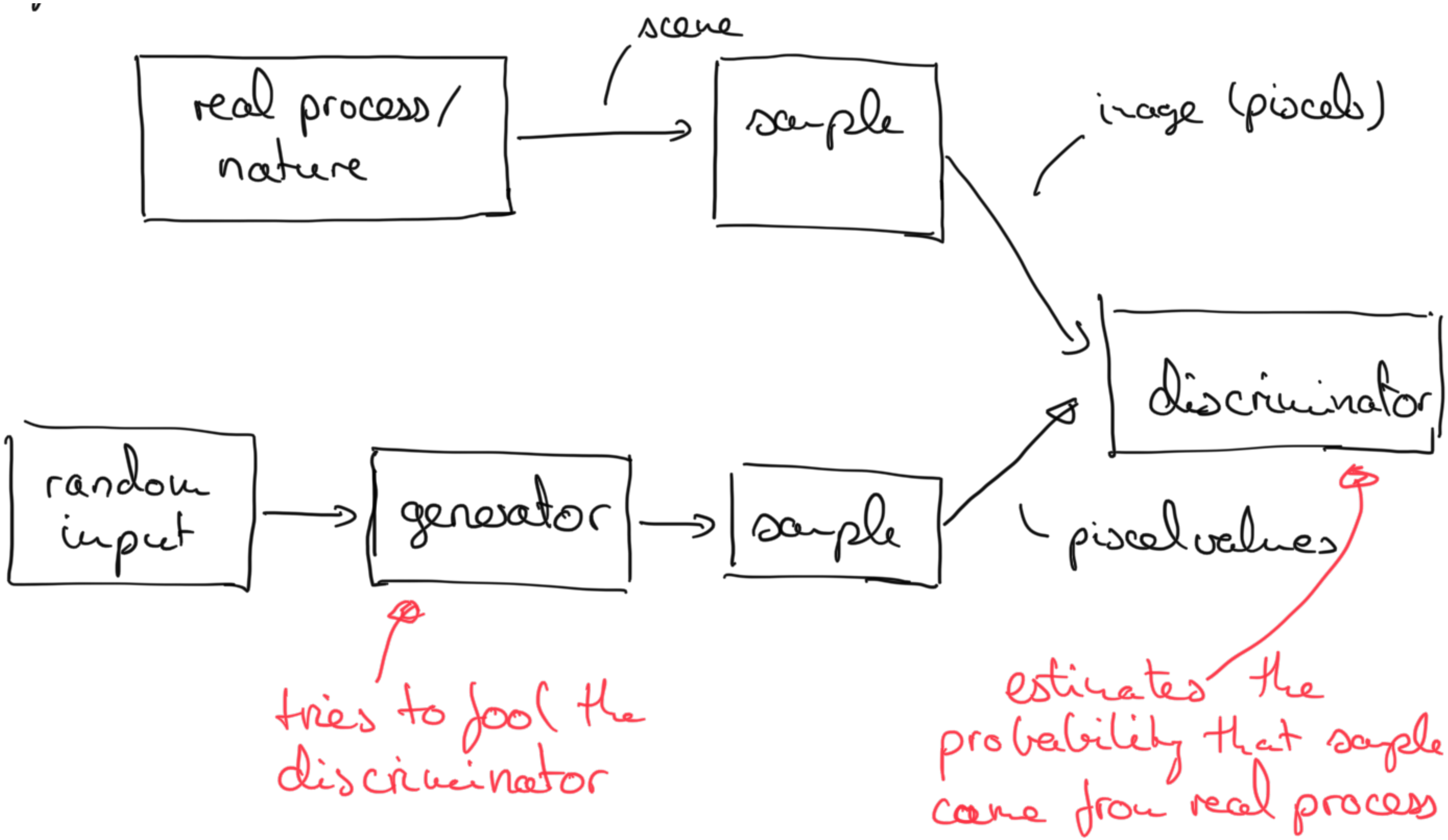
Brief overview of game theory

Literature:
- A. R. Karlin, Y. Peres, "Game Theory, Alive", AMS, 2016
- N. Cesa Bianchi, G. Lugosi, "Predictions, Learning and Games", Cambridge University Press, 2006

1. Motivation

i). learning system is interacting with an environment
$\rightarrow$ for example in the smart grid, different entities make decisions about production, consumption, and storage of energy and have conflicting interests.

ii). generative adversarial networks:



iii). convex optimization : → design an algorithm that works for any function in a given class

→ what is the worst-case rate ?

2. Two-player zero-sum games

- player I can choose between actions $1 \dots m$
- player II can choose between actions $1 \dots n$

→ player I receives a_{ij}

→ player II receives $-a_{ij}$

for choosing action i (player I) and action j (player II)

- a_{ij}, $i = 1, \dots, m$, $j = 1, \dots, n$ is the payoff
- Example

$$\begin{array}{ccccc} & & & \text{p.II} & \\ & & 1 & & 2 \\ \text{p.I} & 1 & 1 & & 2 \\ & 2 & 3 & & 4 \end{array}$$

→ How should they play → think about the worst-case.

→ p. I has an incentive to play 2
→ p. II has an incentive to play 1

→ this leads to an eq. → no player has an incentive to deviate from his worst-case strategy

		p. II: 1	p. II: 2
p. I	1	1	0.5
	2	0	2

→ p. I has an incentive for 1
→ p. II has an incentive for 1

→ this does not lead to an eq.

→ this ...

→ How can we resolve this (from a mathematical point of view) ?

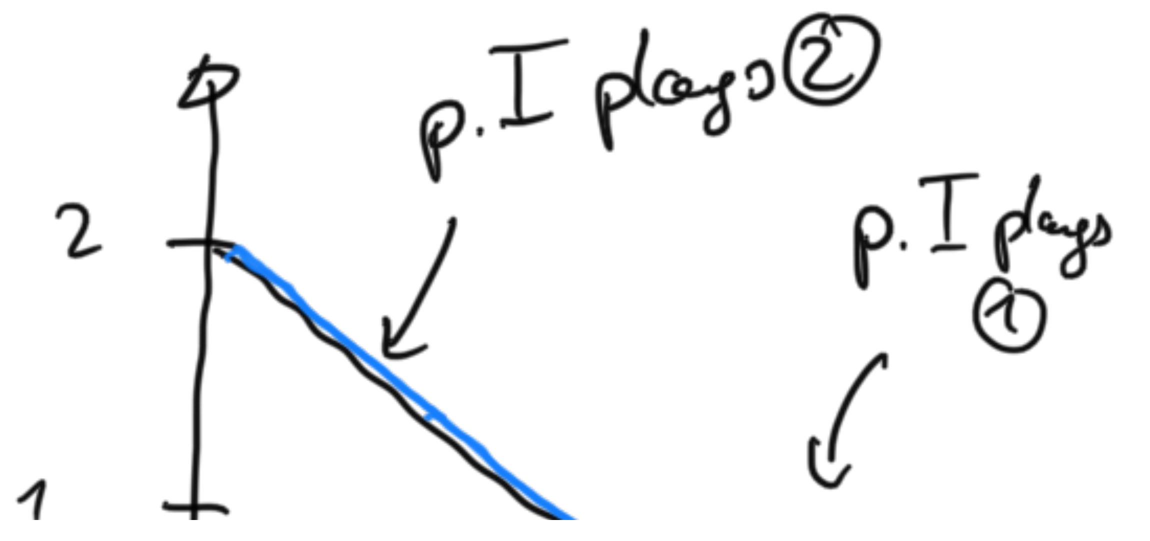
- → p. I chooses action i with prob. x_i
- → p. II chooses action j with prob. y_j

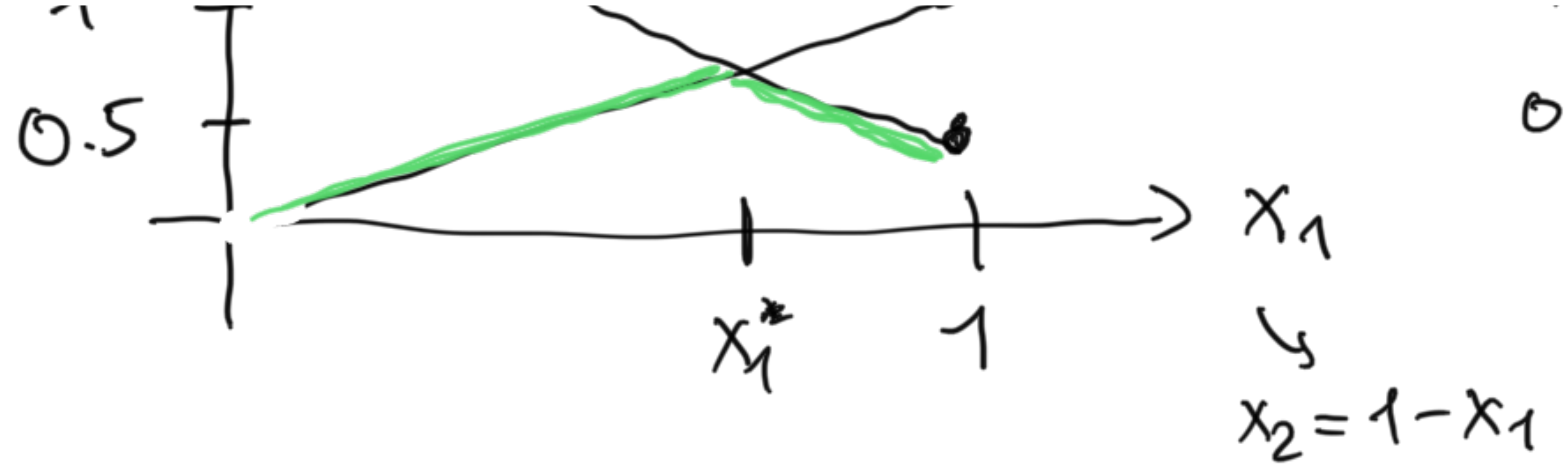
- → $x = (x_1, \dots, x_m) \in \Delta_m$
- → $y = (y_1 \dots, y_n) \in \Delta_n$

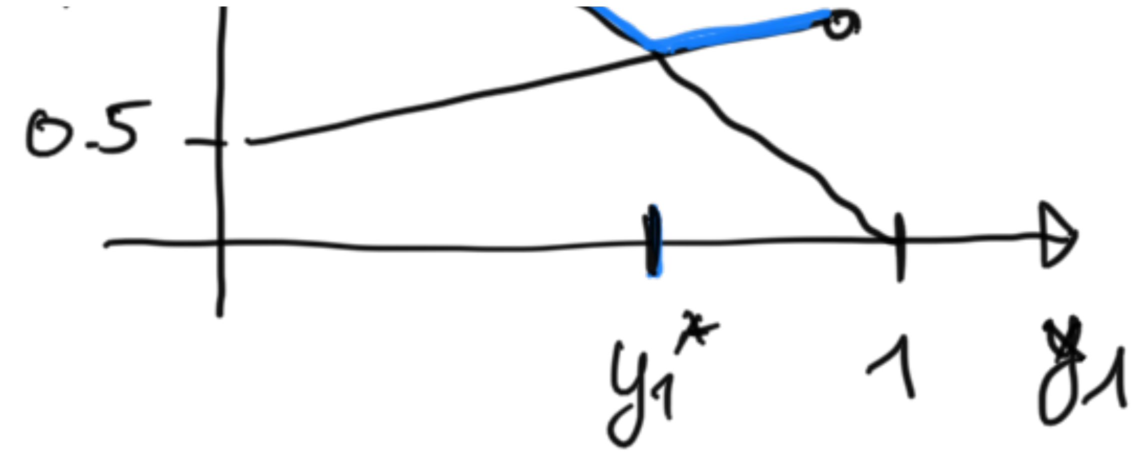
- gain for p. I : $x^T A y$
- gain for p. II : $-x^T A y$

from p. I's perspective

from p II's persp.







$$x_1^* = 2(1 - x_1^*) + 0.5\, x_1^* \rightsquigarrow \begin{array}{l} x_1^* = 4/5 \\ x_2^* = 1/5 \end{array}$$

$$y_1^* = 2/3, \quad y_2^* = 1/3$$

- Assume p. I plays (x_1^*, x_2^*), the payoff for p. II is $-4/5$ for action 1 and $-4/5$ for action 2
  $\to$ p. II has no incentive in deviating
- <u>Def.</u> (x^*, y^*) is a Nash eq. or equilibrium if

$$x^{*T} A y^* \leq x^{*T} A y \qquad \forall y \in \Delta_n$$

$$x^{*T}Ay^{*} \geq x^TAy^{*} \qquad \forall x \in \Delta_m$$

<u>Prop.</u> The following are equivalent:

(i) there exists a Nash eq.

(ii) $$v = \underbrace{\max_{x\in\Delta_m} \underbrace{\min_{y\in\Delta_n} x^TAy}_{\text{worst case from p. I's persp.}}} = \underbrace{\min_{y\in\Delta_n} \underbrace{\max_{x\in\Delta_m} x^TAy}_{\text{worst case from p. II's persp.}}}$$

∘ <u>Remarks</u>: ∘ interpretation of (ii)

$\Rightarrow$ can compute strategies to achieve v

$$\max_{x\in\Delta_m} \underbrace{\min_{y\in\Delta_n} x^TAy}_{\text{p. II knows about}} = \min_{y\in\Delta_n} \underbrace{\max_{x\in\Delta_m} x^TAy}_{\text{p. I knows about}}$$

$$x \qquad\qquad\qquad y$$

$$\rightarrow \underbrace{\min_{y\in\Delta_n} \max_{x\in\Delta_m} x^T A y}_{=v_1} \geq \underbrace{\max_{x\in\Delta_m} \min_{y\in\Delta_n} x^T A y}_{=v_2}$$

- <u>Proof</u> (i) $\Rightarrow$ (ii) (x^*, y^*) is a Nash eq.

$$x^{*T} A y^* = \underbrace{\max_{x\in\Delta_m} x^T A y^*}_{g(y^*)} \geq \underbrace{\min_{y\in\Delta_n} \max_{x\in\Delta_m} x^T A y}_{=v_1}$$

$$x^* A y^* = \underbrace{\min_{y\in\Delta_n} x^{*T} A y}_{\tilde{g}(x^*)} \leq \underbrace{\max_{x\in\Delta_m} \min_{y\in\Delta_n} x^T A y}_{=v_2}$$

$$\Rightarrow \quad v_1 = v_2 .$$

(ii) $\Rightarrow$ (i) $\quad x^* = \underset{x\in\Delta_m}{\operatorname{argmax}} \min_{y\in\Delta_n} x^T A y$

$$y^* = \arg\min_{y \in \Delta_n} \max_{x \in \Delta_m} x^T A y$$

<u>Theorem</u> (Nash) There exists a Nash eq.
(holds also for general sum games)

<u>Proof.</u> Idea: introduce a map

$$T : \Delta_m \times \Delta_n \to \Delta_m \times \Delta_n$$

$$T(x,y) = (x',y'), \text{ such that}$$

(x',y') improves over (x,y).

previous reward

$$\to \text{define } c_i(x,y) = \max\left\{ \sum_{j=1}^{n} a_{ij} y_j - x^T A y, 0 \right\}$$

$$d_j(x,y) = \max\left\{ x^T A y - \sum_{i=1}^{m} a_{ij} x_i , 0 \right\}$$

$x_i + c_i(x,y)$

$$x_i' = \frac{x_i + c_i(x,y)}{\boxed{1 + \sum_{k=1}^{m} c_k(x,y)}} \longrightarrow \text{normalization}$$

$$y_j' = \frac{y_j + d_j(x,y)}{\boxed{1 + \sum_{k=1}^{n} d_k(x,y)}} \longrightarrow \text{norm.}$$

- T is continuous, it maps a compact convex set to itself.

→ Brouwer's theorem to conclude that $\exists\, (x^*, y^*) : T(x^*, y^*) = (x^*, y^*)$.

→ (x^*, y^*) corresponds to a Nash eq.

---

3. Adaptive decision making

5. Adaptive

- Stock market prediction:
  - → different algorithms make predictions about buying/selling stocks
  - → at each day we follow one of the alg. predictions
  - → at the end of the day we observe the outcome of all the different algorithms including the one we decided to follow
  - → 1'500 days horizon → which algorithm should you follow to maximize revenue?

- Formalization:

between

- A decider has to pick between actions $\{1, \dots, n\}$
- At each round t:

  (i) for each action $i \in \{1, \dots, n\}$ the environment generates a loss $\ell_i^t \in [0, 1]$.

  (ii) the decider chooses a prob. distribution $y^t \in \Delta_n$ over actions based on $\ell^1, \ell^2, \dots, \ell^{t-1}$, $y^1, y^2, \dots, y^{t-1}$, where $\ell^j = (\ell_1^j, \ell_2^j, \dots, \ell_n^j)$

  (iii) the decider suffers the loss $\ell^{t\,T} y^t$.

- Performance measure :

$$R_{T_r}(\ell, \mathcal{D}) = \sum_{t=1}^{T_r} y^{t\,T} \ell^t - T_r \min_{y \in \Delta_n} \left(\frac{1}{T_r} \sum_{t=1}^{T_r} \ell^t \right)' y$$

empirical average over the observed losses

$$= \sum_{t=1}^{T_r} y^{t\,T} \ell^t - \min_{i \in \{1, \dots, n\}} \sum_{t=1}^{T_r} \ell_i^t$$

→ Strategy "multiplicative weights"

↳ $R_{T_r}(\ell, \mathcal{D}) \leq \sqrt{\frac{\log(n) T_r}{2}}$