

Solution for Exercise on UCB:

$$a, b \geq 0: \min\{a, b\} \leq \sqrt{ab}$$

$$\bullet \quad \mathbb{E}[N_i(T)] \leq \min\left\{T, \frac{8\ln T}{\Delta_i^2} + 2\right\} \leq \sqrt{T\left(\frac{8\ln T}{\Delta_i^2} + 2\right)}$$

thus
$$\bar{R}_T = \sum_{i: \Delta_i > 0} \Delta_i \mathbb{E}[N_i(T)] \leq \sum_{i: \Delta_i > 0} \sqrt{T(8\ln T + 2\Delta_i^2)} \leq O(K\sqrt{T\ln T})$$

Or a more direct approach:

$$\bar{R}_T = \sum_{i: \Delta_i > \sqrt{\frac{8\ln T}{T}}} \underbrace{\left(2 + \frac{8\ln T}{\Delta_i^2}\right)}_{< 2 + \sqrt{8\ln T}} \Delta_i + \sum_{\substack{i: \Delta_i \leq \sqrt{\frac{8\ln T}{T}} \\ \text{and } \Delta_i > 0}} \Delta_i T \leq K(2 + \sqrt{8\ln T}) = O(K\sqrt{T\ln T})$$

- Where did we fail? We used that $\forall i, \mathbb{E}[N_i(T)] \leq T$ but in fact, a stronger statement holds:

$$\sum_i \mathbb{E}[N_i(T)] = T$$

- The smarter approach is:

$$\begin{aligned} \bar{R}_T &= \sum_{i: \Delta_i > 0} \Delta_i \mathbb{E}[N_i(T)] \\ &\leq \sum_{i: \Delta_i > 0} \Delta_i \min\left\{\mathbb{E}[N_i(T)], \frac{8\ln T}{\Delta_i^2} + 2\right\} && \text{by the Proposition} \\ &\leq \sum_{i: \Delta_i > 0} \sqrt{\mathbb{E}[N_i(T)] \left(\frac{8\ln T}{\Delta_i^2} + 2\right)} && \min\{a, b\} \leq \sqrt{ab} \\ &\leq \sqrt{8\ln T + 2} \sum_{i=1, \dots, K} \sqrt{\mathbb{E}[N_i(T)]} && \sqrt{\cdot} \text{ is concave:} \\ &\leq \sqrt{8\ln T + 2} \sqrt{K \underbrace{\sum_{i=1}^K \mathbb{E}[N_i(T)]}_{=T}} && \text{for } u_1, \dots, u_K \geq 0, \\ &= \sqrt{KT(8\ln T + 2)} && \frac{1}{K} \sum_j \sqrt{u_j} \leq \sqrt{\frac{1}{K} \sum_j u_j} \end{aligned}$$

Exercises with randomized prediction: 1/2

We call "union bound" the fact that $\mathbb{P}\left(\bigcup_{n \geq 1} A_n\right) \leq \sum_{n \geq 1} \mathbb{P}(A_n)$

Choosing $S_T = S/T(T+1)$ for $T \geq 1$,

$$\text{we have } \mathbb{P}\left\{R_T > (M-m)\sqrt{T}\left(\sqrt{\ln N} + \sqrt{\frac{1}{2} \ln \frac{T(T+1)}{S}}\right) + (M-m)\left(2 + \frac{4}{3} \ln N\right)\right\} \stackrel{\text{def}}{=} p(T, S) \leq \frac{S}{T(T+1)}$$

$$\begin{aligned} \text{So that } \mathbb{P}\left\{\exists T \geq 1 \mid R_T > p(T, S)\right\} &\leq S \sum_{T \geq 1} \frac{1}{T(T+1)} \\ &= S \sum_{T \geq 1} \left(\frac{1}{T} - \frac{1}{T+1}\right) = S \end{aligned}$$

That is: $\forall \delta \in (0, 1)$,

with probability at least $1 - \delta$: $\left[\forall T \geq 1, R_T \leq p(T, S) \right]$

$$\text{where } p(T, S) = (M-m)\sqrt{T}\left(\sqrt{\ln N} + \sqrt{\frac{1}{2} \ln \frac{T(T+1)}{S}}\right) + (M-m)\left(2 + \frac{4}{3} \ln N\right).$$

Note: with the techniques of the next exercise, we could find a refined $p(T, S)$ of order $(M-m)\sqrt{T \ln\left(\frac{NT}{S}\right)}$ instead of the $O((M-m)\sqrt{T \ln N})$ we exhibited.

Exercises with randomized prediction: 2/2

(1) Recall that given a filtration $(\mathcal{F}_t)_{t \geq 0}$, and given an adapted process $(S_t)_{t \geq 0}$, we say that:

- $(S_t)_{t \geq 0}$ is a martingale when $\forall 0 \leq t \leq T, \quad X_t = E[X_T | \mathcal{F}_t]$
- $(S_t)_{t \geq 0}$ is a submartingale when $X_t \leq E[X_T | \mathcal{F}_t]$
- $(S_t)_{t \geq 0}$ is a supermartingale when $X_t \geq E[X_T | \mathcal{F}_t]$

By the conditional Jensen's inequality, a convex function of a martingale is a submartingale.

Ex: if $(S_t)_{t \geq 0}$ is a martingale then $(|S_t|)_{t \geq 0}$ and $(\exp(s S_t))_{t \geq 0}$ are submartingales, for all $s \in \mathbb{R}$.

Dob's maximal inequality for non-negative submartingale $(S_t)_{t \geq 0}$:

$$\forall T \geq 0, \quad \forall c > 0, \quad \mathbb{P}\left\{\sup_{0 \leq t \leq T} S_t \geq c\right\} \leq \frac{E[S_T]}{c}$$

A not-so-famous version for non-negative supermartingals $(S_t)_{t \geq 0}$ exists:

$$\forall c > 0, \quad \mathbb{P}\left\{\sup_{t \geq 0} S_t \geq c\right\} \leq \frac{E[S_0]}{c}$$

(2) With the notation of the proof given in class:

$$(S_t)_{t \geq 0} \quad \text{where} \quad S_t = \sum_{G=1}^t X_G - \sum_{G=1}^t E[X_G | \mathcal{F}_{G-1}]$$

is a martingale $(S_0 = 0)$

so that $\forall s \in \mathbb{R}, \quad (e^{s S_t})_{t \geq 0}$ is a non-negative submartingale.

We proved in class (by induction) that $E[e^{s S_T}] \leq \exp\left(\frac{s^2}{8} \sum_{t=1}^T (b_t - a_t)^2\right)$

By Hoeffding - Chernoff:

$$\begin{aligned} \mathbb{P}\left\{\sup_{0 \leq t \leq T} S_t \geq \varepsilon\right\} &= \mathbb{P}\left\{\sup_{0 \leq t \leq T} e^{\lambda S_t} \geq e^{\lambda \varepsilon}\right\} \\ &\stackrel{\text{Doob's maximal inequality}}{\leq} e^{-\lambda \varepsilon} \mathbb{E}[e^{\lambda S_T}] \\ &\leq \exp\left(-\lambda \varepsilon + \frac{\lambda^2}{8} \sum_{t=1}^T (b_t - a_t)^2\right) \\ &= \exp\left(-2\varepsilon^2 / \sum_{t=1}^T (b_t - a_t)^2\right) \end{aligned}$$

for the same $\lambda = \lambda^*$ as in the original proof

Hence the claimed bound by picking

$$\varepsilon = \sqrt{\frac{\sum_{t=1}^T (b_t - a_t)^2}{2} \ln \frac{1}{\delta}}$$

(3) We decompose the regret as:

$$R_T = \sum_{t=1}^T l_{j_t, t} - \min_k \sum_{t=1}^T l_{k, t} = \underbrace{\sum_{t=1}^T l_{j_t, t} - \sum_{t=1}^T \sum_j p_j l_{j, t}}_{= S_T} + \bar{R}_T$$

$\bar{R}_T = O(\sqrt{T \ln N})$ by assumption

We have $\limsup \frac{R_T}{(M-m)\sqrt{T \ln(\ln T)}}$

so that $\limsup \frac{\bar{R}_T}{\sqrt{T \ln(\ln T)}} \leq 0$

$$\leq \limsup_{T \rightarrow \infty} \frac{S_T}{(M-m)\sqrt{T \ln(\ln T)}}$$

controlling this is a purely probabilistic task
* but * we will recycle some ideas seen in class when studying the doubling trick.

We divide S_T in blocks:

$$r \geq 1, \quad \Delta_r \stackrel{\text{def.}}{=} \max_{t \in [2^{r-1}, 2^r]} \sum_{t=2^{r-1}}^t (l_{j_t, t} - \sum_j p_j l_{j, t})$$

$$S_T \leq \underbrace{\sum_{t=1}^T (\ell_{j,t} - \sum_j p_{j,t} \ell_{j,t})}_{\leq 2(M-m)} + \sum_{r=1}^{\lceil \ln T / \ln 2 \rceil - 1} \Delta_r$$

By (2), we have $\mathbb{P}\{\Delta_r > (M-m) \sqrt{\frac{2^r}{2} \ln \frac{1}{\delta_r}}\} \leq \delta_r \quad \forall r \geq 1$

Picking $\delta_r = 1/r^2$ and applying the Borel-Cantelli lemma:

The random variable $R = \max\{r \geq 1 : \Delta_r > (M-m) \sqrt{2^r \ln r}\}$ is such that $R < +\infty$ a.s.

$$\text{Thus, } S_T \leq 2(M-m) + \underbrace{\sum_{r=1}^R 2^r (M-m)}_{\text{trivial bound on } \Delta_r} + \sum_{r=R+1}^{\lceil \ln T / \ln 2 \rceil - 1} \underbrace{(M-m) \sqrt{2^r \ln r}}_{\substack{\text{for } r \geq R+1, \\ \text{we have, by definition of } R, \\ \Delta_r \leq (M-m) \sqrt{2^r \ln r}}}$$

$$S_T \leq \underbrace{(M-m) \left(2^{R+1} - 1 \right)}_{\substack{\uparrow \\ \sum_{r=0}^R 2^r \\ \text{this is} \\ < +\infty \text{ a.s.}}} + (M-m) \underbrace{\sum_{r=0}^{\lceil \ln T / \ln 2 \rceil - 1} (\sqrt{2})^r}_{= \frac{(\sqrt{2})^{\lceil \ln T / \ln 2 \rceil} - 1}{\sqrt{2} - 1}} \times \underbrace{\sqrt{\ln(\lceil \ln T / \ln 2 \rceil - 1)}}_{\sim \sqrt{\ln(\ln T)}}$$

$$\text{and } \limsup_{T \rightarrow +\infty} \frac{2^{R+1}}{\sqrt{T \ln(\ln T)}} = 0$$

$$\begin{aligned} \text{where } (\sqrt{2})^{\lceil \ln T / \ln 2 \rceil} &\leq \sqrt{2}^{1 + \ln T / \ln 2} \\ &= \exp\left(\left(\frac{1}{2} \ln 2\right) \times \left(1 + \frac{\ln T}{\ln 2}\right)\right) \\ &= \exp\left(\frac{1}{2} \ln(2T)\right) = \sqrt{2T} \end{aligned}$$

All in all:

$$\limsup_{T \rightarrow +\infty} \frac{S_T}{(M-m) \sqrt{T \ln(\ln T)}} \leq \frac{\sqrt{2}}{\sqrt{2}-1} \quad \text{a.s.}$$

Which entails the desired result, with $C = \frac{\sqrt{2}}{\sqrt{2}-1}$.

(4) * We took regimes of the form $[2^r+1, 2^{r+1}]$

By taking regimes of successive lengths $\lceil a^r \rceil$

for some $a > 1$, and $\delta_r = \frac{1}{r(\ln r)^2}$ for Borel-Cantelli

$$\limsup_{T \rightarrow +\infty} \frac{\sum_{r=0}^{r'(T)} \sqrt{\frac{a^r}{2}}}{\sqrt{T}} = \frac{1}{\sqrt{2}(\sqrt{a}-1)} \limsup_{T \rightarrow +\infty} \frac{(\sqrt{a})^{r'(T)+1}}{\sqrt{T}}$$

where $r'(T)$ is the smallest $r \geq 1$ such that $T \leq \sum_{r=0}^{r'} \lceil a^r \rceil$

In particular,

$$\sum_{r=0}^{r'(T)-1} \lceil a^r \rceil < T$$

$$\geq \sum_{r=0}^{r'(T)-1} a^r = \frac{a^{r'(T)} - 1}{a - 1}$$

thus:

$$a^{r'(T)} \leq (a-1)T + 1$$

and $(\sqrt{a})^{r'(T)+1} \leq \sqrt{a} \sqrt{a-1} \sqrt{T} + 1$

Finally we get with these regimes:

$$\limsup_T \frac{\sum_{r=0}^{r'(T)} \sqrt{\frac{a^r}{2}}}{\sqrt{T}} \leq \frac{\sqrt{a} \sqrt{a-1}}{\sqrt{2}(\sqrt{a}-1)}$$

denote this C_a

Note For $a=2$, we get $C_2 = \frac{1}{\sqrt{2}-1}$, which is a $\sqrt{2}$ improvement to what we did in (3), due to a better choice of δ_r :

in (3): with $\delta_r = 1/2$: $\ln 1/8^r = 2 \ln r \rightarrow$ additional $\sqrt{2}$ factor

here: $\delta_r = 1/r(\ln r)^2$: $\ln 1/8^r = \ln r + 2 \ln(\ln r)$

Which is the best $a > 1$?

I think it's around $a \approx 2.5$ and it yields a constant of ≈ 2.35

* Let's compare what we get to the law of iterated logarithm:

Let Z_1, Z_2, \dots be iid random variables, such that $E Z_1^2 < +\infty$

Then, denoting $\mu = E Z_1$ and $\sigma^2 = \text{Var } Z_1$, we have

$$\limsup_{T \rightarrow +\infty} \frac{\sum_{t=1}^T (Z_t - \mu)}{\sqrt{T \ln(\ln T)}} = \sigma \sqrt{2} \quad \text{a.s.}$$

Our argument dealt with martingales and can be applied to $\sum_{t=1}^T (Z_t - \mu)$:
Assuming $Z_t \in [m, M]$ as we got by Hoeffding-Azuma + Borel-Cantelli + regimes of size a^t :

$$\limsup_{T \rightarrow +\infty} \frac{\sum_{t=1}^T (Z_t - \mu)}{\sqrt{T \ln(\ln T)}} \leq (M-m) C_1 = (M-m) \frac{\sqrt{a} \sqrt{a-1}}{\sqrt{2}(\sqrt{a}-1)}$$

Are there cases when $\sigma \sqrt{2} = (M-m) \frac{\sqrt{a} \sqrt{a-1}}{\sqrt{2}(\sqrt{a}-1)}$?

We know that $\sigma \leq \frac{M-m}{2}$ (see the proof of Hoeffding's inequality, subgaussian formula for the variance)

↳ Are there cases when

$$\frac{M-m}{2} \sqrt{2} \stackrel{?}{=} (M-m) \frac{\sqrt{a}(\sqrt{a}-1)}{\sqrt{2}(\sqrt{a}-1)}$$

$$\Leftrightarrow \underbrace{\frac{\sqrt{a}(\sqrt{a}-1)}{\sqrt{a}-1}}_{\text{always larger than } \approx 3.33} \stackrel{?}{=} 1$$

There is room for improvement as for the numerical constant is concerned. ~> Any idea?

The $\sqrt{T \ln(\ln T)}$ rate is optimal \hookrightarrow it seems intuitive, ... but

To be complete, we should show that

for all strategies, as, $\liminf_{T \rightarrow +\infty} \frac{R_T}{\sqrt{T \ln(\ln T)}} > 0$

by showing that for all strategies,

$$\forall j \in [0,1] \quad \liminf_{T \rightarrow +\infty} \frac{\sum_{t=1}^T (Z_{j,t} - \sum_j p_j j_t)}{\sqrt{T \ln(\ln T)}} > 0 \quad \text{a.s.}$$

Again: some reward for whoever would write it up!
and again, we did not reach $\approx \text{Ber}(1/2)$