- MAB (Stochastic)
- IID CASE
- LOWER BOUND ON REGRET

Result. Consider a uniformly good rule.
Let the configuration be $C = (\theta_1, \dots, \theta_K)$

For all $\varepsilon > 0$, for all $j \notin \{\sigma(1), \dots, \sigma(\ell)\}$,

$$\lim_{t \to \infty} P_C \left[T_t(j) \geq \frac{(1-\varepsilon)\log t}{I(\theta_j, \theta_{\sigma(1)})} \right] = 1.$$

As a consequence:

$$\liminf_{t \to \infty} \frac{R(t,C)}{\log t} \geq \sum_{j \notin \{\sigma(1), \dots, \sigma(\ell)\}} \frac{\mu(\theta_{\sigma(1)}) - \mu(\theta_j)}{I(\theta_j, \theta_{\sigma(1)})}$$

Proof. $\boxed{\text{Idea}}$ change of "measure" or of configuration.

New configuration $C^* = (\theta_1, \dots, \theta_{j-1}, \lambda, \theta_{j+1}, \dots, \theta_K)$.

λ is chosen so that:

1. The policy selects arm j very often : $\lambda > \theta_{\sigma(1)}$ (the policy is uniformly good)
2. λ is close to $\theta_{\sigma(1)}$. More precisely, for $\rho > 0$ we choose λ :

$$\left| I(\theta_j, \lambda) - I(\theta_j, \theta_{\sigma(1)}) \right| \leq \rho \, I(\theta_j, \theta_{\sigma(1)})$$

Now we relate $P_C[T_t(j) > \beta]$ to $P_{C^*}[T_t(j) > \beta]$

We will show that : for j suboptimal under C,

$$P_C\left[T_t(j) < \frac{\log t}{(1+2\varrho)\, I(\theta_j,\lambda)}\right] \xrightarrow{t\to\infty} 0$$

(we then conclude by observing that

$$I(\theta_j,\lambda) \in \left[(1-\varrho)\, I(\theta_j,\theta_{\sigma(1)}),\ (1+\varrho)\, I(\theta_j,\theta_{\sigma(1)})\right]$$

and by letting ϱ tend to 0^+).

$\boxed{\text{Change of measure}}$ from P_C to P_{C^*}.

Let $j \notin \{\sigma(1),\dots,\sigma(\ell)\}$.

Denote by $Y_1, Y_2, \dots,$ its successive rewards.

Remember that : Y_t density $f(x,\theta_j)\, d\nu(x)$.

Hence :

$$\mathbb{E}_C\left[g(Y_1,\dots,Y_t)\right] = \int\!\cdots\!\int g(y_1,\dots,y_t) \underbrace{\prod_{a=1}^{t} f(y_a,\theta_j)\, d\nu(y_a)}_{dP_C(y_1,\dots,y_t)}$$

$$= \int\!\cdots\!\int \left[g(y_1,\dots,y_t)\, \frac{\prod_{a=1}^{t} f(y_a,\theta_j)}{\prod_{a=1}^{t} f(y_a,\lambda)}\right] dP_{C^*}(y_1,\dots,y_t)$$

$$= \mathbb{E}_{C^*}\left[g(Y_1,\dots,Y_t)\ \underbrace{\frac{\prod_{a=1}^{t} f(Y_a,\theta_j)}{\prod_{a=1}^{t} f(Y_a,\lambda)}}_{\text{"Likelyhood ratio"}}\right]$$

$\boxed{\text{Under } P_{C^*}}$ j is the best arm, so it is played very often, all other arms played very rarely

$$\mathbb{E}_{C^*}\left[t - T_t(j)\right] = o(t^\alpha), \quad \forall \alpha > 0.$$

$$\Rightarrow (t - K\log t)\, P_{C^*}\left[T_t(j) < K\log t\right] = o(t^\alpha), \quad \forall \alpha > 0$$

$$\Rightarrow P_{C^*}\left(T_t(j) < K\log t\right) = o(t^{\alpha-1}). \quad \forall \alpha > 0.$$

$\boxed{\text{Likelyhood ratio}}$

let $L_t = \sum_{a=1}^{t} \log \frac{f(Y_a, \theta_j)}{f(Y_a, \lambda)}$

Law of large numbers: $\frac{L_t}{t} \xrightarrow[t\to\infty]{} I(\theta_j, \lambda)$ a.s. - P_C

$$\Rightarrow \frac{1}{t} \max_{a \le t} L_a \xrightarrow{t\to\infty} I(\theta_j, \lambda) \quad P_C\text{- a.s.}$$

$$\Rightarrow \lim_{t\to\infty} P_C\left[L_a > K(1+e)I(\theta_j, \lambda)\log t, \text{ for } a < K\log t\right] = 0.$$

Now

$$\{T_t(j) < K\log t,\ L_{T_t(j)} \le K(1+e)I(\theta_j,\lambda)\log t\}$$

$$= \bigcup_{a < K\log t} \{T_t(j) = a,\ L_a \le K(1+e)I(\theta_j,\lambda)\log t\}.$$

$$P_{C^*}\left[T_t(j) = a,\ L_a \le K(1+e)I(\theta_j,\lambda)\log t\right]$$

$$= \int_{\{T_t(j)=a,\ L_a \le K(1+e)I(\theta_j,\lambda)\log t\}} e^{-L_a}\, dP_C$$

$$\geq t^{-K(1+e)I(\theta_j,\lambda)} \cdot P_C\left[T_t(j) = a,\ L_a \le K(1+e)I(\theta_j,\lambda)\log t\right]$$

Take $K = \dfrac{1}{(1+2e)\, I(\theta_j, \partial)}$

- We have :

(1) $\displaystyle \lim_{t \to \infty} P_c\left[L_a > \frac{1+e}{1+2e} \log t \text{, for some } a < \frac{\log t}{(1+2e)\, I(\theta_j, \partial)} \right] = 0$

(2) $\displaystyle P_c\left[T_t(j) < \frac{\log t}{(1+2e)\, I(\theta_j, \partial)} ,\ L_{T_t(j)} \leq \frac{1+e}{1+2e} \log t \right]$

$$\leq t^{\frac{1+e}{1+2e}} \underbrace{P_{C^*}\left[T_t(j) < \frac{\log t}{(1+2e)\, I(\theta_j, \partial)} ,\ L_{T_t(j)} \leq \dots \right]}_{= o(t^{\alpha - 1}),\ \forall \alpha > 0}$$

We conclude that :

$$P_c\left[T_t(j) < \frac{\log t}{(1+2e)\, I(\theta_j, \partial)} \right] \xrightarrow{t \to \infty} 0 .$$

$e \to 0$ + choice of ∂ : $\displaystyle P_c\left[T_t(j) < \frac{\log t\,(1-\varepsilon)}{I(\theta_j, \partial)} \right] \xrightarrow{t \to \infty} 0$

$\boxed{\text{Regret bound}}$:

$$\mathbb{E}_c\left[T_t(j) \right] \geqslant \frac{(1-\varepsilon) \log t}{I(\theta_j, \theta_{\sigma(1)})} \cdot P_c\left[T_t(j) \geqslant \frac{(1-\varepsilon)\log t}{I(\theta_j, \theta_{\sigma(1)})} \right] \quad , \forall \varepsilon > 0$$

$$\Rightarrow \liminf_{t \to \infty} \mathbb{E}_c\left[\frac{T_t(j)}{\log t} \right] \geqslant \frac{1}{I(\theta_j, \theta_{\sigma(1)})}$$