# Direct non-regret procedures with random signals

Vianney Perchet

Equipe Combinatoire et Optimisation
Université Pierre et Marie Curie

25 novembre 2008
Dynamic games, Differential games III

# Outline

# Outline

1. External Consistency
   - Definitions
   - An undirect Proof
   - A direct Proof
2. Internal Consistency
   - Definitions
   - Direct Proof
3. Random Signals
   - Model
   - Proof

# 2 players game

2-players repeated game with finite actions sets $I$ and $J$ and a payoff function $\rho : I \times J \to \mathbb{R}$ (extended multilineary).

At round $n$, player 1 chooses $i_n \in I$ and player 2 $j_n \in J$. Player 1 gets $\rho(i_n, j_n)$ as payoff.

Players observe the past actions played by their opponent.

**Definitions**

- $\overline{p}_n = \sum_{m=1}^{n} \rho(i_m, j_m) / n$ the average payoff until stage $n$.
- $\overline{y}_n = \sum_{m=1}^{n} j_m / n$ the empirical mixed action of player 2.

# 2 players game

2-players repeated game with finite actions sets $I$ and $J$ and a payoff function $\rho : I \times J \to \mathbb{R}$ (extended multilineary).

At round $n$, player 1 chooses $i_n \in I$ and player 2 $j_n \in J$. Player 1 gets $\rho(i_n, j_n)$ as payoff.

Players observe the past actions played by their opponent.

**Definitions**

- $\overline{p}_n = \sum_{m=1}^{n} \rho(i_m, j_m) / n$ the average payoff until stage $n$.
- $\overline{y}_n = \sum_{m=1}^{n} j_m / n$ the empirical mixed action of player 2.

# 2 players game

2-players repeated game with finite actions sets $I$ and $J$ and a payoff function $\rho : I \times J \to \mathbb{R}$ (extended multilineary).

At round $n$, player 1 chooses $i_n \in I$ and player 2 $j_n \in J$. Player 1 gets $\rho(i_n, j_n)$ as payoff.

Players observe the past actions played by their opponent.

## Definitions

- $\overline{p}_n = \sum_{m=1}^{n} \rho(i_m, j_m) \big/ n$ the average payoff until stage $n$.
- $\overline{y}_n = \sum_{m=1}^{n} j_m \big/ n$ the empirical mixed action of player 2.

# 2 players game

2-players repeated game with finite actions sets $I$ and $J$ and a payoff function $\rho : I \times J \to \mathbb{R}$ (extended multilineary).

At round $n$, player 1 chooses $i_n \in I$ and player 2 $j_n \in J$. Player 1 gets $\rho(i_n, j_n)$ as payoff.

Players observe the past actions played by their opponent.

**Definitions**

- $\overline{p}_n = \sum_{m=1}^{n} \rho(i_m, j_m) \big/ n$ the average payoff until stage $n$.
- $\overline{y}_n = \sum_{m=1}^{n} j_m \big/ n$ the empirical mixed action of player 2.

# External Consistency

## External Regret (of Player 1)

$$R_n^e = \max_{i \in I} \rho(i, \bar{y}_n) - \bar{p}_n$$

## External Consistency

A strategy $\sigma$ of the player 1 is externally consistent if for every strategy $\tau$ of the second player:

$$\limsup_{n \to \infty} R_n^e \leq 0, (\sigma, \tau)\text{-ps}$$

## Hannan-Blackwell

There exist strategies that are externally consistent.

# External Consistency

## External Regret (of Player 1)

$$R_n^e = \max_{i \in I} \rho(i, \bar{y}_n) - \bar{p}_n$$

## External Consistency

A strategy $\sigma$ of the player 1 is externally consistent if for every strategy $\tau$ of the second player:

$$\limsup_{n \to \infty} R_n^e \leq 0, (\sigma, \tau)\text{-ps}$$

## Hannan-Blackwell

There exist strategies that are externally consistent.

# External Consistency

## External Regret (of Player 1)

$$R_n^e = \max_{i \in I} \rho(i, \overline{y}_n) - \overline{p}_n$$

## External Consistency

A strategy $\sigma$ of the player 1 is externally consistent if for every strategy $\tau$ of the second player:

$$\limsup_{n \to \infty} R_n^e \leq 0, (\sigma, \tau)\text{-ps}$$

## Hannan-Blackwell

There exist strategies that are externally consistent.

# Approachability

2 players game, actions sets $I$ and $J$, payoff $g : I \times J \to \mathbb{R}^k$.
$\overline{g}_n = \sum_{m \leq n} g(i_m, j_m)/n$ is the average payoff at stage $n$.

## Definition

A strategy $\sigma$ of player 1 approaches a set $C$ if for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$, such that for every strategy $\tau$ of player 2 and for every $n \geq N$ :

$$\mathbb{E}_{\sigma,\tau}[d(\overline{g}_n, C)] \leq \varepsilon.$$

## Theorem - Blackwell

A closed convex $C \subset \mathbb{R}^k$ set is either approachable by player 1 or excludable by player 2:

$$\exists y \in \Delta(J), \forall x \in \Delta(I), g(x,y) \notin C$$

# Approachability

2 players game, actions sets $I$ and $J$, payoff $g : I \times J \to \mathbb{R}^k$.
$\overline{g}_n = \sum_{m \leq n} g(i_m, j_m)/n$ is the average payoff at stage $n$.

## Definition

A strategy $\sigma$ of player 1 approaches a set $C$ if for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$, such that for every strategy $\tau$ of player 2 and for every $n \geq N$ :

$$\mathbb{E}_{\sigma,\tau}[d(\overline{g}_n, C)] \leq \varepsilon.$$

## Theorem - Blackwell

A closed convex $C \subset \mathbb{R}^k$ set is either approachable by player 1 or excludable by player 2:

$$\exists y \in \Delta(J), \forall x \in \Delta(I), g(x,y) \notin C$$

# Approachability

2 players game, actions sets $I$ and $J$, payoff $g : I \times J \to \mathbb{R}^k$.
$\overline{g}_n = \sum_{m \leq n} g(i_m, j_m)/n$ is the average payoff at stage $n$.

## Definition

A strategy $\sigma$ of player 1 approaches a set $C$ if for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$, such that for every strategy $\tau$ of player 2 and for every $n \geq N$ :

$$\mathbb{E}_{\sigma,\tau}[d(\overline{g}_n, C)] \leq \varepsilon.$$

## Theorem - Blackwell

A closed convex $C \subset \mathbb{R}^k$ set is either approachable by player 1 or excludable by player 2:

$$\exists y \in \Delta(J), \forall x \in \Delta(I), g(x,y) \notin C$$

# Back to External Regret

$$\limsup_{n\to\infty}\left[\max_{i\in I}\rho(i,\bar{y}_n)-\bar{p}_n\right]\leq 0,(\sigma,\tau)\text{-ps} \quad (1)$$

is equivalent to :

$$\bar{p}_n \to \max_{i\in I}\rho(i,\bar{y}_n)+\mathbb{R}^+,(\sigma,\tau)\text{-ps.} \quad (2)$$

which is implied by the fact that $(\bar{p}_n,\bar{y}_n)$ approaches the closed convex set $C$ :

$$C=\bigcup_{y\in\Delta(J)}\left(\max_{i\in I}\rho(i,y)+\mathbb{R}^+,y\right)\subset\mathbb{R}^{1+J}.$$

If $\sigma$ is such that $(\bar{p}_n,\bar{y}_n)\in\mathbb{R}^{1+J}$ approaches $C$, then $\sigma$ is externally consistent.

# Back to External Regret

$$\limsup_{n\to\infty}\left[\max_{i\in I}\rho(i,\overline{y}_n)-\overline{p}_n\right]\leq 0,(\sigma,\tau)\text{-ps} \quad (1)$$

is equivalent to :

$$\overline{p}_n \to \max_{i\in I}\rho(i,\overline{y}_n)+\mathbb{R}^+,(\sigma,\tau)\text{-ps.} \quad (2)$$

which is implied by the fact that $(\overline{p}_n,\overline{y}_n)$ approaches the closed convex set $C$ :

$$C=\bigcup_{y\in\Delta(J)}\left(\max_{i\in I}\rho(i,y)+\mathbb{R}^+,y\right)\subset\mathbb{R}^{1+J}.$$

If $\sigma$ is such that $(\overline{p}_n,\overline{y}_n)\in\mathbb{R}^{1+J}$ approaches $C$, then $\sigma$ is externally consistent.

# Back to External Regret

$$\limsup_{n\to\infty}\left[\max_{i\in I}\rho(i,\bar{y}_n)-\bar{p}_n\right]\leq 0,(\sigma,\tau)\text{-ps} \tag{1}$$

is equivalent to :

$$\bar{p}_n \to \max_{i\in I}\rho(i,\bar{y}_n)+\mathbb{R}^+,(\sigma,\tau)\text{-ps.} \tag{2}$$

which is implied by the fact that $(\bar{p}_n,\bar{y}_n)$ approaches the closed convex set $C$ :

$$C=\bigcup_{y\in\Delta(J)}\left(\max_{i\in I}\rho(i,y)+\mathbb{R}^+,y\right)\subset\mathbb{R}^{1+J}.$$

If $\sigma$ is such that $(\bar{p}_n,\bar{y}_n)\in\mathbb{R}^{1+J}$ approaches $C$, then $\sigma$ is externally consistent.

# Undirect Proof

Consider the auxiliary 2-player game, with vector-payoff :

$$\gamma(i,j) = (\rho(i,j), \underbrace{0,\ldots,0,1,0,\ldots,0}_{1 \text{ in } j\text{-th coordinates}})$$

In this game the mean payoff is :

$$\overline{\gamma}_n = (\overline{p}_n, \overline{y}_n),$$

and in this game, $C$ is not excludable by player 2:
if he plays $y$, then player 1 plays $i \in \mathrm{BR}(y)$, and $(\rho(i,y),y) \in C$.

$C$ is approachable by player 1, and such a strategy is externally consistent.

# Undirect Proof

Consider the auxiliary 2-player game, with vector-payoff :

$$\gamma(i,j) = (\rho(i,j), \underbrace{0,\ldots,0,1,0,\ldots,0}_{1 \text{ in } j\text{-th coordinates}})$$

In this game the mean payoff is :

$$\overline{\gamma}_n = (\overline{p}_n, \overline{y}_n)\,,$$

and in this game, $C$ is not excludable by player 2:
if he plays $y$, then player 1 plays $i \in \mathrm{BR}(y)$, and $(\rho(i,y),y) \in C$.

$C$ is approachable by player 1, and such a strategy is externally consistent.

# Undirect Proof

Consider the auxiliary 2-player game, with vector-payoff :

$$\gamma(i,j) = (\rho(i,j), \underbrace{0,\ldots,0,1,0,\ldots,0}_{1 \text{ in } j\text{-th coordinates}})$$

In this game the mean payoff is :

$$\overline{\gamma}_n = (\overline{p}_n, \overline{y}_n),$$

and in this game, $C$ is not excludable by player 2:
if he plays $y$, then player 1 plays $i \in \mathrm{BR}(y)$, and $(\rho(i,y),y) \in C$.

$C$ is approachable by player 1, and such a strategy is externally consistent.

# Undirect Proof

Consider the auxiliary 2-player game, with vector-payoff :

$$\gamma(i,j) = (\rho(i,j), \underbrace{0,\ldots,0,1,0,\ldots,0}_{1 \text{ in } j\text{-th coordinates}})$$

In this game the mean payoff is :

$$\overline{\gamma}_n = (\overline{p}_n, \overline{y}_n),$$

and in this game, $C$ is not excludable by player 2:
if he plays $y$, then player 1 plays $i \in \mathrm{BR}(y)$, and $(\rho(i,y),y) \in C$.

$C$ is approachable by player 1, and such a strategy is externally consistent.

# A direct Proof

Consider the auxiliary 2-player game, with vector- payoff :

$$\gamma(i,j) = (\rho(1,j) - \rho(i,j), \ldots, \rho(I,j) - \rho(i,j)) \in \mathbb{R}^I,$$

so that :

$$\overline{\gamma}_n = (\rho(1,\overline{y}_n) - \overline{p}_n, \ldots, \rho(I,\overline{y}_n) - \overline{p}_n).$$

$\overline{\gamma}_n$ approaches the negative orthant $\Rightarrow$ external consistency.

**Definition of the strategy**

At stage $n$ :

- If $\overline{\gamma}_n \notin \mathbb{R}^I_-$, play $x_{n+1}$ proportional to $(\overline{\gamma}_n)^+ = \left(\max\{0, \gamma^i_n\}\right)_{i \in I}$
- Otherwise play anything.

# A direct Proof

Consider the auxiliary 2-player game, with vector- payoff :

$$\gamma(i,j) = (\rho(1,j) - \rho(i,j), \ldots, \rho(I,j) - \rho(i,j)) \in \mathbb{R}^I,$$

so that :

$$\overline{\gamma}_n = (\rho(1,\overline{y}_n) - \overline{p}_n, \ldots, \rho(I,\overline{y}_n) - \overline{p}_n).$$

$\overline{\gamma}_n$ approaches the negative orthant $\Rightarrow$ external consistency.

**Definition of the strategy**

At stage $n$ :

- If $\overline{\gamma}_n \notin \mathbb{R}^I_-$, play $x_{n+1}$ proportional to $(\overline{\gamma}_n)^+ = \left(\max\{0, \gamma^i_n\}\right)_{i \in I}$
- Otherwise play anything.

# A direct Proof

Consider the auxiliary 2-player game, with vector- payoff :

$$\gamma(i,j) = (\rho(1,j) - \rho(i,j), \ldots, \rho(I,j) - \rho(i,j)) \in \mathbb{R}^I,$$

so that :

$$\overline{\gamma}_n = (\rho(1,\overline{y}_n) - \overline{p}_n, \ldots, \rho(I,\overline{y}_n) - \overline{p}_n).$$

$\overline{\gamma}_n$ approaches the negative orthant $\Rightarrow$ external consistency.

## Definition of the strategy

At stage $n$ :

- If $\overline{\gamma}_n \notin \mathbb{R}^I_-$, play $x_{n+1}$ proportional to $(\overline{\gamma}_n)^+ = \left(\max\{0, \gamma_n^i\}\right)_{i \in I}$
- Otherwise play anything.

**Lemma**

The following holds :

$$\langle \mathbb{E}_\sigma \left[ \gamma(i_{n+1}, j_{n+1}) \right] - \Pi_-(\overline{\gamma}_n), \overline{\gamma}_n - \Pi_-(\overline{\gamma}_n) \rangle = 0 \qquad \text{(B)}$$

with $\Pi_-$ the projection on the negative orthant.

Equation (B), which implies the Blackwell condition, ensures that $\overline{\gamma}_n$ converges to $C$.

$\sigma$ is externally consistent.

**Observations**

Note that $\sigma$ actually depends on the value of $\{\rho(i,j_n)\}_{i,n}$, and not on $j_n$.
Same result if player 1 observe $(\rho(1,j_n), \ldots, \rho(I,j_n))$ instead of $j_n$.

## Lemma

The following holds :

$$\langle \mathbb{E}_\sigma \left[\gamma(i_{n+1}, j_{n+1})\right] - \Pi_-(\overline{\gamma}_n), \overline{\gamma}_n - \Pi_-(\overline{\gamma}_n) \rangle = 0 \qquad \text{(B)}$$

with $\Pi_-$ the projection on the negative orthant.

Equation (B), which implies the Blackwell condition, ensures that $\overline{\gamma}_n$ converges to $C$.

**$\sigma$ is externally consistent.**

## Observations

Note that $\sigma$ actually depends on the value of $\{\rho(i, j_n)\}_{i,n}$, and not on $j_n$.

Same result if player 1 observe $(\rho(1, j_n), \ldots, \rho(I, j_n))$ instead of $j_n$.

## Lemma

The following holds :

$$\langle \mathbb{E}_\sigma\left[\gamma(i_{n+1},j_{n+1})\right] - \Pi_-(\overline{\gamma}_n), \overline{\gamma}_n - \Pi_-(\overline{\gamma}_n)\rangle = 0 \qquad \text{(B)}$$

with $\Pi_-$ the projection on the negative orthant.

Equation (B), which implies the Blackwell condition, ensures that $\overline{\gamma}_n$ converges to $C$.

$\sigma$ is externally consistent.

## Observations

Note that $\sigma$ actually depends on the value of $\{\rho(i,j_n)\}_{i,n}$, and not on $j_n$.
Same result if player 1 observe $(\rho(1,j_n),\ldots,\rho(I,j_n))$ instead of $j_n$.

# Outline

# Internal Regret

Player 1 observe $\mathbf{p} \in [-1,1]^I$ an outcome vector chosen by player 2 and gets $\mathbf{p}^i$ if he chooses action $i$.

## Definitions

- $N_n(i) = \{m \in \{1,\dots,n\}, i_m = i\}$ the set of dates of types $i$
- $\overline{\mathbf{p}}_n(i) = \sum_{m \in N_n(i)} \mathbf{p}_n \big/ N_n(i)$ the mean outcome vector on $N_n(i)$.

## Internal Regret

$$R_n(i,k) = \left(\overline{\mathbf{p}}_n^k(i) - \overline{\mathbf{p}}_n^i(i)\right)$$

# Internal Regret

Player 1 observe $\mathbf{p} \in [-1,1]^I$ an outcome vector chosen by player 2 and gets $\mathbf{p}^i$ if he chooses action $i$.

## Definitions

- $N_n(i) = \{m \in \{1,\ldots,n\}, i_m = i\}$ the set of dates of types $i$
- $\overline{\mathbf{p}}_n(i) = \sum_{m \in N_n(i)} \mathbf{p}_n \big/ N_n(i)$ the mean outcome vector on $N_n(i)$.

## Internal Regret

$$R_n(i,k) = \left(\overline{\mathbf{p}}_n^k(i) - \overline{\mathbf{p}}_n^i(i)\right)$$

# Internal Regret

Player 1 observe $\mathbf{p} \in [-1,1]^I$ an outcome vector chosen by player 2 and gets $\mathbf{p}^i$ if he chooses action $i$.

## Definitions

- $N_n(i) = \{m \in \{1,\ldots,n\}, i_m = i\}$ the set of dates of types $i$
- $\overline{\mathbf{p}}_n(i) = \sum_{m \in N_n(i)} \mathbf{p}_n \big/ N_n(i)$ the mean outcome vector on $N_n(i)$.

## Internal Regret

$$R_n(i,k) = \left(\overline{\mathbf{p}}_n^k(i) - \overline{\mathbf{p}}_n^i(i)\right)$$

# Internal Consistency

## Internal Consistency

A strategy $\sigma$ of the player 1 is internally consistent if for every strategy $\tau$ of the second player:

$$\forall i,k \in I, \limsup_{n\to\infty} \frac{N_n(i)}{n} R_n(i,k) \leq 0, (\sigma,\tau)\text{-ps}$$

## Foster Vohra, Hart Mas-Colell, Cover, ...

There exist strategies that are internally consistent.

# Auxiliary Game

Consider the auxiliary 2-player game, with vector payoff :

$$\gamma(i,\mathbf{p}) = \begin{pmatrix} 0 & ,\ldots, & 0 \\ & \vdots & \\ \mathbf{p}^1 - \mathbf{p}^i & ,\ldots, & \mathbf{p}^I - \mathbf{p}^i \\ & \vdots & \\ 0 & ,\ldots, & 0 \end{pmatrix} \in \mathbb{R}^{I \times I}$$

So that :

$$\overline{\gamma}_n = \begin{pmatrix} \frac{N_n(1)}{n}\big(R_n(1,1) & ,\ldots, & R_n(1,I)\big) \\ & \vdots & \\ \frac{N_n(i)}{n}\big(R_n(i,1) & ,\ldots, & R_n(i,I)\big) \\ & \vdots & \\ \frac{N_n(I}{n}\big(R_n(I,1) & ,\ldots, & R_n(I,I)\big) \end{pmatrix} \in \mathbb{R}^{I \times I}$$

# Strategy

$\overline{\gamma}_n$ approaches $\mathbb{R}_-^{I\times I}$ $\Rightarrow$ internally consistency

**Definition of the strategy**

At stage $n$ :

- If $\overline{\gamma}_n \notin \mathbb{R}_-^{I\times I}$, play $x_{n+1}$ proportional to an invariant measure of $(\overline{\gamma}_n)^+$
- Otherwise play anything.

**Lemma**

$$\langle \mathbb{E}_\sigma [\gamma(i_{n+1}, j_{n+1})] - \Pi_-(\overline{\gamma}_n), \overline{\gamma}_n - \Pi_-(\overline{\gamma}_n) \rangle = 0 \quad \text{(B)}$$

$\sigma$ is internally consistent.

# Strategy

$\overline{\gamma}_n$ approaches $\mathbb{R}_-^{I\times I} \Rightarrow$ internally consistency

## Definition of the strategy

At stage $n$ :

- If $\overline{\gamma}_n \notin \mathbb{R}_-^{I\times I}$, play $x_{n+1}$ proportional to an invariant measure of $(\overline{\gamma}_n)^+$
- Otherwise play anything.

## Lemma

$$\langle \mathbb{E}_\sigma [\gamma(i_{n+1}, j_{n+1})] - \Pi_-(\overline{\gamma}_n), \overline{\gamma}_n - \Pi_-(\overline{\gamma}_n) \rangle = 0 \qquad \text{(B)}$$

$\sigma$ is internally consistent.

# Strategy

$\overline{\gamma}_n$ approaches $\mathbb{R}_{-}^{I\times I}$ $\Rightarrow$ internally consistency

## Definition of the strategy

At stage $n$ :

- If $\overline{\gamma}_n \notin \mathbb{R}_{-}^{I\times I}$, play $x_{n+1}$ proportional to an invariant measure of $(\overline{\gamma}_n)^+$
- Otherwise play anything.

## Lemma

$$\langle \mathbb{E}_\sigma \left[\gamma(i_{n+1}, j_{n+1})\right] - \Pi_{-}(\overline{\gamma}_n), \overline{\gamma}_n - \Pi_{-}(\overline{\gamma}_n)\rangle = 0 \qquad \text{(B)}$$

$\sigma$ is internally consistent.

# Strategy

$\overline{\gamma}_n$ approaches $\mathbb{R}_-^{I\times I}$ $\Rightarrow$ internally consistency

## Definition of the strategy

At stage $n$ :

- If $\overline{\gamma}_n \notin \mathbb{R}_-^{I\times I}$, play $x_{n+1}$ proportional to an invariant measure of $(\overline{\gamma}_n)^+$
- Otherwise play anything.

## Lemma

$$\langle \mathbb{E}_\sigma\left[\gamma(i_{n+1},j_{n+1})\right] - \Pi_-(\overline{\gamma}_n), \overline{\gamma}_n - \Pi_-(\overline{\gamma}_n)\rangle = 0 \qquad \text{(B)}$$

$\sigma$ is internally consistent.

# Outline

# Random Signals

When actions $(i,j)$ are played, player 1 receives a signal drawn accordingly to $s(i,j)$ with $s : I \times J \rightarrow \Delta(S)$.

## More Generally

At stage $n$,

- Player 1 chooses $i_n \in I$,
- Player 2 chooses $\mu_n \in \Delta(S)^I$,
- Player 1 receives $s_n$ drawn accordingly to $\mu_n^{i_n} \in \Delta(S)$.

## Evaluation

Player 1 evaluates his payoff through :

$$G : \Delta(I) \times \Delta(S)^I \rightarrow \mathbb{R}$$

Example - Pessimistic evaluation :
$G(x,\mu) = \min\{\rho(x,y),$ st $s(i,y) = \mu^i, \forall i \in I\}$.

# Random Signals

When actions $(i,j)$ are played, player 1 receives a signal drawn accordingly to $s(i,j)$ with $s: I \times J \to \Delta(S)$.

## More Generally

At stage $n$,

- Player 1 chooses $i_n \in I$,
- Player 2 chooses $\mu_n \in \Delta(S)^I$,
- Player 1 receives $s_n$ drawn accordingly to $\mu_n^{i_n} \in \Delta(S)$.

## Evaluation

Player 1 evaluates his payoff through :

$$G : \Delta(I) \times \Delta(S)^I \to \mathbb{R}$$

Example - Pessimistic evaluation :
$G(x,\mu) = \min\{\rho(x,y),$ st $s(i,y) = \mu^i, \forall i \in I\}$.

# Random Signals

When actions $(i,j)$ are played, player 1 receives a signal drawn accordingly to $s(i,j)$ with $s : I \times J \to \Delta(S)$.

## More Generally

At stage $n$,

- Player 1 chooses $i_n \in I$,
- Player 2 chooses $\mu_n \in \Delta(S)^I$,
- Player 1 receives $s_n$ drawn accordingly to $\mu_n^{i_n} \in \Delta(S)$.

## Evaluation

Player 1 evaluates his payoff through :

$$G : \Delta(I) \times \Delta(S)^I \to \mathbb{R}$$

Example - Pessimistic evaluation :
$G(x,\mu) = \min\{\rho(x,y),$ st $s(i,y) = \mu^i, \forall i \in I\}$.

# About the Evaluation Function

Consider the game, no observations are made by player 1.

| | L | R |
|---|---|---|
| T | 1 | 0 |
| B | 0 | 1 |

Player 1 is pessimistic

Its only *good* action is to play repeatedly (1/2,1/2), and receives 1/2.

On the set of stages when he played $T$, his evaluation of payoff is 0 (Player 2 might have played $R$).
On the set of stages when he played $B$, his evaluation of payoff is also 0 (Player 2 might have played $L$).

The evaluation function has to be defined on $\Delta(I)$, the set of mixed actions.

# About the Evaluation Function

Consider the game, no observations are made by player 1.

| | L | R |
|---|---|---|
| T | 1 | 0 |
| B | 0 | 1 |

Player 1 is pessimistic

Its only *good* action is to play repeatedly (1/2,1/2), and receives 1/2.

On the set of stages when he played $T$, his evaluation of payoff is 0 (Player 2 might have played $R$).
On the set of stages when he played $B$, his evaluation of payoff is also 0 (Player 2 might have played $L$).

The evaluation function has to be defined on $\Delta(I)$, the set of mixed actions.

# About the Evaluation Function

Consider the game, no observations are made by player 1.

| | L | R |
|---|---|---|
| T | 1 | 0 |
| B | 0 | 1 |

Player 1 is pessimistic

Its only *good* action is to play repeatedly (1/2,1/2), and receives 1/2.

On the set of stages when he played $T$, his evaluation of payoff is 0 (Player 2 might have played $R$).
On the set of stages when he played $B$, his evaluation of payoff is also 0 (Player 2 might have played $L$).

The evaluation function has to be defined on $\Delta(I)$, the set of mixed actions.

# About the Evaluation Function

Consider the game, no observations are made by player 1.

| | L | R |
|---|---|---|
| T | 1 | 0 |
| B | 0 | 1 |

Player 1 is pessimistic

Its only *good* action is to play repeatedly (1/2,1/2), and receives 1/2.

On the set of stages when he played $T$, his evaluation of payoff is 0 (Player 2 might have played $R$).
On the set of stages when he played $B$, his evaluation of payoff is also 0 (Player 2 might have played $L$).

The evaluation function has to be defined on $\Delta(I)$, the set of mixed actions.

# About the Evaluation Function

Consider the game, no observations are made by player 1.

| | L | R |
|---|---|---|
| T | 1 | 0 |
| B | 0 | 1 |

Player 1 is pessimistic

Its only *good* action is to play repeatedly (1/2,1/2), and receives 1/2.

On the set of stages when he played $T$, his evaluation of payoff is 0 (Player 2 might have played $R$).
On the set of stages when he played $B$, his evaluation of payoff is also 0 (Player 2 might have played $L$).

The evaluation function has to be defined on $\Delta(I)$, the set of mixed actions.

# Strategy

Player 1 will only use a finite number of mixed actions $\{x_l\}_{l\in L}$ : the set $L$ will replace the finite set $I$ in the definition of the regret.

**Timing of the game**

At stage $n$,

- Player 1 chooses (randomly) $l_n \in L$,
- Player 2 chooses $\mu_n \in \Delta(S)^I$,
- A pure action $i_n$ is selected accordingly to $x_{l_n}$,
- Player 1 receives $s_n$ drawn accordingly to $\mu_n^{i_n} \in \Delta(S)$.

With a slight perturbation of Player 1's choices, we can assume that he observes a vector of signal $s_n = (s_n^1, \ldots, s_n^I) \in S^I$ (at each stage he plays with a small probability uniformly, and so he can estimate the vector).

# Strategy

Player 1 will only use a finite number of mixed actions $\{x_l\}_{l \in L}$ : the set $L$ will replace the finite set $I$ in the definition of the regret.

## Timing of the game

At stage $n$,

- Player 1 chooses (randomly) $l_n \in L$,
- Player 2 chooses $\mu_n \in \Delta(S)^I$,
- A pure action $i_n$ is selected accordingly to $x_{l_n}$,
- Player 1 receives $s_n$ drawn accordingly to $\mu_n^{i_n} \in \Delta(S)$.

With a slight perturbation of Player 1's choices, we can assume that he observes a vector of signal $s_n = (s_n^1, \ldots, s_n^I) \in S^I$ (at each stage he plays with a small probability uniformly, and so he can estimate the vector).

# Strategy

Player 1 will only use a finite number of mixed actions $\{x_l\}_{l \in L}$ : the set $L$ will replace the finite set $I$ in the definition of the regret.

## Timing of the game

At stage $n$,

- Player 1 chooses (randomly) $l_n \in L$,
- Player 2 chooses $\mu_n \in \Delta(S)^I$,
- A pure action $i_n$ is selected accordingly to $x_{l_n}$,
- Player 1 receives $s_n$ drawn accordingly to $\mu_n^{i_n} \in \Delta(S)$.

With a slight perturbation of Player 1's choices, we can assume that he observes a vector of signal $s_n = (s_n^1, \ldots, s_n^I) \in S^I$ (at each stage he plays with a small probability uniformly, and so he can estimate the vector).

# Strategy

Player 1 will only use a finite number of mixed actions $\{x_l\}_{l\in L}$ : the set $L$ will replace the finite set $I$ in the definition of the regret.

## Timing of the game

At stage $n$,

- Player 1 chooses (randomly) $l_n \in L$,
- Player 2 chooses $\mu_n \in \Delta(S)^I$,
- A pure action $i_n$ is selected accordingly to $x_{l_n}$,
- Player 1 receives $s_n \in S^I$ drawn accordingly to $\mu_n \in \Delta(S)^I$.

With a slight perturbation of Player 1's choices, we can assume that he observes a vector of signal $s_n = (s_n^1, \ldots, s_n^I) \in S^I$ (at each stage he plays with a small probability uniformly, and so he can estimate the vector).

# Strategy

Player 1 will only use a finite number of mixed actions $\{x_l\}_{l \in L}$ : the set $L$ will replace the finite set $I$ in the definition of the regret.

## Timing of the game

At stage $n$,

- Player 1 chooses (randomly) $l_n \in L$,
- Player 2 chooses $\mu_n \in \Delta(S)^I$,
- A pure action $i_n$ is selected accordingly to $x_{l_n}$,
- Player 1 receives $s_n \in S^I$ drawn accordingly to $\mu_n \in \Delta(S)^I$.

## Definition

- $N_n(l)$ is the set of stages of type $l$,
- $\overline{\mu}_n(l) = \sum_{m \in N_n(l)} s_m / N_n(l)$.

# Consistency

## Mixed-Internally Regret

$$R_n(l) = \max_{x \in \Delta(I)} [G(x, \overline{\mu}_n(l))] - G(x_l, \overline{\mu}_n(l))$$

## $\varepsilon$-mixed internally consistency

A strategy $\sigma$ is $\varepsilon$-mixed internally consistent if, for every strategy $\tau$ of player 2, :

$$\forall l \in L, \limsup_{n \to \infty} \frac{N_n(l)}{n} (R_n(l) - \varepsilon) \leq 0, (\sigma, \tau)\text{-ps}$$

## Theorem

If $\{G(x, \cdot)\}_{x \in \Delta(I)}$ is equicontinuous, then for every $\varepsilon > 0$, there are $\varepsilon$-mixed internally consistent strategies.

# Consistency

## Mixed-Internally Regret

$$R_n(l) = \max_{x \in \Delta(I)} [G(x, \overline{\mu}_n(l))] - G(x_l, \overline{\mu}_n(l))$$

## $\varepsilon$-mixed internally consistency

A strategy $\sigma$ is $\varepsilon$-mixed internally consistent if, for every strategy $\tau$ of player 2, :

$$\forall l \in L, \limsup_{n \to \infty} \frac{N_n(l)}{n} (R_n(l) - \varepsilon) \leq 0, (\sigma, \tau)\text{-ps}$$

## Theorem

If $\{G(x, \cdot)\}_{x \in \Delta(I)}$ is equicontinuous, then for every $\varepsilon > 0$, there are $\varepsilon$-mixed internally consistent strategies.

# Consistency

## Mixed-Internally Regret

$$R_n(l) = \max_{x \in \Delta(I)} [G(x, \overline{\mu}_n(l))] - G(x_l, \overline{\mu}_n(l))$$

## $\varepsilon$-mixed internally consistency

A strategy $\sigma$ is $\varepsilon$-mixed internally consistent if, for every strategy $\tau$ of player 2, :

$$\forall l \in L, \limsup_{n \to \infty} \frac{N_n(l)}{n} (R_n(l) - \varepsilon) \leq 0, (\sigma, \tau)\text{-ps}$$

## Theorem

If $\{G(x, \cdot)\}_{x \in \Delta(I)}$ is equicontinuous, then for every $\varepsilon > 0$, there are $\varepsilon$-mixed internally consistent strategies.

# Consistency

## Mixed-Internally Regret

$$R_n(l) = \max_{x \in \Delta(I)} [G(x, \overline{\mu}_n(l))] - G(x_l, \overline{\mu}_n(l))$$

## $\varepsilon$-mixed internally consistency

A strategy $\sigma$ is $\varepsilon$-mixed internally consistent if, for every strategy $\tau$ of player 2, :

$$\forall l \in L, \limsup_{n \to \infty} \frac{N_n(l)}{n} (R_n(l) - \varepsilon) \leq 0, (\sigma, \tau)\text{-ps}$$

## Theorem

If $\{G(x, \cdot)\}_{x \in \Delta(I)}$ is equicontinuous, then for every $\varepsilon > 0$, there are $\varepsilon$-mixed internally consistent strategies.

# Sketch of Proof :

A strategy will be consistent if :

$$\forall l \in L, x_l \in \mathrm{BR}_\varepsilon(\overline{\mu}_n(l)) \qquad \text{(i)}$$

or equivalently :

$$\forall l \in L, \overline{\mu}_n(l) \in \mathrm{BR}_\varepsilon^{-1}(x_l) \qquad \text{(ii)}$$

Since $G$ is equicontinuous, there exists $\delta > 0$, $\{x_l\}$ and $\{\mu_l\}$ such that :
$\{\mu_l\}_{l\in L}$ is a $\delta$-grid of $\Delta(S)^I$ and $x_l \in \mathrm{BR}_\varepsilon(\mu)$ as soon as $\|\mu - \mu_l\|^2 \leq \delta$.

Then (ii) is implied by the fact that $\|\overline{\mu}_n(l) - \mu_l\|^2 \leq \delta$ or $\overline{\mu}_n(l)$ is closer to $\mu_l$ than to any $\mu_k$.

# Sketch of Proof :

A strategy will be consistent if :

$$\forall l \in L, x_l \in \mathrm{BR}_\varepsilon(\overline{\mu}_n(l)) \quad \text{(i)}$$

or equivalently :

$$\forall l \in L, \overline{\mu}_n(l) \in \mathrm{BR}_\varepsilon^{-1}(x_l) \quad \text{(ii)}$$

Since $G$ is equicontinuous, there exists $\delta > 0$, $\{x_l\}$ and $\{\mu_l\}$ such that :
$\{\mu_l\}_{l\in L}$ is a $\delta$-grid of $\Delta(S)^I$ and $x_l \in \mathrm{BR}_\varepsilon(\mu)$ as soon as $\|\mu - \mu_l\|^2 \leq \delta$.

Then (ii) is implied by the fact that $\|\overline{\mu}_n(l) - \mu_l\|^2 \leq \delta$ or $\overline{\mu}_n(l)$ is closer to $\mu_l$ than to any $\mu_k$.

# Sketch of Proof :

A strategy will be consistent if :

$$\forall l \in L, x_l \in \mathrm{BR}_\varepsilon(\overline{\mu}_n(l)) \qquad \text{(i)}$$

or equivalently :

$$\forall l \in L, \overline{\mu}_n(l) \in \mathrm{BR}_\varepsilon^{-1}(x_l) \qquad \text{(ii)}$$

This property is equivalent to the one in the undirect approach with external regret.

Since $G$ is equicontinuous, there exists $\delta > 0$, $\{x_l\}$ and $\{\mu_l\}$ such that :
$\{\mu_l\}_{l \in L}$ is a $\delta$-grid of $\Delta(S)^I$ and $x_l \in \mathrm{BR}_\varepsilon(\mu)$ as soon as $\|\mu - \mu_l\|^2 \leq \delta$.

Then (ii) is implied by the fact that $\|\overline{\mu}_n(l) - \mu_l\|^2 \leq \delta$ or $\overline{\mu}_n(l)$ is closer to $\mu_l$ than to any $\mu_k$.

# Sketch of Proof :

A strategy will be consistent if :

$$\forall l \in L, x_l \in \mathrm{BR}_\varepsilon(\overline{\mu}_n(l)) \qquad \text{(i)}$$

or equivalently :

$$\forall l \in L, \overline{\mu}_n(l) \in \mathrm{BR}_\varepsilon^{-1}(x_l) \qquad \text{(ii)}$$

Since $G$ is equicontinuous, there exists $\delta > 0$, $\{x_l\}$ and $\{\mu_l\}$ such that :
$\{\mu_l\}_{l\in L}$ is a $\delta$-grid of $\Delta(S)^I$ and $x_l \in \mathrm{BR}_\varepsilon(\mu)$ as soon as $\|\mu - \mu_l\|^2 \leq \delta$.

Then (ii) is implied by the fact that $\|\overline{\mu}_n(l) - \mu_l\|^2 \leq \delta$ or $\overline{\mu}_n(l)$ is closer to $\mu_l$ than to any $\mu_k$.

# Sketch of Proof :

A strategy will be consistent if :

$$\forall l \in L, x_l \in \mathrm{BR}_\varepsilon(\overline{\mu}_n(l)) \tag{i}$$

or equivalently :

$$\forall l \in L, \overline{\mu}_n(l) \in \mathrm{BR}_\varepsilon^{-1}(x_l) \tag{ii}$$

Since $G$ is equicontinuous, there exists $\delta > 0$, $\{x_l\}$ and $\{\mu_l\}$ such that :
$\{\mu_l\}_{l\in L}$ is a $\delta$-grid of $\Delta(S)^I$ and $x_l \in \mathrm{BR}_\varepsilon(\mu)$ as soon as $\|\mu - \mu_l\|^2 \leq \delta$.

Then (ii) is implied by the fact that $\|\overline{\mu}_n(l) - \mu_l\|^2 \leq \delta$ or $\overline{\mu}_n(l)$ is closer to $\mu_l$ than to any $\mu_k$.

# Sketch of Proof 2

$$\forall l, k \in L, \|\overline{\mu}_n(l) - \mu_l\|^2 \leq \|\overline{\mu}_n(l) - \mu_k\|^2 \quad \text{(iii)}$$

This is the definition of calibration : at each stage, player 1 predicts $s \in S^I$ (with $\mu_l$). The prediction are calibrated, if on $N_n(l)$, the average empirical distribution of signal is closer to $\mu_l$ than to any other $\mu_k$.
(iii) is equivalent (by linearity of the scalar product) to

$$\forall l, k \in L, \sum_{m \in N_n(l)} \frac{\|s_m - \mu_l\|^2 - \|s_m - \mu_k\|^2}{N_n(l)} \leq 0. \quad \text{(iv)}$$

(iv) is exactly the definition of internal consistency in a auxiliary game with actions sets $L$ and $S^I$, and the payoff $-\|s_m - \mu_l\|^2$.

Any strategy internally consistent in this auxiliary game will be $\varepsilon$-mixed internally consistent.

# Sketch of Proof 2

$$\forall l,k \in L, \|\overline{\mu}_n(l) - \mu_l\|^2 \leq \|\overline{\mu}_n(l) - \mu_k\|^2 \quad \text{(iii)}$$

This is the definition of calibration : at each stage, player 1 predicts $s \in S^I$ (with $\mu_l$). The prediction are calibrated, if on $N_n(l)$, the average empirical distribution of signal is closer to $\mu_l$ than to any other $\mu_k$.

(iii) is equivalent (by linearity of the scalar product) to

$$\forall l,k \in L, \sum_{m \in N_n(l)} \frac{\|s_m - \mu_l\|^2 - \|s_m - \mu_k\|^2}{N_n(l)} \leq 0. \quad \text{(iv)}$$

(iv) is exactly the definition of internal consistency in a auxiliary game with actions sets $L$ and $S^I$, and the payoff $-\|s_m - \mu_l\|^2$.

Any strategy internally consistent in this auxiliary game will be $\varepsilon$-mixed internally consistent.

# Sketch of Proof 2

$$\forall l,k \in L, \|\overline{\mu}_n(l) - \mu_l\|^2 \leq \|\overline{\mu}_n(l) - \mu_k\|^2 \qquad \text{(iii)}$$

This is the definition of calibration : at each stage, player 1 predicts $s \in S^I$ (with $\mu_l$). The prediction are calibrated, if on $N_n(l)$, the average empirical distribution of signal is closer to $\mu_l$ than to any other $\mu_k$.
(iii) is equivalent (by linearity of the scalar product) to

$$\forall l,k \in L, \sum_{m \in N_n(l)} \frac{\|s_m - \mu_l\|^2 - \|s_m - \mu_k\|^2}{N_n(l)} \leq 0. \qquad \text{(iv)}$$

(iv) is exactly the definition of internal consistency in a auxiliary game with actions sets $L$ and $S^I$, and the payoff $-\|s_m - \mu_l\|^2$.

Any strategy internally consistent in this auxiliary game will be $\varepsilon$-mixed internally consistent.

# Sketch of Proof 2

$$\forall l,k \in L, \|\overline{\mu}_n(l) - \mu_l\|^2 \leq \|\overline{\mu}_n(l) - \mu_k\|^2 \qquad \text{(iii)}$$

This is the definition of calibration : at each stage, player 1 predicts $s \in S^I$ (with $\mu_l$). The prediction are calibrated, if on $N_n(l)$, the average empirical distribution of signal is closer to $\mu_l$ than to any other $\mu_k$.
(iii) is equivalent (by linearity of the scalar product) to

$$\forall l,k \in L, \sum_{m \in N_n(l)} \frac{\|s_m - \mu_l\|^2 - \|s_m - \mu_k\|^2}{N_n(l)} \leq 0. \qquad \text{(iv)}$$

(iv) is exactly the definition of internal consistency in a auxiliary game with actions sets $L$ and $S^I$, and the payoff $-\|s_m - \mu_l\|^2$.

Any strategy internally consistent in this auxiliary game will be $\varepsilon$-mixed internally consistent.

# Sketch of Proof 2

$$\forall l,k \in L, \|\overline{\mu}_n(l) - \mu_l\|^2 \leq \|\overline{\mu}_n(l) - \mu_k\|^2 \qquad \text{(iii)}$$

This is the definition of calibration : at each stage, player 1 predicts $s \in S^I$ (with $\mu_l$). The prediction are calibrated, if on $N_n(l)$, the average empirical distribution of signal is closer to $\mu_l$ than to any other $\mu_k$.
(iii) is equivalent (by linearity of the scalar product) to

$$\forall l,k \in L, \sum_{m \in N_n(l)} \frac{\|s_m - \mu_l\|^2 - \|s_m - \mu_k\|^2}{N_n(l)} \leq 0. \qquad \text{(iv)}$$

(iv) is exactly the definition of internal consistency in a auxiliary game with actions sets $L$ and $S^I$, and the payoff $-\|s_m - \mu_l\|^2$.

Any strategy internally consistent in this auxiliary game will be $\varepsilon$-mixed internally consistent.

# Related Results - Conclusion

External Regret with signals (direct proof) : Rustichini ’99, Lugosi-Mannor-Stoltz ’08

Internal Regret (undirect proof) : Lehrer-Solan ’08

## Conclusion

- Proof in the space of signals
- Gives a direct procedure that leads to internal consistency with imperfect monitoring
- Generalizes the precedent results