At each time $t = 1, 2, 3, \ldots$

- Algorithm chooses an action $I_t \in \{1, \ldots, n\}$
- Observes a reward $X_{I_t,t} \sim P_{I_t}$ where $P_1, \ldots, P_n$ are unknown distributions

Regret

$$\mathbb{E}_{X \sim P_i}[X] = \theta_i \in [0,1]$$

$$R_T = \max_{j=1,\ldots,n} \mathbb{E}\left[\sum_{t=1}^{T} X_{j,t} - \sum_{t=1}^{T} X_{I_t,t}\right]$$

$$= \max_{j=1,\ldots,n} \theta_j^* T - \mathbb{E}\left[\sum_{t=1}^{T} X_{I_t,t}\right]$$

$$= \mathbb{E}\left[\sum_{t=1}^{T} \sum_{i=1}^{n} \mathbb{1}\{I_t = i\} X_{i,t}\right]$$

$$= \sum_{i=1}^{n} \theta_i \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}\{I_t = i\}\right]}_{= T_i(t)}$$

$$= \sum_{i=1}^{n} \theta_i \mathbb{E}[T_i(T)]$$

Assume means sorted $\theta_1 \geq \theta_2 \geq \theta_3 \geq \ldots \geq \theta_n$

$$\Delta_i = \theta_1 - \theta_i$$

$$R_T = \theta_1 T - \sum_{i=1}^{n} \theta_i \mathbb{E}[T_i(T)]$$

$$= \sum_{i=1}^{n} \Delta_i \mathbb{E}[T_i(T)]$$

$$= \sum_{i=2}^{n} \Delta_i \mathbb{E}[T_i(T)]$$

$$\mathbb{P}\left(\frac{1}{s}\sum_{t=1}^{s} Z_t \geq \sqrt{\frac{2\sigma^2 \log(1/\delta)}{s}}\right) \leq \delta$$

$$\mathbb{P}\left(\frac{1}{s}\sum_{t=1}^{s} Z_t > \varepsilon\right) \leq \exp\left(-\frac{s\varepsilon^2}{2\sigma^2}\right) = \delta$$

Corollary 1. *Let $Z_1, Z_2, \dots$ be independent mean-zero σ^2-sub-Gaussian random variables so that $\psi_Z(\lambda) := \log(\mathbb{E}[\exp(\lambda Z_t)]) \leq \exp(\lambda^2\sigma^2/2)$, then for $\tau = \lceil 2\sigma^2\epsilon^{-2}\log(1/\delta)\rceil$ we have $\mathbb{P}(\frac{1}{\tau}\sum_{t=1}^{\tau} Z_t \leq \epsilon) \geq 1-\delta$.*

Lemma 1 (Hoeffding's Lemma)**.** *Let X be an independent random variable with support in $[a,b]$ almost surely and $\mathbb{E}[X] = 0$. Then $\log(\mathbb{E}[\exp(\lambda X)]) \leq (b-a)^2\lambda^2/8$.*

Input: n arms $\mathcal{X} = \{1, \dots, n\}$, confidence level $\delta \in (0,1)$.
Let $\mathcal{X}_1 \leftarrow \mathcal{X}, \ell \leftarrow 1$
while $|\mathcal{X}_\ell| > 1$ **do**
  $\epsilon_\ell = 2^{-\ell}$
  Pull each arm in $\mathcal{X}_\ell$ exactly $\tau_\ell = \lceil 2\epsilon_\ell^{-2}\log(\frac{4\ell^2|\mathcal{X}|}{\delta})\rceil$ times
  Compute the empirical mean of these rewards $\widehat{\theta}_{i,\ell}$ for all $i \in \mathcal{X}_\ell$
  $\mathcal{X}_{\ell+1} \leftarrow \mathcal{X}_\ell \setminus \{i \in \mathcal{X}_\ell : \max_{j\in\mathcal{X}_\ell} \widehat{\theta}_{j,\ell} - \widehat{\theta}_{i,\ell} > 2\epsilon_\ell\}$
  $\ell \leftarrow \ell + 1$
Output: $\mathcal{X}_{\ell+1}$ (or play the last arm forever in the regret setting)

$\sigma = 1$

Lemma 2. *Assume that $\max_{i\in\mathcal{X}} \Delta_i \leq 4$. With probability at least $1-\delta$, we have $1 \in \mathcal{X}_\ell$ and $\max_{i\in\mathcal{X}_\ell} \Delta_i \leq 8\epsilon_\ell$ for all $\ell \in \mathbb{N}$.*

Proof. For any $\ell \in \mathbb{N}$ and $i \in [n]$ define

$$\mathcal{E}_{i,\ell} = \left\{|\widehat{\theta}_{i,\ell} - \theta_i^*| \leq \epsilon_\ell\right\}$$

$\mathcal{E} = \bigcap_{i=1}^{n} \bigcap_{\ell=1}^{\infty} \mathcal{E}_{i,\ell}$ ← Want to show holds w.p. $\geq 1-\delta$.

Equiv. Show that $\mathbb{P}(\mathcal{E}^c) \leq \delta$

$\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$

$$\mathbb{P}(\mathcal{E}^c) = \mathbb{P}\left(\bigcup_{i=1}^{n} \bigcup_{\ell=1}^{\infty} \mathcal{E}_{i,\ell}^c\right) \leq \sum_{i=1}^{n}\sum_{\ell=1}^{\infty} \mathbb{P}(\mathcal{E}_{i,\ell}^c)$$

$$\mathbb{P}(\mathcal{E}_{i,\ell}^c) = \mathbb{P}\left(\left|\frac{1}{\tau_\ell}\sum_{t=1}^{\tau_\ell}(X_{i,t} - \theta_i)\right| > \epsilon_\ell\right)$$
$$\leq \mathbb{P}\left(\frac{1}{\tau_\ell}\sum_{t=1}^{\tau_\ell}(X_{i,t} - \theta_i) > \epsilon_\ell\right) + \mathbb{P}\left(\frac{-1}{\tau_\ell}\sum_{t=1}^{\tau_\ell}(X_{i,t} - \theta_i) > \epsilon_\ell\right)$$
$$\leq 2\exp\left(-\frac{\tau_\ell \epsilon_\ell^2}{2}\right)$$

$$= 2\exp\left(-\log\left(\frac{4|\mathcal{X}|\ell^2}{\delta}\right)\right)$$

$$= \frac{\delta}{2|\mathcal{X}|\ell^2}$$

$|\mathcal{X}| = n$

$$\mathbb{P}(\mathcal{E}^c) \leq \sum_{i=1}^{n} \sum_{\ell=1}^{\infty} \frac{\delta}{2|\mathcal{X}|\ell^2}$$

$$= \sum_{\ell=1}^{\infty} \frac{\delta}{2\ell^2}$$

$$\leq \delta.$$

Show arm 1 never gets discarded.

$$\equiv \quad \max_{j \in \mathcal{X}_\ell} \hat{\theta}_{j,\ell} - \hat{\theta}_{1,\ell} \leq 2\varepsilon_\ell \quad \forall \ell$$

$$\max_{j \in \mathcal{X}_\ell} \hat{\theta}_{j,\ell} - \hat{\theta}_{1,\ell} = \max_{j} \hat{\theta}_{j,\ell} - \theta_j - \hat{\theta}_{1,\ell} + \theta_1 + \theta_j - \theta_1$$

$$\leq \max_{j \in \mathcal{X}_\ell} \theta_j - \theta_1 + 2\varepsilon_\ell$$

$$\leq 2\varepsilon_\ell.$$

Fix some arm $i \in [n]$. When is it discarded?

Choose i s.t. $\Delta_i := \theta_1 - \theta_i > 4\varepsilon_\ell$

$$\max_{j \in \mathcal{X}_\ell} \hat{\theta}_{j,\ell} - \hat{\theta}_{i,\ell} \geq \hat{\theta}_{1,\ell} - \hat{\theta}_{i,\ell}$$

$$\geq \theta_1 - \theta_i - 2\varepsilon_\ell$$

$$> 2\varepsilon_\ell$$

$$\max_{j \in \mathcal{X}_{\ell+1}} \Delta_j \leq 4\varepsilon_\ell = 8\varepsilon_{\ell+1}$$

$\Rightarrow$

$a \vee b = \max\{a, b\}$

$\Delta_i \vee \nu \geq \nu$

Theorem W.P. $\geq 1-\delta$

$$R_T = \sum_{i=1}^{n} \Delta_i T_i \leq \underbrace{\inf_{\nu \geq 0} \nu T + \sum_{i=2}^{n} c (\Delta_i \vee \nu)^{-1} \log\left(\frac{\log((\Delta_i \vee \nu)^{-1}) |\mathcal{X}|}{\delta}\right)}_{r_{T,\delta}}$$

$$\leq \inf_{\nu \geq 0} \nu T + c\, n \nu^{-1} \log\left(\frac{\log(\nu^{-1}) |\mathcal{X}|}{\delta}\right)$$

$$\leq c' \sqrt{nT \log\left(\frac{\log(nT)\, n}{\delta}\right)}$$

$$\mathbb{E}[R_T] = \mathbb{E}[R_T \mathbb{1}\{\mathcal{E}\} + R_T \mathbb{1}\{\mathcal{E}^c\}]$$

$$\leq \mathbb{E}[r_{T,\delta} \mathbb{1}\{\mathcal{E}\}] + \mathbb{E}[R_T \mathbb{1}\{\mathcal{E}^c\}]$$

$$\leq r_{T,\delta} + T \cdot \mathbb{E}[\mathbb{1}\{\mathcal{E}^c\}]$$

$\Delta_i \in (0, 1]$

$$\le r_{T,\delta} + \delta T \qquad \delta = 1/T$$

$$\le 1 + r_{T,\delta}$$

Proof

$$R_T = \sum_{i=1}^{n} \Delta_i T_i$$

$$= \sum_{i:\Delta_i \le \nu} \Delta_i T_i + \sum_{i:\Delta_i > \nu} \Delta_i T_i$$

$$\le \nu T + \sum_{i:\Delta_i > \nu} \Delta_i T_i$$

$$= \nu T + \sum_{i:\Delta_i > \nu} \sum_{\ell=1}^{\infty} \Delta_i \tau_\ell \mathbb{1}\{i \in \mathcal{X}_\ell\}$$

$$\le \nu T + \sum_{i:\Delta_i > \nu} \sum_{\ell=1}^{\infty} \Delta_i \tau_\ell \mathbb{1}\{\Delta_i \le 8\varepsilon_\ell\}$$

$$\le \nu T + \sum_{i=2}^{n} \sum_{\ell=1}^{\infty} \Delta_i \tau_\ell \mathbb{1}\{\underbrace{\Delta_i \vee \nu \le 8\varepsilon_\ell}\} \qquad \varepsilon_\ell = 2^{-\ell}$$

$$\equiv \ell \le \log_2(8(\Delta_i \vee \nu)^{-1})$$

$$\le \nu T + \sum_{i=2}^{n} \sum_{\ell=1}^{\lceil \log_2 8(\nu \vee \Delta_i)^{-1} \rceil} 8\varepsilon_\ell \tau_\ell \qquad \sum_{\ell=0}^{k} a^\ell = \frac{1-a^{k+1}}{1-a}$$

$$= \nu T + \sum_{i=2}^{n} \sum_{\ell=1}^{\lceil \cdot \rceil} c\, \varepsilon_\ell^{-1} \log\left(\frac{|\mathcal{X}|\ell^2}{\delta}\right)$$

$$\le \nu T + \sum_{i=2}^{n} c \log\left(\frac{|\mathcal{X}| \log_2(8(\nu \vee \Delta_i)^{-1})^2}{\delta}\right) \sum_{\ell=1}^{\lceil \log_2 8(\Delta_i \vee \nu)^{-1} \rceil} 2^\ell$$

$$\le \nu T + \sum_{i=2}^{n} c'(\Delta_i \vee \nu)^{-1} \log\left(\frac{|\mathcal{X}| \log(\Delta_i \vee \nu)}{\delta}\right)$$

∎

Roughly $\quad \mathbb{E}[R_T] \leq \min\left\{\sqrt{nT\log T},\ \sum_{i>1} \Delta_i^{-1}\log(T)\right\}.$

<u>Theorem</u>

It can be shown $\exists c>0$ such that

$$\inf_{Alg} \max_{\theta\in[0,1]^n} \mathbb{E}[R_T] \geq c\sqrt{nT}.$$

<u>Theorem</u>

For any alg satisfying $\max_{\theta\in[0,1]^n} \mathbb{E}[R^T] \leq c' T^{\alpha}$ for $\alpha<1$

$$\inf_{T\to\infty} \frac{\mathbb{E}[R_T]}{\log(T)} \geq \sum_{i>1} \Delta_i^{-1}.$$

n arms

each arm i is associated w/ score $\theta_i \in \mathbb{R}$

$$\mathbb{P}(i > j) = \frac{\exp(\theta_i)}{\exp(\theta_i) + \exp(\theta_j)} = \frac{1}{1 + \exp(\theta_j - \theta_i)}$$

Bradly
Terry
Luce

$$S_i = \frac{1}{n-1} \sum_{j \neq i}^{n} \mathbb{P}(i > j)$$

Borda winner

$$i_A = \operatorname{argmax}_i S_i$$

Copeland / Condorcet

$$i_A = \operatorname{argmax}_i \sum_{j=1}^{n} \mathbb{1}\{P_{ij} > \tfrac{1}{2}\}$$
