Warning!

Today's lecture notes are based on an article that I am currently writing.

Some results could be improved, some proofs be simplified.

Please don't be shy and send me your suggestions!
- Minor improvements will lead to additional points at the final exam.
- Major improvements will give you an automatic 18 to 20 grade without the need to take the final exam.
- Outstanding improvements will lead you to even be a co-author of this paper under writing.

In particular, these outstanding improvements could consist in:

(1) a much shorter and much natural proof of the variational formula for $\mathcal{K}_{\inf}$

(2) dropping or weakening the dependency in n in front of $e^{-n\varepsilon}$ in the bound:

$$\mathbb{P}\left\{ \max_{\lambda \in [0,1]} \frac{1}{n} \sum_{k=1}^{n} \ln(1-\lambda G_k) \geq \varepsilon \right\} \leq e(n+1)e^{-n\varepsilon}.$$

(3) get a bound of the form $\dfrac{\ln T}{\mathcal{K}_{\inf}(\nu_a, \mu^*)} + O(\ln \ln T)$

instead of $O\big((\ln T)^{4/5} \ln \ln T\big)$

KL-UCB algorithm, to deal with the model $\mathcal{D} = \mathcal{P}([0,1])$

In this lecture, we will write $\mathcal{K}_{\inf}(\nu_a, \mu)$ instead of $\mathcal{K}_{\inf}(\nu_a, \mu, \mathcal{P}([0,1]))$ as the model $\mathcal{D} = \mathcal{P}([0,1])$ will be fixed throughout.

Parameter: $f : \mathbb{N} \setminus \{1 \ldots (K-1)\} \to \mathbb{R}_+^*$ e.g., $f(t) = \ln t + \ln \ln t$

Initialization: Pull each arm once, ie, $I_t = t$ for $t = 1, 2, \ldots K$

For $t \geqslant K$:

1. Compute for each arm a the index

$$U_a(t) = \sup \left\{ \mu \in [0,1] : \mathcal{K}_{\inf}\big(\hat{\nu}_a(t), \mu\big) \leqslant \frac{f(t)}{N_a(t)} \right\}$$

where $\displaystyle \hat{\nu}_a(t) = \frac{1}{N_a(t)} \sum_{s=1}^{t} \delta_{Y_s} \mathbb{1}_{\{I_s = a\}}$ is the empirical distribution associated with a

2. Play $\displaystyle I_{t+1} \in \operatorname*{argmax}_{a \in \{1..K\}} U_a(t)$ [ties broken arbitrarily]

3. Get and observe a reward Y_{t+1} s.t. $Y_{t+1} \mid I_{t+1} \sim \nu_{I_{t+1}}$

Note: We will analyze this algorithm for $\mathcal{P}([0,1])$ but I believe that the analysis would go through for $\mathcal{P}((-\infty, 1])$.

 Lecture 4 — March 1, 2017 Lecture 4 — March 1, 2017

Analysis: Let a^* be any optimal arm, i.e., such that $E(\nu_{a^*}) = \mu^*$.

$$\{I_{t+1} = a\} \subseteq \{\mu^* - \delta \geq U_{a^*}(t)\} \cup \{\mu^* - \delta < U_{a^*}(t) \text{ and } I_{t+1} = a\}$$

$$\subseteq \{\mu^* - \delta \geq U_{a^*}(t)\} \cup \{\mu^* - \delta < U_a(t) \text{ and } I_{t+1} = a\}$$

since we play a, we have $U_{a^*}(t) \leq U_a(t)$

Now, $U_a(t) > \mu^* - \delta$ means that there exists $\mu \in [0,1]$ s.t. $\mu > \mu^* - \delta$ and $K_{\inf}(\hat{\nu}_a(t), \mu) \leq \dfrac{f(t)}{N_a(t)}$

Since $x \mapsto K_{\inf}(\hat{\nu}_a(t), x)$ is non-decreasing (its defining infimum is taken over a smaller set as x increases), we also have

$$K_{\inf}(\hat{\nu}_a(t), \mu^* - \delta) \leq f(t)/N_a(t)$$

so that finally, keeping in mind the initialization phase, we get:

$$E_\nu[N_a(T)] = 1 + \underbrace{\sum_{t=K}^{T-1} \mathbb{P}\{U_{a^*}(t) \leq \mu^* - \delta\}}_{\text{"}a^*\text{ term"}} + \underbrace{\sum_{t=K}^{T-1} \mathbb{P}\left\{K_{\inf}(\hat{\nu}_a(t), \mu^* - \delta) \leq \frac{f(t)}{N_a(t)} \text{ and } I_{t+1} = a\right\}}_{\text{"}a\text{ term"}}$$

$$\leq 1 + \frac{2e^2}{(1 - e^{-\delta^2/2})^2}(2 + \ln \ln T) + \frac{\ln T + \ln \ln T}{K_{\inf}(\nu_a, \mu^*) - 2\delta/1-\mu^*} + \frac{1}{1 - e^{-\delta^2\sqrt{1-\mu^*}/18}}$$

as shown in the next pages (assuming $\mu^* \in (0,1)$)

for any $\delta \in (0, \mu^*)$

Asymptotic bound:

$$\limsup_{T \to +\infty} \frac{E_\nu[N_a(T)]}{\ln T} \leq \frac{1}{K_{\inf}(\nu_a, \mu^*) - 2\delta/1-\mu^*}$$

letting $\delta \downarrow 0$:

$$\limsup_{T \to +\infty} \frac{E_\nu[N_a(T)]}{\ln T} \leq \frac{1}{K_{\inf}(\nu_a, \mu^*)}$$

which is the optimal bound on $\mathcal{P}([0,1])$

Non-asymptotic bound: Optimize the bound above over $\delta \in (0, \mu^*)$ which requires some preliminary further bounding.

Part 0. Preparation / Useful lemmas on Kinf.

REGULARITY

Lemma: For all $\nu \in \mathcal{P}([0,1])$, all $\mu \in (0,1)$, and all $\varepsilon \in (0,\mu)$:

in all cases: $$\mathrm{Kinf}(\nu, \mu-\varepsilon) \leqslant \mathrm{Kinf}(\nu,\mu) \leqslant \mathrm{Kinf}(\nu, \mu-\varepsilon) + \frac{\varepsilon}{1-\mu} \qquad (1)$$

if $\mu - \varepsilon > E(\nu)$: $$\mathrm{Kinf}(\nu,\mu-\varepsilon) + 2\varepsilon^2 \leqslant \mathrm{Kinf}(\nu,\mu) \qquad (2)$$

Proof: (1) Given the definition of $\mathrm{Kinf}(\nu,\mu-\varepsilon)$, we are interested in distributions ν' such that
$$\nu' \in \mathcal{P}([0,1]), \qquad E(\nu') > \mu-\varepsilon, \qquad \nu \ll \nu' \qquad (*)$$

Since ν' has at most countably many atoms, there exists $x \in (\mu, 1)$ with $\nu'\{x\} = 0$

We denote $\nu''_\alpha = \alpha \delta_x + (1-\alpha)\nu'$ for $\alpha \in (0,1)$

$\nu''_\alpha \in \mathcal{P}([0,1])$,
$$E(\nu''_\alpha) = \alpha x + (1-\alpha)E(\nu') > \alpha x + (1-\alpha)(\mu-\varepsilon) = \mu$$
$\nu \ll \nu' \ll \nu''_\alpha$

for the choice $$\alpha = \frac{\varepsilon}{x-(\mu-\varepsilon)} \in (0,1)$$

indeed, $$\frac{1}{x-(\mu-\varepsilon)}\Big(x\varepsilon + (x-\mu)(\mu-\varepsilon)\Big) = \mu$$

Therefore, $\mathrm{Kinf}(\nu,\mu) \leqslant \mathrm{KL}(\nu, \nu''_\alpha)$, which we compute now.

We have $\nu'\{x\} = 0$ thus $\nu\{x\} = 0$, that is, $\delta_x \perp \nu$

from which we see that $$\frac{d\nu}{d\nu''_\alpha} = \frac{1}{1-\alpha}\frac{d\nu}{d\nu'}$$ (easy exercise → but see details in 2 pages)

Therefore,
$$\mathrm{KL}(\nu, \nu''_\alpha) = \int \left(\frac{d\nu}{d\nu''_\alpha} \ln \frac{d\nu}{d\nu''_\alpha}\right) d\nu''_\alpha = \int \left(\ln \frac{d\nu}{d\nu''_\alpha}\right) d\nu$$
$$= \ln\frac{1}{1-\alpha} + \int\left(\ln \frac{d\nu}{d\nu'}\right) d\nu$$
$$= \ln\frac{1}{1-\alpha} + \mathrm{KL}(\nu,\nu')$$

Putting all inequalities together:

$$\mathrm{Kinf}(\nu,\mu) \leqslant \mathrm{KL}(\nu,\nu''_\alpha) \leqslant \mathrm{KL}(\nu,\nu') + \ln\frac{1}{1-\alpha}$$

Taking the infimum over the ν' such that $(*)$ holds, we get:

$$\mathcal{K}_{\inf}(\nu, \mu) \leqslant \mathcal{K}_{\inf}(\nu, \mu - \varepsilon) + \ln \frac{1}{1-\alpha}$$

We substitute the value of α to conclude:

$$\ln \frac{1}{1-\alpha} = \ln \frac{x - \mu + \varepsilon}{(x - \mu + \varepsilon) - \varepsilon} = \ln\left(1 + \frac{\varepsilon}{x - \mu + \varepsilon}\right) \leqslant \frac{\varepsilon}{x - \mu + \varepsilon} \leqslant \varepsilon / x - \mu$$

Now, x could be chosen arbitrarily close to 1, so we can let $x \to 1$ in the previous expression.

(2) We follow the same methodology: let $\nu' \in \mathcal{P}([0,1])$ s.t. $E(\nu') > \mu$ and $\nu \ll \nu'$.

This time, $\nu''_\alpha = \alpha \nu + (1-\alpha)\nu'$ with $\alpha = \dfrac{\varepsilon}{E(\nu') - E(\nu)}$

α lies in $(0,1)$ as $E(\nu') - \varepsilon > \mu - \varepsilon > E(\nu)$

$$\begin{aligned} E(\nu''_\alpha) &= E(\nu') + \alpha\big(E(\nu) - E(\nu')\big) \\ &= E(\nu') - \varepsilon > \mu - \varepsilon \end{aligned}$$

Thus, $$\mathcal{K}_{\inf}(\nu, \mu - \varepsilon) \leqslant \mathrm{KL}(\nu, \nu''_\alpha) \underset{\substack{\text{by convexity}\\ \text{of KL}}}{\leqslant} \alpha\, \cancel{\mathrm{KL}(\nu, \nu)} + (1-\alpha)\, \mathrm{KL}(\nu, \nu')$$

Also, by Pinsker's inequality: (given that all distributions are over $[0,1]$) $$2\alpha\,\big(E(\nu) - E(\nu')\big)^2 \leqslant \alpha\, \mathrm{KL}(\nu, \nu')$$

Adding the two inequalities: (and substituting the value of α) $$\mathcal{K}_{\inf}(\nu, \mu - \varepsilon) + 2\varepsilon\,\underbrace{\big(E(\nu') - E(\nu)\big)}_{> \varepsilon} \leqslant \mathrm{KL}(\nu, \nu')$$

Taking the infimum over the considered ν': $$\mathcal{K}_{\inf}(\nu, \mu - \varepsilon) + 2\varepsilon^2 \leqslant \mathcal{K}_{\inf}(\nu, \mu)$$

<u>Consequence of (1):</u> $\forall \nu \in \mathcal{P}([0,1])$, $\forall \mu \in (0,1)$,

$$\mathcal{K}_{\inf}(\nu, \mu - \varepsilon) \underset{\varepsilon \searrow 0}{\nearrow} \mathcal{K}_{\inf}(\nu, \mu)$$

ie: left-continuity of $\mathcal{K}_{\inf}$

In particular, we have by a sandwich argument:

$$\mathcal{K}_{\inf}(\nu, \mu) = 0 \quad \text{when } E(\nu) \geqslant \mu$$

 Lecture 4 — March 1, 2017 Lecture 4 — March 1, 2017

Additional details given in class on why in the proof of (i), we may take

$$\frac{d\nu}{d\nu''_\alpha} = \frac{1}{1-\alpha}\,\frac{d\nu}{d\nu'},$$

- First, note that we are interested in this function to compute

$$KL(\nu, \nu''_\alpha) = \int \left(\ln \frac{d\nu}{d\nu''_\alpha} \right) d\nu$$

We therefore only need to determine $\frac{d\nu}{d\nu''_\alpha}$ ν-a.s.

- We show that

$$f : y \longmapsto \begin{cases} \dfrac{1}{1-\alpha}\,\dfrac{d\nu}{d\nu'}(y) & \text{if } y \neq x \\ 0 & \text{if } y = x \end{cases}$$

is a density of ν wrt ν''_α.

Indeed: for all $A \in \mathcal{B}([0,1])$, since $\nu\{x\} = 0$,

$$\nu(A) = \nu(A \setminus \{x\}) \underset{\nu \ll \nu''_\alpha}{=} \int \frac{d\nu}{d\nu'}\,\mathbb{1}_A\, d\nu'$$

$$= \int \left(\frac{1}{1-\alpha}\,\frac{d\nu}{d\nu'} \right) \mathbb{1}_A\,(1-\alpha)\,d\nu'$$

$$= \int f\, \mathbb{1}_A \big((1-\alpha)\,d\nu' + \alpha\, d\delta_x \big) \qquad \text{as } f(x) = 0$$

$$= \int f\, \mathbb{1}_A\, d\nu''_\alpha$$

- Thus, $\frac{d\nu}{d\nu''_\alpha} = f$ $\quad \nu''_\alpha$-as thus ν-as

while $f = \frac{1}{1-\alpha}\,\frac{d\nu}{d\nu'}$, ν-as (given $\nu\{x\} = 0$)

This concludes that $\frac{d\nu}{d\nu''_\alpha} = \frac{1}{1-\alpha}\,\frac{d\nu}{d\nu'}$, ν-as

VARIATIONAL FORMULA.

Will be proved next week:

Lemma: For all $\nu \in \mathcal{P}([0,1])$ and $\mu \in (0,1)$:

$$\mathcal{K}_{\inf}(\nu, \mu) = \max_{\lambda \in [0, 1/(1-\mu)]} \mathbb{E}_\nu\big[\ln\big(1 - \lambda(X-\mu)\big)\big]$$

Part 1. Control of the a^* term.

We write

$$\{ U_{a^*}(t) \leq \mu^* - \delta \} \underset{\substack{\text{by definition}\\ \text{of } U_{a^*}(t)}}{\subset} \left\{ \forall \mu > \mu^* - \delta, \quad K_{\inf}(\hat{\nu}_{a^*}(t), \mu) > \frac{f(t)}{N_{a^*}(t)} \right\}$$

$$\subset \left\{ K_{\inf}(\hat{\nu}_{a^*}(t), \mu^* - \delta/2) > \frac{f(t)}{N_{a^*}(t)} \right\} \qquad \text{pick } \mu = \mu^* - \delta/2$$

either $E(\hat{\nu}_{a^*}(t)) \geq \mu^* - \delta/2$ and $K_{\inf}(\hat{\nu}_{a^*}(t), \mu^* - \delta/2) = 0$ as a consequence of left continuity and the event above is the empty set

or $E(\hat{\nu}_{a^*}(t)) < \mu^* - \delta/2$ and we can apply the second part of the regularity lemma:

$$K_{\inf}(\hat{\nu}_{a^*}(t), \mu^* - \delta/2) + \delta^2/2 \leq K_{\inf}(\hat{\nu}_{a^*}(t), \mu^*)$$

In all cases,

$$\{ U_{a^*}(t) \leq \mu^* - \delta \} \subset \left\{ K_{\inf}(\hat{\nu}_{a^*}(t), \mu^*) > \frac{f(t)}{N_{a^*}(t)} + \frac{\delta^2}{2} \right\}$$

cf. variational formula for Kinf

$$= \left\{ \max_{\lambda \in [0, 1/(1-\mu^*)]} \frac{1}{N_{a^*}(t)} \sum_{s=1}^{t} \ln\left(1 - \lambda(Y_s - \mu^*)\right) \mathbb{1}_{\{I_s = a^*\}} > \frac{f(t)}{N_{a^*}(t)} + \frac{\delta^2}{2} \right\}$$

we may add "and $N_{a^*}(t) \in [1, T-K+1]$", because of the initial exploration

The theorem stated in two pages ensures that the probability of the latter event is smaller than

$$\sum_{n=1}^{T-K+1} e^2 (n+1) \exp\left(-f(t) - n\delta^2/2\right).$$

$f(t) = \ln t + \ln \ln t$

$$= \frac{e^2}{t \ln t} \sum_{n=1}^{T-K+1} (n+1) e^{-n\delta^2/2}.$$

The a^* term is therefore bounded by

$$\sum_{t=K}^{T-1} \mathbb{P}\{ U_{a^*}(t) \leq \mu^* - \delta \} \leq \left(\sum_{t=K}^{T-1} \frac{e^2}{t \ln t} \right) \underbrace{\left(\sum_{n=1}^{T-K+1} (n+1) e^{-n\delta^2/2} \right)}_{\text{we will show that } \leq \frac{2}{(1-e^{-\delta^2/2})^2}}$$

$$\leq \frac{e^2}{2\ln 2} + \int_2^{T-1} \frac{e^2}{t \ln t}\, dt \leq \frac{e^2}{2\ln 2} + \left[e^2 \ln \ln t \right]_2^{T-1} \leq e^2 (2 + \ln \ln T)$$

while $\displaystyle\sum_{n=1}^{T-K+1} (n+1)\, e^{-n\delta^2/2} \leqslant \sum_{n\geqslant 1} (n+1)\, e^{-n\delta^2/2} \leqslant ?$

Easy to compute explicitly:

$x>0$, $\displaystyle\varphi(x) = \sum_{n\geqslant 1} (n+1)\, e^{-nx}$

$$A(x) = \sum_{n\geqslant 1} e^{-nx} = e^{-x} \sum_{n\geqslant 0} e^{-nx} = e^{-x} \frac{1}{1-e^{-x}}$$

A is differentiable, with $\displaystyle A'(x) = \sum_{n\geqslant 1} -n e^{-nx} = \frac{-e^{-x}}{(1-e^{-x})^2}$

And by absolute convergence, $\displaystyle\varphi(x) = -A'(x) + A(x) = \frac{e^{-x}(1-e^{-x}) + e^{-x}}{(1-e^{-x})^2}$

$$= \frac{2e^{-x} - e^{-2x}}{(1-e^{-x})^2} \leqslant \frac{2}{(1-e^{-x})^2}$$

We got $\displaystyle\sum_{n=1}^{T-K+1} (n+1)\, e^{-n\delta^2/2} \leqslant \frac{2}{(1-e^{-\delta^2/2})^2}$

All in all, the a^* term is smaller than:

$$\frac{2e^2}{(1-e^{-\delta^2/2})^2} \left(2 + \ln\ln T\right)$$

Theorem (iid version): Let $C_1, C_2, \ldots$ be independent random variables such that $\begin{cases} \mathbb{E}[C_k] = 0 & \text{for all } k \\ C_k \leqslant 1 & \text{a.s. for all } k \end{cases}$

Then for all $\varepsilon > 0$,

$$\mathbb{P}\left\{ \max_{d \in [0,1]} \frac{1}{n} \sum_{k=1}^{n} \ln(1 - d C_k) \geqslant \varepsilon \right\} \leqslant e^2 (n+1)\, e^{-n\varepsilon}.$$

MAIN CHALLENGE: get rather a $O(e^{-n\varepsilon})$ bound; but I'm interested in any weakening of the $n+1$ factor, e.g., into a $\sqrt{n}$.

Theorem (version with a random number of summands): Let $\underline{n}, \overline{n}$ be two integers $\geqslant 1$.

Let g be any non-negative function, possibly taking t or some $\varepsilon > 0$ as arguments:

$$\mathbb{P}\left\{ \max_{d \in [0, 1/(1-\mu^*)]} \sum_{s=1}^{t} \ln\big(1 - d(Y_s - \mu^*)\big) \mathbb{1}_{\{I_s = a^*\}} \geqslant g(N_{a^*}(t), t, \varepsilon) \text{ and } N_{a^*}(t) \in [\![\underline{n}, \overline{n}]\!] \right\}$$

$$\leqslant \sum_{n=\underline{n}}^{\overline{n}} e^2 (n+1) \exp\big(- g(n, t, \varepsilon)\big).$$

 Lecture 4 — March 1, 2017 Lecture 4 — March 1, 2017

The proof for the iid case will rely on the following lemma:

Lemma: For all $d \in [0,1)$ and $d' \in [0,1)$, such that $\begin{cases} \text{either} & d \leq d' \leq 1/2 \\ \text{or} & 1/2 \leq d' \leq d \end{cases}$

for all $c \in (-\infty, 1]$,

$$\ln(1-dc) - \ln(1-d'c) \leq 2\,|d-d'|.$$

Proof: $\Psi_c : d \in [0,1) \longmapsto \ln(1-dc)$ is concave and differentiable,

it can be increasing or decreasing depending on $c>0$ or $c<0$

with $\Psi'_c(d) = \dfrac{-c}{1-dc}$; in particular, $\Psi'_c(1/2) = \dfrac{-c}{1-c/2} = \dfrac{-2c+4-4}{2-c}$

$$= 2 - \frac{4}{2-c}$$

$c \in (-\infty, 1] \longmapsto \Psi'_c(1/2)$ is decreasing

thus $\forall c \in (-\infty,1]$, $\quad -2 = \Psi'_c(1) \leq \Psi'_c(1/2) \leq \lim\limits_{c \to -\infty} \Psi'_c(1/2) = 2$

Now, by convexity (cf. inequality on the slopes): [we of course exclude the case $d = d'$]

- if $d < d' \leq 1/2$: $\quad \dfrac{\Psi_c(d') - \Psi_c(d)}{d'-d} \geqslant \Psi'_c(1/2) \geqslant -2$

  thus $\quad \Psi_c(d') - \Psi_c(d) \geqslant -2(d'-d)$

  $\Psi_c(d) - \Psi_c(d') \leq 2(d'-d) = 2\,|d-d'|$

- if $1/2 \leq d' < d$: $\quad 2 \geqslant \Psi'_c(1/2) \geqslant \dfrac{\Psi_c(d) - \Psi_c(d')}{d-d'}$ hence $\Psi_c(d) - \Psi_c(d') \leqslant 2(d-d')$

Proof (of the theorem, iid version):

$$\mathbb{P}\left\{ \max_{d \in [0,1]} \frac{1}{n} \sum_{k=1}^{n} \ln(1-dC_k) \geq \varepsilon \right\}$$

for all $c \leq 1$, $\ln(1-dc) \xrightarrow[d \to 1]{} \ln(1-c) \in [-\infty, +\infty)$

$$= \mathbb{P}\left\{ \max_{d \in [0,1)} \frac{1}{n} \sum_{k=1}^{n} \ln(1-dC_k) \geq \varepsilon \right\}$$

Markov-Chernoff $\leq e^{-n\varepsilon}\, \mathbb{E}\left[\max_{d\in[0,1)} \exp\left(\sum_{k=1}^{n} \ln(1-dC_k) \right) \right]$

Let $\gamma \in (0, 1/2)$ and define the grid

$$\mathcal{G}_\gamma = \Bigl(\{1/2\} \cup \{ 1/2 - k\gamma : k \in \{1, \dots, \lfloor 1/2\gamma \rfloor \} \} \cup \{ 1/2 + k\gamma : k \in \{1, \dots, \lfloor 1/2\gamma \rfloor\} \} \Bigr) \setminus \{1\}$$

useful only in the case $1/2\gamma \in \mathbb{N}$

$\forall d \in [0,1) \quad \exists d' \in \mathcal{G}_\gamma \quad | \quad |d-d'| \leq \gamma$ and $\begin{cases} \text{either } d \leq d' \leq 1/2 \\ \text{or } 1/2 \leq d' \leq d \end{cases}$

(thus $d' \in [0,1)$)

Also, $|\mathcal{G}_\gamma| \leq 1 + 2 \times 1/2\gamma = 1 + 1/\gamma$

By the lemma, $\displaystyle \max_{d \in [0,1)} \exp\left(\sum_{k=1}^{n} \ln(1-dC_k) \right) = \exp\left(\max_{d\in[0,1)} \sum_{k=1}^{n} \ln(1-dC_k) \right)$

$$\leq \max_{d' \in \mathcal{G}_\gamma} \exp\left(\sum_{k=1}^{n} \ln(1-d'C_k) + 2\gamma n \right)$$

upper bound "à la Pisier" $\displaystyle \leq \left(\sum_{d'\in\mathcal{G}_\gamma} \underbrace{\prod_{k=1}^{n} (1-d'C_k)}_{} \right) e^{2\gamma n}$

where $\displaystyle \mathbb{E}\left[\prod_{k=1}^{n} (1-d'C_k) \right] = \prod_{k=1}^{n} \bigl(1 - d'\, \mathbb{E}[C_k]\bigr) = 1$

by independence of the C_k and $\mathbb{E}[C_k] = 0$

Putting all elements together,

$$\mathbb{P}\left\{ \max_{d\in[0,1]} \frac{1}{n} \sum_{k=1}^{n} \ln(1-dC_k) \geq \varepsilon \right\} \leq e^{-n\varepsilon}\, \mathbb{E}\left[\max_{d\in[0,1)} e^{\sum_{k=1}^{n} \ln(1-dC_k)} \right]$$

$$\leq e^{-n\varepsilon}\, |\mathcal{G}_\gamma|\, e^{2n\gamma}$$

$$= e^{-n\varepsilon} (1 + 1/\gamma)\, e^{2n\gamma}$$

and we pick $\gamma = 1/n$ to conclude.

Lecture 4 — March 1, 2017 Lecture 4 — March 1, 2017

<u>Proof</u> (of the theorem, random number of summands):

- For each fixed $d \in [0, 1/_{1-\mu^*}]$,

$$M_{d,t} = \exp\left(\sum_{s=1}^{t} \ln\bigl(1 - d(y_s - \mu^*)\bigr) \mathbb{1}_{\{I_s = a^*\}} \right)$$

is a martingale wrt $\mathcal{F}_t = \sigma(y_1 \ldots y_t)$. In particular, $\mathbb{E}[M_{d,t}] = 1$.

<u>Indeed</u>, it suffices to note that

$$\mathbb{E}\Bigl[\exp\bigl(\ln(1 - d(y_t - \mu^*)) \mathbb{1}_{\{I_t = a^*\}} \bigr) \Bigm| \mathcal{F}_{t-1} \Bigr]$$

$$= \mathbb{E}\Bigl[\mathbb{1}_{\{I_t \neq a^*\}} + \mathbb{1}_{\{I_t = a^*\}} \bigl(1 - d(y_t - \mu^*)\bigr) \Bigm| \mathcal{F}_{t-1} \Bigr]$$

(I_t is $\mathcal{F}_{t-1}$-measurable)

$$= \mathbb{1}_{\{I_t \neq a^*\}} + \mathbb{1}_{\{I_t = a^*\}} \Bigl(1 - d\bigl(\mathbb{E}[y_t \mid \mathcal{F}_{t-1}] - \mu^*\bigr)\Bigr)$$

by the bandit model, $\mathbb{E}[y_t \mid \mathcal{F}_{t-1}] = \mu_{I_t}$ as $y_t \mid I_t \sim \nu_{I_t}$ (we integrate over the random draw of y_t given I_t)

$$= \mathbb{1}_{\{I_t \neq a^*\}} + \mathbb{1}_{\{I_t = a^*\}} \bigl(1 - d(\underbrace{\mu_{I_t} - \mu^*}_{=0 \text{ on } \{I_t = a^*\}})\bigr)$$

$$= \mathbb{1}_{\{I_t \neq a^*\}} + \mathbb{1}_{\{I_t = a^*\}}$$

$$= 1.$$

- The probability of interest equals

$$\mathbb{P}\Bigl\{ \max_{d \in [0, 1/_{1-\mu^*}]} \ln M_{d,t} > g(N_{a^*}(t), t, \varepsilon) \text{ and } N_{a^*}(t) \in [\underline{n}, \overline{n}] \Bigr\}$$

$$= \sum_{n=\underline{n}}^{\overline{n}} \mathbb{P}\Bigl\{ \max_{d \in [0, 1/_{1-\mu^*}]} M_{d,t} > \exp\bigl(g(n,t,\varepsilon)\bigr) \text{ and } N_{a^*}(t) = n \Bigr\}$$

Now, setting $C_s = \dfrac{y_s - \mu^*}{1 - \mu^*} \leqslant 1$, we get

$$\max_{d \in [0, 1/_{1-\mu^*}]} M_{d,t} = \max_{d \in [0,1]} \exp\left(\sum_{s=1}^{t} \bigl(\ln(1 - dC_s)\bigr) \mathbb{1}_{\{I_s = a^*\}} \right)$$

(we rescale)

$$\leqslant \max_{d \in \mathcal{G}_\gamma} \exp\left(\sum_{s=1}^{t} \ln(1 - dC_s) \mathbb{1}_{\{I_s = a^*\}} \right) \times \exp\bigl(\gamma N_{a^*}(t)\bigr)$$

(cf. approximation lemma and $d \in [0,1)$ is enough)

so that the probability of interest is bounded, by a Markov–Chernov inequality:

$$\sum_{n=\underline{n}}^{\overline{n}} e^{-g(n,t,\varepsilon)} \, \mathbb{E}\Bigl[\Bigl(\max_{d \in [0, 1/_{1-\mu^*}]} M_{d,t} \Bigr) \mathbb{1}_{\{N_{a^*}(t) = n\}} \Bigr]$$

(we drop the indicator $\mathbb{1}_{\{N_{a^*}(t) = n\}}$)

$$\leqslant \sum_{n=\underline{n}}^{\overline{n}} e^{-g(n,t,\varepsilon)} e^{\gamma n} \, \mathbb{E}\Bigl[\max_{d \in \mathcal{G}_\gamma} \exp\Bigl(\sum_{s=1}^{t} \ln(1 - dC_s) \mathbb{1}_{\{I_s = a^*\}} \Bigr) \Bigr]$$

As shown above, $\mathbb{E}\Big[\exp\Big(\sum_{s=1}^{t} \ln(1-\alpha C_s)\, \mathbb{1}_{\{I_s = a^\star\}} \Big) \Big]$

$= \mathbb{E}\big[M_{\alpha/1-\mu^\star, t} \big] = 1$

So that, bounding "à la Pisier", we see that the probability of interest is smaller than

$$\sum_{n=\underline{n}}^{\overline{n}} e^{-g(n,t,\varepsilon)} \, e^{\gamma n} \, \underbrace{|\mathcal{G}_\gamma|}_{\leq e^2 (n+1)} \quad \text{for } \gamma = 1/n.$$

 Lecture 4 — March 1, 2017 Lecture 4 — March 1, 2017

Part 2. Control of the a term

By the regularity lemma $K_{\inf}(\hat{\nu}_a(t), \mu^*-\delta) + \dfrac{\delta}{1-\mu^*} \geqslant K_{\inf}(\hat{\nu}_a(t), \mu^*)$

So that
$$\sum_{t=K}^{T-1} \mathbb{P}\Big\{ K_{\inf}(\hat{\nu}_a(t), \mu^*-\delta) \leqslant \frac{f(t)}{N_a(t)} \text{ and } I_{t+1}=a \Big\}$$
$$\leqslant \sum_{t=K}^{T-1} \mathbb{P}\Big\{ K_{\inf}(\hat{\nu}_a(t), \mu^*) \leqslant \frac{f(t)}{N_a(t)} + \frac{\delta}{1-\mu^*} \text{ and } I_{t+1}=a \Big\}$$
$$= \sum_{t=K}^{T-1} \sum_{\ell=1}^{T-1} \mathbb{P}\Big\{ K_{\inf}(\hat{\nu}_a(t), \mu^*) \leqslant \frac{f(t)}{\ell} + \frac{\delta}{1-\mu^*} \text{ and } N_a(t)=\ell \text{ and } I_{t+1}=a \Big\}$$

We set $n_0 = \left\lceil \dfrac{f(T)}{K_{\inf}(\nu_a, \mu^*) - 2\delta/1-\mu^*} \right\rceil$ so that for $\ell \geqslant n_0$,
$$\frac{f(t)}{\ell} \leqslant \frac{f(T)}{n_0} \leqslant K_{\inf}(\nu_a, \mu^*) - \frac{2\delta}{1-\mu^*}$$

By the variational lemma, there exists $\tilde{\lambda} \in [0, 1/1-\mu^*]$ such that
$$K_{\inf}(\nu_a, \mu^*) = \mathbb{E}_{\nu_a}\big[\ln(1-\tilde{\lambda}(X-\mu^*)) \big] \quad \text{where } X \sim \nu_a$$

By the variational lemma again,
$$K_{\inf}(\hat{\nu}_a(t), \mu^*) \geqslant \frac{1}{N_a(t)} \sum_{s=1}^{t} \ln\big(1-\tilde{\lambda}(Y_s-\mu^*)\big) \mathbb{1}_{\{I_s=a\}}$$

Therefore, for $\ell \geqslant n_0$,
$$\Big\{ K_{\inf}(\hat{\nu}_a(t), \mu^*) \leqslant \frac{f(t)}{\ell} + \frac{\delta}{1-\mu^*} \Big\} \subset \Big\{ K_{\inf}(\hat{\nu}_a(t), \mu^*) \leqslant K_{\inf}(\nu_a, \mu^*) - \delta/1-\mu^* \Big\}$$
$$\subset \Big\{ \frac{1}{N_a(t)} \sum_{s=1}^{t} \ln\big(1-\tilde{\lambda}(Y_s-\mu^*)\big) \mathbb{1}_{\{I_s=a\}} \leqslant \mathbb{E}_{\nu_a}\big[\ln(1-\tilde{\lambda}(X-\mu^*))\big] - \delta/1-\mu^* \Big\}$$

Putting all bounds together, we see that the a term is smaller than
$$\sum_{t=1}^{K} \Big(\sum_{\ell=1}^{n_0-1} \mathbb{P}\{ N_a(t)=\ell \text{ and } I_{t+1}=a \} + \sum_{\ell=n_0}^{T-1} \mathbb{P}\Big\{ \frac{1}{N_a(t)} \sum_{s=1}^{t} X_s \mathbb{1}_{\{I_s=a\}} - \mathbb{E}Z_a \leqslant -\frac{\delta}{1-\mu^*} \text{ and } N_a(t)=\ell \text{ and } I_{t+1}=a \Big\} \Big)$$

X_s short-hand notation for $\ln(1-\tilde{\lambda}(Y_s-\mu^*))$

short-hand notation for $\mathbb{E}_{\nu_a}[\ln(1-\tilde{\lambda}(X-\mu^*))]$

for all t, there exists at most one ℓ such that $N_a(t)=\ell$ and $I_{t+1}=a$

To continue the bounding, we will use the following lemmas.