Bayesian Lower Bounds for Regret Minimization

Alessio Russo

June 2025

Boston University

Overview i

Introduction

Paper: Regret Lower Bound [Lai, 1987]

The Annals of Statistics
1987, Vol. 15, No. 3, 1091–1114

ADAPTIVE TREATMENT ALLOCATION AND THE MULTI-ARMED BANDIT PROBLEM[1]

BY TZE-LEUNG LAI

Columbia University

A class of simple adaptive allocation rules is proposed for the problem (often called the "multi-armed bandit problem") of sampling $x_1, \dots, x_N$ sequentially from k populations with densities belonging to an exponential family, in order to maximize the expected value of the sum $S_N = x_1 + \cdots + x_N$. These allocation rules are based on certain upper confidence bounds, which are developed from boundary crossing theory, for the k population parameters. The rules are shown to be asymptotically optimal as $N \to \infty$ from both Bayesian and frequentist points of view. Monte Carlo studies show that they also perform very well for moderate values of the horizon N.

Motivation

Today we discuss unstructured multi-armed bandit problems (MAB). We will talk about:

1. Deriving asymptotic instance-dependent regret lower bounds in the Bayesian setting [Lai, 1987, Atsidakou et al., 2023]).
2. Possible extensions and some considerations on this topic.

Motivation

Today we discuss unstructured multi-armed bandit problems (MAB). We will talk about:

1. Deriving asymptotic instance-dependent regret lower bounds in the Bayesian setting [Lai, 1987, Atsidakou et al., 2023]).
2. Possible extensions and some considerations on this topic.

Model and Assumptions

Consider a MAB problem with $a = 1, \dots, K$ arms:

- **Sequential**: In round t the learner pulls arm $a_t \in [K]$ and receives the reward $r_t \sim \theta_{a_t}$.
  - Each arm a is characterized by a density function $f(r; \theta_a)$ with respect to the Lebesgue measure, and $\theta_a \in \Theta$ is an unknown parameter that belongs to some open set $\Theta \subset \mathbb{R}$.
  - (Integrability) For all $\theta \in \Theta$ we assume that $\mu_a(\theta) := \mathbb{E}_{\theta_a}[|R|] = \int_{\mathbb{R}} |r| f(r; \theta_a) \mathrm{d}r < \infty$.
- (Bayesian Prior) We denote by $H = (H_1, \dots, H_K)$ a factorized prior distribution on Θ^K, with density $h(\theta) = \prod_a h_a(\theta_a)$ (note that each h_a may be different).
- We indicate by $\mu^\star(\theta) := \max_a \mu_a(\theta)$ the value of the best arm.

In Bayesian analysis, also the model θ is a random variable.

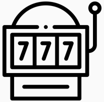
Model and Assumptions

Consider a MAB problem with $a = 1, \dots, K$ arms:

- **Sequential**: In round t the learner pulls arm $a_t \in [K]$ and receives the reward $r_t \sim \theta_{a_t}$.
  - Each arm a is characterized by a density function $f(r; \theta_a)$ with respect to the Lebesgue measure, and $\theta_a \in \Theta$ is an unknown parameter that belongs to some open set $\Theta \subset \mathbb{R}$.
  - (Integrability) For all $\theta \in \Theta$ we assume that $\mu_a(\theta) := \mathbb{E}_{\theta_a}[|R|] = \int_{\mathbb{R}} |r| f(r; \theta_a) \mathrm{d}r < \infty$.
- (Bayesian Prior) We denote by $H = (H_1, \dots, H_K)$ a factorized prior distribution on Θ^K, with density $h(\theta) = \prod_a h_a(\theta_a)$ (note that each h_a may be different).
- We indicate by $\mu^\star(\theta) := \max_a \mu_a(\theta)$ the value of the best arm.

In Bayesian analysis, also the model θ is a random variable.

Model and Assumptions



Consider a MAB problem with $a = 1, \dots, K$ arms:


- **Sequential**: In round t the learner pulls arm $a_t \in [K]$ and receives the reward $r_t \sim \theta_{a_t}$.
  - Each arm a is characterized by a density function $f(r; \theta_a)$ with respect to the Lebesgue measure, and $\theta_a \in \Theta$ is an unknown parameter that belongs to some open set $\Theta \subset \mathbb{R}$.
  - (Integrability) For all $\theta \in \Theta$ we assume that $\mu_a(\theta) := \mathbb{E}_{\theta_a}[|R|] = \int_{\mathbb{R}} |r| f(r; \theta_a) \mathrm{d}r < \infty$.
- (Bayesian Prior) We denote by $H = (H_1, \dots, H_K)$ a factorized prior distribution on Θ^K, with density $h(\theta) = \prod_a h_a(\theta_a)$ (note that each h_a may be different).
- We indicate by $\mu^\star(\theta) := \max_a \mu_a(\theta)$ the value of the best arm.

In Bayesian analysis, also the model θ is a random variable.

Model and Assumptions

Consider a MAB problem with $a = 1, \dots, K$ arms:

- **Sequential**: In round t the learner pulls arm $a_t \in [K]$ and receives the reward $r_t \sim \theta_{a_t}$.
  - Each arm a is characterized by a density function $f(r; \theta_a)$ with respect to the Lebesgue measure, and $\theta_a \in \Theta$ is an unknown parameter that belongs to some open set $\Theta \subset \mathbb{R}$.
  - (Integrability) For all $\theta \in \Theta$ we assume that $\mu_a(\theta) := \mathbb{E}_{\theta_a}[|R|] = \int_{\mathbb{R}} |r| f(r; \theta_a) \mathrm{d}r < \infty$.
- (Bayesian Prior) We denote by $H = (H_1, \dots, H_K)$ a factorized prior distribution on Θ^K, with density $h(\theta) = \prod_a h_a(\theta_a)$ (note that each h_a may be different).
- We indicate by $\mu^\star(\theta) := \max_a \mu_a(\theta)$ the value of the best arm.

In Bayesian analysis, also the model θ is a random variable.

Problems Considered

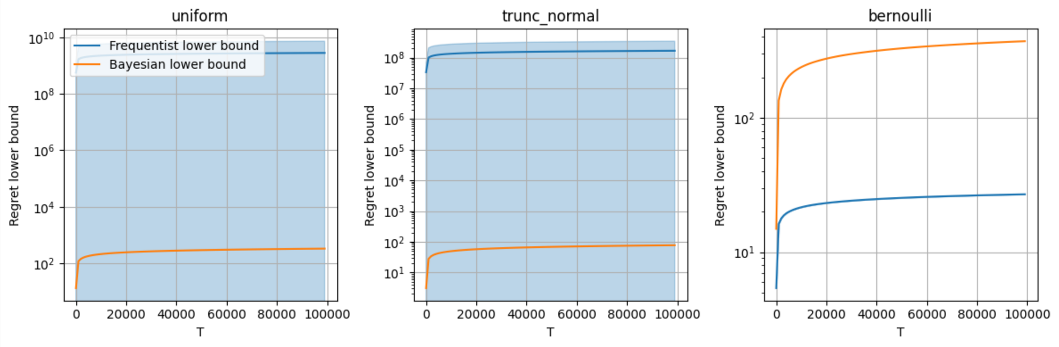




- **Regret Minimization**: Minimize the regret incurred in not choosing the best arm in each time-step over an horizon T.
- **Best Arm Identification objective**: quickly find the optimal arm with confidence at-least $1-\delta, \delta \in (0, 1/2)$.

Bayesian vs Frequentist Regret Lower Bound

- Can we just compute the average frequentist lower bound over many different problems?



Average computed over 3000000 sampled MAB problems. Shaded areas indicate the 95% C.I.

The frequentist lower bound simply explodes with continuous priors.

We consider $f(r;\theta_a)$to belong to the single-parameter exponential family[1]

$$f_a(r;\theta) = \exp(\theta_a r - \psi(\theta_a)),$$

where $\psi(\theta_a)$ is the cumulant generating function[2].

- θ_a is called natural parameter.
- $\dot{\psi}(\theta_a) = \frac{\mathrm{d}\psi}{\mathrm{d}\theta_a} = \mu_a(\theta)$ is the mean value and $\ddot{\psi}(\theta_a) = \mathbb{E}_{r\sim\theta_a}[(r-\dot{\psi}(\theta_a))^2]$ is the variance. We also have that $\dot{\psi}$ is increasing in θ_a, and we let $\theta^\star = \max_a \theta_a$.
- Kullback-Leibler (KL) Divergence defined as

$$D(\theta_a,\theta_a') = (\theta_a - \theta_a')\dot{\psi}(\theta_a) - (\psi(\theta_a) - \psi(\theta_a')) \left(= \int_{\theta_a}^{\theta_a'} (t-\theta_a)\ddot{\psi}(t)\mathrm{d}t\right)$$

[1] Includes Bernoulli, Poisson, Gaussian distribution with known variance, etc. [Efron, 2022].

[2] For a r.v. X the cumulant ψ is defined as $\log \mathbb{E}[e^X]$.

Single-parameter exponential family

We consider $f(r;\theta_a)$to belong to the single-parameter exponential family[1]

$$f_a(r;\theta) = \exp(\theta_a r - \psi(\theta_a)),$$

where $\psi(\theta_a)$ is the cumulant generating function[2].

- θ_a is called natural parameter.
- $\dot{\psi}(\theta_a) = \frac{d\psi}{d\theta_a} = \mu_a(\theta)$ is the mean value and $\ddot{\psi}(\theta_a) = \mathbb{E}_{r\sim\theta_a}[(r - \dot{\psi}(\theta_a))^2]$ is the variance. We also have that $\dot{\psi}$ is increasing in θ_a, and we let $\theta^\star = \max_a \theta_a$.
- Kullback-Leibler (KL) Divergence defined as

$$D(\theta_a, \theta'_a) = (\theta_a - \theta'_a)\dot{\psi}(\theta_a) - (\psi(\theta_a) - \psi(\theta'_a)) \left(= \int_{\theta_a}^{\theta'_a} (t - \theta_a)\ddot{\psi}(t)\mathrm{d}t \right)$$

[1] Includes Bernoulli, Poisson, Gaussian distribution with known variance, etc. [Efron, 2022].

[2] For a r.v. X the cumulant ψ is defined as $\log \mathbb{E}[e^X]$.

Single-parameter exponential family

We consider $f(r;\theta_a)$to belong to the single-parameter exponential family[1]

$$f_a(r;\theta) = \exp(\theta_a r - \psi(\theta_a)),$$

where $\psi(\theta_a)$ is the cumulant generating function[2].

- θ_a is called natural parameter.
- $\dot{\psi}(\theta_a) = \frac{\mathrm{d}\psi}{\mathrm{d}\theta_a} = \mu_a(\theta)$ is the mean value and $\ddot{\psi}(\theta_a) = \mathbb{E}_{r\sim\theta_a}[(r - \dot{\psi}(\theta_a))^2]$ is the variance. We also have that $\dot{\psi}$ is increasing in θ_a, and we let $\theta^\star = \max_a \theta_a$.
- Kullback-Leibler (KL) Divergence defined as

$$D(\theta_a, \theta'_a) = (\theta_a - \theta'_a)\dot{\psi}(\theta_a) - (\psi(\theta_a) - \psi(\theta'_a)) \left(= \int_{\theta_a}^{\theta'_a} (t - \theta_a)\ddot{\psi}(t)\mathrm{d}t \right)$$

[1] Includes Bernoulli, Poisson, Gaussian distribution with known variance, etc. [Efron, 2022].

[2] For a r.v. X the cumulant ψ is defined as $\log \mathbb{E}[e^X]$.

Single-parameter exponential family

We consider $f(r;\theta_a)$to belong to the single-parameter exponential family[1]

$$f_a(r;\theta) = \exp(\theta_a r - \psi(\theta_a)),$$

where $\psi(\theta_a)$ is the cumulant generating function[2].

- θ_a is called natural parameter.
- $\dot{\psi}(\theta_a) = \frac{\mathrm{d}\psi}{\mathrm{d}\theta_a} = \mu_a(\theta)$ is the mean value and $\ddot{\psi}(\theta_a) = \mathbb{E}_{r\sim\theta_a}[(r - \dot{\psi}(\theta_a))^2]$ is the variance. We also have that $\dot{\psi}$ is increasing in θ_a, and we let $\theta^\star = \max_a \theta_a$.
- Kullback-Leibler (KL) Divergence defined as

$$D(\theta_a, \theta_a') = (\theta_a - \theta_a')\dot{\psi}(\theta_a) - (\psi(\theta_a) - \psi(\theta_a'))\left(= \int_{\theta_a}^{\theta_a'} (t - \theta_a)\ddot{\psi}(t)\mathrm{d}t\right)$$

[1] Includes Bernoulli, Poisson, Gaussian distribution with known variance, etc. [Efron, 2022].

[2] For a r.v. X the cumulant ψ is defined as $\log \mathbb{E}[e^X]$.

Properties for single-parameter exponential distributions

Observation. For the single-parameter exponential distribution family we have that for $y > x$, with $x, y \in [a, b]$, $\exists c_1, c_2 > 0$ such that

$$c_1(y-x) \geq \underbrace{[\dot{\psi}(y) - \dot{\psi}(x)]}_{\text{mean values}} \geq 2c_2 \frac{D(x,y)}{(y-x)} > 0.$$

Proof.

1. By the mean value theorem $\exists \xi \in (x, y)$: $\dot{\psi}(y) - \dot{\psi}(x) = \ddot{\psi}(\xi)(y-x) \geq \min_{z\in[a,b]} \ddot{\psi}(z)(y-x)$ (the upper bound follows similarly).
2. Then, recall $D(x,y) = (x-y)\dot{\psi}(x) - (\psi(x) - \psi(y)) = \int_x^y (t-x)\ddot{\psi}(t)\mathrm{d}t$.
3. We use that $\dot{\psi}$ is increasing and differentiable $\Rightarrow \ddot{\psi} > 0$. Thus $D(x,y) \leq \frac{\max_{z\in[a,b]} \ddot{\psi}(z)}{2}(x-y)^2$.

$$\dot{\psi}(y) - \dot{\psi}(x) = \ddot{\psi}(\xi)(y-x) \geq \ddot{\psi}(x)(y-x) \geq \frac{\min_{z_1\in[a,b]} \ddot{\psi}(z_1)}{\max_{z_2\in[a,b]} \ddot{\psi}(z_2)} \cdot \frac{2D(x,y)}{(y-x)}.$$

Properties for single-parameter exponential distributions

Observation. For the single-parameter exponential distribution family we have that for $y > x$, with $x, y \in [a,b]$, $\exists c_1, c_2 > 0$ such that

$$c_1(y-x) \geq \underbrace{[\dot{\psi}(y) - \dot{\psi}(x)]}_{\text{mean values}} \geq 2c_2 \frac{D(x,y)}{(y-x)} > 0.$$

Proof.

1. By the mean value theorem $\exists \xi \in (x,y)$: $\dot{\psi}(y) - \dot{\psi}(x) = \ddot{\psi}(\xi)(y-x) \geq \min_{z \in [a,b]} \ddot{\psi}(z)(y-x)$ (the upper bound follows similarly).
2. Then, recall $D(x,y) = (x-y)\dot{\psi}(x) - (\psi(x) - \psi(y)) = \int_x^y (t-x)\ddot{\psi}(t)\mathrm{d}t$.
3. We use that $\dot{\psi}$ is increasing and differentiable $\Rightarrow \ddot{\psi} > 0$. Thus $D(x,y) \leq \frac{\max_{z\in[a,b]} \ddot{\psi}(z)}{2}(x-y)^2$.

$$\dot{\psi}(y) - \dot{\psi}(x) = \ddot{\psi}(\xi)(y-x) \geq \ddot{\psi}(x)(y-x) \geq \frac{\min_{z_1\in[a,b]} \ddot{\psi}(z_1)}{\max_{z_2\in[a,b]} \ddot{\psi}(z_2)} \cdot \frac{2D(x,y)}{(y-x)}.$$

Properties for single-parameter exponential distributions

Observation. For the single-parameter exponential distribution family we have that for $y > x$, with $x, y \in [a, b]$, $\exists c_1, c_2 > 0$ such that

$$c_1(y - x) \geq \underbrace{[\dot{\psi}(y) - \dot{\psi}(x)]}_{\text{mean values}} \geq 2c_2 \frac{D(x, y)}{(y - x)} > 0.$$

Proof.

1. By the mean value theorem $\exists \xi \in (x, y)$: $\dot{\psi}(y) - \dot{\psi}(x) = \ddot{\psi}(\xi)(y - x) \geq \min_{z \in [a,b]} \ddot{\psi}(z)(y - x)$ (the upper bound follows similarly).
2. Then, recall $D(x, y) = (x - y)\dot{\psi}(x) - (\psi(x) - \psi(y)) = \int_x^y (t - x)\ddot{\psi}(t)\mathrm{d}t$.
3. We use that $\dot{\psi}$ is increasing and differentiable $\Rightarrow \ddot{\psi} > 0$. Thus $D(x, y) \leq \frac{\max_{z \in [a,b]} \ddot{\psi}(z)}{2}(x - y)^2$.

$$\dot{\psi}(y) - \dot{\psi}(x) = \ddot{\psi}(\xi)(y - x) \geq \ddot{\psi}(x)(y - x) \geq \frac{\min_{z_1 \in [a,b]} \ddot{\psi}(z_1)}{\max_{z_2 \in [a,b]} \ddot{\psi}(z_2)} \cdot \frac{2D(x, y)}{(y - x)}.$$

Properties for single-parameter exponential distributions

Observation. For the single-parameter exponential distribution family we have that for $y > x$, with $x, y \in [a, b]$, $\exists c_1, c_2 > 0$ such that

$$c_1(y-x) \geq \underbrace{[\dot{\psi}(y) - \dot{\psi}(x)]}_{\text{mean values}} \geq 2c_2 \frac{D(x,y)}{(y-x)} > 0.$$

Proof.

1. By the mean value theorem $\exists \xi \in (x, y)$: $\dot{\psi}(y) - \dot{\psi}(x) = \ddot{\psi}(\xi)(y-x) \geq \min_{z \in [a,b]} \ddot{\psi}(z)(y-x)$ (the upper bound follows similarly).
2. Then, recall $D(x,y) = (x-y)\dot{\psi}(x) - (\psi(x) - \psi(y)) = \int_x^y (t-x)\ddot{\psi}(t)\mathrm{d}t$.
3. We use that $\dot{\psi}$ is increasing and differentiable $\Rightarrow \ddot{\psi} > 0$. Thus $D(x,y) \leq \frac{\max_{z \in [a,b]} \ddot{\psi}(z)}{2}(x-y)^2$.

$$\dot{\psi}(y) - \dot{\psi}(x) = \ddot{\psi}(\xi)(y-x) \geq \ddot{\psi}(x)(y-x) \geq \frac{\min_{z_1 \in [a,b]} \ddot{\psi}(z_1)}{\max_{z_2 \in [a,b]} \ddot{\psi}(z_2)} \cdot \frac{2D(x,y)}{(y-x)}.$$

Properties for single-parameter exponential distributions

Observation. For the single-parameter exponential distribution family we have that for $y > x$, with $x, y \in [a, b]$, $\exists c_1, c_2 > 0$ such that

$$c_1(y - x) \geq \underbrace{[\dot{\psi}(y) - \dot{\psi}(x)]}_{\text{mean values}} \geq 2c_2 \frac{D(x, y)}{(y - x)} > 0.$$

Proof.

1. By the mean value theorem $\exists \xi \in (x, y)$: $\dot{\psi}(y) - \dot{\psi}(x) = \ddot{\psi}(\xi)(y - x) \geq \min_{z \in [a,b]} \ddot{\psi}(z)(y - x)$ (the upper bound follows similarly).
2. Then, recall $D(x, y) = (x - y)\dot{\psi}(x) - (\psi(x) - \psi(y)) = \int_x^y (t - x)\ddot{\psi}(t)\mathrm{d}t$.
3. We use that $\dot{\psi}$ is increasing and differentiable $\Rightarrow \ddot{\psi} > 0$. Thus $D(x, y) \leq \frac{\max_{z \in [a,b]} \ddot{\psi}(z)}{2}(x - y)^2$.

$$\dot{\psi}(y) - \dot{\psi}(x) = \ddot{\psi}(\xi)(y - x) \geq \ddot{\psi}(x)(y - x) \geq \frac{\min_{z_1 \in [a,b]} \ddot{\psi}(z_1)}{\max_{z_2 \in [a,b]} \ddot{\psi}(z_2)} \cdot \frac{2D(x, y)}{(y - x)}.$$

Regret Minimization: lower bound

Sampling rule or allocation rule

Denote by $\pi = (\pi_t)_{t \geq 1}$ the sampling rule of the learner (a.k.a. allocation rule). Concretely

- π is a sequence of measurable functions, each of which associates past data with an arm, namely

$$a_{t+1} = \pi_t(I_t), \quad \text{where } I_t = (u_0, a_1, r_1, u_1, \ldots, a_t, r_t, u_t), \quad I_0 = u_0.$$

  and $(u_t)_{t \geq 0}$ is a sequence of iid uniform noise, such that u_t is independent of I_{t-1} and (a_t, r_t). Thus $a_t \in \mathcal{F}_{t-1} := \sigma(I_{t-1})$.
- Let $N_a(t) = \sum_{n=1}^{t} \mathbf{1}_{\{a_t = a\}}$ be the number of times we selected arm a up to time t.

Sampling rule or allocation rule

Denote by $\pi = (\pi_t)_{t\geq 1}$ the sampling rule of the learner (a.k.a. allocation rule). Concretely

- π is a sequence of measurable functions, each of which associates past data with an arm, namely

$$a_{t+1} = \pi_t(I_t), \quad \text{where } I_t = (u_0, a_1, r_1, u_1, \ldots, a_t, r_t, u_t), \quad I_0 = u_0.$$

and $(u_t)_{t\geq 0}$ is a sequence of iid uniform noise, such that u_t is independent of I_{t-1} and (a_t, r_t). Thus $a_t \in \mathcal{F}_{t-1} := \sigma(I_{t-1})$.

- Let $N_a(t) = \sum_{n=1}^{t} \mathbf{1}_{\{a_t=a\}}$ be the number of times we selected arm a up to time t.

Sampling rule or allocation rule

Denote by $\pi = (\pi_t)_{t\geq 1}$ the sampling rule of the learner (a.k.a. allocation rule). Concretely

- π is a sequence of measurable functions, each of which associates past data with an arm, namely

$$a_{t+1} = \pi_t(I_t), \quad \text{where } I_t = (u_0, a_1, r_1, u_1, \dots, a_t, r_t, u_t), \quad I_0 = u_0.$$

  and $(u_t)_{t\geq 0}$ is a sequence of iid uniform noise, such that u_t is independent of I_{t-1} and (a_t, r_t). Thus $a_t \in \mathcal{F}_{t-1} \coloneqq \sigma(I_{t-1})$.

- Let $N_a(t) = \sum_{n=1}^{t} \mathbf{1}_{\{a_t = a\}}$ be the number of times we selected arm a up to time t.

Regret Minimization

For a fixed $\theta \in \Theta^K$, define the regret at time T as:

$$\begin{aligned}
\mathrm{Reg}(T;\theta) &= \mathbb{E}_\theta\left[\sum_{t=1}^{T} \mu^\star(\theta) - \mu_{a_t}(\theta)\right], \\
&= T\mu^\star(\theta) - \sum_a \mathbb{E}_\theta\left[\mu_a(\theta) N_a(T)\right], \\
&= \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Bigg[\underbrace{\mu^\star(\theta) - \mu_a(\theta)}_{=:\Delta_a(\theta)}\Bigg] \mathbb{E}_\theta[N_a(T)], && \triangleright \text{ Use that } T = \textstyle\sum_a N_a(T) \\
&= \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)].
\end{aligned}$$

Bayesian Regret (Bayes Risk): $\mathrm{Reg}(T) = \mathbb{E}_{\theta\sim H}[\mathrm{Reg}(T;\theta)] = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}\underbrace{H(\theta)}_{\text{prior}}$.

Regret Minimization

For a fixed $\theta \in \Theta^K$, define the regret at time T as:

$$\begin{aligned}
\mathrm{Reg}(T;\theta) &= \mathbb{E}_\theta\left[\sum_{t=1}^{T} \mu^\star(\theta) - \mu_{a_t}(\theta)\right], \\
&= T\mu^\star(\theta) - \sum_a \mathbb{E}_\theta\left[\mu_a(\theta) N_a(T)\right], \\
&= \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \left[\underbrace{\mu^\star(\theta) - \mu_a(\theta)}_{=:\Delta_a(\theta)}\right] \mathbb{E}_\theta[N_a(T)], \qquad \triangleright \text{ Use that } T = \textstyle\sum_a N_a(T) \\
&= \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)].
\end{aligned}$$

Bayesian Regret (Bayes Risk): $\mathrm{Reg}(T) = \mathbb{E}_{\theta\sim H}[\mathrm{Reg}(T;\theta)] = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}\underbrace{H(\theta)}_{\text{prior}}$.

Regret Minimization

For a fixed $\theta \in \Theta^K$, define the regret at time T as:

$$\begin{aligned}
\mathrm{Reg}(T;\theta) &= \mathbb{E}_\theta\left[\sum_{t=1}^{T} \mu^\star(\theta) - \mu_{a_t}(\theta)\right], \\
&= T\mu^\star(\theta) - \sum_a \mathbb{E}_\theta\left[\mu_a(\theta) N_a(T)\right], \\
&= \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \left[\underbrace{\mu^\star(\theta) - \mu_a(\theta)}_{=:\Delta_a(\theta)}\right] \mathbb{E}_\theta[N_a(T)], \qquad \triangleright \text{ Use that } T = \textstyle\sum_a N_a(T) \\
&= \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)].
\end{aligned}$$

Bayesian Regret (Bayes Risk): $\mathrm{Reg}(T) = \mathbb{E}_{\theta\sim H}[\mathrm{Reg}(T;\theta)] = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}\underbrace{H(\theta)}_{\text{prior}}$.

Regret Minimization

For a fixed $\theta \in \Theta^K$, define the regret at time T as:

$$
\begin{aligned}
\mathrm{Reg}(T;\theta) &= \mathbb{E}_\theta\left[\sum_{t=1}^{T} \mu^\star(\theta) - \mu_{a_t}(\theta)\right], \\
&= T\mu^\star(\theta) - \sum_a \mathbb{E}_\theta\left[\mu_a(\theta) N_a(T)\right], \\
&= \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \left[\underbrace{\mu^\star(\theta) - \mu_a(\theta)}_{=:\Delta_a(\theta)}\right] \mathbb{E}_\theta[N_a(T)], \qquad \triangleright \text{ Use that } T = \textstyle\sum_a N_a(T) \\
&= \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)].
\end{aligned}
$$

Bayesian Regret (Bayes Risk): $\mathrm{Reg}(T) = \mathbb{E}_{\theta\sim H}[\mathrm{Reg}(T;\theta)] = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}\underbrace{H(\theta)}_{\text{prior}}$.

Regret Minimization

For a fixed $\theta \in \Theta^K$, define the regret at time T as:

$$\begin{aligned}
\mathrm{Reg}(T;\theta) &= \mathbb{E}_\theta\left[\sum_{t=1}^{T} \mu^\star(\theta) - \mu_{a_t}(\theta)\right], \\
&= T\mu^\star(\theta) - \sum_a \mathbb{E}_\theta\left[\mu_a(\theta) N_a(T)\right], \\
&= \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Bigg[\underbrace{\mu^\star(\theta) - \mu_a(\theta)}_{=:\Delta_a(\theta)}\Bigg] \mathbb{E}_\theta[N_a(T)], \qquad \triangleright \text{ Use that } T = \textstyle\sum_a N_a(T) \\
&= \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)].
\end{aligned}$$

Bayesian Regret (Bayes Risk): $\mathrm{Reg}(T) = \mathbb{E}_{\theta\sim H}[\mathrm{Reg}(T;\theta)] = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}\underbrace{H(\theta)}_{\text{prior}}$.

Bayesian Regret vs Frequentist Regret

Bayesian Regret $\text{Reg}(T) = \int_{\Theta^K} \text{Reg}(T;\theta)\mathrm{d}H(\theta)$ **vs** Frequentist Regret $\text{Reg}(T;\theta)$

The problem is: do we consider the model θ fixed or not? Why should we average over H if the problem is fixed in reality?[3]

- Note that people have applied Bayesian algorithms to the Frequentist regret: Posterior sampling helps with exploration (epistemic uncertainty).
- Similarly, UCB-designs have been used to find rules that are efficient in the Bayesian regret sense.
- In this work we study the Bayesian regret.

[3]Some people claim that uncertainty in the model can be seen as uncertainty in future data. Personally, I lean more towards the frequentist view.

Bayesian Regret vs Frequentist Regret

Bayesian Regret $\mathrm{Reg}(T) = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta)$ **vs** Frequentist Regret $\mathrm{Reg}(T;\theta)$

The problem is: do we consider the model θ fixed or not? Why should we average over H if the problem is fixed in reality?[3]

- Note that people have applied Bayesian algorithms to the Frequentist regret: Posterior sampling helps with exploration (epistemic uncertainty).
- Similarly, UCB-designs have been used to find rules that are efficient in the Bayesian regret sense.
- In this work we study the Bayesian regret.

[3] Some people claim that uncertainty in the model can be seen as uncertainty in future data. Personally, I lean more towards the frequentist view.

Bayesian Regret vs Frequentist Regret

Bayesian Regret $\mathrm{Reg}(T) = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta)$ **vs** Frequentist Regret $\mathrm{Reg}(T;\theta)$

The problem is: do we consider the model θ fixed or not? Why should we average over H if the problem is fixed in reality?[3]

- Note that people have applied Bayesian algorithms to the Frequentist regret: Posterior sampling helps with exploration (epistemic uncertainty).
- Similarly, UCB-designs have been used to find rules that are efficient in the Bayesian regret sense.
- In this work we study the Bayesian regret.

[3] Some people claim that uncertainty in the model can be seen as uncertainty in future data. Personally, I lean more towards the frequentist view.

Bayesian Regret vs Frequentist Regret

Bayesian Regret $\mathrm{Reg}(T) = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta)$ **vs** Frequentist Regret $\mathrm{Reg}(T;\theta)$

The problem is: do we consider the model θ fixed or not? Why should we average over H if the problem is fixed in reality?[3]

- Note that people have applied Bayesian algorithms to the Frequentist regret: Posterior sampling helps with exploration (epistemic uncertainty).
- Similarly, UCB-designs have been used to find rules that are efficient in the Bayesian regret sense.
- In this work we study the Bayesian regret.

[3] Some people claim that uncertainty in the model can be seen as uncertainty in future data. Personally, I lean more towards the frequentist view.

Optimal solutions

Bayesian Regret: $\text{Reg}(T) = \int_{\Theta^K} \text{Reg}(T;\theta)\mathrm{d}H(\theta)$.

Dynamic programming. In principle, one could use dynamic programming to find a solution. The problem becomes impractical for general MAB problems [Fabius and Zwet, 1970, Lai, 1987]. Simple example:

- Bernoulli bandits with uniform prior on the means. The posterior is a Beta distribution $\text{Beta}(S_a(t)+1, N_a(t)-S_a(t)+1)$, where $N_a(t) = |\{t \in [K] : a_t = a\}|$ and $S_a(t) = |\{t : r_{a_t} = 1\}|$.
- We can define an MDP with state $s_t = ((S_a(t), N_a(t) - S_a(t))_a)$.
- Use Dynamic programming, etc... to solve this MDP [Gittins, 1979].

$\Rightarrow$ great interest to develop simple algorithms that are optimal in the Bayesian sense.

Optimal solutions

Bayesian Regret: $\mathrm{Reg}(T) = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta)$.

Dynamic programming. In principle, one could use dynamic programming to find a solution. The problem becomes impractical for general MAB problems [Fabius and Zwet, 1970, Lai, 1987]. Simple example:

- Bernoulli bandits with uniform prior on the means. The posterior is a Beta distribution $\mathrm{Beta}(S_a(t)+1, N_a(t)-S_a(t)+1)$, where $N_a(t) = |\{t \in [K] : a_t = a\}|$ and $S_a(t) = |\{t : r_{a_t} = 1\}|$.
- We can define an MDP with state $s_t = ((S_a(t), N_a(t) - S_a(t))_a)$.
- Use Dynamic programming, etc... to solve this MDP [Gittins, 1979].

$\Rightarrow$ great interest to develop simple algorithms that are optimal in the Bayesian sense.

Optimal solutions

Bayesian Regret: $\mathrm{Reg}(T) = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta)$.

Dynamic programming. In principle, one could use dynamic programming to find a solution. The problem becomes impractical for general MAB problems [Fabius and Zwet, 1970, Lai, 1987]. Simple example:

- Bernoulli bandits with uniform prior on the means. The posterior is a Beta distribution $\mathrm{Beta}(S_a(t)+1, N_a(t)-S_a(t)+1)$, where $N_a(t) = |\{t \in [K] : a_t = a\}|$ and $S_a(t) = |\{t : r_{a_t} = 1\}$.
- We can define an MDP with state $s_t = ((S_a(t), N_a(t) - S_a(t))_a)$.
- Use Dynamic programming, etc... to solve this MDP [Gittins, 1979].

$\Rightarrow$ great interest to develop simple algorithms that are optimal in the Bayesian sense.

Optimal solutions

Bayesian Regret: $\mathrm{Reg}(T) = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta)$.

Dynamic programming. In principle, one could use dynamic programming to find a solution. The problem becomes impractical for general MAB problems [Fabius and Zwet, 1970, Lai, 1987]. Simple example:

- Bernoulli bandits with uniform prior on the means. The posterior is a Beta distribution $\mathrm{Beta}(S_a(t)+1, N_a(t)-S_a(t)+1)$, where $N_a(t) = |\{t \in [K] : a_t = a\}|$ and $S_a(t) = |\{t : r_{a_t} = 1\}|$.
- We can define an MDP with state $s_t = ((S_a(t), N_a(t) - S_a(t))_a)$.
- Use Dynamic programming, etc... to solve this MDP [Gittins, 1979].

$\Rightarrow$ great interest to develop simple algorithms that are optimal in the Bayesian sense.

Optimal solutions

Bayesian Regret: $\text{Reg}(T) = \int_{\Theta^K} \text{Reg}(T;\theta) \mathrm{d}H(\theta)$.

Dynamic programming. In principle, one could use dynamic programming to find a solution. The problem becomes impractical for general MAB problems [Fabius and Zwet, 1970, Lai, 1987]. Simple example:

- Bernoulli bandits with uniform prior on the means. The posterior is a Beta distribution $\text{Beta}(S_a(t)+1, N_a(t)-S_a(t)+1)$, where $N_a(t) = |\{t \in [K] : a_t = a\}|$ and $S_a(t) = |\{t : r_{a_t} = 1\}$.
- We can define an MDP with state $s_t = ((S_a(t), N_a(t) - S_a(t))_a)$.
- Use Dynamic programming, etc... to solve this MDP [Gittins, 1979].

$\Rightarrow$ great interest to develop simple algorithms that are optimal in the Bayesian sense.

Bayesian Regret Lower Bound: Proof Idea

Bayesian Regret $\mathrm{Reg}(T) = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta)$.

We now derive an asymptotic instance-dependent lower bound for the Bayesian Regret.

Proof idea:

1. Recall that $\mathrm{Reg}(T;\theta) = \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$.
2. Lower bound $\liminf_{T\to\infty} \mathbb{E}_\theta[N_a(T)] \geq k(\theta;T)$.
3. Integrate $\int \Delta_a(\theta)k(\theta;T)\mathrm{d}H(\theta)$.

Bayesian Regret Lower Bound: Proof Idea

Bayesian Regret $\mathrm{Reg}(T) = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta)$.

We now derive an asymptotic instance-dependent lower bound for the Bayesian Regret.

Proof idea:

1. Recall that $\mathrm{Reg}(T;\theta) = \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$.
2. Lower bound $\liminf_{T\to\infty} \mathbb{E}_\theta[N_a(T)] \geq k(\theta;T)$.
3. Integrate $\int \Delta_a(\theta)k(\theta;T)\mathrm{d}H(\theta)$.

Bayesian Regret Lower Bound: Proof Idea

Bayesian Regret $\mathrm{Reg}(T) = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta)$.

We now derive an asymptotic instance-dependent lower bound for the Bayesian Regret.

Proof idea:

1. Recall that $\mathrm{Reg}(T;\theta) = \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$.
2. Lower bound $\liminf_{T\to\infty} \mathbb{E}_\theta[N_a(T)] \geq k(\theta;T)$.
3. Integrate $\int \Delta_a(\theta)k(\theta;T)\mathrm{d}H(\theta)$.

Bayesian Regret Lower Bound: Proof Idea

Bayesian Regret $\mathrm{Reg}(T) = \int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta)$.

We now derive an asymptotic instance-dependent lower bound for the Bayesian Regret.

Proof idea:

1. Recall that $\mathrm{Reg}(T;\theta) = \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$.
2. Lower bound $\liminf_{T\to\infty} \mathbb{E}_\theta[N_a(T)] \geq k(\theta;T)$.
3. Integrate $\int \Delta_a(\theta)k(\theta;T)\mathrm{d}H(\theta)$.

Wrong! We need to integrate and then take the limit. Need to be careful, since the lower bound needs to have some form of uniformity in θ!

Frequentist Regret Minimization

In frequentist regret we usually look at uniformly fast convergent strategies.

Uniformly fast convergent strategies. A strategy π is uniformly fast convergent if for all models θ, for all sub-optimal arms a we have $\mathbb{E}_\theta[N_a(T)] = o(T^\alpha)$ for all $\alpha \in (0,1)$.

Using this result, by a change-of-measure argument, we can show that

$$\mathbb{E}_\theta[N_a(T)] \underset{T\to\infty}{\sim} \frac{\log(T)}{D(\theta_a, \theta^\star)}$$

- However, the condition $\mathbb{E}_\theta[N_a(T)] = o(T^\alpha)$ is not uniform in θ. The convergence is pointwise, and for different θs the convergence speed may be different.[4]
- If $\theta^\star = \theta_a + \epsilon$ with $\epsilon \to 0$, then

$$\frac{\mathbb{E}_\theta[N_a(T)]}{\log T} \sim \frac{1}{D(\theta_a, \theta_a + \epsilon)} \sim \frac{1}{\epsilon^2} \underset{\epsilon\to 0}{\to} \infty.$$

[4] This is important, since we will take an integral over Θ of the regret.

Frequentist Regret Minimization

In frequentist regret we usually look at uniformly fast convergent strategies.

Uniformly fast convergent strategies. A strategy π is uniformly fast convergent if for all models θ, for all sub-optimal arms a we have $\mathbb{E}_\theta[N_a(T)] = o(T^\alpha)$ for all $\alpha \in (0,1)$.

Using this result, by a change-of-measure argument, we can show that

$$\mathbb{E}_\theta[N_a(T)] \underset{T\to\infty}{\sim} \frac{\log(T)}{D(\theta_a, \theta^\star)}$$

- However, the condition $\mathbb{E}_\theta[N_a(T)] = o(T^\alpha)$ is not uniform in θ. The convergence is pointwise, and for different θs the convergence speed may be different.[4]
- If $\theta^\star = \theta_a + \epsilon$ with $\epsilon \to 0$, then

$$\frac{\mathbb{E}_\theta[N_a(T)]}{\log T} \sim \frac{1}{D(\theta_a, \theta_a+\epsilon)} \sim \frac{1}{\epsilon^2} \underset{\epsilon\to 0}{\to} \infty.$$

[4]This is important, since we will take an integral over Θ of the regret.

Frequentist Regret Minimization

In frequentist regret we usually look at uniformly fast convergent strategies.

> **Uniformly fast convergent strategies.** A strategy π is uniformly fast convergent if for all models θ, for all sub-optimal arms a we have $\mathbb{E}_\theta[N_a(T)] = o(T^\alpha)$ for all $\alpha \in (0, 1)$.

Using this result, by a change-of-measure argument, we can show that

$$\mathbb{E}_\theta[N_a(T)] \underset{T\to\infty}{\sim} \frac{\log(T)}{D(\theta_a, \theta^\star)}$$

- However, the condition $\mathbb{E}_\theta[N_a(T)] = o(T^\alpha)$ is not uniform in θ. The convergence is pointwise, and for different θs the convergence speed may be different.[4]
- If $\theta^\star = \theta_a + \epsilon$ with $\epsilon \to 0$, then

$$\frac{\mathbb{E}_\theta[N_a(T)]}{\log T} \sim \frac{1}{D(\theta_a, \theta_a + \epsilon)} \sim \frac{1}{\epsilon^2} \underset{\epsilon\to 0}{\to} \infty.$$

[4] This is important, since we will take an integral over Θ of the regret.

Frequentist Regret Minimization

In frequentist regret we usually look at uniformly fast convergent strategies.

Uniformly fast convergent strategies. A strategy π is uniformly fast convergent if for all models θ, for all sub-optimal arms a we have $\mathbb{E}_\theta[N_a(T)] = o(T^\alpha)$ for all $\alpha \in (0, 1)$.

Using this result, by a change-of-measure argument, we can show that

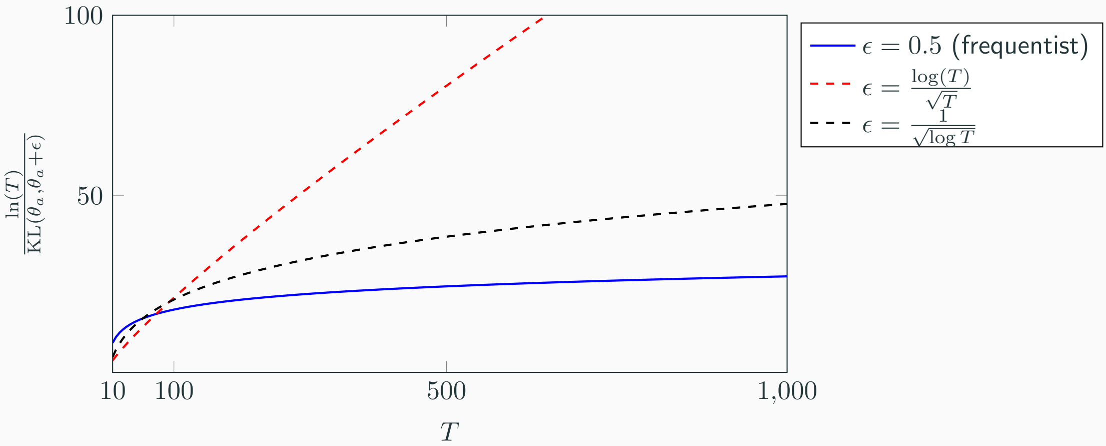
$$\mathbb{E}_\theta[N_a(T)] \underset{T\to\infty}{\sim} \frac{\log(T)}{D(\theta_a, \theta^\star)}$$

- However, the condition $\mathbb{E}_\theta[N_a(T)] = o(T^\alpha)$ is not uniform in θ. The convergence is pointwise, and for different θs the convergence speed may be different.[4]
- If $\theta^\star = \theta_a + \epsilon$ with $\epsilon \to 0$, then

$$\frac{\mathbb{E}_\theta[N_a(T)]}{\log T} \sim \frac{1}{D(\theta_a, \theta_a + \epsilon)} \sim \frac{1}{\epsilon^2} \underset{\epsilon\to 0}{\to} \infty.$$

[4] This is important, since we will take an integral over Θ of the regret.

Frequentist Regret Minimization



Remark: Just saying that π is *uniformly fast convergent strategy* is not enough. We need to guarantee uniform convergence across different values of θ.

Some notation

Notation.

- For an arm a and vector θ, denote by $\theta_{\setminus a} = (\theta_1, \ldots, \theta_{a-1}, \theta_{a+1}, \ldots, \theta_K)$ the vector θ without θ_a. Similarly, we also define $H_{\setminus a}$.
- Let $\theta^\star_{\setminus a} = \max_{j \neq a} \theta_j$ be the best element in $\theta_{\setminus a}$ and recall that $\theta^\star = \max_a \theta_a$.
- Also recall that $\mu_a(\theta) = \dot{\psi}(\theta_a)$, which is increasing in θ_a. Hence $\theta^\star$ corresponds to the parameter of the best arm.

Using this last fact, we also write

$$\mathrm{Reg}(T;\theta) = \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)] = \sum_{a:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$$

Some notation

Notation.

- For an arm a and vector θ, denote by $\theta_{\backslash a} = (\theta_1, \dots, \theta_{a-1}, \theta_{a+1}, \dots, \theta_K)$ the vector θ without θ_a. Similarly, we also define $H_{\backslash a}$.
- Let $\theta^\star_{\backslash a} = \max_{j \neq a} \theta_j$ be the best element in $\theta_{\backslash a}$ and recall that $\theta^\star = \max_a \theta_a$.
- Also recall that $\mu_a(\theta) = \dot{\psi}(\theta_a)$, which is increasing in θ_a. Hence $\theta^\star$ corresponds to the parameter of the best arm.

Using this last fact, we also write

$$\mathrm{Reg}(T;\theta) = \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)] = \sum_{a:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$$

Some notation

Notation.

- For an arm a and vector θ, denote by $\theta_{\setminus a} = (\theta_1, \ldots, \theta_{a-1}, \theta_{a+1}, \ldots, \theta_K)$ the vector θ without θ_a. Similarly, we also define $H_{\setminus a}$.
- Let $\theta^\star_{\setminus a} = \max_{j \neq a} \theta_j$ be the best element in $\theta_{\setminus a}$ and recall that $\theta^\star = \max_a \theta_a$.
- Also recall that $\mu_a(\theta) = \dot{\psi}(\theta_a)$, which is increasing in θ_a. Hence $\theta^\star$ corresponds to the parameter of the best arm.

Using this last fact, we also write

$$\mathrm{Reg}(T;\theta) = \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)] = \sum_{a:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$$

Some notation

Notation.

- For an arm a and vector θ, denote by $\theta_{\backslash a} = (\theta_1, \ldots, \theta_{a-1}, \theta_{a+1}, \ldots, \theta_K)$ the vector θ without θ_a. Similarly, we also define $H_{\backslash a}$.
- Let $\theta^\star_{\backslash a} = \max_{j \neq a} \theta_j$ be the best element in $\theta_{\backslash a}$ and recall that $\theta^\star = \max_a \theta_a$.
- Also recall that $\mu_a(\theta) = \dot{\psi}(\theta_a)$, which is increasing in θ_a. Hence $\theta^\star$ corresponds to the parameter of the best arm.

Using this last fact, we also write

$$\mathrm{Reg}(T;\theta) = \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)] = \sum_{a:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$$

Bayesian Regret Lower Bound: Proof Idea

Proof idea: uniformly lower bound $\Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$ over a small region around $\theta^\star$ (i.e., where the gap is small $\Rightarrow$ this region contributes the most to the regret), and take the limit. Write the regret as follows

$$\mathrm{Reg}(T;\theta) = \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)] = \sum_{a:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$$

Then, we can integrate and note that it is sufficient to consider the case $\theta_a < \theta^\star$:

$$\begin{aligned}\int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta) &= \int_{\Theta^K} \sum_{a:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\,\mathrm{d}H(\theta),\\ &= \int_{\Theta^K} \sum_{a} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\mathbb{1}_{\{\theta_a<\theta^\star\}}\,\mathrm{d}H(\theta),\\ &= \sum_a \int_{\theta\in\Theta^K:\theta_a<\theta^\star_{\setminus a}} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\,\mathrm{d}H(\theta)\end{aligned}$$

Focus on the integral (it's the same integral for each a). And now?

Bayesian Regret Lower Bound: Proof Idea

Proof idea: uniformly lower bound $\Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$ over a small region around $\theta^\star$ (i.e., where the gap is small $\Rightarrow$ this region contributes the most to the regret), and take the limit. Write the regret as follows

$$\mathrm{Reg}(T;\theta) = \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)] = \sum_{a:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$$

Then, we can integrate and note that it is sufficient to consider the case $\theta_a < \theta^\star$:

$$\begin{aligned}\int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta) &= \int_{\Theta^K} \sum_{a:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\, \mathrm{d}H(\theta), \\ &= \int_{\Theta^K} \sum_a \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\mathbb{1}_{\{\theta_a<\theta^\star\}}\, \mathrm{d}H(\theta), \\ &= \sum_a \int_{\theta\in\Theta^K:\theta_a<\theta^\star_{\setminus a}} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\, \mathrm{d}H(\theta)\end{aligned}$$

Focus on the integral (it's the same integral for each a). And now?

Bayesian Regret Lower Bound: Proof Idea

Proof idea: uniformly lower bound $\Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$ over a small region around $\theta^\star$ (i.e., where the gap is small $\Rightarrow$ this region contributes the most to the regret), and take the limit. Write the regret as follows

$$\mathrm{Reg}(T;\theta) = \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)] = \sum_{a:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$$

Then, we can integrate and note that it is sufficient to consider the case $\theta_a < \theta^\star$:

$$\int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta) = \int_{\Theta^K} \sum_{a:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\,\mathrm{d}H(\theta),$$

$$= \int_{\Theta^K} \sum_a \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\mathbf{1}_{\{\theta_a<\theta^\star\}}\,\mathrm{d}H(\theta),$$

$$= \sum_a \int_{\theta\in\Theta^K:\theta_a<\theta^\star_{\setminus a}} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\,\mathrm{d}H(\theta)$$

Focus on the integral (it's the same integral for each a). And now?

Bayesian Regret Lower Bound: Proof Idea

Proof idea: uniformly lower bound $\Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$ over a small region around $\theta^\star$ (i.e., where the gap is small $\Rightarrow$ this region contributes the most to the regret), and take the limit. Write the regret as follows

$$\mathrm{Reg}(T;\theta) = \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)] = \sum_{a:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$$

Then, we can integrate and note that it is sufficient to consider the case $\theta_a < \theta^\star$:

$$\begin{aligned}
\int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta) &= \int_{\Theta^K} \sum_{a:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\,\mathrm{d}H(\theta), \\
&= \int_{\Theta^K} \sum_a \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\mathbf{1}_{\{\theta_a<\theta^\star\}}\,\mathrm{d}H(\theta), \\
&= \sum_a \int_{\theta\in\Theta^K:\theta_a<\theta^\star_{\setminus a}} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\,\mathrm{d}H(\theta)
\end{aligned}$$

Focus on the integral (it's the same integral for each a). And now?

Bayesian Regret Lower Bound: Proof Idea

Proof idea: uniformly lower bound $\Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$ over a small region around $\theta^\star$ (i.e., where the gap is small $\Rightarrow$ this region contributes the most to the regret), and take the limit. Write the regret as follows

$$\mathrm{Reg}(T;\theta) = \sum_{a:\mu_a(\theta)<\mu^\star(\theta)} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)] = \sum_{a:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]$$

Then, we can integrate and note that it is sufficient to consider the case $\theta_a < \theta^\star$:

$$\begin{aligned}
\int_{\Theta^K} \mathrm{Reg}(T;\theta)\mathrm{d}H(\theta) &= \int_{\Theta^K} \sum_{a:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\,\mathrm{d}H(\theta), \\
&= \int_{\Theta^K} \sum_{a} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\mathbf{1}_{\{\theta_a<\theta^\star\}}\,\mathrm{d}H(\theta), \\
&= \sum_a \int_{\theta\in\Theta^K:\theta_a<\theta^\star_{\setminus a}} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\,\mathrm{d}H(\theta)
\end{aligned}$$

Focus on the integral (it's the same integral for each a). And now?

To our help comes Prof. Lai [Lai, 1987]. With a single parameter, he noted that UCB methods (based on the KL divergence[5]) satisfy the following:

> Let $S = \inf\{n \le T : \mathrm{UCB}(n) \le \theta + \epsilon\}$, then, as $T \to \infty$
>
> $$\mathbb{P}_\theta(S \le (1-\gamma)(\log N\epsilon^2)/D(\theta, \theta+\epsilon)) \to 0 \qquad \forall \gamma \in (0,1)$$
>
> uniformly in $\alpha_T \le \epsilon \le \beta_T$, with $\alpha_T \to 0, \beta_T \to \infty$ and $\sqrt{T}\alpha_T \to \infty, \beta_T = o(\sqrt{\log T})$.

▷ for a given θ we need to sample at a rate $\approx \log(T\epsilon^2)/\epsilon^2$ to detect an ϵ difference.

- When integrating over ϵ we get $\int \log(T\epsilon^2)/\epsilon^2 \mathrm{d}\epsilon = -\frac{\ln T}{\epsilon} - 2\frac{1+\ln \epsilon}{\epsilon} + C = -\frac{2+\ln T\epsilon^2}{\epsilon} + C$.

Over a small region, e.g. $(\log^{-1} T, T^{-1/2})$, we get $\sim \log^2(T)$.

[5] $\mathrm{UCB}(n) = \inf\{\theta : \theta \ge \hat{\theta}_n, nD(\hat{\theta}_n, \theta) \ge \log(n/T) + \xi \log\log(n/T)\}$ for some $\xi \in \mathbb{R}$ and $\hat{\theta}_n$ is the MLE in round n.

To our help comes Prof. Lai [Lai, 1987]. With a single parameter, he noted that UCB methods (based on the KL divergence[5]) satisfy the following:

> Let $S = \inf\{n \leq T : \text{UCB}(n) \leq \theta + \epsilon\}$, then, as $T \to \infty$
>
> $$\mathbb{P}_\theta(S \leq (1-\gamma)(\log N\epsilon^2)/D(\theta, \theta+\epsilon)) \to 0 \qquad \forall \gamma \in (0,1)$$
>
> uniformly in $\alpha_T \leq \epsilon \leq \beta_T$, with $\alpha_T \to 0, \beta_T \to \infty$ and $\sqrt{T}\alpha_T \to \infty, \beta_T = o(\sqrt{\log T})$.

▷ for a given θ we need to sample at a rate $\approx \log(T\epsilon^2)/\epsilon^2$ to detect an ϵ difference.

- When integrating over ϵ we get $\int \log(T\epsilon^2)/\epsilon^2 \mathrm{d}\epsilon = -\frac{\ln T}{\epsilon} - 2\frac{1+\ln\epsilon}{\epsilon} + C = -\frac{2+\ln T\epsilon^2}{\epsilon} + C$.

Over a small region, e.g. $(\log^{-1} T, T^{-1/2})$, we get $\sim \log^2(T)$.

[5] $\text{UCB}(n) = \inf\{\theta : \theta \geq \hat{\theta}_n, nD(\hat{\theta}_n, \theta) \geq \log(n/T) + \xi \log\log(n/T)\}$ for some $\xi \in \mathbb{R}$ and $\hat{\theta}_n$ is the MLE in round n.

To our help comes Prof. Lai [Lai, 1987]. With a single parameter, he noted that UCB methods (based on the KL divergence[5]) satisfy the following:

> Let $S = \inf\{n \leq T : \mathrm{UCB}(n) \leq \theta + \epsilon\}$, then, as $T \to \infty$
>
> $$\mathbb{P}_\theta(S \leq (1-\gamma)(\log N\epsilon^2)/D(\theta, \theta+\epsilon)) \to 0 \qquad \forall \gamma \in (0,1)$$
>
> uniformly in $\alpha_T \leq \epsilon \leq \beta_T$, with $\alpha_T \to 0, \beta_T \to \infty$ and $\sqrt{T}\alpha_T \to \infty, \beta_T = o(\sqrt{\log T})$.

▷ for a given θ we need to sample at a rate $\approx \log(T\epsilon^2)/\epsilon^2$ to detect an ϵ difference.

- When integrating over ϵ we get $\int \log(T\epsilon^2)/\epsilon^2 \mathrm{d}\epsilon = -\frac{\ln T}{\epsilon} - 2\frac{1+\ln\epsilon}{\epsilon} + C = -\frac{2+\ln T\epsilon^2}{\epsilon} + C$.

Over a small region, e.g. $(\log^{-1} T, T^{-1/2})$, we get $\sim \log^2(T)$.

[5] $\mathrm{UCB}(n) = \inf\{\theta : \theta \geq \hat{\theta}_n, nD(\hat{\theta}_n, \theta) \geq \log(n/T) + \xi \log\log(n/T)\}$ for some $\xi \in \mathbb{R}$ and $\hat{\theta}_n$ is the MLE in round n.

Bayesian Regret Lower Bound: Assumptions on the Sampling Policy

Using this intuition, [Lai, 1987] derived the following sampling condition [6].

Let $\xi, \gamma \in (0,1)$. π is a Bayes-uniformly fast convergent strategy if

$$\lim_{T\to\infty, \epsilon\to 0, T\epsilon^2\to\infty} \int_{\Theta^{K-1}} \mathbb{P}_\theta \left(N_a(T) \le (1-\gamma) \frac{\log T\epsilon^2}{D(\theta_a, \theta^\star_{\setminus a} + \xi\epsilon)} \right) h_a(\theta_a) \mathrm{d}H(\theta_{\setminus a}) = 0$$

with $\theta_a = \theta^\star_{\setminus a} - \epsilon$.

The probability that we under-sample over regions with small gaps tends to 0.

[6] Recall from the previous slide that we need a sampling rate $\approx \log(T\epsilon^2)/\epsilon^2$. Since $\theta_a = \theta^\star_{\setminus a} - \epsilon$ we have $D(\theta_a, \theta^\star_{\setminus a} + \xi\epsilon) \approx (1+\xi)^2\epsilon^2$.

Bayesian Regret Lower Bound: Final Steps

$$\int_{\theta \in \Theta^K : \theta_a < \theta^\star} \Delta_a(\theta) \mathbb{E}_\theta[N_a(T)] \, \mathrm{d}H(\theta) = \int_{\theta \in \Theta^K : \theta_a < \theta^\star_{\backslash a}} \Delta_a(\theta) \mathbb{E}_\theta[N_a(T)] \, \mathrm{d}H(\theta) = (*)$$

The idea is to consider $\theta_a = \theta^\star_{\backslash a} - \epsilon$, with ϵ belonging to a small region around the maximum. We consider an open set $\mathcal{E}_T \subset \mathbb{R}_+$ such that $\mathcal{E}_T \to \{0\}$ (more details on this later).

$$\begin{aligned}
(*) &= \int_{\theta \in \Theta^{K-1}} \int_{\theta_a < \theta^\star_{\backslash a}} \Delta_a(\theta) \mathbb{E}_\theta[N_a(T)] \mathrm{d}H_a(\theta_a) \, \mathrm{d}H_{\backslash a}(\theta_{\backslash a}), \\
&\geq \int_{\theta \in \Theta^{K-1}} \int_{\theta^\star_{\backslash a} - \theta_a \in \mathcal{E}_T} \Delta_a(\theta) \mathbb{E}_\theta[N_a(T)] \, \mathrm{d}H_a(\theta_a) \mathrm{d}H_{\backslash a}(\theta_{\backslash a}), \\
&= \int_{\theta \in \Theta^{K-1}} \int_{\mathcal{E}_T} (\dot\psi(\theta^\star_{\backslash a}) - \dot\psi(\theta^\star_{\backslash a} - \epsilon)) \mathbb{E}_\theta[N_a(T)] h_a(\theta^\star_{\backslash a} - \epsilon) \, \mathrm{d}\epsilon \, \mathrm{d}H_{\backslash a}(\theta_{\backslash a}),
\end{aligned}$$

where we used that $\Delta_a(\theta) = \max_j \mu_j(\theta) - \mu_a(\theta) = \max_j \dot\psi(\theta_j) - \dot\psi(\theta_a)$ and performed a change of variable $\theta_a = \theta^\star_{\backslash a} - \epsilon, \epsilon \in \mathcal{E}_T$.

Bayesian Regret Lower Bound: Final Steps

$$\int_{\theta\in\Theta^K:\theta_a<\theta^\star} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\,\mathrm{d}H(\theta) = \int_{\theta\in\Theta^K:\theta_a<\theta^\star_{\setminus a}} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\,\mathrm{d}H(\theta) = (*)$$

The idea is to consider $\theta_a = \theta^\star_{\setminus a} - \epsilon$, with ϵ belonging to a small region around the maximum. We consider an open set $\mathcal{E}_T \subset \mathbb{R}_+$ such that $\mathcal{E}_T \to \{0\}$ (more details on this later).

$$\begin{aligned}(*) &= \int_{\theta\in\Theta^{K-1}} \int_{\theta_a<\theta^\star_{\setminus a}} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\mathrm{d}H_a(\theta_a)\,\mathrm{d}H_{\setminus a}(\theta_{\setminus a}),\\ &\geq \int_{\theta\in\Theta^{K-1}} \int_{\theta^\star_{\setminus a}-\theta_a\in\mathcal{E}_T} \Delta_a(\theta)\mathbb{E}_\theta[N_a(T)]\,\mathrm{d}H_a(\theta_a)\mathrm{d}H_{\setminus a}(\theta_{\setminus a}),\\ &= \int_{\theta\in\Theta^{K-1}} \int_{\mathcal{E}_T} (\dot\psi(\theta^\star_{\setminus a}) - \dot\psi(\theta^\star_{\setminus a}-\epsilon))\mathbb{E}_\theta[N_a(T)]h_a(\theta^\star_{\setminus a}-\epsilon)\,\mathrm{d}\epsilon\,\mathrm{d}H_{\setminus a}(\theta_{\setminus a}),\end{aligned}$$

where we used that $\Delta_a(\theta) = \max_j \mu_j(\theta) - \mu_a(\theta) = \max_j \dot\psi(\theta_j) - \dot\psi(\theta_a)$ and performed a change of variable $\theta_a = \theta^\star_{\setminus a} - \epsilon, \epsilon \in \mathcal{E}_T$.

Bayesian Regret Lower Bound: Final Steps

$$\int_{\theta \in \Theta^K : \theta_a < \theta^\star} \Delta_a(\theta) \mathbb{E}_\theta[N_a(T)] \, \mathrm{d}H(\theta) = \int_{\theta \in \Theta^K : \theta_a < \theta^\star_{\setminus a}} \Delta_a(\theta) \mathbb{E}_\theta[N_a(T)] \, \mathrm{d}H(\theta) = (*)$$

The idea is to consider $\theta_a = \theta^\star_{\setminus a} - \epsilon$, with ϵ belonging to a small region around the maximum. We consider an open set $\mathcal{E}_T \subset \mathbb{R}_+$ such that $\mathcal{E}_T \to \{0\}$ (more details on this later).

$$\begin{aligned}
(*) &= \int_{\theta \in \Theta^{K-1}} \int_{\theta_a < \theta^\star_{\setminus a}} \Delta_a(\theta) \mathbb{E}_\theta[N_a(T)] \mathrm{d}H_a(\theta_a) \, \mathrm{d}H_{\setminus a}(\theta_{\setminus a}), \\
&\geq \int_{\theta \in \Theta^{K-1}} \int_{\theta^\star_{\setminus a} - \theta_a \in \mathcal{E}_T} \Delta_a(\theta) \mathbb{E}_\theta[N_a(T)] \, \mathrm{d}H_a(\theta_a) \mathrm{d}H_{\setminus a}(\theta_{\setminus a}), \\
&= \int_{\theta \in \Theta^{K-1}} \int_{\mathcal{E}_T} (\dot\psi(\theta^\star_{\setminus a}) - \dot\psi(\theta^\star_{\setminus a} - \epsilon)) \mathbb{E}_\theta[N_a(T)] h_a(\theta^\star_{\setminus a} - \epsilon) \, \mathrm{d}\epsilon \, \mathrm{d}H_{\setminus a}(\theta_{\setminus a}),
\end{aligned}$$

where we used that $\Delta_a(\theta) = \max_j \mu_j(\theta) - \mu_a(\theta) = \max_j \dot\psi(\theta_j) - \dot\psi(\theta_a)$ and performed a change of variable $\theta_a = \theta^\star_{\setminus a} - \epsilon, \epsilon \in \mathcal{E}_T$.

Bayesian Regret Lower Bound: Final Steps

$$\int_{\theta \in \Theta^{K-1}} \int_{\mathcal{E}_T} (\dot{\psi}(\theta^\star_{\setminus a}) - \dot{\psi}(\theta^\star_{\setminus a} - \epsilon))\mathbb{E}_\theta[N_a(T)] h_a(\theta^\star_{\setminus a} - \epsilon) \, \mathrm{d}\epsilon \, \mathrm{d}H_{\setminus a}(\theta_{\setminus a}) = (*)$$

Let $\gamma \in (0, 1)$. Use that

1. $p_\theta(T, \epsilon) = \mathbb{P}_\theta\left(N_a(T) \leq (1-\gamma)\frac{\log T\epsilon^2}{D(\theta^\star_{\setminus a} - \epsilon, \theta^\star_{\setminus a} + \xi\epsilon)}\right)$ [7] implies

   $$\mathbb{E}[N_a(\theta)] \geq (1 - p_\theta(T, \epsilon))(1-\gamma)\frac{\log T\epsilon^2}{D(\theta^\star_{\setminus a} - \epsilon, \theta^\star_{\setminus a} + \xi\epsilon)}$$

   by Markov's inequality. Note also that $D(\theta^\star_{\setminus a} - \epsilon, \theta^\star_{\setminus a} + \xi\epsilon) \leq c'(1+\xi)^2\epsilon^2/2$

2. By continuity, for ϵ small we can say $h_a(\theta^\star_{\setminus a} - \epsilon) \approx h_a(\theta^\star_{\setminus a})$.
3. Also note that $\dot{\psi}(\theta^\star_{\setminus a}) - \dot{\psi}(\theta^\star_{\setminus a} - \epsilon) \geq c\epsilon$ for $y \approx x$ by continuity.
4. Then, for ϵ, ξ sufficiently small, by continuity we have

$$(\dot{\psi}(\theta^\star_{\setminus a}) - \dot{\psi}(\theta^\star_{\setminus a} - \epsilon))h_a(\theta^\star_{\setminus a} - \epsilon) \geq 2(1-\gamma)h_a(\theta^\star_{\setminus a})\frac{D(\theta^\star_{\setminus a} - \epsilon, \theta^\star_{\setminus a} + \xi\epsilon)}{\epsilon}.$$

[7] This is the quantity appearing in the Bayes-uniformly fast convergent strategy assumption.

Bayesian Regret Lower Bound: Final Steps

$$\int_{\theta\in\Theta^{K-1}}\int_{\mathcal{E}_T}(\dot\psi(\theta^\star_{\setminus a}) - \dot\psi(\theta^\star_{\setminus a}-\epsilon))\mathbb{E}_\theta[N_a(T)]h_a(\theta^\star_{\setminus a}-\epsilon)\,\mathrm{d}\epsilon\,\mathrm{d}H_{\setminus a}(\theta_{\setminus a}) = (*)$$

Let $\gamma\in(0,1)$. Use that

1. $p_\theta(T,\epsilon) = \mathbb{P}_\theta\left(N_a(T)\le(1-\gamma)\frac{\log T\epsilon^2}{D(\theta^\star_{\setminus a}-\epsilon,\theta^\star_{\setminus a}+\xi\epsilon)}\right)$ [7] implies

   $$\mathbb{E}[N_a(\theta)] \ge (1-p_\theta(T,\epsilon))(1-\gamma)\frac{\log T\epsilon^2}{D(\theta^\star_{\setminus a}-\epsilon,\theta^\star_{\setminus a}+\xi\epsilon)}$$

   by Markov's inequality. Note also that $D(\theta^\star_{\setminus a}-\epsilon,\theta^\star_{\setminus a}+\xi\epsilon)\le c'(1+\xi)^2\epsilon^2/2$
2. By continuity, for ϵ small we can say $h_a(\theta^\star_{\setminus a}-\epsilon)\approx h_a(\theta^\star_{\setminus a})$.
3. Also note that $\dot\psi(\theta^\star_{\setminus a}) - \dot\psi(\theta^\star_{\setminus a}-\epsilon)\ge c\epsilon$ for $y\approx x$ by continuity.
4. Then, for ϵ,ξ sufficiently small, by continuity we have

   $$(\dot\psi(\theta^\star_{\setminus a}) - \dot\psi(\theta^\star_{\setminus a}-\epsilon))h_a(\theta^\star_{\setminus a}-\epsilon)\ge 2(1-\gamma)h_a(\theta^\star_{\setminus a})\frac{D(\theta^\star_{\setminus a}-\epsilon,\theta^\star_{\setminus a}+\xi\epsilon)}{\epsilon}.$$

[7] This is the quantity appearing in the Bayes-uniformly fast convergent strategy assumption.

Bayesian Regret Lower Bound: Final Steps

$$\int_{\theta\in\Theta^{K-1}}\int_{\mathcal{E}_T}(\dot\psi(\theta^\star_{\setminus a})-\dot\psi(\theta^\star_{\setminus a}-\epsilon))\mathbb{E}_\theta[N_a(T)]h_a(\theta^\star_{\setminus a}-\epsilon)\,\mathrm{d}\epsilon\,\mathrm{d}H_{\setminus a}(\theta_{\setminus a})=(*)$$

Let $\gamma\in(0,1)$. Use that

1. $p_\theta(T,\epsilon)=\mathbb{P}_\theta\left(N_a(T)\le(1-\gamma)\frac{\log T\epsilon^2}{D(\theta^\star_{\setminus a}-\epsilon,\theta^\star_{\setminus a}+\xi\epsilon)}\right)$ [7] implies

$$\mathbb{E}[N_a(\theta)]\ge(1-p_\theta(T,\epsilon))(1-\gamma)\frac{\log T\epsilon^2}{D(\theta^\star_{\setminus a}-\epsilon,\theta^\star_{\setminus a}+\xi\epsilon)}$$

   by Markov's inequality. Note also that $D(\theta^\star_{\setminus a}-\epsilon,\theta^\star_{\setminus a}+\xi\epsilon)\le c'(1+\xi)^2\epsilon^2/2$
2. By continuity, for ϵ small we can say $h_a(\theta^\star_{\setminus a}-\epsilon)\approx h_a(\theta^\star_{\setminus a})$.
3. Also note that $\dot\psi(\theta^\star_{\setminus a})-\dot\psi(\theta^\star_{\setminus a}-\epsilon)\ge c\epsilon$ for $y\approx x$ by continuity.
4. Then, for ϵ,ξ sufficiently small, by continuity we have

$$(\dot\psi(\theta^\star_{\setminus a})-\dot\psi(\theta^\star_{\setminus a}-\epsilon))h_a(\theta^\star_{\setminus a}-\epsilon)\ge2(1-\gamma)h_a(\theta^\star_{\setminus a})\frac{D(\theta^\star_{\setminus a}-\epsilon,\theta^\star_{\setminus a}+\xi\epsilon)}{\epsilon}.$$

[7] This is the quantity appearing in the Bayes-uniformly fast convergent strategy assumption.

Bayesian Regret Lower Bound: Final Steps

$$\int_{\theta\in\Theta^{K-1}}\int_{\mathcal{E}_T}(\dot\psi(\theta^\star_{\backslash a})-\dot\psi(\theta^\star_{\backslash a}-\epsilon))\mathbb{E}_\theta[N_a(T)]h_a(\theta^\star_{\backslash a}-\epsilon)\,\mathrm{d}\epsilon\,\mathrm{d}H_{\backslash a}(\theta_{\backslash a})=(*)$$

Let $\gamma\in(0,1)$. Use that

1. $p_\theta(T,\epsilon)=\mathbb{P}_\theta\left(N_a(T)\le(1-\gamma)\frac{\log T\epsilon^2}{D(\theta^\star_{\backslash a}-\epsilon,\theta^\star_{\backslash a}+\xi\epsilon)}\right)$ [7] implies

$$\mathbb{E}[N_a(\theta)]\ge(1-p_\theta(T,\epsilon))(1-\gamma)\frac{\log T\epsilon^2}{D(\theta^\star_{\backslash a}-\epsilon,\theta^\star_{\backslash a}+\xi\epsilon)}$$

   by Markov's inequality. Note also that $D(\theta^\star_{\backslash a}-\epsilon,\theta^\star_{\backslash a}+\xi\epsilon)\le c'(1+\xi)^2\epsilon^2/2$
2. By continuity, for ϵ small we can say $h_a(\theta^\star_{\backslash a}-\epsilon)\approx h_a(\theta^\star_{\backslash a})$.
3. Also note that $\dot\psi(\theta^\star_{\backslash a})-\dot\psi(\theta^\star_{\backslash a}-\epsilon)\ge c\epsilon$ for $y\approx x$ by continuity.
4. Then, for ϵ,ξ sufficiently small, by continuity we have

$$(\dot\psi(\theta^\star_{\backslash a})-\dot\psi(\theta^\star_{\backslash a}-\epsilon))h_a(\theta^\star_{\backslash a}-\epsilon)\ge 2(1-\gamma)h_a(\theta^\star_{\backslash a})\frac{D(\theta^\star_{\backslash a}-\epsilon,\theta^\star_{\backslash a}+\xi\epsilon)}{\epsilon}.$$

[7] This is the quantity appearing in the Bayes-uniformly fast convergent strategy assumption.

Bayesian Regret Lower Bound: Final Steps

$$(*) \geq 2(1-\gamma)^2 \int_{\theta \in \Theta^{K-1}} h_a(\theta^\star_{\setminus a}) \int_{\mathcal{E}_T} (1 - p_\theta(T,\epsilon)) \frac{\log T\epsilon^2}{\epsilon} \mathrm{d}\epsilon \, \mathrm{d}H_{\setminus a}(\theta_{\setminus a})$$

Recall

Let $\xi, \gamma \in (0,1)$. π is a Bayes-uniformly fast convergent strategy if

$$\lim_{T\to\infty, \epsilon\to 0, T\epsilon^2\to\infty} \int_{\Theta^{K-1}} \mathbb{P}_\theta\left(N_a(T) \leq (1-\gamma)\frac{\log T\epsilon^2}{D(\theta^\star_{\setminus a} - \epsilon, \theta^\star_{\setminus a} + \xi\epsilon)}\right) h_a(\theta^\star_{\setminus a}) \mathrm{d}H(\theta_{\setminus a}) = 0$$

For now, trust me that the term $\int h_a(\theta^\star_{\setminus a}) \int p_\theta(T,\epsilon) \mathrm{d}\epsilon \mathrm{d}H_{\setminus a}(\theta_{\setminus a})$ tends to 0. Then, asymptotically

$$\mathrm{Reg}(T) \sim 2(1-\gamma)^2 \sum_a \int h_a(\theta^\star_{\setminus a}) \int_{\mathcal{E}_T} \frac{\log T\epsilon^2}{\epsilon} \mathrm{d}\epsilon \mathrm{d}H_{\setminus a}(\theta_{\setminus a})$$

Bayesian Regret Lower Bound: Final Steps

$$(*) \geq 2(1-\gamma)^2 \int_{\theta \in \Theta^{K-1}} h_a(\theta^\star_{\backslash a}) \int_{\mathcal{E}_T} (1 - p_\theta(T, \epsilon)) \frac{\log T\epsilon^2}{\epsilon} \mathrm{d}\epsilon \, \mathrm{d}H_{\backslash a}(\theta_{\backslash a})$$

Recall

Let $\xi, \gamma \in (0,1)$. π is a Bayes-uniformly fast convergent strategy if

$$\lim_{T\to\infty, \epsilon\to 0, T\epsilon^2\to\infty} \int_{\Theta^{K-1}} \mathbb{P}_\theta \left(N_a(T) \leq (1-\gamma) \frac{\log T\epsilon^2}{D(\theta^\star_{\backslash a} - \epsilon, \theta^\star_{\backslash a} + \xi\epsilon)} \right) h_a(\theta^\star_{\backslash a}) \mathrm{d}H(\theta_{\backslash a}) = 0$$

For now, trust me that the term $\int h_a(\theta^\star_{\backslash a}) \int p_\theta(T,\epsilon) \mathrm{d}\epsilon \mathrm{d}H_{\backslash a}(\theta_{\backslash a})$ tends to 0. Then, asymptotically

$$\mathrm{Reg}(T) \sim 2(1-\gamma)^2 \sum_a \int h_a(\theta^\star_{\backslash a}) \int_{\mathcal{E}_T} \frac{\log T\epsilon^2}{\epsilon} \mathrm{d}\epsilon \mathrm{d}H_{\backslash a}(\theta_{\backslash a})$$

Bayesian Regret Lower Bound: Final Steps

$$(*) \geq 2(1-\gamma)^2 \int_{\theta \in \Theta^{K-1}} h_a(\theta^\star_{\backslash a}) \int_{\mathcal{E}_T} (1 - p_\theta(T,\epsilon)) \frac{\log T\epsilon^2}{\epsilon} \mathrm{d}\epsilon \, \mathrm{d}H_{\backslash a}(\theta_{\backslash a})$$

Recall

Let $\xi, \gamma \in (0,1)$. π is a Bayes-uniformly fast convergent strategy if

$$\lim_{T\to\infty, \epsilon\to 0, T\epsilon^2\to\infty} \int_{\Theta^{K-1}} \mathbb{P}_\theta \left(N_a(T) \leq (1-\gamma) \frac{\log T\epsilon^2}{D(\theta^\star_{\backslash a} - \epsilon, \theta^\star_{\backslash a} + \xi\epsilon)} \right) h_a(\theta^\star_{\backslash a}) \mathrm{d}H(\theta_{\backslash a}) = 0$$

For now, trust me that the term $\int h_a(\theta^\star_{\backslash a}) \int p_\theta(T,\epsilon) \mathrm{d}\epsilon \mathrm{d}H_{\backslash a}(\theta_{\backslash a})$ tends to 0. Then, asymptotically

$$\mathrm{Reg}(T) \sim 2(1-\gamma)^2 \sum_a \int h_a(\theta^\star_{\backslash a}) \int_{\mathcal{E}_T} \frac{\log T\epsilon^2}{\epsilon} \mathrm{d}\epsilon \mathrm{d}H_{\backslash a}(\theta_{\backslash a})$$

Bayesian Regret Lower Bound: Final Steps

$$(*) \geq 2(1-\gamma)^2 \int_{\theta \in \Theta^{K-1}} h_a(\theta^\star_{\backslash a}) \int_{\mathcal{E}_T} (1 - p_\theta(T,\epsilon)) \frac{\log T\epsilon^2}{\epsilon} \mathrm{d}\epsilon \, \mathrm{d}H_{\backslash a}(\theta_{\backslash a})$$

Recall

Let $\xi, \gamma \in (0,1)$. π is a Bayes-uniformly fast convergent strategy if

$$\lim_{T\to\infty, \epsilon\to 0, T\epsilon^2\to\infty} \int_{\Theta^{K-1}} \mathbb{P}_\theta \left(N_a(T) \leq (1-\gamma) \frac{\log T\epsilon^2}{D(\theta^\star_{\backslash a} - \epsilon, \theta^\star_{\backslash a} + \xi\epsilon)} \right) h_a(\theta^\star_{\backslash a}) \mathrm{d}H(\theta_{\backslash a}) = 0$$

For now, trust me that the term $\int h_a(\theta^\star_{\backslash a}) \int p_\theta(T,\epsilon) \mathrm{d}\epsilon \mathrm{d}H_{\backslash a}(\theta_{\backslash a})$ tends to 0. Then, asymptotically

$$\mathrm{Reg}(T) \sim 2(1-\gamma)^2 \sum_a \int h_a(\theta^\star_{\backslash a}) \int_{\mathcal{E}_T} \frac{\log T\epsilon^2}{\epsilon} \mathrm{d}\epsilon \mathrm{d}H_{\backslash a}(\theta_{\backslash a})$$

Bayesian Regret Lower Bound: Final Steps

$$\mathrm{Reg}(T) \sim 2(1-\gamma)^2 \sum_a \int_{\Theta^{K-1}} h_a(\theta^\star_{\backslash a}) \int_{\mathcal{E}_T} \frac{\log T\epsilon^2}{\epsilon} \mathrm{d}\epsilon \mathrm{d}H_{\backslash a}(\theta_{\backslash a})$$

How to choose $\mathcal{E}_T$? [Lai, 1987] show that $\inf \mathcal{E}_T = T^{-(1-\gamma)/2}$. Then, consider $\mathcal{E}_T = (T^{-(1-\gamma)/2}, 1/\log T)$ (the upper bound need to be $o(T^\alpha)$ for any $\alpha \in (0,1)$). Then

$$\int_{\mathcal{E}_T} \frac{\log T\epsilon^2}{\epsilon} \mathrm{d}\epsilon = \frac{1}{4} \log^2(T\epsilon^2)\Big|_{T^{-(1-\gamma)/2}}^{1/\log T} = \frac{1-\gamma^2}{4} \log^2 T - O(\log^2 \log^2 T).$$

The end. Letting $\gamma \to 0$ and $T \to \infty$, we obtain

$$\liminf_{T\to\infty} \frac{\mathrm{Reg}(T)}{\log^2 T} \geq \frac{1}{2} \sum_a \int_{\Theta^{K-1}} h_a(\theta^\star_{\backslash a}) \mathrm{d}H_{\backslash a}(\theta_{\backslash a}).$$

Bayesian Regret Lower Bound: Final Steps

$$\mathrm{Reg}(T) \sim 2(1-\gamma)^2 \sum_a \int_{\Theta^{K-1}} h_a(\theta^\star_{\backslash a}) \int_{\mathcal{E}_T} \frac{\log T\epsilon^2}{\epsilon} \mathrm{d}\epsilon \mathrm{d}H_{\backslash a}(\theta_{\backslash a})$$

How to choose $\mathcal{E}_T$? [Lai, 1987] show that $\inf \mathcal{E}_T = T^{-(1-\gamma)/2}$. Then, consider $\mathcal{E}_T = (T^{-(1-\gamma)/2}, 1/\log T)$ (the upper bound need to be $o(T^\alpha)$ for any $\alpha \in (0,1)$). Then

$$\int_{\mathcal{E}_T} \frac{\log T\epsilon^2}{\epsilon}\mathrm{d}\epsilon = \frac{1}{4}\log^2(T\epsilon^2)\Big|_{T^{-(1-\gamma)/2}}^{1/\log T} = \frac{1-\gamma^2}{4}\log^2 T - O(\log^2\log^2 T).$$

The end. Letting $\gamma \to 0$ and $T \to \infty$, we obtain

$$\liminf_{T\to\infty} \frac{\mathrm{Reg}(T)}{\log^2 T} \geq \frac{1}{2}\sum_a \int_{\Theta^{K-1}} h_a(\theta^\star_{\backslash a})\mathrm{d}H_{\backslash a}(\theta_{\backslash a}).$$

Bayesian Regret Lower Bound: Final Steps

$$\mathrm{Reg}(T) \sim 2(1-\gamma)^2 \sum_a \int_{\Theta^{K-1}} h_a(\theta^\star_{\setminus a}) \int_{\mathcal{E}_T} \frac{\log T\epsilon^2}{\epsilon} \mathrm{d}\epsilon \mathrm{d}H_{\setminus a}(\theta_{\setminus a})$$

How to choose $\mathcal{E}_T$? [Lai, 1987] show that $\inf \mathcal{E}_T = T^{-(1-\gamma)/2}$. Then, consider $\mathcal{E}_T = (T^{-(1-\gamma)/2}, 1/\log T)$ (the upper bound need to be $o(T^\alpha)$ for any $\alpha \in (0,1)$). Then

$$\int_{\mathcal{E}_T} \frac{\log T\epsilon^2}{\epsilon} \mathrm{d}\epsilon = \frac{1}{4} \log^2(T\epsilon^2)\Big|_{T^{-(1-\gamma)/2}}^{1/\log T} = \frac{1-\gamma^2}{4} \log^2 T - O(\log^2 \log^2 T).$$

The end. Letting $\gamma \to 0$ and $T \to \infty$, we obtain

$$\liminf_{T\to\infty} \frac{\mathrm{Reg}(T)}{\log^2 T} \geq \frac{1}{2} \sum_a \int_{\Theta^{K-1}} h_a(\theta^\star_{\setminus a}) \mathrm{d}H_{\setminus a}(\theta_{\setminus a}).$$

Discussion

What did we not discuss?

Congrats for reaching this point!

There are still some things we have not yet disucssed...

- What is the intuition behind the lower bound?
- Show how the condition on $p_\theta(T, \epsilon)$ makes $\int \int p_\theta(T, \epsilon) \to 0$ (see appendix).
- Why $\inf \mathcal{E}_T = N^{-(1-\gamma)/2}$ (We will not discuss this)
- How is the definition of a Bayes-uniformly fast convergent strategy derived? (We will not discuss this) .
- Is there a simpler way to derive all of this? (unclear).
- Can we use this result to derive optimal algorithms?

What did we not discuss?

Congrats for reaching this point!

There are still some things we have not yet disucssed...

- What is the intuition behind the lower bound?
- Show how the condition on $p_\theta(T, \epsilon)$ makes $\int \int p_\theta(T, \epsilon) \to 0$ (see appendix).
- Why $\inf \mathcal{E}_T = N^{-(1-\gamma)/2}$ (We will not discuss this)
- How is the definition of a Bayes-uniformly fast convergent strategy derived? (We will not discuss this) .
- Is there a simpler way to derive all of this? (unclear).
- Can we use this result to derive optimal algorithms?

What did we not discuss?

Congrats for reaching this point!

There are still some things we have not yet disucssed...

- What is the intuition behind the lower bound?
- Show how the condition on $p_\theta(T, \epsilon)$ makes $\int \int p_\theta(T, \epsilon) \to 0$ (see appendix).
- Why $\inf \mathcal{E}_T = N^{-(1-\gamma)/2}$ (We will not discuss this)
- How is the definition of a Bayes-uniformly fast convergent strategy derived? (We will not discuss this) .
- Is there a simpler way to derive all of this? (unclear).
- Can we use this result to derive optimal algorithms?

What did we not discuss?

Congrats for reaching this point!

There are still some things we have not yet disucssed...

- What is the intuition behind the lower bound?
- Show how the condition on $p_\theta(T, \epsilon)$ makes $\int \int p_\theta(T, \epsilon) \to 0$ (see appendix).
- Why $\inf \mathcal{E}_T = N^{-(1-\gamma)/2}$ (We will not discuss this)
- How is the definition of a Bayes-uniformly fast convergent strategy derived? (We will not discuss this) .
- Is there a simpler way to derive all of this? (unclear).
- Can we use this result to derive optimal algorithms?

What did we not discuss?

Congrats for reaching this point!

There are still some things we have not yet disucssed...

- What is the intuition behind the lower bound?
- Show how the condition on $p_\theta(T, \epsilon)$ makes $\int \int p_\theta(T, \epsilon) \to 0$ (see appendix).
- Why $\inf \mathcal{E}_T = N^{-(1-\gamma)/2}$ (We will not discuss this)
- How is the definition of a Bayes-uniformly fast convergent strategy derived? (We will not discuss this) .
- Is there a simpler way to derive all of this? (unclear).
- Can we use this result to derive optimal algorithms?

What did we not discuss?

Congrats for reaching this point!

There are still some things we have not yet disucssed...

- What is the intuition behind the lower bound?
- Show how the condition on $p_\theta(T, \epsilon)$ makes $\int\int p_\theta(T, \epsilon) \to 0$ (see appendix).
- Why $\inf \mathcal{E}_T = N^{-(1-\gamma)/2}$ (We will not discuss this)
- How is the definition of a Bayes-uniformly fast convergent strategy derived? (We will not discuss this) .
- Is there a simpler way to derive all of this? (unclear).
- Can we use this result to derive optimal algorithms?

What did we not discuss?

Congrats for reaching this point!

There are still some things we have not yet disucssed...

- What is the intuition behind the lower bound?
- Show how the condition on $p_\theta(T, \epsilon)$ makes $\int \int p_\theta(T, \epsilon) \to 0$ (see appendix).
- Why $\inf \mathcal{E}_T = N^{-(1-\gamma)/2}$ (We will not discuss this)
- How is the definition of a Bayes-uniformly fast convergent strategy derived? (We will not discuss this) .
- Is there a simpler way to derive all of this? (unclear).
- Can we use this result to derive optimal algorithms?

What does it mean?

$$\liminf_{T\to\infty} \frac{\mathrm{Reg}(T)}{\log^2 T} \geq \sum_a \underbrace{\frac{1}{2}\int_{\Theta^{K-1}} h_a(\theta^\star_{\setminus a})\mathrm{d}H_{\setminus a}(\theta_{\setminus a})}_{=:K^\star_a}, \qquad K^\star := \sum_a K^\star_a.$$

- The regret is only characterized by the complexity of the priors!
- $K^\star_a$ denotes the complexity for arm a: large $K^\star_a$ implies a larger likelihood that a is close to optimality (if $h_a(\theta^\star_{\setminus a})$ is large, then it becomes harder to distinguish between a and the other good arm).
- Can we simplify $K^\star_a$?

What does it mean?

$$\liminf_{T\to\infty} \frac{\mathrm{Reg}(T)}{\log^2 T} \geq \sum_a \underbrace{\frac{1}{2}\int_{\Theta^{K-1}} h_a(\theta^\star_{\setminus a}) \mathrm{d}H_{\setminus a}(\theta_{\setminus a})}_{=:K^\star_a}, \qquad K^\star := \sum_a K^\star_a.$$

- The regret is only characterized by the complexity of the priors!
- $K^\star_a$ denotes the complexity for arm a: large $K^\star_a$ implies a larger likelihood that a is close to optimality (if $h_a(\theta^\star_{\setminus a})$ is large, then it becomes harder to distinguish between a and the other good arm).
- Can we simplify $K^\star_a$?

What does it mean?

$$\liminf_{T\to\infty} \frac{\mathrm{Reg}(T)}{\log^2 T} \geq \sum_a \underbrace{\frac{1}{2}\int_{\Theta^{K-1}} h_a(\theta^\star_{\setminus a}) \mathrm{d}H_{\setminus a}(\theta_{\setminus a})}_{=: K^\star_a}, \qquad K^\star := \sum_a K^\star_a.$$

- The regret is only characterized by the complexity of the priors!
- $K^\star_a$ denotes the complexity for arm a: large $K^\star_a$ implies a larger likelihood that a is close to optimality (if $h_a(\theta^\star_{\setminus a})$ is large, then it becomes harder to distinguish between a and the other good arm).
- Can we simplify $K^\star_a$?

What does it mean? Simplification.

$$\liminf_{T\to\infty} \frac{\mathrm{Reg}(T)}{\log^2 T} \geq \sum_a \underbrace{\frac{1}{2}\int_{\Theta^{K-1}} h_a(\theta^\star_{\setminus a})\mathrm{d}H_{\setminus a}(\theta_{\setminus a})}_{=:K^\star_a}, \qquad K^\star := \sum_a K^\star_a.$$

Assume i.i.d. $(\theta_a)_a$ (i.i.d. priors) with $h_a \equiv h \ \forall a$ (sim. $H_a \equiv H$). Then $H_{\setminus a} = H^{K-1}$, implying

$$\mathbb{P}_H(\theta^\star_{\setminus a} \leq x) = \mathbb{P}_H(\theta_1 \leq x, \ldots, \theta_{a-1} \leq x, \theta_{a+1} \leq x, \ldots, \theta_K \leq x) = H^{K-1}(x).$$

$$\Rightarrow K^\star = \frac{K}{2}\int_\Theta h(\theta)\mathrm{d}H^{K-1}(\theta)$$

Since $\mathrm{d}H^{K-1}(\theta) = (K-1)H^{K-2}(\theta)\mathrm{d}H(\theta)$ and $\mathrm{d}H(\theta) = h(\theta)\mathrm{d}\theta$

$$K^\star = \frac{K(K-1)}{2}\int_\Theta h^2(\theta)H^{K-2}(\theta)\mathrm{d}\theta$$

What does it mean? Simplification.

$$\liminf_{T\to\infty} \frac{\mathrm{Reg}(T)}{\log^2 T} \geq \sum_a \underbrace{\frac{1}{2}\int_{\Theta^{K-1}} h_a(\theta^\star_{\setminus a})\mathrm{d}H_{\setminus a}(\theta_{\setminus a})}_{=:K^\star_a}, \qquad K^\star := \sum_a K^\star_a.$$

Assume i.i.d. $(\theta_a)_a$ (i.i.d. priors) with $h_a \equiv h \ \forall a$ (sim. $H_a \equiv H$). Then $H_{\setminus a} = H^{K-1}$, implying

$$\mathbb{P}_H(\theta^\star_{\setminus a} \leq x) = \mathbb{P}_H(\theta_1 \leq x, \ldots, \theta_{a-1} \leq x, \theta_{a+1} \leq x, \ldots, \theta_K \leq x) = H^{K-1}(x).$$

$$\Rightarrow K^\star = \frac{K}{2}\int_\Theta h(\theta)\mathrm{d}H^{K-1}(\theta)$$

Since $\mathrm{d}H^{K-1}(\theta) = (K-1)H^{K-2}(\theta)\mathrm{d}H(\theta)$ and $\mathrm{d}H(\theta) = h(\theta)\mathrm{d}\theta$

$$K^\star = \frac{K(K-1)}{2}\int_\Theta h^2(\theta)H^{K-2}(\theta)\mathrm{d}\theta$$

What does it mean? Simplification.

$$\liminf_{T\to\infty} \frac{\mathrm{Reg}(T)}{\log^2 T} \geq \sum_a \underbrace{\frac{1}{2}\int_{\Theta^{K-1}} h_a(\theta^\star_{\backslash a})\mathrm{d}H_{\backslash a}(\theta_{\backslash a})}_{=:K^\star_a}, \qquad K^\star := \sum_a K^\star_a.$$

Assume i.i.d. $(\theta_a)_a$ (i.i.d. priors) with $h_a \equiv h \ \forall a$ (sim. $H_a \equiv H$). Then $H_{\backslash a} = H^{K-1}$, implying

$$\mathbb{P}_H(\theta^\star_{\backslash a} \leq x) = \mathbb{P}_H(\theta_1 \leq x, \ldots, \theta_{a-1} \leq x, \theta_{a+1} \leq x, \ldots, \theta_K \leq x) = H^{K-1}(x).$$

$$\Rightarrow K^\star = \frac{K}{2}\int_\Theta h(\theta)\mathrm{d}H^{K-1}(\theta)$$

Since $\mathrm{d}H^{K-1}(\theta) = (K-1)H^{K-2}(\theta)\mathrm{d}H(\theta)$ and $\mathrm{d}H(\theta) = h(\theta)\mathrm{d}\theta$

$$K^\star = \frac{K(K-1)}{2}\int_\Theta h^2(\theta)H^{K-2}(\theta)\mathrm{d}\theta$$

What does it mean? Simplification.

$$\liminf_{T\to\infty} \frac{\mathrm{Reg}(T)}{\log^2 T} \geq \sum_a \underbrace{\frac{1}{2}\int_{\Theta^{K-1}} h_a(\theta^\star_{\backslash a})\mathrm{d}H_{\backslash a}(\theta_{\backslash a})}_{=:K^\star_a}, \qquad K^\star := \sum_a K^\star_a.$$

Assume i.i.d. $(\theta_a)_a$ (i.i.d. priors) with $h_a \equiv h \ \forall a$ (sim. $H_a \equiv H$). Then $H_{\backslash a} = H^{K-1}$, implying

$$\mathbb{P}_H(\theta^\star_{\backslash a} \leq x) = \mathbb{P}_H(\theta_1 \leq x, \ldots, \theta_{a-1} \leq x, \theta_{a+1} \leq x, \ldots, \theta_K \leq x) = H^{K-1}(x).$$

$$\Rightarrow K^\star = \frac{K}{2}\int_\Theta h(\theta)\mathrm{d}H^{K-1}(\theta)$$

Since $\mathrm{d}H^{K-1}(\theta) = (K-1)H^{K-2}(\theta)\mathrm{d}H(\theta)$ and $\mathrm{d}H(\theta) = h(\theta)\mathrm{d}\theta$

$$K^\star = \frac{K(K-1)}{2}\int_\Theta h^2(\theta)H^{K-2}(\theta)\mathrm{d}\theta$$

Connections to order statistics

$$K^\star = \frac{K(K-1)}{2} \int_\Theta h^2(\theta) H^{K-2}(\theta) \mathrm{d}\theta$$

there is actually a connection to order statistics.

Order statistics. Given a random vector $\theta = (\theta_1, \ldots, \theta_K)$, sort the components into a vector $(\theta_{(1)}, \ldots, \theta_{(K)})$ satisfying

$$\theta_{(1)} \leq \theta_{(2)} \leq \cdots \leq \theta_{(K)}.$$

This vector is called the order statistics of θ.

- The joint pdf f of (θ_{K-1}, θ_K) (with cdf F) is [Casella and Berger, 2024]

$$f(x, y) = K(K-1) f(x) f(y) F^{K-2}(x)$$

Letting $x \to y$ we find $\lim_{x \to y} f(x, y) = K(K-1) f^2(y) F^{K-2}(y)$.This is the limiting contribution when the two upper-most samples almost tie.

Connections to order statistics

$$K^\star = \frac{K(K-1)}{2} \int_\Theta h^2(\theta) H^{K-2}(\theta) \mathrm{d}\theta$$

there is actually a connection to order statistics.

Order statistics. Given a random vector $\theta = (\theta_1, \ldots, \theta_K)$, sort the components into a vector $(\theta_{(1)}, \ldots, \theta_{(K)})$ satisfying

$$\theta_{(1)} \leq \theta_{(2)} \leq \cdots \leq \theta_{(K)}.$$

This vector is called the order statistics of θ.

- The joint pdf f of (θ_{K-1}, θ_K) (with cdf F) is [Casella and Berger, 2024]

$$f(x, y) = K(K-1) f(x) f(y) F^{K-2}(x)$$

Letting $x \to y$ we find $\lim_{x \to y} f(x, y) = K(K-1) f^2(y) F^{K-2}(y)$.This is the limiting contribution when the two upper-most samples almost tie.

Connections to order statistics

$$K^\star = \frac{K(K-1)}{2} \int_\Theta h^2(\theta) H^{K-2}(\theta) \mathrm{d}\theta$$

there is actually a connection to order statistics.

Order statistics. Given a random vector $\theta = (\theta_1, \dots, \theta_K)$, sort the components into a vector $(\theta_{(1)}, \dots, \theta_{(K)})$ satisfying

$$\theta_{(1)} \leq \theta_{(2)} \leq \cdots \leq \theta_{(K)}.$$

This vector is called the order statistics of θ.

- The joint pdf f of (θ_{K-1}, θ_K) (with cdf F) is [Casella and Berger, 2024]

$$f(x, y) = K(K-1) f(x) f(y) F^{K-2}(x)$$

Letting $x \to y$ we find $\lim_{x \to y} f(x, y) = K(K-1) f^2(y) F^{K-2}(y)$.This is the limiting contribution when the two upper-most samples almost tie.

Other interpreation

Since $\lim_{x \to y} f(x, y) = K(K-1) f^2(y) F^{K-2}(y)$, another connection is to see the overall integral as the chance that the top two samples fall into it a tiny interval

$$\begin{aligned} \mathbb{P}(|\theta_{(K)} - \theta_{(K-1)}| < \epsilon) &= K(K-1) \int_\Theta \int_0^\epsilon h(\theta) h(\theta + \epsilon) H^{K-2}(\theta) \, \mathrm{d}\epsilon \, \mathrm{d}\theta, \\ &= 2K^\star \epsilon + o(\epsilon) \end{aligned}$$

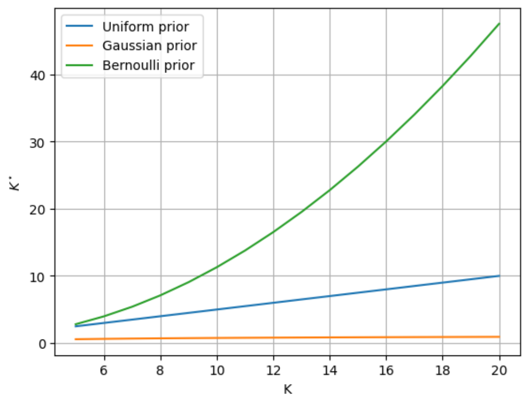
Thus

$$\frac{\mathbb{P}(|\theta_{(K)} - \theta_{(K-1)}| < \epsilon)}{\epsilon} \underset{\epsilon \to 0}{\to} 2K^\star.$$

Scaling of $K^\star$

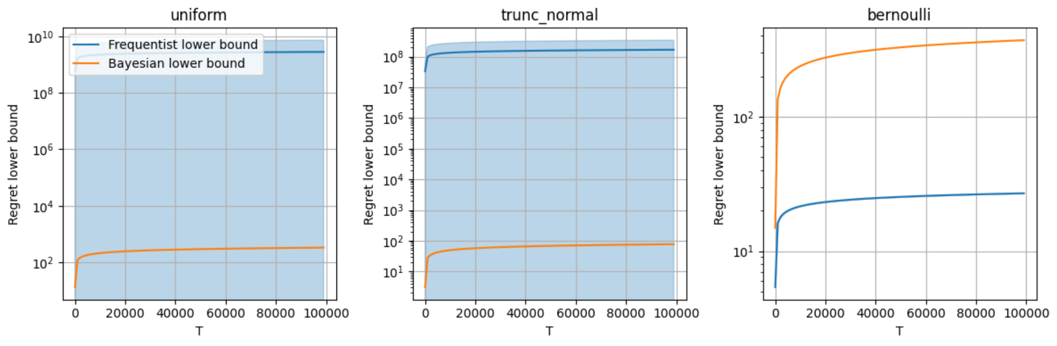
Scaling of $K^\star$ vs K, with: (1) $H = \mathcal{U}([0,1])$; (2) $H = \mathcal{N}(0,1)$ and (3) $H = \mathrm{Ber}(0.5)$ (uniform, Gaussian, Bernoulli).

We consider a MAB problem with K arms and Gaussian rewards $\mathcal{N}(\theta_a, 1)$, with θ_a drawn iid from the prior.



Bayesian vs Frequentist Regret Lower Bound

- Can we just compute the average frequentist lower bound over many different problems?



Same setting as in the previous slide, with $K = 5$. Average computed over 3000000 sampled MAB problems. Shaded areas indicate the 95% C.I.

The frequentist lower bound simply explodes with continuous priors.

Algorithm Design

Can we use the lower bound to design asymptotically optimal algorithms? Probably. I believe the intuition is to solve the following problem

$$\inf_{\eta} \sum_a \int \eta_a \Delta_a(\theta) \mathrm{d}H(\theta) \text{ s.t. } \int_{\Theta^{K-1}} \eta_a D(\theta_a, \theta^\star_{\setminus a}) h_a(\theta_a) \mathrm{d}H_{\setminus a}(\theta_{\setminus a}) \geq 1$$

where $\eta \in \Delta(K)$ represents the proportion of times we should play each arm [8]

[8] This is probably incorrect, but the true problem should vaguely resemble this one.

Algorithm Design

Can we use the lower bound to design asymptotically optimal algorithms? Probably. I believe the intuition is to solve the following problem

$$\inf_{\eta} \sum_a \int \eta_a \Delta_a(\theta) \mathrm{d}H(\theta) \text{ s.t. } \int_{\Theta^{K-1}} \eta_a D(\theta_a, \theta^\star_{\backslash a}) h_a(\theta_a) \mathrm{d}H_{\backslash a}(\theta_{\backslash a}) \geq 1$$

where $\eta \in \Delta(K)$ represents the proportion of times we should play each arm [8]

[8] This is probably incorrect, but the true problem should vaguely resemble this one.

Conclusion

Conclusions[9]

Possible extensions:

- Optimal algorithms based on the lower bound.
- Bayesian regret lower bounds for MDPs.
- A more comprehensive analysis of Bayesian BAI.

[9] Credits to Flaticon.com for some of the logos used in this presentation.

Conclusions[9]

Possible extensions:

- Optimal algorithms based on the lower bound.
- Bayesian regret lower bounds for MDPs.
- A more comprehensive analysis of Bayesian BAI.

[9] Credits to Flaticon.com for some of the logos used in this presentation.

Conclusions[9]

Possible extensions:

- Optimal algorithms based on the lower bound.
- Bayesian regret lower bounds for MDPs.
- A more comprehensive analysis of Bayesian BAI.

[9] Credits to Flaticon.com for some of the logos used in this presentation.

References i

Atsidakou, A., Kveton, B., Katariya, S., Caramanis, C., and Sanghavi, S. (2023).
Finite-time logarithmic bayes regret upper bounds.
In Oh, A., Naumann, T., Globerson, A., Saenko, K., Hardt, M., and Levine, S., editors, *Advances in Neural Information Processing Systems*, volume 36, pages 4331–4350. Curran Associates, Inc.

Casella, G. and Berger, R. (2024).
Statistical inference.
CRC press.

Efron, B. (2022).
Exponential Families in Theory and Practice.
Institute of Mathematical Statistics Textbooks. Cambridge University Press, Cambridge.

Fabius, J. and Zwet, W. R. V. (1970).
Some remarks on the two-armed bandit.
The Annals of Mathematical Statistics, 41(6):1906–1916.

Gittins, J. C. (1979).
Bandit Processes and Dynamic Allocation Indices.
Journal of the Royal Statistical Society. Series B (Methodological), 41(2):148–177.

Lai, T. L. (1987).
Adaptive Treatment Allocation and the Multi-Armed Bandit Problem.
The Annals of Statistics, 15(3):1091–1114.

Appendix

Condition on $p_\theta(T, \epsilon)$

We need to show that

$$(*) = \int_{\Theta^{K-1}} h_a(\theta^\star_{\setminus a}) \int_{\mathcal{E}_T} p_\theta(T, \epsilon)\mathrm{d}\epsilon \mathrm{d}H_{\setminus a}(\theta_{\setminus a}) \to 0,$$

where $(\mathcal{E}_T)_T$ is a sequence of open sets, satisfying such that $\lambda(\mathcal{E}_T) < \infty$ (Lebesgue measure) for all T, with $\sup E_T \leq \sup E_{T-1}, \inf \mathcal{E}_T \leq \inf \mathcal{E}_{T-1}$ and $\mathcal{E}_T \underset{T\to\infty}{\to} \{0\}$.

Note

$$(*) \leq \lambda(\mathcal{E}_T) \int_{\Theta^{K-1}} [\sup_{\epsilon \in \mathcal{E}_T} p_\theta(T, \epsilon)] h_a(\theta^\star_{\setminus a}) \mathrm{d}H_{\setminus a}(\theta_{\setminus a}).$$

We show how we can rewrite the original condition in this form.

Condition on $p_\theta(T, \epsilon)$

We need to show that

$$(*) = \int_{\Theta^{K-1}} h_a(\theta^\star_{\setminus a}) \int_{\mathcal{E}_T} p_\theta(T, \epsilon) \mathrm{d}\epsilon \mathrm{d}H_{\setminus a}(\theta_{\setminus a}) \to 0,$$

where $(\mathcal{E}_T)_T$ is a sequence of open sets, satisfying such that $\lambda(\mathcal{E}_T) < \infty$ (Lebesgue measure) for all T, with $\sup E_T \leq \sup E_{T-1}, \inf \mathcal{E}_T \leq \inf \mathcal{E}_{T-1}$ and $\mathcal{E}_T \underset{T\to\infty}{\to} \{0\}$.

Note

$$(*) \leq \lambda(\mathcal{E}_T) \int_{\Theta^{K-1}} [\sup_{\epsilon \in \mathcal{E}_T} p_\theta(T, \epsilon)] h_a(\theta^\star_{\setminus a}) \mathrm{d}H_{\setminus a}(\theta_{\setminus a}).$$

We show how we can rewrite the original condition in this form.

Bayesian Regret Lower Bound: Final Steps

$$\lim_{T\to\infty,\epsilon\to 0,T\epsilon^2\to\infty} \underbrace{\int_{\Theta^{K-1}} p_\theta(T,\epsilon)h_a(\theta^\star_{\setminus a})\mathrm{d}H(\theta_{\setminus a})}_{g_T(\epsilon)} = 0$$

This limit also implies that

$$\forall\delta>0\ \exists T_\delta\in\mathbb{N},\alpha_\delta,K_\delta\in\mathbb{R}_+ : g_T(\epsilon)<\delta \ \text{ whenever } \ (T,\epsilon)\in\{T,\epsilon : T\geq T_\delta,\epsilon\leq\alpha_\delta,T\epsilon^2\geq K_\delta\}.$$

Consider the set $\mathcal{E}_T = (1/T^{(1-\gamma)/2}, \log^{-1} T)$. Then

- $\forall\alpha_\delta\ \exists T'_\delta :\ \log^{-1} T\leq\alpha_\delta$ whenever $T\geq T'_\delta$.
- $\forall K_\delta\ \exists T''_\delta :\ T^\gamma\geq K_\delta$ whenever $T\geq T''_\delta$.

Set $T^\star_\delta = \max(T_\delta, T'_\delta, T''_\delta)$. Then

$$\forall\delta>0\ \exists T^\star_\delta\in\mathbb{N} : \sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon)<\delta \ \text{ whenever } T\geq T^\star_\delta.$$

Bayesian Regret Lower Bound: Final Steps

$$\lim_{T\to\infty,\epsilon\to 0,T\epsilon^2\to\infty} \underbrace{\int_{\Theta^{K-1}} p_\theta(T,\epsilon)h_a(\theta^\star_{\setminus a})\mathrm{d}H(\theta_{\setminus a})}_{g_T(\epsilon)} = 0$$

This limit also implies that

$\forall\delta>0\ \exists T_\delta\in\mathbb{N},\alpha_\delta,K_\delta\in\mathbb{R}_+ : g_T(\epsilon)<\delta$ whenever $(T,\epsilon)\in\{T,\epsilon : T\geq T_\delta, \epsilon\leq\alpha_\delta, T\epsilon^2\geq K_\delta\}$.

Consider the set $\mathcal{E}_T = (1/T^{(1-\gamma)/2}, \log^{-1} T)$. Then

- $\forall\alpha_\delta\ \exists T'_\delta :\ \log^{-1} T\leq\alpha_\delta$ whenever $T\geq T'_\delta$.
- $\forall K_\delta\ \exists T''_\delta :\ T^\gamma\geq K_\delta$ whenever $T\geq T''_\delta$.

Set $T^\star_\delta = \max(T_\delta, T'_\delta, T''_\delta)$. Then

$$\forall\delta>0\ \exists T^\star_\delta\in\mathbb{N} : \sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon)<\delta \ \text{ whenever } T\geq T^\star_\delta.$$

Bayesian Regret Lower Bound: Final Steps

$$\lim_{T\to\infty,\epsilon\to 0,T\epsilon^2\to\infty} \underbrace{\int_{\Theta^{K-1}} p_\theta(T,\epsilon) h_a(\theta^\star_{\setminus a}) \mathrm{d}H(\theta_{\setminus a})}_{g_T(\epsilon)} = 0$$

This limit also implies that

$$\forall \delta > 0 \; \exists T_\delta \in \mathbb{N}, \alpha_\delta, K_\delta \in \mathbb{R}_+ : g_T(\epsilon) < \delta \;\text{ whenever }\; (T,\epsilon) \in \{T,\epsilon : T \geq T_\delta, \epsilon \leq \alpha_\delta, T\epsilon^2 \geq K_\delta\}.$$

Consider the set $\mathcal{E}_T = (1/T^{(1-\gamma)/2}, \log^{-1} T)$. Then

- $\forall \alpha_\delta \; \exists T'_\delta : \; \log^{-1} T \leq \alpha_\delta$ whenever $T \geq T'_\delta$.
- $\forall K_\delta \; \exists T''_\delta : \; T^\gamma \geq K_\delta$ whenever $T \geq T''_\delta$.

Set $T^\star_\delta = \max(T_\delta, T'_\delta, T''_\delta)$. Then

$$\forall \delta > 0 \; \exists T^\star_\delta \in \mathbb{N} : \sup_{\epsilon \in \mathcal{E}_T} g_T(\epsilon) < \delta \;\text{ whenever }\; T \geq T^\star_\delta.$$

Bayesian Regret Lower Bound: Final Steps

$$\lim_{T\to\infty,\epsilon\to 0,T\epsilon^2\to\infty} \underbrace{\int_{\Theta^{K-1}} p_\theta(T,\epsilon)h_a(\theta^\star_{\backslash a})\mathrm{d}H(\theta_{\backslash a})}_{g_T(\epsilon)} = 0$$

This limit also implies that

$$\forall\delta>0\ \exists T_\delta\in\mathbb{N},\alpha_\delta,K_\delta\in\mathbb{R}_+ : g_T(\epsilon)<\delta \ \text{ whenever } \ (T,\epsilon)\in\{T,\epsilon : T\geq T_\delta, \epsilon\leq\alpha_\delta, T\epsilon^2\geq K_\delta\}.$$

Consider the set $\mathcal{E}_T = (1/T^{(1-\gamma)/2}, \log^{-1} T)$. Then

- $\forall\alpha_\delta\ \exists T'_\delta :\ \log^{-1} T \leq \alpha_\delta$ whenever $T\geq T'_\delta$.
- $\forall K_\delta\ \exists T''_\delta :\ T^\gamma \geq K_\delta$ whenever $T\geq T''_\delta$.

Set $T^\star_\delta = \max(T_\delta, T'_\delta, T''_\delta)$. Then

$$\forall\delta>0\ \exists T^\star_\delta\in\mathbb{N} : \sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon)<\delta \ \text{ whenever } T\geq T^\star_\delta.$$

Bayesian Regret Lower Bound: Final Steps

Therefore

$$\lim_{T\to\infty,\epsilon\to 0,T\epsilon^2\to\infty} g_T(\epsilon) = \lim_{T\to\infty}\sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) = 0.$$

Then, observe that by Tonelli-Fubini

$$\begin{aligned}\lambda(\mathcal{E}_T)\sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) &\geq \int_{\mathcal{E}_T}\int_{\Theta^{K-1}} p_\theta(T,\epsilon)h_a(\theta^\star_{\setminus a})\mathrm{d}H(\theta_{\setminus a})\mathrm{d}\epsilon,\\ &= \int_{\Theta^{K-1}}\int_{\mathcal{E}_T} p_\theta(T,\epsilon)h_a(\theta^\star_{\setminus a})\mathrm{d}\epsilon\mathrm{d}H(\theta_{\setminus a}) \geq 0.\end{aligned}$$

Since $\lambda(\mathcal{E}_T) < \infty$ and both $\sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) \to 0$ and $\lambda(\mathcal{E}_T)\to 0$, then $\lambda(\mathcal{E}_T)\sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) \to 0$,

Bayesian Regret Lower Bound: Final Steps

Therefore

$$\lim_{T\to\infty,\epsilon\to 0,T\epsilon^2\to\infty} g_T(\epsilon) = \lim_{T\to\infty} \sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) = 0.$$

Then, observe that by Tonelli-Fubini

$$\lambda(\mathcal{E}_T) \sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) \geq \int_{\mathcal{E}_T}\int_{\Theta^{K-1}} p_\theta(T,\epsilon) h_a(\theta^\star_{\setminus a}) \mathrm{d}H(\theta_{\setminus a}) \mathrm{d}\epsilon,$$

$$= \int_{\Theta^{K-1}}\int_{\mathcal{E}_T} p_\theta(T,\epsilon) h_a(\theta^\star_{\setminus a}) \mathrm{d}\epsilon \mathrm{d}H(\theta_{\setminus a}) \geq 0.$$

Since $\lambda(\mathcal{E}_T) < \infty$ and both $\sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) \to 0$ and $\lambda(\mathcal{E}_T) \to 0$, then $\lambda(\mathcal{E}_T) \sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) \to 0$,

Bayesian Regret Lower Bound: Final Steps

Therefore

$$\lim_{T\to\infty,\epsilon\to 0,T\epsilon^2\to\infty} g_T(\epsilon) = \lim_{T\to\infty} \sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) = 0.$$

Then, observe that by Tonelli-Fubini

$$\begin{aligned}
\lambda(\mathcal{E}_T) \sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) &\geq \int_{\mathcal{E}_T}\int_{\Theta^{K-1}} p_\theta(T,\epsilon) h_a(\theta^\star_{\setminus a}) \mathrm{d}H(\theta_{\setminus a}) \mathrm{d}\epsilon, \\
&= \int_{\Theta^{K-1}}\int_{\mathcal{E}_T} p_\theta(T,\epsilon) h_a(\theta^\star_{\setminus a}) \mathrm{d}\epsilon \mathrm{d}H(\theta_{\setminus a}) \geq 0.
\end{aligned}$$

Since $\lambda(\mathcal{E}_T) < \infty$ and both $\sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) \to 0$ and $\lambda(\mathcal{E}_T) \to 0$, then $\lambda(\mathcal{E}_T) \sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) \to 0$,

Bayesian Regret Lower Bound: Final Steps

Therefore

$$\lim_{T\to\infty,\epsilon\to 0,T\epsilon^2\to\infty} g_T(\epsilon) = \lim_{T\to\infty} \sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) = 0.$$

Then, observe that by Tonelli-Fubini

$$\begin{aligned}\lambda(\mathcal{E}_T) \sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) &\geq \int_{\mathcal{E}_T} \int_{\Theta^{K-1}} p_\theta(T,\epsilon) h_a(\theta^\star_{\setminus a}) \mathrm{d}H(\theta_{\setminus a})\mathrm{d}\epsilon, \\ &= \int_{\Theta^{K-1}} \int_{\mathcal{E}_T} p_\theta(T,\epsilon) h_a(\theta^\star_{\setminus a}) \mathrm{d}\epsilon\mathrm{d}H(\theta_{\setminus a}) \geq 0.\end{aligned}$$

Since $\lambda(\mathcal{E}_T) < \infty$ and both $\sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) \to 0$ and $\lambda(\mathcal{E}_T) \to 0$, then $\lambda(\mathcal{E}_T) \sup_{\epsilon\in\mathcal{E}_T} g_T(\epsilon) \to 0$,