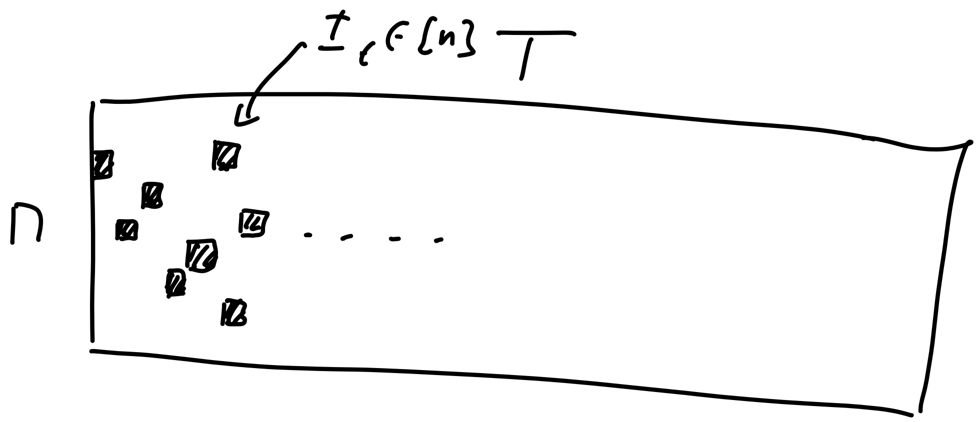
Regret

$$\max_i \sum_{t=1}^{T} X_{i,t} - X_{I_t,t}$$



Stochastic case $\mathbb{E}[X_{i,t}] = \mu_i \quad \forall t$

Adversarial setting: $X_{i,t}$ is arbitrary $\forall i,t$

but chosen <u>before</u> start of game.

<u>Convention</u>: going to use <u>losses</u> instead of rewards.

Adversary chooses losses $\{\ell_t\}_{t=1}^{T} \in [-1,1]^n$

for $t = 1, 2, \ldots, T$

Player chooses $I_t \in [n]$

Gets loss $\ell_{t,I_t} \in [-1,1]$

$$\text{Regret} = \max_i \sum_{t=1}^{T} \ell_{t,I_t} - \ell_{t,i}$$

Suppose $I_t \sim P_t \in \Delta_n$

$$\boxed{\hat{\ell}_{t,i} = \frac{\mathbb{1}\{I_t = i\}}{P_{t,i}} \ell_{t,i}} \qquad \boxed{P_{t,i} > 0 \quad \forall i}$$

$$\mathbb{E}[\hat{\ell}_{t,i} \mid P_t] = \sum_{j=1}^{n} P_{t,j} \frac{\mathbb{1}\{j=i\}}{P_{t,i}} \ell_{t,i} = \ell_{t,i}$$

$$\sum_{t=1}^{T} \hat{\ell}_{t,i}$$

$$\begin{aligned}
\mathbb{V}\mathrm{ar}(\hat{\ell}_{t,i}) &= \mathbb{E}[\hat{\ell}_{t,i}^2] - \mathbb{E}[\hat{\ell}_{t,i}]^2 \\
&= \left(\sum_{j=1}^{n} P_{t,j} \frac{\mathbb{1}\{j=i\}}{P_{t,i}^2} \ell_{t,i}^2\right) - \ell_{t,i}^2 \\
&= \frac{\ell_{t,i}^2}{P_{t,i}} - \ell_{t,i}^2 \\
&= \ell_{t,i}^2 \left(\frac{1}{P_{t,i}} - 1\right)
\end{aligned}$$

$$\mathbb{V}\mathrm{ar}\left(\sum_{t=1}^{T} \hat{\ell}_{t,i}\right) = \sum_{t=1}^{T} \ell_{t,i}^2 \left(\frac{1}{P_{t,i}} - 1\right)$$

EXP3(γ): Exponential Weights for Exploration Exploitation

Input: Time horizon T, n arms, $\eta > 0$, $\gamma \in [0,1]$, $\lambda \in \triangle_n$.

Initialize: Player sets $p_1 = (1/n, \dots, 1/n) \in \triangle_n$. Adversary chooses $\{\ell_t\}_{t=1}^T \subset [-1,1]^n$.

for: $t = 1, \cdots, T$

Player draws $I_t \sim q_t := (1-\gamma)p_t + \gamma\lambda$ and suffers (and observes) loss $\ell(I_t, \ell_t) = \ell_{t,I_t}$

Player computes $\widehat{\ell}_{t,i} = \frac{\mathbf{1}\{I_t = i\}}{q_{t,i}} \ell_{t,i}$

Update iterates:

$$w_{t+1,i} = w_{t,i} \exp(-\eta \widehat{\ell}_{t,i}) \qquad p_{t+1,i} = w_{t+1,i} / \sum_{j=1}^n w_{t+1,j}.$$

Thm) For any loss sequence $\{\ell_t\}_{t=1}^T \in [-1,1]^n$, and $\eta, \gamma \geq 0$ where $-\eta \hat{\ell}_{t,i} \leq 1$ for all i, t then

$$\max_i \mathbb{E}\left[\sum_{t=1}^T \ell_{I_t,t} - \ell_{i,t}\right] \leq 2\gamma T + \frac{\log(n)}{\eta} + (1-\gamma)\eta \, \mathbb{E}\left[\sum_{t=1}^T \sum_{i=1}^n p_{t,i} \hat{\ell}_{t,i}^2\right].$$

Multi-armed bandits:

$$\lambda = \left(\tfrac{1}{n}, \dots, \tfrac{1}{n}\right)$$

$$-\eta \hat{\ell}_{t,i} = -\eta \frac{\mathbb{1}\{I_t = i\}}{q_{t,i}} \ell_{t,i} \leq \eta \frac{1}{\gamma/n} = 1 \quad \text{if } \gamma = \eta n$$

$$\mathbb{E}\left[\sum_{t=1}^T \sum_{i=1}^n p_{t,i} \hat{\ell}_{t,i}^2\right] = \mathbb{E}\left[\sum_{t=1}^T \sum_{i=1}^n p_{t,i} \frac{\mathbb{1}\{I_t = i\}}{q_{t,i}^2} \ell_{t,i}^2\right]$$

$$\leq \frac{1}{(1-\gamma)} \mathbb{E}\left[\sum_{t=1}^T \sum_{i=1}^n \frac{\mathbb{1}\{I_t = i\}}{q_{t,i}} \ell_{t,i}^2\right]$$

$$= \frac{1}{1-\gamma} \sum_{t=1}^T \sum_{i=1}^n \ell_{t,i}^2 \leq \frac{nT}{1-\gamma}.$$

$$\text{Regret} \leq 2\gamma T + \frac{\log(n)}{\varepsilon} + \varepsilon n T$$

$$= 3\varepsilon n T + \frac{\log(n)}{\varepsilon}$$

$$\leq \sqrt{12\, nT \log(n)} \quad w/ \quad \varepsilon = \sqrt{\frac{\log(n)}{3nT}}$$

$$q_{t,i} = (1-\gamma)p_{t,i} + \gamma \frac{1}{n}$$

$$\mathbb{E}\left[\sum_t \ell_{t,I_t} - \ell_{t,i}\right] = \mathbb{E}\left[\sum_{t=1}^{T} \left(\sum_{j=1}^{n} q_{t,j}\ell_{t,j}\right) - \ell_{t,i}\right]$$

$$\leq \gamma T + \mathbb{E}\left[\sum_{t=1}^{T}\left((1-\gamma)\sum_{j=1}^{n} p_{t,j}\ell_{t,j}\right) - \ell_{t,i}\right]$$

$$\leq 2\gamma T + (1-\gamma)\,\mathbb{E}\left[\sum_{t=1}^{T}\left(\sum_{j=1}^{n} p_{t,j}\ell_{t,j}\right) - \ell_{t,i}\right]$$

$$= 2\gamma T + (1-\gamma)\,\mathbb{E}\left[\sum_{t=1}^{T}\left(\sum_{j=1}^{n} p_{t,j}\hat{\ell}_{t,j}\right) - \hat{\ell}_{t,i}\right]$$

$$\ell_{t,i} = (1-\gamma)\ell_{t,i} + \gamma \ell_{t,i}$$

$$W_t = \sum_{i=1}^{n} w_{t,i}$$

$$\log\left(\frac{W_{T+1}}{W_1}\right) \geq -\log(n) + \log(w_{T+1,i})$$

$$= -\log(n) + \log\left(\prod_{t=1}^{T} \exp(-\eta \hat{\ell}_{t,i})\right)$$

$$= -\log(n) - \eta \sum_{t=1}^{T} \hat{\ell}_{t,i}$$

$$\log\left(\frac{W_{T+1}}{W_1}\right) = \sum_{t=1}^{T} \log\left(\frac{W_{t+1}}{W_t}\right)$$

$$= \sum_{t=1}^{T} \log\left(\sum_{i=1}^{n} \frac{w_{t,i} \exp(-\eta \hat{\ell}_{t,i})}{W_t}\right)$$

$$= \sum_{t=1}^{T} \log\left(\sum_{i=1}^{n} p_{t,i} \exp(-\eta \hat{\ell}_{t,i})\right)$$

for $x \leq 1$ $\quad \exp(x) \leq 1 + x + x^2$

$$\leq \sum_{t=1}^{T} \log\left(1 - \eta \sum_{i=1}^{n} p_{t,i} \hat{\ell}_{t,i} + \eta^2 \sum_{i=1}^{n} p_{t,i} \hat{\ell}_{t,i}^2\right)$$

$\forall x,\ 1 + x \leq e^x \quad = \quad x \geq \log(1+x)$

$$\leq -\eta \sum_{t=1}^{T} \sum_{j=1}^{n} p_{t,j} \hat{\ell}_{t,j} + \eta^2 \sum_{t=1}^{T} \sum_{j=1}^{n} p_{t,j} \hat{\ell}_{t,j}^2$$