Log-Optimality & Multi-Armed Sequential Hypothesis Testing

Ian Waudby-Smith

Miller Institute & Department of Statistics
University of California, Berkeley

University of California, Davis Statistics Seminar, 2026

Ricardo J. Sandoval

UC Berkeley

Michael I. Jordan

UC Berkeley & Inria

Outline

1. What is sequential hypothesis testing (by betting)?
2. How are sequential hypothesis tests derived?
3. Defining and deriving optimal sequential tests.
4. Multi-armed sequential hypothesis testing

Outline

1. **What is sequential hypothesis testing (by betting)?**
2. How are sequential hypothesis tests derived?
3. Defining and deriving optimal sequential tests.
4. Multi-armed sequential hypothesis testing

A motivating example to keep in mind: **treatment effect testing**.

Question: Is the new treatment effective?

Question: Is the new treatment effective?

Step 1:

$$H_0 : \text{trt effect} = 0$$
$$H_1 : \text{trt effect} \neq 0$$

$$\alpha := 0.01$$

Step 2:

Recruit n patients and randomize (**trt** or **ctrl**).

Question: Is the new treatment effective?

Step 1:

$$H_0 : \text{trt effect} = 0$$
$$H_1 : \text{trt effect} \neq 0$$

$$\alpha := 0.01$$

Step 2:

Recruit n patients and randomize (**trt** or **ctrl**).

Step 3:

Compute a p-value: $p_n = 0.023$

$\implies$ **Cannot reject** H_0 **at** $\alpha = 0.01$

Question: Is the new treatment effective?

Step 1:

$$H_0 : \text{trt effect} = 0$$
$$H_1 : \text{trt effect} \neq 0$$

$$\alpha := 0.01$$

Step 2:

Recruit n patients and randomize (**trt** or **ctrl**).

Step 3:

Compute a p-value: $p_n = 0.023$

$\implies$ **Cannot reject** H_0 **at** $\alpha = 0.01$

Step 4:

Recruit just a few more patients?

Question: Is the new treatment effective?

Step 1:

$$H_0 : \text{trt effect} = 0$$
$$H_1 : \text{trt effect} \neq 0$$

$$\alpha := 0.01$$

Step 2:

Recruit n patients and randomize (**trt** or **ctrl**).

Step 3:

Compute a p-value: $p_n = 0.023$

$\implies$ **Cannot reject** H_0 **at** $\alpha = 0.01$

Step 4:

Recruit just a few more patients?

No! (This is "p-hacking".)

Alternatively, even if n was large enough to reject H_0, it is possible that $n' \ll n$ could have sufficed (time & money saved).

$\rightarrow$ e.g. $p_n < 0.00001 \ll \alpha = 0.01$.

Alternatively, even if n was large enough to reject H_0, it is possible that $n' \ll n$ could have sufficed (time & money saved).

→ e.g. $p_n < 0.00001 \ll \alpha = 0.01$.

Sequential testing addresses these unsettling scenarios.

What is sequential hypothesis testing?

What is sequential hypothesis testing?

(Foundations laid by Wald in the '40s, Robbins & co in the '60s-'70s)

What is sequential hypothesis testing?

(Foundations laid by Wald in the '40s, Robbins & co in the '60s-'70s)

(Modern breakthroughs by Grünwald, Ramdas, and others in 2010's onwards)

Throughout, fix a composite null $\mathcal{P}$ and a composite alternative $\mathcal{Q}$.

$$\textit{(e.g. } \mathcal{P} = \{P : \textit{trt effect} = 0\} \textit{ versus } \mathcal{Q} = \{P : \textit{trt effect} > 0\}\textit{)}$$

We are tasked with finding a test $\phi_n^{(\alpha)} \equiv \phi^{(\alpha)}(X_1, \ldots, X_n)$ that outputs 1 (rejects $\mathcal{P}$ in favour of $\mathcal{Q}$) with small probability under $\mathcal{P}$.

Throughout, fix a composite null $\mathcal{P}$ and a composite alternative $\mathcal{Q}$.

$$\textit{(e.g. } \mathcal{P} = \{P : \textit{trt effect} = 0\} \textit{ versus } \mathcal{Q} = \{P : \textit{trt effect} > 0\}\textit{)}$$

We are tasked with finding a test $\phi_n^{(\alpha)} \equiv \phi^{(\alpha)}(X_1, \dots, X_n)$ that outputs 1 (rejects $\mathcal{P}$ in favour of $\mathcal{Q}$) with small probability under $\mathcal{P}$.

Fixed-n test: $\quad \forall n \in \mathbb{N}, \ \sup_{\mathsf{P} \in \mathcal{P}} \mathsf{P}\left(\phi_n^{(\alpha)} \text{ rejects}\right) \leqslant \alpha.$

Sequential test: $\quad \sup_{\mathsf{P} \in \mathcal{P}} \mathsf{P}\left(\exists n \in \mathbb{N} : \phi_n^{(\alpha)} \text{ rejects}\right) \leqslant \alpha.$

Throughout, fix a composite null $\mathcal{P}$ and a composite alternative $\mathcal{Q}$.

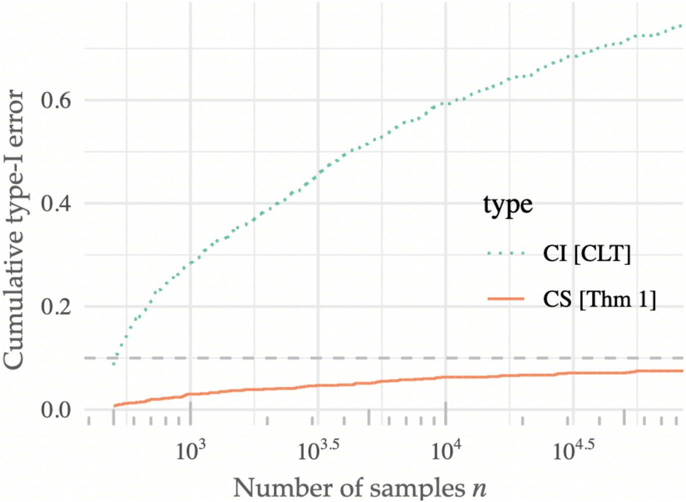
(*e.g.* $\mathcal{P} = \{P : \textit{trt effect} = 0\}$ *versus* $\mathcal{Q} = \{P : \textit{trt effect} > 0\}$)

We are tasked with finding a test $\phi_n^{(\alpha)} \equiv \phi^{(\alpha)}(X_1, \ldots, X_n)$ that outputs 1 (rejects $\mathcal{P}$ in favour of $\mathcal{Q}$) with small probability under $\mathcal{P}$.

Fixed-n test: $\forall n \in \mathbb{N},\ \sup_{\mathsf{P} \in \mathcal{P}} \mathsf{P}\left(\phi_n^{(\alpha)} \text{ rejects}\right) \leqslant \alpha.$

Sequential test: $\sup_{\mathsf{P} \in \mathcal{P}} \mathsf{P}\left(\exists n \in \mathbb{N} : \phi_n^{(\alpha)} \text{ rejects}\right) \leqslant \alpha.$

$$\iff \sup_{\mathsf{P} \in \mathcal{P}} \mathsf{P}\left(\phi_\tau^{(\alpha)} \text{ rejects}\right) \leqslant \alpha \quad \forall \tau.$$



$$\forall n,\ \mathsf{P}(\phi_n^{(\alpha)} = 1) \leqslant \alpha$$

$$\mathsf{P}(\exists n : \phi_n^{(\alpha)} = 1) \leqslant \alpha$$

Outline

1. What is sequential hypothesis testing (by betting)?

2. How are sequential hypothesis tests derived?

3. Defining and deriving optimal sequential tests.

4. Multi-armed sequential hypothesis testing

Suppose data come in an online stream: $X_1, X_2, \cdots \sim \mathsf{P}$, $\mathsf{P} \in \mathcal{P} \cup \mathcal{Q}$.

Suppose data come in an online stream: $X_1, X_2, \cdots \sim \mathsf{P}$, $\mathsf{P} \in \mathcal{P} \cup \mathcal{Q}$.

Sequential tests result from the following two-step procedure:

Suppose data come in an online stream: $X_1, X_2, \cdots \sim \mathsf{P}$, $\mathsf{P} \in \mathcal{P} \cup \mathcal{Q}$.

Sequential tests result from the following two-step procedure:

1. Find $W_n \equiv W(X_1, \ldots, X_n)$ that forms a nonnegative P-supermartingale with mean $\mathbb{E}_{\mathsf{P}}[W_1] \leqslant 1$ for every $\mathsf{P} \in \mathcal{P}$.

   We often call these **test supermartingales**.

Suppose data come in an online stream: $X_1, X_2, \dots \sim \mathsf{P}$, $\mathsf{P} \in \mathcal{P} \cup \mathcal{Q}$.

Sequential tests result from the following two-step procedure:

1. Find $W_n \equiv W(X_1, \dots, X_n)$ that forms a nonnegative P-supermartingale with mean $\mathbb{E}_{\mathsf{P}}[W_1] \leqslant 1$ for every $\mathsf{P} \in \mathcal{P}$.

   We often call these **test supermartingales**.

$$\left(\textit{Example: likelihood ratios. } W_n = \prod_{i=1}^{n} \frac{\mathrm{d}\mathsf{Q}(X_i)}{\mathrm{d}\mathsf{P}(X_i)} \right)$$

Suppose data come in an online stream: $X_1, X_2, \cdots \sim \mathsf{P}$, $\mathsf{P} \in \mathcal{P} \cup \mathcal{Q}$.

Sequential tests result from the following two-step procedure:

1. Find $W_n \equiv W(X_1, \ldots, X_n)$ that forms a nonnegative P-supermartingale with mean $\mathbb{E}_{\mathsf{P}}[W_1] \leqslant 1$ for every $\mathsf{P} \in \mathcal{P}$.

   We often call these **test supermartingales**.

$$\left(\textit{Example: likelihood ratios. } W_n = \prod_{i=1}^{n} \frac{\mathrm{d}\mathsf{Q}(X_i)}{\mathrm{d}\mathsf{P}(X_i)} \right)$$

2. Set the test as $\phi_n^{(\alpha)} := \mathbb{1}\{W_n \geqslant 1/\alpha\}$.

This procedure works because of Ville's inequality which states that for a nonnegative $\mathcal{P}$-supermartingale $(W_n)_{n \in \mathbb{N}}$,

$$\forall x > 0, \quad \mathsf{P}\left(\exists n \in \mathbb{N} : W_n \geqslant x\right) \leqslant \frac{\mathbb{E}_{\mathsf{P}}[W_1]}{x}.$$

This procedure works because of Ville's inequality which states that for a nonnegative $\mathcal{P}$-supermartingale $(W_n)_{n\in\mathbb{N}}$,

$$\forall x > 0, \quad \mathsf{P}\left(\exists n \in \mathbb{N} : W_n \geqslant x\right) \leqslant \frac{\mathbb{E}_{\mathsf{P}}[W_1]}{x}.$$

When $x = 1/\alpha$ and $\mathbb{E}_{\mathsf{P}}[W_1] \leqslant 1$, we get

$$\forall \alpha \in (0,1), \quad \mathsf{P}\Big(\exists n \in \mathbb{N} : W_n \geqslant 1/\alpha\Big) \leqslant \alpha.$$

This procedure works because of Ville's inequality which states that for a nonnegative $\mathcal{P}$-supermartingale $(W_n)_{n\in\mathbb{N}}$,

$$\forall x > 0, \quad \mathsf{P}\left(\exists n \in \mathbb{N} : W_n \geqslant x\right) \leqslant \frac{\mathbb{E}_{\mathsf{P}}[W_1]}{x}.$$

When $x = 1/\alpha$ and $\mathbb{E}_{\mathsf{P}}[W_1] \leqslant 1$, we get

$$\forall \alpha \in (0,1), \quad \mathsf{P}\left(\exists n \in \mathbb{N} : \phi_n^{(\alpha)} = 1\right) \leqslant \alpha.$$

So, does this mean it is “easy” to come up with sequential tests? **No**.

It is often challenging to devise nontrivial supermartingales.

So, does this mean it is “easy” to come up with sequential tests? **No**.

It is often challenging to devise nontrivial supermartingales.

However, lots of progress has been made in recent years.

Example: Testing the mean of a bounded random variables.

(Vovk & Shafer (2001), Hendriks (2018), ***W-S*** *& Ramdas (2023), Orabona & Jun (2023), others)*

Example: Testing the mean of a bounded random variables.

*(Vovk & Shafer (2001), Hendriks (2018), **W-S** & Ramdas (2023), Orabona & Jun (2023), others)*

Let $X_1, X_2, \dots \in [0,1]$. The null is $\mathcal{P} := \{\mathsf{P} : \mathbb{E}_{\mathsf{P}}[X_1] = 1/2\}$.

Example: Testing the mean of a bounded random variables.

(Vovk & Shafer (2001), Hendriks (2018), ***W-S*** *& Ramdas (2023), Orabona & Jun (2023), others)*

Let $X_1, X_2, \dots \in [0,1]$. The null is $\mathcal{P} := \{\mathsf{P} : \mathbb{E}_{\mathsf{P}}[X_1] = 1/2\}$. Then,

$$W_n := \prod_{i=1}^{n} (1 + \gamma_i \cdot (2X_i - 1))$$

forms a test martingale for any $[0,1]$-valued *predictable* $(\gamma_n)_{n \in \mathbb{N}}$.

(*Quick definition of "predictable"*: $\gamma_i \in \sigma(X_1, \dots, X_{i-1})$).

Example: Testing the mean of a bounded random variables.

*(Vovk & Shafer (2001), Hendriks (2018), **W-S** & Ramdas (2023), Orabona & Jun (2023), others)*

Let $X_1, X_2, \dots \in [0,1]$. The null is $\mathcal{P} := \{\mathsf{P} : \mathbb{E}_{\mathsf{P}}[X_1] = 1/2\}$. Then,

$$W_n := \prod_{i=1}^{n} (1 + \gamma_i \cdot (2X_i - 1))$$

forms a test martingale for any $[0,1]$-valued *predictable* $(\gamma_n)_{n \in \mathbb{N}}$.

(*Quick definition of "predictable"*: $\gamma_i \in \sigma(X_1, \dots, X_{i-1})$).

Therefore, $\phi_n^{(\alpha)} := \mathbb{1}\{W_n \geqslant 1/\alpha\}$ yields a sequential test for $\mathcal{P}$.

Example: Testing the mean of a bounded random variables.

(Vovk & Shafer (2001), Hendriks (2018), ***W-S*** *& Ramdas (2023), Orabona & Jun (2023), others)*

Let $X_1, X_2, \dots \in [0,1]$. The null is $\mathcal{P} := \{\mathsf{P} : \mathbb{E}_{\mathsf{P}}[X_1] = 1/2\}$. Then,

$$W_n := \prod_{i=1}^{n} (1 + \gamma_i \cdot (2X_i - 1))$$

forms a test martingale for any $[0,1]$-valued *predictable* $(\gamma_n)_{n\in\mathbb{N}}$.

(*Quick definition of "predictable"*: $\gamma_i \in \sigma(X_1, \dots, X_{i-1})$).

Therefore, $\phi_n^{(\alpha)} := \mathbb{1}\{W_n \geqslant 1/\alpha\}$ yields a sequential test for $\mathcal{P}$.

It is **not** obvious how to choose γ_i... more on that later.

One can re-cast several testing problems *(bounded means, two-sample, independence, equality of bounded tuples, testing randomness online, etc.)* from the literature with a single test supermartingale.

One can re-cast several testing problems *(bounded means, two-sample, independence, equality of bounded tuples, testing randomness online, etc.)* from the literature with a single test supermartingale.

$$W_n := \prod_{i=1}^{n} \lambda_i^\top \mathbf{E}_i,$$

One can re-cast several testing problems *(bounded means, two-sample, independence, equality of bounded tuples, testing randomness online, etc.)* from the literature with a single test supermartingale.

$$W_n := \prod_{i=1}^{n} \lambda_i{}^\top \mathbf{E}_i,$$

$$\text{where} \quad \mathbf{E}_i = (E_0, \dots, E_d) \quad \text{is a tuple of } \mathcal{P}\text{-e-values,}$$

One can re-cast several testing problems *(bounded means, two-sample, independence, equality of bounded tuples, testing randomness online, etc.)* from the literature with a single test supermartingale.

$$W_n := \prod_{i=1}^{n} \lambda_i^\top \mathbf{E}_i,$$

$$\text{where} \quad \mathbf{E}_i = (E_0, \ldots, E_d) \quad \text{is a tuple of } \mathcal{P}\text{-e-values,}$$

$$\text{and} \quad \lambda_i \in \triangle_d \quad \text{depends on } \mathbf{E}_1, \ldots, \mathbf{E}_{i-1}.$$

One can re-cast several testing problems *(bounded means, two-sample, independence, equality of bounded tuples, testing randomness online, etc.)* from the literature with a single test supermartingale.

$$W_n := \prod_{i=1}^{n} \lambda_i^\top \mathbf{E}_i,$$

where $\mathbf{E}_i = (E_0, \ldots, E_d)$ is a tuple of $\mathcal{P}$-e-values,

and $\lambda_i \in \triangle_d$ depends on $\mathbf{E}_1, \ldots, \mathbf{E}_{i-1}$.

Definition: e-value. (See Shafer & Vovk, Grünwald et al., Vovk & Wang)

A nonnegative random variable $E \geqslant 0$ is said to be a $\mathcal{P}$-e-value if

$$\sup_{\mathsf{P} \in \mathcal{P}} \mathbb{E}_{\mathsf{P}}[E] \leqslant 1.$$

Example: For the testing-means-of-bounded-random-variables problem,

$$\prod_{i=1}^{n}(1 + \gamma_i(2X_i - 1))$$

is a special case of

$$\prod_{i=1}^{n} \lambda_i{}^{\top}\mathbf{E}_i$$

when taking $\mathbf{E}_i = (1, 2X_i)$ and $\lambda_i = (1 - \gamma_i, \gamma_i)$.

Claim: The more abstract process $W_n := \prod_{i=1}^{n} \lambda_i{}^{\top} \mathbf{E}_i$ forms a test supermartingale.

Claim: The more abstract process $W_n := \prod_{i=1}^{n} \lambda_i{}^\top \mathbf{E}_i$ forms a test supermartingale.

(1) It is nonnegative because $\lambda_i{}^\top \mathbf{E}_i$ is a convex combination.

Claim: The more abstract process $W_n := \prod_{i=1}^{n} \lambda_i^\top \mathbf{E}_i$ forms a test supermartingale.

(1) It is nonnegative because $\lambda_i^\top \mathbf{E}_i$ is a convex combination.

(2) Need to show $\mathbb{E}_{\mathsf{P}}[W_n \mid \mathbf{E}_1, \ldots, \mathbf{E}_{n-1}] \leqslant W_{n-1}$ for $\mathsf{P} \in \mathcal{P}$.

Claim: The more abstract process $W_n := \prod_{i=1}^{n} \lambda_i^\top \mathbf{E}_i$ forms a test supermartingale.

(1) It is nonnegative because $\lambda_i^\top \mathbf{E}_i$ is a convex combination.

(2) Need to show $\mathbb{E}_{\mathsf{P}}[W_n \mid \mathbf{E}_1, \ldots, \mathbf{E}_{n-1}] \leqslant W_{n-1}$ for $\mathsf{P} \in \mathcal{P}$.

$$\begin{aligned} &\mathbb{E}_{\mathsf{P}}\left[\prod_{i=1}^{n} \lambda_i^\top \mathbf{E}_i \mid \mathbf{E}_1, \ldots, \mathbf{E}_{n-1}\right] \\ = &\left(\prod_{i=1}^{n-1} \lambda_i^\top \mathbf{E}_i\right) \mathbb{E}_{\mathsf{P}}\left[\lambda_n^\top \mathbf{E}_n \mid \mathbf{E}_1, \ldots, \mathbf{E}_{n-1}\right] \\ \leqslant &\ W_{n-1}. \end{aligned}$$

Claim: The more abstract process $W_n := \prod_{i=1}^{n} \lambda_i{}^\top \mathbf{E}_i$ forms a test supermartingale.

(1) It is nonnegative because $\lambda_i{}^\top \mathbf{E}_i$ is a convex combination.

(2) Need to show $\mathbb{E}_{\mathsf{P}}[W_n \mid \mathbf{E}_1, \ldots, \mathbf{E}_{n-1}] \leqslant W_{n-1}$ for $\mathsf{P} \in \mathcal{P}$.

$$\begin{aligned}
&\mathbb{E}_{\mathsf{P}}\left[\prod_{i=1}^{n} \lambda_i{}^\top \mathbf{E}_i \mid \mathbf{E}_1, \ldots, \mathbf{E}_{n-1}\right] \\
&= \left(\prod_{i=1}^{n-1} \lambda_i{}^\top \mathbf{E}_i\right) \mathbb{E}_{\mathsf{P}}\left[\lambda_n{}^\top \mathbf{E}_n \mid \mathbf{E}_1, \ldots, \mathbf{E}_{n-1}\right] \\
&\leqslant W_{n-1}.
\end{aligned}$$

Therefore, W_n forms a test supermartingale. □

Some other special cases found in the literature

One-sided bounded mean testing: Set $\mathbf{E}_i = (1, X_i/\mu_0)$.
Two-sided bounded mean testing: Set $\mathbf{E}_i = ((1 - X_i)/(1 - \mu_0), X_i/\mu_0)$.
Two-sample testing: Set $\mathbf{E}_i = (1, g^\star(X_i) - g^\star(Y_i))$ for a witness f'n $g^\star$.
Testing randomness: Set $\mathbf{E}_i = (2(1 - s_i), 2s_i)$ for a conformal score s_i.

Some other special cases found in the literature

One-sided bounded mean testing: Set $\mathbf{E}_i = (1, X_i/\mu_0)$.
Two-sided bounded mean testing: Set $\mathbf{E}_i = ((1 - X_i)/(1 - \mu_0), X_i/\mu_0)$.
Two-sample testing: Set $\mathbf{E}_i = (1, g^\star(X_i) - g^\star(Y_i))$ for a witness f'n $g^\star$.
Testing randomness: Set $\mathbf{E}_i = (2(1 - s_i), 2s_i)$ for a conformal score s_i.

There has been one lingering question this entire discussion:

How should one choose $(\lambda_n)_{n\in\mathbb{N}}$?

Outline

1. What is sequential hypothesis testing (by betting)?
2. How are sequential hypothesis tests derived?
3. **Defining and deriving optimal sequential tests.**
4. Multi-armed sequential hypothesis testing

“Power” in the traditional sense is not the right concept for sequential tests.

“Power” in the traditional sense is not the right concept for sequential tests.

There are two common power desiderata for sequential tests:

(i) **Growth-rate optimality**

"Power" in the traditional sense is not the right concept for sequential tests.

There are two common power desiderata for sequential tests:

(i) **Growth-rate optimality**

(ii) **Small expected rejection times**

"Power" in the traditional sense is not the right concept for sequential tests.

There are two common power desiderata for sequential tests:

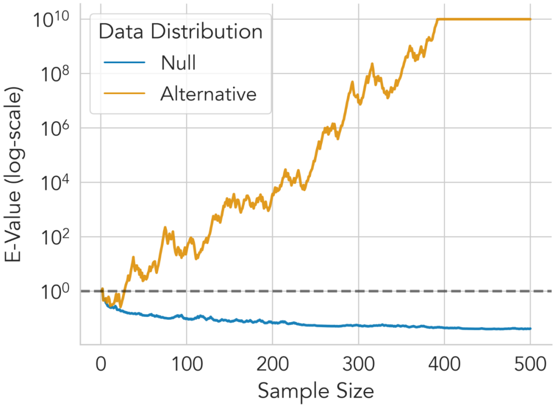
(i) **Growth-rate optimality**

(ii) **Small expected rejection times**

We show that these are optimized via the exact same criterion.

(i) Growth-rate-optimality.

(Kelly ['56], Long Jr. ['90], Grünwald et al. [2024], Larsson et al. [2024])



An e-process is expected to be small under the ***null****;*
we want it to grow large under the ***alternative****.*

Image credit: YJ Choe.

Since we *reject* the null $\mathcal{P}$ when $W_n \geqslant 1/\alpha$, we aim to choose $(\lambda_n)_{n\in\mathbb{N}}$ so that W_n diverges "quickly".

Since we *reject* the null $\mathcal{P}$ when $W_n \geqslant 1/\alpha$, we aim to choose $(\lambda_n)_{n\in\mathbb{N}}$ so that W_n diverges "quickly".

Observe by the strong law of large numbers:

$$\begin{aligned} W_n &= \exp\left\{ n \cdot \frac{1}{n} \sum_{i=1}^{n} \log\left(\lambda_i{}^\top \mathbf{E}_i \right) \right\} \\ &\approx \exp\left\{ n \cdot \mathbb{E}_{\mathsf{Q}} \left[\log\left(\lambda_i{}^\top \mathbf{E}_i \right) \right] + o(n) \right\} \end{aligned}$$

Since we *reject* the null $\mathcal{P}$ when $W_n \geqslant 1/\alpha$, we aim to choose $(\lambda_n)_{n\in\mathbb{N}}$ so that W_n diverges "quickly".

Observe by the strong law of large numbers:

$$\begin{aligned} W_n &= \exp\left\{ n \cdot \frac{1}{n} \sum_{i=1}^{n} \log\left(\lambda_i{}^\top \mathbf{E}_i\right) \right\} \\ &\approx \exp\left\{ n \cdot \mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda_i{}^\top \mathbf{E}_i\right)\right] + o(n) \right\} \end{aligned}$$

So, if data comes from $\mathsf{Q} \in \mathcal{Q}$, it is reasonable to want to choose λ_i so as to maximize:

$$\max_{\lambda \in \triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log(\lambda^\top \mathbf{E}_1)\right].$$

Since we *reject* the null $\mathcal{P}$ when $W_n \geqslant 1/\alpha$, we aim to choose $(\lambda_n)_{n\in\mathbb{N}}$ so that W_n diverges "quickly".

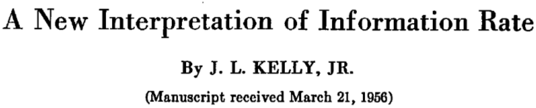
Observe by the strong law of large numbers:

$$W_n = \exp\left\{n \cdot \frac{1}{n}\sum_{i=1}^{n} \log\left(\lambda_i^\top \mathbf{E}_i\right)\right\}$$
$$\approx \exp\left\{n \cdot \mathbb{E}_\mathsf{Q}\left[\log\left(\lambda_i^\top \mathbf{E}_i\right)\right] + o(n)\right\}$$

So, if data comes from $\mathsf{Q} \in \mathcal{Q}$, it is reasonable to want to choose λ_i so as to maximize:

$$\max_{\lambda \in \triangle_d} \mathbb{E}_\mathsf{Q}\left[\log(\lambda^\top \mathbf{E}_1)\right].$$

This is the "Kelly criterion" from gambling / information theory.


A New Interpretation of Information Rate

By J. L. KELLY, JR.

(Manuscript received March 21, 1956)

“Kelly bet”: $\lambda_{\mathsf{Q}} := \underset{\lambda \in \triangle_d}{\operatorname{argmax}} \, \mathbb{E}_{\mathsf{Q}} \left[\log(\lambda^\top \mathbf{E}_1) \right].$

(ii) Measuring optimality through expected rejection times

(Wald 1945, Breiman 1961, Kaufmann, Agrawal, Koolen, others from the bandit literature)

Define the first rejection time for the null hypothesis $\mathcal{P}$:

$$\tau_\alpha := \inf\left\{n \in \mathbb{N} : W_n \geqslant \frac{1}{\alpha}\right\}.$$

Define the first rejection time for the null hypothesis $\mathcal{P}$:

$$\tau_\alpha := \inf\left\{n \in \mathbb{N} : W_n \geqslant \frac{1}{\alpha}\right\}.$$

Since τ_α is a random variable, let us consider its (normalized) Q-expectation: $\mathbb{E}_{\mathsf{Q}}[\tau_\alpha]$.

Define the first rejection time for the null hypothesis $\mathcal{P}$:

$$\tau_\alpha := \inf\left\{n \in \mathbb{N} : W_n \geqslant \frac{1}{\alpha}\right\}.$$

Since τ_α is a random variable, let us consider its (normalized) Q-expectation: $\mathbb{E}_{\mathsf{Q}}[\tau_\alpha]$.

Proposition: Lower bound on the expected stopping time

$$\mathbb{E}_{\mathsf{Q}}[\tau_\alpha] \geqslant \frac{\log(1/\alpha)}{\max_{\lambda \in \triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1\right)\right]}.$$

We have set ourselves up well to optimize this criterion.

Scholars in mathematical finance (and information theory) have studied related problems.

We have set ourselves up well to optimize this criterion.

Scholars in mathematical finance (and information theory) have studied related problems.

Sequential testing as portfolio selection over $d+1$ stocks.

You start with $W_1 = \$1$.

For $n = 1, 2, \ldots$:

1. Choose portfolio $\lambda_n \in \triangle_d$.
2. Observe stock returns $\mathbf{E}_n = (E_n^{(0)}, \ldots, E_n^{(d)}) \in [0, \infty)^d$.
3. $W_n = W_{n-1} \lambda_n^\top \mathbf{E}_n$.

EndFor

We have set ourselves up well to optimize this criterion.

Scholars in mathematical finance (and information theory) have studied related problems.

Sequential testing as portfolio selection over $d + 1$ stocks.

You start with $W_1 = \$1$.

For $n = 1, 2, ...$:

1. Choose portfolio $\lambda_n \in \triangle_d$.
2. Observe stock returns $\mathbf{E}_n = (E_n^{(0)}, \ldots, E_n^{(d)}) \in [0, \infty)^d$.
3. $W_n = W_{n-1} \lambda_n{}^\top \mathbf{E}_n$.

EndFor

Then $W_n = \prod_{i=1}^n \lambda_i{}^\top \mathbf{E}_i$.

Given the form

$$W_n \approx \exp\left\{n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1\right)\right] + o(1)\right\}$$

and the lower bound

$$\mathbb{E}_{\mathsf{Q}}[\tau_\alpha] \geqslant \frac{\log(1/\alpha)}{\max_{\lambda\in\triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1\right)\right]},$$

Given the form

$$W_n \approx \exp\left\{n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1\right)\right] + o(1)\right\}$$

and the lower bound

$$\mathbb{E}_{\mathsf{Q}}[\tau_\alpha] \geqslant \frac{\log(1/\alpha)}{\max_{\lambda\in\triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1\right)\right]},$$

the problem of deriving a "powerful" test has been reduced to the problem of deriving a *log-profitable portfolio* $\lambda_{\mathsf{Q}} \in \triangle_d$.

Given the form

$$W_n \approx \exp\left\{n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1\right)\right] + o(1)\right\}$$

and the lower bound

$$\mathbb{E}_{\mathsf{Q}}[\tau_\alpha] \geqslant \frac{\log(1/\alpha)}{\max_{\lambda \in \triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1\right)\right]},$$

the problem of deriving a "powerful" test has been reduced to the problem of deriving a *log-profitable portfolio* $\lambda_{\mathsf{Q}} \in \triangle_d$.

However, λ_{Q} depends on the alternative Q. Can it be "learned"?

Given the form

$$W_n \approx \exp\left\{n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1\right)\right] + o(1)\right\}$$

and the lower bound

$$\mathbb{E}_{\mathsf{Q}}[\tau_\alpha] \geqslant \frac{\log(1/\alpha)}{\max_{\lambda \in \triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1\right)\right]},$$

the problem of deriving a "powerful" test has been reduced to the problem of deriving a *log-profitable portfolio* $\lambda_{\mathsf{Q}} \in \triangle_d$.

However, λ_{Q} depends on the alternative Q. Can it be "learned"?

Yes; this motivates the definition of *portfolio regret*.

Definition: Portfolio regret.

Define the portfolio regret $\mathcal{R}_n$ of $(\lambda_n)_{n \in \mathbb{N}}$ to be

$$\mathcal{R}_n := \max_{\lambda \in \triangle_d} \sum_{i=1}^{n} \log\left(\lambda^\top \mathbf{E}_i\right) - \sum_{i=1}^{n} \log\left(\lambda_i{}^\top \mathbf{E}_i\right).$$

Definition: Portfolio regret.

Define the portfolio regret $\mathcal{R}_n$ of $(\lambda_n)_{n\in\mathbb{N}}$ to be

$$\mathcal{R}_n := \max_{\lambda\in\triangle_d} \sum_{i=1}^{n} \log\left(\lambda^\top \mathbf{E}_i\right) - \sum_{i=1}^{n} \log\left(\lambda_i{}^\top \mathbf{E}_i\right).$$

This is *precisely* the notion of regret considered by Thomas Cover and Erik Ordentlich in their work on on universal portfolios circa 1990s.

Fortunately for us, Cover and Ordentlich derived an algorithm with *logarithmic* portfolio regret ("Universal Portfolio").

Theorem: Log-optimality via sublinear portfolio regret.

Suppose that $(\lambda_n)_{n \in \mathbb{N}}$ have sublinear portfolio regret. Then for $\mathsf{Q} \in \mathcal{Q}$,

$$\lim_{n \to \infty} \frac{1}{n} \log W_n = \max_{\lambda \in \triangle_d} \mathbb{E}_{\mathsf{Q}} \left[\log \left(\lambda^\top \mathbf{E}_1 \right) \right] \quad \mathsf{Q}\text{-almost surely.}$$

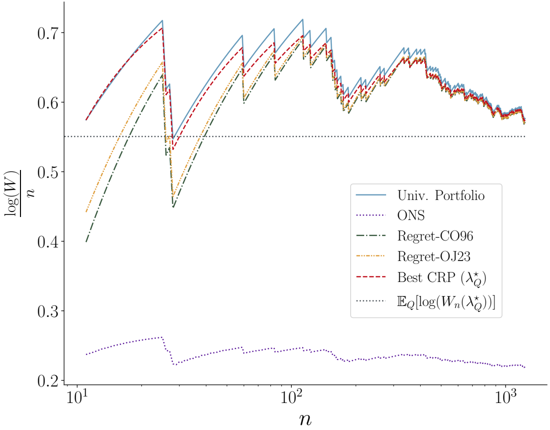
Theorem: Log-optimality via sublinear portfolio regret.

Suppose that $(\lambda_n)_{n\in\mathbb{N}}$ have sublinear portfolio regret. Then for $\mathsf{Q} \in \mathcal{Q}$,

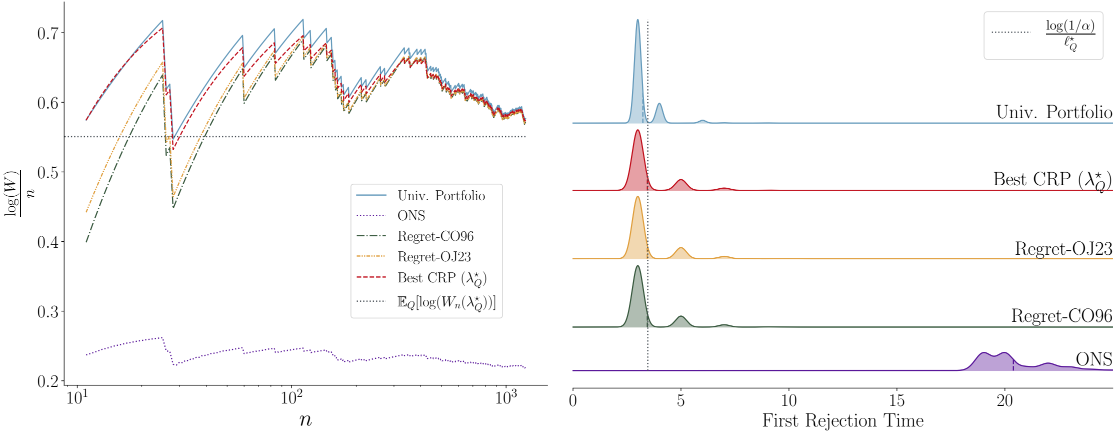
$$\lim_{n\to\infty} \frac{1}{n} \log W_n = \max_{\lambda\in\triangle_d} \mathbb{E}_{\mathsf{Q}} \left[\log\left(\lambda^\top \mathbf{E}_1\right)\right] \quad \mathsf{Q}\text{-almost surely.}$$

Moreover,

$$\lim_{\alpha\to 0^+} \frac{\mathbb{E}_{\mathsf{Q}}\left[\tau_\alpha\right]}{\log(1/\alpha)} = \frac{1}{\max_{\lambda\in\triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1\right)\right]}$$



0.7
0.6
0.5
0.4
0.3
0.2
$\frac{\log(W)}{n}$
10^1
10^2
10^3
n
Univ. Portfolio
ONS
Regret-CO96
Regret-OJ23
Best CRP ($\lambda_Q^\star$)
$\mathbb{E}_Q[\log(W_n(\lambda_Q^\star))]$



0.7
0.6
0.5
0.4
0.3
0.2
$\frac{\log(W)}{n}$
10^1
10^2
10^3
n
Univ. Portfolio
ONS
Regret-CO96
Regret-OJ23
Best CRP ($\lambda_Q^\star$)
$\mathbb{E}_Q[\log(W_n(\lambda_Q^\star))]$
$\frac{\log(1/\alpha)}{\ell_Q^\star}$
Univ. Portfolio
Best CRP ($\lambda_Q^\star$)
Regret-OJ23
Regret-CO96
ONS
0
5
10
15
20
First Rejection Time

Recap so far

1. **Q:** "What is sequential hypothesis testing?"
   A: For all stopping times τ, $\mathsf{P}(\phi_{\tau}^{(\alpha)} \text{ rejects}) \leqslant \alpha$.

Recap so far

1. **Q:** "What is sequential hypothesis testing?"
   A: For all stopping times τ, $\mathsf{P}(\phi_\tau^{(\alpha)} \text{ rejects}) \leqslant \alpha$.

2. **Q:** "How to derive valid sequential tests?"
   A: $W_n \mapsto \mathbb{1}\{W_n \geqslant 1/\alpha\}$ for a test supermartingale $(W_n)_{n\in\mathbb{N}}$.

Recap so far

1. **Q:** "What is sequential hypothesis testing?"
   A: For all stopping times τ, $\mathsf{P}(\phi_\tau^{(\alpha)} \text{ rejects}) \leqslant \alpha$.

2. **Q:** "How to derive valid sequential tests?"
   A: $W_n \mapsto \mathbb{1}\{W_n \geqslant 1/\alpha\}$ for a test supermartingale $(W_n)_{n\in\mathbb{N}}$.

3. **Q:** "How to derive a powerful supermartingale $(W_n)_{n\in\mathbb{N}}$?"
   A: Sublinear portfolio regret (e.g. Universal Portfolio algorithms).

Outline

1. What is sequential hypothesis testing (by betting)?

2. How are sequential hypothesis tests derived?

3. Defining and deriving optimal sequential tests.

4. Multi-armed sequential hypothesis testing

0mg 0.2mg 0.4mg 0.6mg 0.8mg 1mg

0mg 0.2mg 0.4mg 0.6mg 0.8mg 1mg

$\mathcal{P} = \{\mathsf{P} : \textit{All} \text{ dosages are ineffective under } \mathsf{P}\}$

$\mathcal{Q} = \{\mathsf{P} : \textit{At least one} \text{ dosage is effective under } \mathsf{P}\}$

0mg 0.2mg 0.4mg 0.6mg 0.8mg 1mg

$$\mathcal{P} = \{\mathsf{P} : \textit{All} \text{ dosages are ineffective under } \mathsf{P}\}$$

$$\mathcal{Q} = \{\mathsf{P} : \textit{At least one} \text{ dosage is effective under } \mathsf{P}\}$$

How does this complicate type-I error control under $\mathcal{P}$ and/or log-optimality under $\mathcal{Q}$?

Let $\mathcal{A} = \{1, \ldots, K\}$ be the arm set.

Multi-armed sequential testing.

Start with $W_1 = \$1$.

For $n = 1, 2, \ldots$:

1. **Choose arm $A_n \in \mathcal{A}$.**
2. Choose portfolio $\lambda_n \in \triangle_d$.
3. Observe $\mathbf{E}_n(A_n) \sim \mathsf{P}(A_n)$.
4. $W_n = W_{n-1} {\lambda_n}^\top \mathbf{E}_n(A_n)$.

EndFor

Let $\mathcal{A} = \{1, \ldots, K\}$ be the arm set.

Multi-armed sequential testing.

Start with $W_1 = \$1$.

For $n = 1, 2, \ldots$:

1. **Choose arm $A_n \in \mathcal{A}$.**
2. Choose portfolio $\lambda_n \in \triangle_d$.
3. Observe $\mathbf{E}_n(A_n) \sim \mathsf{P}(A_n)$.
4. $W_n = W_{n-1} \lambda_n^\top \mathbf{E}_n(A_n)$.

EndFor

Next: Type-I error control is preserved under *arbitrary* arm selection.

Proposition Type-I error control under $\mathcal{P}$

Suppose that λ_n and $A_n \in \mathcal{A}$ depend only on

$$\mathbf{E}_1(A_1), \ldots, \mathbf{E}_{n-1}(A_{n-1}).$$

Proposition Type-I error control under $\mathcal{P}$

Suppose that λ_n and $A_n \in \mathcal{A}$ depend only on

$$\mathbf{E}_1(A_1), \ldots, \mathbf{E}_{n-1}(A_{n-1}).$$

Then,

$$W_n = \prod_{i=1}^{n} \lambda_i{}^\top \mathbf{E}_i(A_i)$$

forms a test supermartingale

Proposition Type-I error control under $\mathcal{P}$

Suppose that λ_n and $A_n \in \mathcal{A}$ depend only on

$$\mathbf{E}_1(A_1), \ldots, \mathbf{E}_{n-1}(A_{n-1}).$$

Then,

$$W_n = \prod_{i=1}^{n} \lambda_i^\top \mathbf{E}_i(A_i)$$

forms a test supermartingale and hence

$$\phi_n^{(\alpha)} := \mathbb{1}\{W_n \geqslant 1/\alpha\}$$

forms a sequential hypothesis test for the global null $\mathcal{P}$.

The proof proceeds almost identically to before.

Takeaway: type-I error control in the multi-armed setting is easy.

Optimality is a different story.

$$\max_{\lambda \in \triangle_d} \mathbb{E}_{\mathsf{Q}} \left[\log \left(\lambda^\top \mathbf{E}_1 \right) \right].$$

$$\max_{(a,\lambda)\in\mathcal{A}\times\triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right].$$

$$\max_{(a,\lambda)\in\mathcal{A}\times\triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right].$$

Let us now see how $\frac{1}{n}\sum_{i=1}^n \log\left(\lambda_i{}^\top \mathbf{E}_i(A_i)\right)$ might concentrate around the above.

$$
\max_{(a,\lambda)\in\mathcal{A}\times\triangle_d} n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right] - \sum_{i=1}^{n} \log\left({\lambda_i}^\top \mathbf{E}_i(A_i)\right)
$$

$$
= \underbrace{\max_{(a,\lambda)\in\mathcal{A}\times\triangle_d} n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right] - \sum_{i=1}^{n} \mathbb{E}_{\mathsf{Q}}\left[\log(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}(A_i))\right]}_{\textbf{Allocation (multi-armed bandit) regret}}
$$

$$
+ \underbrace{\sum_{i=1}^{n} \mathbb{E}_{\mathsf{Q}}\left[\log(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i))\right] - \sum_{i=1}^{n} \log\left(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i)\right)}_{\approx 0 \ \mathsf{Q}\text{-a.s. by the strong law of large numbers.}}
$$

$$
+ \underbrace{\sum_{i=1}^{n} \log\left(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i)\right) - \sum_{i=1}^{n} \log\left({\lambda_i}^\top \mathbf{E}_i(A_i)\right)}_{\text{Arm-wise portfolio regret}}
$$

$$
\begin{aligned}
&\max_{(a,\lambda)\in\mathcal{A}\times\triangle_d} n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right] - \sum_{i=1}^n \log\left(\lambda_i^\top \mathbf{E}_i(A_i)\right) \\
= &\underbrace{\max_{(a,\lambda)\in\mathcal{A}\times\triangle_d} n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right] - \sum_{i=1}^n \mathbb{E}_{\mathsf{Q}}\left[\log(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}(A_i))\right]}_{\textbf{Allocation (multi-armed bandit) regret}} \\
+ &\underbrace{\sum_{i=1}^n \mathbb{E}_{\mathsf{Q}}\left[\log(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i))\right] - \sum_{i=1}^n \log\left(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i)\right)}_{\approx 0 \text{ Q-a.s. by the strong law of large numbers.}} \\
+ &\underbrace{\sum_{i=1}^n \log\left(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i)\right) - \sum_{i=1}^n \log\left(\lambda_i^\top \mathbf{E}_i(A_i)\right)}_{\text{Arm-wise portfolio regret}}
\end{aligned}
$$

$$\max_{(a,\lambda)\in\mathcal{A}\times\triangle_d} n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right] - \sum_{i=1}^{n} \log\left(\lambda_i^\top \mathbf{E}_i(A_i)\right)$$

$$= \underbrace{\max_{(a,\lambda)\in\mathcal{A}\times\triangle_d} n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right] - \sum_{i=1}^{n} \mathbb{E}_{\mathsf{Q}}\left[\log(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}(A_i))\right]}_{\textbf{Allocation (multi-armed bandit) regret}}$$

$$+ \underbrace{\sum_{i=1}^{n} \mathbb{E}_{\mathsf{Q}}\left[\log(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i))\right] - \sum_{i=1}^{n} \log\left(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i)\right)}_{\approx 0 \text{ } \mathsf{Q}\text{-a.s. by the strong law of large numbers.}}$$

$$+ \underbrace{\sum_{i=1}^{n} \log\left(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i)\right) - \sum_{i=1}^{n} \log\left(\lambda_i^\top \mathbf{E}_i(A_i)\right)}_{\text{Arm-wise portfolio regret}}$$

$$
\begin{aligned}
&\max_{(a,\lambda)\in\mathcal{A}\times\triangle_d} n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right] - \sum_{i=1}^{n} \log\left(\lambda_i^\top \mathbf{E}_i(A_i)\right) \\
= &\underbrace{\max_{(a,\lambda)\in\mathcal{A}\times\triangle_d} n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right] - \sum_{i=1}^{n} \mathbb{E}_{\mathsf{Q}}\left[\log(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}(A_i))\right]}_{\textbf{Allocation (multi-armed bandit) regret}} \\
+ &\underbrace{\sum_{i=1}^{n} \mathbb{E}_{\mathsf{Q}}\left[\log(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i))\right] - \sum_{i=1}^{n} \log\left(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i)\right)}_{\approx\, 0 \text{ Q-a.s. by the strong law of large numbers.}} \\
+ &\underbrace{\sum_{i=1}^{n} \log\left(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i)\right) - \sum_{i=1}^{n} \log\left(\lambda_i^\top \mathbf{E}_i(A_i)\right)}_{\text{Arm-wise portfolio regret}}
\end{aligned}
$$

$$\max_{(a,\lambda)\in\mathcal{A}\times\triangle_d} n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right] - \sum_{i=1}^{n} \log\left(\lambda_i^\top \mathbf{E}_i(A_i)\right)$$

$$= \underbrace{\max_{(a,\lambda)\in\mathcal{A}\times\triangle_d} n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right] - \sum_{i=1}^{n} \mathbb{E}_{\mathsf{Q}}\left[\log(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}(A_i))\right]}_{\textbf{Allocation (multi-armed bandit) regret}}$$

$$+ \underbrace{\sum_{i=1}^{n} \mathbb{E}_{\mathsf{Q}}\left[\log(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i))\right] - \sum_{i=1}^{n} \log\left(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i)\right)}_{\approx 0 \text{ Q-a.s. by the strong law of large numbers.}}$$

$$+ \underbrace{\sum_{i=1}^{n} \log\left(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}_i(A_i)\right) - \sum_{i=1}^{n} \log\left(\lambda_i^\top \mathbf{E}_i(A_i)\right)}_{\text{Arm-wise portfolio regret}}$$

We are tasked with minimizing

$$\underbrace{\max_{a\in\mathcal{A}} n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda_{\mathsf{Q}}^{\top}\mathbf{E}_1(a)\right)\right] - \sum_{i=1}^{n}\mathbb{E}_{\mathsf{Q}}\left[\log(\lambda_{\mathsf{Q}}(A_i)^{\top}\mathbf{E}(A_i))\right]}_{\textbf{Allocation (multi-armed bandit) regret}}.$$

We are tasked with minimizing

$$\underbrace{\max_{a\in\mathcal{A}} n\mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda_{\mathsf{Q}}^{\top}\mathbf{E}_1(a)\right)\right] - \sum_{i=1}^{n}\mathbb{E}_{\mathsf{Q}}\left[\log(\lambda_{\mathsf{Q}}(A_i)^{\top}\mathbf{E}(A_i))\right]}_{\textbf{Allocation (multi-armed bandit) regret}}.$$

This is *almost exactly* the type of regret considered in stochastic multi-armed bandits.

A one-slide crash course on upper-confidence-bound algorithms and the "principle of optimism" (Lai & Robbins, 1985).

A one-slide crash course on upper-confidence-bound algorithms and the "principle of optimism" (Lai & Robbins, 1985).

Suppose that I maintain confidence intervals

$$\left(L_n^{(1)}, U_n^{(1)}\right), \ldots, \left(L_n^{(K)}, U_n^{(K)}\right)$$

for the means of all K arms.

A one-slide crash course on upper-confidence-bound algorithms and the "principle of optimism" (Lai & Robbins, 1985).

Suppose that I maintain confidence intervals

$$\left(L_n^{(1)}, U_n^{(1)}\right), \ldots, \left(L_n^{(K)}, U_n^{(K)}\right)$$

for the means of all K arms.

There are two reasons why $U_n^{(a)}$ for $a \in \mathcal{A}$ could be large.

1. The mean of arm a is large (should exploit).
2. I am uncertain about how large its mean is (should explore).

A one-slide crash course on upper-confidence-bound algorithms and the "principle of optimism" (Lai & Robbins, 1985).

Suppose that I maintain confidence intervals

$$\left(L_n^{(1)}, U_n^{(1)}\right), \ldots, \left(L_n^{(K)}, U_n^{(K)}\right)$$

for the means of all K arms.

There are two reasons why $U_n^{(a)}$ for $a \in \mathcal{A}$ could be large.

1. The mean of arm a is large (should exploit).
2. I am uncertain about how large its mean is (should explore).

Illustration (credits: Adrien Prevost, INRIA)

</detour>

"Nice" confidence intervals $\implies$ allocation regret minimization.

How are CIs in the bandit literature usually derived?

How are CIs in the bandit literature usually derived?

1. Assume sub-Gaussianity, and
2. Use the fact that $X_1(a), \ldots, X_n(a) \stackrel{\text{iid}}{\sim} \mathsf{P}(a)$ for each arm $a \in \mathcal{A}$.

How are CIs in the bandit literature usually derived?

1. Assume sub-Gaussianity, and
2. Use the fact that $X_1(a), \ldots, X_n(a) \overset{\text{iid}}{\sim} \mathsf{P}(a)$ for each arm $a \in \mathcal{A}$.

But recall the allocation regret:

$$\max_{a \in \mathcal{A}} n \mathbb{E}_{\mathsf{Q}} \left[\log \left(\lambda_{\mathsf{Q}}^\top \mathbf{E}_1(a) \right) \right] - \sum_{i=1}^{n} \mathbb{E}_{\mathsf{Q}} \left[\log(\lambda_{\mathsf{Q}}(A_i)^\top \mathbf{E}(A_i)) \right].$$

So for us, $X_i = \log \left(\lambda_{\mathsf{Q}}(a)^\top \mathbf{E}_i(a) \right)$.

How are CIs in the bandit literature usually derived?

1. Assume sub-Gaussianity, and
2. Use the fact that $X_1(a), \ldots, X_n(a) \overset{\text{iid}}{\sim} \mathsf{P}(a)$ for each arm $a \in \mathcal{A}$.

But recall the allocation regret:

$$\max_{a \in \mathcal{A}} n \mathbb{E}_{\mathsf{Q}} \left[\log \left(\lambda_{\mathsf{Q}}^{\top} \mathbf{E}_1(a) \right) \right] - \sum_{i=1}^{n} \mathbb{E}_{\mathsf{Q}} \left[\log(\lambda_{\mathsf{Q}}(A_i)^{\top} \mathbf{E}(A_i)) \right].$$

So for us, $X_i = \log \left(\lambda_{\mathsf{Q}}(a)^{\top} \mathbf{E}_i(a) \right)$.

These are *neither* sub-Gaussian *nor* observed iid from any distribution.

Let $N_a(n) := \sum_{i=1}^{n} \mathbb{1}\{A_i = a\}$ be the number of times arm a was pulled up until time n.

Let $N_a(n) := \sum_{i=1}^{n} \mathbb{1}\{A_i = a\}$ be the number of times arm a was pulled up until time n.

Define the upper confidence bound

$$\mathrm{UCB}_a(n) := \max_{\lambda \in \triangle_d} \frac{1}{N_a(n-1)} \sum_{i=1}^{n} \mathbb{1}\{A_i = a\} \log(\lambda^\top \mathbf{E}_i(A_i)) + \sqrt{\frac{8 \log(n+1)}{N_a(n-1)}} + \frac{4 \log(n+1)}{N_a(n-1)}.$$

Let $N_a(n) := \sum_{i=1}^n \mathbb{1}\{A_i = a\}$ be the number of times arm a was pulled up until time n.

Define the upper confidence bound

$$\mathrm{UCB}_a(n) := \max_{\lambda \in \triangle_d} \frac{1}{N_a(n-1)} \sum_{i=1}^{n} \mathbb{1}\{A_i = a\} \log(\lambda^\top \mathbf{E}_i(A_i)) + \sqrt{\frac{8\log(n+1)}{N_a(n-1)}} + \frac{4\log(n+1)}{N_a(n-1)}.$$

This turns out to have all the "nice" properties needed for UCB-style analyses. In particular, it suffices for

$$\text{Allocation regret}_n = O\left(\log(n)\right).$$

All of this work was for the following theorem.

All of this work was for the following theorem.

Theorem: Multi-armed log-optimality

Fix $\mathsf{Q} \in \mathcal{Q}$. Suppose $(\lambda_n(a))_{n\in\mathbb{N}}$ is chosen using universal portfolio for each arm $a \in \mathcal{A}$.

Also, suppose $(A_n)_{n\in\mathbb{N}}$ are chosen via our UCB allocation.

All of this work was for the following theorem.

Theorem: Multi-armed log-optimality

Fix $\mathsf{Q} \in \mathcal{Q}$. Suppose $(\lambda_n(a))_{n\in\mathbb{N}}$ is chosen using universal portfolio for each arm $a \in \mathcal{A}$.

Also, suppose $(A_n)_{n\in\mathbb{N}}$ are chosen via our UCB allocation. Then,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \log\left(\lambda_i(a)^\top \mathbf{E}_i(A_i)\right) = \max_{(a,\lambda)\in\mathcal{A}\times\triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right]$$

with Q-probability one.

All of this work was for the following theorem.

Theorem: Multi-armed log-optimality

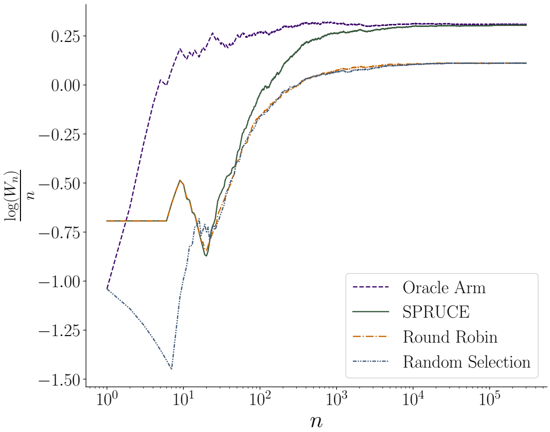
Fix $\mathsf{Q} \in \mathcal{Q}$. Suppose $(\lambda_n(a))_{n \in \mathbb{N}}$ is chosen using universal portfolio for each arm $a \in \mathcal{A}$.

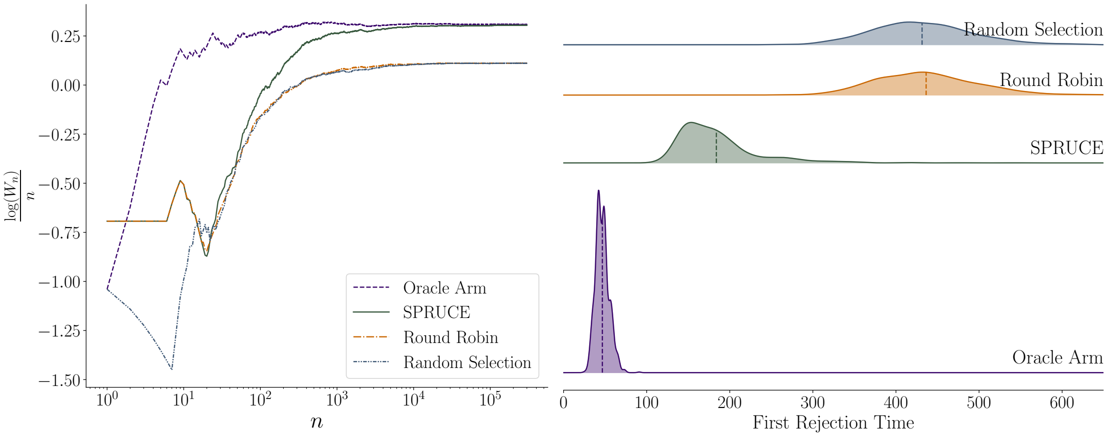
Also, suppose $(A_n)_{n \in \mathbb{N}}$ are chosen via our UCB allocation. Then,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \log\left(\lambda_i(a)^\top \mathbf{E}_i(A_i)\right) = \max_{(a,\lambda) \in \mathcal{A} \times \triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right]$$

with Q-probability one. Furthermore,

$$\lim_{\alpha \to 0^+} \frac{\mathbb{E}_{\mathsf{Q}}[\tau_\alpha]}{\log(1/\alpha)} = \left(\max_{(a,\lambda) \in \mathcal{A} \times \triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right]\right)^{-1}$$


0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1.25
−1.50
$\frac{\log(W_n)}{n}$
10^0 10^1 10^2 10^3 10^4 10^5
n
Oracle Arm
SPRUCE
Round Robin
Random Selection


0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1.25
−1.50
$\frac{\log(W_n)}{n}$
10^0
10^1
10^2
10^3
10^4
10^5
n
Oracle Arm
SPRUCE
Round Robin
Random Selection
Random Selection
Round Robin
SPRUCE
Oracle Arm
0
100
200
300
400
500
600
First Rejection Time

Summary

1. **Q:** "What is sequential hypothesis testing?"
   A: For all stopping times τ, $\mathsf{P}(\phi_\tau^{(\alpha)}$ rejects$) \leqslant \alpha$.

Summary

1. **Q:** "What is sequential hypothesis testing?"
   A: For all stopping times τ, $\mathsf{P}(\phi_\tau^{(\alpha)} \text{ rejects}) \leqslant \alpha$.

2. **Q:** "How to derive valid sequential tests?"
   A: $W_n \mapsto \mathbb{1}\{W_n \geqslant 1/\alpha\}$ for a test supermartingale $(W_n)_{n\in\mathbb{N}}$.

Summary

1. **Q:** "What is sequential hypothesis testing?"
   A: For all stopping times τ, $\mathsf{P}(\phi_{\tau}^{(\alpha)} \text{ rejects}) \leqslant \alpha$.

2. **Q:** "How to derive valid sequential tests?"
   A: $W_n \mapsto \mathbb{1}\{W_n \geqslant 1/\alpha\}$ for a test supermartingale $(W_n)_{n\in\mathbb{N}}$.

3. **Q:** "How to derive a powerful supermartingale $(W_n)_{n\in\mathbb{N}}$?"
   A: Sublinear portfolio regret (e.g. Universal Portfolio algorithms).

Summary

1. **Q:** "What is sequential hypothesis testing?"
   A: For all stopping times τ, $\mathsf{P}(\phi_\tau^{(\alpha)} \text{ rejects}) \leqslant \alpha$.

2. **Q:** "How to derive valid sequential tests?"
   A: $W_n \mapsto \mathbb{1}\{W_n \geqslant 1/\alpha\}$ for a test supermartingale $(W_n)_{n\in\mathbb{N}}$.

3. **Q:** "How to derive a powerful supermartingale $(W_n)_{n\in\mathbb{N}}$?"
   A: Sublinear portfolio regret (e.g. Universal Portfolio algorithms).

4. **Q:** "How to do the same when multiple arms are present?"
   A: Bespoke upper-confidence-bound algorithms.

Thank you

ianws.com