Multi-agent learning

Hypothesis Testing

Gerard Vreeswijk, Intelligent Systems Group, Computer Science Department, Faculty of Sciences, Utrecht University, The Netherlands.

Wednesday 12th June, 2019

Plan for today

Plan for today

Basic concepts

Plan for today

Basic concepts

1. Hypothesis testing

Plan for today

Basic concepts

1. Hypothesis testing
2. A simple hypothesis testing algorithm that leads to ϵ-Nash $1 - \epsilon$ of the time.

Plan for today

Basic concepts

1. Hypothesis testing
2. A simple hypothesis testing algorithm that leads to ϵ-Nash $1 - \epsilon$ of the time.
3. Generalisation of (2).

Plan for today

Basic concepts

1. Hypothesis testing
2. A simple hypothesis testing algorithm that leads to ϵ-Nash $1 - \epsilon$ of the time.
3. Generalisation of (2).

Results

Plan for today

Basic concepts

1. Hypothesis testing
2. A simple hypothesis testing algorithm that leads to ϵ-Nash $1 - \epsilon$ of the time.
3. Generalisation of (2).

Results

1. Comparison with Bayesian play

Plan for today

Basic concepts

1. Hypothesis testing
2. A simple hypothesis testing algorithm that leads to ϵ-Nash $1 - \epsilon$ of the time.
3. Generalisation of (2).

Results

1. Comparison with Bayesian play
2. Learning to predict with bounded recall.

Plan for today

Basic concepts

1. Hypothesis testing
2. A simple hypothesis testing algorithm that leads to ϵ-Nash $1 - \epsilon$ of the time.
3. Generalisation of (2).

Results

1. Comparison with Bayesian play
2. Learning to predict with bounded recall.
3. Tightening learning parameters (= simulated annealing).

Part I:
Motivation

Hypothesis testing: motivation

Hypothesis testing: motivation

- In approaches previously discussed, agents maintain a vector of

Hypothesis testing: motivation

- In approaches previously discussed, agents maintain a vector of

    1. Past rewards
       (reinforcement).

Hypothesis testing: motivation

- In approaches previously discussed, agents maintain a vector of

  1. Past rewards
     (reinforcement).
  2. Past actions (no-regret).

Hypothesis testing: motivation

- In approaches previously discussed, agents maintain a vector of

  1. Past rewards
     (reinforcement).
  2. Past actions (no-regret).
  3. Past action profiles
     (conditionalised no-regret).

Hypothesis testing: motivation

- In approaches previously discussed, agents maintain a vector of

1. Past rewards (reinforcement).
2. Past actions (no-regret).
3. Past action profiles (conditionalised no-regret).
4. Mixed strategies (fictitious play).

Hypothesis testing: motivation

- In approaches previously discussed, agents maintain a vector of

1. Past rewards (reinforcement).
2. Past actions (no-regret).
3. Past action profiles (conditionalised no-regret).
4. Mixed strategies (fictitious play).
5. Conditional probabilities (forecasting, Bayesian play).

Hypothesis testing: motivation

- In approaches previously discussed, agents maintain a vector of

  1. Past rewards (reinforcement).
  2. Past actions (no-regret).
  3. Past action profiles (conditionalised no-regret).
  4. Mixed strategies (fictitious play).
  5. Conditional probabilities (forecasting, Bayesian play).

  Such a vector typically changes continuously and gradually on the basis of individual incoming stimuli.

Hypothesis testing: motivation

- In approaches previously discussed, agents maintain a vector of

  1. Past rewards (reinforcement).
  2. Past actions (no-regret).
  3. Past action profiles (conditionalised no-regret).
  4. Mixed strategies (fictitious play).
  5. Conditional probabilities (forecasting, Bayesian play).

  Such a vector typically changes continuously and gradually on the basis of individual incoming stimuli.

- Rather than maintaining a "docile" opponent model, why not adhere to a hypothesis of opponent play for as long as possible, until it really is no longer tenable due to strong counter-evidence.

Hypothesis testing: motivation

- In approaches previously discussed, agents maintain a vector of

  1. Past rewards (reinforcement).
  2. Past actions (no-regret).
  3. Past action profiles (conditionalised no-regret).
  4. Mixed strategies (fictitious play).
  5. Conditional probabilities (forecasting, Bayesian play).

  Such a vector typically changes continuously and gradually on the basis of individual incoming stimuli.

- Rather than maintaining a "docile" opponent model, why not adhere to a hypothesis of opponent play for as long as possible, until it really is no longer tenable due to strong counter-evidence.

  A hypothesis $\approx$ an opponent model formed on the basis of a large and (therefore hopefully) representative sample of stimuli.

Tim Salmon's experiment

Tim Salmon's experiment (2004)

		Player B			
		a	b	c	d
Player A	a	$2, 0$	$2, 0$	$0, 2$	$0, 2$
	b	$0, 2$	$2, 0$	$0, 2$	$2, 0$
	c	$2, 0$	$0, 2$	$2, 0$	$0, 2$
	d	$0, 2$	$0, 2$	$2, 0$	$2, 0$

Tim Salmon's experiment (2004)

Player B

		a	b	c	d
Player A	a	$2, 0$	$2, 0$	$0, 2$	$0, 2$
	b	$0, 2$	$2, 0$	$0, 2$	$2, 0$
	c	$2, 0$	$0, 2$	$2, 0$	$0, 2$
	d	$0, 2$	$0, 2$	$2, 0$	$2, 0$

- Mixed equilibria $\Rightarrow$ Pointless to stick to one strategy.

Tim Salmon's experiment (2004)

Player B

		a	b	c	d
Player A	a	$2,0$	$2,0$	$0,2$	$0,2$
	b	$0,2$	$2,0$	$0,2$	$2,0$
	c	$2,0$	$0,2$	$2,0$	$0,2$
	d	$0,2$	$0,2$	$2,0$	$2,0$

- Mixed equilibria $\Rightarrow$ Pointless to stick to one strategy.
- Constant sum $\Rightarrow$ Folk theorem doesn't apply $\Rightarrow$ repetition does not introduce new strategic opportunities.

Tim Salmon's experiment (2004)

		Player B			
		a	b	c	d
	a	2, 0	2, 0	0, 2	0, 2
	b	0, 2	2, 0	0, 2	2, 0
Player A	c	2, 0	0, 2	2, 0	0, 2
	d	0, 2	0, 2	2, 0	2, 0

- Mixed equilibria $\Rightarrow$ Pointless to stick to one strategy.
- Constant sum $\Rightarrow$ Folk theorem doesn't apply $\Rightarrow$ repetition does not introduce new strategic opportunities.
- Experiments $\Rightarrow$ participants form hypotheses about opponent behaviour.

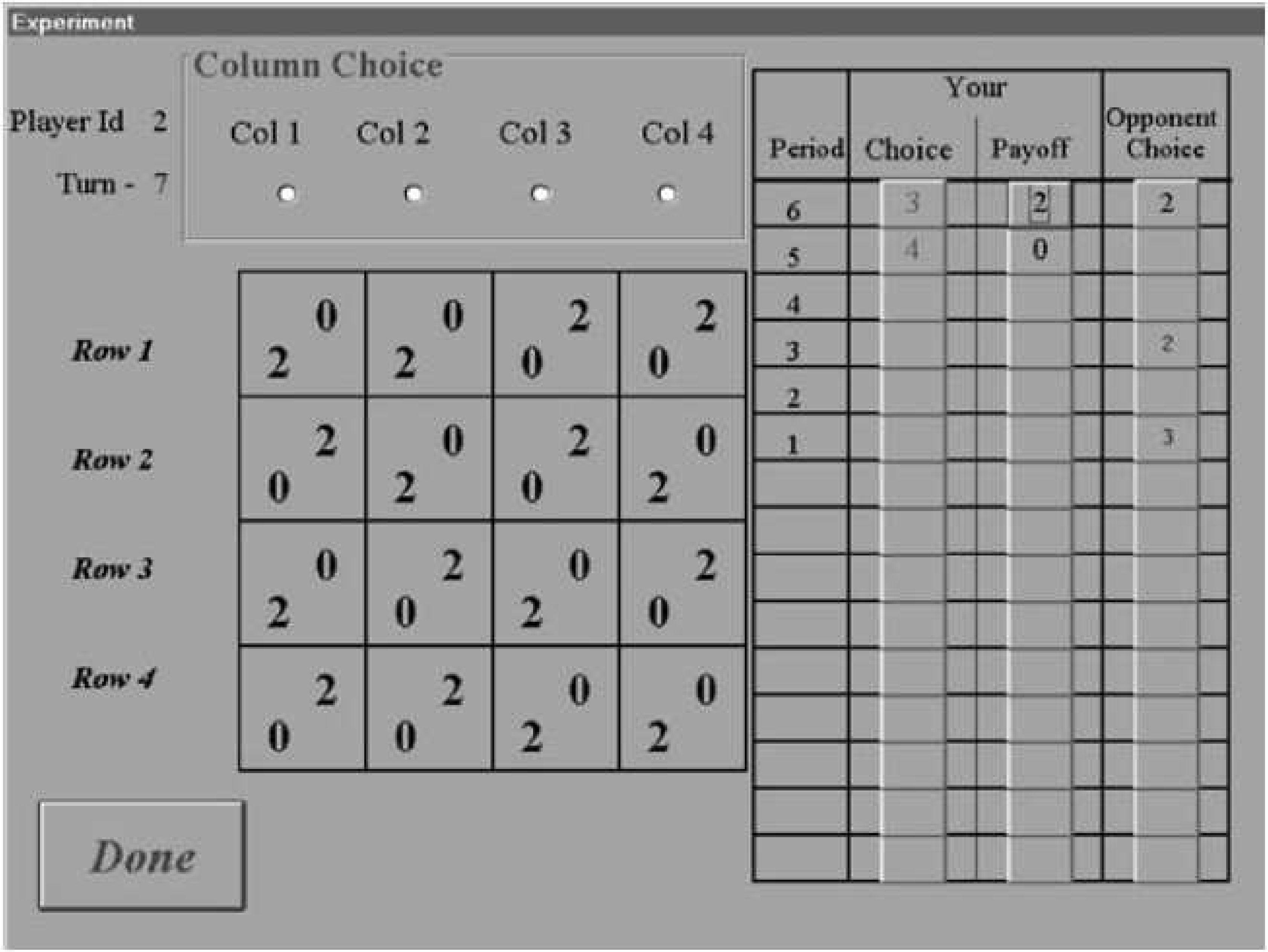
Tim Salmon's experiment (2004)

		Player B			
		a	b	c	d
	a	2, 0	2, 0	0, 2	0, 2
Player A	b	0, 2	2, 0	0, 2	2, 0
	c	2, 0	0, 2	2, 0	0, 2
	d	0, 2	0, 2	2, 0	2, 0

- Mixed equilibria $\Rightarrow$ Pointless to stick to one strategy.
- Constant sum $\Rightarrow$ Folk theorem doesn't apply $\Rightarrow$ repetition does not introduce new strategic opportunities.
- Experiments $\Rightarrow$ participants form hypotheses about opponent behaviour.

T. Salmon (2004). "Evidence for Learning to Learn Behavior in Normal Form Games," in: *Theory and Decision*, Vol. **56**, No. 4, pp. 367-404.

Tim Salmon's experiment (user interface)



Tim Salmon's experiment (transcript of subject 5.7)

Tim Salmon's experiment (transcript of subject 5.7)

Round 10: Previous to the last choice I had chosen Column two three times in a row, hoping that once the other player won on that column it or she would choose a different row. However, when that player kept picking either 1 or 2, I decided to switch to Column three, where the layout was the exact opposite, hoping that my hypothesis that the other player was either staying on or switching between one and two was correct. I think it was.

Tim Salmon's experiment (transcript of subject 5.7)

Round 10: Previous to the last choice I had chosen Column two three times in a row, hoping that once the other player won on that column it or she would choose a different row. However, when that player kept picking either 1 or 2, I decided to switch to Column three, where the layout was the exact opposite, hoping that my hypothesis that the other player was either staying on or switching between one and two was correct. I think it was.

Round 20: I am trying out a theory that the other player may be playing his or her columns that are 0-2-0-2 in alternation. I cannot tell if this is true yet.

Tim Salmon’s experiment (transcript of subject 5.7)

Round 10: Previous to the last choice I had chosen Column two three times in a row, hoping that once the other player won on that column it or she would choose a different row. However, when that player kept picking either 1 or 2, I decided to switch to Column three, where the layout was the exact opposite, hoping that my hypothesis that the other player was either staying on or switching between one and two was correct. I think it was.

Round 20: I am trying out a theory that the other player may be playing his or her columns that are 0-2-0-2 in alternation. I cannot tell if this is true yet.

Round 30: I’m just looking at the opponents history of choices and trying to make an educated guess as to what it or she will pick. I noticed they hadn’t used three in a while, so I picked a column that did not have a payoff for me in three.

Tim Salmon's experiment (transcript of subject 5.7)

Round 10: Previous to the last choice I had chosen Column two three times in a row, hoping that once the other player won on that column it or she would choose a different row. However, when that player kept picking either 1 or 2, I decided to switch to Column three, where the layout was the exact opposite, hoping that my hypothesis that the other player was either staying on or switching between one and two was correct. I think it was.

Round 20: I am trying out a theory that the other player may be playing his or her columns that are 0-2-0-2 in alternation. I cannot tell if this is true yet.

Round 30: I'm just looking at the opponents history of choices and trying to make an educated guess as to what it or she will pick. I noticed they hadn't used three in a while, so I picked a column that did not have a payoff for me in three.

Round 41: By now I had realized that it was only playing 1 and 4 because they're was no column I could pick that would cover both of them, so it was just a matter of guessing how long it would stay on one or the other. Even though it kept losing by choosing row 1 I thought it might keep trying figuring I would eventually not want to test my luck any more.

Part II:
Hypothesis testing

Global scheme

Global scheme

Global scheme

1. Adopt a tentative model of the opponent's behaviour.

Global scheme

1. Adopt a tentative model of the opponent's behaviour.
2. After enough experience, test the model against actual behaviour.

Global scheme

1. Adopt a tentative model of the opponent's behaviour.
2. After enough experience, test the model against actual behaviour. Keep the model if it fits, else go to Step 1.

Global scheme

1. Adopt a tentative model of the opponent's behaviour.
2. After enough experience, test the model against actual behaviour. Keep the model if it fits, else go to Step 1.

1. Restrict models to behavioural strategies of bounded recall. (This rules out Bayesian play.)

Global scheme

1. Adopt a tentative model of the opponent's behaviour.
2. After enough experience, test the model against actual behaviour. Keep the model if it fits, else go to Step 1.

1. Restrict models to behavioural strategies of bounded recall.
   (This rules out Bayesian play.)
   Examples:

Global scheme

1. Adopt a tentative model of the opponent's behaviour.
2. After enough experience, test the model against actual behaviour. Keep the model if it fits, else go to Step 1.

1. Restrict models to behavioural strategies of bounded recall.
   (This rules out Bayesian play.)
   Examples:
   — Empirical frequency of play.

Global scheme

1. Adopt a tentative model of the opponent's behaviour.
2. After enough experience, test the model against actual behaviour.
   Keep the model if it fits, else go to Step 1.

1. Restrict models to behavioural
   strategies of bounded recall.
   (This rules out Bayesian play.)
   Examples:
   — Empirical frequency of play.
   — A predicted mixed strategy
   (FP with bounded recall).

Global scheme

1. Adopt a tentative model of the opponent's behaviour.
2. After enough experience, test the model against actual behaviour.
   Keep the model if it fits, else go to Step 1.

1. Restrict models to behavioural strategies of bounded recall. (This rules out Bayesian play.) Examples:
   — Empirical frequency of play.
   — A predicted mixed strategy (FP with bounded recall).
2. The choice of the new model is partly based on opponent's behaviour, but always has randomness in it. Examples:

Global scheme

1. Adopt a tentative model of the opponent's behaviour.
2. After enough experience, test the model against actual behaviour. Keep the model if it fits, else go to Step 1.

1. Restrict models to behavioural strategies of bounded recall. (This rules out Bayesian play.) Examples:
   — Empirical frequency of play.
   — A predicted mixed strategy (FP with bounded recall).
2. The choice of the new model is partly based on opponent's behaviour, but always has randomness in it. Examples:
   - Almost always, the next model is as accurate as posssible. Else it is chosen randomly.

Global scheme

1. Adopt a tentative model of the opponent's behaviour.
2. After enough experience, test the model against actual behaviour. Keep the model if it fits, else go to Step 1.

1. Restrict models to behavioural strategies of bounded recall. (This rules out Bayesian play.) Examples:
   — Empirical frequency of play.
   — A predicted mixed strategy (FP with bounded recall).
2. The choice of the new model is partly based on opponent's behaviour, but always has randomness in it. Examples:
   - Almost always, the next model is as accurate as posssible. Else it is chosen randomly.
   - The next model is chosen randomly all of the time.

Hypothesis testing

Hypothesis testing

1. Adopt a hypothesis, H.

Hypothesis testing

1. Adopt a hypothesis, H.
2. Collect data.

Hypothesis testing

1. Adopt a hypothesis, H.
2. Collect data.
3. Compare H against the data.

Hypothesis testing

1. Adopt a hypothesis, H.
2. Collect data.
3. Compare H against the data.
4. Decide whether to accept H, given the data.

Hypothesis testing

1. Adopt a hypothesis, H.
2. Collect data.
3. Compare H against the data.
4. Decide whether to accept H, given the data.

	H is accepted	H is rejected
H is true	OK	Type I error
H is false	Type II error	OK

Hypothesis testing

1. Adopt a hypothesis, H.
2. Collect data.
3. Compare H against the data.
4. Decide whether to accept H, given the data.

	H is accepted	H is rejected
H is true	OK	Type I error
H is false	Type II error	OK

1. Type I error: a correct hypothesis is rejected (false positive).

Hypothesis testing

1. Adopt a hypothesis, H.
2. Collect data.
3. Compare H against the data.
4. Decide whether to accept H, given the data.

	H is accepted	H is rejected
H is true	OK	Type I error
H is false	Type II error	OK

1. Type I error: a correct hypothesis is rejected (false positive).
2. Type II error: an incorrect hypothesis is maintained (false negative).

Hypothesis testing

1. Adopt a hypothesis, H.
2. Collect data.
3. Compare H against the data.
4. Decide whether to accept H, given the data.

	H is accepted	H is rejected
H is true	OK	Type I error
H is false	Type II error	OK

1. Type I error: a correct hypothesis is rejected (false positive).
2. Type II error: an incorrect hypothesis is maintained (false negative).

Confidence. (Also: level of significance.) The probability α that a Type I error (false positive) will not be made.

Hypothesis testing

1. Adopt a hypothesis, H.
2. Collect data.
3. Compare H against the data.
4. Decide whether to accept H, given the data.

	H is accepted	H is rejected
H is true	OK	Type I error
H is false	Type II error	OK

1. Type I error: a correct hypothesis is rejected (false positive).
2. Type II error: an incorrect hypothesis is maintained (false negative).

Confidence. (Also: level of significance.) The probability α that a Type I error (false positive) will not be made.

Power. (Dutch: onderscheidend vermogen). The probability β that a Type II error (false negative) will not be made.

Part III:
Naive hypothesis testing

Naive hypothesis testing: concepts and parameters

Naive hypothesis testing: concepts and parameters

Concepts

Naive hypothesis testing: concepts and parameters

Concepts

- Two agents, $i = 1, 2$.

Naive hypothesis testing: concepts and parameters

Concepts

- Two agents, $i = 1, 2$.
- Two action sets, X_1 and X_2.

Naive hypothesis testing: concepts and parameters

Concepts

- Two agents, $i = 1, 2$.
- Two action sets, X_1 and X_2.
- Hypothesis: a mixed strategy of the opponent(s) play

$$p^{-it} \in \Delta_{-i}.$$

Naive hypothesis testing: concepts and parameters

Concepts

- Two agents, $i = 1, 2$.
- Two action sets, X_1 and X_2.
- Hypothesis: a mixed strategy of the opponent(s) play

$$p^{-it} \in \Delta_{-i}.$$

- Data: empirical frequency of play, measured after a sample period: ϕ^{-it}.

Naive hypothesis testing: concepts and parameters

Concepts

- Two agents, $i = 1, 2$.
- Two action sets, X_1 and X_2.
- Hypothesis: a mixed strategy of the opponent(s) play

$$p^{-it} \in \Delta_{-i}.$$

- Data: empirical frequency of play, measured after a sample period: ϕ^{-it}.

Parameters

Naive hypothesis testing: concepts and parameters

Concepts

- Two agents, $i = 1, 2$.
- Two action sets, X_1 and X_2.
- Hypothesis: a mixed strategy of the opponent(s) play

$$p^{-it} \in \Delta_{-i}.$$

- Data: empirical frequency of play, measured after a sample period: ϕ^{-it}.

Parameters

- A smoothing factor γ_i for playing a smoothed reply.

Naive hypothesis testing: concepts and parameters

Concepts

- Two agents, $i = 1, 2$.
- Two action sets, X_1 and X_2.
- Hypothesis: a mixed strategy of the opponent(s) play

$$p^{-it} \in \Delta_{-i}.$$

- Data: empirical frequency of play, measured after a sample period: ϕ^{-it}.

Parameters

- A smoothing factor γ_i for playing a smoothed reply.
- A tolerance level τ_i for rejecting hypotheses.

Naive hypothesis testing: concepts and parameters

Concepts

- Two agents, $i = 1, 2$.
- Two action sets, X_1 and X_2.
- Hypothesis: a mixed strategy of the opponent(s) play

$$p^{-it} \in \Delta_{-i}.$$

- Data: empirical frequency of play, measured after a sample period: ϕ^{-it}.

Parameters

- A smoothing factor γ_i for playing a smoothed reply.
- A tolerance level τ_i for rejecting hypotheses.
- A sample size s_i. Has two uses:

Naive hypothesis testing: concepts and parameters

Concepts

- Two agents, $i = 1, 2$.
- Two action sets, X_1 and X_2.
- Hypothesis: a mixed strategy of the opponent(s) play

$$p^{-it} \in \Delta_{-i}.$$

- Data: empirical frequency of play, measured after a sample period: ϕ^{-it}.

Parameters

- A smoothing factor γ_i for playing a smoothed reply.
- A tolerance level τ_i for rejecting hypotheses.
- A sample size s_i. Has two uses:

1. Sample size, in rounds.

Naive hypothesis testing: concepts and parameters

Concepts

- Two agents, $i = 1, 2$.
- Two action sets, X_1 and X_2.
- Hypothesis: a mixed strategy of the opponent(s) play

$$p^{-it} \in \Delta_{-i}.$$

- Data: empirical frequency of play, measured after a sample period: ϕ^{-it}.

Parameters

- A smoothing factor γ_i for playing a smoothed reply.
- A tolerance level τ_i for rejecting hypotheses.
- A sample size s_i. Has two uses:
  1. Sample size, in rounds.
  2. The probability $1/s_i$ of starting a sample period (when not sampling).

Naive hypothesis testing: concepts and parameters

Concepts

- Two agents, $i = 1, 2$.
- Two action sets, X_1 and X_2.
- Hypothesis: a mixed strategy of the opponent(s) play

$$p^{-it} \in \Delta_{-i}.$$

- Data: empirical frequency of play, measured after a sample period: ϕ^{-it}.

Parameters

- A smoothing factor γ_i for playing a smoothed reply.
- A tolerance level τ_i for rejecting hypotheses.
- A sample size s_i. Has two uses:
  1. Sample size, in rounds.
  2. The probability $1/s_i$ of starting a sample period (when not sampling).
- A randomisation factor ρ_i for determining the opponent's model.

Naive hypothesis testing: algorithm

Naive hypothesis testing: algorithm

1. A hypothesis $p^{-i} \in \Delta_{-i}$ of opponent's play is chosen randomly.

Naive hypothesis testing: algorithm

1. A hypothesis $p^{-i} \in \Delta_{-i}$ of opponent's play is chosen randomly.
2. Repeat:

Naive hypothesis testing: algorithm

1. A hypothesis $p^{-i} \in \Delta_{-i}$ of opponent's play is chosen randomly.
2. Repeat:
   (a) Play a smooth response to the current hypothesis (through, e.g., soft-max).

Naive hypothesis testing: algorithm

1. A hypothesis $p^{-i} \in \Delta_{-i}$ of opponent's play is chosen randomly.
2. Repeat:
   (a) Play a smooth response to the current hypothesis (through, e.g., soft-max).
   (b) When not sampling, a new sample period of length s_i is started with probability $1/s_i$.

Naive hypothesis testing: algorithm

1. A hypothesis $p^{-i} \in \Delta_{-i}$ of opponent's play is chosen randomly.
2. Repeat:
   (a) Play a smooth response to the current hypothesis (through, e.g., soft-max).
   (b) When not sampling, a new sample period of length s_i is started with probability $1/s_i$.
   (c) When sampling is over, the data is formed by the empirical frequency of play ϕ^{-i}, sampled over s_i rounds.

Naive hypothesis testing: algorithm

1. A hypothesis $p^{-i} \in \Delta_{-i}$ of opponent's play is chosen randomly.
2. Repeat:
   (a) Play a smooth response to the current hypothesis (through, e.g., soft-max).

   (b) When not sampling, a new sample period of length s_i is started with probability $1/s_i$.

   (c) When sampling is over, the data is formed by the empirical frequency of play ϕ^{-i}, sampled over s_i rounds.

   The hypothesis p^{-i} is compared with the data ϕ^{-i}.

Naive hypothesis testing: algorithm

1. A hypothesis $p^{-i} \in \Delta_{-i}$ of opponent's play is chosen randomly.
2. Repeat:
   (a) Play a smooth response to the current hypothesis (through, e.g., soft-max).
   (b) When not sampling, a new sample period of length s_i is started with probability $1/s_i$.
   (c) When sampling is over, the data is formed by the empirical frequency of play ϕ^{-i}, sampled over s_i rounds.

   The hypothesis p^{-i} is compared with the data ϕ^{-i}.

   If $\| p^{-i} - \phi^{-i} \| \leq \tau_i$, then the hypothesis is preserved.

Naive hypothesis testing: algorithm

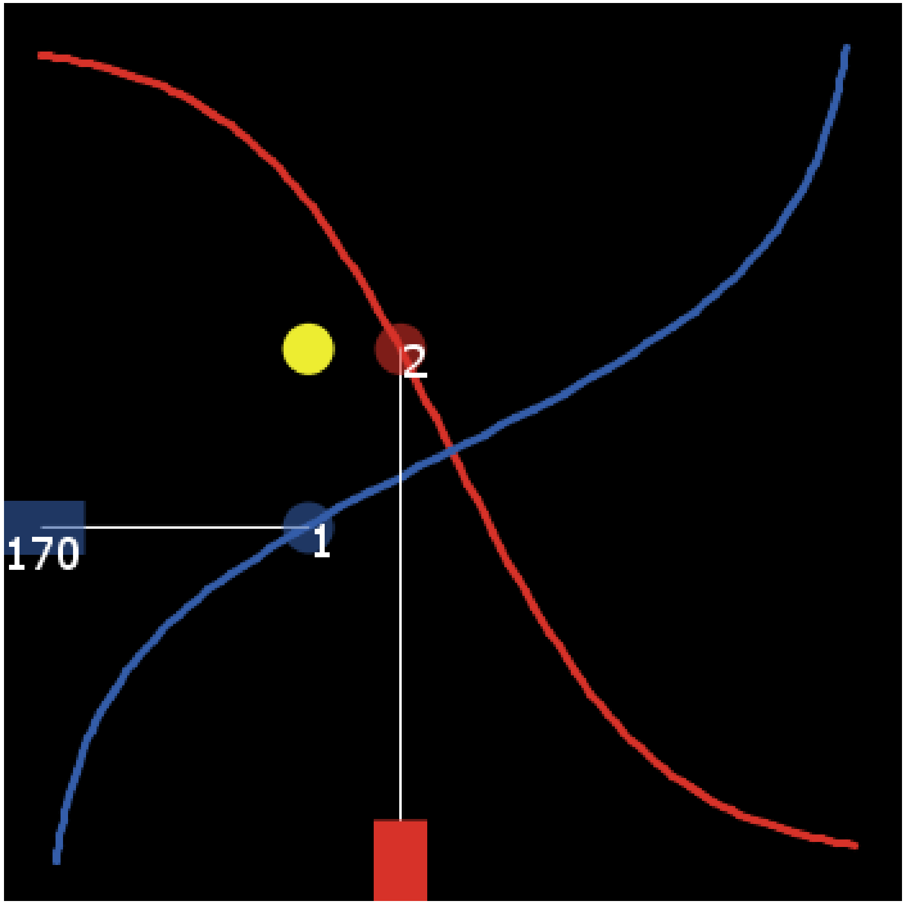
1. A hypothesis $p^{-i} \in \Delta_{-i}$ of opponent's play is chosen randomly.
2. Repeat:
   (a) Play a smooth response to the current hypothesis (through, e.g., soft-max).
   (b) When not sampling, a new sample period of length s_i is started with probability $1/s_i$.
   (c) When sampling is over, the data is formed by the empirical frequency of play ϕ^{-i}, sampled over s_i rounds.

   The hypothesis p^{-i} is compared with the data ϕ^{-i}.

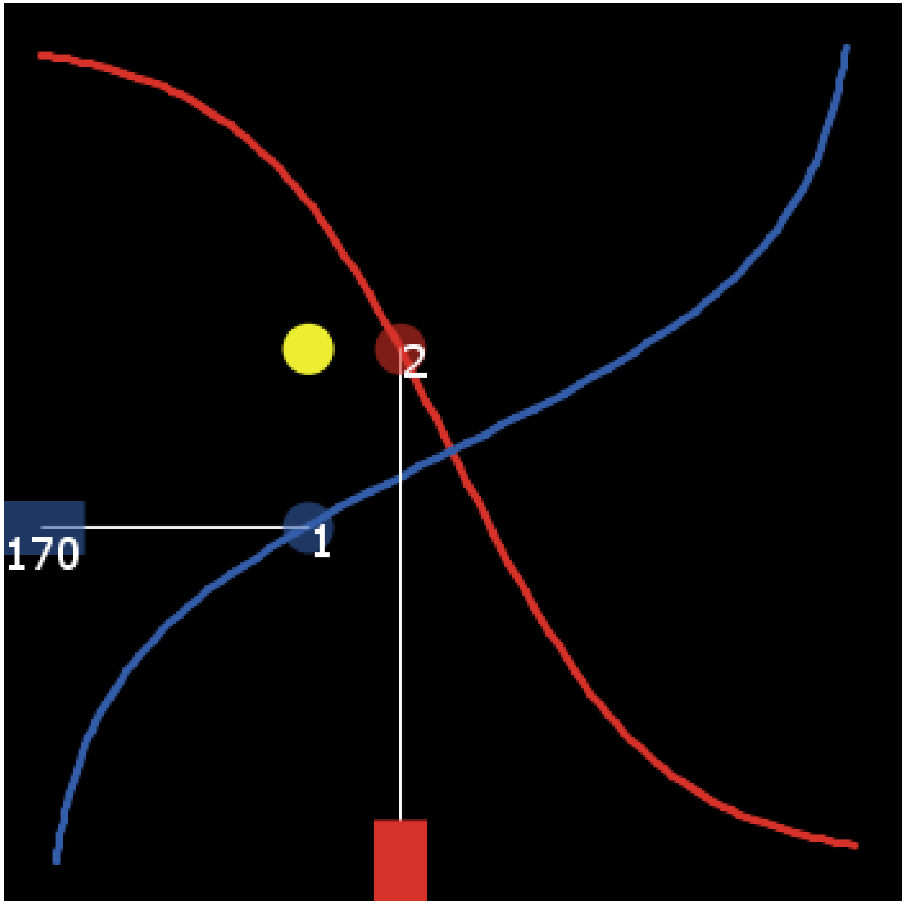
   If $\| p^{-i} - \phi^{-i} \| \leq \tau_i$, then the hypothesis is preserved.

   Else, the hypothesis is rejected, and a new hypothesis is generated. With probability $1 - \rho$ the new hypothesis is taken equal to ϕ^{-it}. Else, it is a random element of Δ_{-i}.

Demo: the matching pennies game

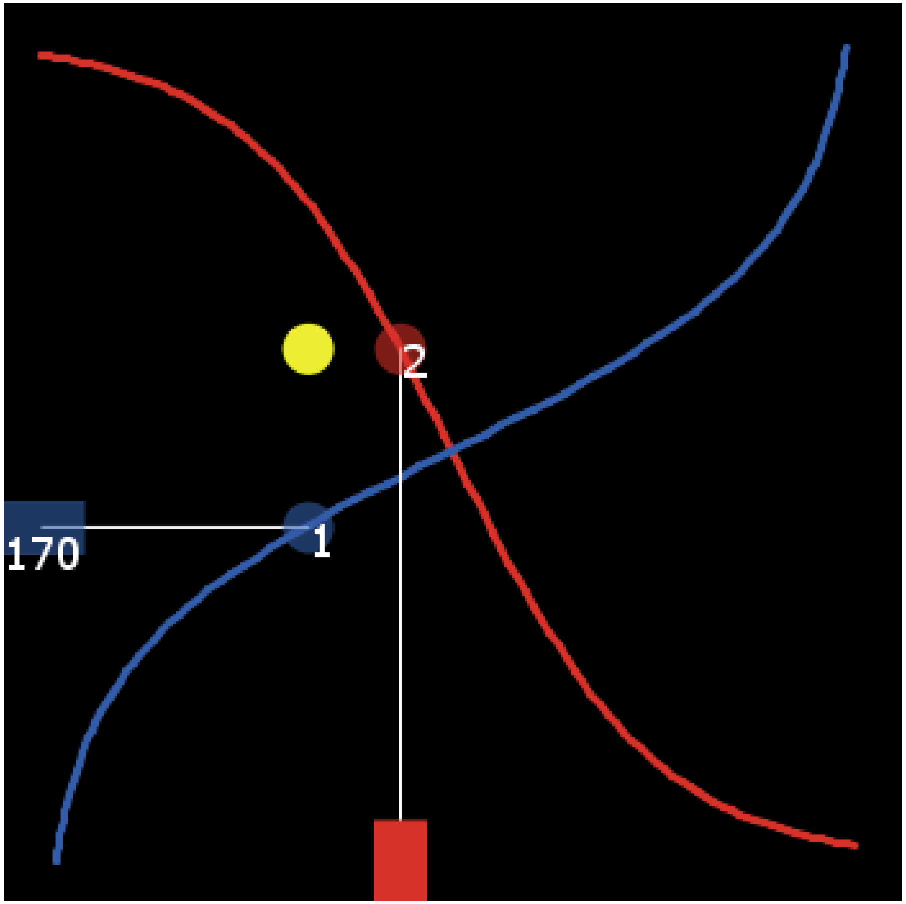


Demo: the matching pennies game



- Player 1: blue, Player 2: red.

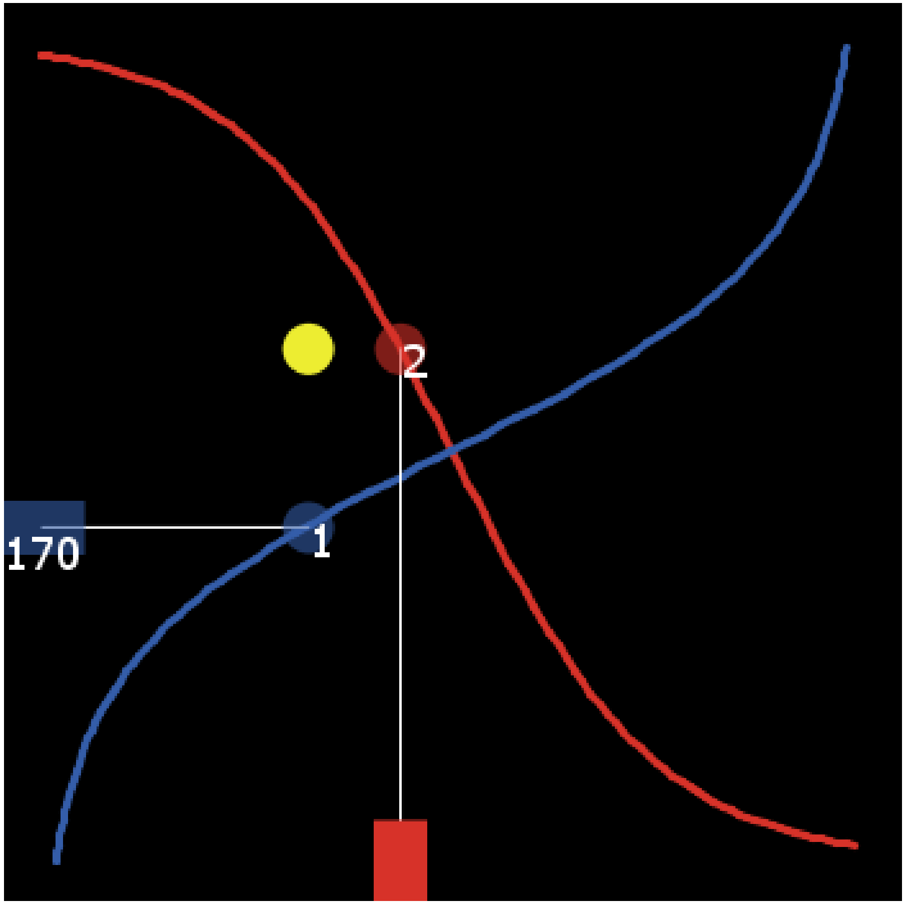
Demo: the matching pennies game



- Player 1: blue, Player 2: red.
- — Horizontal axis **rectangle**: Player 2's model of Player 1's

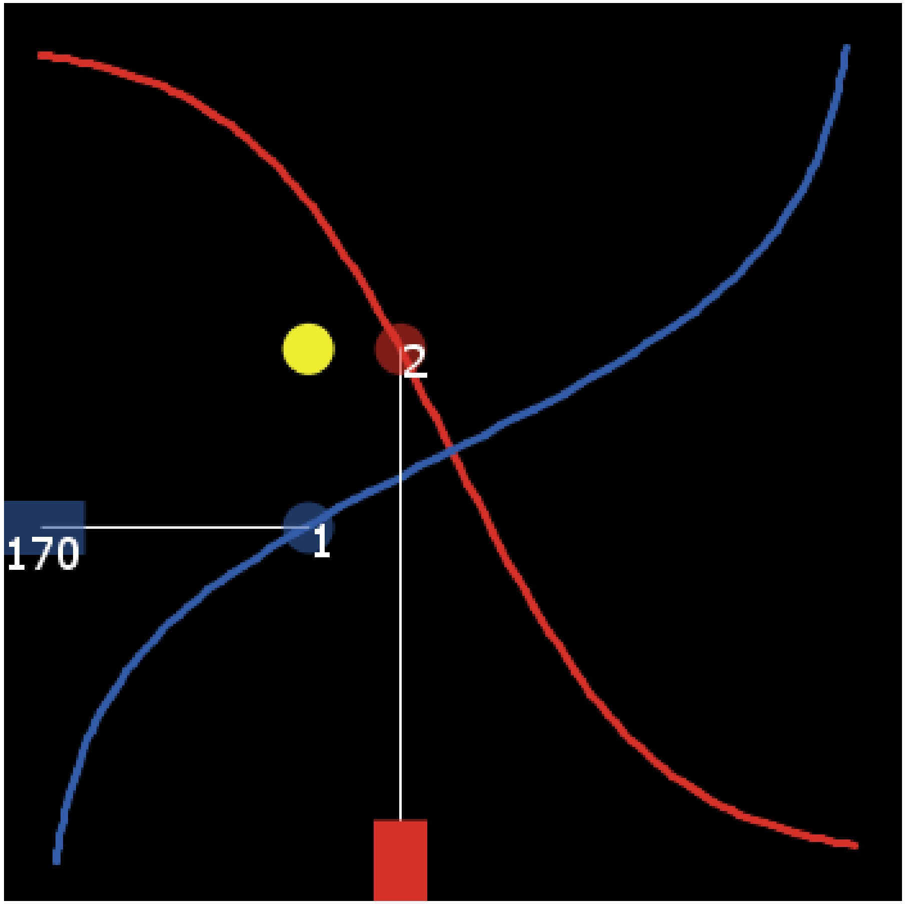
actions = empirical frequency of Player 1.

Demo: the matching pennies game



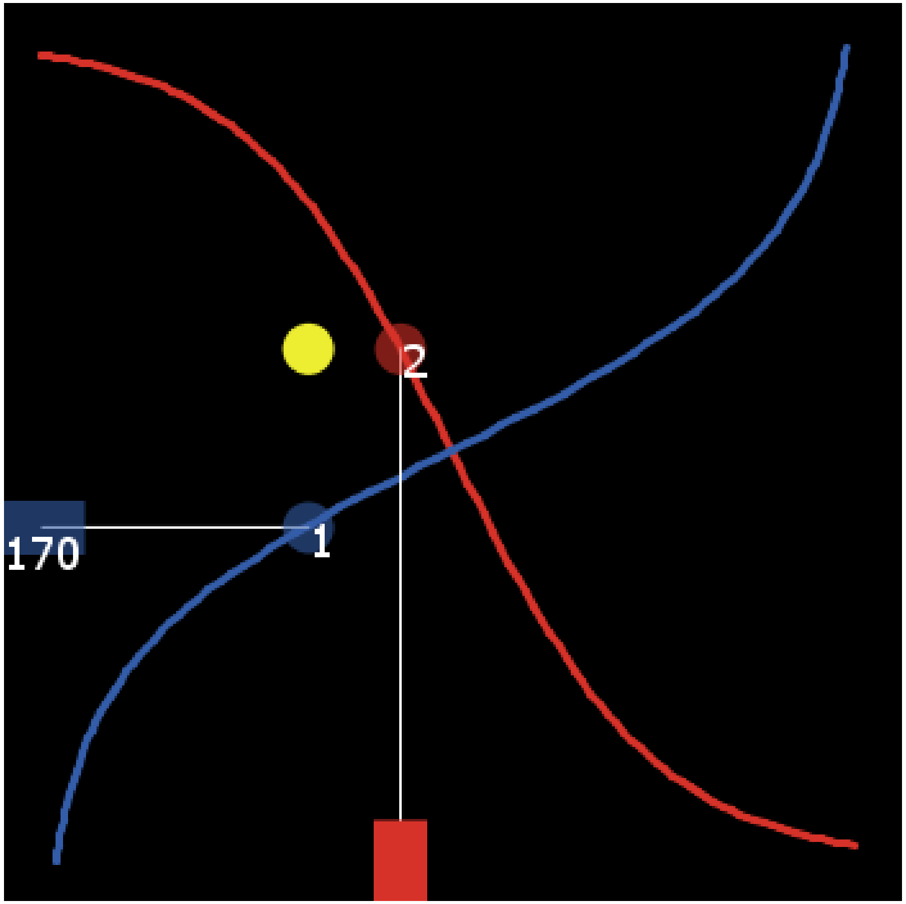
- Player 1: blue, Player 2: red.
- — Horizontal axis **rectangle**: Player 2's model of Player 1's actions = empirical frequency of Player 1. Sleeps (= not collecting data).

Demo: the matching pennies game



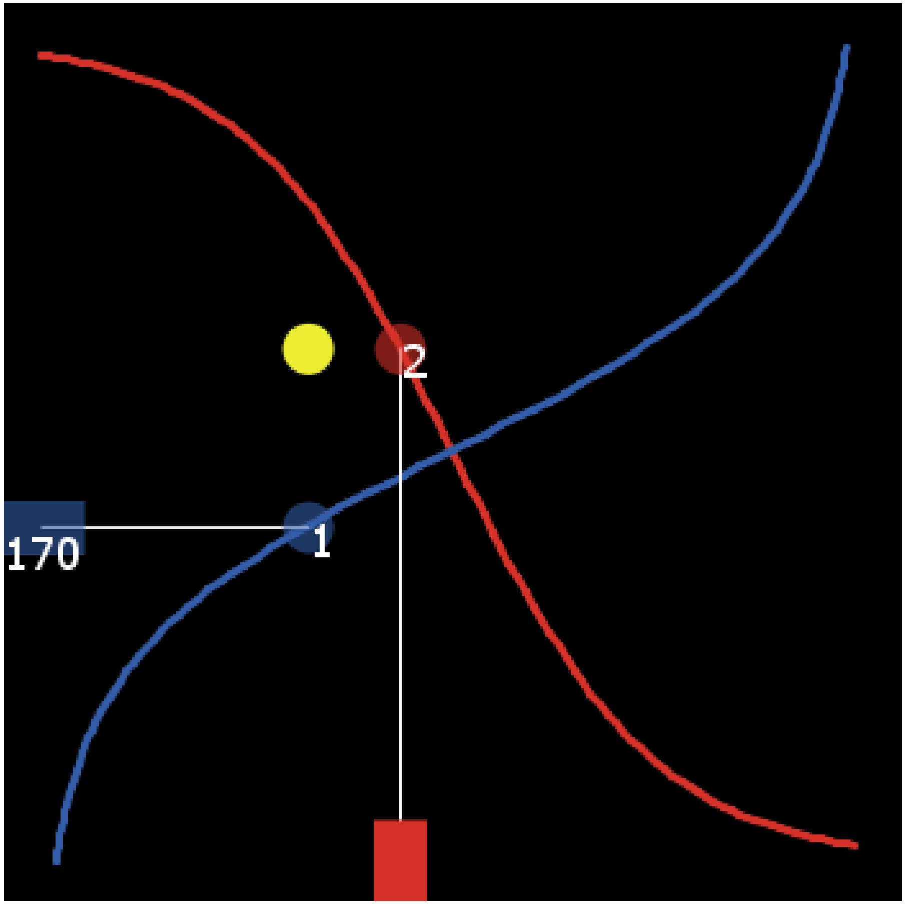
- Player 1: blue, Player 2: red.
- — Horizontal axis **rectangle**: Player 2's model of Player 1's actions = empirical frequency of Player 1. Sleeps (= not collecting data).
— Vertical axis **rectangle**: Player 1's model of Player 2's actions = empirical frequency of Player 2.

Demo: the matching pennies game



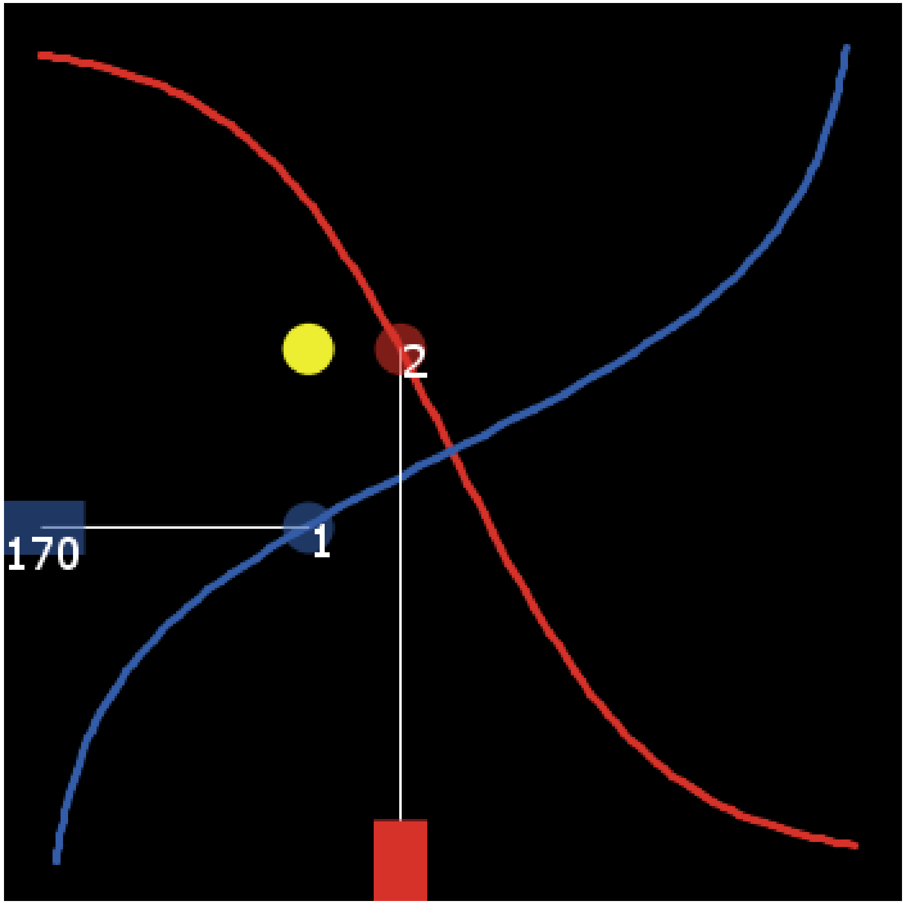
- Player 1: blue, Player 2: red.
- — Horizontal axis **rectangle**: Player 2's model of Player 1's actions = empirical frequency of Player 1. Sleeps (= not collecting data).
  — Vertical axis **rectangle**: Player 1's model of Player 2's actions = empirical frequency of Player 2. In 170th round of collecting data.
- — Red **ball**: Player 2's smoothed reply to what it observes.

Demo: the matching pennies game



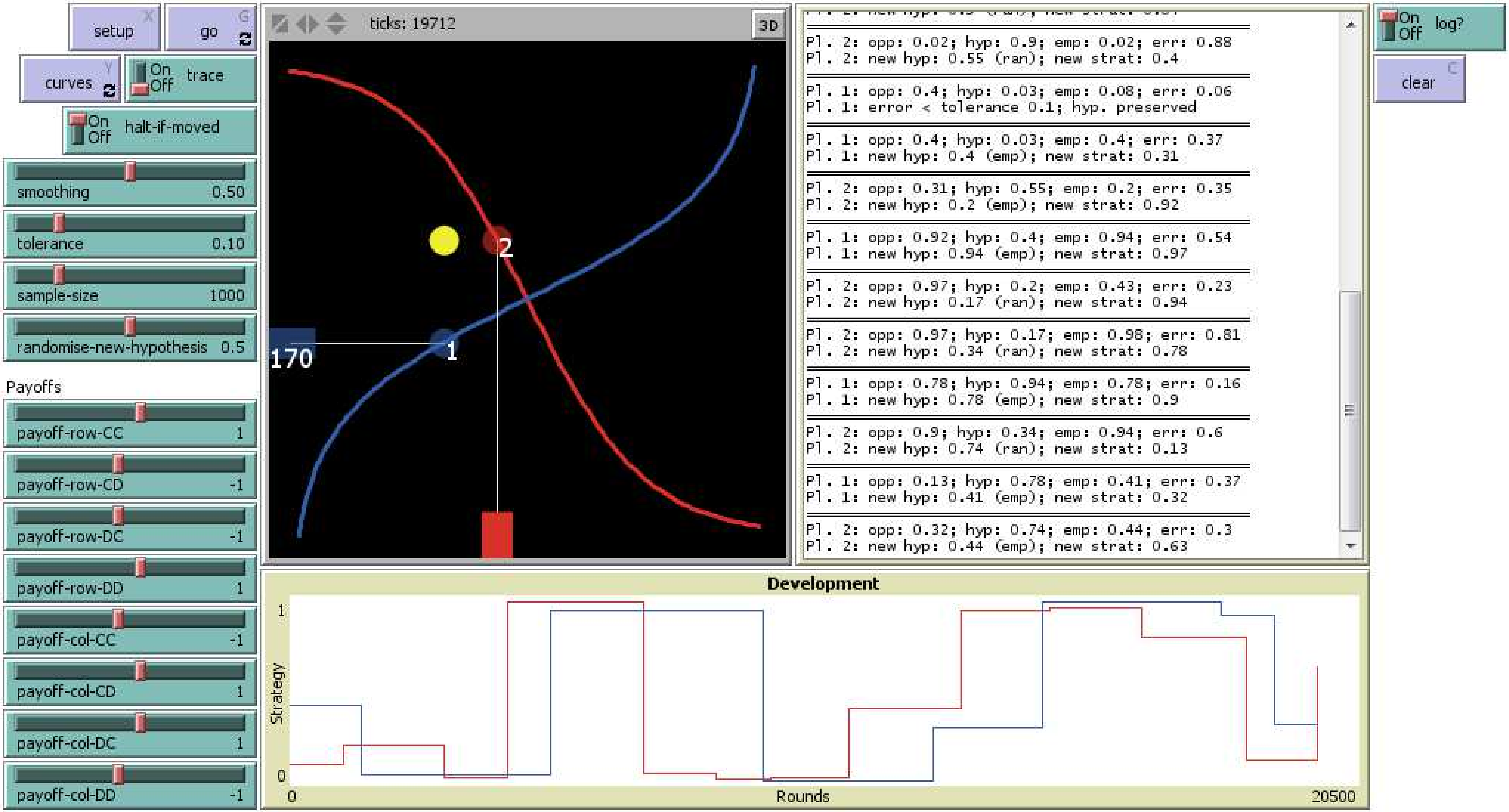
- Player 1: blue, Player 2: red.
- — Horizontal axis **rectangle**: Player 2's model of Player 1's actions = empirical frequency of Player 1. Sleeps (= not collecting data).
  — Vertical axis **rectangle**: Player 1's model of Player 2's actions = empirical frequency of Player 2. In 170th round of collecting data.
- — Red **ball**: Player 2's smoothed reply to what it observes.
  — Blue **ball**: Player 1's smoothed reply to what it observes.

Demo: the matching pennies game

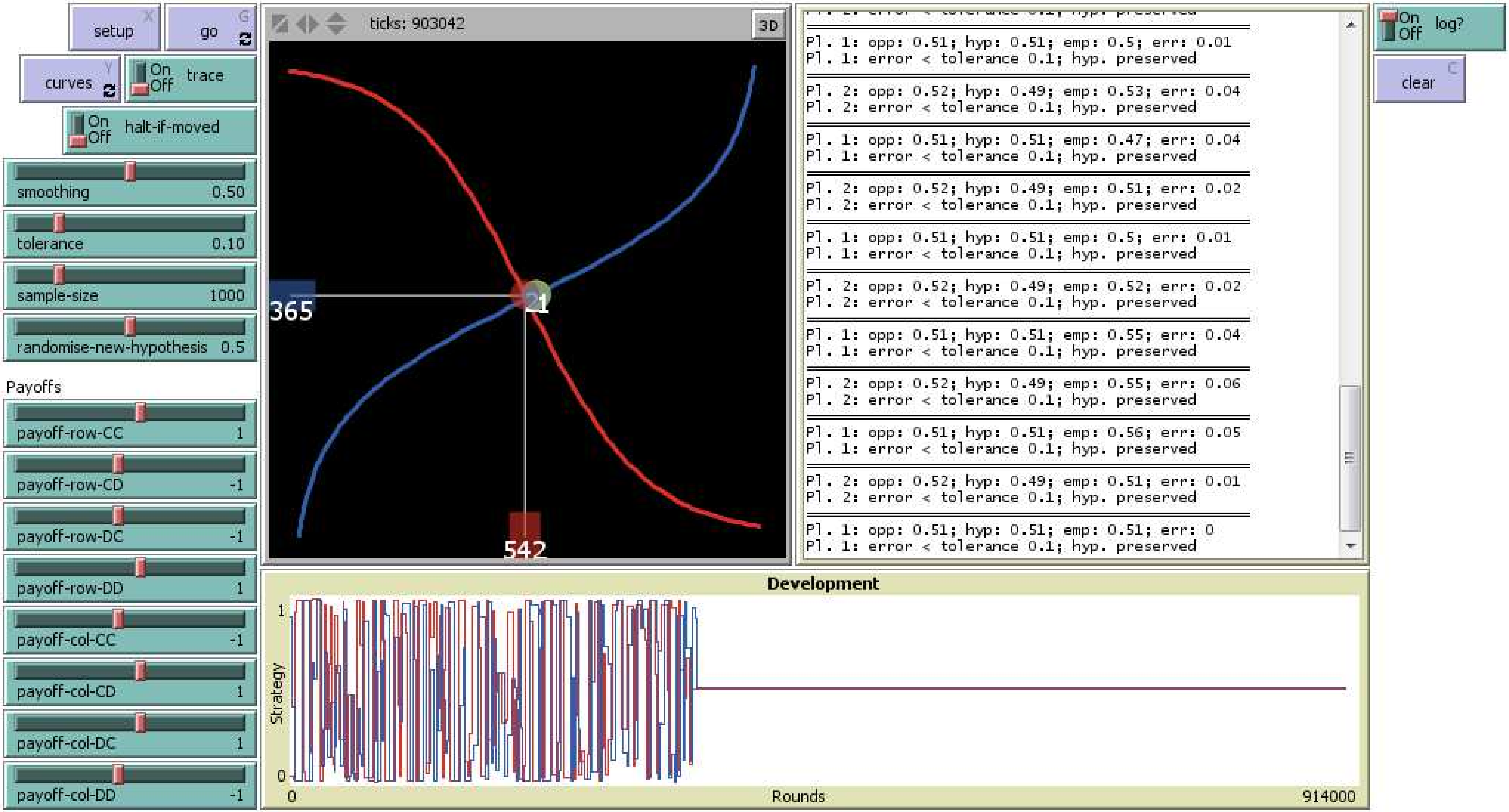


- Player 1: blue, Player 2: red.
- — Horizontal axis **rectangle**: Player 2's model of Player 1's actions = empirical frequency of Player 1. Sleeps (= not collecting data).
  — Vertical axis **rectangle**: Player 1's model of Player 2's actions = empirical frequency of Player 2. In 170th round of collecting data.
- — Red **ball**: Player 2's smoothed reply to what it observes.
  — Blue **ball**: Player 1's smoothed reply to what it observes.
- Yellow ball: strategy profile.

Demo: the matching pennies game



Demo: the matching pennies game (ϵ-equilibrium)



Naive testing: sampling a changing opponent

Naive testing: sampling a changing opponent

- Suppose $s_i = 1000$, $i = 1, 2$.

Naive testing: sampling a changing opponent

- Suppose $s_i = 1000$, $i = 1, 2$.

  1. Opponent's strategy: 70% heads.

Naive testing: sampling a changing opponent

- Suppose $s_i = 1000$, $i = 1, 2$.
  1. Opponent's strategy: 70% heads.
  2. Your model of opponent's behaviour: 73% heads.

Naive testing: sampling a changing opponent

- Suppose $s_i = 1000$, $i = 1, 2$.

  1. Opponent's strategy: 70% heads.
  2. Your model of opponent's behaviour: 73% heads.

- Smoothed fictitious play prescribes

Naive testing: sampling a changing opponent

- Suppose $s_i = 1000$, $i = 1, 2$.
  1. Opponent's strategy: 70% heads.
  2. Your model of opponent's behaviour: 73% heads.
- Smoothed fictitious play prescribes
  1. You respond with heads most of the time (depends on γ_i), say 80%.

Naive testing: sampling a changing opponent

- Suppose $s_i = 1000$, $i = 1, 2$.
    1. Opponent's strategy: 70% heads.
    2. Your model of opponent's behaviour: 73% heads.
- Smoothed fictitious play prescribes
    1. You respond with heads most of the time (depends on γ_i), say 80%.
- At round 745, your opponent starts sampling.

Naive testing: sampling a changing opponent

- Suppose $s_i = 1000$, $i = 1, 2$.
  1. Opponent's strategy: 70% heads.
  2. Your model of opponent's behaviour: 73% heads.
- Smoothed fictitious play prescribes
  1. You respond with heads most of the time (depends on γ_i), say 80%.
- At round 745, your opponent starts sampling.
- At round 811, you start sampling.

Naive testing: sampling a changing opponent

- Suppose $s_i = 1000$, $i = 1, 2$.
  1. Opponent's strategy: 70% heads.
  2. Your model of opponent's behaviour: 73% heads.
- Smoothed fictitious play prescribes
  1. You respond with heads most of the time (depends on γ_i), say 80%.
- At round 745, your opponent starts sampling.
- At round 811, you start sampling.
- At round 1746, your opponent completes sampling.

Naive testing: sampling a changing opponent

- Suppose $s_i = 1000$, $i = 1, 2$.
  1. Opponent's strategy: 70% heads.
  2. Your model of opponent's behaviour: 73% heads.
- Smoothed fictitious play prescribes
  1. You respond with heads most of the time (depends on γ_i), say 80%.
- At round 745, your opponent starts sampling.
- At round 811, you start sampling.
- At round 1746, your opponent completes sampling.
  - His data says you play heads 79% of the time, say.

Naive testing: sampling a changing opponent

- Suppose $s_i = 1000$, $i = 1, 2$.
  1. Opponent's strategy: 70% heads.
  2. Your model of opponent's behaviour: 73% heads.
- Smoothed fictitious play prescribes
  1. You respond with heads most of the time (depends on γ_i), say 80%.
- At round 745, your opponent starts sampling.
- At round 811, you start sampling.
- At round 1746, your opponent completes sampling.
  - His data says you play heads 79% of the time, say.
  - Opponent chooses new model: 15% heads.

Naive testing: sampling a changing opponent

- Suppose $s_i = 1000$, $i = 1, 2$.
  1. Opponent's strategy: 70% heads.
  2. Your model of opponent's behaviour: 73% heads.
- Smoothed fictitious play prescribes
  1. You respond with heads most of the time (depends on γ_i), say 80%.
- At round 745, your opponent starts sampling.
- At round 811, you start sampling.
- At round 1746, your opponent completes sampling.
  - His data says you play heads 79% of the time, say.
  - Opponent chooses new model: 15% heads.
- Meanwhile your are still sampling.

  Notice different opponent behaviour in $[1, 1746]$ and $[1746, 1812]$.

Observations

Observations

Suppose empirical data is based on realised behaviour $H, T, T, H, T \ldots$.

Observations

Suppose empirical data is based on realised behaviour $H, T, T, H, T \ldots$.

1. If a player is unlucky, a Type I error occurs:

Observations

Suppose empirical data is based on realised behaviour $H, T, T, H, T \ldots$.

1. If a player is unlucky, a Type I error occurs:
   - Player maintains a correct model p^{-i} of the opponent.

Observations

Suppose empirical data is based on realised behaviour $H, T, T, H, T \ldots$.

1. If a player is unlucky, a Type I error occurs:
   - Player maintains a correct model p^{-i} of the opponent.
   - Due to bad sampling, the acquired data ϕ^{-i} misrepresents p^{-i}.

Observations

Suppose empirical data is based on realised behaviour $H, T, T, H, T \ldots$.

1. If a player is unlucky, a Type I error occurs:
   - Player maintains a correct model p^{-i} of the opponent.
   - Due to bad sampling, the acquired data ϕ^{-i} misrepresents p^{-i}.

   Consequence: player unjustly drops a correct hypothesis.

Observations

Suppose empirical data is based on realised behaviour $H, T, T, H, T \ldots$.

1. If a player is unlucky, a Type I error occurs:
   - Player maintains a correct model p^{-i} of the opponent.
   - Due to bad sampling, the acquired data ϕ^{-i} misrepresents p^{-i}.

   Consequence: player unjustly drops a correct hypothesis.
2. If a player is unlucky, a Type II error occurs:

Observations

Suppose empirical data is based on realised behaviour $H, T, T, H, T \dots$.

1. If a player is unlucky, a Type I error occurs:
   - Player maintains a correct model p^{-i} of the opponent.
   - Due to bad sampling, the acquired data ϕ^{-i} misrepresents p^{-i}.

   Consequence: player unjustly drops a correct hypothesis.
2. If a player is unlucky, a Type II error occurs:
   - Player maintains an incorrect model p^{-i} of the opponent.

Observations

Suppose empirical data is based on realised behaviour $H, T, T, H, T \dots$.

1. If a player is unlucky, a Type I error occurs:
   - Player maintains a correct model p^{-i} of the opponent.
   - Due to bad sampling, the acquired data ϕ^{-i} misrepresents p^{-i}.

   Consequence: player unjustly drops a correct hypothesis.
2. If a player is unlucky, a Type II error occurs:
   - Player maintains an incorrect model p^{-i} of the opponent.
   - Due to bad sampling, the acquired data ϕ^{-i} unfortunately coincidentally matches p^{-i}.

Observations

Suppose empirical data is based on realised behaviour $H, T, T, H, T \dots$.

1. If a player is unlucky, a Type I error occurs:
   - Player maintains a correct model p^{-i} of the opponent.
   - Due to bad sampling, the acquired data ϕ^{-i} misrepresents p^{-i}.

   Consequence: player unjustly drops a correct hypothesis.

2. If a player is unlucky, a Type II error occurs:
   - Player maintains an incorrect model p^{-i} of the opponent.
   - Due to bad sampling, the acquired data ϕ^{-i} unfortunately coincidentally matches p^{-i}.

   Consequence: player unjustly sticks to an incorrect hypothesis.

Part IV:
Towards ϵ-Nash equilibria

Part IV:
Towards ϵ-Nash equilibria
$1 - \epsilon$ of the time

Foster and Young's first Theorem

Foster and Young's first Theorem

- A property P is said to hold at least α of the time on a given realisation ω if
$$\liminf_{t\to\infty} \frac{|\{t' \leq t \mid P \text{ holds at time } t' \text{ on } \omega\}|}{t} \geq \alpha.$$

Foster and Young's first Theorem

- A property P is said to hold at least α of the time on a given realisation ω if
$$\liminf_{t\to\infty} \frac{|\{t' \leq t \mid P \text{ holds at time } t' \text{ on } \omega\}|}{t} \geq \alpha.$$

- Players $1, \dots, n$ are said to use comparable amounts of data if no one's sample size is at least the square of another's.

Foster and Young's first Theorem

- A property P is said to hold at least α of the time on a given realisation ω if
$$\liminf_{t\to\infty} \frac{|\{t' \leq t \mid P \text{ holds at time } t' \text{ on } \omega\}|}{t} \geq \alpha.$$

- Players $1, \dots, n$ are said to use comparable amounts of data if no one's sample size is at least the square of another's.

Foster and Young's first Theorem

- A property P is said to hold at least α of the time on a given realisation ω if
$$\liminf_{t\to\infty} \frac{|\{t' \leq t \mid P \text{ holds at time } t' \text{ on } \omega\}|}{t} \geq \alpha.$$

- Players $1, \dots, n$ are said to use comparable amounts of data if no one's sample size is at least the square of another's.

Theorem (Foster and Young, 2003).

Foster and Young's first Theorem

- A property P is said to hold at least α of the time on a given realisation ω if
$$\liminf_{t\to\infty} \frac{|\{t' \leq t \mid P \text{ holds at time } t' \text{ on } \omega\}|}{t} \geq \alpha.$$

- Players $1, \dots, n$ are said to use comparable amounts of data if no one's sample size is at least the square of another's.

Theorem (Foster and Young, 2003). Let G be a finite game where players enjoy normalised payoffs and engage in naive hypothesis testing with soft-max responses.

Foster and Young's first Theorem

- A property P is said to hold at least α of the time on a given realisation ω if
$$\liminf_{t\to\infty} \frac{|\{t' \leq t \mid P \text{ holds at time } t' \text{ on } \omega\}|}{t} \geq \alpha.$$

- Players $1, \dots, n$ are said to use comparable amounts of data if no one's sample size is at least the square of another's.

Theorem (Foster and Young, 2003). Let G be a finite game where players enjoy normalised payoffs and engage in naive hypothesis testing with soft-max responses. For every $\epsilon > 0$

Foster and Young's first Theorem

- A property P is said to hold at least α of the time on a given realisation ω if

$$\liminf_{t\to\infty} \frac{|\{t' \leq t \mid P \text{ holds at time } t' \text{ on } \omega\}|}{t} \geq \alpha.$$

- Players $1, \ldots, n$ are said to use comparable amounts of data if no one's sample size is at least the square of another's.

Theorem (Foster and Young, 2003). Let G be a finite game where players enjoy normalised payoffs and engage in naive hypothesis testing with soft-max responses. For every $\epsilon > 0$, if the responses are sufficiently rational

Foster and Young's first Theorem

- A property P is said to hold at least α of the time on a given realisation ω if
$$\liminf_{t\to\infty} \frac{|\{t' \leq t \mid P \text{ holds at time } t' \text{ on } \omega\}|}{t} \geq \alpha.$$

- Players $1, \ldots, n$ are said to use comparable amounts of data if no one's sample size is at least the square of another's.

Theorem (Foster and Young, 2003). Let G be a finite game where players enjoy normalised payoffs and engage in naive hypothesis testing with soft-max responses. For every $\epsilon > 0$, if the responses are sufficiently rational, the test tolerances are sufficiently fine (given the responses)

Foster and Young's first Theorem

- A property P is said to hold at least α of the time on a given realisation ω if
$$\liminf_{t\to\infty} \frac{|\{t' \leq t \mid P \text{ holds at time } t' \text{ on } \omega\}|}{t} \geq \alpha.$$

- Players $1, \ldots, n$ are said to use comparable amounts of data if no one's sample size is at least the square of another's.

Theorem (Foster and Young, 2003). Let G be a finite game where players enjoy normalised payoffs and engage in naive hypothesis testing with soft-max responses. For every $\epsilon > 0$, if the responses are sufficiently rational, the test tolerances are sufficiently fine (given the responses), and the amounts of data collected are comparable and sufficiently large (given the tolerances and the responses)

Foster and Young's first Theorem

- A property P is said to hold at least α of the time on a given realisation ω if
$$\liminf_{t\to\infty} \frac{|\{t' \leq t \mid P \text{ holds at time } t' \text{ on } \omega\}|}{t} \geq \alpha.$$

- Players $1, \dots, n$ are said to use comparable amounts of data if no one's sample size is at least the square of another's.

Theorem (Foster and Young, 2003). Let G be a finite game where players enjoy normalised payoffs and engage in naive hypothesis testing with soft-max responses. For every $\epsilon > 0$, if the responses are sufficiently rational, the test tolerances are sufficiently fine (given the responses), and the amounts of data collected are comparable and sufficiently large (given the tolerances and the responses), then the responses constitute an ϵ-equilibrium of G at least $1 - \epsilon$ of the time

Foster and Young's first Theorem

- A property P is said to hold at least α of the time on a given realisation ω if

$$\liminf_{t\to\infty} \frac{|\{t' \leq t \mid P \text{ holds at time } t' \text{ on } \omega\}|}{t} \geq \alpha.$$

- Players $1, \ldots, n$ are said to use comparable amounts of data if no one's sample size is at least the square of another's.

Theorem (Foster and Young, 2003). Let G be a finite game where players enjoy normalised payoffs and engage in naive hypothesis testing with soft-max responses. For every $\epsilon > 0$, if the responses are sufficiently rational, the test tolerances are sufficiently fine (given the responses), and the amounts of data collected are comparable and sufficiently large (given the tolerances and the responses), then the responses constitute an ϵ-equilibrium of G at least $1 - \epsilon$ of the time a.s.

Part V: Generalisation

Five assumptions

Five assumptions

1. **Models.** All players' models are of bounded memory m.

Five assumptions

1. **Models.** All players' models are of bounded memory m.
2. **Responses.** Each player i has a family of response functions that are continuous, diffuse, and γ_i-optimal.

Five assumptions

1. **Models.** All players' models are of bounded memory m.
2. **Responses.** Each player i has a family of response functions that are continuous, diffuse, and γ_i-optimal. (Think soft-max.)

Five assumptions

1. **Models.** All players' models are of bounded memory m.
2. **Responses.** Each player i has a family of response functions that are continuous, diffuse, and γ_i-optimal. (Think soft-max.)
3. **Tests.** Each player i uses a powerful family of tests[1] with sample size s_i and test tolerance τ_i.

Five assumptions

1. **Models.** All players' models are of bounded memory m.
2. **Responses.** Each player i has a family of response functions that are continuous, diffuse, and γ_i-optimal. (Think soft-max.)
3. **Tests.** Each player i uses a powerful family of tests[1] with sample size s_i and test tolerance τ_i.
4. **Test frequencies.** If player i is not currently in a test phase, it enters a test phase with probability $1/s_i$, collects data for s_i periods, then tests.

Five assumptions

1. **Models.** All players' models are of bounded memory m.
2. **Responses.** Each player i has a family of response functions that are continuous, diffuse, and γ_i-optimal. (Think soft-max.)
3. **Tests.** Each player i uses a powerful family of tests[1] with sample size s_i and test tolerance τ_i.
4. **Test frequencies.** If player i is not currently in a test phase, it enters a test phase with probability $1/s_i$, collects data for s_i periods, then tests. While gathering data his hypothesis and response remain unchanged.

Five assumptions

1. **Models.** All players' models are of bounded memory m.
2. **Responses.** Each player i has a family of response functions that are continuous, diffuse, and γ_i-optimal. (Think soft-max.)
3. **Tests.** Each player i uses a powerful family of tests[1] with sample size s_i and test tolerance τ_i.
4. **Test frequencies.** If player i is not currently in a test phase, it enters a test phase with probability $1/s_i$, collects data for s_i periods, then tests. While gathering data his hypothesis and response remain unchanged.
5. **Model revision.** Every player is flexible in his choice of a new model when a hypothesis test fails.

Five assumptions

1. **Models.** All players' models are of bounded memory m.
2. **Responses.** Each player i has a family of response functions that are continuous, diffuse, and γ_i-optimal. (Think soft-max.)
3. **Tests.** Each player i uses a powerful family of tests[1] with sample size s_i and test tolerance τ_i.
4. **Test frequencies.** If player i is not currently in a test phase, it enters a test phase with probability $1/s_i$, collects data for s_i periods, then tests. While gathering data his hypothesis and response remain unchanged.
5. **Model revision.** Every player is flexible in his choice of a new model when a hypothesis test fails.

[1] Such families of tests would discriminate very sharply between models that are near the truth (nearer than τ) and those that are far from the truth (further than τ).

I) : Models are of "bounded recall"

I) : Models are of "bounded recall"

- A memory-m response is a function $X^m \to \Delta_i$, i.e., an element of

  $$\Delta_i^M$$

  where M is shorthand for X^m.

I) : Models are of “bounded recall”

- A memory-m response is a function $X^m \to \Delta_i$, i.e., an element of
$$\Delta_i^M$$
where M is shorthand for X^m.

- A memory-m opponent model is an element of
$$\Pi_{j \neq i} \Delta_j^M,$$
i.e., an $n-1$ tuple of memory-m responses.

I) : Models are of "bounded recall"

- A memory-m response is a function $X^m \to \Delta_i$, i.e., an element of

$$\Delta_i^M$$

where M is shorthand for X^m.

- A memory-m opponent model is an element of

$$\Pi_{j \neq i} \Delta_j^M,$$

i.e., an $n-1$ tuple of memory-m responses.

Because opponent models are independent, each row adds up to one. I.e., probabailities are on rows.

I) : Models are of "bounded recall"

- A memory-m response is a function $X^m \rightarrow \Delta_i$, i.e., an element of

$$\Delta_i^M$$

where M is shorthand for X^m.

- A memory-m opponent model is an element of

$$\Pi_{j \neq i} \Delta_j^M,$$

i.e., an $n-1$ tuple of memory-m responses.

Because opponent models are independent, each row adds up to one. I.e., probabailities are on rows. (If opponent models were dependent, probabilities would be on matrices.)

I) : Models are of "bounded recall"

- A memory-m response is a function $X^m \to \Delta_i$, i.e., an element of
  $$\Delta_i^M$$
  where M is shorthand for X^m.
- A memory-m opponent model is an element of
  $$\Pi_{j \neq i} \Delta_j^M,$$

i.e., an $n-1$ tuple of memory-m responses.

Because opponent models are independent, each row adds up to one. I.e., probabailities are on rows. (If opponent models were dependent, probabilities would be on matrices.)

I.e.,

$$(x^{t-m+1}, \dots, x^t) \mapsto \left.\underbrace{\begin{pmatrix} p_{1,1} & \dots & p_{k,1} \\ \vdots & \ddots & \vdots \\ p_{1,n} & \dots & p_{k,n} \end{pmatrix}}_{\#\text{actions}}\right\} \#\text{players} - 1$$

II) : Responses are continuous, diffuse, and γ_i-optimal

II) : Responses are continuous, diffuse, and γ_i-optimal

- Recall that soft-max (mixed logit) possesses the following properties.

II) : Responses are continuous, diffuse, and γ_i-optimal

- Recall that soft-max (mixed logit) possesses the following properties.
  - Continuous.

II) : Responses are continuous, diffuse, and γ_i-optimal

- Recall that soft-max (mixed logit) possesses the following properties.
  - Continuous.
  - Gives positive probability to all actions.

II) : Responses are continuous, diffuse, and γ_i-optimal

- Recall that soft-max (mixed logit) possesses the following properties.
  - Continuous.
  - Gives positive probability to all actions.
  - Depending on γ_i, the expected utility of a response can be made arbitrary close to the expected utility of a best response (indeed, let $\gamma_i \downarrow 0$).

II) : Responses are continuous, diffuse, and γ_i-optimal

- Recall that soft-max (mixed logit) possesses the following properties.
  - Continuous.
  - Gives positive probability to all actions.
  - Depending on γ_i, the expected utility of a response can be made arbitrary close to the expected utility of a best response (indeed, let $\gamma_i \downarrow 0$).
- Generalise soft-max to any memory-m response $X^m \to \Delta_i$ that likewise possesses the following properties.

II) : Responses are continuous, diffuse, and γ_i-optimal

- Recall that soft-max (mixed logit) possesses the following properties.
  - Continuous.
  - Gives positive probability to all actions.
  - Depending on γ_i, the expected utility of a response can be made arbitrary close to the expected utility of a best response (indeed, let $\gamma_i \downarrow 0$).
- Generalise soft-max to any memory-m response $X^m \to \Delta_i$ that likewise possesses the following properties.
  - Continuous.

II) : Responses are continuous, diffuse, and γ_i-optimal

- Recall that soft-max (mixed logit) possesses the following properties.
  - Continuous.
  - Gives positive probability to all actions.
  - Depending on γ_i, the expected utility of a response can be made arbitrary close to the expected utility of a best response (indeed, let $\gamma_i \downarrow 0$).
- Generalise soft-max to any memory-m response $X^m \to \Delta_i$ that likewise possesses the following properties.
  - Continuous.
  - Gives positive probability to all histories ("diffuse").

II) : Responses are continuous, diffuse, and γ_i-optimal

- Recall that soft-max (mixed logit) possesses the following properties.
  - Continuous.
  - Gives positive probability to all actions.
  - Depending on γ_i, the expected utility of a response can be made arbitrary close to the expected utility of a best response (indeed, let $\gamma_i \downarrow 0$).
- Generalise soft-max to any memory-m response $X^m \to \Delta_i$ that likewise possesses the following properties.
  - Continuous.
  - Gives positive probability to all histories ("diffuse").
  - Depending on γ_i, the expected utility of a response can be made arbitrary close to the expected utility of a best response.

III): Tests are "powerful"

III): Tests are "powerful"

- — A memory-m response a_i is a function $X^m \rightarrow \Delta_i$.

III): Tests are "powerful"

- — A memory-m response a_i is a function $X^m \to \Delta_i$.
  — A memory-m opponent model b_i is a function $X^m \to \Delta_{-i}$.

III): Tests are "powerful"

- — A memory-m response a_i is a function $X^m \to \Delta_i$.
  — A memory-m opponent model b_i is a function $X^m \to \Delta_{-i}$.
  — A memory-m hypothesis (a_i, b_i) is a function $X^m \to \Delta_i \times \Delta_{-i}$.

III): Tests are "powerful"

- — A memory-m response a_i is a function $X^m \to \Delta_i$.
  — A memory-m opponent model b_i is a function $X^m \to \Delta_{-i}$.
  — A memory-m hypothesis (a_i, b_i) is a function $X^m \to \Delta_i \times \Delta_{-i}$.
- Every memory-m hypothesis of i generates a subjective distribution over realisations of play Ω.

III): Tests are "powerful"

- — A memory-m response a_i is a function $X^m \to \Delta_i$.
  — A memory-m opponent model b_i is a function $X^m \to \Delta_{-i}$.
  — A memory-m hypothesis (a_i, b_i) is a function $X^m \to \Delta_i \times \Delta_{-i}$.
- Every memory-m hypothesis of i generates a subjective distribution over realisations of play Ω. (Recall the inductive construction of a probability distribution over histories H, then over continuations of histories, $C(H)$, then over Ω.)

III): Tests are "powerful"

- — A memory-m response a_i is a function $X^m \to \Delta_i$.
  — A memory-m opponent model b_i is a function $X^m \to \Delta_{-i}$.
  — A memory-m hypothesis (a_i, b_i) is a function $X^m \to \Delta_i \times \Delta_{-i}$.
- Every memory-m hypothesis of i generates a subjective distribution over realisations of play Ω. (Recall the inductive construction of a probability distribution over histories H, then over continuations of histories, $C(H)$, then over Ω.)
- There is also the actual distribution of play, formed by $a^t = (a_1^t, \ldots, a_n^t)$.

III): Tests are "powerful"

- — A memory-m response a_i is a function $X^m \to \Delta_i$.
  — A memory-m opponent model b_i is a function $X^m \to \Delta_{-i}$.
  — A memory-m hypothesis (a_i, b_i) is a function $X^m \to \Delta_i \times \Delta_{-i}$.
- Every memory-m hypothesis of i generates a subjective distribution over realisations of play Ω. (Recall the inductive construction of a probability distribution over histories H, then over continuations of histories, $C(H)$, then over Ω.)
- There is also the actual distribution of play, formed by $a^t = (a_1^t, \dots, a_n^t)$.
- Let τ (the critical region) be determined.

III): Tests are "powerful"

- — A memory-m response a_i is a function $X^m \to \Delta_i$.
  — A memory-m opponent model b_i is a function $X^m \to \Delta_{-i}$.
  — A memory-m hypothesis (a_i, b_i) is a function $X^m \to \Delta_i \times \Delta_{-i}$.
- Every memory-m hypothesis of i generates a subjective distribution over realisations of play Ω. (Recall the inductive construction of a probability distribution over histories H, then over continuations of histories, $C(H)$, then over Ω.)
- There is also the actual distribution of play, formed by $a^t = (a_1^t, \ldots, a_n^t)$.
- Let τ (the critical region) be determined.
  - Let α_t be the Type I error

III): Tests are "powerful"

- — A memory-m response a_i is a function $X^m \to \Delta_i$.
  — A memory-m opponent model b_i is a function $X^m \to \Delta_{-i}$.
  — A memory-m hypothesis (a_i, b_i) is a function $X^m \to \Delta_i \times \Delta_{-i}$.
- Every memory-m hypothesis of i generates a subjective distribution over realisations of play Ω. (Recall the inductive construction of a probability distribution over histories H, then over continuations of histories, $C(H)$, then over Ω.)
- There is also the actual distribution of play, formed by $a^t = (a_1^t, \dots, a_n^t)$.
- Let τ (the critical region) be determined.
  - Let α_t be the Type I error (probability of rejecting a hypothesis when it lies within tolerance τ of the actual distribution).

III): Tests are "powerful"

- — A memory-m response a_i is a function $X^m \to \Delta_i$.
  — A memory-m opponent model b_i is a function $X^m \to \Delta_{-i}$.
  — A memory-m hypothesis (a_i, b_i) is a function $X^m \to \Delta_i \times \Delta_{-i}$.
- Every memory-m hypothesis of i generates a subjective distribution over realisations of play Ω. (Recall the inductive construction of a probability distribution over histories H, then over continuations of histories, $C(H)$, then over Ω.)
- There is also the actual distribution of play, formed by $a^t = (a_1^t, \dots, a_n^t)$.
- Let τ (the critical region) be determined.
  - Let α_t be the Type I error (probability of rejecting a hypothesis when it lies within tolerance τ of the actual distribution).
  - Let β_t be the Type II error

III): Tests are "powerful"

- — A memory-m response a_i is a function $X^m \to \Delta_i$.
  — A memory-m opponent model b_i is a function $X^m \to \Delta_{-i}$.
  — A memory-m hypothesis (a_i, b_i) is a function $X^m \to \Delta_i \times \Delta_{-i}$.
- Every memory-m hypothesis of i generates a subjective distribution over realisations of play Ω. (Recall the inductive construction of a probability distribution over histories H, then over continuations of histories, $C(H)$, then over Ω.)
- There is also the actual distribution of play, formed by $a^t = (a_1^t, \dots, a_n^t)$.
- Let τ (the critical region) be determined.
  - Let α_t be the Type I error (probability of rejecting a hypothesis when it lies within tolerance τ of the actual distribution).
  - Let β_t be the Type II error (probability of accepting a hypothesis when it is not tolerated by the actual distribution).

III) : Tests are “powerful”

III) : Tests are "powerful"

A family of tests is said to be powerful for τ and s if for every $\tau > 0$ there exist $0 < c(\tau) < \tau$ and positive functions r and k such that

$$\alpha_{c(\tau)}, \beta_\tau \leq k(\tau)e^{-r(\tau)s}.$$

III) : Tests are "powerful"

A family of tests is said to be powerful for τ and s if for every $\tau > 0$ there exist $0 < c(\tau) < \tau$ and positive functions r and k such that

$$\alpha_{c(\tau)}, \beta_\tau \leq k(\tau) e^{-r(\tau)s}.$$

Explanation:

III) : Tests are "powerful"

A family of tests is said to be powerful for τ and s if for every $\tau > 0$ there exist $0 < c(\tau) < \tau$ and positive functions r and k such that

$$\alpha_{c(\tau)}, \beta_\tau \leq k(\tau) e^{-r(\tau)s}.$$

Explanation:

- Whenever the hypothesis is within $c(\tau)$ of the truth, the probability of rejection is small.

III) : Tests are "powerful"

A family of tests is said to be powerful for τ and s if for every $\tau > 0$ there exist $0 < c(\tau) < \tau$ and positive functions r and k such that

$$\alpha_{c(\tau)}, \beta_\tau \leq k(\tau)e^{-r(\tau)s}.$$

Explanation:

■
- Whenever the hypothesis is within $c(\tau)$ of the truth, the probability of rejection is small.
- Whenever the hypothesis more than τ from the truth, the probability of acceptance is small.

III) : Tests are "powerful"

A family of tests is said to be powerful for τ and s if for every $\tau > 0$ there exist $0 < c(\tau) < \tau$ and positive functions r and k such that

$$\alpha_{c(\tau)}, \beta_\tau \leq k(\tau)e^{-r(\tau)s}.$$

Explanation:

- ■ • Whenever the hypothesis is within $c(\tau)$ of the truth, the probability of rejection is small.
  - • Whenever the hypothesis more than τ from the truth, the probability of acceptance is small.
- ■ If both α and β are small, a test discriminates sharply between models that are near the truth and those that are far from the truth.

III) : Tests are “powerful”

A family of tests is said to be powerful for τ and s if for every $\tau > 0$ there exist $0 < c(\tau) < \tau$ and positive functions r and k such that

$$\alpha_{c(\tau)}, \beta_\tau \leq k(\tau)e^{-r(\tau)s}.$$

Explanation:

- - Whenever the hypothesis is within $c(\tau)$ of the truth, the probability of rejection is small.
  - Whenever the hypothesis more than τ from the truth, the probability of acceptance is small.
- If both α and β are small, a test discriminates sharply between models that are near the truth and those that are far from the truth.
- This rate of distinction is related exponentially to s and τ.

V) : Model revision is "flexible"

V) : Model revision is "flexible"

- In the simple framework, a revised model is chosen randomly with probability ρ.

V) : Model revision is "flexible"

- In the simple framework, a revised model is chosen randomly with probability ρ.
- **Definition**. A set $A \subseteq B$ lies dense in B if every $b \in B$ is in A or arbitrarily close to a member of A. Example: $\mathbb{Q}$ lies dense in $\mathbb{R}$.

V) : Model revision is "flexible"

- In the simple framework, a revised model is chosen randomly with probability ρ.
- **Definition**. A set $A \subseteq B$ lies dense in B if every $b \in B$ is in A or arbitrarily close to a member of A. Example: $\mathbb{Q}$ lies dense in $\mathbb{R}$.
- In the generalised framework, the following condition must hold.

V) : Model revision is "flexible"

- In the simple framework, a revised model is chosen randomly with probability ρ.
- **Definition**. A set $A \subseteq B$ lies dense in B if every $b \in B$ is in A or arbitrarily close to a member of A. Example: $\mathbb{Q}$ lies dense in $\mathbb{R}$.
- In the generalised framework, the following condition must hold.
  - The distribution of revised models lies dense in Δ^{M}_{-i}.

V) : Model revision is "flexible"

- In the simple framework, a revised model is chosen randomly with probability ρ.
- **Definition**. A set $A \subseteq B$ lies dense in B if every $b \in B$ is in A or arbitrarily close to a member of A. Example: $\mathbb{Q}$ lies dense in $\mathbb{R}$.
- In the generalised framework, the following condition must hold.
  - The distribution of revised models lies dense in Δ^M_{-i}.
  - I.e., for every $\epsilon > 0$ and every opponent model p_i, the probability of choosing a new model within ϵ from p_i is non-zero.

V) : Model revision is “flexible”

- In the simple framework, a revised model is chosen randomly with probability ρ.
- **Definition**. A set $A \subseteq B$ lies dense in B if every $b \in B$ is in A or arbitrarily close to a member of A. Example: $\mathbb{Q}$ lies dense in $\mathbb{R}$.
- In the generalised framework, the following condition must hold.
  - The distribution of revised models lies dense in Δ^M_{-i}.
  - I.e., for every $\epsilon > 0$ and every opponent model p_i, the probability of choosing a new model within ϵ from p_i is non-zero.
  - I.e., for every $\epsilon > 0$ there is a $\delta > 0$ such that for every $p_i \in \Delta^M_{-i}$

$$\Pr\{\|p_i - p^i_t\| \leq \epsilon\} \geq \delta.$$

V) : Model revision is "flexible"

- In the simple framework, a revised model is chosen randomly with probability ρ.
- **Definition**. A set $A \subseteq B$ lies dense in B if every $b \in B$ is in A or arbitrarily close to a member of A. Example: $\mathbb{Q}$ lies dense in $\mathbb{R}$.
- In the generalised framework, the following condition must hold.
  - The distribution of revised models lies dense in Δ_{-i}^{M}.
  - I.e., for every $\epsilon > 0$ and every opponent model p_i, the probability of choosing a new model within ϵ from p_i is non-zero.
  - I.e., for every $\epsilon > 0$ there is a $\delta > 0$ such that for every $p_i \in \Delta_{-i}^{M}$

$$\Pr\{\|p_i - p_t^i\| \leq \epsilon\} \geq \delta.$$

  Such a model revision process is said to be flexible.

Generalised theorem

Generalised theorem

Theorem (Foster and Young, 2003).

Generalised theorem

Theorem (Foster and Young, 2003). Let G be a finite game where n players enjoy normalised payoffs and use a hypothesis testing procedure satisfying conditions I-V

Generalised theorem

Theorem (Foster and Young, 2003). Let G be a finite game where n players enjoy normalised payoffs and use a hypothesis testing procedure satisfying conditions I-V with memory m

Generalised theorem

Theorem (Foster and Young, 2003). Let G be a finite game where n players enjoy normalised payoffs and use a hypothesis testing procedure satisfying conditions I-V with memory m and learning parameters γ_i, τ_i, and s_i.

Generalised theorem

Theorem (Foster and Young, 2003). Let G be a finite game where n players enjoy normalised payoffs and use a hypothesis testing procedure satisfying conditions I-V with memory m and learning parameters γ_i, τ_i, and s_i. For every $\epsilon > 0$

Generalised theorem

Theorem (Foster and Young, 2003). Let G be a finite game where n players enjoy normalised payoffs and use a hypothesis testing procedure satisfying conditions I-V with memory m and learning parameters γ_i, τ_i, and s_i. For every $\epsilon > 0$, if the responses are sufficiently rational (the γ_i are small)

Generalised theorem

Theorem (Foster and Young, 2003). Let G be a finite game where n players enjoy normalised payoffs and use a hypothesis testing procedure satisfying conditions I-V with memory m and learning parameters γ_i, τ_i, and s_i. For every $\epsilon > 0$, if the responses are sufficiently rational (the γ_i are small), and the test tolerances τ_i are sufficiently fine (given the γ_i)

Generalised theorem

Theorem (Foster and Young, 2003). Let G be a finite game where n players enjoy normalised payoffs and use a hypothesis testing procedure satisfying conditions I-V with memory m and learning parameters γ_i, τ_i, and s_i. For every $\epsilon > 0$, if the responses are sufficiently rational (the γ_i are small), and the test tolerances τ_i are sufficiently fine (given the γ_i), and the amounts of data collected s_i are comparable and sufficiently large (given γ_i and τ_i)

Generalised theorem

Theorem (Foster and Young, 2003). Let G be a finite game where n players enjoy normalised payoffs and use a hypothesis testing procedure satisfying conditions I-V with memory m and learning parameters γ_i, τ_i, and s_i. For every $\epsilon > 0$, if the responses are sufficiently rational (the γ_i are small), and the test tolerances τ_i are sufficiently fine (given the γ_i), and the amounts of data collected s_i are comparable and sufficiently large (given γ_i and τ_i), then the the responses constitute an ϵ-equilibrium of the repeated game at least $1 - \epsilon$ of the time

Generalised theorem

Theorem (Foster and Young, 2003). Let G be a finite game where n players enjoy normalised payoffs and use a hypothesis testing procedure satisfying conditions I-V with memory m and learning parameters γ_i, τ_i, and s_i. For every $\epsilon > 0$, if the responses are sufficiently rational (the γ_i are small), and the test tolerances τ_i are sufficiently fine (given the γ_i), and the amounts of data collected s_i are comparable and sufficiently large (given γ_i and τ_i), then the the responses constitute an ϵ-equilibrium of the repeated game at least $1 - \epsilon$ of the time a.s.

Generalised theorem

Theorem (Foster and Young, 2003). Let G be a finite game where n players enjoy normalised payoffs and use a hypothesis testing procedure satisfying conditions I-V with memory m and learning parameters γ_i, τ_i, and s_i. For every $\epsilon > 0$, if the responses are sufficiently rational (the γ_i are small), and the test tolerances τ_i are sufficiently fine (given the γ_i), and the amounts of data collected s_i are comparable and sufficiently large (given γ_i and τ_i), then the the responses constitute an ϵ-equilibrium of the repeated game at least $1 - \epsilon$ of the time a.s.

Informally: more or less any reasonable hypothesis testing strategy 'works' in the sense that learned behaviours are eventually close to Nash equilibrium most of the time.

Generalised theorem: towards a proof

Generalised theorem: towards a proof

What needs to be shown, in a proof, is the following.

Generalised theorem: towards a proof

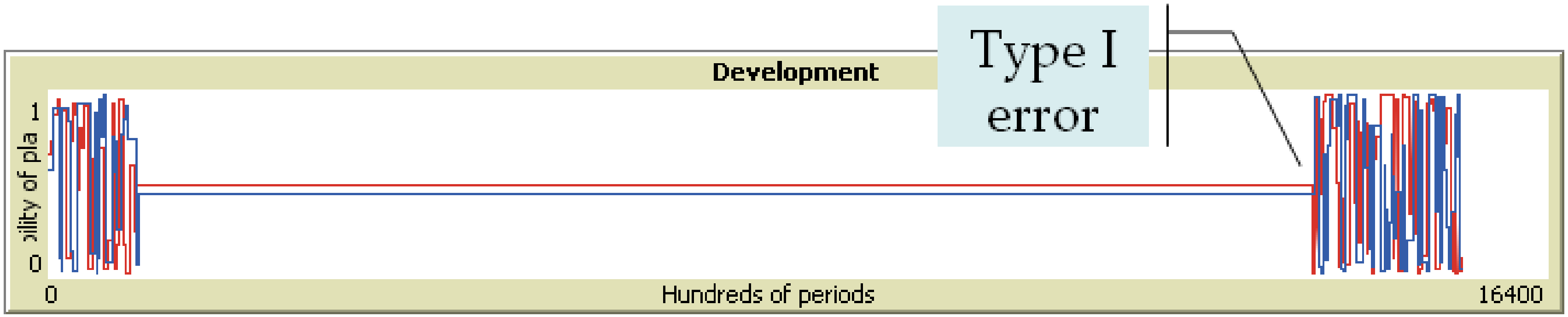
What needs to be shown, in a proof, is the following.

1. The probability of making a Type I error is small relative to the time it takes for a player to come close to the true model.

Generalised theorem: towards a proof

What needs to be shown, in a proof, is the following.

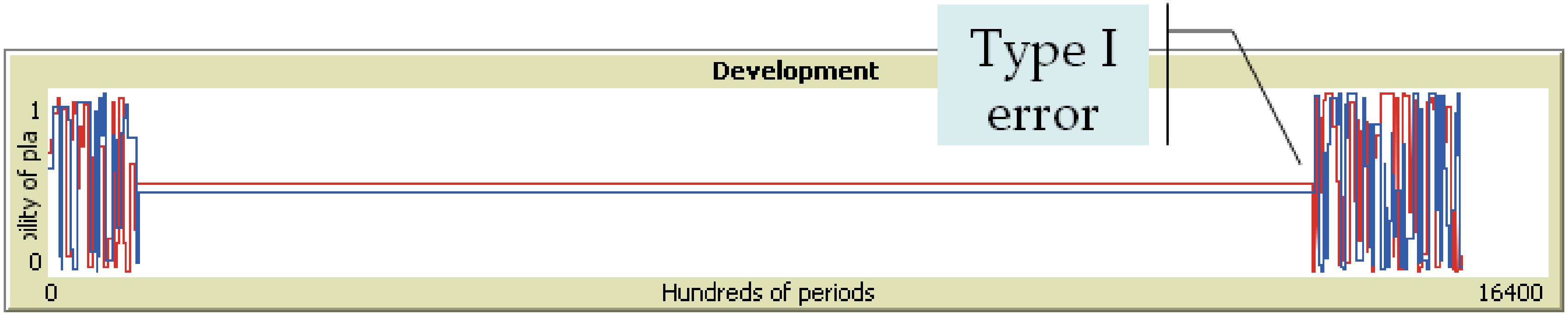
1. The probability of making a Type I error is small relative to the time it takes for a player to come close to the true model.



Generalised theorem: towards a proof

What needs to be shown, in a proof, is the following.

1. The probability of making a Type I error is small relative to the time it takes for a player to come close to the true model.

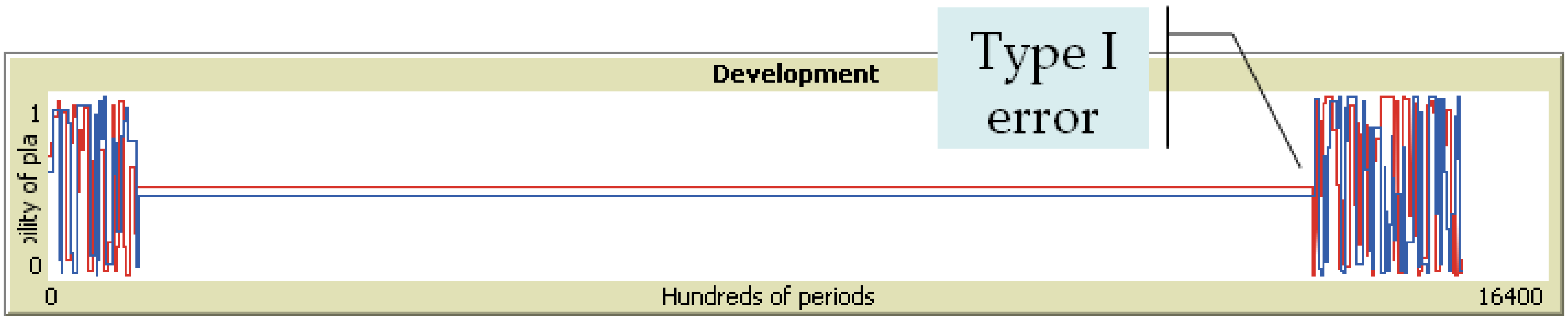


Preferably exponentially small.

Generalised theorem: towards a proof

What needs to be shown, in a proof, is the following.

1. The probability of making a Type I error is small relative to the time it takes for a player to come close to the true model.



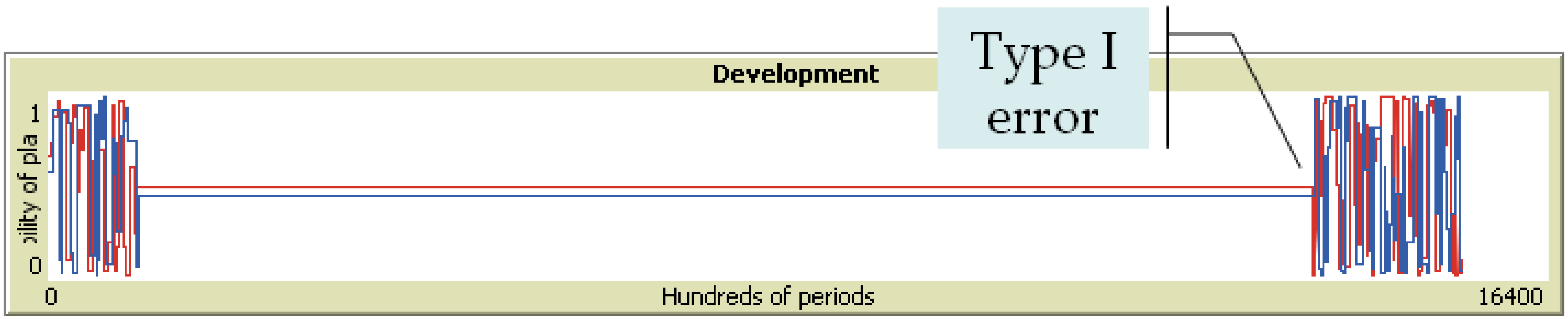
Preferably exponentially small.

2. There must be sufficiently many times where all hypotheses jointly approach the true strategy profile.

Generalised theorem: towards a proof

What needs to be shown, in a proof, is the following.

1. The probability of making a Type I error is small relative to the time it takes for a player to come close to the true model.



   Preferably exponentially small.

2. There must be sufficiently many times where all hypotheses jointly approach the true strategy profile.

H. Peyton Young, Dean P. Foster (2003). "Learning, Hypothesis Testing, and Nash Equilibrium," in *Games and Economic Behavior*, Vol. **45**, pp. 73-96.

Part VI:
Observations

Some observations

Some observations

- **Vs. Bayesian play.** Theorem (Kalai and Lehrer). *Let μ represent the true distribution of play, and let μ_i represent i's view on the distribution of play. If $\mu \ll \mu_i$, then μ_i tends to play like μ.*

  $\mu \ll \mu_i$ is bold assumption.

Some observations

- **Vs. Bayesian play.** Theorem (Kalai and Lehrer). *Let μ represent the true distribution of play, and let μ_i represent i's view on the distribution of play. If $\mu \ll \mu_i$, then μ_i tends to play like μ.*

  $\mu \ll \mu_i$ is bold assumption. It assumes that Player i's priors must contain a grain of truth about the eventual course of play.

Some observations

- **Vs. Bayesian play.** Theorem (Kalai and Lehrer). *Let μ represent the true distribution of play, and let μ_i represent i's view on the distribution of play. If $\mu \ll \mu_i$, then μ_i tends to play like μ.*

  $\mu \ll \mu_i$ is bold assumption. It assumes that Player i's priors must contain a grain of truth about the eventual course of play.

- **Young's notion of a "Hunch".** A continuous, diffuse and γ_i-optimal response is determined by two factors.

Some observations

- **Vs. Bayesian play.** Theorem (Kalai and Lehrer). *Let μ represent the true distribution of play, and let μ_i represent i's view on the distribution of play. If $\mu \ll \mu_i$, then μ_i tends to play like μ.*

  $\mu \ll \mu_i$ is bold assumption. It assumes that Player i's priors must contain a grain of truth about the eventual course of play.

- **Young's notion of a "Hunch".** A continuous, diffuse and γ_i-optimal response is determined by two factors.

  - History of play.

Some observations

- **Vs. Bayesian play.** Theorem (Kalai and Lehrer). *Let μ represent the true distribution of play, and let μ_i represent i's view on the distribution of play. If $\mu \ll \mu_i$, then μ_i tends to play like μ.*

  $\mu \ll \mu_i$ is bold assumption. It assumes that Player i's priors must contain a grain of truth about the eventual course of play.

- **Young's notion of a "Hunch".** A continuous, diffuse and γ_i-optimal response is determined by two factors.

  - History of play.
  - Coincidence.

Some observations

- **Vs. Bayesian play.** Theorem (Kalai and Lehrer). *Let μ represent the true distribution of play, and let μ_i represent i's view on the distribution of play. If $\mu \ll \mu_i$, then μ_i tends to play like μ.*

  $\mu \ll \mu_i$ is bold assumption. It assumes that Player i's priors must contain a grain of truth about the eventual course of play.

- **Young's notion of a "Hunch".** A continuous, diffuse and γ_i-optimal response is determined by two factors.
  - History of play.
  - Coincidence.

- **Annealing.** Tighten learning parameters slowly $\Rightarrow$ convergence in probability to the set of Nash equilibria.

Part VII:
Extensions

Convergence in probability

Convergence in probability

- Bayesian play already shows that the “convergence to Nash” program has it limits.

Convergence in probability

- Bayesian play already shows that the "convergence to Nash" program has it limits.
- Given $\epsilon > 0$, the previous theorem ensures that we can choose γ, τ and s such that the strategies are ϵ-Nash $1 - \epsilon$ of the time.

Convergence in probability

- Bayesian play already shows that the "convergence to Nash" program has it limits.
- Given $\epsilon > 0$, the previous theorem ensures that we can choose γ, τ and s such that the strategies are ϵ-Nash $1 - \epsilon$ of the time.
- Assume that at the end of each period ϵ is halved with some probability depending on ϵ itself.

Convergence in probability

- Bayesian play already shows that the "convergence to Nash" program has it limits.
- Given $\epsilon > 0$, the previous theorem ensures that we can choose γ, τ and s such that the strategies are ϵ-Nash $1 - \epsilon$ of the time.
- Assume that at the end of each period ϵ is halved with some probability depending on ϵ itself.
- If $T_k =_{Def}$ "first time at which some player tightens ϵ for the kth time"

Convergence in probability

- Bayesian play already shows that the "convergence to Nash" program has it limits.
- Given $\epsilon > 0$, the previous theorem ensures that we can choose γ, τ and s such that the strategies are ϵ-Nash $1 - \epsilon$ of the time.
- Assume that at the end of each period ϵ is halved with some probability depending on ϵ itself.
- If $T_k =_{Def}$ "first time at which some player tightens ϵ for the kth time", then it can be shown that $t \geq T_k \Rightarrow$ strategies are in $\epsilon/2^{k-1}$-Nash equilibrium $1 - \epsilon/2^{k-1}$ of the time.

Convergence in probability

- Bayesian play already shows that the "convergence to Nash" program has it limits.
- Given $\epsilon > 0$, the previous theorem ensures that we can choose γ, τ and s such that the strategies are ϵ-Nash $1 - \epsilon$ of the time.
- Assume that at the end of each period ϵ is halved with some probability depending on ϵ itself.
- If $T_k =_{Def}$ "first time at which some player tightens ϵ for the kth time", then it can be shown that $t \geq T_k \Rightarrow$ strategies are in $\epsilon/2^{k-1}$-Nash equilibrium $1 - \epsilon/2^{k-1}$ of the time.

Theorem (Foster and Young, 2003).

Convergence in probability

- Bayesian play already shows that the "convergence to Nash" program has it limits.
- Given $\epsilon > 0$, the previous theorem ensures that we can choose γ, τ and s such that the strategies are ϵ-Nash $1 - \epsilon$ of the time.
- Assume that at the end of each period ϵ is halved with some probability depending on ϵ itself.
- If $T_k =_{Def}$ "first time at which some player tightens ϵ for the kth time", then it can be shown that $t \geq T_k \Rightarrow$ strategies are in $\epsilon/2^{k-1}$-Nash equilibrium $1 - \epsilon/2^{k-1}$ of the time.

Theorem (Foster and Young, 2003). Let G be a finite n-person game.

Convergence in probability

- Bayesian play already shows that the "convergence to Nash" program has it limits.
- Given $\epsilon > 0$, the previous theorem ensures that we can choose γ, τ and s such that the strategies are ϵ-Nash $1 - \epsilon$ of the time.
- Assume that at the end of each period ϵ is halved with some probability depending on ϵ itself.
- If $T_k =_{Def}$ "first time at which some player tightens ϵ for the kth time", then it can be shown that $t \geq T_k \Rightarrow$ strategies are in $\epsilon/2^{k-1}$-Nash equilibrium $1 - \epsilon/2^{k-1}$ of the time.

Theorem (Foster and Young, 2003). Let G be a finite n-person game. If the hypothesis testing parameters are tightened sufficiently slowly

Convergence in probability

- Bayesian play already shows that the "convergence to Nash" program has it limits.
- Given $\epsilon > 0$, the previous theorem ensures that we can choose γ, τ and s such that the strategies are ϵ-Nash $1 - \epsilon$ of the time.
- Assume that at the end of each period ϵ is halved with some probability depending on ϵ itself.
- If $T_k =_{Def}$ "first time at which some player tightens ϵ for the kth time", then it can be shown that $t \geq T_k \Rightarrow$ strategies are in $\epsilon/2^{k-1}$-Nash equilibrium $1 - \epsilon/2^{k-1}$ of the time.

Theorem (Foster and Young, 2003). Let G be a finite n-person game. If the hypothesis testing parameters are tightened sufficiently slowly, the players' responses converge set-wise[a] in probability[b] to the set of Nash equilibria $\mathcal{N}$ of the repeated game.

[a] Not necessarily pointwise.

[b] For all $\epsilon > 0$, $\lim_{n\to\infty} \Pr\{|X_n - X| > \epsilon\} = 0$.

Learning to predict

Learning to predict

- Player i's predictive learning rule yields a forecast $p_t^{-i} \in \Delta_{-i}$.

Learning to predict

- Player i's predictive learning rule yields a forecast $p_t^{-i} \in \Delta_{-i}$.
- The actual reply rules g induce the distribution proper, $q_t^{-i} \in \Delta_{-i}$.

Learning to predict

- Player i's predictive learning rule yields a forecast $p_t^{-i} \in \Delta_{-i}$.
- The actual reply rules g induce the distribution proper, $q_t^{-i} \in \Delta_{-i}$.
- Player i is said to be a good predictor if

$$\lim_{t\to\infty} \frac{1}{t} \sum_{s\leq t} \| q_s^{-i} - p_s^{-i} \|^2 \qquad (1)$$

goes to zero a.s.

Learning to predict

- Player i's predictive learning rule yields a forecast $p_t^{-i} \in \Delta_{-i}$.
- The actual reply rules g induce the distribution proper, $q_t^{-i} \in \Delta_{-i}$.
- Player i is said to be a good predictor if

$$\lim_{t \to \infty} \frac{1}{t} \sum_{s \leq t} \| q_s^{-i} - p_s^{-i} \|^2 \qquad (1)$$

goes to zero a.s.

- Player i is said to be an ϵ-good predictor if (1) $\leq \epsilon$ a.s.

Learning to predict

- Player i's predictive learning rule yields a forecast $p_t^{-i} \in \Delta_{-i}$.
- The actual reply rules g induce the distribution proper, $q_t^{-i} \in \Delta_{-i}$.
- Player i is said to be a good predictor if

$$\lim_{t\to\infty} \frac{1}{t} \sum_{s\leq t} \|q_s^{-i} - p_s^{-i}\|^2 \qquad (1)$$

  goes to zero a.s.
- Player i is said to be an ϵ-good predictor if (1) $\leq \epsilon$ a.s.
- It is said that prediction matters by at most ϵ to i if i has a memory-m strategy that is an ϵ-best response to every memory-m model of his opponents.

Learning to predict

- Player i's predictive learning rule yields a forecast $p_t^{-i} \in \Delta_{-i}$.
- The actual reply rules g induce the distribution proper, $q_t^{-i} \in \Delta_{-i}$.
- Player i is said to be a good predictor if

$$\lim_{t\to\infty} \frac{1}{t} \sum_{s \leq t} \| q_s^{-i} - p_s^{-i} \|^2 \qquad (1)$$

  goes to zero a.s.
- Player i is said to be an ϵ-good predictor if (1) $\leq \epsilon$ a.s.
- It is said that prediction matters by at most ϵ to i if i has a memory-m strategy that is an ϵ-best response to every memory-m model of his opponents.

Theorem (Foster and Young, 2003).

Learning to predict

- Player i's predictive learning rule yields a forecast $p_t^{-i} \in \Delta_{-i}$.
- The actual reply rules g induce the distribution proper, $q_t^{-i} \in \Delta_{-i}$.
- Player i is said to be a good predictor if

$$\lim_{t\to\infty} \frac{1}{t} \sum_{s\leq t} \|q_s^{-i} - p_s^{-i}\|^2 \qquad (1)$$

goes to zero a.s.

- Player i is said to be an ϵ-good predictor if (1) $\leq \epsilon$ a.s.
- It is said that prediction matters by at most ϵ to i if i has a memory-m strategy that is an ϵ-best response to every memory-m model of his opponents.

Theorem (Foster and Young, 2003). For every $\epsilon > 0$ the learning parameters may be chosen so that all players for whom prediction matters by at least ϵ, are ϵ-good predictors.

Learning to predict

- Player i's predictive learning rule yields a forecast $p_t^{-i} \in \Delta_{-i}$.
- The actual reply rules g induce the distribution proper, $q_t^{-i} \in \Delta_{-i}$.
- Player i is said to be a good predictor if

$$\lim_{t\to\infty} \frac{1}{t} \sum_{s\leq t} \|q_s^{-i} - p_s^{-i}\|^2 \qquad (1)$$

goes to zero a.s.

- Player i is said to be an ϵ-good predictor if (1) $\leq \epsilon$ a.s.
- It is said that prediction matters by at most ϵ to i if i has a memory-m strategy that is an ϵ-best response to every memory-m model of his opponents.

Theorem (Foster and Young, 2003). For every $\epsilon > 0$ the learning parameters may be chosen so that all players for whom prediction matters by at least ϵ, are ϵ-good predictors. Further, the learning parameters can be tightened at a rate such that all players are good predictors.