\Rightarrow\quad$ $\pi \vee \hat{\pi} \sim \hat{\pi}$
- π' is (very) informative given $\hat{\pi}$ $\quad\Rightarrow\quad$ $\pi' \vee \hat{\pi} \succ \hat{\pi}$

Blackwell's Order on Dichotomies

Difficulties of Blackwell's Ranking

When $|\Theta| \geq 3$ Blackwell's ranking is hard to verify:

> *"For two given experiments ... neither (the definition of more informative) ... nor any of the six equivalent conditions ... yields a systematic method of determining whether (one experiment is more informative than another)."*
>
> *– Blackwell and Girshick (1954)*

Example:

- Unknown state $\theta = (\mu, \sigma^2)$
- Two signals: $X \sim N(\mu, \sigma^2)$ and $Y \sim N(\mu, 0.25 \cdot \sigma^2)$
  - Not comparable in the Blackwell sense

Difficulties of Blackwell's Ranking

When $|\Theta| \geq 3$ Blackwell's ranking is hard to verify:

> *"For two given experiments ... neither (the definition of more informative) ... nor any of the six equivalent conditions ... yields a systematic method of determining whether (one experiment is more informative than another)."*
>
> *– Blackwell and Girshick (1954)*

Example:

- Unknown state $\theta = (\mu, \sigma^2)$
- Two signals: $X \sim N(\mu, \sigma^2)$ and $Y \sim N(\mu, 0.25 \cdot \sigma^2)$
  - Not comparable in the Blackwell sense

Blackwell's Ranking with Binary States

When $|\Theta| = 2$ Blackwell's ranking is easy to verify:

Fix distributions F and $\widetilde{F}$ on $[0, 1]$.

Lemma

F is MPS of $\widetilde{F}$ if and only if:

1. Same posterior mean: $\mathbb{E}_F[\boldsymbol{Z}] = \mathbb{E}_{\widetilde{F}}[\widetilde{\boldsymbol{Z}}]$
2. Second-order stochastic dominance: for each $z \in [0, 1]$

$$\int_0^t F(z)dz \geq \int_0^t \widetilde{F}(z)dz$$

Proof: $\Rightarrow$

Assume F is a MPS of $\widetilde{F}$: There is some probability space with $\boldsymbol{Z} \sim F$ and $\widetilde{\boldsymbol{Z}} \sim \widetilde{F}$

- Part (1): Law of iterated expectations

$$\mathbb{E}[\boldsymbol{Z}] = \mathbb{E}[\mathbb{E}[\boldsymbol{Z} \mid \widetilde{\boldsymbol{Z}}]] = \mathbb{E}[\widetilde{\boldsymbol{Z}}]$$

- Part (2): Recall MPS is equivalent to dominance in the convex order
    - The function $\phi_t(x) := (t-x) \cdot \mathbb{1}[x \leq t]$ is convex
    - Dominance in the convex order: $\displaystyle\int_0^1 \phi_t(x) dFx \geq \int_0^1 \phi_t(x) d\widetilde{F}x$
    - Integration by parts: $\displaystyle\int_0^t F(z)dz = \int_0^t (t-x)dFx = \int_0^1 \phi_t(x)dFx$

Proof: $\Rightarrow$

Exercise (Proposition 3 in Tilman's notes)

If (1) F and $\widetilde{F}$ have the same mean, and

(2) F second-order stochastically dominates $\widetilde{F}$,

Then F is dominates $\widetilde{F}$ in the convex order.

By previous slide (2) is equivalent to:

$$\int_0^1 \phi_t(x) dFx \geq \int_0^1 \phi_t(x) d\widetilde{F}x \qquad \text{for each } t \in [0,1]$$

To show: F dominates $\widetilde{F}$ in the convex order

$$\int_0^1 h(x) dFx \geq \int_0^1 h(x) d\widetilde{F}x \text{ for each convex function } h$$

Proof: $\Rightarrow$

Exercise (Proposition 3 in Tilman's notes)

If (1) F and $\widetilde{F}$ have the same mean, and

(2) F second-order stochastically dominates $\widetilde{F}$,

Then F is dominates $\widetilde{F}$ in the convex order.

By previous slide (2) is equivalent to:

$$\int_0^1 \phi_t(x) dFx \geq \int_0^1 \phi_t(x) d\widetilde{F}x \qquad \text{for each } t \in [0,1]$$

To show: F dominates $\widetilde{F}$ in the convex order

$$\int_0^1 h(x) dFx \geq \int_0^1 h(x) d\widetilde{F}x \text{ for each convex function } h$$

Blackwell Ranking in Dichotomies

For each experiment $E = (S, \pi)$ and each subset $\Omega \subset \Theta$, write $E_\Omega = (S, \pi_\Omega)$ for the experiment restricted to the state space Ω.

Definition

An experiment $E = (S, \pi)$ dominates $E' = (S', \pi')$ in the **Blackwell on dichotomies order** if E_Ω dominates E'_Ω for each dichotomous subset $\Omega = \{\theta, \theta'\} \subset \Theta$.

- Blackwell dominance implies Blackwell dominance on dichotomies
  - Why?
- Blackwell's ranking on dichotomies is (strictly) weaker than Blackwell's ranking
  - Example? Exercise

Blackwell Ranking in Dichotomies

For each experiment $E = (S, \pi)$ and each subset $\Omega \subset \Theta$, write $E_\Omega = (S, \pi_\Omega)$ for the experiment restricted to the state space Ω.

Definition

An experiment $E = (S, \pi)$ dominates $E' = (S', \pi')$ in the **Blackwell on dichotomies order** if E_Ω dominates E'_Ω for each dichotomous subset $\Omega = \{\theta, \theta'\} \subset \Theta$.

- Blackwell dominance implies Blackwell dominance on dichotomies
  - Why?
- Blackwell's ranking on dichotomies is (strictly) weaker than Blackwell's ranking
  - Example? **Exercise**

Blackwell Ranking in Dichotomies

For each experiment $E = (S, \pi)$ and each subset $\Omega \subset \Theta$, write $E_\Omega = (S, \pi_\Omega)$ for the experiment restricted to the state space Ω.

Definition

An experiment $E = (S, \pi)$ dominates $E' = (S', \pi')$ in the **Blackwell on dichotomies order** if E_Ω dominates E'_Ω for each dichotomous subset $\Omega = \{\theta, \theta'\} \subset \Theta$.

- Blackwell dominance implies Blackwell dominance on dichotomies
  - Why?
- Blackwell's ranking on dichotomies is (strictly) weaker than Blackwell's ranking
  - Example? **Exercise**

Open Question?

- Blackwell ranking ranks information for **all** decision problems
- Blackwell ranking in dichotomies ranks information for **some** decision problems
- **Open question:** What is this subclass of decision problems?
  - It is simple to characterize?
  - Nice applications?

Hint?

Fix a binary state space $\Theta = \{\underline{\theta}, \overline{\theta}\}$. Call a decision problem $DP = \{A, u, p\}$ **simple** $A = \{L, H\}$ and there is some $k \in \mathbb{R}$ so that

	$\underline{\theta}$	$\overline{\theta}$
H	0	$-k$
L	-1	0

Exercise (See Tilman's Notes)

Assume Θ is binary. Experiment E dominates $\widetilde{E}$ in the Blackwell order if and only if

$$V(E, \mathrm{DP}) \geq V(\widetilde{E}, \mathrm{DP})$$

for each **simple** decision problem DP.

Lehmann's Ranking of Experiments

Restrictions on Information Comparison

- Blackwell order compares among **all** experiments, and **all** decision problems
- What can we say if we restrict the environment?

Restrictions on Environment

- Restrict to **“monotone” decision problems**
  - $\Theta = [\underline{\theta}, \overline{\theta}] \subset \mathbb{R}, \quad A \subset \mathbb{R}.$
  - $U : \Theta \times A \to \mathbb{R}$ satisfies single crossing condition / KR monotonicity

    “Higher optimal actions for higher states”

- Restrict to **“monotone” experiments**
  - $F : \Theta \to \Delta(X), \quad X = [\underline{x}, \overline{x}] \subset \mathbb{R}$
    - Interpret $F(\cdot \mid \theta)$ as a CDF
  - F satisfies monotone likelihood ratio property (MLRP)

    “Higher signals indicate higher states”

Monotone Decision problems

Fix $\Theta \subset \mathbb{R}$ and $A \subset \mathbb{R}$ and utility function $u : \Theta \times A \to \mathbb{R}$

Definition (Karlin and Rubin,1956)

The utility function u is **KR-monotone** if it is quasiconvex in a, each action is best for some state, and the set of best actions given θ is increasing in strict set order in θ.

Definition

The utility $u : \Theta \times A \to \mathbb{R}$ satisfies the **single crossing property** if for any actions $a_\ell < a_h \in A$ and states $\theta_\ell < \theta_h \in \Theta$, the following holds:

$$u(\theta_\ell, a_h) \geq u(\theta_\ell, a_\ell) \quad \text{implies} \quad u(\theta_h, a_h) \geq u(\theta_h, a_\ell)$$

- KR-monotonicity is not weaker nor stronger than the single crossing condition.

Monotone Experiments

Interpret $F(\cdot \mid \theta)$ as the cdf of the signal x in terms of the state θ and $f(\cdot \mid \theta)$ the associated density

Definition

The distribution F satisfies the **monotone likelihood ratio property (MLRP)** if for each $\theta_\ell < \theta_h$ and for all x in the support of F,

$$\frac{f(x \mid \theta_h)}{f(x \mid \theta_\ell)} \text{ is non-decreasing in } x.$$

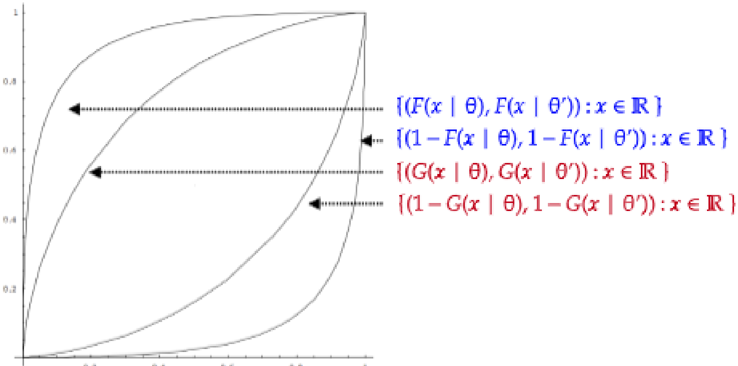
Implies first-order stochastic dominance of $\{F(\cdot \mid \theta)\}$ over θ

Interpret $F(\cdot \mid \theta)$ as the cdf of the signal x in terms of the state θ

Definition

F dominates G in **Lehmann's ranking** if for each $\theta < \theta'$ and each $x \in \mathbb{R}$

$$F^{-1}(G(x \mid \theta) \mid \theta) \leq F^{-1}(G(x \mid \theta') \mid \theta')$$



Lehmann/Persico Theorem

Theorem - (Lehmann/Persico)

Suppose F and G satisfy the monotone likelihood property and F dominates G according to the Lehman's ranking. Then, for each **monotone** decision problem DP,

$$V(F, \mathrm{DP}) \geq V(G, \mathrm{DP})$$

Relation between Lehmann and Blackwell

Theorem - Jewitt(2007)

Restricted to experiments that satisfy the **MLRP** condition, **Lehmann's ranking** is equivalent to the **Blackwell ranking on dichotomies**

- Blackwell ranking on dichotomies is easy to verify.
- **Remark:** This holds only over the MLRP condition.

Relation between Lehmann and Blackwell

Theorem - Jewitt(2007)

Restricted to experiments that satisfy the **MLRP** condition, **Lehmann's ranking** is equivalent to the **Blackwell ranking on dichotomies**

- Blackwell ranking on dichotomies is easy to verify.
- **Remark:** This holds only over the MLRP condition.