Solution for Exercise on UCB :

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

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

thus

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

Or a more direct approach:

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

- 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 :

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

Exercices wrt randomized prediction: 1/2

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

Choosing $\delta_T = \delta / T(T+1)$ for $T \geq 1$,

we have
$$\mathbb{P}\Big\{ R_T > \underbrace{(M-m)\sqrt{T}\Big(\sqrt{\ln N} + \sqrt{\tfrac{1}{2}\ln\tfrac{T(T+1)}{\delta}}\Big) + (M-m)\Big(2+\tfrac{4}{3}\ln N\Big)}_{\overset{\text{def}}{=}\rho(T,\delta)} \Big\} \leq \frac{\delta}{T(T+1)}$$

So that
$$\mathbb{P}\Big\{ \exists T\geq 1 \;\big|\; R_T > \rho(T,\delta) \Big\} \leq \delta \sum_{T\geq 1} \frac{1}{T(T+1)} = \delta \sum_{T\geq 1}\Big(\frac{1}{T} - \frac{1}{T+1}\Big) = \delta$$

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

with probability at least $1-\delta$: $\Big[\, \forall T \geq 1, \;\; R_T \leq \rho(T,\delta) \,\Big]$

where
$$\rho(T,\delta) = (M-m)\sqrt{T}\Big(\sqrt{\ln N} + \sqrt{\tfrac{1}{2}\ln\tfrac{T(T+1)}{\delta}}\Big) + (M-m)\Big(2+\tfrac{4}{3}\ln N\Big).$$

Note: with the techniques of the next exercise, we could find a refined $\rho(T,\delta)$ of order $(M-m)\sqrt{T\ln\big(\frac{\ln T}{\delta}\big)}$ instead of the $\mathcal{O}\big((M-m)\sqrt{T\ln T/\delta}\big)$ we exhibited.

Exercises wrt randomized prediction: 2/2

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

- $(S_t)_{t \geqslant 0}$ is a martingale when $\forall\, 0 \leqslant t \leqslant T$, $X_t = \mathbb{E}[X_T \mid \mathcal{F}_t]$
- $(S_t)_{t \geqslant 0}$ is a submartingale when $X_t \leqslant \mathbb{E}[X_T \mid \mathcal{F}_t]$
- $(S_t)_{t \geqslant 0}$ is a supermartingale when $X_t \geqslant \mathbb{E}[X_T \mid \mathcal{F}_t]$

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

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

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

$$\forall T \geqslant 0, \quad \forall c > 0, \qquad \mathbb{P}\Big\{ \sup_{0 \leqslant t \leqslant T} S_t \geqslant c \Big\} \leqslant \frac{\mathbb{E}[S_T]}{c}$$

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

$$\forall c > 0, \qquad \mathbb{P}\Big\{ \sup_{t \geqslant 0} S_t \geqslant c \Big\} \leqslant \frac{\mathbb{E}[S_0]}{c}$$

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

$(S_t)_{t \geqslant 0}$ where $S_t = \sum_{\tau=1}^{t} X_\tau - \sum_{\tau=1}^{t} \mathbb{E}[X_\tau \mid \mathcal{F}_{\tau-1}]$ is a martingale $(S_0 = 0)$

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

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

By Markov – Chernov :

$\forall s>0 : x \mapsto e^{sx}$ is increasing

$$\mathbb{P}\left\{ \sup_{0\leq t\leq T} S_t \geq \varepsilon \right\} = \mathbb{P}\left\{ \sup_{0\leq t\leq T} e^{sS_T} \geq e^{s\varepsilon} \right\}$$

Doob's maximal inequality

$$\leq e^{-s\varepsilon}\, \mathbb{E}\left[e^{sS_T} \right]$$

$$\leq \exp\left(-s\varepsilon + \frac{s^2}{8} \sum_{t=1}^{T} (b_t - a_t)^2 \right)$$

$$= \exp\left(-2\varepsilon^2 \Big/ \sum_{t\leq T} (b_t - a_t)^2 \right)$$

for the same $s = s^*$ as in the original proof

Hence the claimed bound by picking

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

(3) We decompose the regret as :

$$R_T = \sum_{t=1}^{T} \ell_{J_t,t} - \min_k \sum_{t=1}^{T} \ell_{k,t} = \underbrace{\sum_{t=1}^{T} \ell_{J_t,t} - \sum_{t=1}^{T} \sum_j p_{j,t} \ell_{j,t}}_{= S_T} + \overline{R}_T$$

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

so that

$$\limsup \frac{\overline{R}_T}{\sqrt{T \ln(\ln T)}} \leq 0$$

We have

$$\limsup \frac{R_T}{(M-m)\sqrt{T \ln(\ln T)}} \leq \limsup_{T\to+\infty} \frac{S_T}{(M-m)\sqrt{T \ln(\ln T)}}$$

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 :

$$r \geq 1, \qquad \Delta_r \stackrel{\text{def.}}{=} \max_{t \in [2^r+1,\, 2^{r+1}]} \sum_{\tau = 2^r+1}^{t} \left(\ell_{J_t,t} - \sum_j p_{j,t} \ell_{j,t} \right)$$

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

By (2), we have $\mathbb{P}\left\{ \Delta_r > (M-m)\sqrt{\frac{2^r}{2}\ln\frac{1}{\delta_r}} \right\} \leqslant \delta_r \qquad \forall r \geqslant 1$

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

The random variable $R = \max\left\{ r \geqslant 1 : \ \Delta_r > (M-m)\sqrt{2^r \ln r} \right\}$

is such that $R < +\infty$ a.s.

Thus, $$S_T \leqslant 2(M-m) + \sum_{r=1}^{R} \underbrace{2^r (M-m)}_{\text{trivial bound on } \Delta_r} + \sum_{r=R+1}^{\lceil \ln T/\ln 2 \rceil - 1} \underbrace{(M-m)\sqrt{2^r \ln r}}$$

for $r \geqslant R+1$, we have, by definition of R, $\Delta_r \leqslant (M-m)\sqrt{2^r \ln r}$

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

cf. $\sum_{r=0}^{R} 2^r$

$$= \frac{\left(\sqrt{2}\right)^{\lceil \ln T/\ln 2 \rceil} - 1}{\sqrt{2} - 1} \qquad \sim \sqrt{\ln(\ln T)}$$

and $\limsup_{T \to +\infty} \dfrac{2^{R+1}}{\sqrt{T \ln(\ln T)}} = 0$

where $\left(\sqrt{2}\right)^{\lceil \ln T/\ln 2 \rceil}$

$$\leqslant \sqrt{2}^{\,1 + \ln T/\ln 2}$$

$$= \exp\left(\left(\tfrac{1}{2}\ln 2\right) \times \left(1 + \frac{\ln T}{\ln 2}\right) \right)$$

$$= \exp\left(\tfrac{1}{2} \ln(2T) \right) = \sqrt{2T}$$

All in all:

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

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

For question (4) I provide two answers:

- My original answer, where I perform a doubling trick with regimes of lengths given by the integer part of a^r instead of 2^r; the constant may be improved but I explain why we still have a gap w.r.t. law of the iterated logarithm
- An answer by Dau Hai Dang (a student who took the course in Spring 2019), where he explains how a modification of the Borel-Cantelli lemma, based on a doubling trick (!), does the job

This all should be some food for thought!

And maybe a clearer summary can be written (also with lower bounds). Please send me your notes if they are worth it!

(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 $\delta_r = \dfrac{1}{r(\ln r)^2}$ for Borel-Cantelli

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

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

In particular,

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

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

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

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

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

Which is the best $a > 1$? I think it's around $a \simeq 2{,}5$ and it yields a constant of $\simeq 2{,}35$

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

Let $Z_1, Z_2, \ldots$ be iid random variables, such that $\mathbb{E} Z_1^2 < +\infty$

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

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

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

$$\limsup_{T\to+\infty} \frac{\sum_{t=1}^{T} (Z_t - \mu)}{\sqrt{T \ln(\ln T)}} \leqslant (M-m)\, C_a = (M-m) \frac{\sqrt{a}\,\sqrt{a-1}}{\sqrt{2}\,(\sqrt{a}-1)}$$

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

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

↳ Are there cases when

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

$$\iff \underbrace{\frac{\sqrt{a}\,(\sqrt{a-1})}{\sqrt{a}-1}}_{\text{always larger than } \simeq 3{,}33} \overset{?}{=} 1$$

There is room for improvement as far as the numerical constant is concerned. ⇝ Any idea?

The $\sqrt{T \ln(\ln T)}$ rate is optimal

↳ it seems intuitive, … but

To be complete, we should show that

for all strategies, a.s., $\displaystyle \liminf_{T\to+\infty} \frac{R_T}{\sqrt{T \ln \ln T}} > 0$

and again, we would pick $\ell_{j,t}$ iid $\sim \mathrm{Ber}(1/2)$

by showing that for all strategies,

$$\forall\, \ell_{j,t} \in [0,1] \qquad \liminf \frac{\sum_{t=1}^{T} \left(\ell_{J_t,t} - \sum_j p_{j,t}\,\ell_{j,t} \right)}{\sqrt{T \ln(\ln T)}} > 0 \quad \text{a.s.}$$

 Lecture 4 — Spring 2021 term Lecture 4 — Spring 2021 term

Constante optimale pour la borne de regret \ 2

SPDG, supposons que $M - m = 1$ et on souhaite ~~contrôler~~ démontrer que

$$\limsup_{t \to +\infty} \frac{S_t}{\sqrt{t \log\log t}} \leq C =: \frac{1}{\sqrt{2}} \quad \text{p.s.} \qquad (1)$$

où $S_t = \sum_{s=1}^{t} \Big(\ell_{J_s,s} - \mathbb{E}[\ell_{J_s,s} \mid \mathcal{F}_{s-1}] \Big)$.

Rappelons que par l'inégalité de Doob, on a

$$\mathbb{P}\Big(\sup_{t \leq T} S_t \geq \varepsilon \Big) \leq \exp\Big(-\frac{2\varepsilon^2}{T} \Big). \qquad (2).$$

Maintenant, fixons un $\varepsilon > 0$ et posons V_t l'événement suivant

$$V_t = \Big\{ S_t \leq (C+\varepsilon)\sqrt{t \log\log t} \Big\}.$$

<u>Lemme (Borel-Cantelli modifié)</u> Pour démontrer (1), il suffit de démontrer que, pour tout $\varepsilon > 0$ et pour un $a > 1$ quelconque, on a

$$\sum_r \mathbb{P}\Big[\big(V_{\lfloor a^r + 1 \rfloor} \cap V_{\lfloor a^r + 2 \rfloor} \cap \dots \cap V_{\lfloor a^{r+1} \rfloor} \big)^c \Big] < +\infty$$

où B^c signifie le complément de l'événement B.

<u>Preuve du lemme</u> (exactement comme la preuve de Borel-Cantelli).

L'inégalité signifie que $\mathbb{E}\Big[\sum_r \mathbb{1}_{(V_{\lfloor a^r+1 \rfloor} \cap \dots \cap V_{\lfloor a^{r+1} \rfloor})^c} \Big] < +\infty$

donc $\sum_r \mathbb{1}_{(V_{\lfloor a^r+1 \rfloor} \cap \dots \cap V_{\lfloor a^{r+1} \rfloor})^c} < +\infty$ ps

$\Rightarrow \forall r > r_0(\omega) : \ \omega \in V_{\lfloor a^r + 1 \rfloor} \cap \dots \cap V_{\lfloor a^{r+1} \rfloor}$

$\Rightarrow \forall t > t_0(\omega) : \ \omega \in V_t$. On conclut par définition de V_t □

<u>Retour à la preuve de (1)</u>

$$\begin{aligned}
&\mathbb{P}\Big[\big(V_{\lfloor a^r+1 \rfloor} \cap V_{\lfloor a^r+2 \rfloor} \cap \dots \cap V_{\lfloor a^{r+1} \rfloor} \big)^c \Big] \\
&= \mathbb{P}\Big(\exists t : \lfloor a^r + 1 \rfloor \leq t \leq \lfloor a^{r+1} \rfloor \text{ tq } S_t > (C+\varepsilon)\sqrt{t \log\log t} \Big) \\
&\leq \mathbb{P}\Big(\exists t : \lfloor a^r + 1 \rfloor \leq t \leq \lfloor a^{r+1} \rfloor \text{ tq } S_t > (C+\varepsilon)\sqrt{a^r \log\log(a^r)} \Big) \\
&\leq \mathbb{P}\Big(\exists t : \cancel{\lfloor a^r+1 \rfloor}\ 1 \leq t \leq \lfloor a^{r+1} \rfloor \text{ tq } S_t > (C+\varepsilon)\sqrt{a^r \log\log(a^r)} \Big) \\
&\overset{(2)}{\leq} \exp\left(-\frac{2(C+\varepsilon)^2 a^r \log\log(a^r)}{\lfloor a^{r+1} \rfloor} \right) \leq \exp\left(-\frac{2(C+\varepsilon)^2 a^r \log\log(a^r)}{a^{r+1}} \right)
\end{aligned}$$

 Lecture 4 — Spring 2021 term Lecture 4 — Spring 2021 term

$$\leq \exp\left(-\frac{2(C+\varepsilon)^2\log(r\log a)}{a}\right) = \exp\left(-\frac{2(C+\varepsilon)^2\log\log a}{a}\right) r^{-\frac{2(C+\varepsilon)^2}{a}}$$

3

Il suffit donc de choisir $a > 1$ tel que

$$\sum_r r^{-\frac{2(C+\varepsilon)^2}{a}} < +\infty .$$

Or, comme $C = \frac{1}{\sqrt{2}}$, un tel a existe toujours $\square$.

<u>Remarque</u>. La constante $C = \frac{1}{\sqrt{2}}$ est optimale, comme vous avez dit dans le corrigé.