On Information Inequalities and the Complexity of Sequential Decision Problems

Orsay Machine Learning Master Seminar

Aurélien Garivier

November 5th, 2019

Table of contents

1

Introduction: Information for Deviation Lower Bounds

Chernoff Bound for Bernoulli variables

Let $\mu \in (0,1)$. Let $X_1, X_2, \ldots X_n \overset{iid}{\sim} \mathcal{B}(\mu)$, and $\bar{X}_n = (X_1 + \cdots + X_n)/n$.

Theorem

For all $\mu \leq x \leq 1$,

$$\mathbb{P}_\mu \left(\bar{X}_n \geq x \right) \leq e^{-n \, \mathrm{kl}(x, \mu)}$$

where $\mathrm{kl}(x, y) = x \log \dfrac{x}{y} + (1 - x) \log \dfrac{1 - x}{1 - y}$ is the binary relative entropy. Similary, for all $0 \leq x \leq \mu$,

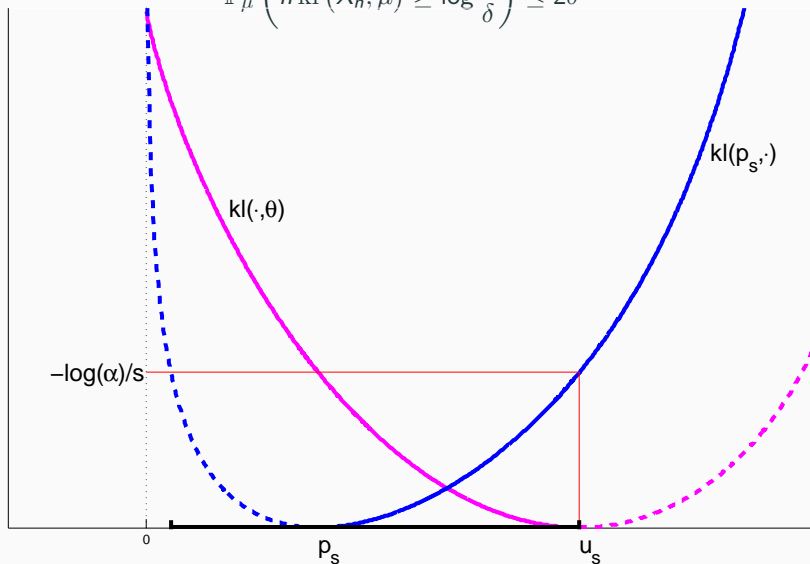$$\mathbb{P}_\mu \left(\bar{X}_n \leq x \right) \leq e^{-n \, \mathrm{kl}(x, \mu)} \; .$$

Corollary

For every $\delta > 0$,

$$\mathbb{P}_\mu \left(n \, \mathrm{kl} \left(\bar{X}_n, \mu \right) \geq \log \frac{1}{\delta} \right) \leq 2\delta \; .$$

$$\mathbb{P}_\mu \left(n \operatorname{kl} \left(\bar{X}_n, \mu \right) \geq \log \frac{1}{\delta} \right) \leq 2\delta$$

kl($\cdot$,θ)

kl(p_s,$\cdot$)

$-\log(\alpha)/s$

0   p_s u_s

3

Proof: Fenchel-Legendre transform of log-Laplace

For every $\lambda > 0$,

$$\mathbb{P}_\mu \left(\bar{X}_n \geq x \right) = \mathbb{P}_\mu \left(e^{\lambda(X_1 + \cdots + X_n)} \geq e^{\lambda n x} \right)$$

$$\leq \frac{\mathbb{E}_\mu \left[e^{\lambda(X_1 + \cdots + X_n)} \right]}{e^{\lambda n x}} \qquad \text{by Markov's inequality}$$

$$= e^{-n\left(\lambda x - \log \mathbb{E}_\mu [\exp \lambda X_1] \right)} .$$

Thus, $\quad -\dfrac{1}{n} \log \mathbb{P}_\mu \left(\bar{X}_n \geq x \right) \geq \sup_{\lambda > 0} \left\{ \lambda x - \log \mathbb{E}_\mu [\exp \lambda X_1]) \right\}$

$$= \sup_{\lambda > 0} \left\{ \lambda x - \log \left(1 - \mu + \mu e^\lambda \right) \right\}$$

$$= \mathrm{kl}(x, \mu) .$$

kl = binary Kullback-Leibler divergence: $\mathrm{kl}(x, \mu) = \mathrm{KL}\big(\mathcal{B}(x), \mathcal{B}(\mu)\big)$

$$\text{where} \quad \mathrm{KL}(P, Q) = \begin{cases} \int \log \frac{dP}{dQ} dP & \text{if } P \ll Q , \\ +\infty & \text{otherwise.} \end{cases}$$

Properties: $0 \leq \mathrm{KL}(P, Q) \leq +\infty$, and $\mathrm{KL}(P, Q) = 0$ iff $P = Q$.

Lower Bound: Change of Measure

For all $\epsilon > 0$ and all $\alpha > 0$,

$$\mathbb{P}_\mu\left(\bar{X}_n \geq x\right) = \mathbb{E}_\mu\left[\mathbb{1}\{\bar{X}_n \geq x\}\right]$$

$$= \mathbb{E}_{x+\epsilon}\left[\mathbb{1}\{\bar{X}_n \geq x\} \times \frac{d\mathbb{P}_\mu}{d\mathbb{P}_{x+\epsilon}}(X_1, \ldots, X_n)\right]$$

$$= \mathbb{E}_{x+\epsilon}\left[\mathbb{1}\{\bar{X}_n \geq x\} \times e^{-\sum_{i=1}^n \log \frac{dP_{x+\epsilon}}{dP_\mu}(X_i)}\right]$$

$$\geq \mathbb{E}_{x+\epsilon}\left[\mathbb{1}\{\bar{X}_n \geq x\} \, \mathbb{1}\left\{\frac{1}{n}\sum_{i=1}^n \log \frac{dP_{x+\epsilon}}{dP_\mu}(X_i) \leq \mathbb{E}_{x+\epsilon}\left[\log \frac{dP_{x+\epsilon}}{dP_\mu}(X_1)\right] + \alpha\right\}\right.$$

$$\left. \times \, e^{-\sum_{i=1}^n \log \frac{dP_{x+\epsilon}}{dP_\mu}(X_i)}\right]$$

$$\geq e^{-n\left\{\mathbb{E}_{x+\epsilon}\left[\log \frac{dP_{x+\epsilon}}{dP_\mu}(X_1)\right] + \alpha\right\}}\left[1 - \mathbb{P}_{x+\epsilon}\left(\bar{X}_n < x\right)\right.$$

$$\left. - \mathbb{P}_{x+\epsilon}\left(\frac{1}{n}\sum_{i=1}^n \log \frac{dP_{x+\epsilon}}{dP_\mu}(X_i) > \mathbb{E}_{x+\epsilon}\left[\log \frac{dP_{x+\epsilon}}{dP_\mu}(X_1)\right] + \alpha\right)\right]$$

$$= e^{-n\left\{\mathrm{kl}(x+\epsilon,\mu) + \alpha\right\}}\left(1 - o_n(1)\right).$$

Lower Bound: Change of Measure

For all $\epsilon > 0$ and all $\alpha > 0$,

$$\mathbb{P}_\mu\left(\bar{X}_n \geq x\right) = \mathbb{E}_\mu\left[\mathbb{1}\{\bar{X}_n \geq x\}\right]$$

$$\geq \mathbb{E}_{x+\epsilon}\left[\mathbb{1}\{\bar{X}_n \geq x\}\,\mathbb{1}\left\{\frac{1}{n}\sum_{i=1}^n \log\frac{dP_{x+\epsilon}}{dP_\mu}(X_i) \leq \mathbb{E}_{x+\epsilon}\left[\log\frac{dP_{x+\epsilon}}{dP_\mu}(X_1)\right] + \alpha\right\}\right.$$

$$\left.\times\, e^{-\sum_{i=1}^n \log\frac{dP_{x+\epsilon}}{dP_\mu}(X_i)}\right]$$

$$\geq e^{-n\left\{\mathbb{E}_{x+\epsilon}\left[\log\frac{dP_{x+\epsilon}}{dP_\mu}(X_1)\right]+\alpha\right\}}\left[1 - \mathbb{P}_{x+\epsilon}\left(\bar{X}_n < x\right)\right.$$

$$\left. - \mathbb{P}_{x+\epsilon}\left(\frac{1}{n}\sum_{i=1}^n \log\frac{dP_{x+\epsilon}}{dP_\mu}(X_i) > \mathbb{E}_{x+\epsilon}\left[\log\frac{dP_{x+\epsilon}}{dP_\mu}(X_1)\right] + \alpha\right)\right]$$

$$= e^{-n\left\{\mathrm{kl}(x+\epsilon,\mu)+\alpha\right\}}\left(1 - o_n(1)\right).$$

Asymptotic Optimality (Large Deviation Lower Bound)

$$\liminf_n \frac{1}{n}\log\mathbb{P}_\mu\left(\bar{X}_n \geq x\right) \geq -\mathrm{kl}(x,\mu).$$

Lower Bound: Change of Measure

For all $\epsilon > 0$ and all $\alpha > 0$,

$$\mathbb{P}_\mu \left(\bar{X}_n \geq x \right) = \mathbb{E}_\mu \left[\mathbb{1}\{\bar{X}_n \geq x\} \right]$$

$$\geq \mathbb{E}_{x+\epsilon} \left[\mathbb{1}\{\bar{X}_n \geq x\} \, \mathbb{1}\left\{ \frac{1}{n} \sum_{i=1}^{n} \log \frac{dP_{x+\epsilon}}{dP_\mu}(X_i) \leq \mathbb{E}_{x+\epsilon} \left[\log \frac{dP_{x+\epsilon}}{dP_\mu}(X_1) \right] + \alpha \right\} \right.$$

$$\left. \times \; e^{-\sum_{i=1}^{n} \log \frac{dP_{x+\epsilon}}{dP_\mu}(X_i)} \right]$$

$$\geq e^{-n\left\{ \mathbb{E}_{x+\epsilon}\left[\log \frac{dP_{x+\epsilon}}{dP_\mu}(X_1) \right] + \alpha \right\}} \left[1 - \mathbb{P}_{x+\epsilon}\left(\bar{X}_n < x \right) \right.$$

$$\left. - \mathbb{P}_{x+\epsilon}\left(\frac{1}{n} \sum_{i=1}^{n} \log \frac{dP_{x+\epsilon}}{dP_\mu}(X_i) > \mathbb{E}_{x+\epsilon}\left[\log \frac{dP_{x+\epsilon}}{dP_\mu}(X_1) \right] + \alpha \right) \right]$$

$$= e^{-n\left\{ \mathrm{kl}(x+\epsilon,\mu) + \alpha \right\}} \left(1 - o_n(1) \right) .$$

Asymptotic Optimality (Large Deviation Principle)

$$\frac{1}{n} \log \mathbb{P}_\mu \left(\bar{X}_n \geq x \right) \xrightarrow[n\to\infty]{} -\mathrm{kl}(x, \mu) .$$

Lower Bound: the Entropic Way

Let $\Omega = \{0,1\}^n$, $X_i(\omega) = \omega_i$

Probability laws on Ω: $\mathbb{P}_p = \mathcal{B}(p)^{\otimes n}$.

For all $\epsilon > 0$,

$$n \, \mathrm{kl}(x + \epsilon, \mu) = \mathrm{KL}\left(\mathbb{P}_{x+\epsilon}, \mathbb{P}_\mu\right) \qquad \scriptstyle \mathrm{KL}(P \otimes P', \, Q \otimes Q') = \mathrm{KL}(P, Q) + \mathrm{KL}(P', Q')$$

$$\geq \mathrm{KL}\left(\mathbb{P}_{x+\epsilon}^{\mathbb{1}\left\{\bar{X}_n \geq x\right\}}, \, \mathbb{P}_\mu^{\mathbb{1}\left\{\bar{X}_n \geq x\right\}}\right) \qquad \begin{array}{c} \scriptstyle \mathrm{KL}(P, Q) \geq \mathrm{KL}(P^X, Q^X) \\ \scriptstyle \text{contraction of entropy} \\ \scriptstyle = \text{data-processing inequality} \end{array}$$

$$= \mathrm{kl}\left(\mathbb{P}_{x+\epsilon}\left(\bar{X}_n \geq x\right), \, \mathbb{P}_\mu\left(\bar{X}_n \geq x\right)\right)$$

$$\geq \mathbb{P}_{x+\epsilon}\left(\bar{X}_n \geq x\right) \log \frac{1}{\mathbb{P}_\mu\left(\bar{X}_n \geq x\right)} - \log(2) \qquad \scriptstyle \mathrm{kl}(p, q) \geq p \log \frac{1}{q} \, - \log 2$$

A non-asymptotic lower bound

$$\forall \epsilon > 0, \qquad \mathbb{P}_\mu\left(\bar{X}_n \geq x\right) \geq e^{-\frac{n \, \mathrm{kl}(x+\epsilon, \, \mu) + \log(2)}{1 - e^{-2n\epsilon^2}}} \, .$$
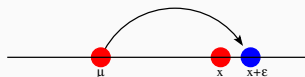
Intermediate: Sequential Test

$X_1, X_2, \ldots$ independent random variables with distribution $\mathcal{B}(\mu)$.

- Goal: say if $\mu > 1/2$ or $\mu < 1/2$.
- Sequential: stopping time τ, decision rule $\hat{a}_\tau = \in \{+, -\}$ is $\mathcal{F}_\tau$-measurable, where $\mathcal{F}_t = \sigma(X_1, \ldots, X_t)$.
- δ-correctness constraint:

  $$\forall \mu < 1/2, \ \mathbb{P}_\mu(\hat{a}_\tau = +) \le \delta \quad \text{and} \quad \forall \mu > 1/2, \ \mathbb{P}_\mu(\hat{a}_\tau = -) \le \delta \, .$$

- a good procedure minimizes $\mathbb{E}_\mu[\tau]$ for all $\mu \ne 1/2$.

Entropic Lower Bound on the Sample Complexity

If $\mu > 1/2$:

$$
\begin{aligned}
\mathbb{E}_\mu[\tau]\,\mathrm{kl}\left(\mu,\frac{1}{2}-\epsilon\right) &= \mathrm{KL}\left(\mathbb{P}_\mu^{(X_1,\ldots,X_\tau)}, \mathbb{P}_{\frac{1}{2}-\epsilon}^{(X_1,\ldots,X_\tau)}\right) \qquad \text{\small tensorization works with stopping times}\\
&\geq \mathrm{KL}\left(\mathbb{P}_\mu^{\mathbb{1}\{\hat{a}_\tau=+\}}, \mathbb{P}_{\frac{1}{2}-\epsilon}^{\mathbb{1}\{\hat{a}_\tau=+\}}\right) \qquad \text{\small contraction just like before}\\
&= \mathrm{kl}\big(\mathbb{P}_\mu(\hat{a}_\tau=+), \mathbb{P}_{\frac{1}{2}-\epsilon}(\hat{a}_\tau=+)\big)\\
&\geq \mathrm{kl}(1-\delta,\delta) \qquad \text{\small by δ-correctness.}
\end{aligned}
$$

Theorem

For every δ-correct stopping time τ and every $\mu \in [0,1]$,

$$
\mathbb{E}_\mu[\tau] \geq \frac{1}{\mathrm{kl}\left(\mu,\frac{1}{2}\right)}\,\mathrm{kl}(1-\delta,\delta)\,.
$$

Remark: $\mathrm{kl}(\delta,1-\delta) \underset{\delta\to 0}{\sim} \log\frac{1}{\delta}$ and $\mathrm{kl}(\delta,1-\delta) \geq \log\frac{1}{2.4\delta}$.

Tight: possible to choose τ s.t. $\mathbb{E}_\mu[\tau] \leq \dfrac{\log\frac{1}{\delta}}{\mathrm{kl}\left(\mu,\frac{1}{2}\right)}\Big(1 + o_\delta\big(\log\frac{1}{\delta}\big)\Big).$

Active Learning: Finding the Distribution with Largest Mean

Best-Arm Identification with Fixed Confidence

K options $= \left(\mathcal{B}(\mu_a) \right)_{1 \le a \le K}$. Parameter $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K)$.

| μ_1 | μ_2 | μ_3 | μ_4 | μ_5 |

At round t, you may:

- choose an option $A_t = \phi_t \left(A_1, X_1, \ldots, A_{t-1}, X_{t-1} \right) \in \{1, \ldots, K\}$
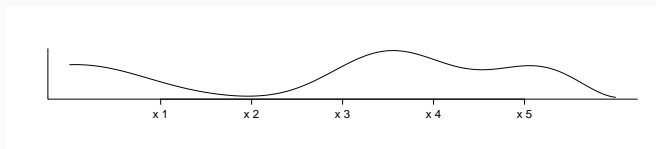- observe a new independent sample $X_t \sim \mathcal{B}(\mu_{A_t})$

so as to identify the best arm $a^*(\boldsymbol{\mu}) = \text{argmax}_a \ \mu_a$ as fast as possible: stopping time τ, decision $\hat{a}_\tau$.

Goal: minimize $\mathbb{E}_{\boldsymbol{\mu}}[\tau]$

under δ-correctness constraint: $\forall \boldsymbol{\mu}, \mathbb{P}_{\boldsymbol{\mu}} \left(\hat{a}_\tau \ne a^*(\boldsymbol{\mu}) \right) \le \delta$.

Best-Arm Identification with Fixed Confidence

K options $= \left(\mathcal{B}(\mu_a)\right)_{1 \leq a \leq K}$. Parameter $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K)$.

At round t, you may:

- choose an option $A_t = \phi_t\left(A_1, X_1, \ldots, A_{t-1}, X_{t-1}\right) \in \{1, \ldots, K\}$
- observe a new independent sample $X_t \sim \mathcal{B}\left(\mu_{A_t}\right)$

so as to identify the best arm $a^*(\boldsymbol{\mu}) = \text{argmax}_a \; \mu_a$ as fast as possible: stopping time τ, decision $\hat{a}_\tau$.

Goal: minimize $\mathbb{E}_{\boldsymbol{\mu}}[\tau]$
under δ-correctness constraint: $\forall \boldsymbol{\mu}, \mathbb{P}_{\boldsymbol{\mu}}\left(\hat{a}_\tau \neq a^*(\boldsymbol{\mu})\right) \leq \delta$.

$\mathcal{R} := \{1, \dots, K\}$ set of remaining arms.

$r := 0$ current round

while $|\mathcal{R}| > 1$

- $r := r + 1$
- draw each $a \in \mathcal{R}$, compute $\hat{\mu}_{a,r}$, the empirical mean of the r samples observed so far
- compute the empirical best and empirical worst arms:

$$b_r = \underset{a \in \mathcal{R}}{\operatorname{argmax}} \, \hat{\mu}_{a,r} \qquad w_r = \underset{a \in \mathcal{R}}{\operatorname{argmin}} \, \hat{\mu}_{a,r}$$
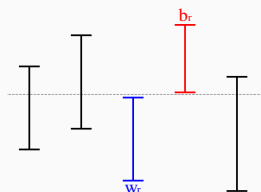
- Elimination step: if

$$\ell_{b_r}(r) > u_{w_r}(r),$$

then eliminate w_r : $\mathcal{R} := \mathcal{R} \backslash \{w_r\}$

end

Output: $\hat{a}$ the single element in $\mathcal{R}$.

10

Entropic Lower Bound

Let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K)$ and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_K)$ be two parameters with different maxima: $a^*(\boldsymbol{\mu}) \neq a^*(\boldsymbol{\lambda})$. Then

$$\sum_{a=1}^{K} \mathbb{E}_{\boldsymbol{\mu}}\big[N_a(\tau)\big]\,\mathrm{kl}(\mu_a, \lambda_a) = \mathrm{KL}\Big(\mathbb{P}_{\boldsymbol{\mu}}^{(X_1,\ldots,X_\tau)}, \mathbb{P}_{\boldsymbol{\lambda}}^{(X_1,\ldots,X_\tau)}\Big)$$

$$\geq \mathrm{KL}\big(\mathbb{P}_{\boldsymbol{\mu}}^{\mathbb{1}\{\hat{a}_\tau = a^*(\boldsymbol{\mu})\}}, \mathbb{P}_{\boldsymbol{\lambda}}^{\mathbb{1}\{\hat{a}_\tau = a^*(\boldsymbol{\mu})\}}\big)$$

$$\geq \mathrm{kl}\Big(\mathbb{P}_{\boldsymbol{\mu}}\big(\hat{a}_\tau = a^*(\boldsymbol{\mu})\big), \mathbb{P}_{\boldsymbol{\lambda}}\big(\hat{a}_\tau = a^*(\boldsymbol{\mu})\big)\Big)$$

$$\geq \mathrm{kl}(1 - \delta, \delta) \,.$$

Entropic Lower Bound

[Kaufmann, Cappé, G.'15],[G., Ménard, Stoltz '16]

For every δ-correct procedure, if $a^*(\boldsymbol{\mu}) \neq a^*(\boldsymbol{\lambda})$ then

$$\sum_{a=1}^{K} \mathbb{E}_{\boldsymbol{\mu}}\big[N_a(\tau)\big]\,\mathrm{kl}(\mu_a, \lambda_a) \geq \mathrm{kl}\big(1 - \delta, \delta\big) \,.$$

Using the Entropic Lower Bound

Let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_K)$ and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_K)$ be two bandit models.

Entropic Lower Bound

If $a^*(\boldsymbol{\mu}) \neq a^*(\boldsymbol{\lambda})$, any δ-correct algorithm satisfies

$$\sum_{a=1}^{K} \mathbb{E}_{\boldsymbol{\mu}}\left[N_a(\tau)\right] \mathrm{kl}(\mu_a, \lambda_a) \geq \mathrm{kl}(\delta, 1-\delta) .$$

Using it for each arm separately, one obtains:

Proposition [Kaufmann, Cappé, G.'15]
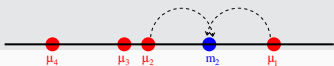
For any δ-correct algorithm,

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \geq \left(\frac{1}{\mathrm{kl}(\mu_1, \mu_2)} + \sum_{a=2}^{K} \frac{1}{\mathrm{kl}(\mu_a, \mu_1)} \right) \mathrm{kl}(\delta, 1-\delta) .$$

Entropic Lower Bound

If $a^*(\mu) \neq a^*(\lambda)$, any δ-correct algorithm satisfies

$$\sum_{a=1}^{K} \mathbb{E}_\mu \left[N_a(\tau) \right] \mathrm{kl}(\mu_a, \lambda_a) \geq \mathrm{kl}(\delta, 1 - \delta).$$

Let $\mathrm{Alt}(\mu) = \left\{ \lambda : a^*(\lambda) \neq a^*(\mu) \right\}$.

$$\inf_{\lambda \in \mathrm{Alt}(\mu)} \sum_{a=1}^{K} \mathbb{E}_\mu[N_a(\tau)] \, \mathrm{kl}(\mu_a, \lambda_a) \;\geq\; \mathrm{kl}(\delta, 1 - \delta)$$

$$\mathbb{E}_\mu[\tau] \times \inf_{\lambda \in \mathrm{Alt}(\mu)} \sum_{a=1}^{K} \frac{\mathbb{E}_\mu[N_a(\tau)]}{\mathbb{E}_\mu[\tau]} \, \mathrm{kl}(\mu_a, \lambda_a) \;\geq\; \mathrm{kl}(\delta, 1 - \delta)$$
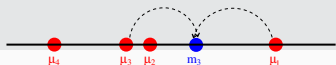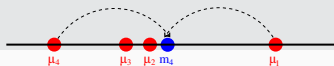
$$\mathbb{E}_\mu[\tau] \times \left(\sup_{w \in \Sigma_K} \inf_{\lambda \in \mathrm{Alt}(\mu)} \sum_{a=1}^{K} w_a \, \mathrm{kl}(\mu_a, \lambda_a) \right) \;\geq\; \mathrm{kl}(\delta, 1 - \delta)$$

Entropic Lower Bound

If $a^*(\boldsymbol{\mu}) \neq a^*(\boldsymbol{\lambda})$, any δ-correct algorithm satisfies

$$\sum_{a=1}^{K} \mathbb{E}_{\boldsymbol{\mu}}\left[N_a(\tau)\right] \mathrm{kl}(\mu_a, \lambda_a) \geq \mathrm{kl}(\delta, 1 - \delta).$$

Let $\mathrm{Alt}(\boldsymbol{\mu}) = \{\boldsymbol{\lambda} : a^*(\boldsymbol{\lambda}) \neq a^*(\boldsymbol{\mu})\}$.

$$\inf_{\boldsymbol{\lambda} \in \mathrm{Alt}(\boldsymbol{\mu})} \sum_{a=1}^{K} \mathbb{E}_{\boldsymbol{\mu}}[N_a(\tau)] \, \mathrm{kl}(\mu_a, \lambda_a) \;\geq\; \mathrm{kl}(\delta, 1 - \delta)$$

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \times \inf_{\boldsymbol{\lambda} \in \mathrm{Alt}(\boldsymbol{\mu})} \sum_{a=1}^{K} \frac{\mathbb{E}_{\boldsymbol{\mu}}[N_a(\tau)]}{\mathbb{E}_{\boldsymbol{\mu}}[\tau]} \, \mathrm{kl}(\mu_a, \lambda_a) \;\geq\; \mathrm{kl}(\delta, 1 - \delta)$$

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \times \left(\sup_{w \in \Sigma_K} \inf_{\boldsymbol{\lambda} \in \mathrm{Alt}(\boldsymbol{\mu})} \sum_{a=1}^{K} w_a \, \mathrm{kl}(\mu_a, \lambda_a) \right) \;\geq\; \mathrm{kl}(\delta, 1 - \delta)$$

Entropic Lower Bound

If $a^*(\boldsymbol{\mu}) \neq a^*(\boldsymbol{\lambda})$, any δ-correct algorithm satisfies

$$\sum_{a=1}^{K} \mathbb{E}_{\boldsymbol{\mu}}\left[N_a(\tau)\right] \mathrm{kl}(\mu_a, \lambda_a) \geq \mathrm{kl}(\delta, 1-\delta).$$

Let $\mathrm{Alt}(\boldsymbol{\mu}) = \left\{\boldsymbol{\lambda} : a^*(\boldsymbol{\lambda}) \neq a^*(\boldsymbol{\mu})\right\}$.

$$\inf_{\boldsymbol{\lambda} \in \mathrm{Alt}(\boldsymbol{\mu})} \sum_{a=1}^{K} \mathbb{E}_{\boldsymbol{\mu}}[N_a(\tau)] \, \mathrm{kl}(\mu_a, \lambda_a) \;\geq\; \mathrm{kl}(\delta, 1-\delta)$$

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \times \inf_{\boldsymbol{\lambda} \in \mathrm{Alt}(\boldsymbol{\mu})} \sum_{a=1}^{K} \frac{\mathbb{E}_{\boldsymbol{\mu}}[N_a(\tau)]}{\mathbb{E}_{\boldsymbol{\mu}}[\tau]} \, \mathrm{kl}(\mu_a, \lambda_a) \;\geq\; \mathrm{kl}(\delta, 1-\delta)$$

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \times \left(\sup_{w \in \Sigma_K} \inf_{\boldsymbol{\lambda} \in \mathrm{Alt}(\boldsymbol{\mu})} \sum_{a=1}^{K} w_a \, \mathrm{kl}(\mu_a, \lambda_a) \right) \;\geq\; \mathrm{kl}(\delta, 1-\delta)$$

13

Entropic Lower Bound

If $a^*(\mu) \neq a^*(\lambda)$, any δ-correct algorithm satisfies

$$\sum_{a=1}^{K} \mathbb{E}_{\mu}\left[N_a(\tau)\right] \mathrm{kl}(\mu_a, \lambda_a) \geq \mathrm{kl}(\delta, 1 - \delta).$$

Let $\mathrm{Alt}(\mu) = \left\{\lambda : a^*(\lambda) \neq a^*(\mu)\right\}$.

$$\inf_{\lambda \in \mathrm{Alt}(\mu)} \sum_{a=1}^{K} \mathbb{E}_{\mu}[N_a(\tau)] \, \mathrm{kl}(\mu_a, \lambda_a) \;\geq\; \mathrm{kl}(\delta, 1 - \delta)$$

$$\mathbb{E}_{\mu}[\tau] \times \inf_{\lambda \in \mathrm{Alt}(\mu)} \sum_{a=1}^{K} \frac{\mathbb{E}_{\mu}[N_a(\tau)]}{\mathbb{E}_{\mu}[\tau]} \, \mathrm{kl}(\mu_a, \lambda_a) \;\geq\; \mathrm{kl}(\delta, 1 - \delta)$$

$$\mathbb{E}_{\mu}[\tau] \times \left(\sup_{w \in \Sigma_K} \inf_{\lambda \in \mathrm{Alt}(\mu)} \sum_{a=1}^{K} w_a \, \mathrm{kl}(\mu_a, \lambda_a)\right) \;\geq\; \mathrm{kl}(\delta, 1 - \delta)$$

Lower Bound: the Complexity of BAI

Theorem [G. and Kaufmann, 2016]

For any δ-correct algorithm,

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \geq T^*(\boldsymbol{\mu})\,\mathrm{kl}(\delta, 1-\delta)\,,$$

where

$$T^*(\boldsymbol{\mu})^{-1} = \sup_{w \in \Sigma_K} \inf_{\lambda \in \mathrm{Alt}(\boldsymbol{\mu})} \sum_{a=1}^{K} w_a\,\mathrm{kl}(\mu_a, \lambda_a)$$

$$= \max_{w \in \Sigma_K} \min_{b \neq a^*} \inf_{\mu_b \leq \lambda \leq \mu_{a^*}} w_{a^*}\,\mathrm{kl}(\mu_{a^*}, \lambda) + w_b\,\mathrm{kl}(\mu_b, \lambda)\,.$$

- A kind of game : you choose the proportions of draws $(w_a)_a$, the opponent chooses the alternative.
➙ the optimal proportions of arm draws are

$$w^*(\boldsymbol{\mu}) = \operatorname*{argmax}_{w \in \Sigma_K} \inf_{\lambda \in \mathrm{Alt}(\boldsymbol{\mu})} \sum_{a=1}^{K} w_a\,\mathrm{kl}(\mu_a, \lambda_a)\,.$$

14

What if I want only an ϵ-optimal distribution?

Now, at round t, you may:

- choose an option $A_t = \phi_t(A_1, X_1, \ldots, A_{t-1}, X_{t-1}) \in \{1, \ldots, K\}$
- observe a new independent sample $X_t \sim \nu_{A_t}$

so as to identify any ϵ-optimal $a \in \mathcal{A}_\epsilon = \{a : \mu_a \geq \mu^* - \epsilon\}$ where $\mu^* = \max_a \mu_a$ as fast as possible: stopping time $\tau_{\delta,\epsilon}$.

$\implies$ minimize $\mathbb{E}[\tau_{\delta,\epsilon}]$ under the PAC constraint $\mathbb{P}_\mu(\mu_{\hat{a}_\tau} < \mu^* - \epsilon) \leq \delta$.

- PAC constraint: Probably Approximately Correct;
- more natural objective, especially in the context of discretized optimization;
- permits to avoid infinite loops in case of draws.

Apparently more complicated

Theorem [G. and Kaufmann'19]

For any ϵ, δ-PAC algorithm with converging proportions of draws,

$$\liminf_{\delta \to 0} \frac{\mathbb{E}_{\boldsymbol{\mu}}[\tau]}{\log \frac{1}{\delta}} \geq T_{\epsilon}^*(\boldsymbol{\mu}),$$

where

$$T_{\epsilon}^*(\boldsymbol{\mu})^{-1} = \sup_{w \in \Sigma_K} \max_{a \in \mathcal{A}_{\epsilon}} \min_{b \neq a} \inf_{(\lambda_a, \lambda_b): \lambda_a \leq \lambda_b - \epsilon} w_a \, \mathrm{kl}(\mu_a, \lambda_a) + w_b \, \mathrm{kl}(\mu_b, \lambda_b) \,.$$

- Asymptotic result (only).
- Assumption on the algorithm (convergence of the proportions of draws) – can be dropped.
- We do not manage to use the information-theoretic technique! We have to go back to the change of measure... even for the sequential test with multiple correct answers!