Networks Seminar 2019

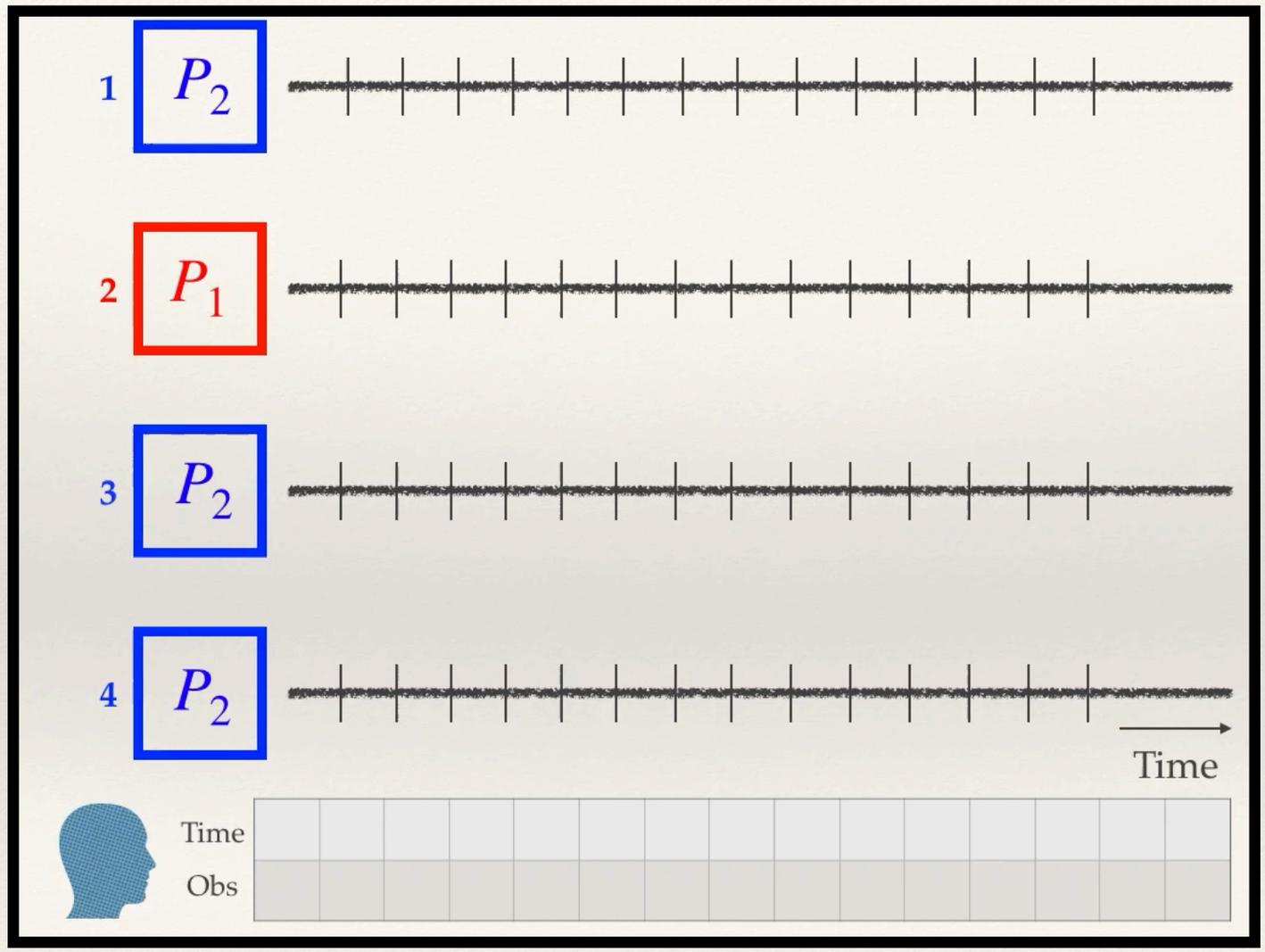
Nov 13 2019

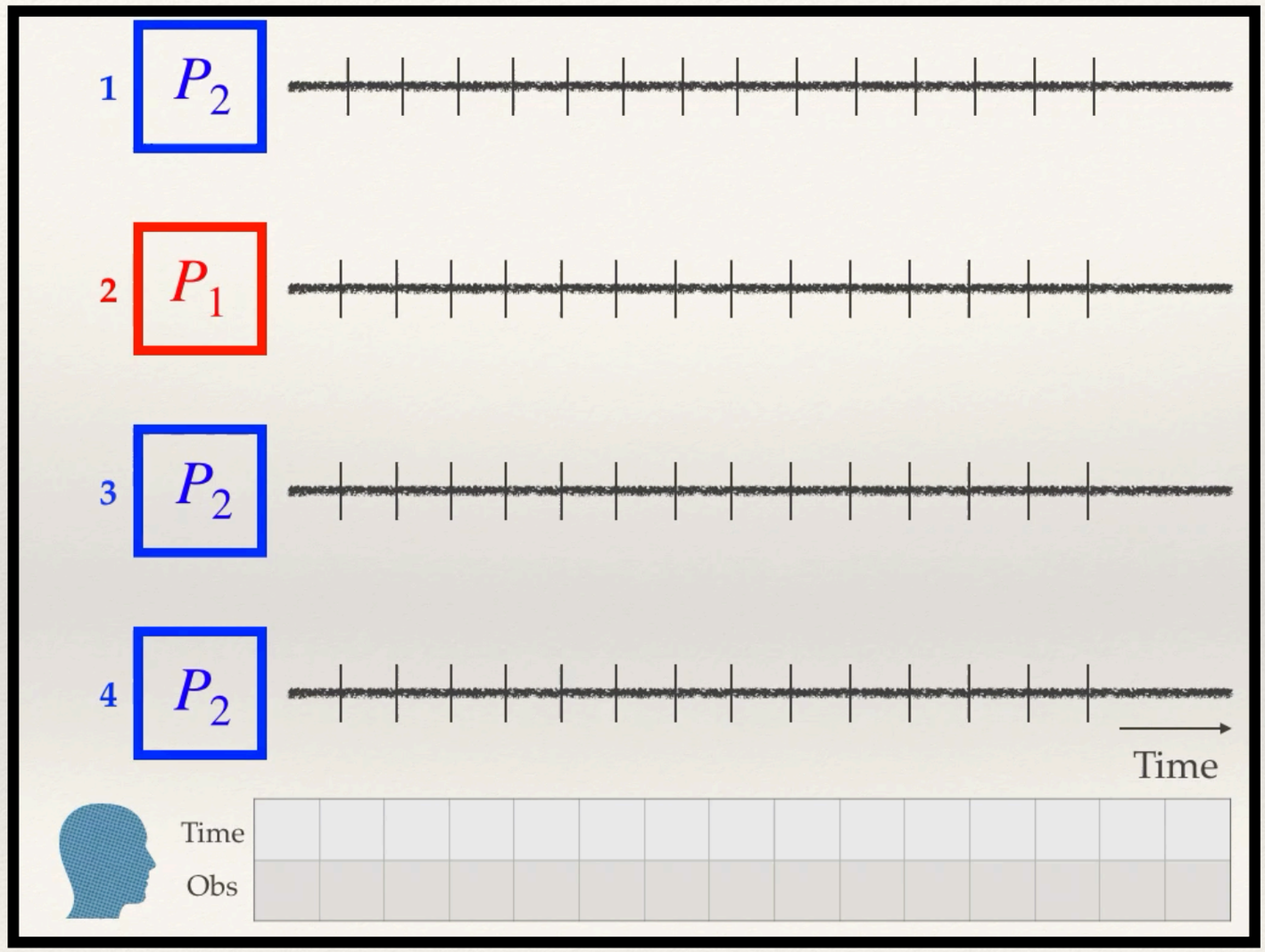
On Detecting An Anomalous Arm in Multi-armed Bandits with Markov Observations

Karthik PN

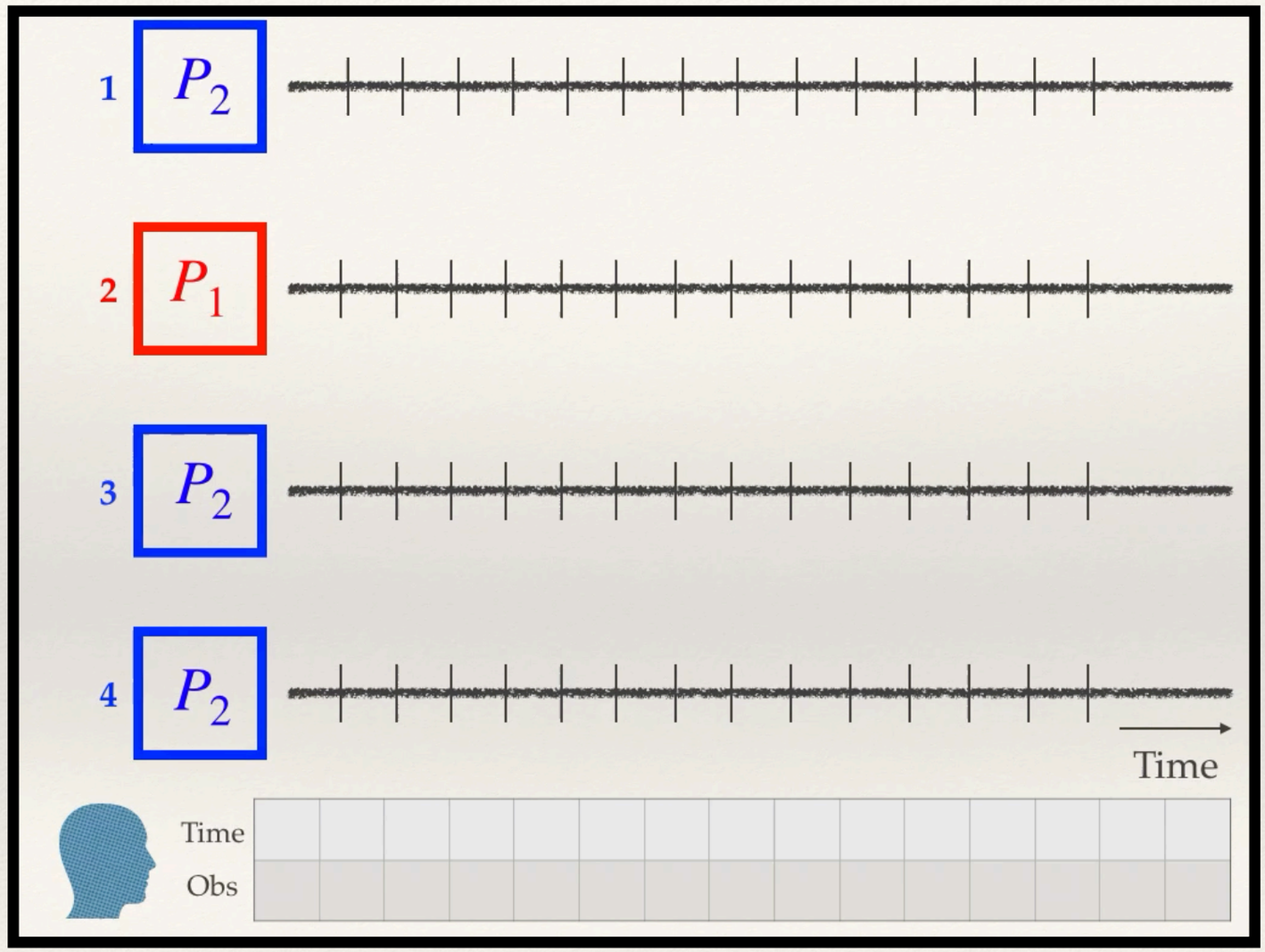
Joint Work with Prof. Rajesh Sundaresan

Presented in part at the 2019 IEEE International Symposium on Information Theory (ISIT)

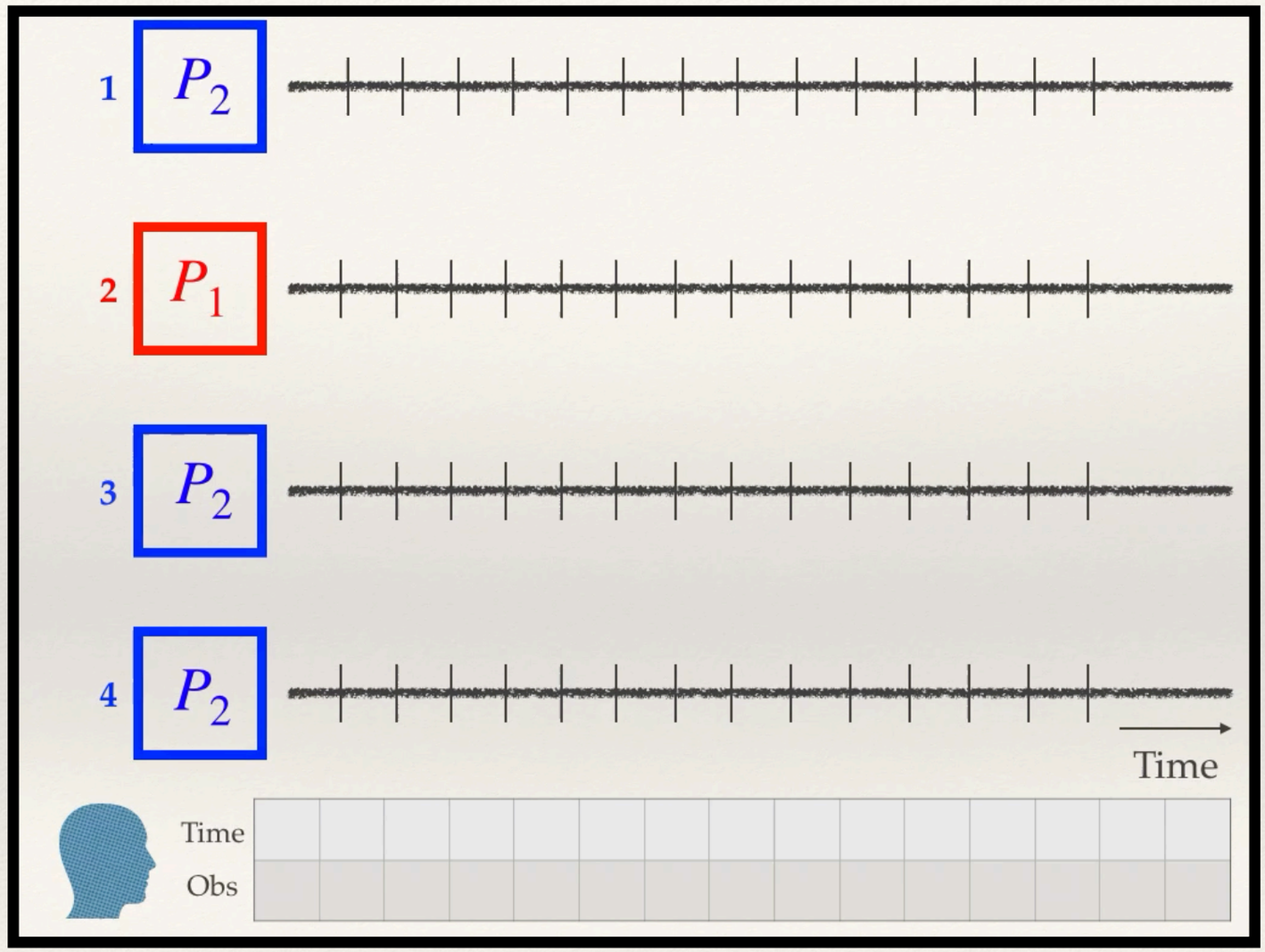

1
P_2
2
P_1
3
P_2
4
P_2
Time
Time
Obs


1
P_2
2
P_1
3
P_2
4
P_2
Time
Time
Obs

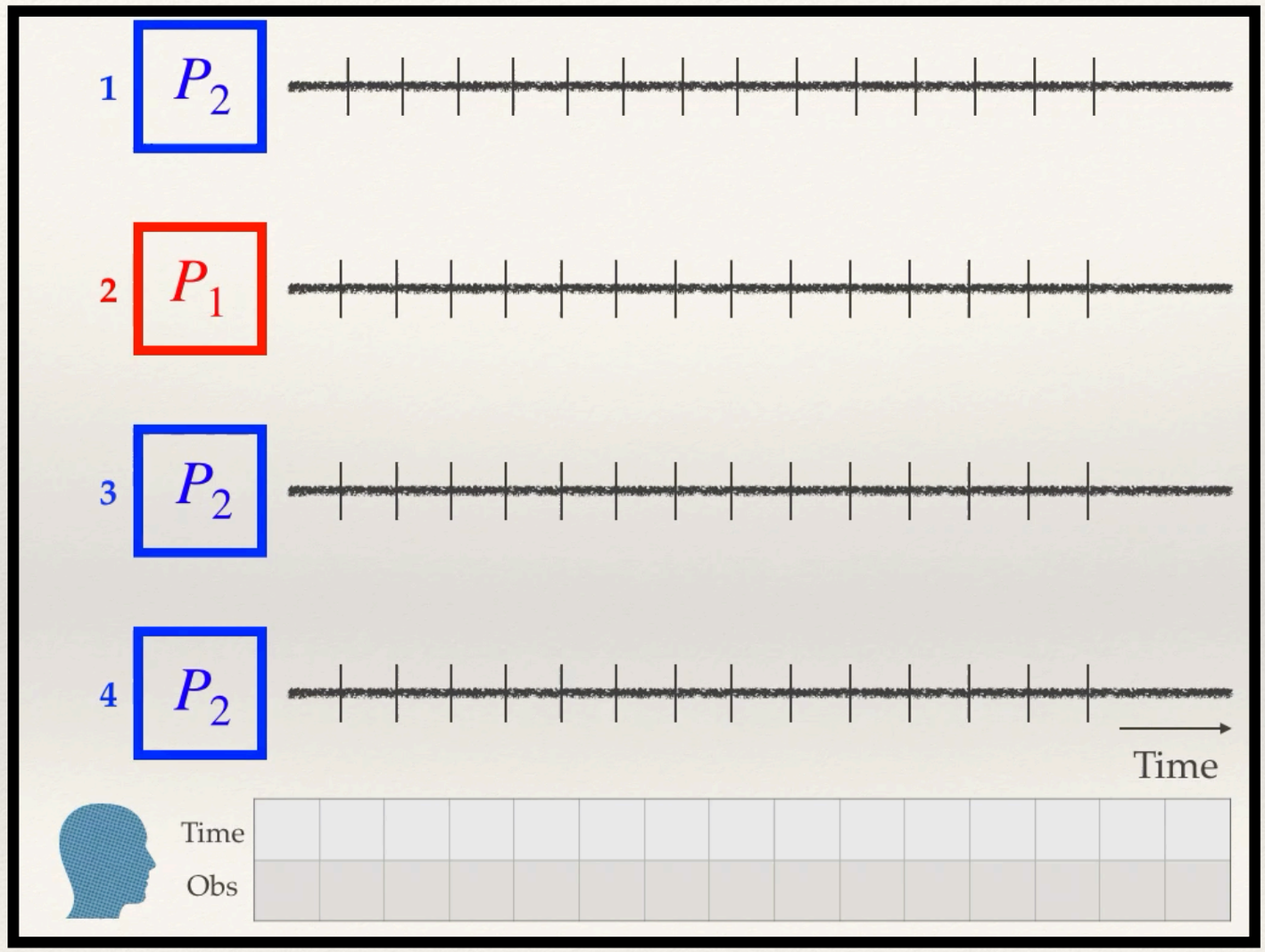
- A multi-armed bandit with K independent arms



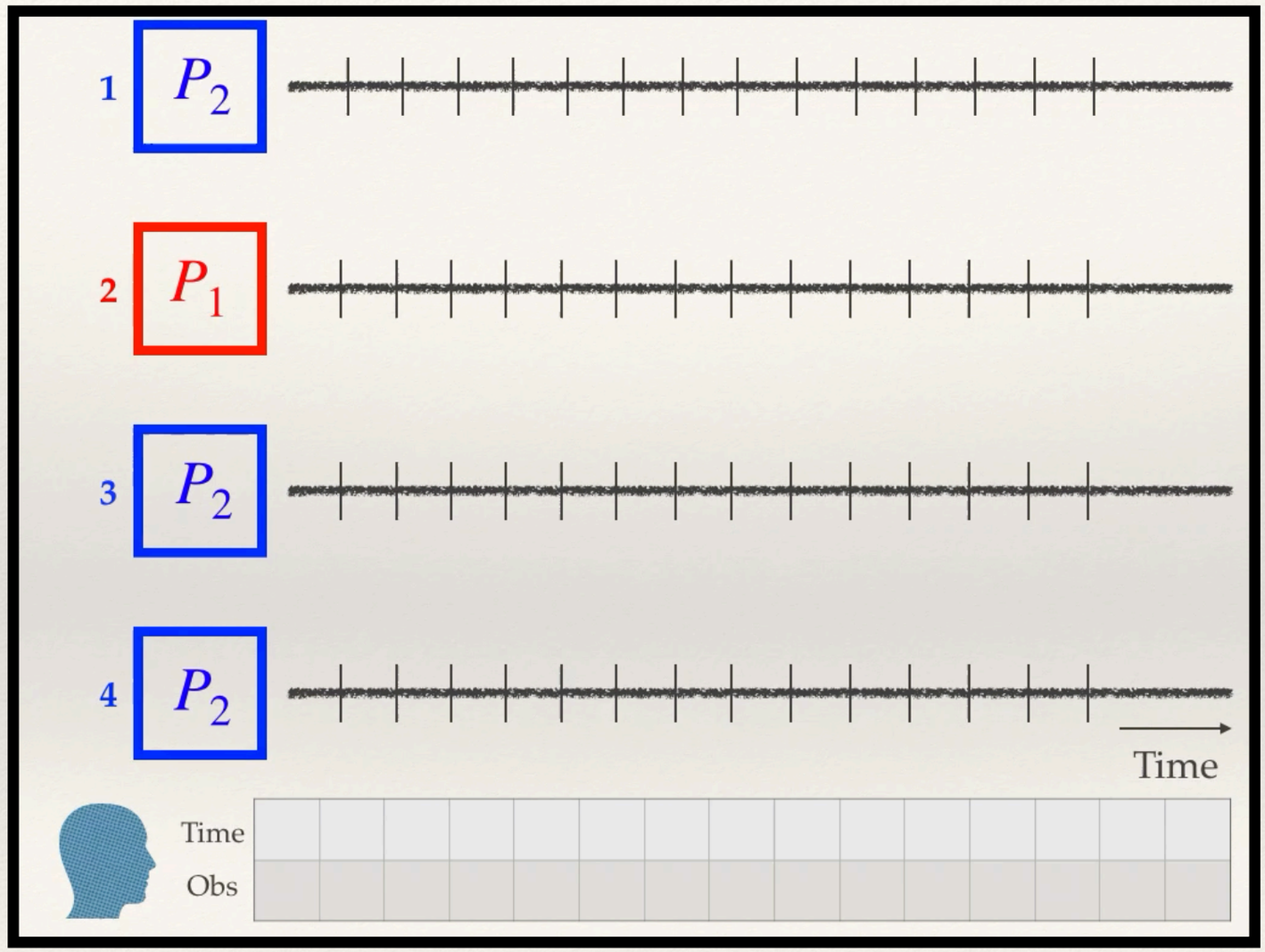
- A multi-armed bandit with K independent arms
- Each arm is a homogeneous and ergodic Markov chain on a finite state space



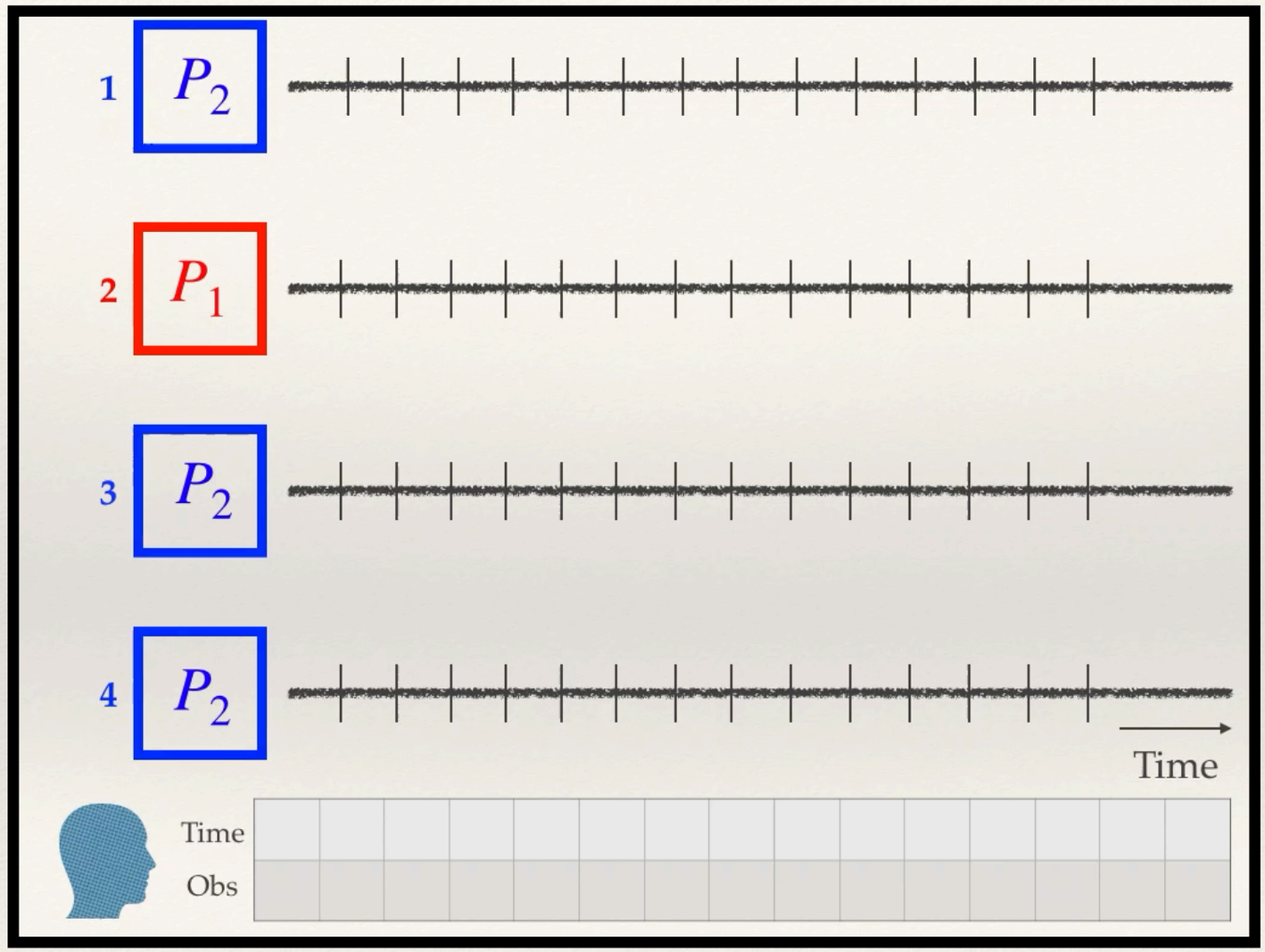
- A multi-armed bandit with K independent arms
- Each arm is a homogeneous and ergodic Markov chain on a finite state space
- State space common to all the arms



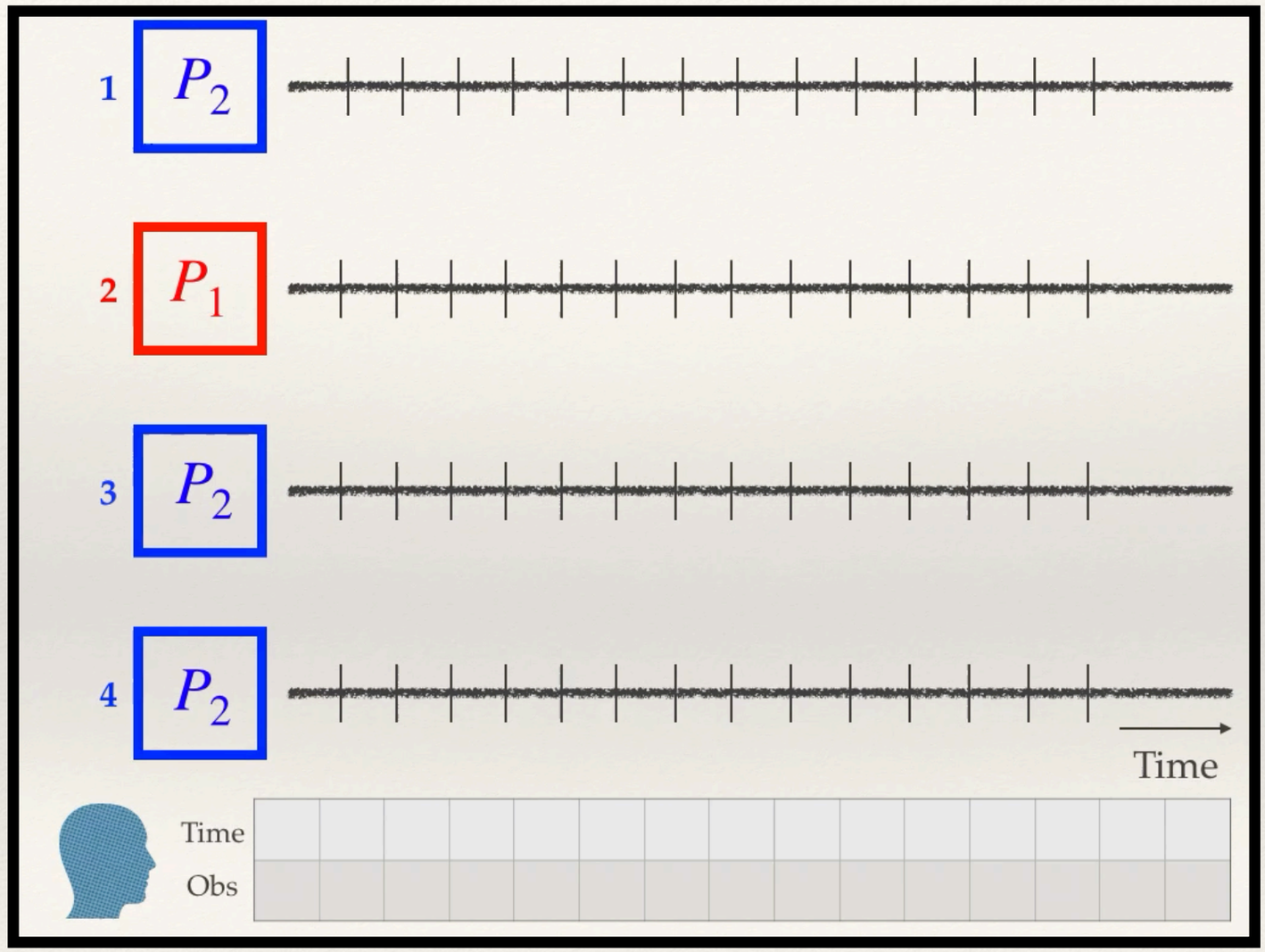
- A multi-armed bandit with K independent arms
- Each arm is a homogeneous and ergodic Markov chain on a finite state space
- State space common to all the arms
- One of the arms has TPM P_1, rest have TPM P_2, where $P_2 \neq P_1$



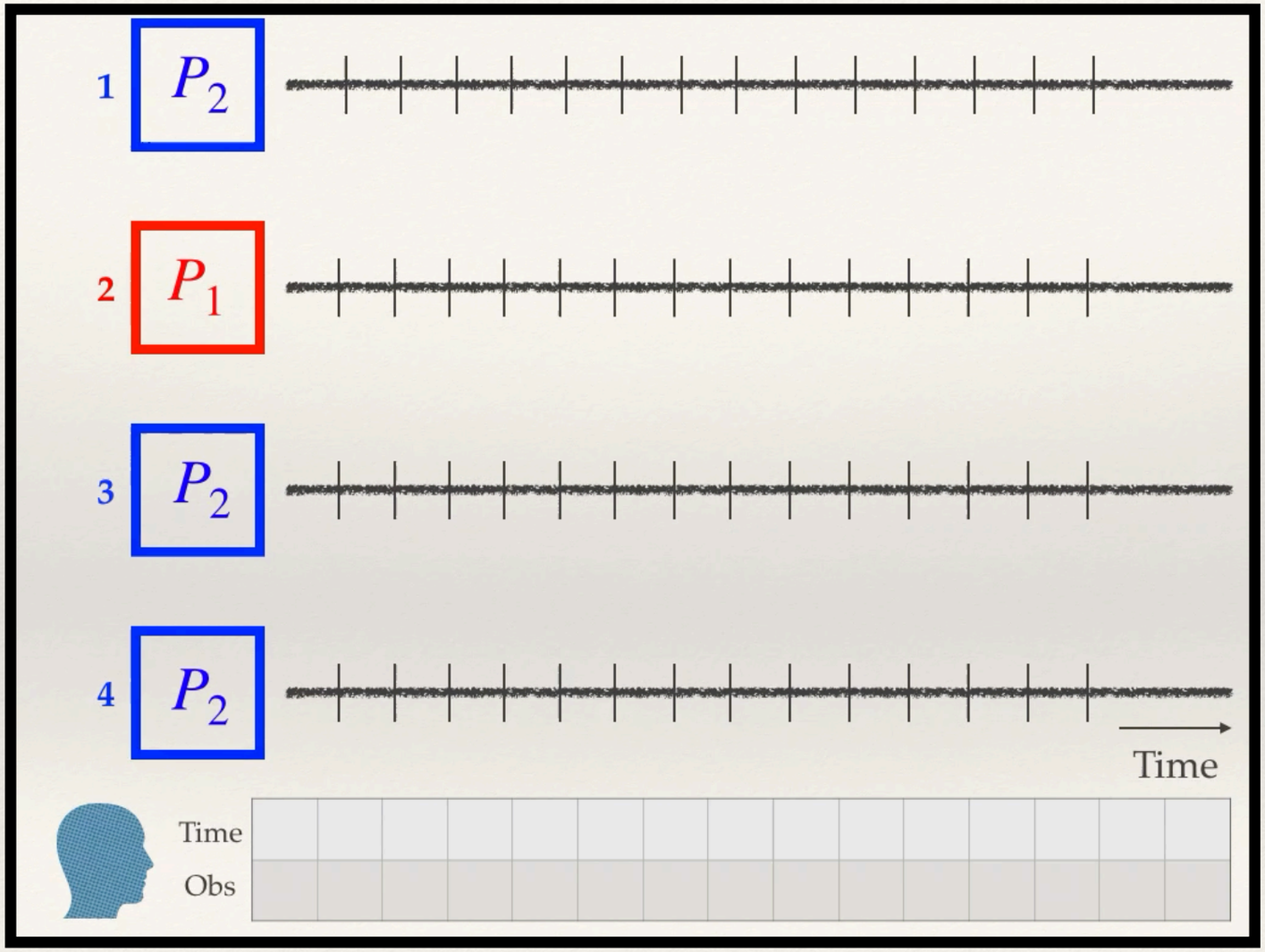
- A multi-armed bandit with K independent arms
- Each arm is a homogeneous and ergodic Markov chain on a finite state space
- State space common to all the arms
- One of the arms has TPM P_1, rest have TPM P_2, where $P_2 \neq P_1$
- You (learner) are given the task to figure out which arm has P_1 as quickly as possible under the **PAC framework**



- A multi-armed bandit with K independent arms
- Each arm is a homogeneous and ergodic Markov chain on a finite state space
- State space common to all the arms
- One of the arms has TPM P_1, rest have TPM P_2, where $P_2 \neq P_1$
- You (learner) are given the task to figure out which arm has P_1 as quickly as possible under the **PAC framework**
- Can select one arm at each time to observe (**sequential** selection)

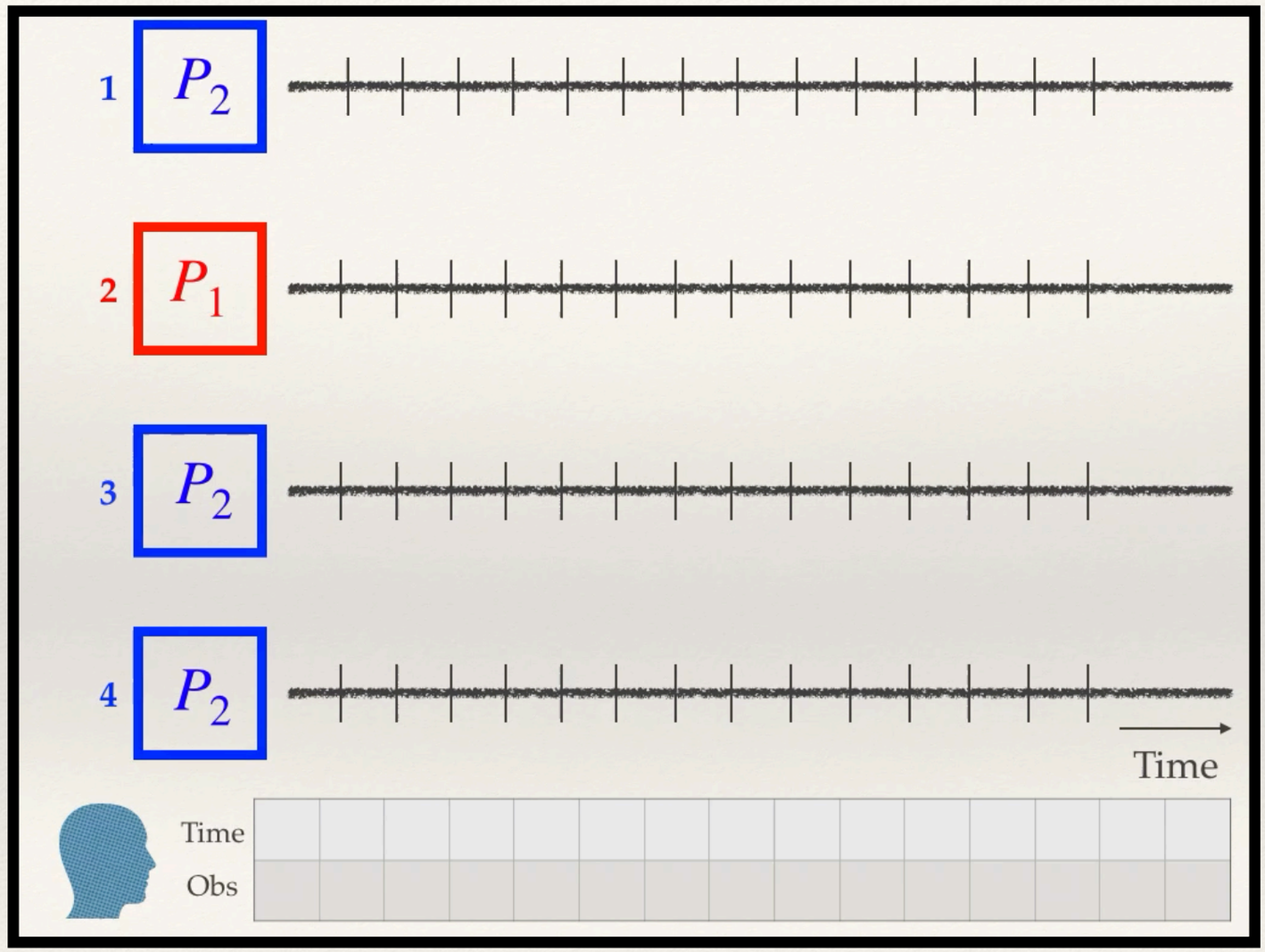


- A multi-armed bandit with K independent arms
- Each arm is a homogeneous and ergodic Markov chain on a finite state space
- State space common to all the arms
- One of the arms has TPM P_1, rest have TPM P_2, where $P_2 \neq P_1$
- You (learner) are given the task to figure out which arm has P_1 as quickly as possible under the **PAC framework**
- Can select one arm at each time to observe (**sequential** selection)
- Unobserved arms continue to evolve (**restless** arms)



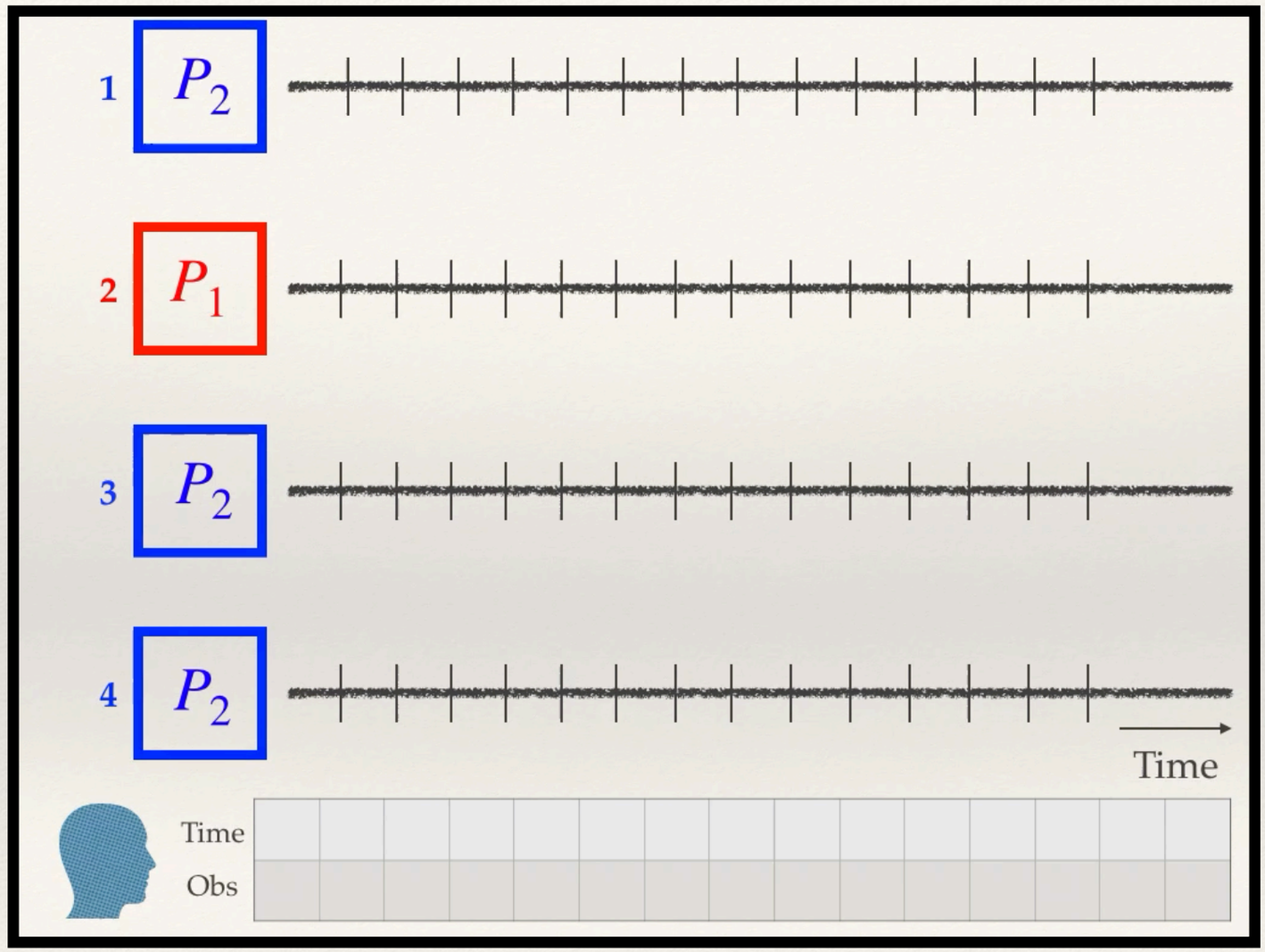
MAB with Restless Markov Arms

- A multi-armed bandit with K independent arms
- Each arm is a homogeneous and ergodic Markov chain on a finite state space
- State space common to all the arms
- One of the arms has TPM P_1, rest have TPM P_2, where $P_2 \neq P_1$
- You (learner) are given the task to figure out which arm has P_1 as quickly as possible under the **PAC framework**
- Can select one arm at each time to observe (**sequential** selection)
- Unobserved arms continue to evolve (**restless** arms)



MAB with Restless Markov Arms

- A multi-armed bandit with K independent arms
- Each arm is a homogeneous and ergodic Markov chain on a finite state space
- State space common to all the arms
- One of the arms has TPM P_1, rest have TPM P_2, where $P_2 \neq P_1$
- You (learner) are given the task to figure out which arm has P_1 as quickly as possible under the **PAC framework**
- Can select one arm at each time to observe (**sequential** selection)
- Unobserved arms continue to evolve (**restless** arms)



Odd arm identification (OAI) in multi-armed bandits with Markov Observations

MAB with Restless Arms: Examples in Case

- Q. Zhao and B. M. Sadler, “A survey of dynamic spectrum access,” IEEE Signal Process. Mag., vol. 24, no. 3, pp. 79–89, May 2007.
- R. Meshram, D. Manjunath, and Aditya Gopalan. "On the whittle index for restless multi-armed hidden Markov bandits." IEEE Transactions on Automatic Control 63.9 (2018): 3046-3053.

MAB with Restless Arms: Examples in Case

- Dynamic spectrum access in cognitive radio networks
  - Identifying the free channel as quickly as possible

- Q. Zhao and B. M. Sadler, “A survey of dynamic spectrum access,” IEEE Signal Process. Mag., vol. 24, no. 3, pp. 79–89, May 2007.
- R. Meshram, D. Manjunath, and Aditya Gopalan. "On the whittle index for restless multi-armed hidden Markov bandits." IEEE Transactions on Automatic Control 63.9 (2018): 3046-3053.

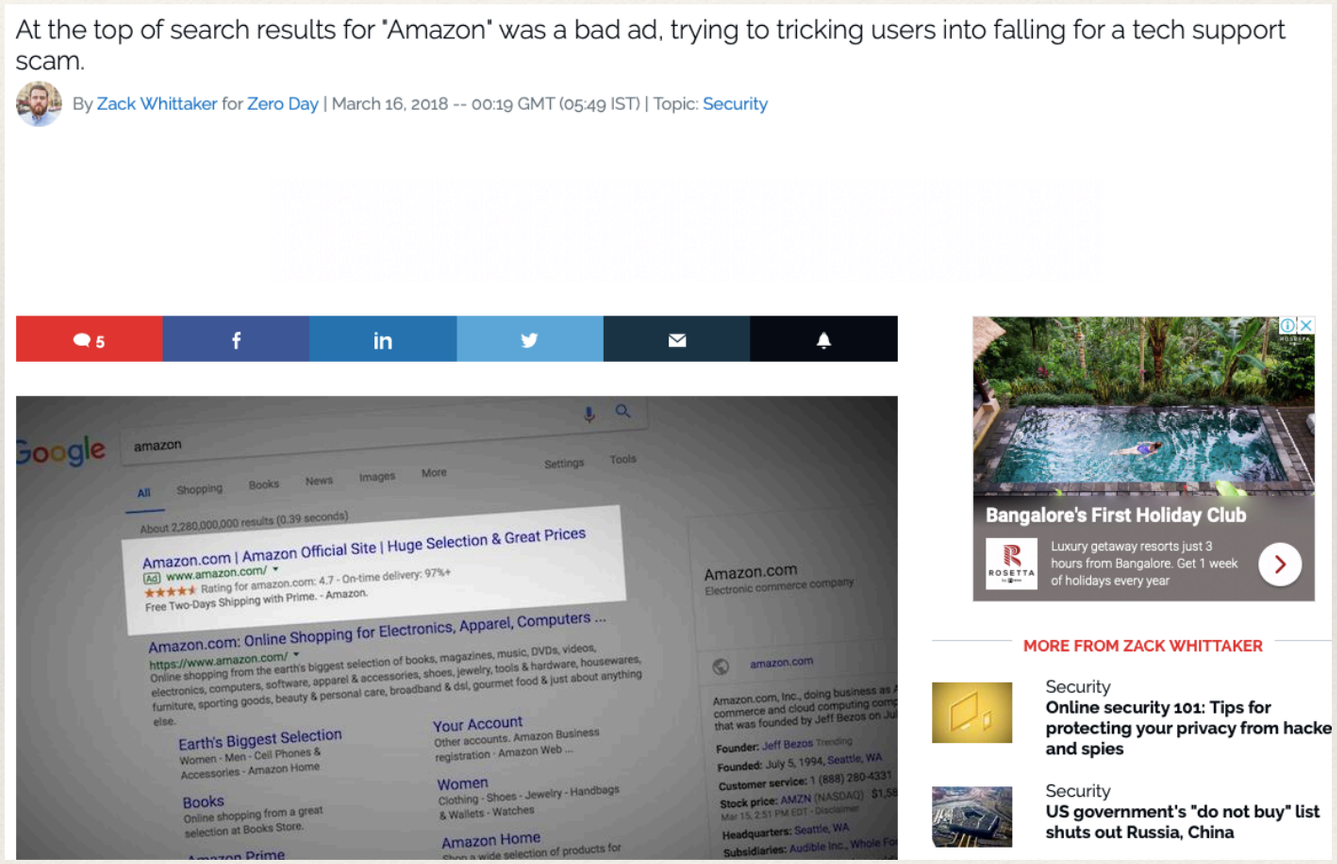
MAB with Restless Arms: Examples in Case

- Dynamic spectrum access in cognitive radio networks
  - Identifying the free channel as quickly as possible
- Ad placement system (APS) for a user in a web browsing session
  - Identifying a `bad' ad as quickly as possible

- Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," IEEE Signal Process. Mag., vol. 24, no. 3, pp. 79–89, May 2007.
- R. Meshram, D. Manjunath, and Aditya Gopalan. "On the whittle index for restless multi-armed hidden Markov bandits." IEEE Transactions on Automatic Control 63.9 (2018): 3046-3053.

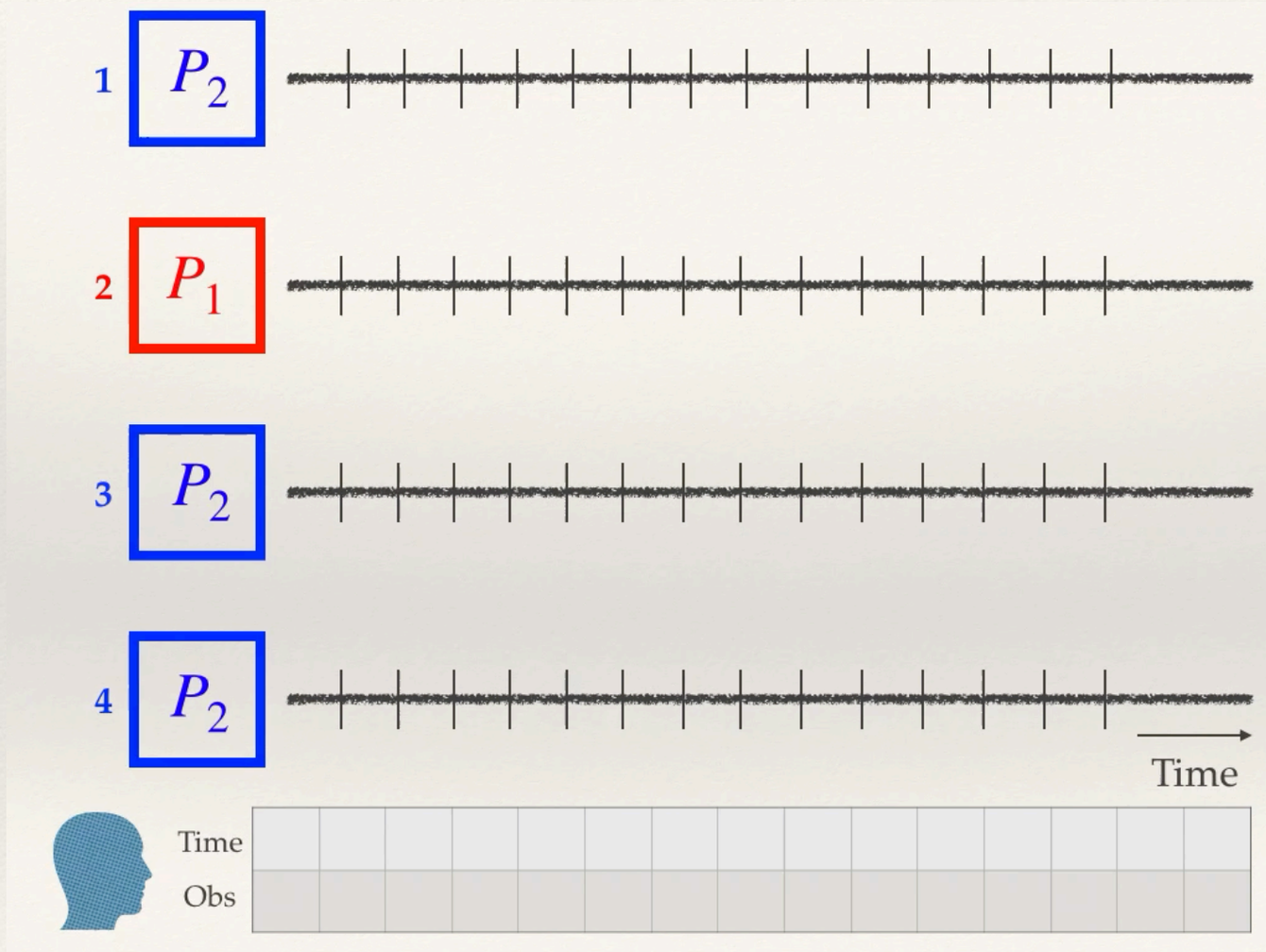
MAB with Restless Arms: Examples in Case

- Dynamic spectrum access in cognitive radio networks
  - Identifying the free channel as quickly as possible
- Ad placement system (APS) for a user in a web browsing session
  - Identifying a `bad' ad as quickly as possible

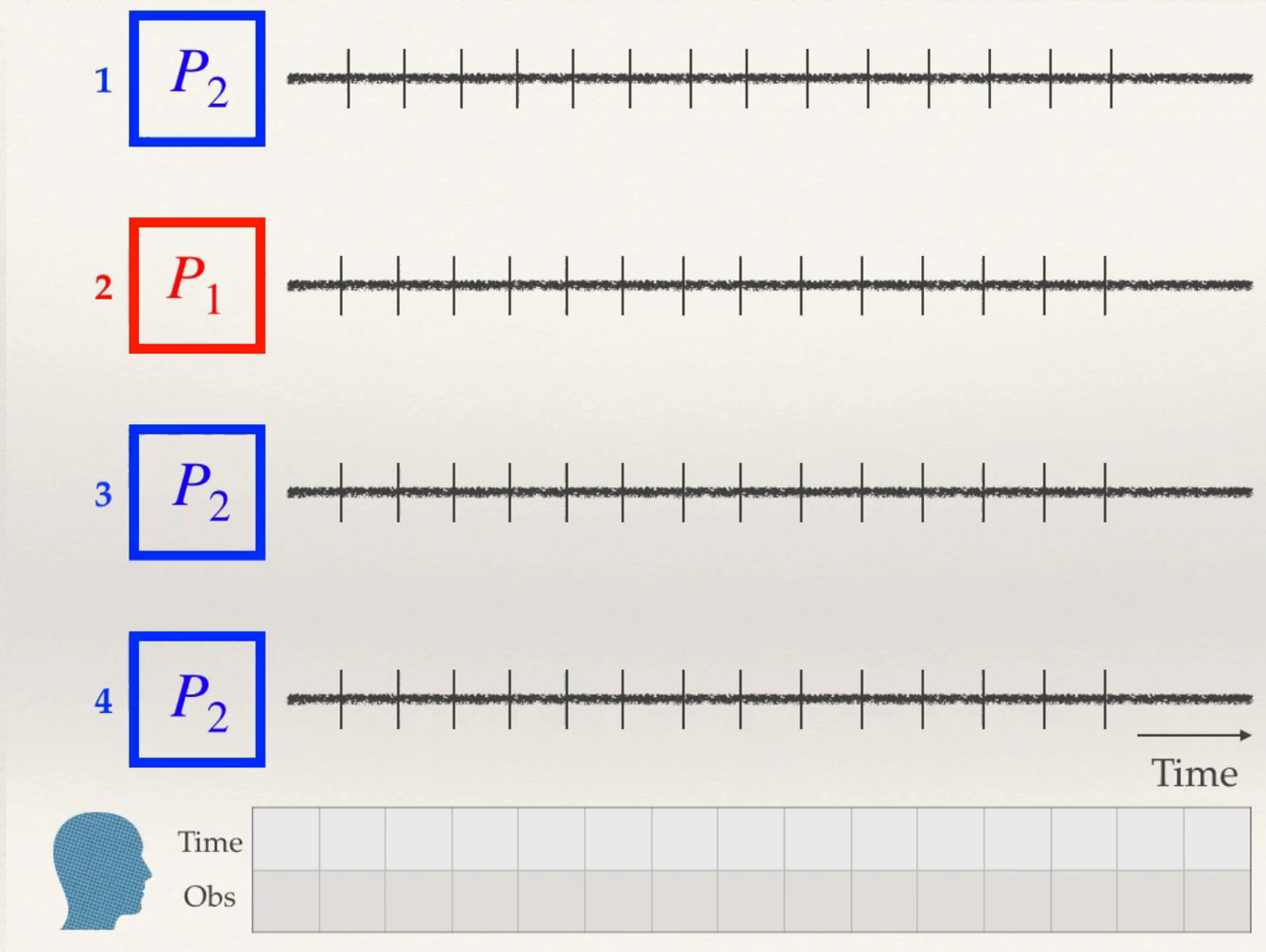


- Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," IEEE Signal Process. Mag., vol. 24, no. 3, pp. 79–89, May 2007.
- R. Meshram, D. Manjunath, and Aditya Gopalan. "On the whittle index for restless multi-armed hidden Markov bandits." IEEE Transactions on Automatic Control 63.9 (2018): 3046-3053.

A Simplification: MAB with Rested Arms

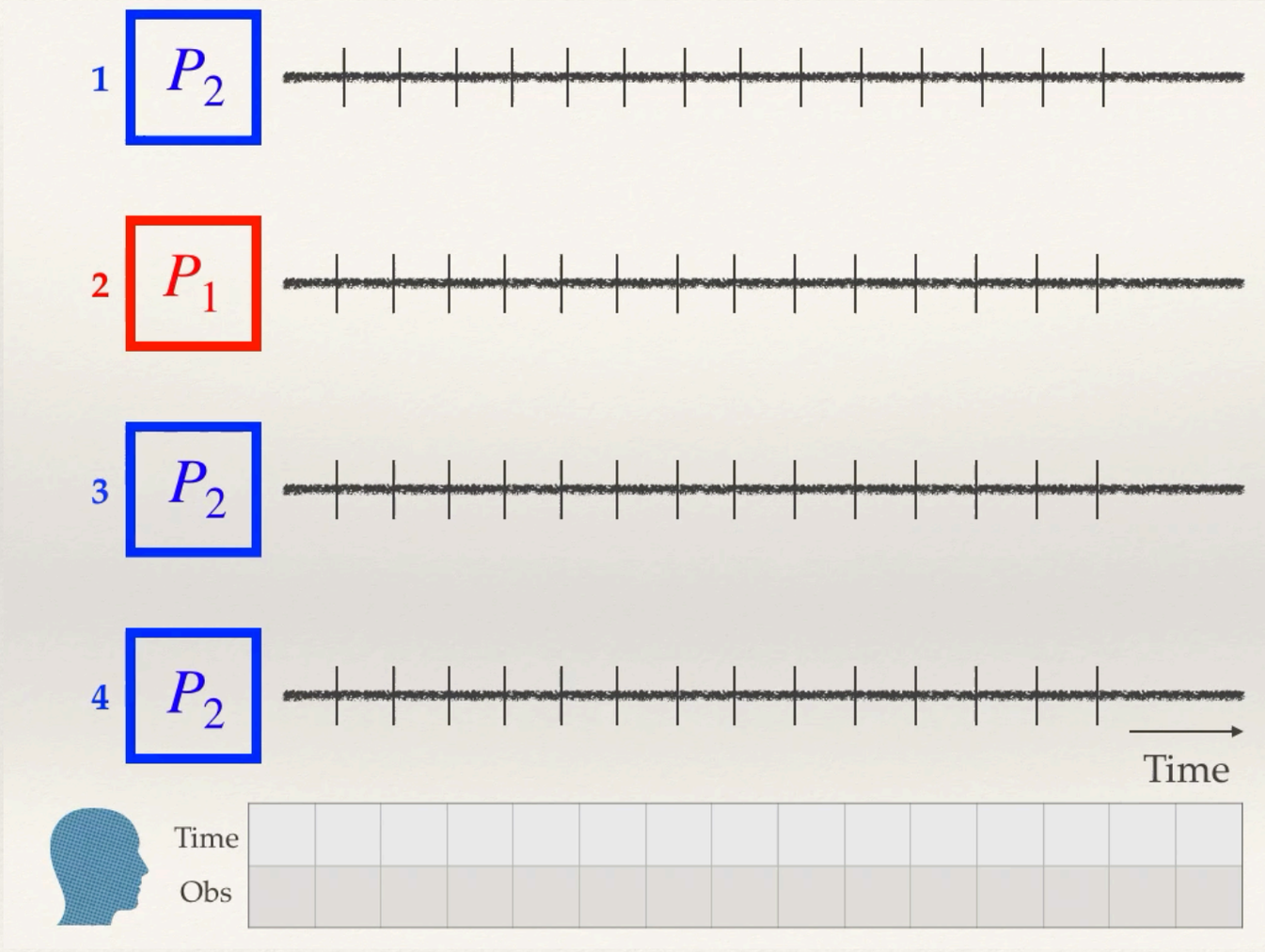


A Simplification: MAB with Rested Arms



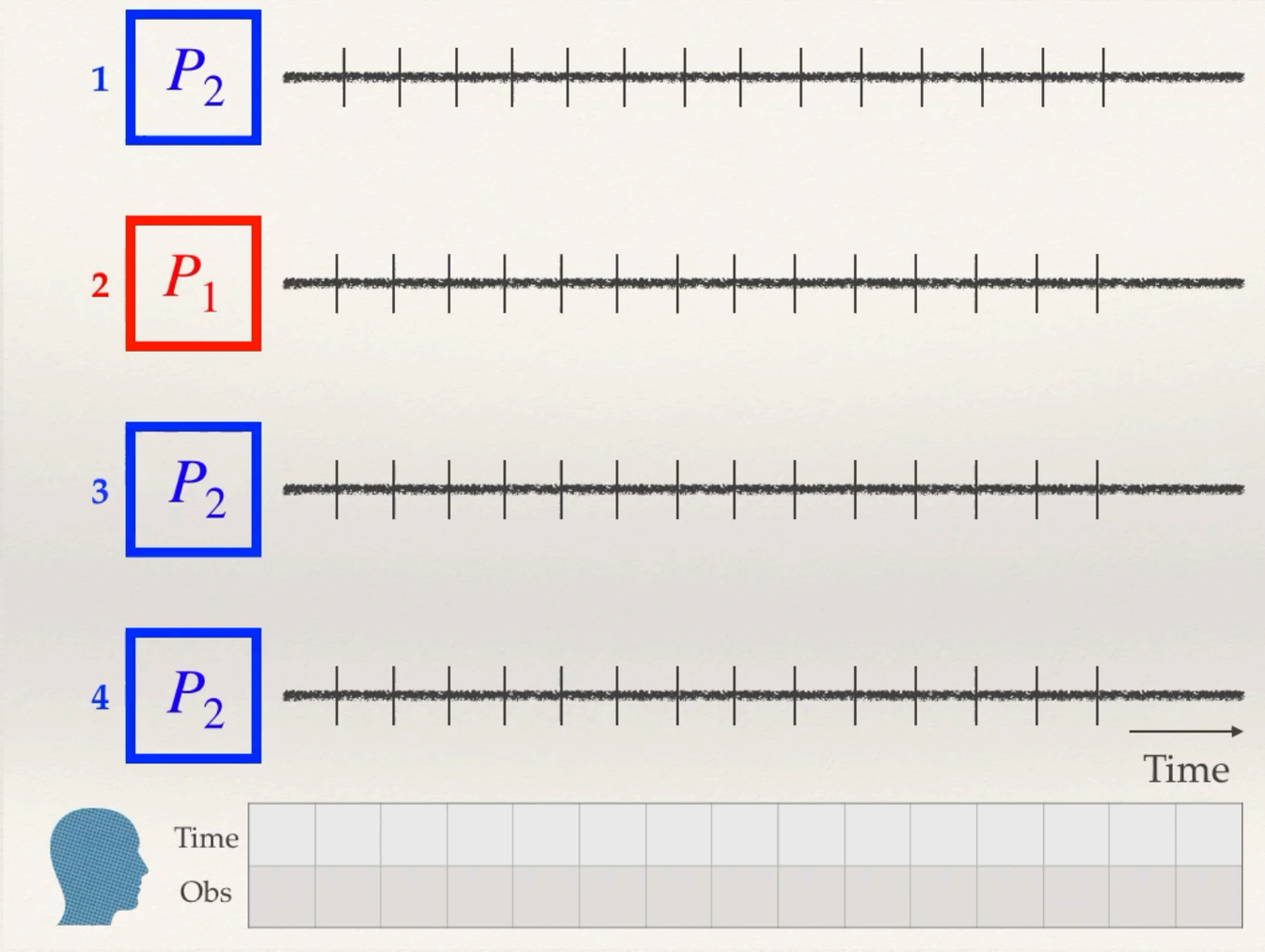
A Simplification: MAB with Rested Arms

- Unobserved arms remain frozen at their last observed states



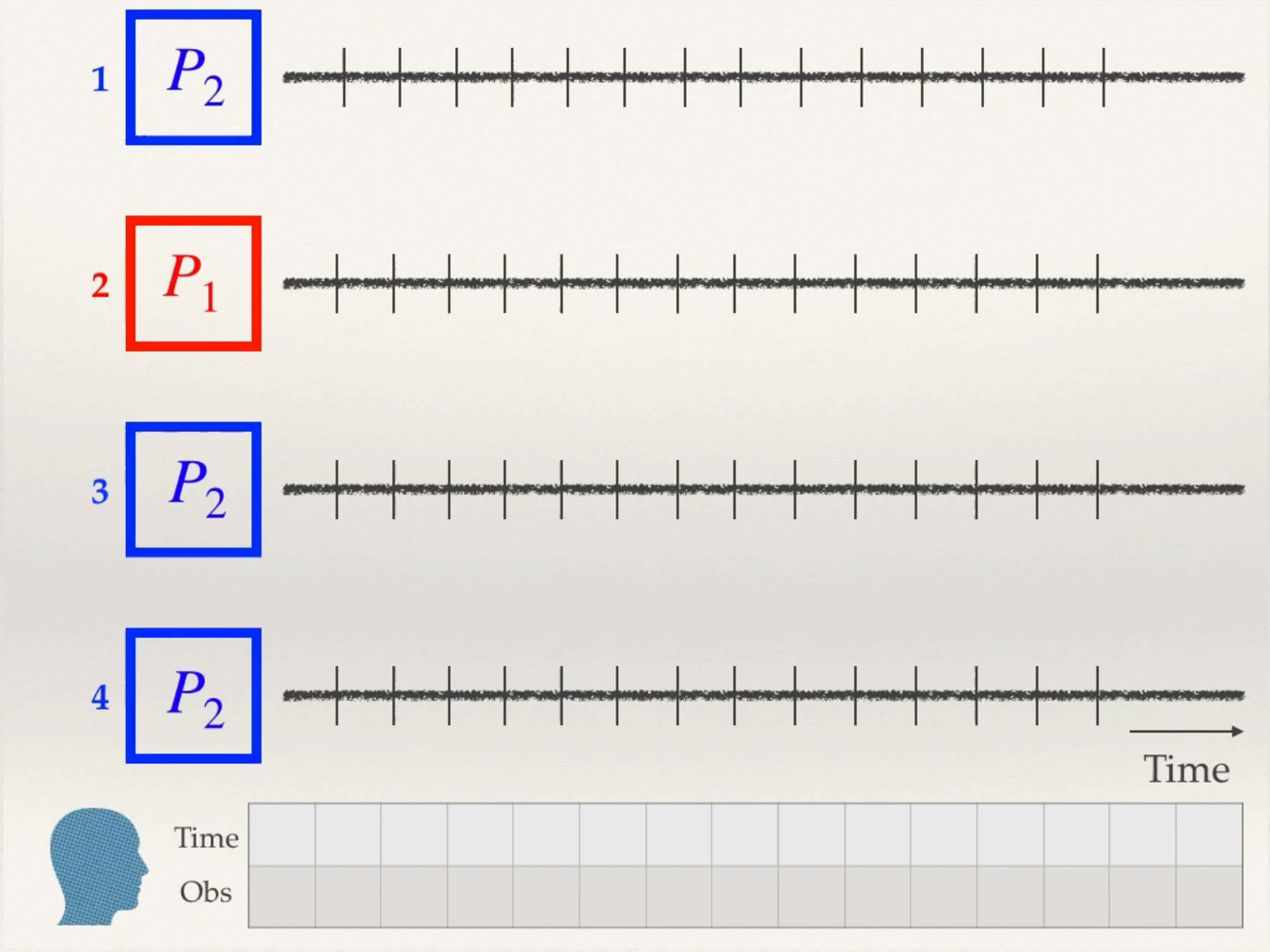
A Simplification: MAB with Rested Arms

- Unobserved arms remain frozen at their last observed states
- Simpler to analyse



A Simplification: MAB with Rested Arms

- Unobserved arms remain frozen at their last observed states
- Simpler to analyse
- Insights derived may serve as pointers to solve the more difficult case of restless arms



Prior Works on Rested Markov Bandits and OAI

		Problem Setting: Regret Minimisation	Problem Setting: Optimal Stopping	Nature of observations: IID	Nature of observations: Markov	Arm distributions: Known	Arm distributions: Unknown	
Rested arms	Gittins[1]	✓	✗	✗	✓	✓	✗	
	Agarwal et al.[2] Anantharam et al.[3]	✓	✗	✗	✓	✗	✓	Finite parameter space
OAI	Vaidhiyan et al.[4] Prabhu et al.[5]	✗	✓	✓	✗	✗	✓	Uncountable parameter space
	Current work	✗	✓	✗	✓	✗	✓	

[1] J. C. Gittins, "Bandit processes and dynamic allocation indices," Journal of the Royal Statistical Society. Series B (Methodological), pp. 148–177, 1979.

[2] R. Agrawal, D. Teneketzis, and V. Anantharam, "Asymptotically efficient adaptive allocation schemes for controlled Markov chains: Finite parameter space," IEEE Trans. on Automatic Control, 1989.

[3] Anantharam V, Varaiya P, Walrand J. Asymptotically efficient allocation rules for the multi-armed bandit problem with multiple plays-Part II: Markovian rewards. IEEE Trans. on Automatic Control, 1987.

[4] N. K. Vaidhiyan and R. Sundaresan, "Learning to detect an oddball target," IEEE Trans. on Information Theory, vol. 64, no. 2, pp. 831–852, 2018.

[5] G. R. Prabhu, S. Bhashyam, A. Gopalan, and R. Sundaresan, "Learning to detect an oddball target with observations from an exponential family," 2017.

Humble Beginnings: Wald and Chernoff

Humble Beginnings: Wald and Chernoff

- Sequential Probability Ratio Test (SPRT)

Humble Beginnings: Wald and Chernoff

- Sequential Probability Ratio Test (SPRT)
- Only one experiment (arm) to choose

Humble Beginnings: Wald and Chernoff

- Sequential Probability Ratio Test (SPRT)
- Only one experiment (arm) to choose
- Optimal in the Bayesian framework

Humble Beginnings: Wald and Chernoff

- Sequential Probability Ratio Test (SPRT)
- Only one experiment (arm) to choose
- Optimal in the Bayesian framework

- Procedure-A

Humble Beginnings: Wald and Chernoff

- Sequential Probability Ratio Test (SPRT)
- Only one experiment (arm) to choose
- Optimal in the Bayesian framework

- Procedure-A
- Multiple experiments (arms) to choose (active sequential hypothesis testing (ASHT))

Humble Beginnings: Wald and Chernoff

- Sequential Probability Ratio Test (SPRT)
- Only one experiment (arm) to choose
- Optimal in the Bayesian framework

- Procedure-A
- Multiple experiments (arms) to choose (active sequential hypothesis testing (ASHT))
- Asymptotically optimal

Humble Beginnings: Wald and Chernoff

- Sequential Probability Ratio Test (SPRT)
- Only one experiment (arm) to choose
- Optimal in the Bayesian framework

- Procedure-A
- Multiple experiments (arms) to choose (active sequential hypothesis testing (ASHT))
- Asymptotically optimal

All's well that starts well

Wald's SPRT

Wald's SPRT

- IID observations $X_1, X_2, \ldots \sim P_\theta$

Wald's SPRT

- IID observations $X_1, X_2, \ldots \sim P_\theta$
- To test

$$\mathcal{H}_0 : \quad \theta = \theta_0$$

vs

$$\mathcal{H}_1 : \quad \theta = \theta_1$$

Wald's SPRT

- IID observations $X_1, X_2, \ldots \sim P_\theta$
- To test

$$\mathcal{H}_0 : \quad \theta = \theta_0$$

vs

$$\mathcal{H}_1 : \quad \theta = \theta_1$$

- P_{θ_0} and P_{θ_1} are known

Wald's SPRT

- IID observations $X_1, X_2, \ldots \sim P_\theta$
- To test

$$\mathcal{H}_0 : \quad \theta = \theta_0$$
$$\text{vs}$$
$$\mathcal{H}_1 : \quad \theta = \theta_1$$

- P_{θ_0} and P_{θ_1} are known
- Priors:

$$P(\mathcal{H}_0) = w, \quad P(\mathcal{H}_1) = 1 - w$$

Wald's SPRT

- IID observations $X_1, X_2, \ldots \sim P_\theta$
- To test

$$\mathcal{H}_0 : \quad \theta = \theta_0$$
$$\text{vs}$$
$$\mathcal{H}_1 : \quad \theta = \theta_1$$

- P_{θ_0} and P_{θ_1} are known
- Priors:

$$P(\mathcal{H}_0) = w, \quad P(\mathcal{H}_1) = 1 - w$$

- Cost c per round of experiment

Wald's SPRT

- IID observations $X_1, X_2, \ldots \sim P_\theta$
- To test

$$\mathcal{H}_0 : \quad \theta = \theta_0$$
$$\text{vs}$$
$$\mathcal{H}_1 : \quad \theta = \theta_1$$

- P_{θ_0} and P_{θ_1} are known
- Priors:

$$P(\mathcal{H}_0) = w, \quad P(\mathcal{H}_1) = 1 - w$$

- Cost c per round of experiment
- At time n:
  - Decide to stop and declare which hypothesis is true
  - Decide to perform the experiment again (take one more observation)

Wald's SPRT

- IID observations $X_1, X_2, \ldots \sim P_\theta$
- To test

$$\mathcal{H}_0 : \quad \theta = \theta_0$$

vs

$$\mathcal{H}_1 : \quad \theta = \theta_1$$

- P_{θ_0} and P_{θ_1} are known
- Priors:

$$P(\mathcal{H}_0) = w, \quad P(\mathcal{H}_1) = 1 - w$$

- Cost c per round of experiment
- At time n:
  - Decide to stop and declare which hypothesis is true
  - Decide to perform the experiment again (take one more observation)
- Average risks at stopping time N:

$$R_0 = P_{FA} + c\, E[N \mid \mathcal{H}_0], \qquad R_1 = P_{MD} + c\, E[N \mid \mathcal{H}_1]$$

Wald's SPRT

- IID observations $X_1, X_2, \ldots \sim P_\theta$
- To test

$$\mathcal{H}_0 : \quad \theta = \theta_0$$
$$\text{vs}$$
$$\mathcal{H}_1 : \quad \theta = \theta_1$$

- P_{θ_0} and P_{θ_1} are known
- Priors:

$$P(\mathcal{H}_0) = w, \quad P(\mathcal{H}_1) = 1 - w$$

- Cost c per round of experiment
- At time n:
  - Decide to stop and declare which hypothesis is true
  - Decide to perform the experiment again (take one more observation)
- Average risks at stopping time N:

$$R_0 = P_{FA} + c\, E[N \mid \mathcal{H}_0], \qquad R_1 = P_{MD} + c\, E[N \mid \mathcal{H}_1]$$

- Goal:

  Minimize $\quad wR_0 + (1 - w)R_1$

Wald's SPRT

- IID observations $X_1, X_2, \ldots \sim P_\theta$
- To test

$$\mathcal{H}_0 : \quad \theta = \theta_0$$
$$\text{vs}$$
$$\mathcal{H}_1 : \quad \theta = \theta_1$$

- P_{θ_0} and P_{θ_1} are known
- Priors:

$$P(\mathcal{H}_0) = w, \quad P(\mathcal{H}_1) = 1 - w$$

- Cost c per round of experiment
- At time n:
  - Decide to stop and declare which hypothesis is true
  - Decide to perform the experiment again (take one more observation)
- Average risks at stopping time N:

$$R_0 = P_{FA} + c\, E[N \mid \mathcal{H}_0], \qquad R_1 = P_{MD} + c\, E[N \mid \mathcal{H}_1]$$

- Goal:

  Minimize $\quad wR_0 + (1 - w)R_1$

- Wald's idea: Construct the test statistic

$$S_n = \log \frac{P_{\theta_1}(X_1, \ldots, X_n)}{P_{\theta_0}(X_1, \ldots . X_n)}$$

Wald's SPRT

- IID observations $X_1, X_2, \ldots \sim P_\theta$
- To test

$$\mathcal{H}_0 : \quad \theta = \theta_0$$
$$\text{vs}$$
$$\mathcal{H}_1 : \quad \theta = \theta_1$$

- P_{θ_0} and P_{θ_1} are known
- Priors:

$$P(\mathcal{H}_0) = w, \quad P(\mathcal{H}_1) = 1 - w$$

- Cost c per round of experiment
- At time n:
  - Decide to stop and declare which hypothesis is true
  - Decide to perform the experiment again (take one more observation)
- Average risks at stopping time N:

$$R_0 = P_{FA} + c\, E[N \mid \mathcal{H}_0], \qquad R_1 = P_{MD} + c\, E[N \mid \mathcal{H}_1]$$

- Goal:

  Minimize $\quad wR_0 + (1 - w)R_1$

- Wald's idea: Construct the test statistic

$$S_n = \log \frac{P_{\theta_1}(X_1, \ldots, X_n)}{P_{\theta_0}(X_1, \ldots . X_n)}$$

- $E[S_n \mid \mathcal{H}_1] = nD(P_{\theta_1} \| P_{\theta_0}), \quad E[S_n \mid \mathcal{H}_0] = -\, nD(P_{\theta_0} \| P_{\theta_1})$

Wald's SPRT

- IID observations $X_1, X_2, \ldots \sim P_\theta$
- To test

$$\mathcal{H}_0 : \quad \theta = \theta_0$$
$$\text{vs}$$
$$\mathcal{H}_1 : \quad \theta = \theta_1$$

- P_{θ_0} and P_{θ_1} are known
- Priors:

$$P(\mathcal{H}_0) = w, \quad P(\mathcal{H}_1) = 1 - w$$

- Cost c per round of experiment
- At time n:
  - Decide to stop and declare which hypothesis is true
  - Decide to perform the experiment again (take one more observation)
- Average risks at stopping time N:

$$R_0 = P_{FA} + c\, E[N \mid \mathcal{H}_0], \qquad R_1 = P_{MD} + c\, E[N \mid \mathcal{H}_1]$$

- Goal:

  Minimize $\quad wR_0 + (1 - w)R_1$

- Wald's idea: Construct the test statistic

$$S_n = \log \frac{P_{\theta_1}(X_1, \ldots, X_n)}{P_{\theta_0}(X_1, \ldots . X_n)}$$

- $E[S_n \mid \mathcal{H}_1] = nD(P_{\theta_1} \| P_{\theta_0}), \quad E[S_n \mid \mathcal{H}_0] = -\, nD(P_{\theta_0} \| P_{\theta_1})$
- By the law of large numbers, we expect the following behaviour:

Wald's SPRT

- IID observations $X_1, X_2, \ldots \sim P_\theta$
- To test

$$\mathcal{H}_0 : \quad \theta = \theta_0$$
$$\text{vs}$$
$$\mathcal{H}_1 : \quad \theta = \theta_1$$

- P_{θ_0} and P_{θ_1} are known
- Priors:

$$P(\mathcal{H}_0) = w, \quad P(\mathcal{H}_1) = 1 - w$$

- Cost c per round of experiment
- At time n:
  - Decide to stop and declare which hypothesis is true
  - Decide to perform the experiment again (take one more observation)
- Average risks at stopping time N:

$$R_0 = P_{FA} + c\,E[N \mid \mathcal{H}_0], \qquad R_1 = P_{MD} + c\,E[N \mid \mathcal{H}_1]$$

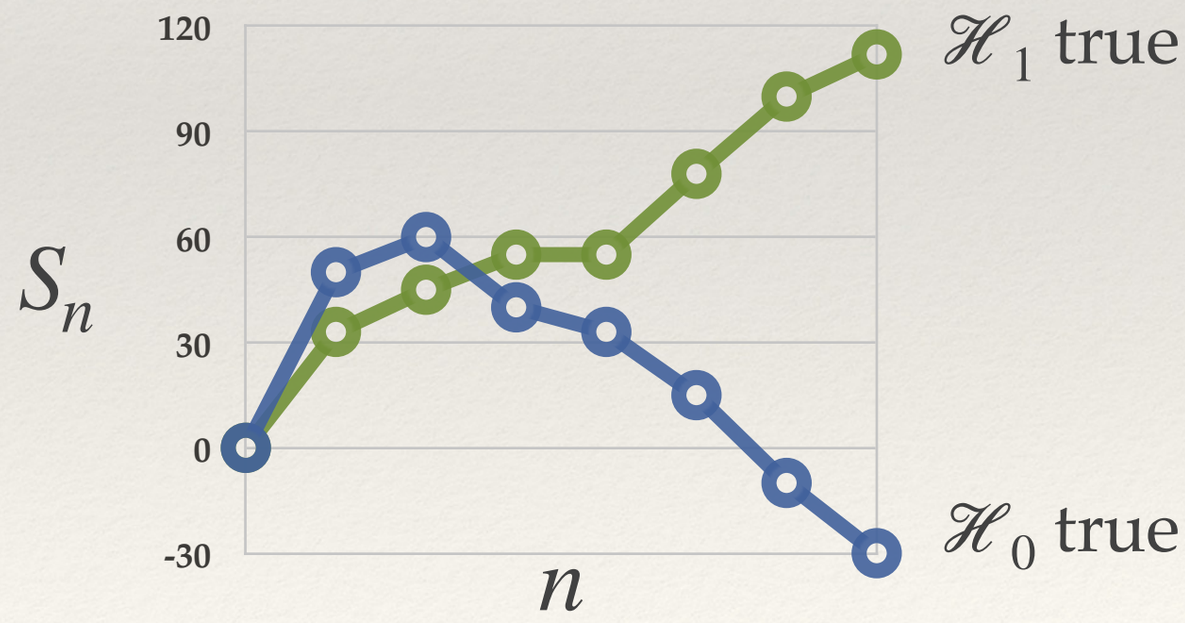
- Goal:

  Minimize $\quad wR_0 + (1-w)R_1$

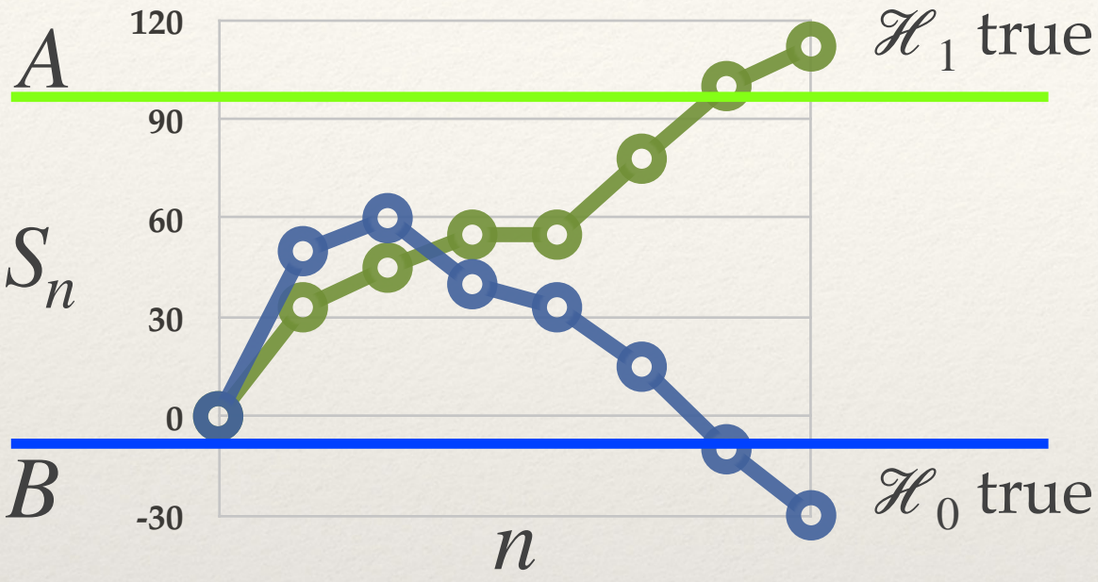
- Wald's idea: Construct the test statistic

$$S_n = \log \frac{P_{\theta_1}(X_1, \ldots, X_n)}{P_{\theta_0}(X_1, \ldots . X_n)}$$

- $E[S_n \mid \mathcal{H}_1] = nD(P_{\theta_1} \| P_{\theta_0}), \quad E[S_n \mid \mathcal{H}_0] = -nD(P_{\theta_0} \| P_{\theta_1})$
- By the law of large numbers, we expect the following behaviour:



Wald's SPRT

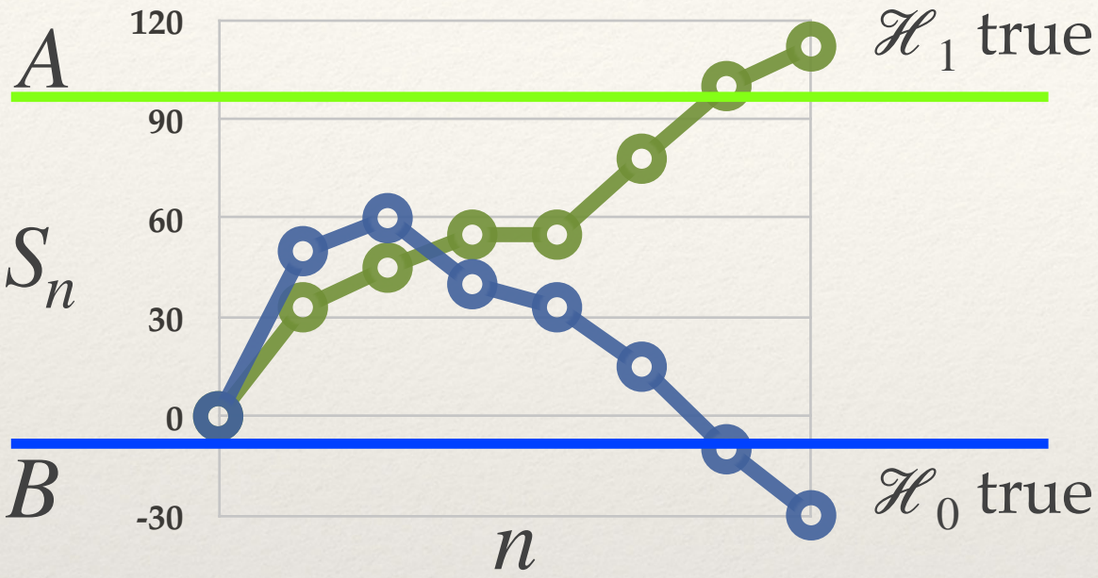


- Wald's idea: Construct the test statistic

$$S_n = \log \frac{P_{\theta_1}(X_1, \ldots, X_n)}{P_{\theta_0}(X_1, \ldots, X_n)}$$

- $E[S_n \mid \mathcal{H}_1] = nD(P_{\theta_1} \| P_{\theta_0}), \quad E[S_n \mid \mathcal{H}_0] = -nD(P_{\theta_0} \| P_{\theta_1})$
- Set two thresholds: $A > 0$ and $B < 0$ (these depend on the priors)
- At time n:
  - Stop and declare $\mathcal{H}_1$ true if $S_n \geq A$
  - Stop and declare $\mathcal{H}_0$ true if $S_n \leq B$
  - Continue taking observations if $B < S_n < A$

Wald's SPRT



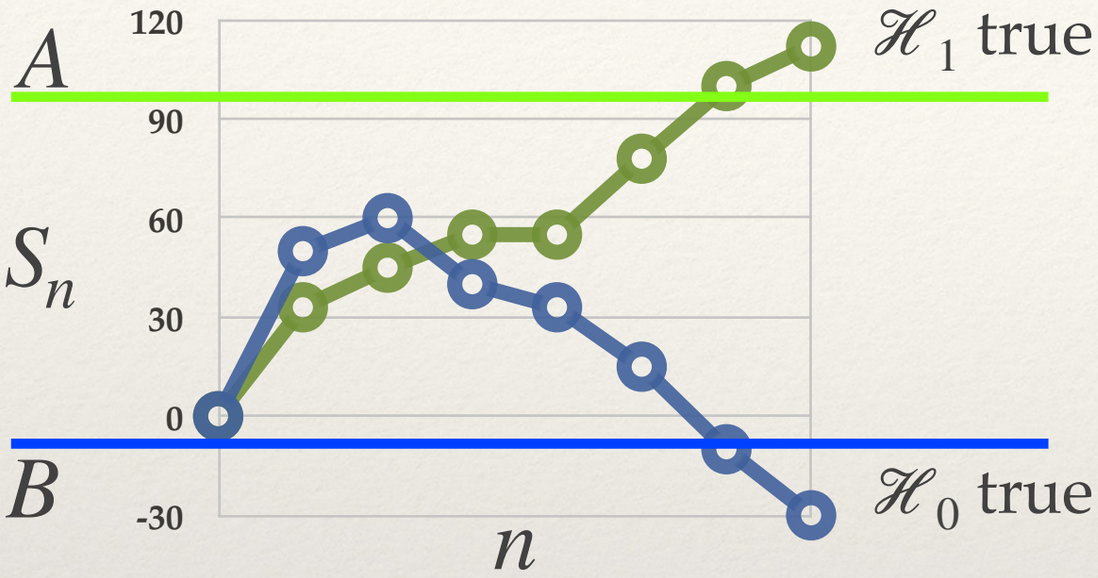
- Wald's idea: Construct the test statistic

$$S_n = \log \frac{P_{\theta_1}(X_1, \ldots, X_n)}{P_{\theta_0}(X_1, \ldots, X_n)}$$

- $E[S_n \mid \mathcal{H}_1] = nD(P_{\theta_1} \| P_{\theta_0}), \quad E[S_n \mid \mathcal{H}_0] = -nD(P_{\theta_0} \| P_{\theta_1})$
- Set two thresholds: $A > 0$ and $B < 0$ (these depend on the priors)
- At time n:
  - Stop and declare $\mathcal{H}_1$ true if $S_n \geq A$
  - Stop and declare $\mathcal{H}_0$ true if $S_n \leq B$
  - Continue taking observations if $B < S_n < A$

- For very small values of cost c:

Wald's SPRT



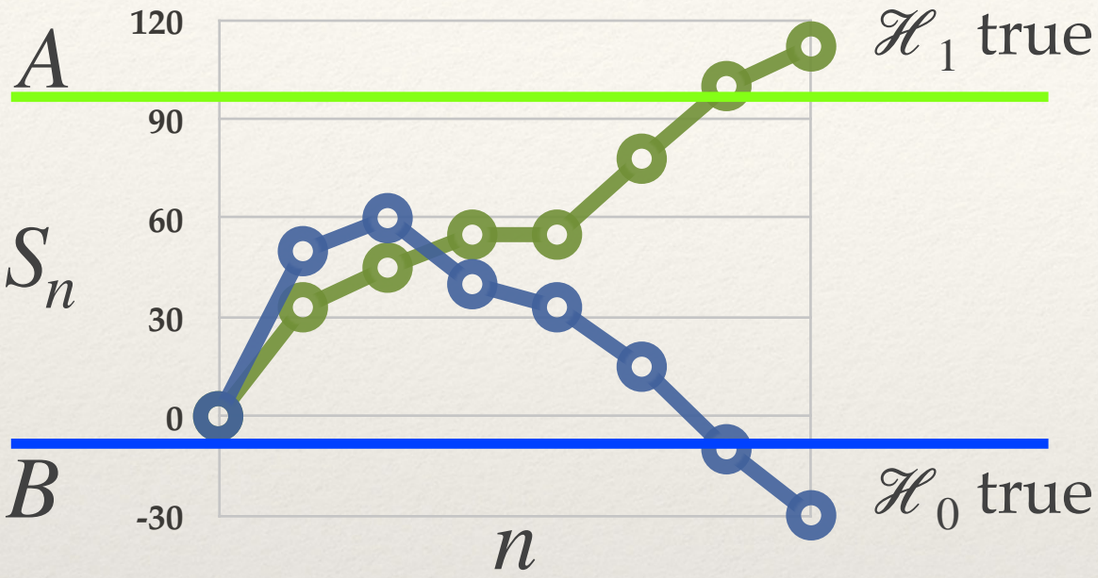
- Wald's idea: Construct the test statistic

$$S_n = \log \frac{P_{\theta_1}(X_1, \ldots, X_n)}{P_{\theta_0}(X_1, \ldots, X_n)}$$

- $E[S_n \mid \mathcal{H}_1] = nD(P_{\theta_1} \| P_{\theta_0}), \quad E[S_n \mid \mathcal{H}_0] = -nD(P_{\theta_0} \| P_{\theta_1})$
- Set two thresholds: $A > 0$ and $B < 0$ (these depend on the priors)
- At time n:
  - Stop and declare $\mathcal{H}_1$ true if $S_n \geq A$
  - Stop and declare $\mathcal{H}_0$ true if $S_n \leq B$
  - Continue taking observations if $B < S_n < A$

- For very small values of cost c:
  - $A \approx -\log c, \quad B \approx \log c$

Wald's SPRT



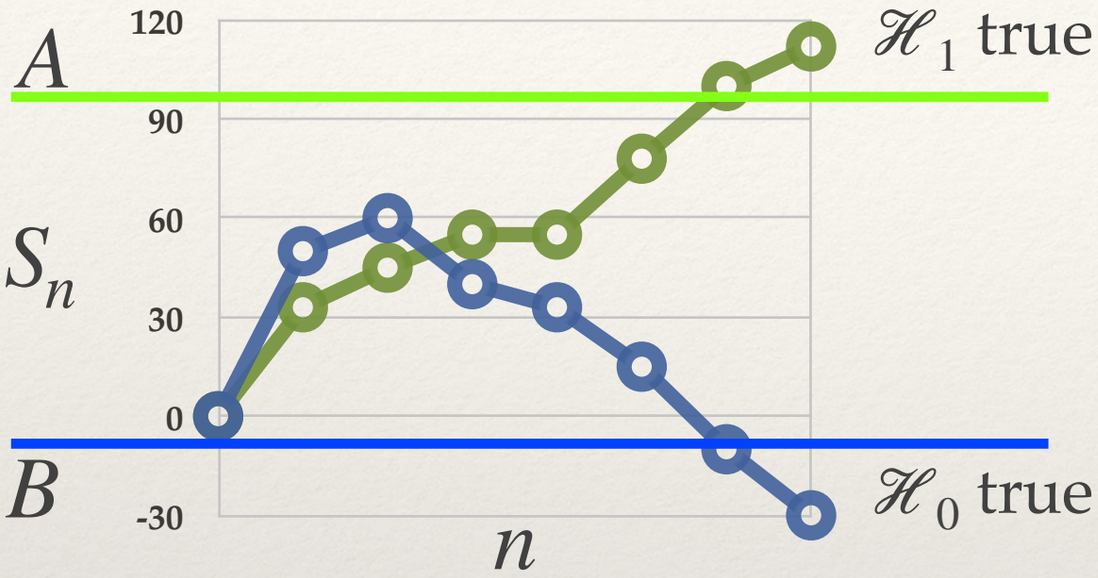
- Wald's idea: Construct the test statistic

$$S_n = \log \frac{P_{\theta_1}(X_1, \ldots, X_n)}{P_{\theta_0}(X_1, \ldots, X_n)}$$

- $E[S_n \mid \mathcal{H}_1] = nD(P_{\theta_1} \| P_{\theta_0}), \quad E[S_n \mid \mathcal{H}_0] = -nD(P_{\theta_0} \| P_{\theta_1})$
- Set two thresholds: $A > 0$ and $B < 0$ (these depend on the priors)
- At time n:
  - Stop and declare $\mathcal{H}_1$ true if $S_n \geq A$
  - Stop and declare $\mathcal{H}_0$ true if $S_n \leq B$
  - Continue taking observations if $B < S_n < A$

- For very small values of cost c:
  - $A \approx -\log c, \quad B \approx \log c$
  - $S_N \approx -\log c$ if $\mathcal{H}_1$ true
    $S_N \approx \log c$ if $\mathcal{H}_0$ true

Wald's SPRT



- Wald's idea: Construct the test statistic

$$S_n = \log \frac{P_{\theta_1}(X_1, \ldots, X_n)}{P_{\theta_0}(X_1, \ldots, X_n)}$$

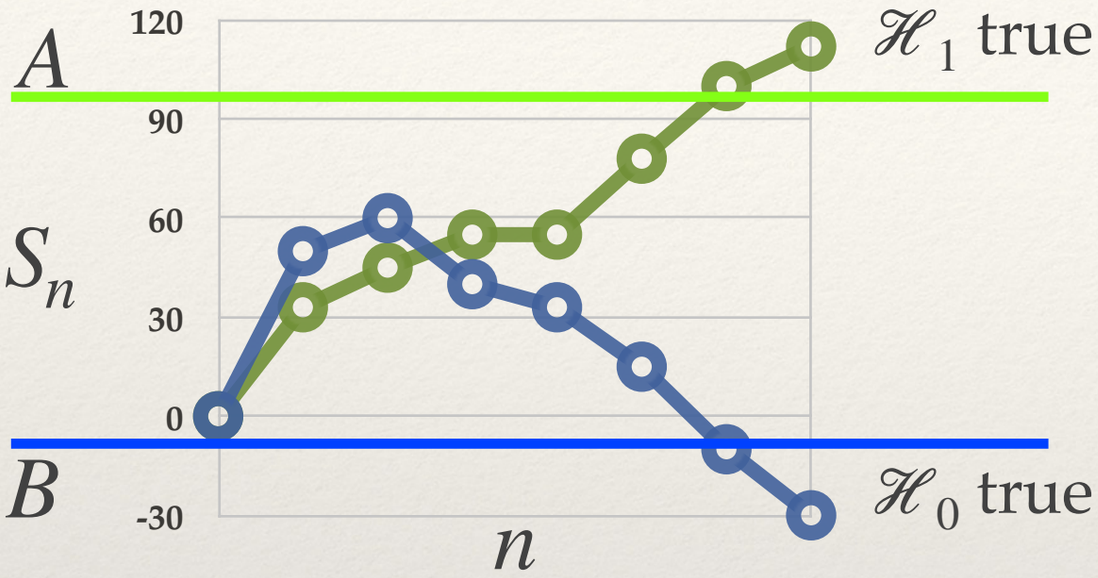
- $E[S_n \mid \mathcal{H}_1] = nD(P_{\theta_1} \| P_{\theta_0}), \quad E[S_n \mid \mathcal{H}_0] = -nD(P_{\theta_0} \| P_{\theta_1})$
- Set two thresholds: $A > 0$ and $B < 0$ (these depend on the priors)
- At time n:
  - Stop and declare $\mathcal{H}_1$ true if $S_n \geq A$
  - Stop and declare $\mathcal{H}_0$ true if $S_n \leq B$
  - Continue taking observations if $B < S_n < A$

- For very small values of cost c:
  - $A \approx -\log c, \quad B \approx \log c$
  - $S_N \approx -\log c$ if $\mathcal{H}_1$ true
    $S_N \approx \log c$ if $\mathcal{H}_0$ true
  - By Wald's identity,

$$E[N \mid \mathcal{H}_1] \approx \frac{-\log c}{D(P_{\theta_1} \| P_{\theta_0})}$$

$$E[N \mid \mathcal{H}_0] \approx \frac{-\log c}{D(P_{\theta_0} \| P_{\theta_1})}$$

Wald's SPRT



- Wald's idea: Construct the test statistic

$$S_n = \log \frac{P_{\theta_1}(X_1, \ldots, X_n)}{P_{\theta_0}(X_1, \ldots, X_n)}$$

- $E[S_n \mid \mathcal{H}_1] = nD(P_{\theta_1} \| P_{\theta_0}), \quad E[S_n \mid \mathcal{H}_0] = -nD(P_{\theta_0} \| P_{\theta_1})$
- Set two thresholds: $A > 0$ and $B < 0$ (these depend on the priors)
- At time n:
  - Stop and declare $\mathcal{H}_1$ true if $S_n \geq A$
  - Stop and declare $\mathcal{H}_0$ true if $S_n \leq B$
  - Continue taking observations if $B < S_n < A$

- For very small values of cost c:
  - $A \approx -\log c, \quad B \approx \log c$
  - $S_N \approx -\log c$ if $\mathcal{H}_1$ true
    $S_N \approx \log c$ if $\mathcal{H}_0$ true
  - By Wald's identity,

$$E[N \mid \mathcal{H}_1] \approx \frac{-\log c}{D(P_{\theta_1} \| P_{\theta_0})}$$

$$E[N \mid \mathcal{H}_0] \approx \frac{-\log c}{D(P_{\theta_0} \| P_{\theta_1})}$$

Wald's Identity

$$E[S_N \mid \mathcal{H}_1] = E[N \mid \mathcal{H}_1] \cdot D(P_{\theta_1} \| P_{\theta_0})$$

$$E[S_N \mid \mathcal{H}_0] = E[N \mid \mathcal{H}_0] \cdot (-D(P_{\theta_0} \| P_{\theta_1}))$$

Handling Multiple Hypotheses in Wald's Formulation

- Objective: to test

$$\begin{aligned} \mathcal{H}_1 &: \quad \theta = \theta_1 \\ \mathcal{H}_2 &: \quad \theta = \theta_2 \\ &\vdots \\ \mathcal{H}_K &: \quad \theta = \theta_K \end{aligned}$$

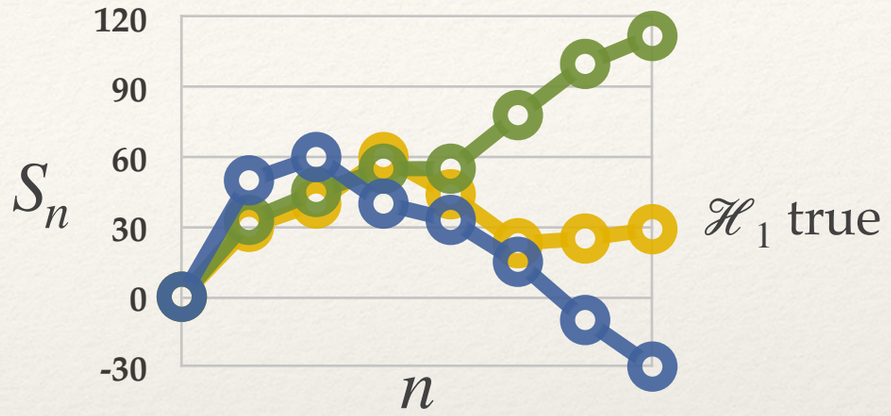
- IID observations $X_1, X_2, \ldots \sim P_{\theta_i}$ under hypothesis $\mathcal{H}_i$
- P_{θ_i} known for all i
- How to devise a test to determine if $\mathcal{H}_0$ is true or $\mathcal{H}_1$ is true?

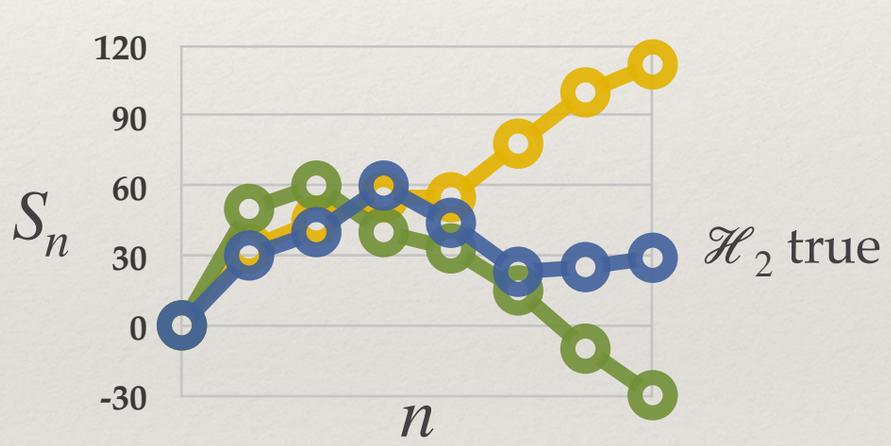
Handling Multiple Hypotheses in Wald's Formulation

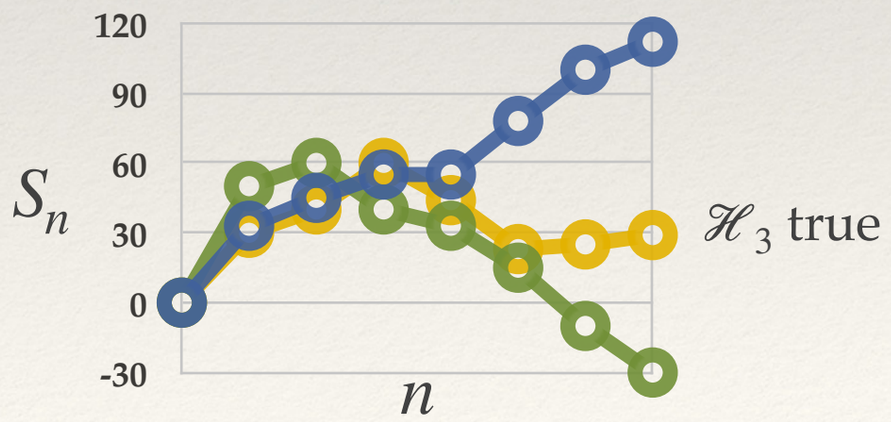
- Objective: to test

$$\begin{aligned} \mathcal{H}_1 &: \quad \theta = \theta_1 \\ \mathcal{H}_2 &: \quad \theta = \theta_2 \\ &\vdots \\ \mathcal{H}_K &: \quad \theta = \theta_K \end{aligned}$$

- IID observations $X_1, X_2, \ldots \sim P_{\theta_i}$ under hypothesis $\mathcal{H}_i$
- P_{θ_i} known for all i
- How to devise a test to determine if $\mathcal{H}_0$ is true or $\mathcal{H}_1$ is true?







Handling Multiple Hypotheses in Wald's Formulation

- Objective: to test

$$\begin{aligned} \mathcal{H}_1 &: \quad \theta = \theta_1 \\ \mathcal{H}_2 &: \quad \theta = \theta_2 \\ &\vdots \\ \mathcal{H}_K &: \quad \theta = \theta_K \end{aligned}$$

- IID observations $X_1, X_2, \ldots \sim P_{\theta_i}$ under hypothesis $\mathcal{H}_i$
- P_{θ_i} known for all i
- How to devise a test to determine if $\mathcal{H}_0$ is true or $\mathcal{H}_1$ is true?

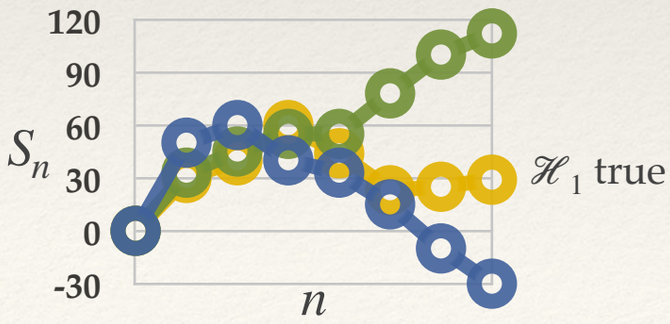
- A natural choice for the test statistic is as follows:

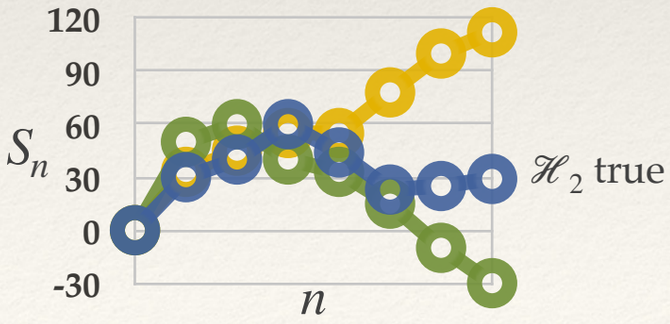
$$S_i(n) = \log \frac{P_{\theta_i}(X_1, \ldots, X_n)}{\max_{j \neq i} P_{\theta_j}(X_1, \ldots, X_n)}$$

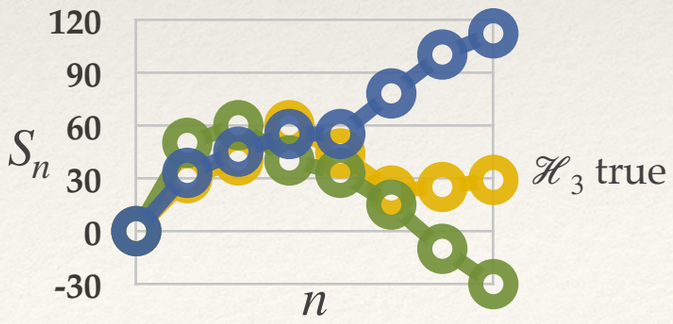
$$S(n) = \max\{S_1(n), \ldots, S_K(n)\}$$

$$= \max_i \min_{j \neq i} \log \frac{P_{\theta_i}(X_1, \ldots, X_n)}{P_{\theta_j}(X_1, \ldots, X_n)}$$

- At time n:
  - Let $i^*(n) = \arg\max_i S_i(n)$
  - Stop and declare $\mathcal{H}_{i^*(n)}$ true if $S_{i^*(n)} \geq -\log c$
  - Else continue experimentation (taking one more observation)







Handling Multiple Hypotheses in Wald's Formulation

- Objective: to test

$$\begin{aligned} \mathcal{H}_1 &: \quad \theta = \theta_1 \\ \mathcal{H}_2 &: \quad \theta = \theta_2 \\ &\vdots \\ \mathcal{H}_K &: \quad \theta = \theta_K \end{aligned}$$

- IID observations $X_1, X_2, \ldots \sim P_{\theta_i}$ under hypothesis $\mathcal{H}_i$
- P_{θ_i} known for all i
- How to devise a test to determine if $\mathcal{H}_0$ is true or $\mathcal{H}_1$ is true?

- A natural choice for the test statistic is as follows:

$$S_i(n) = \log \frac{P_{\theta_i}(X_1, \ldots, X_n)}{\max\limits_{j \neq i} P_{\theta_j}(X_1, \ldots, X_n)}$$

$$S(n) = \max\{S_1(n), \ldots, S_K(n)\}$$

$$= \max_i \min_{j \neq i} \log \frac{P_{\theta_i}(X_1, \ldots, X_n)}{P_{\theta_j}(X_1, \ldots, X_n)}$$

- At time n:
  - Let $i^*(n) = \arg\max\limits_i S_i(n)$
  - Stop and declare $\mathcal{H}_{i^*(n)}$ true if $S_{i^*(n)} \geq -\log c$
  - Else continue experimentation (taking one more observation)

$$E[N \mid \mathcal{H}_i] \approx \frac{-\log c}{\min\limits_{j \neq i} D(P_{\theta_i} \| P_{\theta_j})}$$

From Wald to Chernoff: Active Sequential Hypothesis Testing (ASHT)

From Wald to Chernoff: Active Sequential Hypothesis Testing (ASHT)

- To test

$$\begin{aligned}\mathscr{H}_1 &: \quad \theta = \theta_1 \\ \mathscr{H}_2 &: \quad \theta = \theta_2 \\ &\vdots \\ \mathscr{H}_K &: \quad \theta = \theta_K\end{aligned}$$

From Wald to Chernoff: Active Sequential Hypothesis Testing (ASHT)

- To test

$$\begin{aligned}\mathscr{H}_1 &: \quad \theta = \theta_1 \\ \mathscr{H}_2 &: \quad \theta = \theta_2 \\ &\vdots \\ \mathscr{H}_K &: \quad \theta = \theta_K\end{aligned}$$

- A set of experiments $E_1, \ldots, E_m$ available (these will correspond to arms later on)

From Wald to Chernoff: Active Sequential Hypothesis Testing (ASHT)

- To test

$$\begin{aligned}\mathscr{H}_1 &: \quad \theta = \theta_1 \\ \mathscr{H}_2 &: \quad \theta = \theta_2 \\ &\vdots \\ \mathscr{H}_K &: \quad \theta = \theta_K\end{aligned}$$

- A set of experiments $E_1, \ldots, E_m$ available (these will correspond to arms later on)
- IID observations $X_1, X_2, \ldots \sim P_{(\theta_i, E_j)}$ under hypothesis $\mathscr{H}_i$ and experiment E_j

From Wald to Chernoff: Active Sequential Hypothesis Testing (ASHT)

- To test

$$\begin{aligned}\mathscr{H}_1 &: \quad \theta = \theta_1 \\ \mathscr{H}_2 &: \quad \theta = \theta_2 \\ &\vdots \\ \mathscr{H}_K &: \quad \theta = \theta_K\end{aligned}$$

- A set of experiments $E_1, \ldots, E_m$ available (these will correspond to arms later on)
- IID observations $X_1, X_2, \ldots \sim P_{(\theta_i, E_j)}$ under hypothesis $\mathscr{H}_i$ and experiment E_j
- $P_{(\theta_i, E_j)}$ known for all i, j

From Wald to Chernoff: Active Sequential Hypothesis Testing (ASHT)

- To test

$$\begin{aligned}\mathscr{H}_1 &: \quad \theta = \theta_1 \\ \mathscr{H}_2 &: \quad \theta = \theta_2 \\ &\vdots \\ \mathscr{H}_K &: \quad \theta = \theta_K\end{aligned}$$

- A set of experiments $E_1, \ldots, E_m$ available (these will correspond to arms later on)
- IID observations $X_1, X_2, \ldots \sim P_{(\theta_i, E_j)}$ under hypothesis $\mathscr{H}_i$ and experiment E_j
- $P_{(\theta_i, E_j)}$ known for all i, j
- At each time, any one of the experiments may be chosen

From Wald to Chernoff: Active Sequential Hypothesis Testing (ASHT)

- To test

$$\begin{aligned}\mathscr{H}_1 : &\quad \theta = \theta_1 \\ \mathscr{H}_2 : &\quad \theta = \theta_2 \\ &\vdots \\ \mathscr{H}_K : &\quad \theta = \theta_K\end{aligned}$$

- A set of experiments $E_1, \ldots, E_m$ available (these will correspond to arms later on)
- IID observations $X_1, X_2, \ldots \sim P_{(\theta_i, E_j)}$ under hypothesis $\mathscr{H}_i$ and experiment E_j
- $P_{(\theta_i, E_j)}$ known for all i, j
- At each time, any one of the experiments may be chosen
- Cost c per round of experimentation

From Wald to Chernoff: Active Sequential Hypothesis Testing (ASHT)

- To test

$$\begin{aligned}\mathscr{H}_1 : &\quad \theta = \theta_1 \\ \mathscr{H}_2 : &\quad \theta = \theta_2 \\ &\vdots \\ \mathscr{H}_K : &\quad \theta = \theta_K\end{aligned}$$

- A set of experiments $E_1, \ldots, E_m$ available (these will correspond to arms later on)
- IID observations $X_1, X_2, \ldots \sim P_{(\theta_i, E_j)}$ under hypothesis $\mathscr{H}_i$ and experiment E_j
- $P_{(\theta_i, E_j)}$ known for all i, j
- At each time, any one of the experiments may be chosen
- Cost c per round of experimentation
- How to devise a test to determine which hypothesis is true?

Chernoff's *Procedure A*

Chernoff's *Procedure A*

- Suppose the *n*th experiment (arm chosen) is A_n

Chernoff's *Procedure A*

- Suppose the *n*th experiment (arm chosen) is A_n
- Then, borrowing from Wald, construct the following test statistic.

Chernoff's *Procedure A*

- Suppose the *n*th experiment (arm chosen) is A_n
- Then, borrowing from Wald, construct the following test statistic.

$$S_i(n) = \log \frac{P_{\theta_i}(X_1, A_1 \ldots, X_n, A_n)}{\max\limits_{j \neq i} P_{\theta_j}(X_1, A_1, \ldots, X_n, A_n)}$$

$$i^*(n) = \arg\max\{S_1(n), \ldots, S_K(n)\}$$

Chernoff's *Procedure A*

- Suppose the nth experiment (arm chosen) is A_n
- Then, borrowing from Wald, construct the following test statistic.

$$S_i(n) = \log \frac{P_{\theta_i}(X_1, A_1 \ldots, X_n, A_n)}{\max\limits_{j \neq i} P_{\theta_j}(X_1, A_1, \ldots, X_n, A_n)}$$

$$i^*(n) = \arg\max\{S_1(n), \ldots, S_K(n)\}$$

- Stop and declare $\mathscr{H}_{i^*(n)}$ true if $S_{i^*(n)} \geq -\log c$. Else choose an experiment and continue.

Chernoff's Contributions

Chernoff's Contributions

- Chernoff's results were presented in two stages:

Chernoff's Contributions

- Chernoff's results were presented in two stages:
  - An asymptotic lower bound on the expected number of samples

$$E[N \mid \mathscr{H}_i] \gtrapprox \frac{-\log c}{D_i^*}, \qquad D_i^* = \max_{\lambda} \min_{j \neq i} \sum_{k=1}^{m} \lambda(E_k)\, D(P_{(\theta_i, E_k)} \| P_{(\theta_j, E_k)})$$

Chernoff's Contributions

- Chernoff's results were presented in two stages:
  - An asymptotic lower bound on the expected number of samples

$$E[N \mid \mathcal{H}_i] \gtrapprox \frac{-\log c}{D_i^*}, \qquad D_i^* = \max_{\lambda} \min_{j \neq i} \sum_{k=1}^{m} \lambda(E_k)\, D(P_{(\theta_i, E_k)} \| P_{(\theta_j, E_k)})$$

  - An achievability proof that *Procedure A* meets the lower bound asymptotically

Chernoff's Contributions

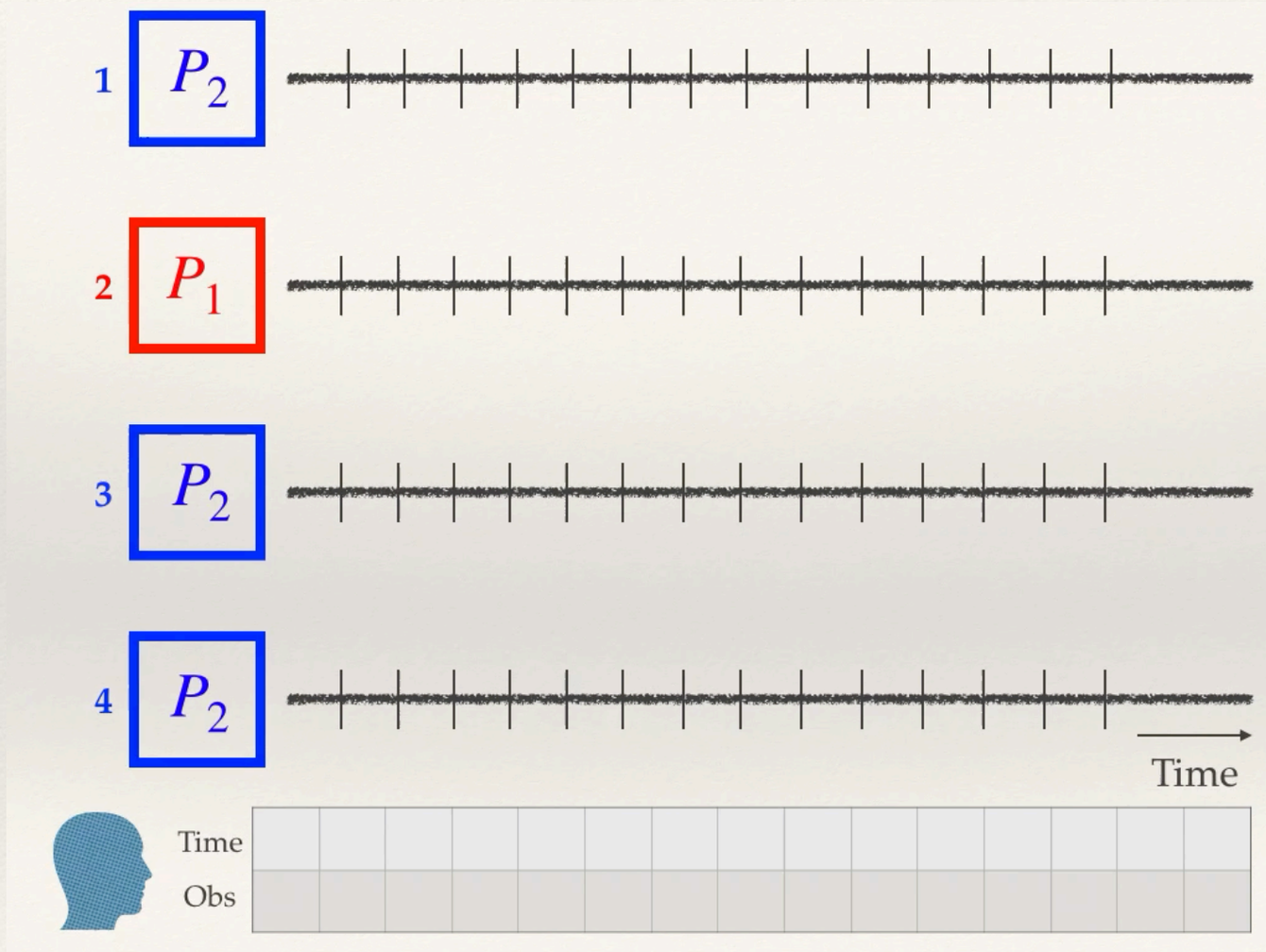
- Chernoff's results were presented in two stages:
  - An asymptotic lower bound on the expected number of samples

$$E[N \mid \mathscr{H}_i] \gtrapprox \frac{-\log c}{D_i^*}, \qquad D_i^* = \max_{\lambda} \min_{j \neq i} \sum_{k=1}^{m} \lambda(E_k)\, D(P_{(\theta_i, E_k)} \| P_{(\theta_j, E_k)})$$

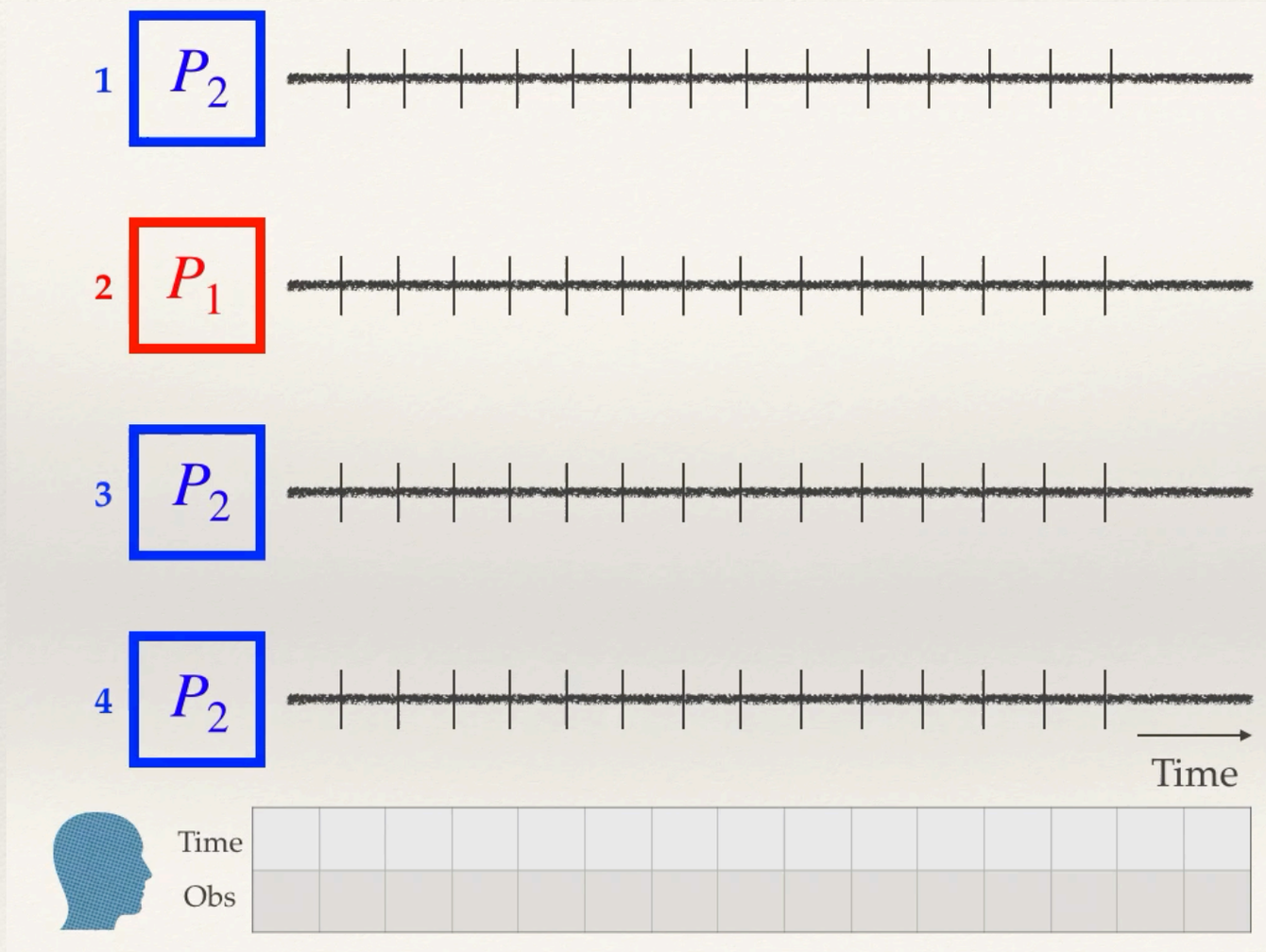
  - An achievability proof that *Procedure A* meets the lower bound asymptotically
- Here, asymptotics is as the cost of experimentation $c \downarrow 0$

Back to Rested Markov Bandits

Recall the Setup

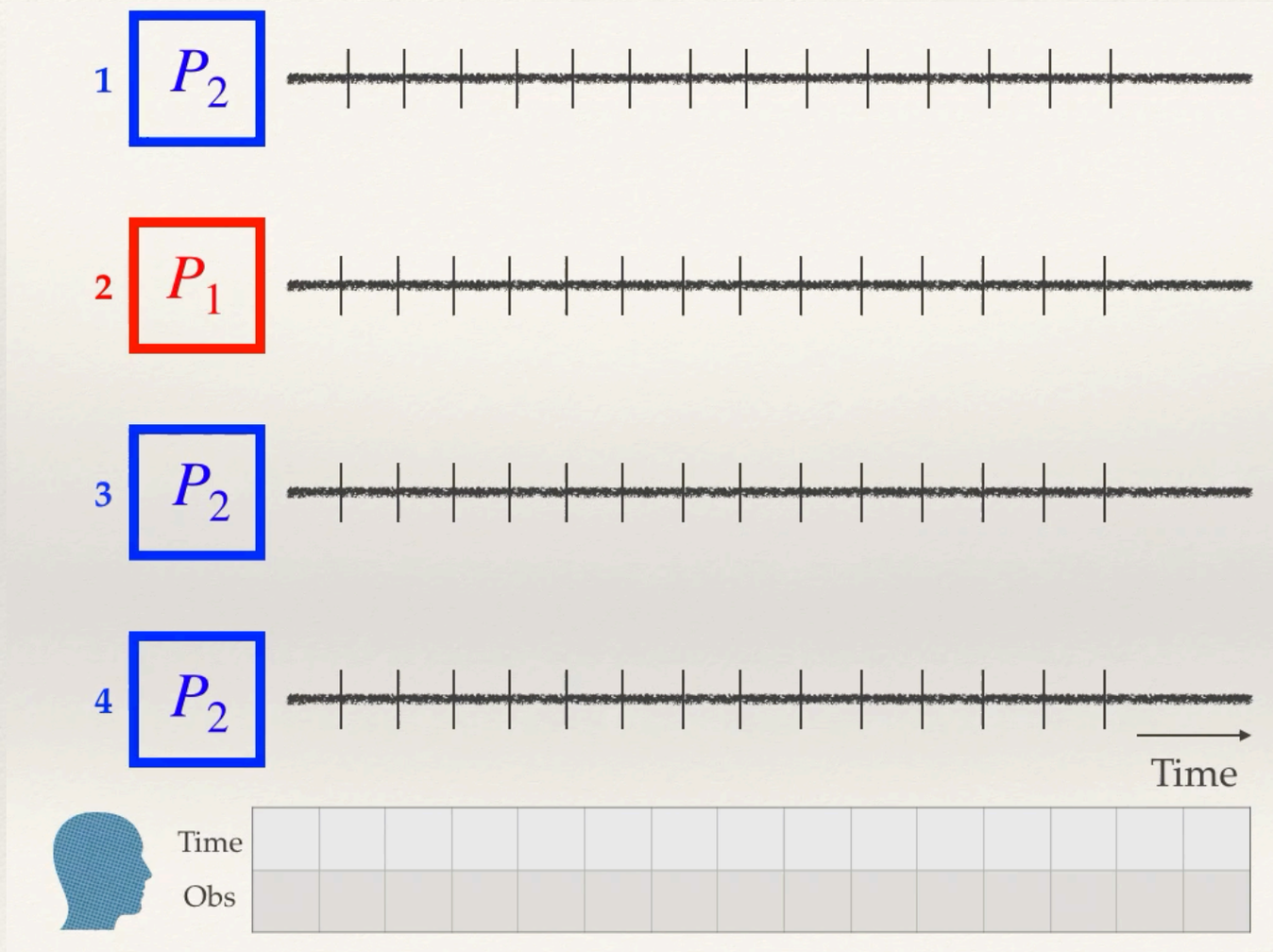


Recall the Setup



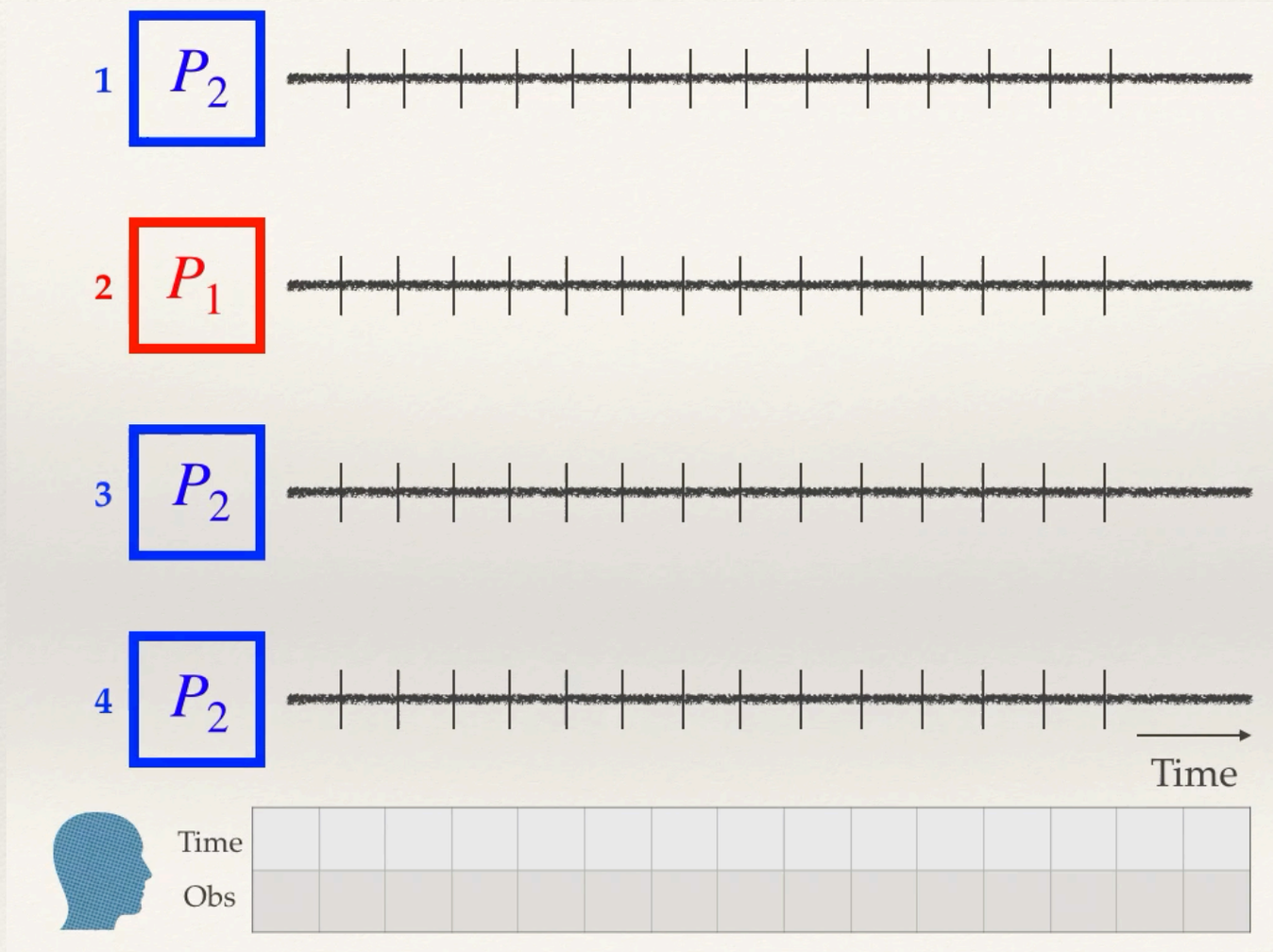
Recall the Setup

- A multi-armed bandit with K independent arms



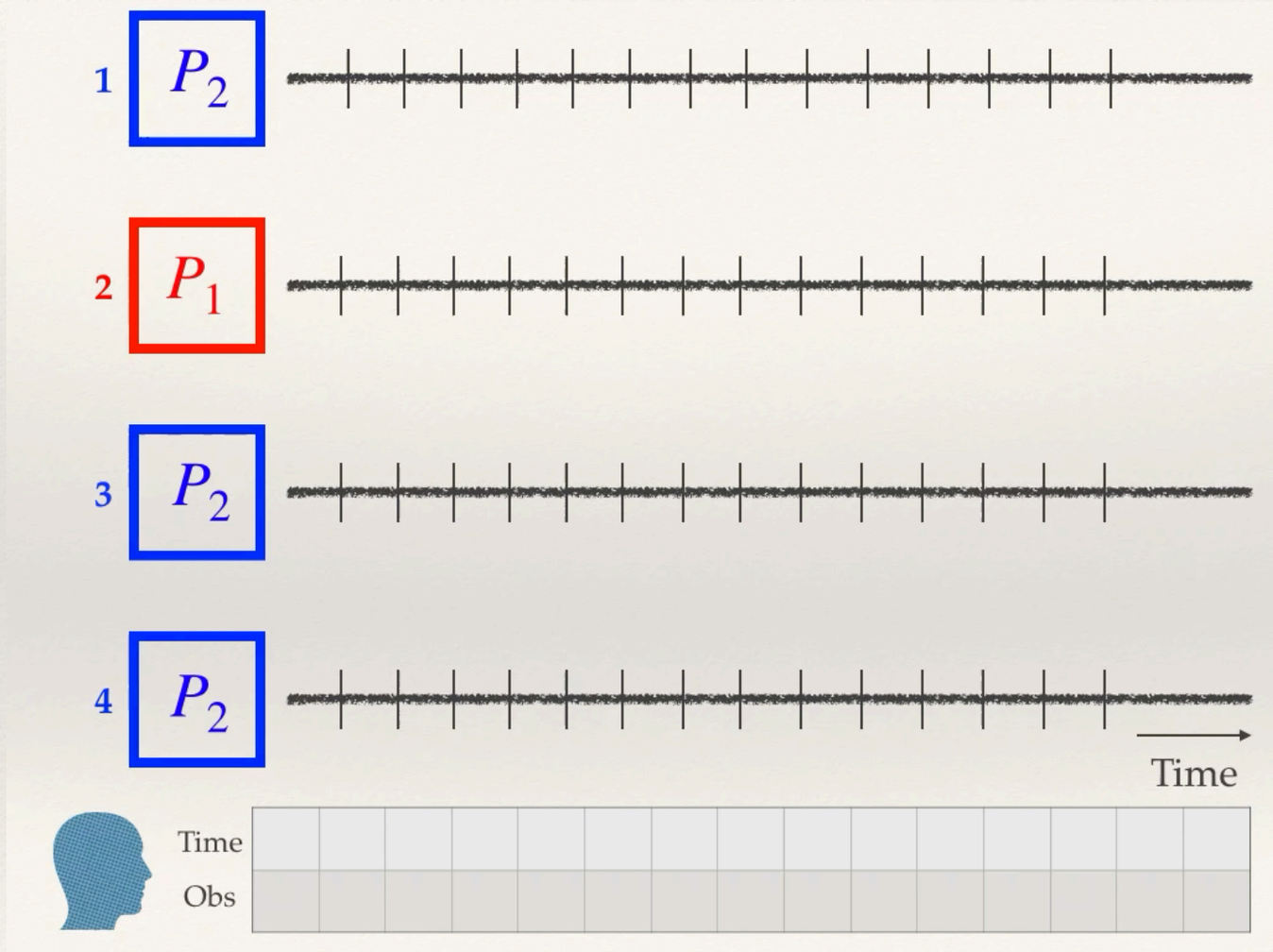
Recall the Setup

- A multi-armed bandit with K independent arms
- Each arm is a homogeneous and ergodic Markov chain on a finite state space



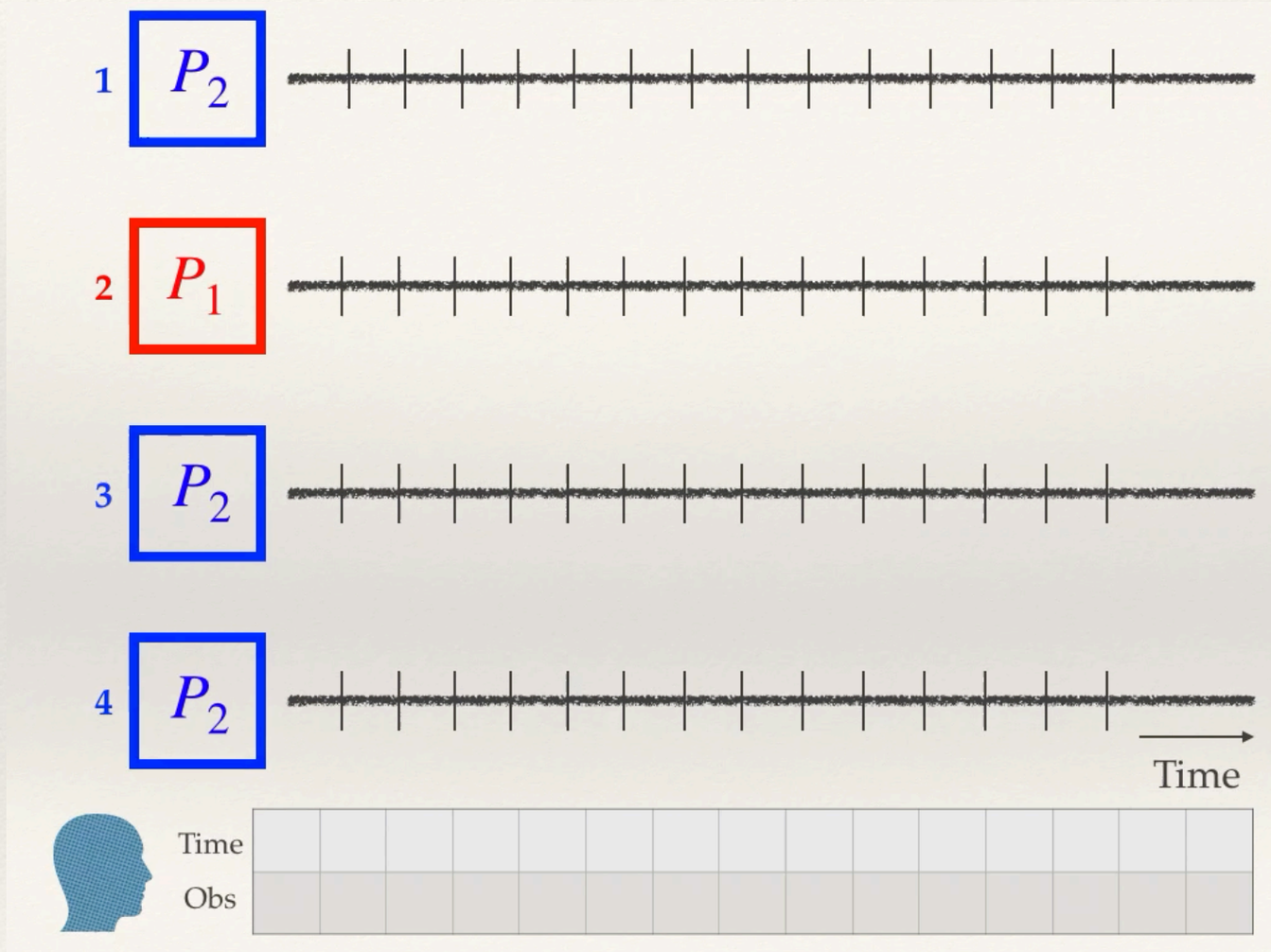
Recall the Setup

- A multi-armed bandit with K independent arms
- Each arm is a homogeneous and ergodic Markov chain on a finite state space
- State space common to all the arms



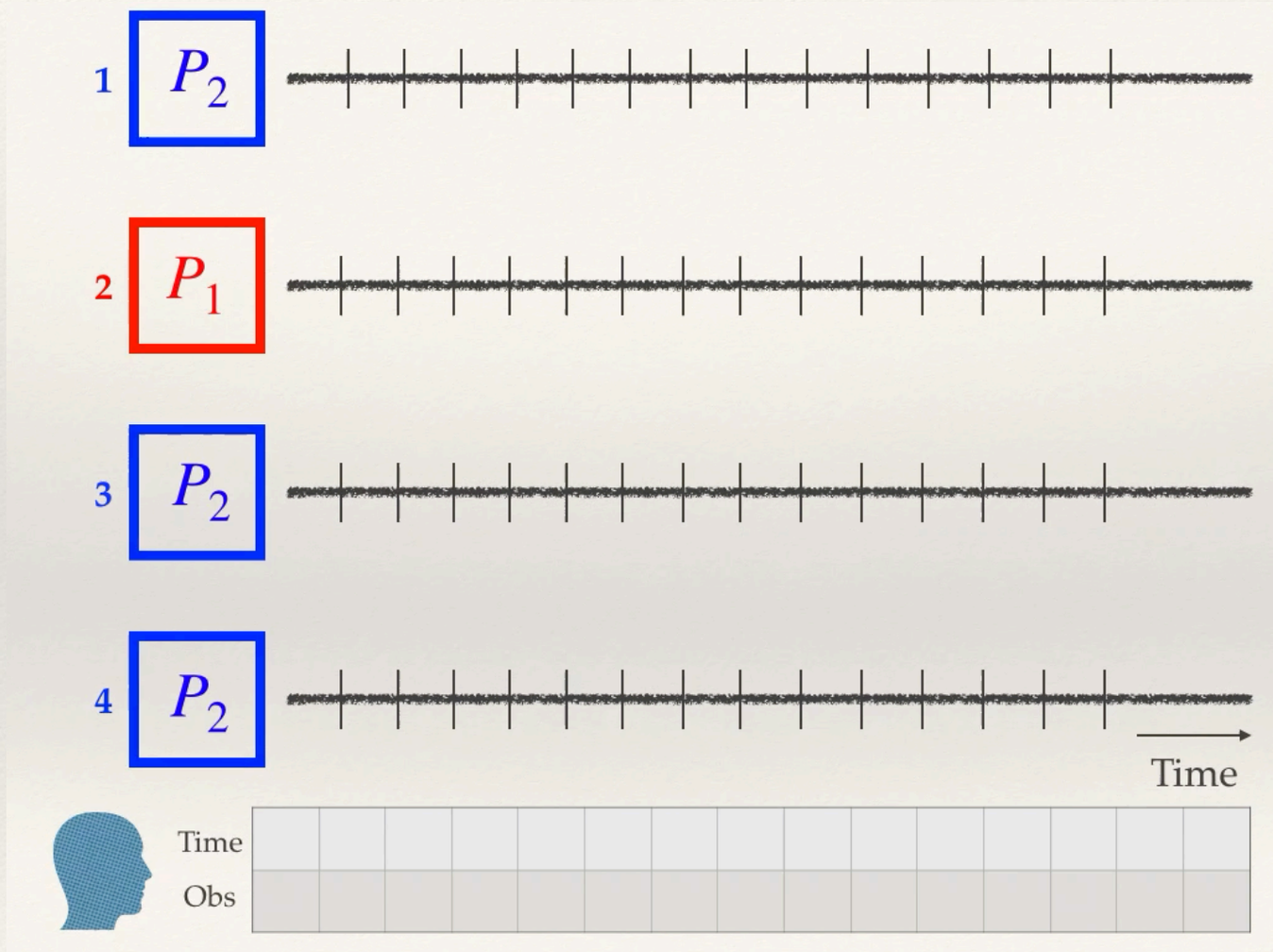
Recall the Setup

- A multi-armed bandit with K independent arms
- Each arm is a homogeneous and ergodic Markov chain on a finite state space
- State space common to all the arms
- One of the arms has TPM P_1, rest have TPM P_2, where $P_2 \neq P_1$



Recall the Setup

- A multi-armed bandit with K independent arms
- Each arm is a homogeneous and ergodic Markov chain on a finite state space
- State space common to all the arms
- One of the arms has TPM P_1, rest have TPM P_2, where $P_2 \neq P_1$
- You (learner) are given the task to figure out which arm has P_1 as quickly as possible under the **PAC framework**

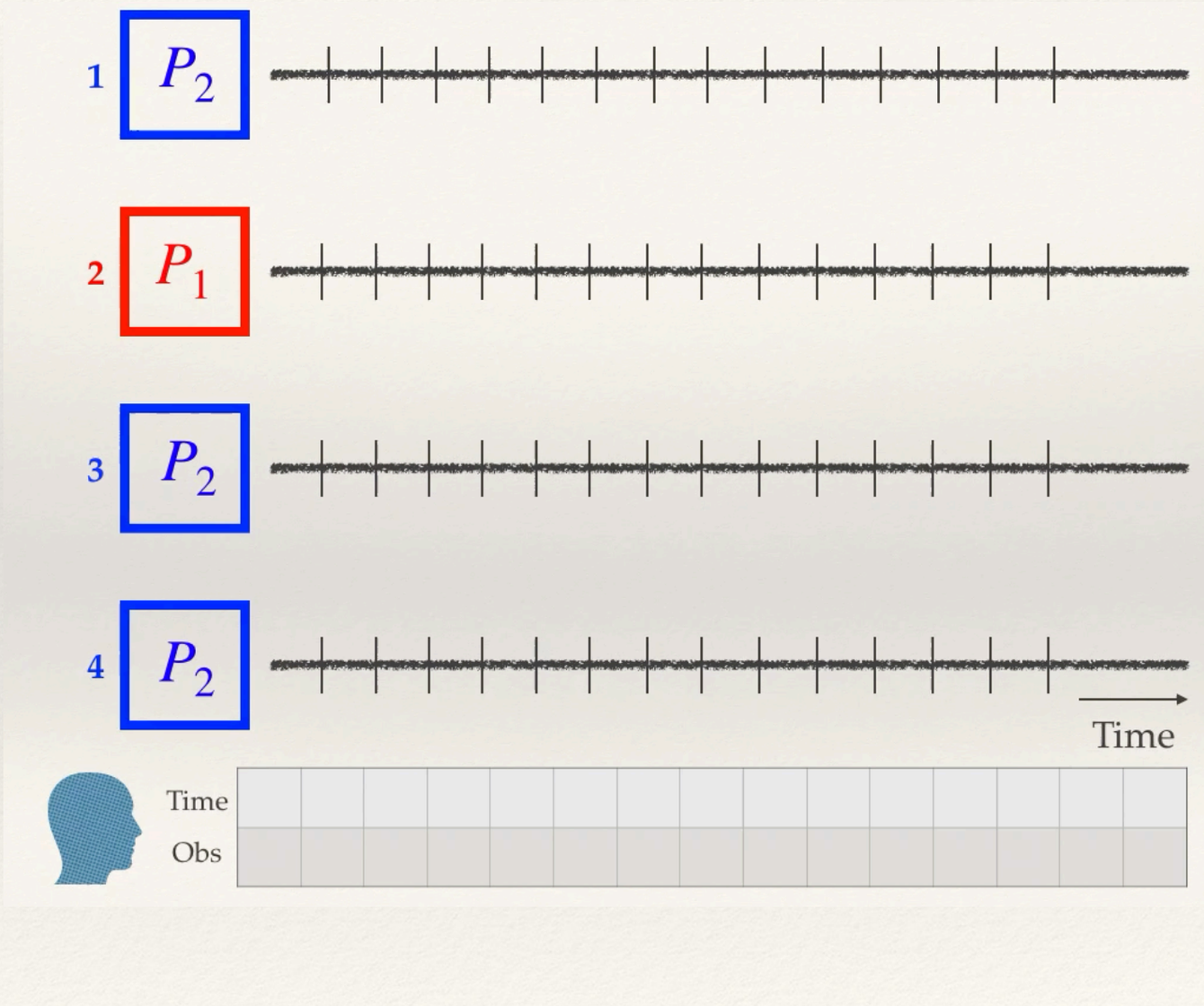


Recall the Setup

- A multi-armed bandit with K independent arms
- Each arm is a homogeneous and ergodic Markov chain on a finite state space
- State space common to all the arms
- One of the arms has TPM P_1, rest have TPM P_2, where $P_2 \neq P_1$
- You (learner) are given the task to figure out which arm has P_1 as quickly as possible under the **PAC framework**
- Can select one arm at each time to observe (**sequential** selection)

Recall the Setup

- A multi-armed bandit with K independent arms
- Each arm is a homogeneous and ergodic Markov chain on a finite state space
- State space common to all the arms
- One of the arms has TPM P_1, rest have TPM P_2, where $P_2 \neq P_1$
- You (learner) are given the task to figure out which arm has P_1 as quickly as possible under the **PAC framework**
- Can select one arm at each time to observe (**sequential** selection)
- Unobserved arms remain frozen (**rested** arms)



Embedding Our Problem in Chernoff's Framework

Embedding Our Problem in Chernoff's Framework

- Arms ≡ experiments

Embedding Our Problem in Chernoff's Framework

- Arms ≡ experiments
- Cost of experimentation $c \downarrow 0 \equiv$ error probability vanishing

Embedding Our Problem in Chernoff's Framework

- Arms ≡ experiments
- Cost of experimentation $c \downarrow 0 \equiv$ error probability vanishing
- Hypotheses:

Embedding Our Problem in Chernoff's Framework

- Arms ≡ experiments
- Cost of experimentation $c \downarrow 0 \equiv$ error probability vanishing
- Hypotheses:
  - If P_1, P_2 are known:

$$\begin{aligned} \mathscr{H}_1 &: \theta = \text{odd arm is } 1 \\ \mathscr{H}_2 &: \theta = \text{odd arm is } 2 \\ &\vdots \\ \mathscr{H}_K &: \theta = \text{odd arm is } K \end{aligned}$$

Embedding Our Problem in Chernoff's Framework

- Arms ≡ experiments
- Cost of experimentation $c \downarrow 0 \equiv$ error probability vanishing
- Hypotheses:
  - If P_1, P_2 are known:

$$\begin{aligned}
\mathscr{H}_1 &: \theta = \text{odd arm is } 1 \\
\mathscr{H}_2 &: \theta = \text{odd arm is } 2 \\
&\vdots \\
\mathscr{H}_K &: \theta = \text{odd arm is } K
\end{aligned}$$

  - If P_1, P_2 are not known:

$$\begin{aligned}
\mathscr{H}_1 &: \theta = (\text{odd arm}, P_1, P_2) = (1, \cdot, \cdot) \\
\mathscr{H}_2 &: \theta = (\text{odd arm}, P_1, P_2) = (2, \cdot, \cdot) \\
&\vdots \\
\mathscr{H}_K &: \theta = (\text{odd arm}, P_1, P_2) = (K, \cdot, \cdot)
\end{aligned}$$

Embedding Our Problem in Chernoff's Framework

- Arms ≡ experiments
- Cost of experimentation $c \downarrow 0 \equiv$ error probability vanishing
- Hypotheses:
  - If P_1, P_2 are known:

$$\begin{aligned}
\mathcal{H}_1 &: \theta = \text{odd arm is } 1 \\
\mathcal{H}_2 &: \theta = \text{odd arm is } 2 \\
&\vdots \\
\mathcal{H}_K &: \theta = \text{odd arm is } K
\end{aligned}$$

K simple hypotheses

  - If P_1, P_2 are not known:

$$\begin{aligned}
\mathcal{H}_1 &: \theta = (\text{odd arm}, P_1, P_2) = (1, \cdot, \cdot) \\
\mathcal{H}_2 &: \theta = (\text{odd arm}, P_1, P_2) = (2, \cdot, \cdot) \\
&\vdots \\
\mathcal{H}_K &: \theta = (\text{odd arm}, P_1, P_2) = (K, \cdot, \cdot)
\end{aligned}$$

K composite hypotheses

Embedding Our Problem in Chernoff's Framework

- Arms ≡ experiments
- Cost of experimentation $c \downarrow 0 \equiv$ error probability vanishing
- Hypotheses:
  - If P_1, P_2 are known:

$$\begin{aligned}
\mathcal{H}_1 &: \theta = \text{odd arm is } 1 \\
\mathcal{H}_2 &: \theta = \text{odd arm is } 2 \\
&\vdots \\
\mathcal{H}_K &: \theta = \text{odd arm is } K
\end{aligned}$$

K simple hypotheses

  - If P_1, P_2 are not known:

$$\begin{aligned}
\mathcal{H}_1 &: \theta = (\text{odd arm}, P_1, P_2) = (1, \cdot, \cdot) \\
\mathcal{H}_2 &: \theta = (\text{odd arm}, P_1, P_2) = (2, \cdot, \cdot) \\
&\vdots \\
\mathcal{H}_K &: \theta = (\text{odd arm}, P_1, P_2) = (K, \cdot, \cdot)
\end{aligned}$$

K composite hypotheses

We obtain results for the composite hypotheses case

Our Contributions

- We obtain an asymptotic[1] lower bound on the expected number of samples required to identify the odd arm as a function of error probability
- We propose an asymptotically optimal scheme which meets the desired error probability criterion
- The scheme is a modification of the classical GLRT with forced exploration
- Key challenges in the Markov setting identified
- We believe our results constitute a key first step towards solving the more difficult case of "restless" Markov arms

[1]The asymptotics is as the error tolerance vanishes.

Notations

Notations

- $(A_k)_{k\geq 0}$: arm selection process

Notations

- $(A_k)_{k\geq 0}$: arm selection process
- $(X_k)_{k\geq 0}$: observation process

Notations

- $(A_k)_{k\geq 0}$: arm selection process
- $(X_k)_{k\geq 0}$: observation process
- TPM of odd arm: P_1; TPM of other arms: P_2 (both unknown)

Notations

- $(A_k)_{k\geq 0}$: arm selection process
- $(X_k)_{k\geq 0}$: observation process
- TPM of odd arm: P_1; TPM of other arms: P_2 (both unknown)
- $C = (i, P_1, P_2)$ denotes a config. of arms in which i is odd arm, P_1 is the TPM of arm i and P_2 is the TPM of the rest

Notations

- $(A_k)_{k\geq 0}$: arm selection process
- $(X_k)_{k\geq 0}$: observation process
- TPM of odd arm: P_1; TPM of other arms: P_2 (both unknown)
- $C = (i, P_1, P_2)$ denotes a config. of arms in which i is odd arm, P_1 is the TPM of arm i and P_2 is the TPM of the rest
- π denotes a policy

Notations

- $(A_k)_{k\geq 0}$: arm selection process
- $(X_k)_{k\geq 0}$: observation process
- TPM of odd arm: P_1; TPM of other arms: P_2 (both unknown)
- $C = (i, P_1, P_2)$ denotes a config. of arms in which i is odd arm, P_1 is the TPM of arm i and P_2 is the TPM of the rest
- π denotes a policy
- $\tau(\pi)$: stopping time of policy π

Notations

- $(A_k)_{k\geq 0}$: arm selection process
- $(X_k)_{k\geq 0}$: observation process
- TPM of odd arm: P_1; TPM of other arms: P_2 (both unknown)
- $C = (i, P_1, P_2)$ denotes a config. of arms in which i is odd arm, P_1 is the TPM of arm i and P_2 is the TPM of the rest
- π denotes a policy
- $\tau(\pi)$: stopping time of policy π
- $P^{\pi}(\cdot \mid C), E^{\pi}[\cdot \mid C]$: probabilities and expectations under policy π and config C

Notations

- $(A_k)_{k\geq 0}$: arm selection process
- $(X_k)_{k\geq 0}$: observation process
- TPM of odd arm: P_1; TPM of other arms: P_2 (both unknown)
- $C = (i, P_1, P_2)$ denotes a config. of arms in which i is odd arm, P_1 is the TPM of arm i and P_2 is the TPM of the rest
- π denotes a policy
- $\tau(\pi)$: stopping time of policy π
- $P^{\pi}(\cdot \mid C), E^{\pi}[\cdot \mid C]$: probabilities and expectations under policy π and config C

$$Z_{ij}^{\pi}(n) = \log \frac{P_i^{\pi}(X_0, A_0, \ldots, X_n, A_n)}{P_j^{\pi}(X_0, A_0, \ldots, X_n, A_n)}$$

Basic entity to work with

Lower Bound

Lower Bound

$$E^{\pi}[\tau(\pi) \mid C = (i, P_1, P_2)] \gtrapprox \frac{-\log \epsilon}{D^*(i, P_1, P_2)}$$

for all π whose error prob. is $\leq \epsilon$

$$D^*(i, P_1, P_2) = \max_{\lambda} \min_{C'=(j,P_1',P_2'),\ j \neq i} \sum_{a=1}^{K} \lambda(a)\, D(P_i^a \| P_j^a \mid \mu_i^a)$$

Lower Bound

$$E^{\pi}[\tau(\pi) \mid C = (i, P_1, P_2)] \gtrapprox \frac{-\log \epsilon}{D^*(i, P_1, P_2)}$$

for all π whose error prob. is $\leq \epsilon$

$$D^*(i, P_1, P_2) = \max_{\lambda} \min_{C'=(j,P_1',P_2'),\ j \neq i} \sum_{a=1}^{K} \lambda(a)\, D(P_i^a \| P_j^a \mid \mu_i^a)$$

$$D_i^* = \max_{\lambda} \min_{j \neq i} \sum_{k=1}^{m} \lambda(E_k)\, D(P_{(\theta_i, E_k)} \| P_{(\theta_j, E_k)})$$

Chernoff's lower bound

Lower Bound

$$E^{\pi}[\tau(\pi) \mid C = (i, P_1, P_2)] \gtrapprox \frac{-\log \epsilon}{D^*(i, P_1, P_2)}$$

for all π whose error prob. is $\leq \epsilon$

$$D^*(i, P_1, P_2) = \max_{\lambda} \min_{C'=(j,P_1',P_2'),\ j\neq i} \sum_{a=1}^{K} \lambda(a)\, D(P_i^a \| P_j^a \mid \mu_i^a)$$

Finitely many alternatives

Uncountably many alternatives

$$D_i^* = \max_{\lambda} \min_{j\neq i} \sum_{k=1}^{m} \lambda(E_k)\, D(P_{(\theta_i,E_k)} \| P_{(\theta_j,E_k)})$$

Chernoff's lower bound

Lower Bound: Guarding Against The Nearest Incorrect Alternative

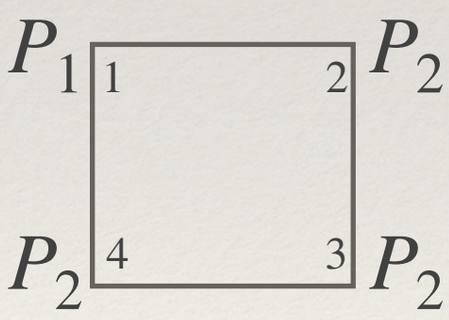
$$E^{\pi}[\tau(\pi) \mid C = (i, P_1, P_2)] \gtrapprox \frac{-\log \epsilon}{D^*(i, P_1, P_2)}$$

$$D^*(i, P_1, P_2) = \max_{\lambda} \min_{C'=(j,P_1',P_2'),\ j \neq i} \sum_{a=1}^{K} \lambda(a)\, D(P_i^a \| P_j^a \mid \mu_i^a)$$

Lower Bound: Guarding Against The Nearest Incorrect Alternative

$$E^{\pi}[\tau(\pi) \mid C = (i, P_1, P_2)] \gtrapprox \frac{-\log \epsilon}{D^*(i, P_1, P_2)}$$

$$D^*(i, P_1, P_2) = \max_{\lambda} \min_{C'=(j,P_1',P_2'),\ j \neq i} \sum_{a=1}^{K} \lambda(a)\, D(P_i^a \| P_j^a \mid \mu_i^a)$$

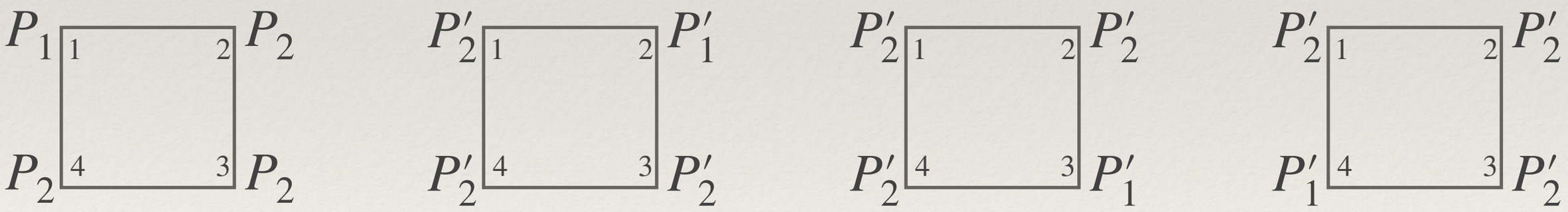


True Configuration

Lower Bound: Guarding Against The Nearest Incorrect Alternative

$$E^{\pi}[\tau(\pi) \mid C = (i, P_1, P_2)] \gtrapprox \frac{-\log \epsilon}{D^*(i, P_1, P_2)}$$

$$D^*(i, P_1, P_2) = \max_{\lambda} \min_{C'=(j,P_1',P_2'),\ j \neq i} \sum_{a=1}^{K} \lambda(a)\, D(P_i^a \| P_j^a \mid \mu_i^a)$$



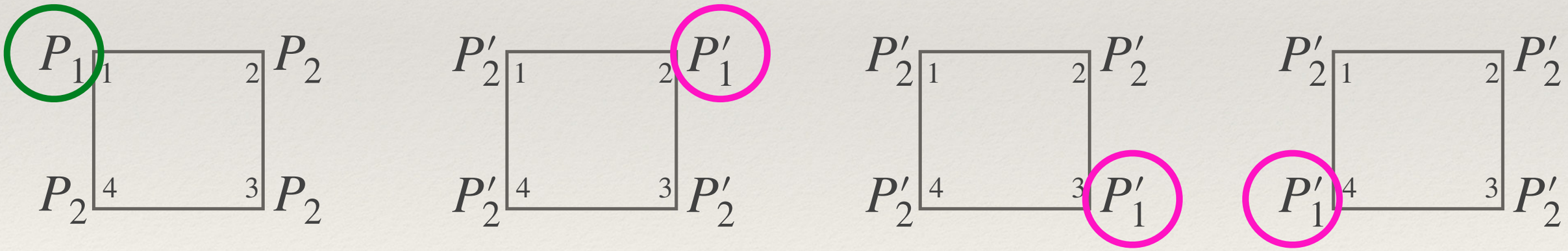
True Configuration

Possible Alternative Configurations

Lower Bound: Guarding Against The Nearest Incorrect Alternative

$$E^{\pi}[\tau(\pi) \mid C = (i, P_1, P_2)] \gtrapprox \frac{-\log \epsilon}{D^*(i, P_1, P_2)}$$

$$D^*(i, P_1, P_2) = \max_{\lambda} \min_{C'=(j,P_1',P_2'),\ j \neq i} \sum_{a=1}^{K} \lambda(a)\, D(P_i^a \| P_j^a \mid \mu_i^a)$$



True Configuration

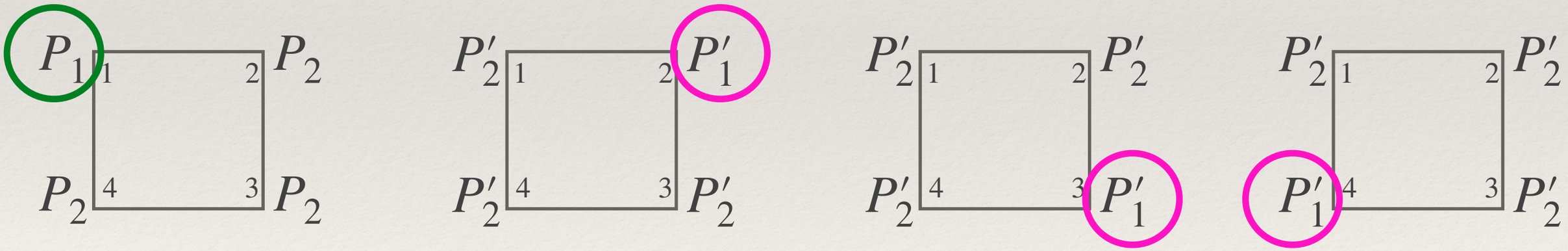
Possible Alternative Configurations

Lower Bound: Guarding Against The Nearest Incorrect Alternative

$$E^{\pi}[\tau(\pi) \mid C = (i, P_1, P_2)] \gtrapprox \frac{-\log \epsilon}{D^*(i, P_1, P_2)}$$

$$D^*(i, P_1, P_2) = \max_{\lambda} \min_{C'=(j,P_1',P_2'),\ j \neq i} \sum_{a=1}^{K} \lambda(a)\, D(P_i^a \| P_j^a \mid \mu_i^a)$$

Which alternative configuration is the `nearest' to the true configuration?



True Configuration

Possible Alternative Configurations

Lower Bound: Guarding Against The Nearest Incorrect Alternative

$$E^{\pi}[\tau(\pi) \mid C = (i, P_1, P_2)] \gtrapprox \frac{-\log \epsilon}{D^*(i, P_1, P_2)}$$

$$D^*(i, P_1, P_2) = \max_{\lambda} \min_{C'=(j,P_1',P_2'),\ j \neq i} \sum_{a=1}^{K} \lambda(a)\, D(P_i^a \| P_j^a \mid \mu_i^a)$$

Lower Bound: Guarding Against The Nearest Incorrect Alternative

$$E^{\pi}[\tau(\pi) \mid C = (i, P_1, P_2)] \gtrapprox \frac{-\log \epsilon}{D^*(i, P_1, P_2)}$$

$$D^*(i, P_1, P_2) = \max_{\lambda} \min_{C'=(j,P_1',P_2'),\ j \neq i} \sum_{a=1}^{K} \lambda(a)\, D(P_i^a \| P_j^a \mid \mu_i^a)$$

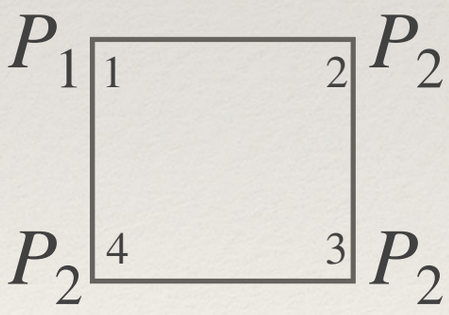
Turns out…..

Lower Bound: Guarding Against The Nearest Incorrect Alternative

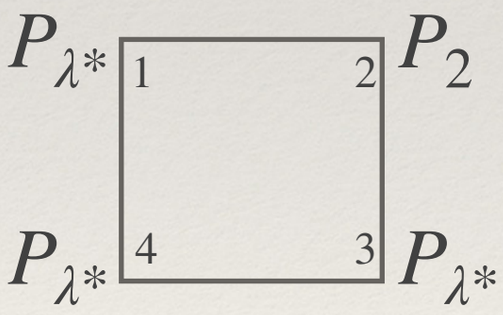
$$E^{\pi}[\tau(\pi) \mid C = (i, P_1, P_2)] \gtrapprox \frac{-\log \epsilon}{D^*(i, P_1, P_2)}$$

$$D^*(i, P_1, P_2) = \max_{\lambda} \min_{C'=(j,P_1',P_2'),\ j \neq i} \sum_{a=1}^{K} \lambda(a)\, D(P_i^a \| P_j^a \mid \mu_i^a)$$

Turns out…..



True Configuration



`Nearest' Alternative Configuration

$$\lambda^* = \lambda^*(i, P_1, P_2) = \arg \max_{0 \leq \lambda_1 \leq 1} \left\{ \lambda_1\, D(P_1 || P_{\lambda_1} \mid \mu_1) + (1 - \lambda_1) \frac{(K-2)}{(K-1)} D(P_2 || P_{\lambda_1} \mid \mu_2) \right\}$$

$$P_{\lambda_1}(y \mid x) = \frac{\lambda_1 \mu_1(x) P_1(y \mid x) + (1 - \lambda_1) \frac{(K-2)}{(K-1)} \mu_2(x) P_2(y \mid x)}{\lambda_1 \mu_1(x) + (1 - \lambda_1) \frac{(K-2)}{(K-1)} \mu_2(x)}$$

Lower Bound: Key Ideas

[1] E. Kaufmann, O. Cappe, and A. Garivier, "On the complexity of best-arm identification in multi-armed bandit models," The Journal of Machine Learning Research, vol. 17, no. 1, pp. 1–42, 2016.

Lower Bound: Key Ideas

- If $C = (i, P_1, P_2)$ then using a result of Kaufmann et al.[1],

$$\min_{C'=(j,P_1',P_2'),\ j\neq i} E^{\pi}[Z_{ij}^{\pi}(\tau(\pi)) \,|\, C] \quad \geq \quad \underbrace{\epsilon \log \frac{\epsilon}{1-\epsilon} + (1-\epsilon) \log \frac{1-\epsilon}{\epsilon}}_{d(\epsilon,1-\epsilon)}$$

[1] E. Kaufmann, O. Cappe, and A. Garivier, "On the complexity of best-arm identification in multi-armed bandit models," The Journal of Machine Learning Research, vol. 17, no. 1, pp. 1–42, 2016.

Lower Bound: Key Ideas

- If $C = (i, P_1, P_2)$ then using a result of Kaufmann et al.[1],

$$\min_{C'=(j,P_1',P_2'),\ j\neq i} E^{\pi}[Z_{ij}^{\pi}(\tau(\pi)) \,|\, C] \quad \geq \quad \underbrace{\epsilon \log \frac{\epsilon}{1-\epsilon} + (1-\epsilon) \log \frac{1-\epsilon}{\epsilon}}_{d(\epsilon,1-\epsilon)}$$

- Wald's identity not applicable. A generalisation of a change of measure argument in Kaufmann et al.[1] to Markov processes used

[1] E. Kaufmann, O. Cappe, and A. Garivier, "On the complexity of best-arm identification in multi-armed bandit models," The Journal of Machine Learning Research, vol. 17, no. 1, pp. 1–42, 2016.

Lower Bound: Key Ideas

- If $C = (i, P_1, P_2)$ then using a result of Kaufmann et al.[1],

$$\min_{C'=(j,P_1',P_2'),\ j\neq i} E^{\pi}[Z_{ij}^{\pi}(\tau(\pi)) \,|\, C] \quad \geq \quad \underbrace{\epsilon \log \frac{\epsilon}{1-\epsilon} + (1-\epsilon) \log \frac{1-\epsilon}{\epsilon}}_{d(\epsilon,1-\epsilon)}$$

- Wald's identity not applicable. A generalisation of a change of measure argument in Kaufmann et al.[1] to Markov processes used
- For any given arm, the long-term fraction of exits from a state i is equal to the long-term fraction of entries to i. This common fraction is the stationary probability of observing i. **This is a consequence of the rested nature of the arms**

[1] **E. Kaufmann, O. Cappe, and A. Garivier, "On the complexity of best-arm identification in multi-armed bandit models," The Journal of Machine Learning Research, vol. 17, no. 1, pp. 1–42, 2016.**

Lower Bound: Key Ideas

- If $C = (i, P_1, P_2)$ then using a result of Kaufmann et al.[1],

$$\min_{C'=(j,P_1',P_2'),\ j\neq i} E^{\pi}[Z_{ij}^{\pi}(\tau(\pi)) \,|\, C] \quad \geq \quad \underbrace{\epsilon \log \frac{\epsilon}{1-\epsilon} + (1-\epsilon) \log \frac{1-\epsilon}{\epsilon}}_{d(\epsilon,1-\epsilon)}$$

- Wald's identity not applicable. A generalisation of a change of measure argument in Kaufmann et al.[1] to Markov processes used
- For any given arm, the long-term fraction of exits from a state i is equal to the long-term fraction of entries to i. This common fraction is the stationary probability of observing i. **This is a consequence of the rested nature of the arms**
- Nearest alternative configuration: $P_1' = P_2,\ P_2' = P_{\lambda^*}$

[1] E. Kaufmann, O. Cappe, and A. Garivier, "On the complexity of best-arm identification in multi-armed bandit models," The Journal of Machine Learning Research, vol. 17, no. 1, pp. 1–42, 2016.

Lower Bound: Key Ideas

- If $C = (i, P_1, P_2)$ then using a result of Kaufmann et al.[1],

$$\min_{C'=(j,P_1',P_2'),\ j\neq i} E^{\pi}[Z_{ij}^{\pi}(\tau(\pi)) \,|\, C] \quad \geq \quad \underbrace{\epsilon \log \frac{\epsilon}{1-\epsilon} + (1-\epsilon) \log \frac{1-\epsilon}{\epsilon}}_{d(\epsilon,1-\epsilon)}$$

- Wald's identity not applicable. A generalisation of a change of measure argument in Kaufmann et al.[1] to Markov processes used
- For any given arm, the long-term fraction of exits from a state i is equal to the long-term fraction of entries to i. This common fraction is the stationary probability of observing i. **This is a consequence of the rested nature of the arms**
- Nearest alternative configuration: $P_1' = P_2,\ P_2' = P_{\lambda^*}$
- $\dfrac{d(\epsilon,1-\epsilon)}{-\log \epsilon} \to 1 \text{ as } \epsilon \downarrow 0$

[1] E. Kaufmann, O. Cappe, and A. Garivier, "On the complexity of best-arm identification in multi-armed bandit models," The Journal of Machine Learning Research, vol. 17, no. 1, pp. 1–42, 2016.

Achievability: Key Ideas

Achievability: Key Ideas

- If the TPMs P_1 and P_2 were known:

Achievability: Key Ideas

- If the TPMs P_1 and P_2 were known:
  - The most natural test statistic (borrowing from Chernoff) would be

$$Z_i(n) = \log \frac{P_i(X_0, A_0, \ldots, X_n, A_n)}{\max\limits_{j \neq i} \; P_j(X_0, A_0, \ldots, X_n, A_n)}$$

$$i^*(n) = \arg\max\{Z_1(n), \ldots, Z_K(n)\}$$

Achievability: Key Ideas

- If the TPMs P_1 and P_2 were known:
  - The most natural test statistic (borrowing from Chernoff) would be

$$Z_i(n) = \log \frac{P_i(X_0, A_0, \ldots, X_n, A_n)}{\max\limits_{j \neq i} \; P_j(X_0, A_0, \ldots, X_n, A_n)}$$

$$i^*(n) = \arg\max\{Z_1(n), \ldots, Z_K(n)\}$$

- The natural algorithm would be: at time n,

Achievability: Key Ideas

- If the TPMs P_1 and P_2 were known:
  - The most natural test statistic (borrowing from Chernoff) would be

$$Z_i(n) = \log \frac{P_i(X_0, A_0, \ldots, X_n, A_n)}{\max\limits_{j \neq i} \; P_j(X_0, A_0, \ldots, X_n, A_n)}$$

$$i^*(n) = \arg\max\{Z_1(n), \ldots, Z_K(n)\}$$

- The natural algorithm would be: at time n,
  - Stop and declare $\mathscr{H}_{i^*(n)}$ true if $Z_{i^*(n)}(n) \geq -\log \epsilon + \log \alpha$
    (the extra $\log \alpha$ term to ensure prob. of error is $\leq \epsilon$)

Achievability: Key Ideas

- If the TPMs P_1 and P_2 were known:
  - The most natural test statistic (borrowing from Chernoff) would be

$$Z_i(n) = \log \frac{P_i(X_0, A_0, \ldots, X_n, A_n)}{\max\limits_{j \neq i} \; P_j(X_0, A_0, \ldots, X_n, A_n)}$$

$$i^*(n) = \arg\max\{Z_1(n), \ldots, Z_K(n)\}$$

- The natural algorithm would be: at time n,
  - Stop and declare $\mathscr{H}_{i^*(n)}$ true if $Z_{i^*(n)}(n) \geq -\log \epsilon + \log \alpha$ (the extra $\log \alpha$ term to ensure prob. of error is $\leq \epsilon$)
  - Else, continue experimentation (select any one arm and take one more observation)

Achievability: Key Ideas

- If the TPMs P_1 and P_2 were known:
  - The most natural test statistic (borrowing from Chernoff) would be

$$Z_i(n) = \log \frac{P_i(X_0, A_0, \ldots, X_n, A_n)}{\max\limits_{j \neq i} \; P_j(X_0, A_0, \ldots, X_n, A_n)}$$

$$i^*(n) = \arg\max\{Z_1(n), \ldots, Z_K(n)\}$$

  - The natural algorithm would be: at time n,
    - Stop and declare $\mathscr{H}_{i^*(n)}$ true if $Z_{i^*(n)}(n) \geq -\log \epsilon + \log \alpha$ (the extra $\log \alpha$ term to ensure prob. of error is $\leq \epsilon$)
    - Else, continue experimentation (select any one arm and take one more observation)
- But P_1 and P_2 are not known

Achievability: Key Ideas

Achievability: Key Ideas

- Note that

$$P_i(X_0, A_0, \ldots, X_n, A_n) \propto \prod_{x,y} \left\{ (P_1(y \mid x))^{N_i(n,x,y)} \quad (P_2(y \mid x))^{\sum_{j \neq i} N_j(n,x,y)} \right\}$$

Achievability: Key Ideas

- Note that

Number of $x \to y$ transitions on arm i up to time n

$$P_i(X_0, A_0, \ldots, X_n, A_n) \propto \prod_{x,y} \left\{ (P_1(y \mid x))^{N_i(n,x,y)} \; (P_2(y \mid x))^{\sum_{j \neq i} N_j(n,x,y)} \right\}$$

Achievability: Key Ideas

- Note that

Number of $x \to y$ transitions on arm i up to time n

$$P_i(X_0, A_0, \ldots, X_n, A_n) \propto \prod_{x,y} \left\{ (P_1(y \mid x))^{N_i(n,x,y)} \; (P_2(y \mid x))^{\sum_{j \neq i} N_j(n,x,y)} \right\}$$

- We do not know P_1 and P_2. One option is to replace P_1 and P_2 above by their ML estimates under hypothesis $\mathscr{H}_i$ for each i. This leads to the classical generalised likelihood ratio test (GLRT) statistic

Achievability: Key Ideas

- Note that

Number of $x \to y$ transitions on arm i up to time n

$$P_i(X_0, A_0, \ldots, X_n, A_n) \propto \prod_{x,y} \left\{ (P_1(y\,|\,x))^{N_i(n,x,y)} \; (P_2(y\,|\,x))^{\sum_{j \neq i} N_j(n,x,y)} \right\}$$

- We do not know P_1 and P_2. One option is to replace P_1 and P_2 above by their ML estimates under hypothesis $\mathscr{H}_i$ for each i. This leads to the classical generalised likelihood ratio test (GLRT) statistic

- Instead of replacing the numerator of $Z_i(n)$ by max as in GLR statistic, we replace it by an *average* computed with respect to an artificial prior on P_1, P_2. We call this *modified GLR statistic*

Policy $\pi^\star(\delta)$: Modified GLR Statistic with Forced Exploration

Policy $\pi^\star(\delta)$: Modified GLR Statistic with Forced Exploration

- So, instead of $Z_i(n)$, we have modified GLR statistic $M_i(n)$ of arm i at time n. Let

  $i^*(n) = \arg\max\{M_1(n), \ldots, M_K(n)\}$

  $\hat{P}_1^n$ = ML estimate of P_1 at time n

  $\hat{P}_2^n$ = ML estimate of P_2 at time n

Policy $\pi^\star(\delta)$: Modified GLR Statistic with Forced Exploration

- So, instead of $Z_i(n)$, we have modified GLR statistic $M_i(n)$ of arm i at time n. Let

  $i^*(n) = \arg\max\{M_1(n), \ldots, M_K(n)\}$

  $\hat{P}_1^n$ = ML estimate of P_1 at time n

  $\hat{P}_2^n$ = ML estimate of P_2 at time n

- At time n:

Policy $\pi^\star(\delta)$: Modified GLR Statistic with Forced Exploration

- So, instead of $Z_i(n)$, we have modified GLR statistic $M_i(n)$ of arm i at time n. Let

  $i^*(n) = \arg\max\{M_1(n), \ldots, M_K(n)\}$

  $\hat{P}_1^n$ = ML estimate of P_1 at time n

  $\hat{P}_2^n$ = ML estimate of P_2 at time n

- At time n:

  - Stop and declare Stop and declare $\mathscr{H}_{i^*(n)}$ true if
    $M_{i^*(n)}(n) \geq -\log\epsilon + \log(K-1)$

Policy $\pi^\star(\delta)$: Modified GLR Statistic with Forced Exploration

- So, instead of $Z_i(n)$, we have modified GLR statistic $M_i(n)$ of arm i at time n. Let

  $i^*(n) = \arg\max\{M_1(n), \ldots, M_K(n)\}$

  $\hat{P}_1^n$ = ML estimate of P_1 at time n

  $\hat{P}_2^n$ = ML estimate of P_2 at time n

- At time n:

  - Stop and declare Stop and declare $\mathscr{H}_{i^*(n)}$ true if
    $M_{i^*(n)}(n) \geq -\log\epsilon + \log(K-1)$

  - Else, continue experimentation by tossing a coin with prob of heads
    $\delta > 0$.

Policy $\pi^{\star}(\delta)$: Modified GLR Statistic with Forced Exploration

- So, instead of $Z_i(n)$, we have modified GLR statistic $M_i(n)$ of arm i at time n. Let

  $i^*(n) = \arg\max\{M_1(n), \ldots, M_K(n)\}$

  $\hat{P}_1^n$ = ML estimate of P_1 at time n

  $\hat{P}_2^n$ = ML estimate of P_2 at time n

- At time n:

  - Stop and declare Stop and declare $\mathcal{H}_{i^*(n)}$ true if
    $M_{i^*(n)}(n) \geq -\log \epsilon + \log(K-1)$ Forced exploration parameter

  - Else, continue experimentation by tossing a coin with prob of heads
    $\delta > 0$.

Policy $\pi^\star(\delta)$: Modified GLR Statistic with Forced Exploration

- So, instead of $Z_i(n)$, we have modified GLR statistic $M_i(n)$ of arm i at time n. Let

  $i^*(n) = \arg\max\{M_1(n), \ldots, M_K(n)\}$

  $\hat{P}_1^n$ = ML estimate of P_1 at time n

  $\hat{P}_2^n$ = ML estimate of P_2 at time n

- At time n:

  - Stop and declare Stop and declare $\mathcal{H}_{i^*(n)}$ true if
    $M_{i^*(n)}(n) \geq -\log \epsilon + \log(K-1)$ Forced exploration parameter

  - Else, continue experimentation by tossing a coin with prob of heads
    $\delta > 0$.

    - If coin falls heads, select the next arm uniformly at random

Policy $\pi^\star(\delta)$: Modified GLR Statistic with Forced Exploration

- So, instead of $Z_i(n)$, we have modified GLR statistic $M_i(n)$ of arm i at time n. Let

  $i^*(n) = \arg\max\{M_1(n), \ldots, M_K(n)\}$

  $\hat{P}_1^n$ = ML estimate of P_1 at time n

  $\hat{P}_2^n$ = ML estimate of P_2 at time n

- At time n:

  - Stop and declare Stop and declare $\mathcal{H}_{i^*(n)}$ true if
    $M_{i^*(n)}(n) \geq -\log \epsilon + \log(K-1)$ Forced exploration parameter

  - Else, continue experimentation by tossing a coin with prob of heads
    $\delta > 0$.

    - If coin falls heads, select the next arm uniformly at random

    - If coin falls tails, select the next arm according to the optimal distribution for the configuration $(i^*(n), \hat{P}_1^n, \hat{P}_2^n)$

Features of Policy $\pi^\star(\delta)$

Features of Policy $\pi^\star(\delta)$

- The policy stops in finite time a.s.

Features of Policy $\pi^\star(\delta)$

- The policy stops in finite time a.s.
- If $C = (i, P_1, P_2)$ is the actual configuration, then a.s.:

$$i^*(n) \to i$$

$$\hat{P}_1^n(y\,|\,x) \to P_1(y\,|\,x) \quad \text{for all } x, y$$

$$\hat{P}_2^n(y\,|\,x) \to P_2(y\,|\,x) \quad \text{for all } x, y$$

Features of Policy $\pi^\star(\delta)$

- The policy stops in finite time a.s.
- If $C = (i, P_1, P_2)$ is the actual configuration, then a.s.:

$$i^*(n) \to i$$

$$\hat{P}_1^n(y\,|\,x) \to P_1(y\,|\,x) \quad \text{for all } x, y$$

$$\hat{P}_2^n(y\,|\,x) \to P_2(y\,|\,x) \quad \text{for all } x, y$$

- Given error probability $\epsilon > 0$, for any $\delta \in (0,1)$,

$$P_e(\pi^\star(\delta)) \leq \epsilon$$

Features of Policy $\pi^\star(\delta)$

- The policy stops in finite time a.s.
- If $C = (i, P_1, P_2)$ is the actual configuration, then a.s.:

$$i^*(n) \to i$$

$$\hat{P}_1^n(y\,|\,x) \to P_1(y\,|\,x) \quad \text{for all } x, y$$

$$\hat{P}_2^n(y\,|\,x) \to P_2(y\,|\,x) \quad \text{for all } x, y$$

- Given error probability $\epsilon > 0$, for any $\delta \in (0,1)$,

$$P_e(\pi^\star(\delta)) \leq \epsilon$$

- **Asymptotic optimality**:

$$\lim_{\delta \downarrow 0} \lim_{\epsilon \downarrow 0} \frac{E[\tau(\pi^\star(\delta))\,|\,(i, P_1, P_2)]}{-\log \epsilon} \leq \frac{1}{D^*(i, P_1, P_2)}$$

What Next?

What Next?

- Restless arms

What Next?

- Restless arms
- Characterising higher order dependencies:

$$E^{\pi}[\tau(\pi) \mid (i, P_1, P_2)] = D^*(i, P_1, P_2) \cdot |\log \epsilon| + ??$$

What Next?

- Restless arms
- Characterising higher order dependencies:

$$E^{\pi}[\tau(\pi)\,|\,(i,P_1,P_2)] = D^*(i,P_1,P_2)\cdot|\log\epsilon| + \text{??}$$

$$C^*(i,P_1,P_2)\cdot|\log|\log\epsilon||$$
$$+o(\log|\log\epsilon|)$$

What Next?

- Restless arms
- Characterising higher order dependencies:

$$E^{\pi}[\tau(\pi)\,|\,(i, P_1, P_2)] = D^*(i, P_1, P_2) \cdot |\log \epsilon| + ??$$

$$C^*(i, P_1, P_2) \cdot ||\log|\log \epsilon|| + o(\log|\log \epsilon|)$$

- Countable and uncountable state spaces

What Next?

- Restless arms
- Characterising higher order dependencies:

$$E^{\pi}[\tau(\pi)\,|\,(i,P_1,P_2)] = D^*(i,P_1,P_2)\cdot|\log\epsilon| + ??$$

$$C^*(i,P_1,P_2)\cdot|\log|\log\epsilon||$$
$$+o(\log|\log\epsilon|)$$

- Countable and uncountable state spaces
- Continuous-time Markov chains (relevant in CR systems)

What Next?

- Restless arms
- Characterising higher order dependencies:

$$E^{\pi}[\tau(\pi)\,|\,(i, P_1, P_2)] = D^*(i, P_1, P_2) \cdot |\log \epsilon| + ??$$

$C^*(i, P_1, P_2) \cdot ||\log|\log \epsilon||$

$+o(\log|\log \epsilon|)$

- Countable and uncountable state spaces
- Continuous-time Markov chains (relevant in CR systems)
- Lower bound as a function of K

What Next?

Sequential Controlled Sensing for Composite Multihypothesis Testing

Aditya Deshmukh Srikrishna Bhashyam Venugopal V. Veeravalli

Abstract

The problem of multi-hypothesis testing with controlled sensing of observations is considered. The distribution of observations collected under each control is assumed to follow a single-parameter exponential family distribution. The goal is to design a policy to find the true hypothesis with minimum expected delay while ensuring that probability of error is below a given constraint. The decision maker can control the delay by intelligently choosing the control for observation collection in each time slot. We derive a policy that satisfies the given constraint on the error probability. We also show that the policy is asymptotically optimal in the sense that it asymptotically achieves an information-theoretic lower bound on the expected delay.

24 Oct 2019 (ArXiv)

What Next?

Sequential Controlled Sensing for Composite Multihypothesis Testing

Aditya Deshmukh Srikrishna Bhashyam Venugopal V. Veeravalli

Abstract

The problem of multi-hypothesis testing with controlled sensing of observations is considered. The distribution of observations collected under each control is assumed to follow a single-parameter exponential family distribution. The goal is to design a policy to find the true hypothesis with minimum expected delay while ensuring that probability of error is below a given constraint. The decision maker can control the delay by intelligently choosing the control for observation collection in each time slot. We derive a policy that satisfies the given constraint on the error probability. We also show that the policy is asymptotically optimal in the sense that it asymptotically achieves an information-theoretic lower bound on the expected delay.

24 Oct 2019 (ArXiv)

- A general framework for dealing with sequential hypothesis testing in multi-armed bandits

What Next?

Sequential Controlled Sensing for Composite Multihypothesis Testing

Aditya Deshmukh Srikrishna Bhashyam Venugopal V. Veeravalli

Abstract

The problem of multi-hypothesis testing with controlled sensing of observations is considered. The distribution of observations collected under each control is assumed to follow a single-parameter exponential family distribution. The goal is to design a policy to find the true hypothesis with minimum expected delay while ensuring that probability of error is below a given constraint. The decision maker can control the delay by intelligently choosing the control for observation collection in each time slot. We derive a policy that satisfies the given constraint on the error probability. We also show that the policy is asymptotically optimal in the sense that it asymptotically achieves an information-theoretic lower bound on the expected delay.

24 Oct 2019 (ArXiv)

- A general framework for dealing with sequential hypothesis testing in multi-armed bandits
- Generalises Chernoff's 1959 framework to include a large class of problems such as OAI, best-arm identification, 2nd best arm identification and so on, when observations are iid and coming from an exponential family

What Next?

Sequential Controlled Sensing for Composite Multihypothesis Testing

Aditya Deshmukh Srikrishna Bhashyam Venugopal V. Veeravalli

Abstract

The problem of multi-hypothesis testing with controlled sensing of observations is considered. The distribution of observations collected under each control is assumed to follow a single-parameter exponential family distribution. The goal is to design a policy to find the true hypothesis with minimum expected delay while ensuring that probability of error is below a given constraint. The decision maker can control the delay by intelligently choosing the control for observation collection in each time slot. We derive a policy that satisfies the given constraint on the error probability. We also show that the policy is asymptotically optimal in the sense that it asymptotically achieves an information-theoretic lower bound on the expected delay.

24 Oct 2019 (ArXiv)

- A general framework for dealing with sequential hypothesis testing in multi-armed bandits
- Generalises Chernoff's 1959 framework to include a large class of problems such as OAI, best-arm identification, 2nd best arm identification and so on, when observations are iid and coming from an exponential family
- Extensions to Markov (rested/restless)?

What Next?

Sequential Controlled Sensing for Composite Multihypothesis Testing

Aditya Deshmukh Srikrishna Bhashyam Venugopal V. Veeravalli

Abstract

The problem of multi-hypothesis testing with controlled sensing of observations is considered. The distribution of observations collected under each control is assumed to follow a single-parameter exponential family distribution. The goal is to design a policy to find the true hypothesis with minimum expected delay while ensuring that probability of error is below a given constraint. The decision maker can control the delay by intelligently choosing the control for observation collection in each time slot. We derive a policy that satisfies the given constraint on the error probability. We also show that the policy is asymptotically optimal in the sense that it asymptotically achieves an information-theoretic lower bound on the expected delay.

24 Oct 2019 (ArXiv)

- A general framework for dealing with sequential hypothesis testing in multi-armed bandits
- Generalises Chernoff's 1959 framework to include a large class of problems such as OAI, best-arm identification, 2nd best arm identification and so on, when observations are iid and coming from an exponential family
- Extensions to Markov (rested/restless)?

Thank you!