 Lecture 6 — Winter 2019 term Lecture 6 — Winter 2019 term

Solution for Exercise $1/\mathcal{K}_{\inf}$ vs. $8/\Delta_a^2$ for UCB

1) For $q \in (0,1)$ ie, excluding for the time being $q=0$ or $q=1$:

$$p \in [0,1] \mapsto \mathrm{kl}(p,q) = p \ln \frac{p}{q} + (1-p) \ln \frac{1-p}{1-q}$$ is twice differentiable at least on $(0,1)$,

with $$\frac{\partial \mathrm{kl}}{\partial p}(p,q) = \ln\left(\frac{p(1-q)}{q(1-p)}\right) \quad \text{and} \quad \frac{\partial^2}{\partial p^2}\mathrm{kl}(p,q) = \frac{1}{p(1-p)} \geqslant 4$$

so that a second-order Taylor expansion ensures:

$$\exists\, r \in [\min(p,q), \max(p,q)] \;\Big|\; \mathrm{kl}(p,q) = \cancel{\mathrm{kl}(q,q)} + \cancel{(p-q)\frac{\partial}{\partial p}\mathrm{kl}(q,q)} + \frac{(p-q)^2}{2}\frac{\partial^2}{\partial p^2}\mathrm{kl}(r,q)$$
$$\geqslant 2(p-q)^2.$$

We deal separately with $p \in \{0,1\}$ or $q \in \{0,1\}$.

Note: Sharper Pinsker lower bounds are possible, see, eg, the excerpt of an article reproving the (optimal in some sense) refined Pinsker's bound by Ordentlich and Weinberger (2005)

2) By the data-processing inequality with expectations of $[0,1]$-valued random variables:

$$\mathrm{KL}(\mathbb{P},\mathbb{Q}) \geqslant \mathrm{kl}\big(\mathbb{E}_{\mathbb{P}}[Z], \mathbb{E}_{\mathbb{Q}}[Z]\big) \underset{\text{by 1)}}{\geqslant} 2\big(\mathbb{E}_{\mathbb{P}}[Z] - \mathbb{E}_{\mathbb{Q}}[Z]\big)^2.$$

Rearranging and taking the sup over Z leads to the desired bound.

3) Note that $E(\nu) = \mathbb{E}_\nu[X]$ where $X \sim \nu$ is indeed a $[0,1]$-valued random variable given that we consider the model $\mathcal{P}([0,1])$

Thus $$\mathrm{KL}(\nu_a, \nu) \geqslant 2\big(\underbrace{E(\nu_a)}_{=\mu_a<\mu^*} - \underbrace{E(\nu)}_{>\mu^*}\big)^2 > 2(\mu^* - \mu_a)^2 = 2\Delta_a^2$$

Taking the infimum over $\nu \in \mathcal{P}([0,1])$ with $E(\nu) > \mu^*$: $\quad \mathcal{K}_{\inf}(\nu_a, \mu^*, \mathcal{P}([0,1])) \geqslant 2\Delta_a^2$

Thus the $\dfrac{\ln T}{\mathcal{K}_{\inf}(\nu_a, \mu^*, \mathcal{P}([0,1]))}$ bound is better than the UCB bounds of the form $\dfrac{c}{\Delta_a^2}$

where actually, after double-checking, it turns out that c can be made arbitrarily close to $1/2$, up to degrading the constant additive term.

Question (answer on next page):

Would you have examples of $\mathcal{K}_{\inf}(\nu_a, \mu^*, \mathcal{P}([0,1])) > 2\Delta_a^2$?

Yes: pick $\nu_a = \mathrm{Ber}(\mu_a)$ with $0 < \mu_a < \mu^*$

Then for all $\nu' \in \mathcal{P}([0,1])$ with $E(\nu') > \mu^*$:

$$\mathrm{KL}(\mathrm{Ber}(\mu_a), \nu') \geqslant \mathrm{KL}(\mathrm{Ber}(\mu_a), \mathrm{Ber}(E(\nu'))) \geqslant kl(\mu_a, \mu^*)$$

data-processing inequality with expectations

$kl(\mu_a, \cdot)$ is increasing on $[\mu_a, 1)$, as it is a strictly convex function with minimum achieved at μ_a

Thus, the "worst-case" ν' are given by Bernoulli distributions.

We proved:

$$\mathcal{K}_{\inf}(\mathrm{Ber}(\mu_a), \mu^*) \geqslant kl(\mu_a, \mu^*) > 2(\mu^* - \mu_a)^2 = 2\Delta_a^2$$

in general (I think: in all cases but $\mu^* = \mu_a$)

The next theorem is a stronger version of Pinsker's inequality for Bernoulli distributions, that was proved[2] by Ordentlich and Weinberger [2005]. Indeed, note that the function φ defined below satisfies $\min \varphi = 2$, so that the next theorem always yields an improvement over the most classical version of Pinsker's inequality: $\mathrm{kl}(p,q) \geqslant 2(p-q)^2$.

We provide below an alternative elementary proof for Bernoulli distributions of this refined Pinsker's inequality. The extension to the case of general distributions, via the contraction-of-entropy property, is stated at the end of this section.

Theorem 15 (A refined Pinsker's inequality by Ordentlich and Weinberger [2005])**.** *For all $p, q \in [0,1]$,*

$$\mathrm{kl}(p,q) \geqslant \frac{\ln\big((1-q)/q\big)}{1-2q}\,(p-q)^2 \stackrel{\mathrm{def}}{=} \varphi(q)\,(p-q)^2\,,$$

where the multiplicative factor $\varphi(q) = (1-2q)^{-1}\ln\big((1-q)/q\big)$ is defined for all $q \in [0,1]$ by extending it by continuity as $\varphi(1/2) = 2$ and $\varphi(0) = \varphi(1) = +\infty$.

The proof shows that $\varphi(q)$ is the optimal multiplicative factor in front of $(p-q)^2$ when the bounds needs to hold for all $p \in [0,1]$; the proof also provides a natural explanation for the value of φ.

Proof: The stated inequality is satisfied for $q \in \{0,1\}$ as $\mathrm{kl}(p,q) = +\infty$ in these cases unless $p = q$. The special case $q = 1/2$ is addressed at the end of the proof. We thus fix $q \in (0,1)\setminus\{1/2\}$ and set $f(p) = \mathrm{kl}(p,q)/(p-q)^2$ for $p \neq q$, with a continuity extension at $p = q$. We exactly show that f attains its minimum at $p = 1-q$, from which the result (and its optimality) follow by noting that

$$f(1-q) = \frac{\mathrm{kl}(1-q,q)}{(1-2q)^2} = \frac{\ln\big((1-q)/q\big)}{1-2q} = \varphi(q)\,.$$

Given the form of f, it is natural to perform a second-order Taylor expansion of $\mathrm{kl}(p,q)$ around q. We have

$$\frac{\partial}{\partial p}\,\mathrm{kl}(p,q) = \ln\!\left(\frac{p(1-q)}{(1-p)q}\right) \qquad \text{and} \qquad \frac{\partial^2}{\partial^2 p}\,\mathrm{kl}(p,q) = \frac{1}{p(1-p)} \stackrel{\mathrm{def}}{=} \psi(p)\,, \tag{41}$$

so that Taylor's formula with integral remainder reveals that for $p \neq q$,

$$f(p) = \frac{\mathrm{kl}(p,q)}{(p-q)^2} = \frac{1}{(p-q)^2}\int_q^p \frac{\psi(t)}{1!}(p-t)^1\,\mathrm{d}t = \int_0^1 \psi\big(q + u(p-q)\big)(1-u)\,\mathrm{d}u\,.$$

This rewriting of f shows that f is strictly convex (as ψ is so). Its global minimum is achieved at the unique point where its derivative vanishes. But by differentiating under the integral sign, we have, at $p = 1-q$,

$$f'(1-q) = \int_0^1 \psi'\big(q + u(1-2q)\big)\,u(1-u)\,\mathrm{d}u = 0\,;$$

the equality to 0 follows from the fact that the function $u \mapsto \psi'\big(q+u(1-2q)\big)u(1-u)$ is antisymmetric around $u = 1/2$ (essentially because ψ' is antisymmetric itself around $1/2$). As a consequence, the convex function f attains its global minimum at $1-q$, which concludes the proof for the case where $q \in (0,1)\setminus\{1/2\}$.

It only remains to deal with $q = 1/2$: we use the continuity of $\mathrm{kl}(p,\,\cdot\,)$ and φ to extend the obtained inequality from $q \in [0,1]\setminus\{1/2\}$ to $q = 1/2$. □

We now prove the second inequality of (13). A picture is helpful, see Figure 1.

[2] We also refer the reader to Kearns and Saul [1998, Lemma 1] and Berend and Kontorovich [2013, Theorem 3.2] for dual inequalities upper bounding the moment-generating function of the Bernoulli distributions.

Solution for Exercise « All algorithms explore much »

(1) * We extract from the proof of question 1 in Exercise 1 that

$$\forall (p,q) \in (0,1)^2 \text{ s.t. } p < q,$$
$$\exists\, r \in (p,q) \text{ s.t.}$$
$$\mathrm{kl}(p,q) = \frac{1}{2r(1-r)} (p-q)^2$$

hence the first stated bound for $p > 0$ and $q < 1$.

The cases $p = 0$ and $q = 1$ follow by continuity here as well.

* The second is because $\max\limits_{x \in [p,q]} x(1-x) \leq \max\limits_{x \in [p,q]} x = q$.

* Stronger than the global version because we replaced the $2 = \dfrac{1}{2 \times 1/4}$ by $\dfrac{1}{2 \max\limits_{x \in [p,q]} x(1-x)}$

↳ we bounded the 2nd derivative of interest $\dfrac{1}{x(1-x)}$ not globally but on the interval $[p,q]$ at hand.

(2) Again, if $\nexists\, \nu'_j \in \mathcal{D}$ with $E(\nu'_j) > \mu^*$, then $K_{\inf}(\nu_j, \mu^*) = +\infty$
for a suboptimal arm j and $KL^* = +\infty$.

So, with no loss of generality, we may assume that: for any suboptimal arm j, there exists such a ν'_j and we get again, considering $\nu' = (\nu'_k)_{k \leq K}$ with $\begin{cases} \nu'_k = \nu_k & \text{if } k \neq j \\ \nu'_j & \text{s.t. } E(\nu'_j) > \mu^* \end{cases}$

$$\mathbb{E}_{\nu}[N_j(T)]\ \mathrm{KL}(\nu_j, \nu'_j) \geq \mathrm{kl}\big(\mathbb{E}_{\nu}[N_j(T)/T],\ \mathbb{E}_{\nu'}[N_j(T)/T]\big)$$

Since the strategy is smarter than the uniform strategy, and j is optimal under ν^j,

$$\mathbb{E}_{\nu^j}[N_j(T)/T] \geqslant \frac{1}{K}$$

If $\mathbb{E}_{\nu}[N_j(T)/T] \geqslant 1/K$ then the desired bound is true.

Otherwise: by strict convexity of $q \mapsto kl(\mathbb{E}_{\nu}[N_j(T)/T], q)$ this function is increasing on $[\mathbb{E}_{\nu}[N_j(T)/T], 1]$, so that

$$kL(\mathbb{E}_{\nu}[N_j(T)/T], \mathbb{E}_{\nu^j}[N_j(T)/T]) \geqslant kl(\mathbb{E}_{\nu}[N_j(T)/T], 1/K)$$

which by (i) is larger than:

$$kl(\mathbb{E}_{\nu}[N_j(T)/T], 1/K) \geqslant \frac{1}{2 \times 1/K}\left(\mathbb{E}_{\nu}[N_j(T)/T] - 1/K\right)^2$$

Recall that we are in the case where $p = \mathbb{E}_{\nu}[N_j(T)/T] < 1/K = q$; we proved so far

$$\frac{T}{K} KL(\nu_j, \nu'_j) > \mathbb{E}_{\nu}[N_j(T)] \, KL(\nu_j, \nu'_j)$$

$$\geqslant \frac{K}{2}\left(\underbrace{\mathbb{E}_{\nu}[N_j(T)/T] - 1/K}_{<1/K}\right)^2$$

and thus get

$$\mathbb{E}_{\nu}[N_j(T)/T] \geqslant \underbrace{\frac{1}{K} - \sqrt{\frac{2T\, KL(\nu_j, \nu'_j)}{K^2}}}_{\text{independent of } \nu'_j \text{ s.t. } \nu'_j \in \mathcal{D},\ E(\nu'_j) > \mu^*}$$

therefore, taking the supremum over these ν'_j:

$$\mathbb{E}[N_j(T)/T] \geqslant 1/K - \frac{\sqrt{2T\, K_{\inf}(\nu_j, \mu^*)}}{K} \geqslant \frac{1}{K}\left(\underbrace{1 - \sqrt{2T\, KL^*}}_{\text{is } \geqslant 1/2 \text{ for } T \leqslant 1/(8\, KL^*)}\right)$$