Log-Optimality & Multi-Armed Sequential Hypothesis Testing

Ian Waudby-Smith

Miller Institute & Department of Statistics
University of California, Berkeley

London School of Economics, 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: $p_n = 0.023$

$\Longrightarrow$ **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: $p_n = 0.023$

$\Longrightarrow$ **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".

The previous slide illustrates the issue of n being too *small*.

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

The previous slide illustrates the issue of n being too *small*.

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. $p_n < 0.00001 \ll \alpha = 0.01$.

The previous slide illustrates the issue of n being too *small*.

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. $p_n < 0.00001 \ll \alpha = 0.01$.

Sequential testing provides one solution to these 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 in recent years; textbook by Ramdas & Wang [2025])

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 $\mathcal{Q}$.

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

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

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

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

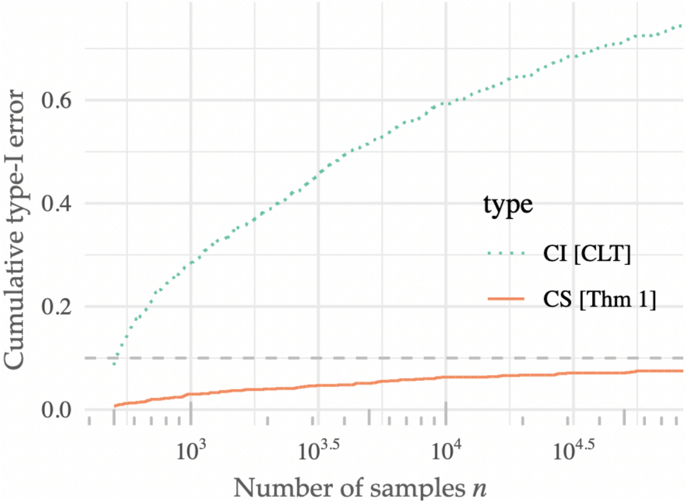
$$(e.g.\ \mathcal{P} = \{P : \mathit{trt\ effect} = 0\}\ \mathit{versus}\ \mathcal{Q} = \{P : \mathit{trt\ effect} > 0\})$$

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

Sequential tests result from the following two-step procedure:

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**.

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**.

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

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

Why does this procedure work?

Why does this procedure work? Ville's inequality: For a nonnegative P-supermartingale $(W_n)_{n\in\mathbb{N}}$,

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

Why does this procedure work? Ville's inequality: For a nonnegative P-supermartingale $(W_n)_{n\in\mathbb{N}}$,

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

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

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

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.

Claim: $W_n = \prod_{i=1}^{n}(1 + \gamma_i(2X_i - 1))$ forms a test martingale.

Claim: $W_n = \prod_{i=1}^n (1 + \gamma_i(2X_i - 1))$ forms a test martingale.

(1) Nonnegativity follows from $1 - \gamma_i + \gamma_i 2X_i$ being a cvx combination of nonnegative rvs.

Claim: $W_n = \prod_{i=1}^n (1 + \gamma_i(2X_i - 1))$ forms a test martingale.

(1) Nonnegativity follows from $1 - \gamma_i + \gamma_i 2X_i$ being a cvx combination of nonnegative rvs.

(2) Need to show $\mathbb{E}_{\mathsf{P}}[W_n \mid X_1, \ldots, X_{n-1}] = W_{n-1}$ for $\mathsf{P} \in \mathcal{P}$.

Claim: $W_n = \prod_{i=1}^n (1 + \gamma_i(2X_i - 1))$ forms a test martingale.

(1) Nonnegativity follows from $1 - \gamma_i + \gamma_i 2X_i$ being a cvx combination of nonnegative rvs.

(2) Need to show $\mathbb{E}_{\mathsf{P}}[W_n \mid X_1, \dots, X_{n-1}] = W_{n-1}$ for $\mathsf{P} \in \mathcal{P}$.

$$
\begin{aligned}
&\mathbb{E}_{\mathsf{P}}\left[\prod_{i=1}^n (1 + \gamma_i(2X_i - 1)) \mid X_1, \dots, X_{n-1}\right] \\
&= W_{n-1}\mathbb{E}_{\mathsf{P}}\left[1 - \gamma_n(2X_i - 1) \mid X_1, \dots, X_{n-1}\right] \\
&= W_{n-1}\left(1 - \gamma_n(2\mathbb{E}_{\mathsf{P}}[X_n] - 1)\right) \\
&= W_{n-1}.
\end{aligned}
$$

Claim: $W_n = \prod_{i=1}^n (1 + \gamma_i(2X_i - 1))$ forms a test martingale.

(1) Nonnegativity follows from $1 - \gamma_i + \gamma_i 2X_i$ being a cvx combination of nonnegative rvs.

(2) Need to show $\mathbb{E}_{\mathsf{P}}[W_n \mid X_1, \ldots, X_{n-1}] = W_{n-1}$ for $\mathsf{P} \in \mathcal{P}$.

$$\begin{aligned}
&\mathbb{E}_{\mathsf{P}}\left[\prod_{i=1}^n (1 + \gamma_i(2X_i - 1)) \mid X_1, \ldots, X_{n-1}\right] \\
&= W_{n-1}\mathbb{E}_{\mathsf{P}}\left[1 - \gamma_n(2X_i - 1) \mid X_1, \ldots, X_{n-1}\right] \\
&= W_{n-1}\left(1 - \gamma_n(2\mathbb{E}_{\mathsf{P}}[X_n] - 1)\right) \\
&= W_{n-1}.
\end{aligned}$$

Therefore, W_n forms a test martingale. □

The previous slide implicitly made use of "e-values". Let us make that explicit.

The previous slide implicitly made use of "e-values". Let us make that explicit.

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

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

The previous slide implicitly made use of "e-values". Let us make that explicit.

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

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

The process W_n is a product of convex combinations of e-values:

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

The previous slide implicitly made use of "e-values". Let us make that explicit.

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

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

The process W_n is a product of convex combinations of e-values:

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

One can re-cast several testing problems *(bounded means, two-sample, independence, equality of bounded tuples, testing randomness online, etc.)* from the literature with convex combinations of carefully chosen e-values.

Throughout, consider the process $(W_n)_{n\in\mathbb{N}}$ given by (predictable) convex combination of $d+1$ e-values:

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

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

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

Throughout, consider the process $(W_n)_{n \in \mathbb{N}}$ given by (predictable) convex combination of $d+1$ e-values:

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

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

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

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 $g^\star$.

There has been one lingering question this entire discussion:

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

There has been one lingering question this entire discussion:

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

Indeed, $\prod_{i=1}^{n} {\lambda_i}^\top \mathbf{E}_i$ yields a valid sequential test under $\mathcal{P}$ for *any* choice of $(\lambda_n)_{n\in\mathbb{N}}$.

Are there "better" choices than others under $\mathcal{Q}$?

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:

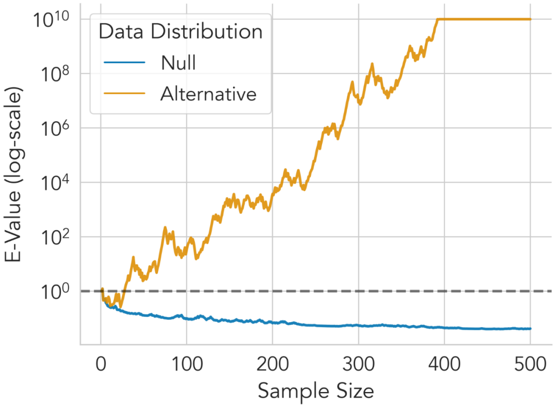
(i) **Growth-rate optimality**

(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 larger values of W_n signify more evidence against $\mathcal{P}$, we aim to choose $(\lambda_n)_{n \in \mathbb{N}}$ so that W_n diverges "quickly".

Since larger values of W_n signify more evidence against $\mathcal{P}$, 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 larger values of W_n signify more evidence against $\mathcal{P}$, 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 larger values of W_n signify more evidence against $\mathcal{P}$, 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:

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

This is an instantiation of the "Kelly criterion" (J.L. Kelly Jr., [1956]).

Proposition: Upper bound on the asymptotic growth rate

Suppose $\mathbf{E}_1, \mathbf{E}_2, \cdots \sim \mathsf{Q} \in \mathcal{Q}$. For any predictable sequence $(\lambda_n)_{n\in\mathbb{N}}$, we have that

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \log\left({\lambda_i}^\top \mathbf{E}_i\right) \leqslant \max_{\lambda \in \triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log(\lambda^\top \mathbf{E}_1)\right]$$

with Q-probability one.

Proposition: Upper bound on the asymptotic growth rate

Suppose $\mathbf{E}_1, \mathbf{E}_2, \cdots \sim \mathsf{Q} \in \mathcal{Q}$. For any predictable sequence $(\lambda_n)_{n \in \mathbb{N}}$, we have that

$$\limsup_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \log\left(\lambda_i{}^{\top} \mathbf{E}_i\right) \leqslant \max_{\lambda \in \triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log(\lambda^{\top} \mathbf{E}_1)\right]$$

with Q-probability one.

One can also motivate the Kelly criterion from a more “purely statistical” perspective.

(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

For any choice of $(\lambda_n)_{n\in\mathbb{N}}$, let τ_α be the resulting first hitting time of $W_n := \prod_{i=1}^n \lambda_i^\top \mathbf{E}_i$. Then,

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

Let $\lambda_{\mathsf{Q}} := \operatorname{argmax}_{\lambda \in \triangle_d} \mathbb{E}_{\mathsf{Q}}[\log(\lambda^\top \mathbf{E}_1)]$. Given the Q-a.s. upper bound

$$\limsup_{n\to\infty} \frac{1}{n} \log(W_n) \leqslant \mathbb{E}_{\mathsf{Q}} \left[\log \left(\lambda_{\mathsf{Q}}^\top \mathbf{E}_1\right)\right]$$

and the rejection time lower bound

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

it is desirable to seek out tests that behave as if we knew λ_{Q} (the "Kelly bet").

Recall, we are in a setting with a *composite* alternative $\mathcal{Q}$. We do have access to $\mathcal{Q}$ but we do not know Q nor λ_{Q}.

Are these "learnable", and under what conditions on Q?

Recall, we are in a setting with a *composite* alternative $\mathcal{Q}$. We do have access to $\mathcal{Q}$ but we do not know Q nor λ_{Q}.

Are these "learnable", and under what conditions on Q?

I do not know for sure, but this will not matter.

Recall, we are in a setting with a *composite* alternative $\mathcal{Q}$. We do have access to $\mathcal{Q}$ but we do not know Q nor λ_{Q}.

Are these "learnable", and under what conditions on Q?

I do not know for sure, but this will not matter.

For the purposes of deriving an optimal test, it suffices to choose $(\lambda_n)_{n \in \mathbb{N}}$ with *sublinear 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.

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.

Note, this is defined pathwise (irrespective of $\mathsf{Q}, \mathcal{Q}, \mathcal{P}$*, etc.). It is not obvious whether this is a desirable quantity to minimize.*

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,

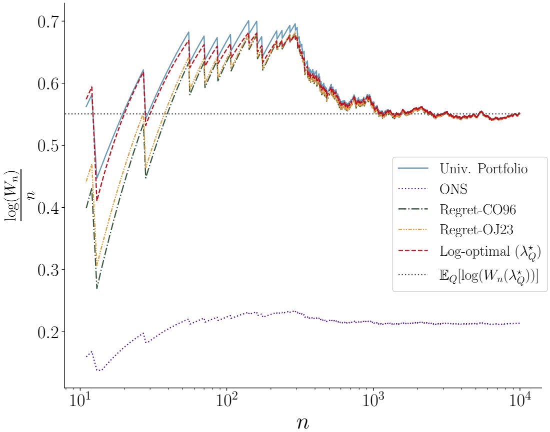
$$\lim_{\alpha\to 0^+} \frac{\mathbb{E}_{\mathsf{Q}}\left[\tau_\alpha\right]}{\log(1/\alpha)} = \frac{1}{\max_{\lambda\in\triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1\right)\right]}.$$

In particular, one could take λ_n according to Cover's universal portfolio algorithm:

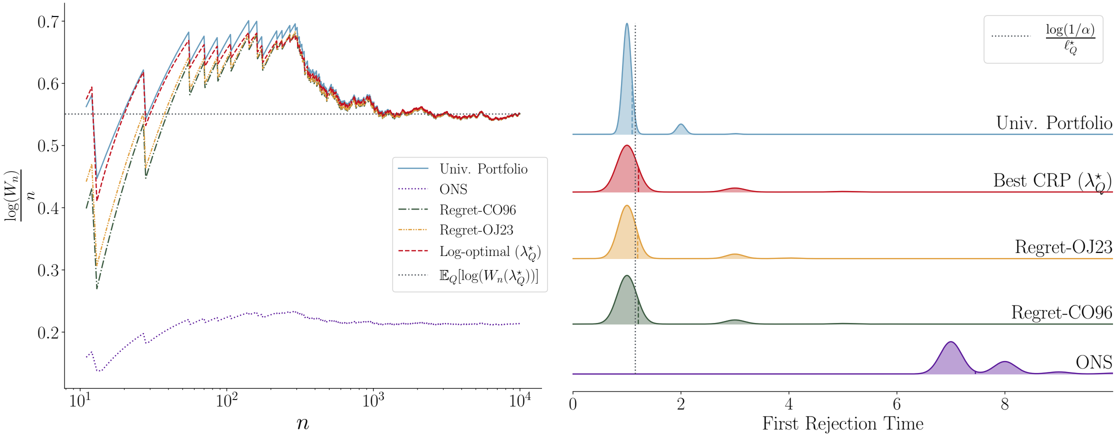
In particular, one could take λ_n according to Cover's universal portfolio algorithm:

$$\lambda_n := \frac{\int_{\lambda \in \triangle_d} \lambda W_{n-1}(\lambda) \mathrm{d}F(\lambda)}{\int_{\lambda \in \triangle_d} W_{n-1}(\lambda) \mathrm{d}F(\lambda)},$$

in which case, the portfolio regret is *logarithmic*: $\mathcal{R}_n = O(\log n)$.



0.7
0.6
0.5
0.4
0.3
0.2
$\frac{\log(W_n)}{n}$
10^1
10^2
10^3
10^4
n
Univ. Portfolio
ONS
Regret-CO96
Regret-OJ23
Log-optimal ($\lambda_Q^\star$)
$\mathbb{E}_Q[\log(W_n(\lambda_Q^\star))]$



0.7
0.6
0.5
0.4
0.3
0.2
$\frac{\log(W_n)}{n}$
10^1
10^2
10^3
10^4
n
Univ. Portfolio
ONS
Regret-CO96
Regret-OJ23
Log-optimal ($\lambda_Q^\star$)
$\mathbb{E}_Q[\log(W_n(\lambda_Q^\star))]$
$\frac{\log(1/\alpha)}{\ell_Q^\star}$
Univ. Portfolio
Best CRP ($\lambda_Q^\star$)
Regret-OJ23
Regret-CO96
ONS
0
2
4
6
8
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

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

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

No matter how λ_n and $A_n \in \mathcal{A}$ are chosen* ,

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

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 $\mathcal{P}$.

* *formally, as long as they are measurable with respect to* $\mathbf{E}_1(A_1), \ldots, \mathbf{E}_{n-1}(A_{n-1})$

Takeaway: type-I error control in the multi-armed setting is obtained for free.

Optimality is a different story.

Takeaway: type-I error control in the multi-armed setting is obtained for free.

Optimality is a different story.

Let us inspect the result first, and discuss how it is achieved later.

Theorem: Multi-armed log-optimality

Fix $\mathsf{Q} \in \mathcal{Q}$ inducing a distribution on all arms $\mathcal{A}$.

Choose $(\lambda_n(a))_{n\in\mathbb{N}}$ with sublinear portfolio regret for each $a \in \mathcal{A}$.

Choose $(A_n)_{n\in\mathbb{N}}$ according to a bespoke upper-confidence-bound-type allocation (more details later...)

Theorem: Multi-armed log-optimality

Fix $\mathsf{Q} \in \mathcal{Q}$ inducing a distribution on all arms $\mathcal{A}$.
Choose $(\lambda_n(a))_{n\in\mathbb{N}}$ with sublinear portfolio regret for each $a \in \mathcal{A}$.
Choose $(A_n)_{n\in\mathbb{N}}$ according to a bespoke upper-confidence-bound-type allocation (more details later...) 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.

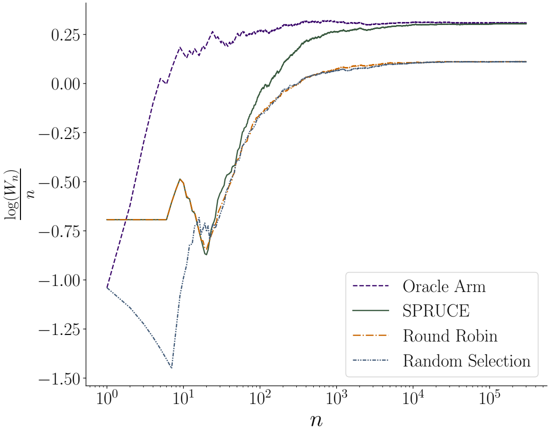
Theorem: Multi-armed log-optimality

Fix $\mathsf{Q} \in \mathcal{Q}$ inducing a distribution on all arms $\mathcal{A}$.
Choose $(\lambda_n(a))_{n\in\mathbb{N}}$ with sublinear portfolio regret for each $a \in \mathcal{A}$.
Choose $(A_n)_{n\in\mathbb{N}}$ according to a bespoke upper-confidence-bound-type allocation (more details later...) Then,

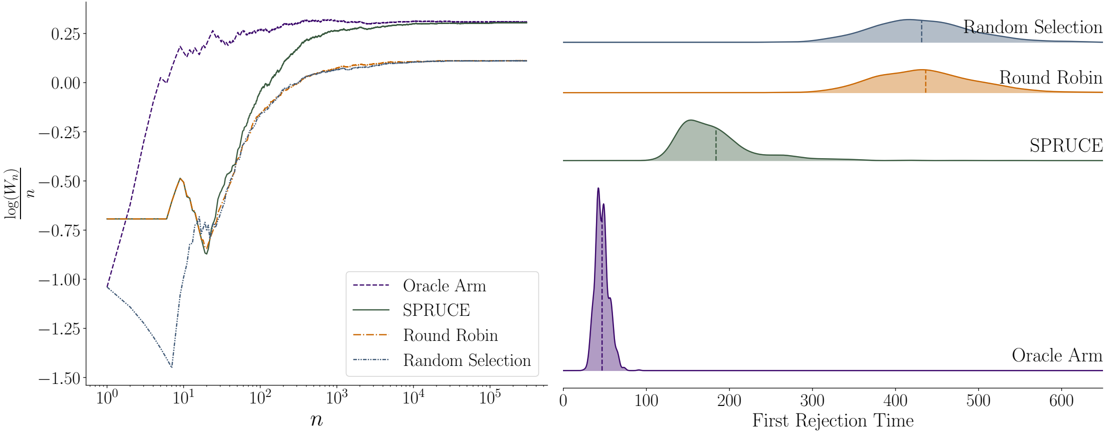
$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \log\left(\lambda_i(a)^\top \mathbf{E}_i(A_i)\right) = \max_{(a,\lambda)\in\mathcal{A}\times\triangle_d} \mathbb{E}_{\mathsf{Q}}\left[\log\left(\lambda^\top \mathbf{E}_1(a)\right)\right]$$

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

What is the idea behind the algorithm and proof?

We are tasked with bounding the difference

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

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

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

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

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

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

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

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

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

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

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

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

We are now focused on 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 now focused on 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, which can typically be $O(\log n)$ when using upper-confidence-bound-type algorithms.

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.

The upper confidence bound (UCB) algorithm says to pull $\operatorname{argmax}_a U_n^{(a)}$ at time n.

Why? Either the mean of arm k is large, or it has not been sufficiently explored.

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.

The upper confidence bound (UCB) algorithm says to pull $\operatorname{argmax}_a U_n^{(a)}$ at time n.

Why? Either the mean of arm k is large, or it has not been sufficiently explored.

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 $X_1(a), \ldots, X_n(a) \overset{\text{iid}}{\sim} \mathsf{P}(a)$ 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 $X_1(a), \dots, X_n(a) \overset{\text{iid}}{\sim} \mathsf{P}(a)$ 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

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

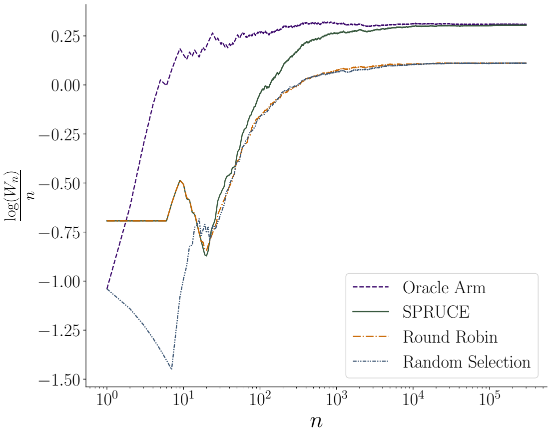
Let $N_a(n) := \sum_{i=1}^{n} \mathbb{1}\{A_i = a\}$ be the number of times arm a was pulled up until time n.

Define the upper confidence bound

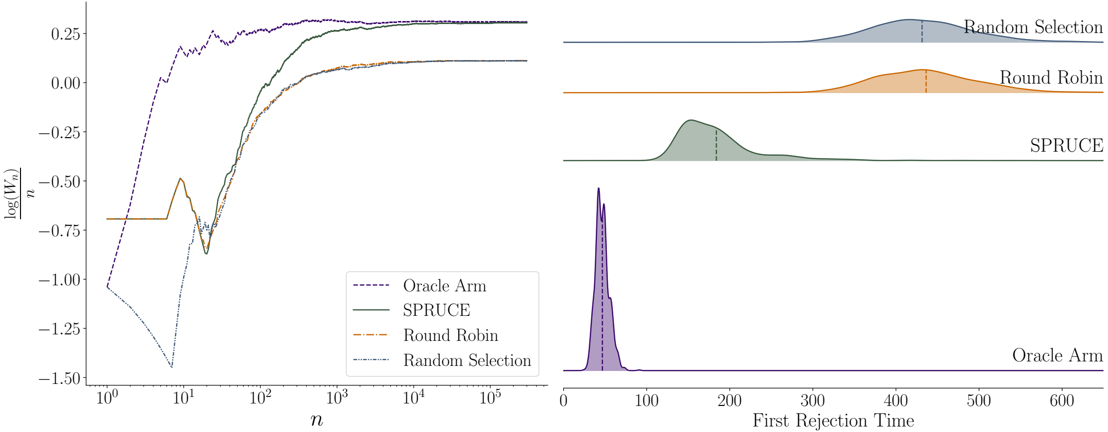
$$\mathrm{UCB}_a(n) := \max_{\lambda \in \triangle_d} \frac{1}{N_a(n-1)} \sum_{i=1}^{n} \mathbb{1}\{A_i = a\} \log(\lambda^\top \mathbf{E}_i(A_i)) + \sqrt{\frac{8\log(n+1)}{N_a(n-1)}} + \frac{4\log(n+1)}{N_a(n-1)}.$$

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

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



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



$\frac{\log(W_n)}{n}$
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1.25
−1.50
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, $\mathsf{P}(\phi_{\tau}^{(\alpha)} \text{ rejects}) \leqslant \alpha$.

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