Tutorial: Sample Complexity Lower Bounds

The i.i.d. case

Alessio Russo

September 2024

Boston University

Overview i

Introduction

Sample complexity and lower bounds

- Sample complexity refers to the "number of samples" you need to achieve some learning objective.
- A lower bound characterizes the minimum number of samples you need to achieve such objective.
  - These bounds can be minimax or instance-specific.
- In this tutorial we will study a general recipe to derive instance-specific lower bounds in the I.I.D. setting.
  - The idea is to study the problem from a hypothesis testing perspective.
  - Long history around these ideas [Wal47, Lor71, Lai81, LR85, BK96, GL97, BK97, Lai98, GK16, GMS19].

Sample complexity and lower bounds

- Sample complexity refers to the "number of samples" you need to achieve some learning objective.
- A lower bound characterizes the minimum number of samples you need to achieve such objective.
  - These bounds can be minimax or instance-specific.
- In this tutorial we will study a general recipe to derive instance-specific lower bounds in the I.I.D. setting.
  - The idea is to study the problem from a hypothesis testing perspective.
  - Long history around these ideas [Wal47, Lor71, Lai81, LR85, BK96, GL97, BK97, Lai98, GK16, GMS19].

Sample complexity and lower bounds

- Sample complexity refers to the "number of samples" you need to achieve some learning objective.
- A lower bound characterizes the minimum number of samples you need to achieve such objective.
  - These bounds can be minimax or instance-specific.
- In this tutorial we will study a general recipe to derive instance-specific lower bounds in the I.I.D. setting.
  - The idea is to study the problem from a hypothesis testing perspective.
  - Long history around these ideas [Wal47, Lor71, Lai81, LR85, BK96, GL97, BK97, Lai98, GK16, GMS19].

A brief introduction

Consider the following simple example:

- You have created a new drug  , and want to determine its efficacy.
    - You test the drug on T users, and check the effect of this drug.
    - We assume to observe IID Bernoulli rewards $X_1, X_2, \ldots, X_T$, with $X_i \sim \mathrm{Ber}(\mu)$.
- You believe the expected efficacy of your drug to be at-least $\lambda \in (0, 1)$.
- How many samples T do you need to refuse this hypothesis, if it's not true, with some confidence δ?

Sample complexity lower bounds help us understand, on average, how many samples we need for some value of δ.

A brief introduction

Consider the following simple example:

- You have created a new drug  , and want to determine its efficacy.
  - You test the drug on T users, and check the effect of this drug.
  - We assume to observe IID Bernoulli rewards $X_1, X_2, \dots, X_T$, with $X_i \sim \mathrm{Ber}(\mu)$.
- You believe the expected efficacy of your drug to be at-least $\lambda \in (0, 1)$.
- How many samples T do you need to refuse this hypothesis, if it's not true, with some confidence δ?

Sample complexity lower bounds help us understand, on average, how many samples we need for some value of δ.

A brief introduction

Consider the following simple example:

- You have created a new drug  , and want to determine its efficacy.
  - You test the drug on T users, and check the effect of this drug.
  - We assume to observe IID Bernoulli rewards $X_1, X_2, \ldots, X_T$, with $X_i \sim \mathrm{Ber}(\mu)$.
- You believe the expected efficacy of your drug to be at-least $\lambda \in (0, 1)$.
- How many samples T do you need to refuse this hypothesis, if it's not true, with some confidence δ?

Sample complexity lower bounds help us understand, on average, how many samples we need for some value of δ.

A brief introduction

Consider the following simple example:

- You have created a new drug , and want to determine its efficacy.
  - You test the drug on T users, and check the effect of this drug.
  - We assume to observe IID Bernoulli rewards $X_1, X_2, \ldots, X_T$, with $X_i \sim \mathrm{Ber}(\mu)$.
- You believe the expected efficacy of your drug to be at-least $\lambda \in (0, 1)$.
- How many samples T do you need to refuse this hypothesis, if it's not true, with some confidence δ?

Sample complexity lower bounds help us understand, on average, how many samples we need for some value of δ.

A brief introduction

Consider the following simple example:

- You have created a new drug  , and want to determine its efficacy.
  - You test the drug on T users, and check the effect of this drug.
  - We assume to observe IID Bernoulli rewards $X_1, X_2, \ldots, X_T$, with $X_i \sim \mathrm{Ber}(\mu)$.
- You believe the expected efficacy of your drug to be at-least $\lambda \in (0, 1)$.
- How many samples T do you need to refuse this hypothesis, if it's not true, with some confidence δ?

Sample complexity lower bounds help us understand, on average, how many samples we need for some value of δ.

A brief introduction (cont.)

- The expected efficacy μ is unknown. You believe it to be at-least λ.
- Frame the problem as a statistical hypothesis testing problem.

Hypothesis Testing Problem

$$H_0 : \mu \geq \lambda \text{ (null hypothesis)} \quad \text{vs} \quad H_1 : \mu < \lambda \text{ (alternative hypothesis)}$$

How many samples T do you need to refuse H_0 if H_1 is true?

A brief introduction (cont.)

- The expected efficacy μ is unknown. You believe it to be at-least λ.
- Frame the problem as a statistical hypothesis testing problem.

Hypothesis Testing Problem

$$H_0 : \mu \geq \lambda \text{ (null hypothesis)} \quad \text{vs} \quad H_1 : \mu < \lambda \text{ (alternative hypothesis)}$$

How many samples T do you need to refuse H_0 if H_1 is true?

A brief introduction (cont.)

Hypothesis Testing Problem

$$H_0 : \mu \geq \lambda \text{ (null hypothesis)} \quad \text{vs} \quad H_1 : \mu < \lambda \text{ (alternative hypothesis)}$$

- Let $\hat{\mu}_t = \frac{1}{t}\sum_{i=1}^{t} X_i$.
- Define the gap $\Delta := |\lambda - \mu|$ and consider the event $\{\hat{\mu}_T \geq \lambda\}$ under H_1.

$$\mathbb{P}_{H_1}(\hat{\mu}_T \geq \lambda) = \mathbb{P}_{H_1}(\hat{\mu}_T - \mu \geq \Delta) \leq \exp(-2T\Delta^2) = \delta \text{ if } T = \frac{1}{2}\frac{\ln(1/\delta)}{\Delta^2}.$$

where we used Hoeffding inequality.

- Under H_1 we need $T = \frac{1}{2}\frac{\ln(1/\delta)}{\Delta^2}$ samples to refuse the null hypothesis with some confidence δ.
- The power[1] of the test is $\geq 1 - \delta$, and δ upper bounds the type II error (false negative).

[1] The power of a test is the probability that the test correctly rejects H_0 when H_1 is true.

A brief introduction (cont.)

Hypothesis Testing Problem

$$H_0 : \mu \geq \lambda \text{ (null hypothesis)} \quad \text{vs} \quad H_1 : \mu < \lambda \text{ (alternative hypothesis)}$$

- Let $\hat{\mu}_t = \frac{1}{t}\sum_{i=1}^{t} X_i$.
- Define the gap $\Delta := |\lambda - \mu|$ and consider the event $\{\hat{\mu}_T \geq \lambda\}$ under H_1.

$$\mathbb{P}_{H_1}(\hat{\mu}_T \geq \lambda) = \mathbb{P}_{H_1}(\hat{\mu}_T - \mu \geq \Delta) \leq \exp(-2T\Delta^2) = \delta \text{ if } T = \frac{1}{2}\frac{\ln(1/\delta)}{\Delta^2}.$$

where we used Hoeffding inequality.

- Under H_1 we need $T = \frac{1}{2}\frac{\ln(1/\delta)}{\Delta^2}$ samples to refuse the null hypothesis with some confidence δ.
- The power[1] of the test is $\geq 1 - \delta$, and δ upper bounds the type II error (false negative).

[1] The power of a test is the probability that the test correctly rejects H_0 when H_1 is true.

A brief introduction (cont.)

Hypothesis Testing Problem

$$H_0 : \mu \geq \lambda \text{ (null hypothesis)} \quad \text{vs} \quad H_1 : \mu < \lambda \text{ (alternative hypothesis)}$$

- Let $\hat{\mu}_t = \frac{1}{t}\sum_{i=1}^{t} X_i$.
- Define the gap $\Delta := |\lambda - \mu|$ and consider the event $\{\hat{\mu}_T \geq \lambda\}$ under H_1.

$$\mathbb{P}_{H_1}(\hat{\mu}_T \geq \lambda) = \mathbb{P}_{H_1}(\hat{\mu}_T - \mu \geq \Delta) \leq \exp(-2T\Delta^2) = \delta \text{ if } T = \frac{1}{2}\frac{\ln(1/\delta)}{\Delta^2}.$$

  where we used Hoeffding inequality.
- Under H_1 we need $T = \frac{1}{2}\frac{\ln(1/\delta)}{\Delta^2}$ samples to refuse the null hypothesis with some confidence δ.
- The power[1] of the test is $\geq 1 - \delta$, and δ upper bounds the type II error (false negative).

[1] The power of a test is the probability that the test correctly rejects H_0 when H_1 is true.

A brief introduction (cont.)

Hypothesis Testing Problem

$$H_0 : \mu \geq \lambda \text{ (null hypothesis)} \quad \text{vs} \quad H_1 : \mu < \lambda \text{ (alternative hypothesis)}$$

- Let $\hat{\mu}_t = \frac{1}{t}\sum_{i=1}^{t} X_i$.
- Define the gap $\Delta := |\lambda - \mu|$ and consider the event $\{\hat{\mu}_T \geq \lambda\}$ under H_1.

$$\mathbb{P}_{H_1}(\hat{\mu}_T \geq \lambda) = \mathbb{P}_{H_1}(\hat{\mu}_T - \mu \geq \Delta) \leq \exp(-2T\Delta^2) = \delta \text{ if } T = \frac{1}{2}\frac{\ln(1/\delta)}{\Delta^2}.$$

  where we used Hoeffding inequality.
- Under H_1 we need $T = \frac{1}{2}\frac{\ln(1/\delta)}{\Delta^2}$ samples to refuse the null hypothesis with some confidence δ.
- The power[1] of the test is $\geq 1 - \delta$, and δ upper bounds the type II error (false negative).

[1] The power of a test is the probability that the test correctly rejects H_0 when H_1 is true.

A brief introduction (cont.)

Hypothesis Testing Problem

$$H_0 : \mu \geq \lambda \text{ (null hypothesis)} \quad \text{vs} \quad H_1 : \mu < \lambda \text{ (alternative hypothesis)}$$

However...

- We assumed to have collected T samples, and to test after T samples.
  1. What if you want to assess at every time-step t? $\Rightarrow$ sequential hypothesis testing problem.
- What about Type I errors? (false positives).
- The variance of the process is not taken into account in this simple example.
- Is it possible to do better? How can we prove a lower bound?

A brief introduction (cont.)

Hypothesis Testing Problem

$$H_0 : \mu \geq \lambda \text{ (null hypothesis)} \quad \text{vs} \quad H_1 : \mu < \lambda \text{ (alternative hypothesis)}$$

However...

- We assumed to have collected T samples, and to test after T samples.
  1. What if you want to assess at every time-step t? $\Rightarrow$ sequential hypothesis testing problem.
- What about Type I errors? (false positives).
- The variance of the process is not taken into account in this simple example.
- Is it possible to do better? How can we prove a lower bound?

A brief introduction (cont.)

Hypothesis Testing Problem

$$H_0 : \mu \geq \lambda \text{ (null hypothesis)} \quad \text{vs} \quad H_1 : \mu < \lambda \text{ (alternative hypothesis)}$$

However...

- We assumed to have collected T samples, and to test after T samples.
    1. What if you want to assess at every time-step t? $\Rightarrow$ sequential hypothesis testing problem.
- What about Type I errors? (false positives).
- The variance of the process is not taken into account in this simple example.
- Is it possible to do better? How can we prove a lower bound?

A brief introduction: lower bound

We look at the previous problem from a sequential perspective: you start to collect data, and stop as soon as the null hypothesis is rejected.

- Let τ be a stopping time that tells you when to stop (it's a random variable w.r.t. $\mathcal{F}_t = \sigma(X_1, \dots, X_t)$).
- To find a lower bound, we start by considering the log-likelihood ratio (LLR) of the data observed up to time t between two models $(\mu_1, \mu_0) \in [0,1]^2$:

$$Z_t(\mu_1, \mu_0) := \ln \frac{\mathrm{d}\mathbb{P}_{\mu_1}(X_1, \dots, X_t)}{\mathrm{d}\mathbb{P}_{\mu_0}(X_1, \dots, X_t)}.$$

- What is $\mathbb{E}_{\mu_1}[Z_\tau(\mu_1, \mu_0)]$?

A brief introduction: lower bound

We look at the previous problem from a sequential perspective: you start to collect data, and stop as soon as the null hypothesis is rejected.

- Let τ be a stopping time that tells you when to stop (it's a random variable w.r.t. $\mathcal{F}_t = \sigma(X_1, \ldots, X_t)$).
- To find a lower bound, we start by considering the log-likelihood ratio (LLR) of the data observed up to time t between two models $(\mu_1, \mu_0) \in [0,1]^2$:

$$Z_t(\mu_1, \mu_0) := \ln \frac{\mathrm{d}\mathbb{P}_{\mu_1}(X_1, \ldots, X_t)}{\mathrm{d}\mathbb{P}_{\mu_0}(X_1, \ldots, X_t)}.$$

- What is $\mathbb{E}_{\mu_1}[Z_\tau(\mu_1, \mu_0)]$?

A brief introduction: lower bound

We look at the previous problem from a sequential perspective: you start to collect data, and stop as soon as the null hypothesis is rejected.

- Let τ be a stopping time that tells you when to stop (it's a random variable w.r.t. $\mathcal{F}_t = \sigma(X_1, \dots, X_t)$).
- To find a lower bound, we start by considering the log-likelihood ratio (LLR) of the data observed up to time t between two models $(\mu_1, \mu_0) \in [0,1]^2$:

$$Z_t(\mu_1, \mu_0) := \ln \frac{\mathrm{d}\mathbb{P}_{\mu_1}(X_1, \dots, X_t)}{\mathrm{d}\mathbb{P}_{\mu_0}(X_1, \dots, X_t)}.$$

- What is $\mathbb{E}_{\mu_1}[Z_\tau(\mu_1, \mu_0)]$?

A brief introduction: lower bound

We look at the previous problem from a sequential perspective: you start to collect data, and stop as soon as the null hypothesis is rejected.

- Let τ be a stopping time that tells you when to stop (it's a random variable w.r.t. $\mathcal{F}_t = \sigma(X_1, \ldots, X_t)$).
- To find a lower bound, we start by considering the log-likelihood ratio (LLR) of the data observed up to time t between two models $(\mu_1, \mu_0) \in [0,1]^2$:

$$Z_t(\mu_1, \mu_0) := \ln \frac{\mathrm{d}\mathbb{P}_{\mu_1}(X_1, \ldots, X_t)}{\mathrm{d}\mathbb{P}_{\mu_0}(X_1, \ldots, X_t)}.$$

- What is $\mathbb{E}_{\mu_1}[Z_\tau(\mu_1, \mu_0)]$?

A brief introduction: lower bound (cont.)

At the stopping time we have:

$$\begin{aligned}
\mathbb{E}_{\mu_1}[Z_\tau(\mu_1,\mu_0)] &= \mathbb{E}_{\mu_1}\left[\ln \frac{\mathrm{d}\mathbb{P}_{\mu_1}(X_1,\dots,X_\tau)}{\mathrm{d}\mathbb{P}_{\mu_0}(X_1,\dots,X_\tau)}\right], \\
&= \mathbb{E}_{\mu_1}\left[\sum_{t=1}^{\tau} X_i \ln\left(\frac{\mu_1}{\mu_0}\right) + (1-X_i)\ln\left(\frac{1-\mu_1}{1-\mu_0}\right)\right], \\
&= \mathbb{E}_{\mu_1}[\tau]\,\mu_1 \ln\left(\frac{\mu_1}{\mu_0}\right) + \mathbb{E}_{\mu_1}[\tau]\,(1-\mu_1)\ln\left(\frac{1-\mu_1}{1-\mu_0}\right), \\
&= \mathbb{E}_{\mu_1}[\tau]\,\mathrm{kl}(\mu_1,\mu_0).
\end{aligned}$$

where $\mathrm{kl}(x,y) \coloneqq x\ln(x/y) + (1-x)\ln((1-x)/(1-y))$ is the Bernoulli KL divergence.

A brief introduction: lower bound (cont.)

At the stopping time we have:

$$
\begin{aligned}
\mathbb{E}_{\mu_1}[Z_\tau(\mu_1,\mu_0)] &= \mathbb{E}_{\mu_1}\left[\ln \frac{\mathrm{d}\mathbb{P}_{\mu_1}(X_1,\dots,X_\tau)}{\mathrm{d}\mathbb{P}_{\mu_0}(X_1,\dots,X_\tau)}\right], \\
&= \mathbb{E}_{\mu_1}\left[\sum_{t=1}^{\tau} X_i \ln\left(\frac{\mu_1}{\mu_0}\right) + (1-X_i)\ln\left(\frac{1-\mu_1}{1-\mu_0}\right)\right], \\
&= \mathbb{E}_{\mu_1}[\tau]\,\mu_1 \ln\left(\frac{\mu_1}{\mu_0}\right) + \mathbb{E}_{\mu_1}[\tau](1-\mu_1)\ln\left(\frac{1-\mu_1}{1-\mu_0}\right), \\
&= \mathbb{E}_{\mu_1}[\tau]\mathrm{kl}(\mu_1,\mu_0).
\end{aligned}
$$

where $\mathrm{kl}(x,y) \coloneqq x\ln(x/y) + (1-x)\ln((1-x)/(1-y))$ is the Bernoulli KL divergence.

A brief introduction: lower bound (cont.)

At the stopping time we have:

$$\begin{aligned}
\mathbb{E}_{\mu_1}[Z_\tau(\mu_1,\mu_0)] &= \mathbb{E}_{\mu_1}\left[\ln \frac{\mathrm{d}\mathbb{P}_{\mu_1}(X_1,\dots,X_\tau)}{\mathrm{d}\mathbb{P}_{\mu_0}(X_1,\dots,X_\tau)}\right], \\
&= \mathbb{E}_{\mu_1}\left[\sum_{t=1}^{\tau} X_i \ln\left(\frac{\mu_1}{\mu_0}\right) + (1-X_i)\ln\left(\frac{1-\mu_1}{1-\mu_0}\right)\right], \\
&= \mathbb{E}_{\mu_1}[\tau]\,\mu_1 \ln\left(\frac{\mu_1}{\mu_0}\right) + \mathbb{E}_{\mu_1}[\tau]\,(1-\mu_1)\ln\left(\frac{1-\mu_1}{1-\mu_0}\right), \\
&= \mathbb{E}_{\mu_1}[\tau]\mathrm{kl}(\mu_1,\mu_0).
\end{aligned}$$

where $\mathrm{kl}(x,y) \coloneqq x\ln(x/y) + (1-x)\ln((1-x)/(1-y))$ is the Bernoulli KL divergence.

A brief introduction: lower bound (cont.)

At the stopping time we have:

$$\begin{aligned}
\mathbb{E}_{\mu_1}[Z_\tau(\mu_1,\mu_0)] &= \mathbb{E}_{\mu_1}\left[\ln \frac{\mathrm{d}\mathbb{P}_{\mu_1}(X_1,\dots,X_\tau)}{\mathrm{d}\mathbb{P}_{\mu_0}(X_1,\dots,X_\tau)}\right], \\
&= \mathbb{E}_{\mu_1}\left[\sum_{t=1}^{\tau} X_i \ln\left(\frac{\mu_1}{\mu_0}\right) + (1-X_i)\ln\left(\frac{1-\mu_1}{1-\mu_0}\right)\right], \\
&= \mathbb{E}_{\mu_1}[\tau]\,\mu_1 \ln\left(\frac{\mu_1}{\mu_0}\right) + \mathbb{E}_{\mu_1}[\tau]\,(1-\mu_1)\ln\left(\frac{1-\mu_1}{1-\mu_0}\right), \\
&= \mathbb{E}_{\mu_1}[\tau]\mathrm{kl}(\mu_1,\mu_0).
\end{aligned}$$

where $\mathrm{kl}(x,y) := x\ln(x/y) + (1-x)\ln((1-x)/(1-y))$ is the Bernoulli KL divergence.

A brief introduction: lower bound (cont.)

We use the following result, which is a form of data processing inequality.

Lemma (Fundamental inequality [GMS19])

For any $\mathcal{F}_\tau$-measurable r.v. $Y \in [0,1]$ we have $\mathbb{E}_{\mu_1}[Z_\tau(\mu_1,\mu_0)] \geq \mathrm{kl}\left(\mathbb{E}_{\mu_1}[Y], \mathbb{E}_{\mu_0}[Y]\right).$

- Assume $\mu_1 < \lambda$ and $\mu_0 \geq \lambda$, and let $\Delta := \lambda - \mu_1$.
- Consider an algorithm with type I error rate $\leq \alpha$ and type II error rate $\leq \beta$.
- For such algorithm, for $Y = \mathbf{1}_{\mathcal{E}}$, with $\mathcal{E} = \{\text{reject } H_0\}$, we have $\mathbb{P}_{\mu_1}(\mathcal{E}) \geq 1-\beta$ and $\mathbb{P}_{\mu_0}(\mathcal{E}) \leq \alpha$. Hence

$$\mathbb{E}_{\mu_1}[Z_\tau(\mu_1,\mu_0)] = \mathbb{E}_{\mu_1}[\tau]\mathrm{kl}(\mu_1,\mu_0) \geq \mathrm{kl}(1-\beta,\alpha).$$

which follows from the monotonicity properties of the KL divergence.

A brief introduction: lower bound (cont.)

We use the following result, which is a form of data processing inequality.

Lemma (Fundamental inequality [GMS19])

For any $\mathcal{F}_\tau$-measurable r.v. $Y \in [0,1]$ we have $\mathbb{E}_{\mu_1}[Z_\tau(\mu_1,\mu_0)] \geq \mathrm{kl}\left(\mathbb{E}_{\mu_1}[Y], \mathbb{E}_{\mu_0}[Y]\right).$

- Assume $\mu_1 < \lambda$ and $\mu_0 \geq \lambda$, and let $\Delta := \lambda - \mu_1$.
- Consider an algorithm with type I error rate $\leq \alpha$ and type II error rate $\leq \beta$.
- For such algorithm, for $Y = \mathbf{1}_{\mathcal{E}}$, with $\mathcal{E} = \{\text{reject } H_0\}$, we have $\mathbb{P}_{\mu_1}(\mathcal{E}) \geq 1-\beta$ and $\mathbb{P}_{\mu_0}(\mathcal{E}) \leq \alpha$. Hence

$$\mathbb{E}_{\mu_1}[Z_\tau(\mu_1,\mu_0)] = \mathbb{E}_{\mu_1}[\tau]\mathrm{kl}(\mu_1,\mu_0) \geq \mathrm{kl}(1-\beta,\alpha).$$

which follows from the monotonicity properties of the KL divergence.

A brief introduction: lower bound (cont.)

We use the following result, which is a form of data processing inequality.

Lemma (Fundamental inequality [GMS19])

For any $\mathcal{F}_\tau$-measurable r.v. $Y \in [0,1]$ we have $\mathbb{E}_{\mu_1}[Z_\tau(\mu_1,\mu_0)] \geq \mathrm{kl}\left(\mathbb{E}_{\mu_1}[Y], \mathbb{E}_{\mu_0}[Y]\right).$

- Assume $\mu_1 < \lambda$ and $\mu_0 \geq \lambda$, and let $\Delta := \lambda - \mu_1$.
- Consider an algorithm with type I error rate $\leq \alpha$ and type II error rate $\leq \beta$.
- For such algorithm, for $Y = \mathbf{1}_{\mathcal{E}}$, with $\mathcal{E} = \{\text{reject } H_0\}$, we have $\mathbb{P}_{\mu_1}(\mathcal{E}) \geq 1-\beta$ and $\mathbb{P}_{\mu_0}(\mathcal{E}) \leq \alpha$. Hence

$$\mathbb{E}_{\mu_1}[Z_\tau(\mu_1,\mu_0)] = \mathbb{E}_{\mu_1}[\tau]\mathrm{kl}(\mu_1,\mu_0) \geq \mathrm{kl}(1-\beta,\alpha).$$

which follows from the monotonicity properties of the KL divergence.

A brief introduction: lower bound (cont.)

We use the following result, which is a form of data processing inequality.

Lemma (Fundamental inequality [GMS19])

For any $\mathcal{F}_\tau$-measurable r.v. $Y \in [0,1]$ we have $\mathbb{E}_{\mu_1}[Z_\tau(\mu_1, \mu_0)] \geq \mathrm{kl}\left(\mathbb{E}_{\mu_1}[Y], \mathbb{E}_{\mu_0}[Y]\right).$

- Assume $\mu_1 < \lambda$ and $\mu_0 \geq \lambda$, and let $\Delta := \lambda - \mu_1$.
- Consider an algorithm with type I error rate $\leq \alpha$ and type II error rate $\leq \beta$.
- For such algorithm, for $Y = \mathbf{1}_{\mathcal{E}}$, with $\mathcal{E} = \{\text{reject } H_0\}$, we have $\mathbb{P}_{\mu_1}(\mathcal{E}) \geq 1 - \beta$ and $\mathbb{P}_{\mu_0}(\mathcal{E}) \leq \alpha$. Hence

$$\mathbb{E}_{\mu_1}[Z_\tau(\mu_1, \mu_0)] = \mathbb{E}_{\mu_1}[\tau]\mathrm{kl}(\mu_1, \mu_0) \geq \mathrm{kl}(1 - \beta, \alpha).$$

  which follows from the monotonicity properties of the KL divergence.

A brief introduction: lower bound (final)

Since $\lambda \leq \mu_0$, we also have that

$$\mathbb{E}_{\mu_1}[\tau]\mathrm{kl}(\mu_1, \lambda) \geq \mathbb{E}_{\mu_1}[\tau]\mathrm{kl}(\mu_1, \mu_0) \geq \mathrm{kl}(1-\beta, \alpha).$$

Hence, under H_1, since $\mu_1 = \lambda - \Delta$, we have that[2][3]

$$\mathbb{E}_{H_1}[\tau] \geq \underbrace{\frac{1}{\mathrm{kl}(\lambda - \Delta, \lambda)}}_{\text{information rate}^4} \cdot \mathrm{kl}(1-\beta, \alpha) \approx \frac{1}{\Delta^2} \cdot (1-\beta)\ln(1/\alpha).$$

---

[2] It's also interesting to note that the bound scales differently for different values of (α, β).

[3] A model with average $\lambda - \Delta$ is the one that is statistically close to the null hypothesis with accuracy $\Delta \Rightarrow$ confusing model.

[4] $\mathrm{kl}(\lambda - \Delta, \lambda)$ is a.k.a. evidence rate. The inverse $1/\mathrm{kl}(\lambda - \Delta, \lambda)$ is the *characteristic time*.

A brief introduction: lower bound (final)

Since $\lambda \leq \mu_0$, we also have that

$$\mathbb{E}_{\mu_1}[\tau]\mathrm{kl}(\mu_1, \lambda) \geq \mathbb{E}_{\mu_1}[\tau]\mathrm{kl}(\mu_1, \mu_0) \geq \mathrm{kl}(1-\beta, \alpha).$$

Hence, under H_1, since $\mu_1 = \lambda - \Delta$, we have that[2][3]

$$\mathbb{E}_{H_1}[\tau] \geq \underbrace{\frac{1}{\mathrm{kl}(\lambda - \Delta, \lambda)}}_{\text{information rate}^4} \cdot \mathrm{kl}(1-\beta, \alpha) \approx \frac{1}{\Delta^2} \cdot (1-\beta)\ln(1/\alpha).$$

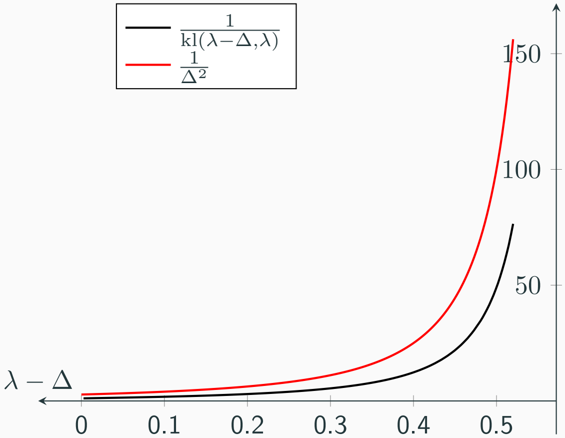
[2] It's also interesting to note that the bound scales differently for different values of (α, β).

[3] A model with average $\lambda - \Delta$ is the one that is statistically close to the null hypothesis with accuracy $\Delta \Rightarrow$ confusing model.

[4] $\mathrm{kl}(\lambda - \Delta, \lambda)$ is a.k.a. evidence rate. The inverse $1/\mathrm{kl}(\lambda - \Delta, \lambda)$ is the *characteristic time*.

Example with $\lambda = 0.6$

Inverse of the information rate with $\lambda = 0.6$ (characteristic time)



Detecting a change in a stream of data as quickly as possible

Quickest Change Detection

We now look at a different problem, called Quickest Change Detection.

- Suppose you observe a stream of random variables $X_1, X_2, X_3 \ldots$.
  - $X_1, \ldots, X_{\nu-1}$ are i.i.d., distributed according to F_0, while $X_\nu, X_{\nu+1}, \ldots$ are i.i.d, distributed according to F_1.
- ν is an unknown change-time.
  - For $\nu = 1, 2, \ldots$ we let $\mathbb{P}_\nu$ denote the probability measure of the sequence $X_1, X_2, \ldots$ where X_ν is the first term distributed as F_1.
- We measure the performance of a detection algorithm using the worst average detection delay (WADD). Let τ be the stopping time of the algorithm, then

$$\bar{E}(\tau) = \sup_{\nu \geq 1} \operatorname{ess\,sup} \mathbb{E}_\nu \left[(\tau - \nu)^+ | X_1, \ldots, X_{\nu-1} \right].$$

- What is the minimum number of samples needed to detect a change?

Quickest Change Detection

We now look at a different problem, called Quickest Change Detection.

- Suppose you observe a stream of random variables $X_1, X_2, X_3 \ldots$.
  - $X_1, \ldots, X_{\nu-1}$ are i.i.d., distributed according to F_0, while $X_\nu, X_{\nu+1}, \ldots$ are i.i.d, distributed according to F_1.
- ν is an unknown change-time.
  - For $\nu = 1, 2, \ldots$ we let $\mathbb{P}_\nu$ denote the probability measure of the sequence $X_1, X_2, \ldots$ where X_ν is the first term distributed as F_1.
- We measure the performance of a detection algorithm using the worst average detection delay (WADD). Let τ be the stopping time of the algorithm, then

$$\bar{E}(\tau) = \sup_{\nu \geq 1} \operatorname{ess\,sup} \mathbb{E}_\nu \left[(\tau - \nu)^+ | X_1, \ldots, X_{\nu-1} \right].$$

- What is the minimum number of samples needed to detect a change?

Quickest Change Detection

We now look at a different problem, called Quickest Change Detection.

- Suppose you observe a stream of random variables $X_1, X_2, X_3 \ldots$.
  - $X_1, \ldots, X_{\nu-1}$ are i.i.d., distributed according to F_0, while $X_\nu, X_{\nu+1}, \ldots$ are i.i.d, distributed according to F_1.
- ν is an unknown change-time.
  - For $\nu = 1, 2, \ldots$ we let $\mathbb{P}_\nu$ denote the probability measure of the sequence $X_1, X_2, \ldots$ where X_ν is the first term distributed as F_1.
- We measure the performance of a detection algorithm using the worst average detection delay (WADD). Let τ be the stopping time of the algorithm, then

$$\bar{E}(\tau) = \sup_{\nu \geq 1} \operatorname{ess\,sup} \mathbb{E}_\nu \left[(\tau - \nu)^+ | X_1, \ldots, X_{\nu-1} \right].$$

- What is the minimum number of samples needed to detect a change?

Quickest Change Detection (cont.)

Hypothesis Testing Problem

$$H_0 : \text{ accept } F_0 \quad \text{vs} \quad H_1 : \text{ accept } F_1$$

- Denote by $\mathbb{P}_\infty$ the measure when there is no change.
- Ideally, we want an algorithm with a certain false alarm rate (type I error), i.e.,

$$E_\infty[\tau] \geq \frac{1}{\alpha}, \text{ (Average time to false alarm)},$$

  for $\alpha > 0$.
- However, it is not possible to directly apply the same arguments as before. Moreover, it is not clear how to choose the event $\mathcal{E}$ (we can't really control the type II error here).

Quickest Change Detection (cont.)

Hypothesis Testing Problem

$$H_0: \text{ accept } F_0 \quad \text{vs} \quad H_1: \text{ accept } F_1$$

- Denote by $\mathbb{P}_\infty$ the measure when there is no change.
- Ideally, we want an algorithm with a certain false alarm rate (type I error), i.e.,

$$E_\infty[\tau] \geq \frac{1}{\alpha}, \text{ (Average time to false alarm)},$$

  for $\alpha > 0$.
- However, it is not possible to directly apply the same arguments as before. Moreover, it is not clear how to choose the event $\mathcal{E}$ (we can't really control the type II error here).

Quickest Change Detection (cont.)

Hypothesis Testing Problem

$$H_0: \text{ accept } F_0 \quad \text{vs} \quad H_1: \text{ accept } F_1$$

- Denote by $\mathbb{P}_\infty$ the measure when there is no change.
- Ideally, we want an algorithm with a certain false alarm rate (type I error), i.e.,

$$E_\infty[\tau] \geq \frac{1}{\alpha}, \text{ (Average time to false alarm)},$$

  for $\alpha > 0$.
- However, it is not possible to directly apply the same arguments as before. Moreover, it is not clear how to choose the event $\mathcal{E}$ (we can't really control the type II error here).

Quickest Change Detection: lower bound

Let

$$(T^\star)^{-1} := \lim_{n\to\infty} \frac{1}{n} \sum_{t=\nu}^{\nu+n} \ln \frac{F_1(X_t)}{F_0(X_t)} = \mathrm{KL}(F_1, F_0),$$

be the information rate. It represents the average amount of information (evidence) that you gain at each time-step to discern between the two hypotheses. The idea is to show the following for any $\delta \in (0,1)$:

$$\lim_{\alpha\to 0} \mathbb{P}_\nu \left(\tau - \nu \le (T^\star)^{-1}(1-\delta)\ln(1/\alpha) \mid \tau \ge \nu\right) = 0,$$

which also implies that[5]

$$\lim_{\alpha\to 0} \frac{\mathbb{E}_\nu[\tau - \nu | \tau \ge \nu]}{\ln(1/\alpha)} \ge T^\star.$$

[5] This would conclude the proof since $\bar{E}(\tau) \ge \mathbb{E}_\nu[\tau - \nu | \tau \ge \nu]$.

Quickest Change Detection: lower bound

Let

$$(T^\star)^{-1} := \lim_{n\to\infty} \frac{1}{n} \sum_{t=\nu}^{\nu+n} \ln \frac{F_1(X_t)}{F_0(X_t)} = \mathrm{KL}(F_1, F_0),$$

be the information rate. It represents the average amount of information (evidence) that you gain at each time-step to discern between the two hypotheses. The idea is to show the following for any $\delta \in (0,1)$:

$$\lim_{\alpha\to 0} \mathbb{P}_\nu \left(\tau - \nu \le (T^\star)^{-1}(1-\delta) \ln(1/\alpha) \mid \tau \ge \nu \right) = 0,$$

which also implies that[5]

$$\lim_{\alpha\to 0} \frac{\mathbb{E}_\nu[\tau - \nu | \tau \ge \nu]}{\ln(1/\alpha)} \ge T^\star.$$

[5] This would conclude the proof since $\bar{E}(\tau) \ge \mathbb{E}_\nu[\tau - \nu | \tau \ge \nu]$.

Quickest Change Detection: lower bound (cont.)

Lemma (Lower-level form of the fundamental inequality [KCG16, GK21])

For all $x \in \mathbb{R}, t \in \mathbb{N}$ and all event $\mathcal{E} \in \mathcal{F}_t$ we have

$$\textit{(Change of measure trick)} \quad \mathbb{P}_\nu(\mathcal{E}) \leq \mathbb{P}_\nu(Z_t \geq x) + e^x \mathbb{P}_\infty(\mathcal{E}),$$

where $Z_t = \ln \frac{\mathrm{d}\mathbb{P}_\nu(X_1,\ldots,X_t)}{\mathrm{d}\mathbb{P}_\infty(X_1,\ldots,X_t)}$ is the log-likelihood ratio.

Let $t = n_\alpha$, where $n_\alpha = T^\star(1-\delta)\ln(1/\alpha)$ with $\delta \in (0,1)$, and $\mathcal{E} = \{\tau - \nu \leq n_\alpha\}$. Then $\mathcal{E} \in \mathcal{F}_{n_\alpha}$. For sufficiently small α one can prove $\mathbb{P}_\infty(\mathcal{E} | \tau \geq \nu) \leq [\ln(1/\alpha)]^2 \alpha$ [6].

$$\Rightarrow \mathbb{P}_\nu(\mathcal{E} \mid \tau \geq \nu) \leq [\ln(1/\alpha)]^2 \alpha e^x + \mathbb{P}_\nu(Z_{n_\alpha} \geq x \mid \tau \geq \nu).$$

Choose $x = (1-\delta/2)\ln(1/\alpha)$. Then

$$\mathbb{P}_\nu(\mathcal{E} \mid \tau \geq \nu) \leq [\ln(1/\alpha)]^2 \alpha^{\delta/2} + \mathbb{P}_\nu\left(Z_{n_\alpha} \geq (1-\delta)\ln(1/\alpha)\frac{n_\alpha}{n_\alpha} \middle| \tau \geq \nu\right).$$

[6] See appendix for a proof.

Quickest Change Detection: lower bound (cont.)

Lemma (Lower-level form of the fundamental inequality [KCG16, GK21])

For all $x \in \mathbb{R}, t \in \mathbb{N}$ and all event $\mathcal{E} \in \mathcal{F}_t$ we have

$$\text{(Change of measure trick)} \quad \mathbb{P}_\nu(\mathcal{E}) \leq \mathbb{P}_\nu(Z_t \geq x) + e^x \mathbb{P}_\infty(\mathcal{E}),$$

where $Z_t = \ln \frac{\mathrm{d}\mathbb{P}_\nu(X_1,\ldots,X_t)}{\mathrm{d}\mathbb{P}_\infty(X_1,\ldots,X_t)}$ is the log-likelihood ratio.

Let $t = n_\alpha$, where $n_\alpha = T^\star(1-\delta)\ln(1/\alpha)$ with $\delta \in (0,1)$, and $\mathcal{E} = \{\tau - \nu \leq n_\alpha\}$. Then $\mathcal{E} \in \mathcal{F}_{n_\alpha}$. For sufficiently small α one can prove $\mathbb{P}_\infty(\mathcal{E} | \tau \geq \nu) \leq [\ln(1/\alpha)]^2 \alpha$ [6].

$$\Rightarrow \mathbb{P}_\nu(\mathcal{E} \mid \tau \geq \nu) \leq [\ln(1/\alpha)]^2 \alpha e^x + \mathbb{P}_\nu(Z_{n_\alpha} \geq x \mid \tau \geq \nu).$$

Choose $x = (1-\delta/2)\ln(1/\alpha)$. Then

$$\mathbb{P}_\nu(\mathcal{E} \mid \tau \geq \nu) \leq [\ln(1/\alpha)]^2 \alpha^{\delta/2} + \mathbb{P}_\nu\left(Z_{n_\alpha} \geq (1-\delta)\ln(1/\alpha)\frac{n_\alpha}{n_\alpha} \middle| \tau \geq \nu\right).$$

[6] See appendix for a proof.

Quickest Change Detection: lower bound (cont.)

Lemma (Lower-level form of the fundamental inequality [KCG16, GK21])

For all $x \in \mathbb{R}, t \in \mathbb{N}$ and all event $\mathcal{E} \in \mathcal{F}_t$ we have

$$\textit{(Change of measure trick)} \quad \mathbb{P}_\nu(\mathcal{E}) \leq \mathbb{P}_\nu(Z_t \geq x) + e^x \mathbb{P}_\infty(\mathcal{E}),$$

where $Z_t = \ln \frac{\mathrm{d}\mathbb{P}_\nu(X_1,\ldots,X_t)}{\mathrm{d}\mathbb{P}_\infty(X_1,\ldots,X_t)}$ is the log-likelihood ratio.

Let $t = n_\alpha$, where $n_\alpha = T^\star(1-\delta)\ln(1/\alpha)$ with $\delta \in (0,1)$, and $\mathcal{E} = \{\tau - \nu \leq n_\alpha\}$. Then $\mathcal{E} \in \mathcal{F}_{n_\alpha}$. For sufficiently small α one can prove $\mathbb{P}_\infty(\mathcal{E}|\tau \geq \nu) \leq [\ln(1/\alpha)]^2\alpha$ [6].

$$\Rightarrow \mathbb{P}_\nu(\mathcal{E} \mid \tau \geq \nu) \leq [\ln(1/\alpha)]^2 \alpha e^x + \mathbb{P}_\nu(Z_{n_\alpha} \geq x \mid \tau \geq \nu).$$

Choose $x = (1-\delta/2)\ln(1/\alpha)$. Then

$$\mathbb{P}_\nu(\mathcal{E} \mid \tau \geq \nu) \leq [\ln(1/\alpha)]^2 \alpha^{\delta/2} + \mathbb{P}_\nu\left(Z_{n_\alpha} \geq (1-\delta)\ln(1/\alpha)\frac{n_\alpha}{n_\alpha}\middle|\tau \geq \nu\right).$$

[6] See appendix for a proof.

Quickest Change Detection: lower bound (final)

$$\mathbb{P}_\nu(\mathcal{E} \mid \tau \geq \nu) \leq [\ln(1/\alpha)]^2 \alpha^{\delta/2} + \mathbb{P}_\nu\left(Z_{n_\alpha} \geq (1-\delta)\ln(1/\alpha)\frac{n_\alpha}{n_\alpha}\middle|\tau \geq \nu\right).$$

By the l.l.n. we have $\frac{Z_{n_\alpha}}{n_\alpha} \to (T^\star)^{-1}$ as $\alpha \to 0$. Then, as $\alpha \to 0$ we have that the inequality

$$(T^\star)^{-1} \geq \frac{(1-\delta/2)\ln(1/\alpha)}{n_\alpha} = \frac{1-\delta/2}{T^\star(1-\delta)} \Rightarrow 1 \geq \frac{1-\delta/2}{1-\delta},$$

is not true. Hence $\lim_{\alpha\to 0}\mathbb{P}_\nu(\mathcal{E} \mid \tau \geq \nu) = 0$, which concludes the proof. Therefore

$$\inf_{\tau:\mathbb{E}_\infty(\tau<\infty)\geq 1/\alpha} \frac{\bar{E}(\tau)}{\ln(1/\alpha)} \geq T^\star \text{ as } \alpha \to 0.$$

Quickest Change Detection: lower bound (final)

$$\mathbb{P}_\nu(\mathcal{E} \mid \tau \geq \nu) \leq [\ln(1/\alpha)]^2 \alpha^{\delta/2} + \mathbb{P}_\nu\left(Z_{n_\alpha} \geq (1-\delta)\ln(1/\alpha)\frac{n_\alpha}{n_\alpha}\middle|\tau \geq \nu\right).$$

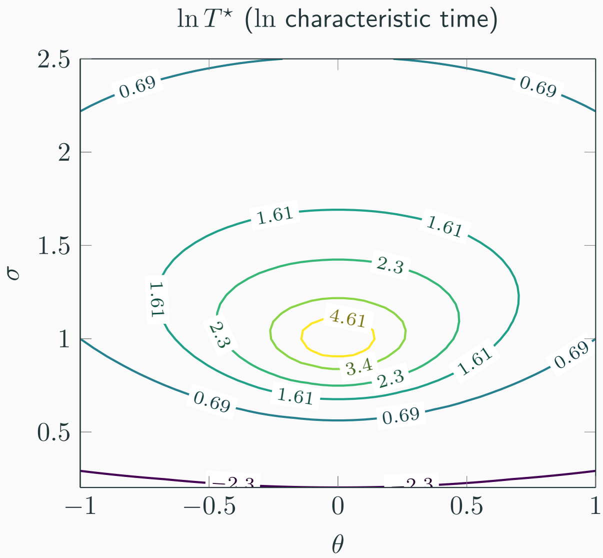
By the l.l.n. we have $\frac{Z_{n_\alpha}}{n_\alpha} \to (T^\star)^{-1}$ as $\alpha \to 0$. Then, as $\alpha \to 0$ we have that the inequality

$$(T^\star)^{-1} \geq \frac{(1-\delta/2)\ln(1/\alpha)}{n_\alpha} = \frac{1-\delta/2}{T^\star(1-\delta)} \Rightarrow 1 \geq \frac{1-\delta/2}{1-\delta},$$

is not true. Hence $\lim_{\alpha\to 0} \mathbb{P}_\nu(\mathcal{E} \mid \tau \geq \nu) = 0$, which concludes the proof. Therefore

$$\inf_{\tau:\mathbb{E}_\infty(\tau<\infty)\geq 1/\alpha} \frac{\bar{E}(\tau)}{\ln(1/\alpha)} \geq T^\star \text{ as } \alpha \to 0.$$

Example with $F_0 = \mathcal{N}(0,1)$ and $F_1 = \mathcal{N}(\theta, \sigma^2)$



Best Arm Identification with Fixed Confidence: MAB with finite number of arms

Best Arm Identification with Fixed Confidence: introduction

- Consider a Multi-Armed Bandit model (MAB) μ with K arms: at each time-step t an agent chooses an arm A_t and observes a random reward $R_t \sim F_{A_t}$ (the rewards are i.i.d.).
- Assume there is only one unique best arm $a^\star(\mu) = \arg\max_k \mu_k$.

Estimate which arm is optimal as quickly as possible with confidence $\delta \in (0, 1/2)$.

- As usual, define τ to be the stopping time of the algorithm.
- Let $\hat{a}_\tau$ be the optimal arm estimated by the algorithm at the stopping time.
- We say that an algorithm is δ-PC (Probably Correct) if $\mathbb{P}_\mu(\tau < \infty, \hat{a}_\tau = a^\star(\mu)) \geq 1 - \delta$ for all possible models μ satisfying the uniqueness of the best arm.

Best Arm Identification with Fixed Confidence: introduction

- Consider a Multi-Armed Bandit model (MAB) μ with K arms: at each time-step t an agent chooses an arm A_t and observes a random reward $R_t \sim F_{A_t}$ (the rewards are i.i.d.).
- Assume there is only one unique best arm $a^\star(\mu) = \arg\max_k \mu_k$.

Estimate which arm is optimal as quickly as possible with confidence $\delta \in (0, 1/2)$.

- As usual, define τ to be the stopping time of the algorithm.
- Let $\hat{a}_\tau$ be the optimal arm estimated by the algorithm at the stopping time.
- We say that an algorithm is δ-PC (Probably Correct) if $\mathbb{P}_\mu(\tau < \infty, \hat{a}_\tau = a^\star(\mu)) \geq 1 - \delta$ for all possible models μ satisfying the uniqueness of the best arm.

Best Arm Identification with Fixed Confidence: introduction

- Consider a Multi-Armed Bandit model (MAB) μ with K arms: at each time-step t an agent chooses an arm A_t and observes a random reward $R_t \sim F_{A_t}$ (the rewards are i.i.d.).
- Assume there is only one unique best arm $a^\star(\mu) = \arg\max_k \mu_k$.

Estimate which arm is optimal as quickly as possible with confidence $\delta \in (0, 1/2)$.

- As usual, define τ to be the stopping time of the algorithm.
- Let $\hat{a}_\tau$ be the optimal arm estimated by the algorithm at the stopping time.
- We say that an algorithm is δ-PC (Probably Correct) if $\mathbb{P}_\mu(\tau < \infty, \hat{a}_\tau = a^\star(\mu)) \geq 1 - \delta$ for all possible models μ satisfying the uniqueness of the best arm.

Best Arm Identification with Fixed Confidence: introduction

- Consider a Multi-Armed Bandit model (MAB) μ with K arms: at each time-step t an agent chooses an arm A_t and observes a random reward $R_t \sim F_{A_t}$ (the rewards are i.i.d.).
- Assume there is only one unique best arm $a^\star(\mu) = \arg\max_k \mu_k$.

Estimate which arm is optimal as quickly as possible with confidence $\delta \in (0, 1/2)$.

- As usual, define τ to be the stopping time of the algorithm.
- Let $\hat{a}_\tau$ be the optimal arm estimated by the algorithm at the stopping time.
- We say that an algorithm is δ-PC (Probably Correct) if $\mathbb{P}_\mu(\tau < \infty, \hat{a}_\tau = a^\star(\mu)) \geq 1 - \delta$ for all possible models μ satisfying the uniqueness of the best arm.

Best Arm Identification with Fixed Confidence: introduction

- Consider a Multi-Armed Bandit model (MAB) μ with K arms: at each time-step t an agent chooses an arm A_t and observes a random reward $R_t \sim F_{A_t}$ (the rewards are i.i.d.).
- Assume there is only one unique best arm $a^\star(\mu) = \arg\max_k \mu_k$.

Estimate which arm is optimal as quickly as possible with confidence $\delta \in (0, 1/2)$.

- As usual, define τ to be the stopping time of the algorithm.
- Let $\hat{a}_\tau$ be the optimal arm estimated by the algorithm at the stopping time.
- We say that an algorithm is δ-PC (Probably Correct) if $\mathbb{P}_\mu(\tau < \infty, \hat{a}_\tau = a^\star(\mu)) \geq 1 - \delta$ for all possible models μ satisfying the uniqueness of the best arm.

Best Arm Identification with Fixed Confidence: lower bound

Hypotheses

Here we have K hypotheses:

$$H_1: \text{ optimal arm is } a_1, \quad H_2: \text{ optimal arm is } a_2, \quad H_3: \dots$$

- Similar to the initial example but with multiple hypotheses.
- Recall that in the initial example we compared the true model with the worst confusing model (λ vs $\lambda - \Delta$).
- We need to do something similar here. Define the set of confusing models

$$\mathrm{Alt}(\mu) := \{\mu' : \arg\max_a \mu'_a \neq a^\star(\mu), |\arg\max_a \mu'_a| = 1\}$$

This is the set of models where the optimal (unique) arm is not a_1.

Best Arm Identification with Fixed Confidence: lower bound

Hypotheses

Here we have K hypotheses:

$$H_1: \text{ optimal arm is } a_1, \quad H_2: \text{ optimal arm is } a_2, \quad H_3: \ldots$$

- Similar to the initial example but with multiple hypotheses.
- Recall that in the initial example we compared the true model with the worst confusing model (λ vs $\lambda - \Delta$).
- We need to do something similar here. Define the set of confusing models

$$\mathrm{Alt}(\mu) := \{\mu' : \arg\max_a \mu'_a \neq a^\star(\mu), |\arg\max_a \mu'_a| = 1\}$$

This is the set of models where the optimal (unique) arm is not a_1.

Best Arm Identification with Fixed Confidence: lower bound

Hypotheses

Here we have K hypotheses:

$$H_1: \text{ optimal arm is } a_1, \quad H_2: \text{ optimal arm is } a_2, \quad H_3: \dots$$

- Similar to the initial example but with multiple hypotheses.
- Recall that in the initial example we compared the true model with the worst confusing model (λ vs $\lambda - \Delta$).
- We need to do something similar here. Define the set of confusing models

$$\mathrm{Alt}(\mu) := \{\mu' : \arg\max_a \mu'_a \neq a^\star(\mu), |\arg\max_a \mu'_a| = 1\}$$

This is the set of models where the optimal (unique) arm is not a_1.

Best Arm Identification with Fixed Confidence: lower bound (cont.)

- The event we are interested in is $\{\hat{a}_\tau \neq a^\star(\mu)\}$, that motivates why we consider this particular set of confusing models (the optimal arm in those models is not $a^\star(\mu)$).
- Consider then the log-likelihood ratio $Z_t = \ln \frac{d\mathbb{P}_\mu(A_1,R_1,\dots,A_t,R_t)}{d\mathbb{P}_{\mu'}(A_1,R_1,\dots,A_t,R_t)}$ between μ and $\mu' \in \mathrm{Alt}(\mu)$. Then:

$$\mathbb{E}_\mu[Z_t] = \mathbb{E}_\mu\left[\sum_{n=1}^{t}\sum_{a\in[K]} \mathbf{1}_{\{A_n=a\}} \ln \frac{f_{\mu_a}(Y_n)}{f_{\mu'_a}(Y_n)}\right] = \sum_a \mathbb{E}_\mu\left[N_a(t)\right] \mathrm{KL}(f_{\mu_a}, f_{\mu'_a}),$$

where f_{μ_a} is the reward PDF for action a.

Best Arm Identification with Fixed Confidence: lower bound (cont.)

- The event we are interested in is $\{\hat{a}_\tau \neq a^\star(\mu)\}$, that motivates why we consider this particular set of confusing models (the optimal arm in those models is not $a^\star(\mu)$).
- Consider then the log-likelihood ratio $Z_t = \ln \frac{\mathrm{d}\mathbb{P}_\mu(A_1,R_1,\dots,A_t,R_t)}{\mathrm{d}\mathbb{P}_{\mu'}(A_1,R_1,\dots,A_t,R_t)}$ between μ and $\mu' \in \mathrm{Alt}(\mu)$. Then:

$$\mathbb{E}_\mu[Z_t] = \mathbb{E}_\mu\left[\sum_{n=1}^{t}\sum_{a\in[K]} \mathbf{1}_{\{A_n=a\}} \ln \frac{f_{\mu_a}(Y_n)}{f_{\mu'_a}(Y_n)}\right] = \sum_a \mathbb{E}_\mu\left[N_a(t)\right] \mathrm{KL}(f_{\mu_a}, f_{\mu'_a}),$$

where f_{μ_a} is the reward PDF for action a.

Best Arm Identification with Fixed Confidence: lower bound (cont.)

Let $t = \tau$

$$\mathbb{E}_\mu[Z_\tau] = \sum_a \mathbb{E}_\mu\left[N_a(\tau)\right] \mathrm{KL}(f_{\mu_a}, f_{\mu'_a}).$$

Recall the fundamental inequality (high-level form)

Lemma (Fundamental inequality [GMS19])

For any $\mathcal{F}_\tau$-measurable r.v. $Y \in [0,1]$ we have $\mathbb{E}_{\mu_1}[Z_\tau(\mu_1, \mu_0)] \geq \mathrm{kl}\left(\mathbb{E}_{\mu_1}[Y], \mathbb{E}_{\mu_0}[Y]\right).$

We apply it and choose $Y = \mathbf{1}_\mathcal{E}, \mathcal{E} = \{\hat{a}_\tau = a^\star(\mu)\}$:

$$\sum_a \mathbb{E}_\mu\left[N_a(\tau)\right] \mathrm{KL}(f_{\mu_a}, f_{\mu'_a}) \geq \mathrm{kl}(\mathbb{P}_\mu(\mathcal{E}), \mathbb{P}_{\mu'}(\mathcal{E})) \geq \mathrm{kl}(1-\delta, \delta).$$

since $\mathbb{P}_\mu(\mathcal{E}) \geq 1-\delta$ and $\mathbb{P}_{\mu'}(\mathcal{E}) \leq \delta$ from the fact that $\mathcal{E} \subset \{\hat{a}_\tau \neq a^\star(\mu')\}$ under $\mathbb{P}_{\mu'}$

Best Arm Identification with Fixed Confidence: lower bound (cont.)

Let $t = \tau$

$$\mathbb{E}_\mu[Z_\tau] = \sum_a \mathbb{E}_\mu\left[N_a(\tau)\right] \mathrm{KL}(f_{\mu_a}, f_{\mu'_a}).$$

Recall the fundamental inequality (high-level form)

Lemma (Fundamental inequality [GMS19])

For any $\mathcal{F}_\tau$-measurable r.v. $Y \in [0,1]$ we have $\mathbb{E}_{\mu_1}[Z_\tau(\mu_1,\mu_0)] \geq \mathrm{kl}\left(\mathbb{E}_{\mu_1}[Y], \mathbb{E}_{\mu_0}[Y]\right).$

We apply it and choose $Y = \mathbf{1}_\mathcal{E}, \mathcal{E} = \{\hat{a}_\tau = a^\star(\mu)\}$:

$$\sum_a \mathbb{E}_\mu\left[N_a(\tau)\right] \mathrm{KL}(f_{\mu_a}, f_{\mu'_a}) \geq \mathrm{kl}(\mathbb{P}_\mu(\mathcal{E}), \mathbb{P}_{\mu'}(\mathcal{E})) \geq \mathrm{kl}(1-\delta, \delta).$$

since $\mathbb{P}_\mu(\mathcal{E}) \geq 1-\delta$ and $\mathbb{P}_{\mu'}(\mathcal{E}) \leq \delta$ from the fact that $\mathcal{E} \subset \{\hat{a}_\tau \neq a^\star(\mu')\}$ under $\mathbb{P}_{\mu'}$

Best Arm Identification with Fixed Confidence: lower bound (cont.)

Let $t = \tau$

$$\mathbb{E}_\mu[Z_\tau] = \sum_a \mathbb{E}_\mu\left[N_a(\tau)\right] \mathrm{KL}(f_{\mu_a}, f_{\mu'_a}).$$

Recall the fundamental inequality (high-level form)

Lemma (Fundamental inequality [GMS19])

For any $\mathcal{F}_\tau$-measurable r.v. $Y \in [0,1]$ we have $\mathbb{E}_{\mu_1}[Z_\tau(\mu_1, \mu_0)] \geq \mathrm{kl}\left(\mathbb{E}_{\mu_1}[Y], \mathbb{E}_{\mu_0}[Y]\right).$

We apply it and choose $Y = \mathbf{1}_{\mathcal{E}}, \mathcal{E} = \{\hat{a}_\tau = a^\star(\mu)\}$:

$$\sum_a \mathbb{E}_\mu\left[N_a(\tau)\right] \mathrm{KL}(f_{\mu_a}, f_{\mu'_a}) \geq \mathrm{kl}(\mathbb{P}_\mu(\mathcal{E}), \mathbb{P}_{\mu'}(\mathcal{E})) \geq \mathrm{kl}(1-\delta, \delta).$$

since $\mathbb{P}_\mu(\mathcal{E}) \geq 1-\delta$ and $\mathbb{P}_{\mu'}(\mathcal{E}) \leq \delta$ from the fact that $\mathcal{E} \subset \{\hat{a}_\tau \neq a^\star(\mu')\}$ under $\mathbb{P}_{\mu'}$

Best Arm Identification with Fixed Confidence: lower bound (cont.)

We can take the infimum over the set of confusing models:

$$\inf_{\mu' \in \mathrm{Alt}(\mu)} \sum_a \mathbb{E}_\mu \left[N_a(\tau)\right] \mathrm{KL}(f_{\mu_a}, f_{\mu'_a}) \geq \mathrm{kl}(1-\delta, \delta),$$

which yields the most confusing model.

Divide and multiply the left hand-side by $\mathbb{E}_\mu[\tau]$ and let $w_a := \mathbb{E}_\mu[N_a(\tau)]/\mathbb{E}_\mu[\tau]$:

$$\mathbb{E}_\mu[\tau] \inf_{\mu' \in \mathrm{Alt}(\mu)} \sum_a w_a \mathrm{KL}(f_{\mu_a}, f_{\mu'_a}) \geq \mathrm{kl}(1-\delta, \delta).$$

Therefore, we conclude by optimizing over $w_a \in \Delta(K)$ (the simplex over $1, \dots, K$):

$$\mathbb{E}_\mu[\tau] \underbrace{\sup_{w \in \Delta(K)} \inf_{\mu' \in \mathrm{Alt}(\mu)} \sum_a w_a \mathrm{KL}(f_{\mu_a}, f_{\mu'_a})}_{=:(T^\star)^{-1}} \geq \mathrm{kl}(1-\delta, \delta).$$

Best Arm Identification with Fixed Confidence: lower bound (cont.)

We can take the infimum over the set of confusing models:

$$\inf_{\mu' \in \mathrm{Alt}(\mu)} \sum_a \mathbb{E}_\mu \left[N_a(\tau)\right] \mathrm{KL}(f_{\mu_a}, f_{\mu'_a}) \geq \mathrm{kl}(1-\delta, \delta),$$

which yields the most confusing model.

Divide and multiply the left hand-side by $\mathbb{E}_\mu[\tau]$ and let $w_a := \mathbb{E}_\mu[N_a(\tau)]/\mathbb{E}_\mu[\tau]$:

$$\mathbb{E}_\mu[\tau] \inf_{\mu' \in \mathrm{Alt}(\mu)} \sum_a w_a \mathrm{KL}(f_{\mu_a}, f_{\mu'_a}) \geq \mathrm{kl}(1-\delta, \delta).$$

Therefore, we conclude by optimizing over $w_a \in \Delta(K)$ (the simplex over $1, \ldots, K$):

$$\mathbb{E}_\mu[\tau] \underbrace{\sup_{w \in \Delta(K)} \inf_{\mu' \in \mathrm{Alt}(\mu)} \sum_a w_a \mathrm{KL}(f_{\mu_a}, f_{\mu'_a})}_{=:(T^\star)^{-1}} \geq \mathrm{kl}(1-\delta, \delta).$$

Best Arm Identification with Fixed Confidence: lower bound (cont.)

We can take the infimum over the set of confusing models:

$$\inf_{\mu' \in \mathrm{Alt}(\mu)} \sum_a \mathbb{E}_\mu \left[N_a(\tau)\right] \mathrm{KL}(f_{\mu_a}, f_{\mu'_a}) \geq \mathrm{kl}(1-\delta, \delta),$$

which yields the most confusing model.

Divide and multiply the left hand-side by $\mathbb{E}_\mu[\tau]$ and let $w_a := \mathbb{E}_\mu[N_a(\tau)]/\mathbb{E}_\mu[\tau]$:

$$\mathbb{E}_\mu[\tau] \inf_{\mu' \in \mathrm{Alt}(\mu)} \sum_a w_a \mathrm{KL}(f_{\mu_a}, f_{\mu'_a}) \geq \mathrm{kl}(1-\delta, \delta).$$

Therefore, we conclude by optimizing over $w_a \in \Delta(K)$ (the simplex over $1, \ldots, K$):

$$\mathbb{E}_\mu[\tau] \underbrace{\sup_{w \in \Delta(K)} \inf_{\mu' \in \mathrm{Alt}(\mu)} \sum_a w_a \mathrm{KL}(f_{\mu_a}, f_{\mu'_a})}_{=:(T^\star)^{-1}} \geq \mathrm{kl}(1-\delta, \delta).$$

Best Arm Identification with Fixed Confidence: lower bound (final)

Therefore

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

- where $T^\star = \sup_{w \in \Delta(K)} \inf_{\mu' \in \mathrm{Alt}(\mu)} \sum_a w_a \mathrm{KL}(f_{\mu_a}, f_{\mu'_a})$ is the characteristic time (and the inverse the information rate).
- The worst confusing model minimizes the information rate.
- The *allocation* w denotes the optimal exploration strategy (the one that maximizes the evidence).
- Can be seen as a zero-sum game between a player playing an exploration strategy, and a player playing a confusing model!

Best Arm Identification with Fixed Confidence: lower bound (final)

Therefore

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

- where $T^\star = \sup_{w \in \Delta(K)} \inf_{\mu' \in \mathrm{Alt}(\mu)} \sum_a w_a \mathrm{KL}(f_{\mu_a}, f_{\mu'_a})$ is the characteristic time (and the inverse the information rate).
- The worst confusing model minimizes the information rate.
- The *allocation* w denotes the optimal exploration strategy (the one that maximizes the evidence).
- Can be seen as a zero-sum game between a player playing an exploration strategy, and a player playing a confusing model!

Best Arm Identification with Fixed Confidence: lower bound (final)

Therefore

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

- where $T^\star = \sup_{w\in\Delta(K)} \inf_{\mu'\in\mathrm{Alt}(\mu)} \sum_a w_a \mathrm{KL}(f_{\mu_a}, f_{\mu'_a})$ is the characteristic time (and the inverse the information rate).
- The worst confusing model minimizes the information rate.
- The *allocation* w denotes the optimal exploration strategy (the one that maximizes the evidence).
- Can be seen as a zero-sum game between a player playing an exploration strategy, and a player playing a confusing model!

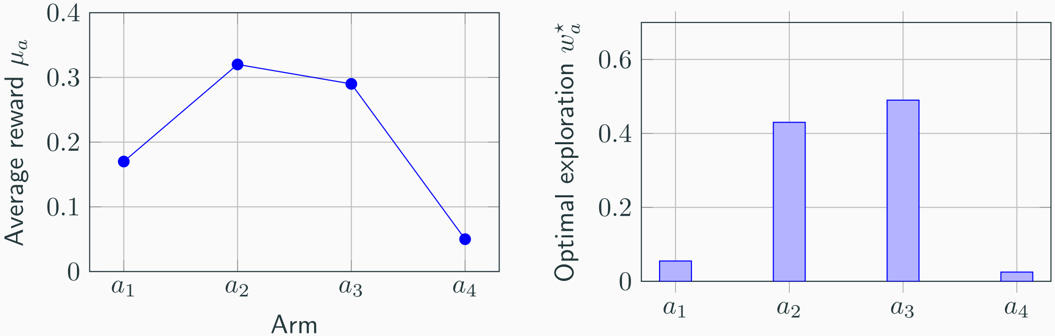
Best Arm Identification with Fixed Confidence: lower bound (final)

Therefore

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

- where $T^\star = \sup_{w \in \Delta(K)} \inf_{\mu' \in \mathrm{Alt}(\mu)} \sum_a w_a \mathrm{KL}(f_{\mu_a}, f_{\mu'_a})$ is the characteristic time (and the inverse the information rate).
- The worst confusing model minimizes the information rate.
- The *allocation* w denotes the optimal exploration strategy (the one that maximizes the evidence).
- Can be seen as a zero-sum game between a player playing an exploration strategy, and a player playing a confusing model!

Example with 4 arms



It's interesting to compare with the solution to the regret minimization problem. Here we spend most of the time sampling the top two arms[7].

[7] Which motivates why top-two algorithms [Rus16, JDB+22] are good for this type of problem.

Bonus: proof using the low-level form of the fundamental inequality

Lemma (Low-level form of the fundamental inequality [KCG16, GK21])

For all $x \in \mathbb{R}, t \in \mathbb{N}$ and all event $\mathcal{E} \in \mathcal{F}_\tau$ we have

$$\mathbb{P}_\mu(\mathcal{E}) \leq \mathbb{P}_\mu(Z_t \geq x) + e^x \mathbb{P}_{\mu'}(\mathcal{E}),$$

where $Z_t = \ln \frac{\mathrm{d}\mathbb{P}_\mu(X_1,\ldots,X_t)}{\mathrm{d}\mathbb{P}_{\mu'}(X_1,\ldots,X_t)}$ is the log-likelihood ratio.

- Let $t = n_\delta = T^\star(1-\beta)\ln(1/\delta)$ with $\beta \in (0,1)$ and define $\mathcal{E} = \{\tau \leq n_\delta\}$. We need to bound $\mathbb{P}_{\mu'}(\mathcal{E})$:

$$\mathbb{P}_{\mu'}(\mathcal{E}) = \sum_a \mathbb{P}_{\mu'}(\mathcal{E} \cap \{\hat{a}_\tau = a\}) \leq \delta + \mathbb{P}_{\mu'}(\mathcal{E} \cap \{\hat{a}_\tau = a^\star(\mu')\}).$$

- Apply the lemma to the latter term on the right hand side with $Z_{n'_\delta} = \ln \frac{\mathrm{d}\mathbb{P}_{\mu'}(X_1,\ldots,X_t)}{\mathrm{d}\mathbb{P}_{\mu}(X_1,\ldots,X_t)}$ and $n'_\delta = T^\star(\mu')(1-\beta)\ln(1/\delta)$, where $T^\star(\mu')^{-1}$ is the characteristic time in μ'. Then

$$\mathbb{P}_{\mu'}(\mathcal{E} \cap \{\hat{a}_\tau = a^\star(\mu')\}) \leq \mathbb{P}_{\mu'}(Z_{n'_\delta} \geq x) + e^x \delta.$$

Bonus: proof using the low-level form of the fundamental inequality

Lemma (Low-level form of the fundamental inequality [KCG16, GK21])

For all $x \in \mathbb{R}, t \in \mathbb{N}$ and all event $\mathcal{E} \in \mathcal{F}_\tau$ we have

$$\mathbb{P}_\mu(\mathcal{E}) \leq \mathbb{P}_\mu(Z_t \geq x) + e^x \mathbb{P}_{\mu'}(\mathcal{E}),$$

where $Z_t = \ln \frac{\mathrm{d}\mathbb{P}_\mu(X_1,\ldots,X_t)}{\mathrm{d}\mathbb{P}_{\mu'}(X_1,\ldots,X_t)}$ is the log-likelihood ratio.

- Let $t = n_\delta = T^\star(1-\beta)\ln(1/\delta)$ with $\beta \in (0,1)$ and define $\mathcal{E} = \{\tau \leq n_\delta\}$. We need to bound $\mathbb{P}_{\mu'}(\mathcal{E})$:

$$\mathbb{P}_{\mu'}(\mathcal{E}) = \sum_a \mathbb{P}_{\mu'}(\mathcal{E} \cap \{\hat{a}_\tau = a\}) \leq \delta + \mathbb{P}_{\mu'}(\mathcal{E} \cap \{\hat{a}_\tau = a^\star(\mu')\}).$$

- Apply the lemma to the latter term on the right hand side with $Z_{n'_\delta} = \ln \frac{\mathrm{d}\mathbb{P}_{\mu'}(X_1,\ldots,X_t)}{\mathrm{d}\mathbb{P}_{\mu}(X_1,\ldots,X_t)}$ and $n'_\delta = T^\star(\mu')(1-\beta)\ln(1/\delta)$, where $T^\star(\mu')^{-1}$ is the characteristic time in μ'. Then

$$\mathbb{P}_{\mu'}(\mathcal{E} \cap \{\hat{a}_\tau = a^\star(\mu')\}) \leq \mathbb{P}_{\mu'}(Z_{n'_\delta} \geq x) + e^x \delta.$$

Bonus: proof using the low-level form of the fundamental inequality

Lemma (Low-level form of the fundamental inequality [KCG16, GK21])

For all $x \in \mathbb{R}, t \in \mathbb{N}$ and all event $\mathcal{E} \in \mathcal{F}_\tau$ we have

$$\mathbb{P}_\mu(\mathcal{E}) \leq \mathbb{P}_\mu(Z_t \geq x) + e^x \mathbb{P}_{\mu'}(\mathcal{E}),$$

where $Z_t = \ln \frac{\mathrm{d}\mathbb{P}_\mu(X_1,\dots,X_t)}{\mathrm{d}\mathbb{P}_{\mu'}(X_1,\dots,X_t)}$ is the log-likelihood ratio.

- Let $t = n_\delta = T^\star(1-\beta)\ln(1/\delta)$ with $\beta \in (0,1)$ and define $\mathcal{E} = \{\tau \leq n_\delta\}$. We need to bound $\mathbb{P}_{\mu'}(\mathcal{E})$:

$$\mathbb{P}_{\mu'}(\mathcal{E}) = \sum_a \mathbb{P}_{\mu'}(\mathcal{E} \cap \{\hat{a}_\tau = a\}) \leq \delta + \mathbb{P}_{\mu'}(\mathcal{E} \cap \{\hat{a}_\tau = a^\star(\mu')\}).$$

- Apply the lemma to the latter term on the right hand side with $Z_{n'_\delta} = \ln \frac{\mathrm{d}\mathbb{P}_{\mu'}(X_1,\dots,X_t)}{\mathrm{d}\mathbb{P}_\mu(X_1,\dots,X_t)}$ and $n'_\delta = T^\star(\mu')(1-\beta)\ln(1/\delta)$, where $T^\star(\mu')^{-1}$ is the characteristic time in μ'. Then

$$\mathbb{P}_{\mu'}(\mathcal{E} \cap \{\hat{a}_\tau = a^\star(\mu')\}) \leq \mathbb{P}_{\mu'}(Z_{n'_\delta} \geq x) + e^x \delta.$$

Bonus: proof using the low-level form of the fundamental inequality (cont.)

- Let $x = (1 - \beta/2)\ln(1/\delta)$.

$$\mathbb{P}_{\mu'}(\mathcal{E} \cap \{\hat{a}_\tau = a^\star(\mu')\}) \leq \underbrace{\mathbb{P}_{\mu'}\left(\frac{Z_{n'_\delta}}{n'_\delta} \geq \frac{(1-\beta/2)}{(1-\beta)T^\star(\mu')}\right)}_{\to 0} + \delta^\beta \to 0 \text{ as } \delta \to 0.$$

- Similarly, one can show that $\mathbb{P}_\mu(Z_{n_\delta} \geq (1-\beta/2)\ln(1/\delta)) \to 0$ as $\delta \to 0$, with $Z_t = \ln \frac{\mathrm{d}\mathbb{P}_\mu(X_1,\ldots,X_t)}{\mathrm{d}\mathbb{P}_{\mu'}(X_1,\ldots,X_t)}$ (reversed numerator and denominator).

Bonus: proof using the low-level form of the fundamental inequality (cont.)

- Let $x = (1 - \beta/2)\ln(1/\delta)$.

$$\mathbb{P}_{\mu'}(\mathcal{E} \cap \{\hat{a}_\tau = a^\star(\mu')\}) \le \underbrace{\mathbb{P}_{\mu'}\left(\frac{Z_{n'_\delta}}{n'_\delta} \ge \frac{(1-\beta/2)}{(1-\beta)T^\star(\mu')}\right)}_{\to 0} + \delta^\beta \to 0 \text{ as } \delta \to 0.$$

- Similarly, one can show that $\mathbb{P}_\mu(Z_{n_\delta} \ge (1-\beta/2)\ln(1/\delta)) \to 0$ as $\delta \to 0$, with $Z_t = \ln \frac{\mathrm{d}\mathbb{P}_\mu(X_1,\ldots,X_t)}{\mathrm{d}\mathbb{P}_{\mu'}(X_1,\ldots,X_t)}$ (reversed numerator and denominator).

Bonus: proof using the low-level form of the fundamental inequality (final)

- Hence $\mathbb{P}_\mu(\mathcal{E}) \to 0$ as $\delta \to 0$, with $\mathcal{E} = \{\tau \le T^\star(1-\beta)\ln(1/\delta)\}$
- Therefore $\forall \beta \in (0,1)$ we have $\mathbb{P}_\mu(\tau \ge T^\star(1-\beta)\ln(1/\delta)) \to 1$, and we conclude that

$$\liminf_{\delta \to 0} \frac{\mathbb{E}_\mu[\tau]}{\ln(1/\delta)} \ge T^\star.$$

- This technique can be more technical but usually gives asymptotic bounds.

Bonus: proof using the low-level form of the fundamental inequality (final)

- Hence $\mathbb{P}_\mu(\mathcal{E}) \to 0$ as $\delta \to 0$, with $\mathcal{E} = \{\tau \leq T^\star (1-\beta)\ln(1/\delta)\}$
- Therefore $\forall \beta \in (0,1)$ we have $\mathbb{P}_\mu(\tau \geq T^\star(1-\beta)\ln(1/\delta)) \to 1$, and we conclude that

$$\liminf_{\delta \to 0} \frac{\mathbb{E}_\mu[\tau]}{\ln(1/\delta)} \geq T^\star.$$

- This technique can be more technical but usually gives asymptotic bounds.

Bonus: proof using the low-level form of the fundamental inequality (final)

- Hence $\mathbb{P}_\mu(\mathcal{E}) \to 0$ as $\delta \to 0$, with $\mathcal{E} = \{\tau \leq T^\star(1-\beta)\ln(1/\delta)\}$
- Therefore $\forall \beta \in (0,1)$ we have $\mathbb{P}_\mu(\tau \geq T^\star(1-\beta)\ln(1/\delta)) \to 1$, and we conclude that

$$\liminf_{\delta \to 0} \frac{\mathbb{E}_\mu[\tau]}{\ln(1/\delta)} \geq T^\star.$$

- This technique can be more technical but usually gives asymptotic bounds.

Best Arm Identification with Fixed Confidence in Linear Bandits Models

Best Arm Identification with Fixed Confidence in Linear Bandits: introduction

The setting is similar to the previous one with Gaussian rewards:

- Similarly as before, we consider a bandit model with a finite number of arms $\{x_1, \dots, x_K\} \in \mathbb{R}^d$.
- The model is characterized by a parameter vector $\mu \in \mathbb{R}^d$.
- We denote by $Y_t = \mu^\top X_t + \eta_t$ the reward observed at time t, and X_t is the arm chosen at time t, and $\eta_t \sim \mathcal{N}(0, 1)$ (i.i.d.).
- We consider models μ with a unique best arm, $a^\star(\mu) = \arg\max_k \mu^\top x_k$.

Best Arm Identification with Fixed Confidence in Linear Bandits: lower bound

To derive a lower bound, define the set of confusing models

$$\mathrm{Alt}(\mu) = \left\{ \mu' : a^\star(\mu) \neq a^\star(\mu'), |\arg\max_k (\mu')^\top x_k| = 1 \right\}.$$

Compute the log-likelihood ratio between μ and $\mu' \in \mathrm{Alt}(\mu)$:

$$\mathbb{E}_\mu[Z_t] = \mathbb{E}_\mu\left[\ln \frac{\mathrm{d}\mathbb{P}_\mu(X_1, Y_1, \ldots, X_t, Y_t)}{\mathrm{d}\mathbb{P}_{\mu'}(X_1, Y_1, \ldots, X_t, Y_t)}\right] = \mathbb{E}_\mu\left[\sum_{n=1}^{t} \ln \frac{f_\mu(Y_n|X_n)}{f_{\mu'}(Y_n|X_n)}\right],$$

$$= \sum_a \mathbb{E}_\mu\left[N_a(t)\right] \mathrm{KL}\left(\mathcal{N}(\mu^\top x_a, 1), \mathcal{N}((\mu')^\top x_a, 1)\right),$$

$$= \frac{1}{2} \sum_a \mathbb{E}_\mu\left[N_a(t)\right] ((\mu - \mu')^\top x_a)^2,$$

$$= \frac{1}{2} (\mu - \mu')^\top \left(\sum_a \mathbb{E}_\mu\left[N_a(t)\right] x_a x_a^\top\right) (\mu - \mu').$$

Best Arm Identification with Fixed Confidence in Linear Bandits: lower bound

To derive a lower bound, define the set of confusing models

$$\mathrm{Alt}(\mu) = \left\{ \mu' : a^\star(\mu) \neq a^\star(\mu'), |\arg\max_k (\mu')^\top x_k| = 1 \right\} .$$

Compute the log-likelihood ratio between μ and $\mu' \in \mathrm{Alt}(\mu)$:

$$\begin{aligned}
\mathbb{E}_\mu[Z_t] &= \mathbb{E}_\mu \left[\ln \frac{\mathrm{d}\mathbb{P}_\mu(X_1, Y_1, \dots, X_t, Y_t)}{\mathrm{d}\mathbb{P}_{\mu'}(X_1, Y_1, \dots, X_t, Y_t)} \right] = \mathbb{E}_\mu \left[\sum_{n=1}^{t} \ln \frac{f_\mu(Y_n|X_n)}{f_{\mu'}(Y_n|X_n)} \right] , \\
&= \sum_a \mathbb{E}_\mu \left[N_a(t) \right] \mathrm{KL} \left(\mathcal{N}(\mu^\top x_a, 1), \mathcal{N}((\mu')^\top x_a, 1) \right) , \\
&= \frac{1}{2} \sum_a \mathbb{E}_\mu \left[N_a(t) \right] ((\mu - \mu')^\top x_a)^2 , \\
&= \frac{1}{2} (\mu - \mu')^\top \left(\sum_a \mathbb{E}_\mu \left[N_a(t) \right] x_a x_a^\top \right) (\mu - \mu') .
\end{aligned}$$

Best Arm Identification with Fixed Confidence in Linear Bandits: lower bound

To derive a lower bound, define the set of confusing models

$$\mathrm{Alt}(\mu) = \left\{\mu' : a^\star(\mu) \neq a^\star(\mu'), |\arg\max_k (\mu')^\top x_k| = 1\right\}.$$

Compute the log-likelihood ratio between μ and $\mu' \in \mathrm{Alt}(\mu)$:

$$\begin{aligned}
\mathbb{E}_\mu[Z_t] &= \mathbb{E}_\mu\left[\ln \frac{\mathrm{d}\mathbb{P}_\mu(X_1, Y_1, \ldots, X_t, Y_t)}{\mathrm{d}\mathbb{P}_{\mu'}(X_1, Y_1, \ldots, X_t, Y_t)}\right] = \mathbb{E}_\mu\left[\sum_{n=1}^{t} \ln \frac{f_\mu(Y_n|X_n)}{f_{\mu'}(Y_n|X_n)}\right], \\
&= \sum_a \mathbb{E}_\mu\left[N_a(t)\right] \mathrm{KL}\left(\mathcal{N}(\mu^\top x_a, 1), \mathcal{N}((\mu')^\top x_a, 1)\right), \\
&= \frac{1}{2} \sum_a \mathbb{E}_\mu\left[N_a(t)\right] ((\mu - \mu')^\top x_a)^2, \\
&= \frac{1}{2} (\mu - \mu')^\top \left(\sum_a \mathbb{E}_\mu\left[N_a(t)\right] x_a x_a^\top\right) (\mu - \mu').
\end{aligned}$$

Best Arm Identification with Fixed Confidence in Linear Bandits: lower bound

To derive a lower bound, define the set of confusing models

$$\mathrm{Alt}(\mu) = \left\{\mu' : a^\star(\mu) \neq a^\star(\mu'), |\arg\max_k (\mu')^\top x_k| = 1\right\}.$$

Compute the log-likelihood ratio between μ and $\mu' \in \mathrm{Alt}(\mu)$:

$$\begin{aligned}
\mathbb{E}_\mu[Z_t] &= \mathbb{E}_\mu\left[\ln \frac{\mathrm{d}\mathbb{P}_\mu(X_1, Y_1, \ldots, X_t, Y_t)}{\mathrm{d}\mathbb{P}_{\mu'}(X_1, Y_1, \ldots, X_t, Y_t)}\right] = \mathbb{E}_\mu\left[\sum_{n=1}^t \ln \frac{f_\mu(Y_n|X_n)}{f_{\mu'}(Y_n|X_n)}\right], \\
&= \sum_a \mathbb{E}_\mu\left[N_a(t)\right] \mathrm{KL}\left(\mathcal{N}(\mu^\top x_a, 1), \mathcal{N}((\mu')^\top x_a, 1)\right), \\
&= \frac{1}{2}\sum_a \mathbb{E}_\mu\left[N_a(t)\right]((\mu - \mu')^\top x_a)^2, \\
&= \frac{1}{2}(\mu - \mu')^\top \left(\sum_a \mathbb{E}_\mu\left[N_a(t)\right] x_a x_a^\top\right)(\mu - \mu').
\end{aligned}$$

Best Arm Identification with Fixed Confidence in Linear Bandits: lower bound (final)

Therefore, as before we apply the fundamental inequality at the stopping time and consider the worst confusing model:

$$\frac{1}{2} \inf_{\mu' \in \mathrm{Alt}(\mu)} (\mu - \mu')^\top \left(\sum_a \mathbb{E}_\mu \left[N_a(\tau) \right] x_a x_a^\top \right) (\mu - \mu') \geq \mathrm{kl}(1-\delta, \delta).$$

Let $w_a = \mathbb{E}_\mu[N_a(\tau)]/\mathbb{E}_\mu[\tau]$ and take the supremum over $w \in \Delta(K)$

$$\frac{\mathbb{E}_\mu[\tau]}{2} \sup_{w \in \Delta(K)} \inf_{\mu' \in \mathrm{Alt}(\mu)} (\mu - \mu')^\top \left(\sum_a w_a x_a x_a^\top \right) (\mu - \mu') \geq \mathrm{kl}(1-\delta, \delta).$$

We conclude that

$$\mathbb{E}_\mu[\tau] \geq T^\star \mathrm{kl}(1-\delta, \delta), \quad (T^\star)^{-1} := \frac{1}{2} \sup_{w \in \Delta(K)} \inf_{\mu' \in \mathrm{Alt}(\mu)} (\mu - \mu')^\top \left(\sum_a w_a x_a x_a^\top \right) (\mu - \mu').$$

Best Arm Identification with Fixed Confidence in Linear Bandits: lower bound (final)

Therefore, as before we apply the fundamental inequality at the stopping time and consider the worst confusing model:

$$\frac{1}{2}\inf_{\mu'\in\mathrm{Alt}(\mu)}(\mu-\mu')^\top\left(\sum_a \mathbb{E}_\mu\left[N_a(\tau)\right]x_a x_a^\top\right)(\mu-\mu')\geq \mathrm{kl}(1-\delta,\delta).$$

Let $w_a=\mathbb{E}_\mu[N_a(\tau)]/\mathbb{E}_\mu[\tau]$ and take the supremum over $w\in\Delta(K)$

$$\frac{\mathbb{E}_\mu[\tau]}{2}\sup_{w\in\Delta(K)}\inf_{\mu'\in\mathrm{Alt}(\mu)}(\mu-\mu')^\top\left(\sum_a w_a x_a x_a^\top\right)(\mu-\mu')\geq \mathrm{kl}(1-\delta,\delta).$$

We conclude that

$$\mathbb{E}_\mu[\tau]\geq T^\star \mathrm{kl}(1-\delta,\delta),\quad (T^\star)^{-1}:=\frac{1}{2}\sup_{w\in\Delta(K)}\inf_{\mu'\in\mathrm{Alt}(\mu)}(\mu-\mu')^\top\left(\sum_a w_a x_a x_a^\top\right)(\mu-\mu').$$

Best Arm Identification with Fixed Confidence in Linear Bandits: lower bound (final)

Therefore, as before we apply the fundamental inequality at the stopping time and consider the worst confusing model:

$$\frac{1}{2} \inf_{\mu' \in \mathrm{Alt}(\mu)} (\mu - \mu')^\top \left(\sum_a \mathbb{E}_\mu \left[N_a(\tau) \right] x_a x_a^\top \right) (\mu - \mu') \geq \mathrm{kl}(1-\delta, \delta).$$

Let $w_a = \mathbb{E}_\mu[N_a(\tau)]/\mathbb{E}_\mu[\tau]$ and take the supremum over $w \in \Delta(K)$

$$\frac{\mathbb{E}_\mu[\tau]}{2} \sup_{w \in \Delta(K)} \inf_{\mu' \in \mathrm{Alt}(\mu)} (\mu - \mu')^\top \left(\sum_a w_a x_a x_a^\top \right) (\mu - \mu') \geq \mathrm{kl}(1-\delta, \delta).$$

We conclude that

$$\mathbb{E}_\mu[\tau] \geq T^\star \mathrm{kl}(1-\delta, \delta), \quad (T^\star)^{-1} := \frac{1}{2} \sup_{w \in \Delta(K)} \inf_{\mu' \in \mathrm{Alt}(\mu)} (\mu - \mu')^\top \left(\sum_a w_a x_a x_a^\top \right) (\mu - \mu').$$

Bonus: Best Arm Identification with Fixed Confidence for Multi-Task Bandit (with Task Selection)

Best Arm Identification for Multi-Task Bandit: introduction

We now look at a more complex case, a multi-task multi-armed bandit model [RP23].

- Consider a model with X tasks. There are G global arms and H local arms.
- At each time-step t the agent selects a task $x_t \in [X]$ and an arm $(g_t, h_t) \in [G] \times [H]$ and observes a Bernoulli reward $Y_t \sim \mathrm{Ber}(\mu_{x_t,g_t,h_t})$.
- μ is the parameter of the model, and $\mu_{x,g,h}$ is the average reward for task x, and arm (g, h).
- The structure of this model is pretty unique. We assume there exists $g^\star \in [G]$ and $h_x^\star$ for each task such that

$$\mu_{x,g^\star,h_x^\star} > \mu_{x,g,h} \quad \forall x, g \neq g^\star, h \neq h_x^\star$$

  In a sense, there exists a global optimal arm $g^\star$ and a local optimal arm $h_x^\star$ that is task specific.

Best Arm Identification for Multi-Task Bandit: introduction

We now look at a more complex case, a multi-task multi-armed bandit model [RP23].

- Consider a model with X tasks. There are G global arms and H local arms.
- At each time-step t the agent selects a task $x_t \in [X]$ and an arm $(g_t, h_t) \in [G] \times [H]$ and observes a Bernoulli reward $Y_t \sim \mathrm{Ber}(\mu_{x_t, g_t, h_t})$.
- μ is the parameter of the model, and $\mu_{x,g,h}$ is the average reward for task x, and arm (g, h).
- The structure of this model is pretty unique. We assume there exists $g^\star \in [G]$ and $h_x^\star$ for each task such that

$$\mu_{x,g^\star,h_x^\star} > \mu_{x,g,h} \quad \forall x, g \neq g^\star, h \neq h_x^\star$$

In a sense, there exists a global optimal arm $g^\star$ and a local optimal arm $h_x^\star$ that is task specific.

Best Arm Identification for Multi-Task Bandit: introduction

We now look at a more complex case, a multi-task multi-armed bandit model [RP23].

- Consider a model with X tasks. There are G global arms and H local arms.
- At each time-step t the agent selects a task $x_t \in [X]$ and an arm $(g_t, h_t) \in [G] \times [H]$ and observes a Bernoulli reward $Y_t \sim \mathrm{Ber}(\mu_{x_t, g_t, h_t})$.
- μ is the parameter of the model, and $\mu_{x,g,h}$ is the average reward for task x, and arm (g, h).
- The structure of this model is pretty unique. We assume there exists $g^\star \in [G]$ and $h_x^\star$ for each task such that

$$\mu_{x,g^\star,h_x^\star} > \mu_{x,g,h} \quad \forall x, g \neq g^\star, h \neq h_x^\star$$

In a sense, there exists a global optimal arm $g^\star$ and a local optimal arm $h_x^\star$ that is task specific.

Best Arm Identification for Multi-Task Bandit: introduction

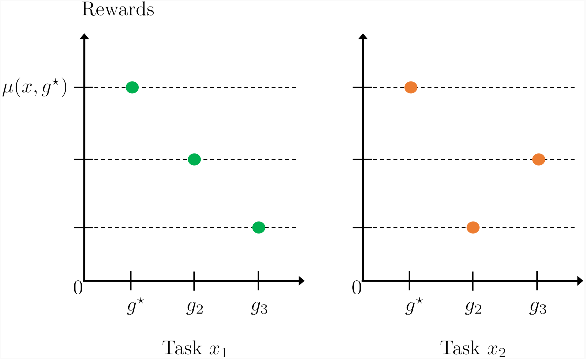
We now look at a more complex case, a multi-task multi-armed bandit model [RP23].

- Consider a model with X tasks. There are G global arms and H local arms.
- At each time-step t the agent selects a task $x_t \in [X]$ and an arm $(g_t, h_t) \in [G] \times [H]$ and observes a Bernoulli reward $Y_t \sim \mathrm{Ber}(\mu_{x_t, g_t, h_t})$.
- μ is the parameter of the model, and $\mu_{x,g,h}$ is the average reward for task x, and arm (g, h).
- The structure of this model is pretty unique. We assume there exists $g^\star \in [G]$ and $h_x^\star$ for each task such that

$$\mu_{x,g^\star,h_x^\star} > \mu_{x,g,h} \quad \forall x, g \neq g^\star, h \neq h_x^\star$$

  In a sense, there exists a global optimal arm $g^\star$ and a local optimal arm $h_x^\star$ that is task specific.

Best Arm Identification for Multi-Task Bandit: introduction (cont.)



Example with $X = 2, G = 3, H = 1$.

(δ_G, δ_H)-PC algorithm

We say that an algorithm is (δ_G, δ_H)-PC if $\mathbb{P}_\mu(\hat{g} \neq g^\star) \leq \delta_G$ and $\mathbb{P}_\mu(\hat{h}_x \neq h^\star_x, \hat{g} = g^\star) \leq \delta_H$.

Best Arm Identification for Multi-Task Bandit: lower bound

Assumption

$$\mu_{x,g^\star,h_x^\star} > \mu_{x,g,h} \quad \forall x, g \neq g^\star, h \neq h_x^\star.$$

Let $g^\star(\mu)$ denote the optimal global arm for a model μ. Similarly $h_x^\star(\mu)$. Then, the first step is to define the set of confusing models:

$$\mathrm{Alt}(\mu) = \{\mu' : (g^\star, h_1^\star, \ldots, h_X^\star)(\mu) \neq (g^\star, h_1^\star, \ldots, h_X^\star)(\mu')\}$$

Secondly, we note that

$$\mathrm{Alt}(\mu) = \mathrm{Alt}_1(\mu) \cup [\cup_x \mathrm{Alt}_2(\mu, x)]$$

where $\mathrm{Alt}_1(\mu) = \{\mu' : g^\star(\mu) \neq g^\star(\mu')\}$ and
$\mathrm{Alt}_2(\mu, x) = \{\mu' : g^\star(\mu) = g^\star(\mu'), h_x^\star(\mu) \neq h_x^\star(\mu')\}$.

Best Arm Identification for Multi-Task Bandit: lower bound

Assumption

$$\mu_{x,g^\star,h_x^\star} > \mu_{x,g,h} \quad \forall x, g \neq g^\star, h \neq h_x^\star.$$

Let $g^\star(\mu)$ denote the optimal global arm for a model μ. Similarly $h_x^\star(\mu)$. Then, the first step is to define the set of confusing models:

$$\mathrm{Alt}(\mu) = \{\mu' : (g^\star, h_1^\star, \dots, h_X^\star)(\mu) \neq (g^\star, h_1^\star, \dots, h_X^\star)(\mu')\}$$

Secondly, we note that

$$\mathrm{Alt}(\mu) = \mathrm{Alt}_1(\mu) \cup [\cup_x \mathrm{Alt}_2(\mu, x)]$$

where $\mathrm{Alt}_1(\mu) = \{\mu' : g^\star(\mu) \neq g^\star(\mu')\}$ and
$\mathrm{Alt}_2(\mu, x) = \{\mu' : g^\star(\mu) = g^\star(\mu'), h_x^\star(\mu) \neq h_x^\star(\mu')\}$.

Best Arm Identification for Multi-Task Bandit: lower bound (cont.)

Hence, we can find the worst confusing model over each set.

- First, consider the log-likelihood ratio, and consider the worst model over $\mathrm{Alt}_1(\mu) = \{\mu' : g^\star(\mu) \neq g^\star(\mu')\}$ [8]

$$\inf_{\mu' \in \mathrm{Alt}_1(\mu)} \sum_{x,g,h} \mathbb{E}_\mu[N_{x,g,h}(\tau)] \mathrm{kl}(\mu_{x,g,h}, \mu'_{x,g,h}) \geq \mathrm{kl}(1-\delta_G, \delta_G).$$

One can immediately see that it can be rewritten as

$$\min_{\bar{g} \neq g^\star, \bar{\mathbf{h}} \in [H]^X} \sum_x \inf_{\mu' \in \mathrm{Alt}(\mu, \bar{g}, \bar{\mathbf{h}})} \sum_{g,h} \mathbb{E}_\mu[N_{x,g,h}(\tau)] \mathrm{kl}(\mu_{x,g,h}, \mu'_{x,g,h}) \geq \mathrm{kl}(1-\delta_G, \delta_G).$$

where $\mathrm{Alt}(\mu, \bar{g}, \bar{\mathbf{h}}) = \{\mu' : (g^\star, h_1^\star, \ldots, h_X^\star)(\mu') = (\bar{g}, \bar{\mathbf{h}})\}$.

[8] When applying the fundamental lemma we consider the event $\mathcal{E} = \{\hat{g} \neq g^\star\}$.

Best Arm Identification for Multi-Task Bandit: lower bound (cont.)

Hence, we can find the worst confusing model over each set.

- First, consider the log-likelihood ratio, and consider the worst model over $\mathrm{Alt}_1(\mu) = \{\mu' : g^\star(\mu) \neq g^\star(\mu')\}$ [8]

$$\inf_{\mu' \in \mathrm{Alt}_1(\mu)} \sum_{x,g,h} \mathbb{E}_\mu[N_{x,g,h}(\tau)]\mathrm{kl}(\mu_{x,g,h}, \mu'_{x,g,h}) \geq \mathrm{kl}(1-\delta_G, \delta_G).$$

One can immediately see that it can be rewritten as

$$\min_{\bar{g} \neq g^\star, \bar{\mathbf{h}} \in [H]^X} \sum_x \inf_{\mu' \in \mathrm{Alt}(\mu, \bar{g}, \bar{\mathbf{h}})} \sum_{g,h} \mathbb{E}_\mu[N_{x,g,h}(\tau)]\mathrm{kl}(\mu_{x,g,h}, \mu'_{x,g,h}) \geq \mathrm{kl}(1-\delta_G, \delta_G).$$

where $\mathrm{Alt}(\mu, \bar{g}, \bar{\mathbf{h}}) = \{\mu' : (g^\star, h_1^\star, \dots, h_X^\star)(\mu') = (\bar{g}, \bar{\mathbf{h}})\}$.

---

[8] When applying the fundamental lemma we consider the event $\mathcal{E} = \{\hat{g} \neq g^\star\}$.

Best Arm Identification for Multi-Task Bandit: lower bound (cont.)

- Secondly, consider the log-likelihood ratio, and consider the worst model over $\mathrm{Alt}_2(\mu, x) = \{\mu' : g^\star(\mu) = g^\star(\mu'), h^\star_x(\mu) \neq h^\star_x(\mu')\}$ for a task x[9]

$$
\begin{aligned}
&\inf_{\mu' \in \mathrm{Alt}_2(\mu,x)} \sum_{x,g,h} \mathbb{E}_\mu[N_{x,g,h}(\tau)]\mathrm{kl}(\mu_{x,g,h}, \mu'_{x,g,h}) \\
&\quad = \inf_{\mu' \in \mathrm{Alt}_2(\mu,x)} \sum_{h} \mathbb{E}_\mu[N_{x,g^\star,h}(\tau)]\mathrm{kl}(\mu_{x,g^\star,h}, \mu'_{x,g^\star,h}), \\
&\quad \geq \mathrm{kl}(1-\delta_H, \delta_H).
\end{aligned}
$$

[9] When applying the fundamental lemma we consider the event $\mathcal{E} = \{\hat{h}_x \neq h^\star_x, \hat{g} = g^\star\}$.

Best Arm Identification for Multi-Task Bandit: lower bound (final)

Putting everything together, and letting $\eta_{x,g,h} = \mathbb{E}[N_{x,g,h}(\tau)]$, we get that the sample complexity lower bound is

$$\begin{aligned}
\min_{\eta} \quad & \sum_{x,g,h} \eta_{x,g,h} \\
\text{s.t.} \quad & \inf_{\mu' \in \mathrm{Alt}_2(\mu,x)} \sum_{h} \eta_{x,g^\star,h} \mathrm{kl}(\mu_{x,g^\star,h}, \mu'_{x,g^\star,h}) \geq \mathrm{kl}(1-\delta_H, \delta_H) \quad \forall x, \\
& \min_{\bar{g} \neq g^\star, \bar{\mathbf{h}} \in [H]^X} \sum_{x} \inf_{\mu' \in \mathrm{Alt}(\mu,\bar{g},\bar{\mathbf{h}})} \sum_{g,h} \eta_{x,g,h} \mathrm{kl}(\mu_{x,g,h}, \mu'_{x,g,h}) \geq \mathrm{kl}(1-\delta_G, \delta_G).
\end{aligned}$$

and the optimal exploration strategy is given by $w_{x,g,h} = \eta_{x,g,h} / \sum_{x,g,h} \eta_{x,g,h}$.

Bonus: Why extending to ϵ-best arm is not trivial

Extending BAI to ϵ-BAI 1/2

- While extending the classical BAI framework to the ϵ-best arm identification case, one (see, e.g., [TJP23]) may be tempted to write the set of confusing models as

$$\mathrm{Alt}(\mu) = \{\mu' : i^\star_\epsilon(\mu) \cap i^\star_\epsilon(\mu') = \emptyset\},$$

where $i^\star_\epsilon(\mu) = \{a : \mu_a \geq \max_{a'} \mu_{a'} - \epsilon\}$ is the set of ϵ-best arms in μ.

- Is it tight?
- Hard to prove the δ-error rate property:

$$\{\hat{a}_\tau \notin i^\star_\epsilon(\mu)\} \subset \{\hat{\mu}_t \in \mathrm{Alt}(\mu)\}$$

does not necessarily hold (the set of ϵ-best arm in $\hat{\mu}_t$ may intersect with $i^\star_\epsilon(\mu)$).

- For whoever is interested, probably the approach in [DK19] is the right one, but the stopping condition is not practically usable(computational issues).

Extending BAI to ϵ-BAI 1/2

- While extending the classical BAI framework to the ϵ-best arm identification case, one (see, e.g., [TJP23]) may be tempted to write the set of confusing models as

$$\mathrm{Alt}(\mu) = \{\mu' : i^\star_\epsilon(\mu) \cap i^\star_\epsilon(\mu') = \emptyset\},$$

  where $i^\star_\epsilon(\mu) = \{a : \mu_a \geq \max_{a'} \mu_{a'} - \epsilon\}$ is the set of ϵ-best arms in μ.
- Is it tight?
- Hard to prove the δ-error rate property:

$$\{\hat{a}_\tau \notin i^\star_\epsilon(\mu)\} \subset \{\hat{\mu}_t \in \mathrm{Alt}(\mu)\}$$

  does not necessarily hold (the set of ϵ-best arm in $\hat{\mu}_t$ may intersect with $i^\star_\epsilon(\mu)$).
- For whoever is interested, probably the approach in [DK19] is the right one, but the stopping condition is not practically usable(computational issues).

Extending BAI to ϵ-BAI 1/2

- While extending the classical BAI framework to the ϵ-best arm identification case, one (see, e.g., [TJP23]) may be tempted to write the set of confusing models as

$$\mathrm{Alt}(\mu) = \{\mu' : i^\star_\epsilon(\mu) \cap i^\star_\epsilon(\mu') = \emptyset\},$$

  where $i^\star_\epsilon(\mu) = \{a : \mu_a \geq \max_{a'} \mu_{a'} - \epsilon\}$ is the set of ϵ-best arms in μ.
- Is it tight?
- Hard to prove the δ-error rate property:

$$\{\hat{a}_\tau \notin i^\star_\epsilon(\mu)\} \subset \{\hat{\mu}_t \in \mathrm{Alt}(\mu)\}$$

  does not necessarily hold (the set of ϵ-best arm in $\hat{\mu}_t$ may intersect with $i^\star_\epsilon(\mu)$).
- For whoever is interested, probably the approach in [DK19] is the right one, but the stopping condition is not practically usable(computational issues).

Extending BAI to ϵ-BAI 1/2

- While extending the classical BAI framework to the ϵ-best arm identification case, one (see, e.g., [TJP23]) may be tempted to write the set of confusing models as

$$\mathrm{Alt}(\mu) = \{\mu' : i^\star_\epsilon(\mu) \cap i^\star_\epsilon(\mu') = \emptyset\},$$

  where $i^\star_\epsilon(\mu) = \{a : \mu_a \geq \max_{a'} \mu_{a'} - \epsilon\}$ is the set of ϵ-best arms in μ.
- Is it tight?
- Hard to prove the δ-error rate property:

$$\{\hat{a}_\tau \notin i^\star_\epsilon(\mu)\} \subset \{\hat{\mu}_t \in \mathrm{Alt}(\mu)\}$$

  does not necessarily hold (the set of ϵ-best arm in $\hat{\mu}_t$ may intersect with $i^\star_\epsilon(\mu)$).
- For whoever is interested, probably the approach in [DK19] is the right one, but the stopping condition is not practically usable(computational issues).

Extending BAI to ϵ-BAI 2/2

- A tighter bound is obtained by instead writing

$$\mathrm{Alt}(\mu) = \{\mu' : \exists a \in i^\star_\epsilon(\mu') : a \notin i^\star_\epsilon(\mu)\}.$$

- However, now it becomes combinatorially hard to optimize $\inf_{\mu' \in \mathrm{Alt}(\mu')} \sum_a w_a \mathrm{KL}(\mu_a, \mu'_a)$.
- Sometimes [AMKG22] people write

$$\mathrm{Alt}(\mu) = \{\mu' : i^\star_\epsilon(\mu) \neq i^\star_\epsilon(\mu')\}$$

which is not correct, since that may imply $i^\star_\epsilon(\mu') \subset i^\star_\epsilon(\mu)$ (while the other way around is fine).
- For more info, refer to [DK19, GK21, JDK23].

Extending BAI to ϵ-BAI 2/2

- A tighter bound is obtained by instead writing

$$\mathrm{Alt}(\mu) = \{\mu' : \exists a \in i^\star_\epsilon(\mu') : a \notin i^\star_\epsilon(\mu)\}.$$

- However, now it becomes combinatorially hard to optimize $\inf_{\mu' \in \mathrm{Alt}(\mu')} \sum_a w_a \mathrm{KL}(\mu_a, \mu'_a)$.
- Sometimes [AMKG22] people write

$$\mathrm{Alt}(\mu) = \{\mu' : i^\star_\epsilon(\mu) \neq i^\star_\epsilon(\mu')\}$$

which is not correct, since that may imply $i^\star_\epsilon(\mu') \subset i^\star_\epsilon(\mu)$ (while the other way around is fine).
- For more info, refer to [DK19, GK21, JDK23].

Extending BAI to ϵ-BAI 2/2

- A tighter bound is obtained by instead writing

$$\mathrm{Alt}(\mu) = \{\mu' : \exists a \in i^\star_\epsilon(\mu') : a \notin i^\star_\epsilon(\mu)\}.$$

- However, now it becomes combinatorially hard to optimize $\inf_{\mu' \in \mathrm{Alt}(\mu')} \sum_a w_a \mathrm{KL}(\mu_a, \mu'_a)$.
- Sometimes [AMKG22] people write

$$\mathrm{Alt}(\mu) = \{\mu' : i^\star_\epsilon(\mu) \neq i^\star_\epsilon(\mu')\}$$

  which is not correct, since that may imply $i^\star_\epsilon(\mu') \subset i^\star_\epsilon(\mu)$ (while the other way around is fine).
- For more info, refer to [DK19, GK21, JDK23].

Conclusions

- These lower bounds help understanding the difficulty of a problem.
- *Usually* the optimization problems are convex (but the solution is not necessarily unique!).
- Not clear how to extend the technique to continuous set of arms.
- The fixed budget setting is in general much harder to deal with, and one requires different techniques.

Thanks for your attention!

References i

- Aymen Al Marjani, Tomas Kocak, and Aurélien Garivier, *On the complexity of all ε-best arms identification*, Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Springer, 2022, pp. 317–332.
- Apostolos N Burnetas and Michael N Katehakis, *Optimal adaptive policies for sequential allocation problems*, Advances in Applied Mathematics **17** (1996), no. 2, 122–142.
- ———, *Optimal adaptive policies for markov decision processes*, Mathematics of Operations Research **22** (1997), no. 1, 222–255.
- Rémy Degenne and Wouter M Koolen, *Pure exploration with multiple correct answers*, Advances in Neural Information Processing Systems **32** (2019).
- Aurélien Garivier and Emilie Kaufmann, *Optimal best arm identification with fixed confidence*, Conference on Learning Theory, PMLR, 2016, pp. 998–1027.
- ———, *Nonasymptotic sequential tests for overlapping hypotheses applied to near-optimal arm identification in bandit models*, Sequential Analysis **40** (2021), no. 1, 61–96.

- Todd L Graves and Tze Leung Lai, *Asymptotically efficient adaptive choice of control laws incontrolled markov chains*, SIAM journal on control and optimization **35** (1997), no. 3, 715–743.
- Aurélien Garivier, Pierre Ménard, and Gilles Stoltz, *Explore first, exploit next: The true shape of regret in bandit problems*, Mathematics of Operations Research **44** (2019), no. 2, 377–399.
- Marc Jourdan, Rémy Degenne, Dorian Baudry, Rianne de Heide, and Emilie Kaufmann, *Top two algorithms revisited*, Advances in Neural Information Processing Systems **35** (2022), 26791–26803.
- Marc Jourdan, Rémy Degenne, and Emilie Kaufmann, *An ϵ-best-arm identification algorithm for fixed-confidence and beyond*, Advances in Neural Information Processing Systems **36** (2023), 16578–16649.

References iii

- Emilie Kaufmann, Olivier Cappé, and Aurélien Garivier, *On the complexity of best-arm identification in multi-armed bandit models*, The Journal of Machine Learning Research **17** (2016), no. 1, 1–42.
- Tze Leung Lai, *Asymptotic optimality of invariant sequential probability ratio tests*, The Annals of Statistics (1981), 318–333.
- ———, *Information bounds and quick detection of parameter changes in stochastic systems*, IEEE Transactions on Information theory **44** (1998), no. 7, 2917–2929.
- Gary Lorden, *Procedures for reacting to a change in distribution*, The annals of mathematical statistics (1971), 1897–1908.
- Tze Leung Lai and Herbert Robbins, *Asymptotically efficient adaptive allocation rules*, Advances in applied mathematics **6** (1985), no. 1, 4–22.

References iv

- Alessio Russo and Alexandre Proutiere, *On the sample complexity of representation learning in multi-task bandits with global and local structure*, Proceedings of the AAAI Conference on Artificial Intelligence, vol. 37, 2023, pp. 9658–9667.
- Daniel Russo, *Simple bayesian algorithms for best arm identification*, Conference on Learning Theory, PMLR, 2016, pp. 1417–1418.
- Jérôme Taupin, Yassir Jedra, and Alexandre Proutiere, *Best policy identification in discounted linear mdps*, Sixteenth European Workshop on Reinforcement Learning, 2023.
- Abraham Wald, *Sequential analysis.*

Appendix

Proof of the false alarm rate

Consider the quickest change detection problem.

- Let $t = n_\alpha$, where $n_\alpha = T^\star(1-\delta)\ln(1/\alpha)$ with $\delta \in (0,1)$, and $\mathcal{E} = \{\tau - \nu \le n_\alpha\}$. Then $\mathcal{E} \in \mathcal{F}_{n_\alpha}$.
- Consider the following lemma (proof in the next page)

Lemma

Consider a stopping time τ and assume that $\mathbb{E}[\tau] \ge 1/\alpha$ for some $\alpha > 0$. Let k be an integer s.t. $k < 1/\alpha$. Then for some $\nu \ge 1$ we have $\mathbb{P}(\tau \ge \nu) > 0$ and $\mathbb{P}(\tau < \nu + k | \tau \ge \nu) \le k\alpha$.

- Use the previous lemma with $k = n_\alpha$. For α sufficiently small we have $T^\star(1-\delta) \le \ln(1/\alpha)$, hence $n_\alpha \le [\ln(1/\alpha)]^2$. Furthermore, for α small we also have $[\ln(1/\alpha)]^2 \le 1/\alpha$, satisfying the condition of the lemma.
- Therefore

$$\mathbb{P}_\infty(\mathcal{E} | \tau \ge \nu) \le n_\alpha \alpha \le [\ln(1/\alpha)]^2 \alpha.$$

Proof of the false alarm rate

Lemma

Consider a stopping time τ *and assume that* $\mathbb{E}[\tau] \geq 1/\alpha$ *for some* $\alpha > 0$. *Let* k *be an integer s.t.* $k < 1/\alpha$. *Then for some* $\nu \geq 1$ *we have* $\mathbb{P}(\tau \geq \nu) > 0$ *and* $\mathbb{P}(\tau < \nu + k | \tau \geq \nu) \leq k\alpha$.

Proof.

The first part is trivial, while the second one can be proved by contradiction. If that was not the case, then we would have $\mathbb{P}(\tau \geq \nu + k | \tau \geq \nu) < 1 - k\alpha$ for all $\nu \geq 1$. Hence

$$
\begin{aligned}
\mathbb{P}(\tau \geq \nu + 2k) &= \mathbb{P}(\tau \geq \nu + 2k | \tau \geq \nu + k)\mathbb{P}(\tau \geq \nu + k), \\
&< \mathbb{P}(\tau \geq \nu + 2k | \tau \geq \nu + k)(1 - k\alpha)\mathbb{P}(\tau \geq \nu), \\
&< (1 - k\alpha)^2 \mathbb{P}(\tau \geq \nu).
\end{aligned}
$$

Thus, by induction $\mathbb{P}(\tau \geq \nu + nk) < (1 - k\alpha)^n \mathbb{P}(\tau \geq \nu)$. Hence
$\mathbb{E}[\tau] = \sum_{i=1}^{\infty} \mathbb{P}(\tau \geq i) = \sum_{n=0}^{\infty} \sum_{j=1}^{k} \mathbb{P}(\tau \geq kn + j) < \sum_{n=0}^{\infty} \sum_{j=1}^{k} (1 - k\alpha)^n \mathbb{P}(\tau \geq j) \leq \frac{1}{\alpha}$,
which is a contradiction. □