Confidence sequences for sequential experimentation

Steve Howard
Joint work with Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon

August 5, 2022

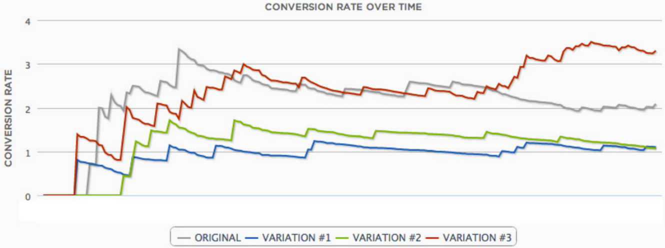
Sequential monitoring of experiment results is problematic.

Click on the button

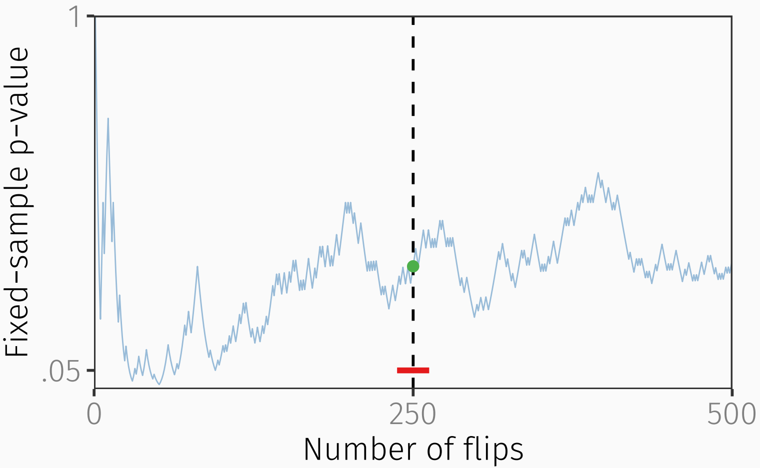
The percentage of visitors who clicked on a tracked element.

Variation #3 is beating Original by +58.0%.

VARIATIONS	VISITORS	CONVERSIONS	CONVERSION RATE	IMPROVEMENT	CHANCE TO BEAT BASELINE
Variation #3	970	32	3.3% (±1.12%)	+58.0%	95.2%
Original BASELINE	1,006	21	2.1% (±0.88%)	---	---
Variation #1	999	11	1.1% (±0.65%)	-47.3%	3.9%
Variation #2	1,027	11	1.1% (±0.63%)	-48.7%	3.3%

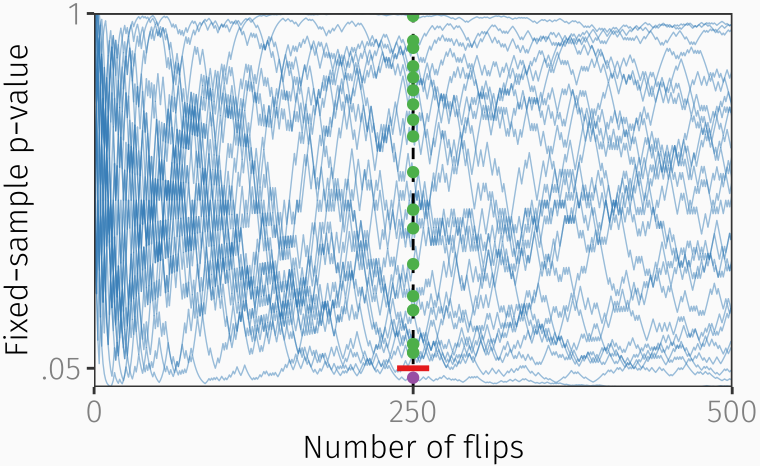


One path of p-values from a fair coin



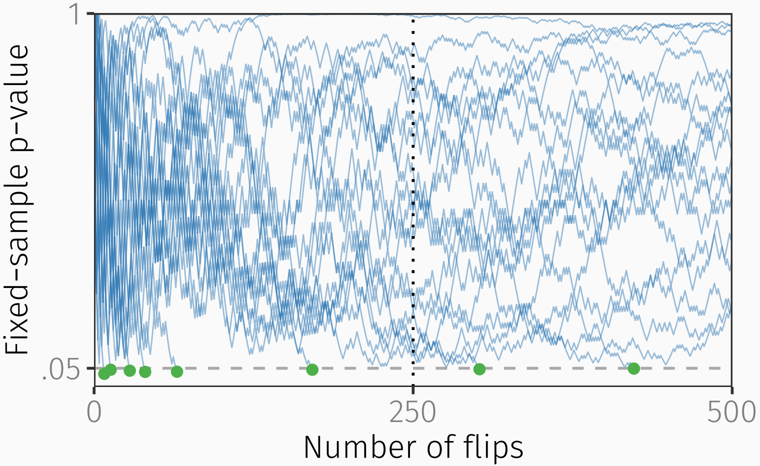
Let's look at many such paths...

With no bias, we only rarely conclude the coin is biased.



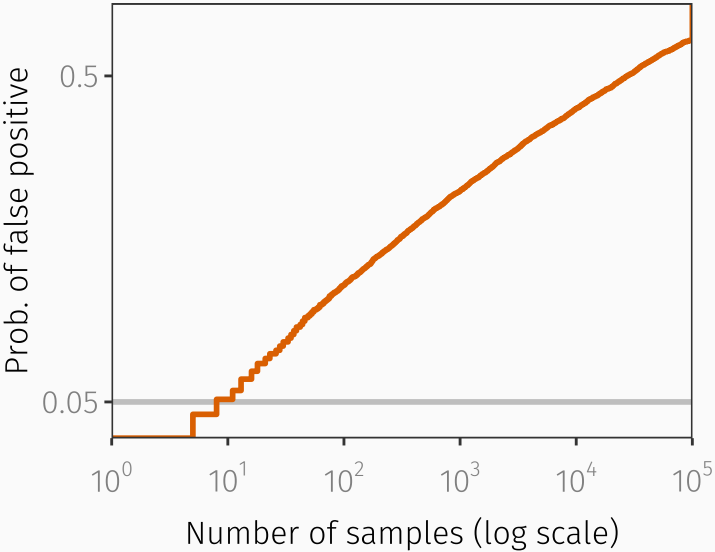
Just one out of 25 p-values is below 0.05.

Continuous monitoring of fixed-sample p-values inflates the false positive rate.



Here, with a fair coin, eight out of 25 paths reach significance.

The false positive rate grows arbitrarily large with enough flips.



Confidence sequences solve the problem of continuous monitoring.

A **confidence sequence** for $(\theta_t)_{t=1}^{\infty}$ is a sequence of intervals $(\text{CI}_t)_{t=1}^{\infty}$ satisfying

$$\mathbb{P}(\theta_t \in \text{CI}_t \text{ for all } t \in \mathbb{N}) \geq 1 - \alpha.$$

[Darling and Robbins 1967, Lai 1984, Jennison and Turnbull 1989, Johari et al. 2015, H. et al. 2021]

Confidence sequences solve the problem of continuous monitoring.

A **confidence sequence** for $(\theta_t)_{t=1}^{\infty}$ is a sequence of intervals $(\mathrm{CI}_t)_{t=1}^{\infty}$ satisfying

$$\mathbb{P}(\theta_t \in \mathrm{CI}_t \text{ for all } t \in \mathbb{N}) \geq 1 - \alpha.$$

[Darling and Robbins 1967, Lai 1984, Jennison and Turnbull 1989, Johari et al. 2015, H. et al. 2021]

Much stronger than the fixed-sample guarantee:

$$\text{For all } t \in \mathbb{N}, \mathbb{P}(\theta_t \in \mathrm{CI}_t) \geq 1 - \alpha.$$

Confidence sequences solve the problem of continuous monitoring.

A **confidence sequence** for $(\theta_t)_{t=1}^{\infty}$ is a sequence of intervals $(\mathrm{CI}_t)_{t=1}^{\infty}$ satisfying

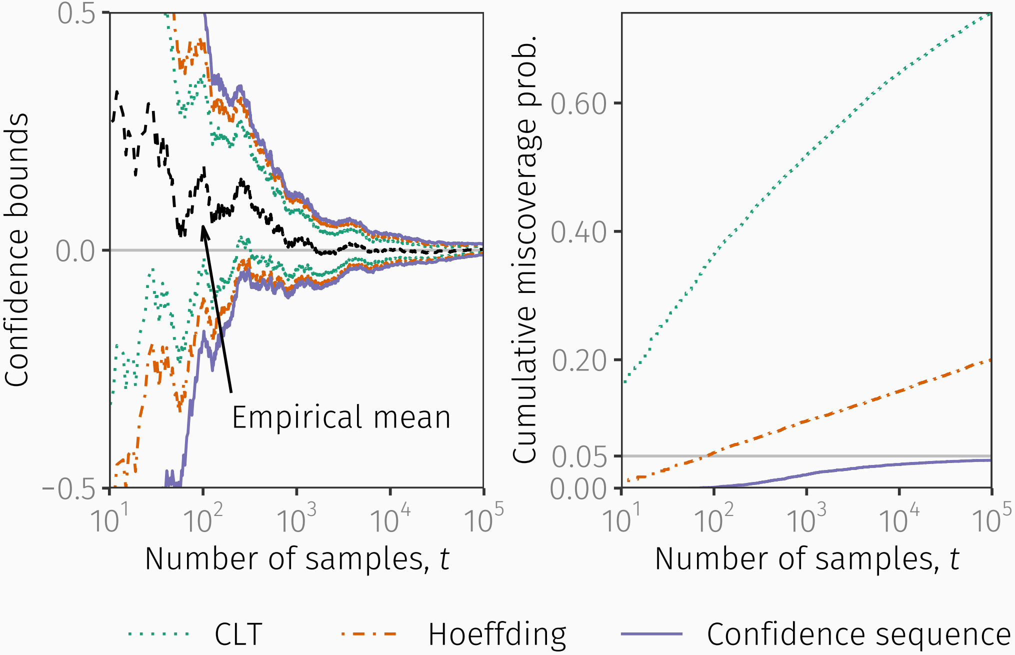
$$\mathbb{P}(\theta_t \in \mathrm{CI}_t \text{ for all } t \in \mathbb{N}) \geq 1 - \alpha.$$

[Darling and Robbins 1967, Lai 1984, Jennison and Turnbull 1989, Johari et al. 2015, H. et al. 2021]

Much stronger than the fixed-sample guarantee:

$$\text{For all } t \in \mathbb{N}, \mathbb{P}(\theta_t \in \mathrm{CI}_t) \geq 1 - \alpha.$$

In short, we achieve this by making confidence intervals wider.



0.5
0.0
−0.5
Confidence bounds
Empirical mean
10^1 10^2 10^3 10^4 10^5
Number of samples, t
Cumulative miscoverage prob.
0.60
0.40
0.20
0.05
0.00
10^1 10^2 10^3 10^4 10^5
Number of samples, t
CLT
Hoeffding
Confidence sequence

Outline

Some key results

Frequently asked questions

Confidence sequence for regression coefficients

Some key results

Empirical-Bernstein confidence sequence for bounded random variables

Theorem (H., Ramdas, McAuliffe, Sekhon 2021)

Suppose X_i are independent and $[a, b]$-valued for all i. Let $\widehat{X}_i$ be any predictable sequence and u_α be any sub-exponential uniform boundary with scale $b - a$. Then

$$\mathbb{P}\left(\left|\bar{X}_t - \mathbb{E}\bar{X}_t\right| < \frac{u_\alpha\left(\sum_{i=1}^{t}(X_i - \widehat{X}_i)^2\right)}{t} \text{ for all } t \in \mathbb{N} \right) \geq 1 - 2\alpha.$$

Empirical-Bernstein confidence sequence for bounded random variables

Theorem (H., Ramdas, McAuliffe, Sekhon 2021)

Suppose X_i are independent and $[a, b]$-valued for all i. Let $\widehat{X}_i$ be any predictable sequence and u_α be any sub-exponential uniform boundary with scale $b - a$. Then

$$\mathbb{P}\left(\left| \bar{X}_t - \mathbb{E}\bar{X}_t \right| < \frac{u_\alpha\left(\sum_{i=1}^t (X_i - \widehat{X}_i)^2 \right)}{t} \text{ for all } t \in \mathbb{N} \right) \geq 1 - 2\alpha.$$

Empirical-Bernstein confidence sequence for bounded random variables

Theorem (H., Ramdas, McAuliffe, Sekhon 2021)

Suppose X_i are independent and $[a, b]$-valued for all i. Let $\widehat{X}_i$ be any predictable sequence and u_α be any sub-exponential uniform boundary with scale $b - a$. Then

$$\mathbb{P}\left(\left|\bar{X}_t - \mathbb{E}\bar{X}_t\right| < \frac{u_\alpha\left(\sum_{i=1}^{t}(X_i - \widehat{X}_i)^2\right)}{t} \text{ for all } t \in \mathbb{N} \right) \geq 1 - 2\alpha.$$

Empirical-Bernstein confidence sequence for bounded random variables

Theorem (H., Ramdas, McAuliffe, Sekhon 2021)

Suppose X_i are independent and $[a, b]$-valued for all i. Let $\widehat{X}_i$ be any predictable sequence and u_α be any sub-exponential uniform boundary with scale $b - a$. Then

$$\mathbb{P}\left(\left|\bar{X}_t - \mathbb{E}\bar{X}_t\right| < \frac{u_\alpha\left(\sum_{i=1}^t (X_i - \widehat{X}_i)^2\right)}{t} \text{ for all } t \in \mathbb{N} \right) \geq 1 - 2\alpha.$$

Here $u_\alpha(v)$ is $\mathcal{O}(\sqrt{v \log v})$ or $\mathcal{O}(\sqrt{v \log \log v})$.

What about that boundedness assumption?

Without some similar assumption, it is impossible to construct a confidence interval.

The problem: one outlier can have arbitrarily large influence, e.g.

$$X_i = \begin{cases} 0, & \text{with probability } 1 - \epsilon, \\ 1/\epsilon, & \text{with probability } \epsilon. \end{cases} \tag{1}$$

Expectation is always one, but you need $\sim 1/\epsilon$ observations to have any idea about this.

Asymptotic arguments often sweep this issue under the rug. (In practice, though, they're usually satisfactory.)

Choosing a uniform boundary in practice

Some advice for choosing the uniform boundary function $u_\alpha(\cdot)$:

- Two-sided normal mixture is a nice starting point.
  - Simple closed form: eq. (14) of `https://arxiv.org/pdf/1810.08240.pdf`
  - Asymptotic justification: Waudby-Smith et al. (2021) "Time-uniform central limit theory with applications to anytime-valid causal inference"
- For choosing the tuning parameter ρ:
  - Sec. 3.5 of `https://arxiv.org/pdf/1810.08240.pdf`
  - Sec. 5 of `https://arxiv.org/pdf/1906.09712.pdf`
- For non-asymptotic guarantees, need to go deeper
- State-of-the-art for bounded random variables: Waudby-Smith & Ramdas (2022) "Estimating means of bounded random variables by betting"

Average treatment effect: setup

Unit i has fixed potential outcomes $Y_i(0), Y_i(1)$, for $i = 1, 2, \ldots$

Average treatment effect: setup

Unit i has fixed potential outcomes $Y_i(0), Y_i(1)$, for $i = 1, 2, \ldots$

We assign unit i randomly to treatment with probability p or control with probability $1 - p$, and observe $Y_i(1)$ or $Y_i(0)$ accordingly. [Neyman 1923, Rubin 1974]

Average treatment effect: setup

Unit i has fixed potential outcomes $Y_i(0), Y_i(1)$, for $i = 1, 2, \ldots$

We assign unit i randomly to treatment with probability p or control with probability $1 - p$, and observe $Y_i(1)$ or $Y_i(0)$ accordingly. [Neyman 1923, Rubin 1974]

Assumption: no interference

Average treatment effect: setup

Unit i has fixed potential outcomes $Y_i(0), Y_i(1)$, for $i = 1, 2, \ldots$

We assign unit i randomly to treatment with probability p or control with probability $1 - p$, and observe $Y_i(1)$ or $Y_i(0)$ accordingly. [Neyman 1923, Rubin 1974]

Assumption: no interference

Our goal: after observing units $1, \ldots, t$, we'd like to estimate

$$\mathrm{ATE}_t := \frac{1}{t} \sum_{i=1}^{t} [Y_i(1) - Y_i(0)].$$

Average treatment effect: setup

Unit i has fixed potential outcomes $Y_i(0), Y_i(1)$, for $i = 1, 2, \ldots$

We assign unit i randomly to treatment with probability p or control with probability $1 - p$, and observe $Y_i(1)$ or $Y_i(0)$ accordingly. [Neyman 1923, Rubin 1974]

Assumption: no interference

Our goal: after observing units $1, \ldots, t$, we'd like to estimate

$$\text{ATE}_t := \frac{1}{t} \sum_{i=1}^{t} [Y_i(1) - Y_i(0)].$$

Assumption: $Y_i(k) \in [0, 1]$ for $k = 0, 1$, all i.

Average treatment effect: theorem

For each unit i, X_i is unbiased for the individual treatment effect $Y_i(1) - Y_i(0)$,

$$X_i := \widehat{Y}_i(1) - \widehat{Y}_i(0) + \left(\frac{Z_i - p}{p(1-p)}\right)\left(Y_i^{\text{obs}} - \widehat{Y}_i(Z_i)\right)$$

(augmented inverse propensity weighting).

Average treatment effect: theorem

For each unit i, X_i is unbiased for the individual treatment effect $Y_i(1) - Y_i(0)$,

$$X_i := \widehat{Y}_i(1) - \widehat{Y}_i(0) + \left(\frac{Z_i - p}{p(1-p)}\right)\left(Y_i^{\text{obs}} - \widehat{Y}_i(Z_i)\right)$$

(augmented inverse propensity weighting).

Theorem (H., Ramdas, McAuliffe, Sekhon 2021)

Assume no interference and $Y_t(k) \in [0,1]$ for all k, t. Let u_α be any sub-exponential uniform boundary with scale $2/\min\{p, 1-p\}$. Then

$$\mathbb{P}\left(\left|\bar{X}_t - \text{ATE}_t\right| < \frac{u_\alpha\left(\sum_{i=1}^t (X_i - \widehat{X}_t)^2\right)}{t} \text{ for all } t \in \mathbb{N}\right) \geq 1 - \alpha.$$

The point: we can reduce ATE estimation to bounded mean estimation.

General treatment of adaptive allocation

What if assignment probability is time-changing (but predictable) P_t? Just replace p with P_t:

$$X_t := \widehat{Y}_t(1) - \widehat{Y}_t(0) + \left(\frac{Z_t - P_t}{P_t(1-P_t)}\right)\left(Y_t^{\mathrm{obs}} - \widehat{Y}_t(Z_t)\right).$$

The rest of the argument goes through fine.

General treatment of adaptive allocation

What if assignment probability is time-changing (but predictable) P_t? Just replace p with P_t:

$$X_t := \widehat{Y}_t(1) - \widehat{Y}_t(0) + \left(\frac{Z_t - P_t}{P_t(1 - P_t)}\right)\left(Y_t^{\text{obs}} - \widehat{Y}_t(Z_t)\right).$$

The rest of the argument goes through fine.

Works for any adaptive allocation scheme with exploration, i.e., all arms have positive probability of being assigned at each time. (For Thompson sampling, computing P_t may require some work.)

General treatment of adaptive allocation

What if assignment probability is time-changing (but predictable) P_t? Just replace p with P_t:

$$X_t := \widehat{Y}_t(1) - \widehat{Y}_t(0) + \left(\frac{Z_t - P_t}{P_t(1 - P_t)}\right)\left(Y_t^{\mathrm{obs}} - \widehat{Y}_t(Z_t)\right).$$

The rest of the argument goes through fine.

Works for any adaptive allocation scheme with exploration, i.e., all arms have positive probability of being assigned at each time. (For Thompson sampling, computing P_t may require some work.)

If we assume a stationary mean, can just use ordinary confidence sequence and ignore assignment probabilities.

Quantile estimation

$X_1, X_2, \ldots$ i.i.d. from any distribution F. Let q be the p^{th} quantile of F, let $\widehat{Q}_t(p)$ denote the p^{th} sample quantile at time t.

Quantile estimation

$X_1, X_2, \ldots$ i.i.d. from any distribution F. Let q be the p^{th} quantile of F, let $\widehat{Q}_t(p)$ denote the p^{th} sample quantile at time t.

Theorem

Suppose X_i are i.i.d. from any distribution F. Let $u_{\alpha,p}$ be an appropriately scaled sub-Bernoulli uniform boundary. Then

$$\mathbb{P}\left(\widehat{Q}_t\left(p - \frac{u_{\alpha,1-p}(t)}{t}\right) \leq q \leq \widehat{Q}_t\left(p + \frac{u_{\alpha,p}(t)}{t}\right) \text{ for all } t \in \mathbb{N}\right) \geq 1 - \alpha.$$

Quantile estimation

$X_1, X_2, \ldots$ i.i.d. from any distribution F. Let q be the p^{th} quantile of F, let $\widehat{Q}_t(p)$ denote the p^{th} sample quantile at time t.

Theorem

Suppose X_i are i.i.d. from any distribution F. Let $u_{\alpha,p}$ be an appropriately scaled sub-Bernoulli uniform boundary. Then

$$\mathbb{P}\left(\widehat{Q}_t\left(p - \frac{u_{\alpha,1-p}(t)}{t}\right) \leq q \leq \widehat{Q}_t\left(p + \frac{u_{\alpha,p}(t)}{t}\right) \text{ for all } t \in \mathbb{N}\right) \geq 1 - \alpha.$$

No assumption on the distribution F.

Quantile estimation

Theorem

Suppose X_i are i.i.d. from any distribution F. Let $u_{\alpha,p}$ be an appropriately scaled sub-Bernoulli uniform boundary. Then

$$\mathbb{P}\left(\widehat{Q}_t\left(p - \frac{u_{\alpha,1-p}(t)}{t}\right) \le q \le \widehat{Q}_t\left(p + \frac{u_{\alpha,p}(t)}{t}\right) \text{ for all } t \in \mathbb{N}\right) \ge 1 - \alpha.$$

Using package confseq

```
u(t) = confseq.boundaries.beta_binomial_mixture_bound(
    p * (1 - p) * t, alpha,
    g = 1 - p, h = p,
    v_opt = ...)
```

Estimation of a cumulative distribution function

Theorem

Suppose X_i are i.i.d. from any distribution F. Let $\widehat{F}_t$ denote the empirical cumulative distribution function at time t. Then

$$\mathbb{P}\left(\left\|\widehat{F}_t - F\right\|_\infty \leq A\sqrt{\frac{\log\log(et) + C}{t}} \text{ for all } t \in \mathbb{N}\right) \geq 1 - e^{-\mathcal{O}(A^2C)}.$$

Estimation of a cumulative distribution function

Theorem

Suppose X_i are i.i.d. from any distribution F. Let $\widehat{F}_t$ denote the empirical cumulative distribution function at time t. Then

$$\mathbb{P}\left(\left\|\widehat{F}_t - F\right\|_\infty \le A\sqrt{\frac{\log\log(et) + C}{t}} \text{ for all } t \in \mathbb{N}\right) \ge 1 - e^{-\mathcal{O}(A^2 C)}.$$

Using package `confseq`

```
u(t) = confseq.quantiles.empirical_process_lil_bound(
  t, alpha, t_min=1)
```

Frequently asked questions

What about the intersection CI?

If $\mathbb{P}(\theta \in \mathrm{CI}_t \text{ for all } t \in \mathbb{N}) \geq 1 - \alpha$, then it also holds that

$$\mathbb{P}\left(\theta \in \bigcap_{s \leq t} \mathrm{CI}_s \text{ for all } t \in \mathbb{N}\right) \geq 1 - \alpha, \tag{2}$$

Pro: smaller CIs

What about the intersection CI?

If $\mathbb{P}(\theta \in \text{CI}_t \text{ for all } t \in \mathbb{N}) \geq 1 - \alpha$, then it also holds that

$$\mathbb{P}\left(\theta \in \bigcap_{s \leq t} \text{CI}_s \text{ for all } t \in \mathbb{N}\right) \geq 1 - \alpha, \tag{2}$$

Pro: smaller CIs

Cons:

- Not valid for time-varying estimands (θ_t)
- Can become empty
- Not a function of sufficient statistics at time t

See sec. 6 of `https://arxiv.org/pdf/1810.08240.pdf`.

What about group sequential methods?

Key distinguishing features of group sequential methods:

- Require interim analysis schedule specifying a small number of “peeks”, planned in advance
- Justified by central limit theorem

Pro: smaller CIs

What about group sequential methods?

Key distinguishing features of group sequential methods:

- Require interim analysis schedule specifying a small number of “peeks”, planned in advance
- Justified by central limit theorem

Pro: smaller CIs

Cons:

- Require more planning and expertise
- Only valid asymptotically
- Typically involve a finite endpoint
- Not appropriate for continuous monitoring

What about Bayesian methods?

It is sometimes said that Bayesian methods are immune to peeking.

This is true only because Bayesian methods rewrite the interpretation of probability.

What about Bayesian methods?

It is sometimes said that Bayesian methods are immune to peeking.

This is true only because Bayesian methods rewrite the interpretation of probability.

- Bayesian methods will not, in general, achieve the frequentist guarantee $\mathbb{P}(\theta_t \in \mathrm{CI}_t \text{ for all } t \in \mathbb{N}) \geq 1 - \alpha$.

What about Bayesian methods?

It is sometimes said that Bayesian methods are immune to peeking.

This is true only because Bayesian methods rewrite the interpretation of probability.

- Bayesian methods will not, in general, achieve the frequentist guarantee $\mathbb{P}(\theta_t \in \mathrm{CI}_t \text{ for all } t \in \mathbb{N}) \geq 1 - \alpha$.
- Some confidence sequence methods have a Bayesian interpretation.

What about Bayesian methods?

It is sometimes said that Bayesian methods are immune to peeking.

This is true only because Bayesian methods rewrite the interpretation of probability.

- Bayesian methods will not, in general, achieve the frequentist guarantee $\mathbb{P}(\theta_t \in \mathrm{CI}_t \text{ for all } t \in \mathbb{N}) \geq 1 - \alpha$.
- Some confidence sequence methods have a Bayesian interpretation.
- Bayesian methods are a good way to introduce shrinkage, which can help with selection bias. (But they come with no guarantees about bias, which is a frequentist concept.)

What about Bayesian methods?

It is sometimes said that Bayesian methods are immune to peeking.

This is true only because Bayesian methods rewrite the interpretation of probability.

- Bayesian methods will not, in general, achieve the frequentist guarantee $\mathbb{P}(\theta_t \in \mathrm{CI}_t \text{ for all } t \in \mathbb{N}) \geq 1 - \alpha$.
- Some confidence sequence methods have a Bayesian interpretation.
- Bayesian methods are a good way to introduce shrinkage, which can help with selection bias. (But they come with no guarantees about bias, which is a frequentist concept.)

Dawid (1994) "Selection paradoxes of Bayesian inference" is a nice reference.

Confidence sequences for regression coefficients

Basic mechanism for mean estimation

- Suppose $X_1, X_2, \ldots$ are i.i.d. and we wish to estimate $\mathbb{E}X_1$.
- $S_t(\theta) := \sum_{i=1}^{t}(X_i - \theta)$ is a martingale when $\theta = \mathbb{E}X_1$.
- Uniform martingale concentration $\Rightarrow |S_t(\theta)| \leq u_t$ for all t with high probability.
    - We need some extra information about $X_i - \theta$, for example boundedness.
- Invert to get confidence sequence: $\mathrm{CI}_t := \{\theta : |S_t(\theta)| \leq u_t\}$.

Martingales from univariate score functions

We just need a function $g(x, \theta)$ such that $\mathbb{E}g(X_i, \theta) = 0$ at the desired value of θ.

- $g(X_i, \theta) = X_i - \theta$ estimates the mean $\mathbb{E}X_i$
- $g(X_i, \theta) = 1_{X_i \leq \theta} - p$ estimates the p-quantile
- In general, $g(X_i, \theta) = f'(X_i, \theta)$ estimates the population minimizer of the smooth, convex loss f.
    - Confidence sequence for any M-estimator.

Vector-valued martingales from multivariate score functions

There are uniform concentration results for vector-valued martingales, e.g.,

$$\|S_t(\theta)\|_2 \leq u_t \text{ for all } t \text{ with high probability,} \tag{3}$$

where $S_t(\theta)$ takes values in R^d.

Vector-valued martingales from multivariate score functions

There are uniform concentration results for vector-valued martingales, e.g.,

$$\|S_t(\theta)\|_2 \leq u_t \text{ for all } t \text{ with high probability,} \tag{3}$$

where $S_t(\theta)$ takes values in R^d.

Let

$$f(y_i, X_i, \theta) = \frac{1}{2}\|y_i - X_i^T\theta\|_2^2 \tag{4}$$

$$g(y_i, X_i, \theta) = \nabla f(y_i, X_i, \theta) = X_i^T(y_i - X_t^T\theta). \tag{5}$$

Then $S_t(\theta) = \sum_{i=1}^t g(y_i, X_i, \theta)$ is a martingale when θ is the population OLS coefficients.

Vector-valued martingales from multivariate score functions

There are uniform concentration results for vector-valued martingales, e.g.,

$$\|S_t(\theta)\|_2 \leq u_t \text{ for all } t \text{ with high probability,} \tag{3}$$

where $S_t(\theta)$ takes values in R^d.

Let

$$f(y_i, X_i, \theta) = \frac{1}{2}\|y_i - X_i^T\theta\|_2^2 \tag{4}$$

$$g(y_i, X_i, \theta) = \nabla f(y_i, X_i, \theta) = X_i^T(y_i - X_t^T\theta). \tag{5}$$

Then $S_t(\theta) = \sum_{i=1}^t g(y_i, X_i, \theta)$ is a martingale when θ is the population OLS coefficients.

(Straightforward in principle; implementation may be challenging.)

Summary

- Peeking is a serious problem, and confidence sequences are a flexible solution which allow continuous monitoring.

Summary

- Peeking is a serious problem, and confidence sequences are a flexible solution which allow continuous monitoring.
- We derive useful confidence sequences in a variety of nonparametric settings, including for estimating average treatment effects under adaptive allocation, quantiles, and CDFs.

Summary

- Peeking is a serious problem, and confidence sequences are a flexible solution which allow continuous monitoring.
- We derive useful confidence sequences in a variety of nonparametric settings, including for estimating average treatment effects under adaptive allocation, quantiles, and CDFs.
- Estimating vector-valued minimizers of general loss functions is a frontier.
  - For examples of vector-valued confidence sequences, see Abbasi-Yadkori et al. (2011) "Improved Algorithms for Linear Stochastic Bandits", or Corollary 10 of `https://arxiv.org/pdf/1808.03204.pdf`.

Thank you!

- steve@stevehoward.org
- Time-uniform Chernoff bounds via nonnegative supermartingales (2020): https://arxiv.org/abs/1808.03204
- Time-uniform, nonparametric, nonasymptotic confidence sequences (2021): https://arxiv.org/abs/1810.08240
- Sequential estimation of quantiles with applications to A/B-testing and best-arm identification (2022): https://arxiv.org/abs/1906.09712
- Some state-of-the-art papers on the websites of Aaditya Ramdas and Ian Waudby-Smith.
- Implementations of many uniform boundaries and confidence sequences: https://github.com/gostevehoward/confseq
- Slides: stevehoward.org

Appendix

We construct an unbiased estimator of each individual treatment effect.

We require predictions $\widehat{Y}_t(1)$ and $\widehat{Y}_t(0)$ for $Y_t(1)$ and $Y_t(0)$.

We construct an unbiased estimator of each individual treatment effect.

We require predictions $\widehat{Y}_t(1)$ and $\widehat{Y}_t(0)$ for $Y_t(1)$ and $Y_t(0)$.

For each unit t, we construct an estimator X_t of the individual treatment effect $Y_t(1) - Y_t(0)$.

X_t depends on our predictions as well as the observed treatment and outcome.

We construct an unbiased estimator of each individual treatment effect.

We require predictions $\widehat{Y}_t(1)$ and $\widehat{Y}_t(0)$ for $Y_t(1)$ and $Y_t(0)$.

For each unit t, we construct an estimator X_t of the individual treatment effect $Y_t(1) - Y_t(0)$.

X_t depends on our predictions as well as the observed treatment and outcome.

Two key properties of X_t:

1. **Unbiased**: $\mathbb{E}X_t = Y_t(1) - Y_t(0)$
2. **Variance** of X_t depends on **prediction errors** $(Y_t(1) - \widehat{Y}_t(1))^2$ and $(Y_t(0) - \widehat{Y}_t(0))^2$.

We construct an unbiased estimator of each individual treatment effect.

We require predictions $\widehat{Y}_t(1)$ and $\widehat{Y}_t(0)$ for $Y_t(1)$ and $Y_t(0)$.

For each unit t, we construct an estimator X_t of the individual treatment effect $Y_t(1) - Y_t(0)$.

X_t depends on our predictions as well as the observed treatment and outcome.

Two key properties of X_t:

1. **Unbiased**: $\mathbb{E}X_t = Y_t(1) - Y_t(0)$
2. **Variance** of X_t depends on **prediction errors** $(Y_t(1) - \widehat{Y}_t(1))^2$ and $(Y_t(0) - \widehat{Y}_t(0))^2$.

Let $S_t = \sum_{i=1}^{t} X_i$. Then S_t/t is unbiased for ATE_t.

CLT confidence bounds

$$\frac{S_t}{t} \pm \frac{1.96\sqrt{\sum_{i=1}^{t}(X_i - \bar{X}_t)^2}}{t}$$

CLT confidence bounds

$$\frac{S_t}{t} \pm \frac{1.96\sqrt{\sum_{i=1}^{t}(X_i - \bar{X}_t)^2}}{t}$$

Uniform, non-asymptotic confidence bounds

$$\frac{S_t}{t} \pm \frac{u_\alpha\left(\sum_{i=1}^{t}(X_i - \widehat{X}_i)^2\right)}{t} \quad \text{where } \widehat{X}_i = \widehat{Y}_i(1) - \widehat{Y}_i(0).$$

CLT confidence bounds

$$\frac{S_t}{t} \pm \frac{1.96\sqrt{\sum_{i=1}^{t}(X_i - \bar{X}_t)^2}}{t}$$

Uniform, non-asymptotic confidence bounds

$$\frac{S_t}{t} \pm \frac{u_\alpha\left(\sum_{i=1}^{t}(X_i - \widehat{X}_i)^2\right)}{t} \quad \text{where } \widehat{X}_i = \widehat{Y}_i(1) - \widehat{Y}_i(0).$$

- $\sum_{i=1}^{t}(X_i - \widehat{X}_i)^2$ is an “online” estimate of $\mathrm{Var}(S_t)$.
- Estimation precision depends on prediction accuracy.

CLT confidence bounds

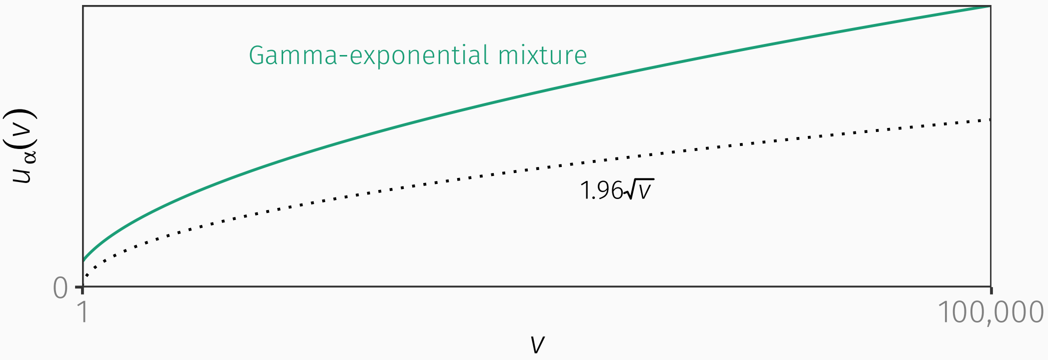
$$\frac{S_t}{t} \pm \frac{1.96\sqrt{\sum_{i=1}^{t}(X_i - \bar{X}_t)^2}}{t}$$

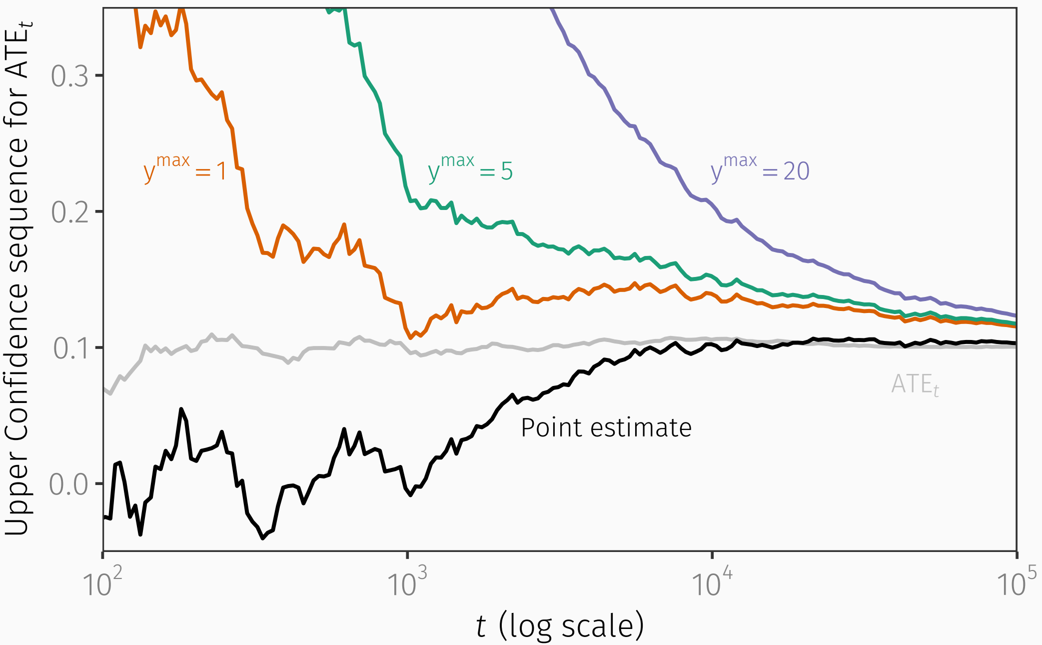
Uniform, non-asymptotic confidence bounds

$$\frac{S_t}{t} \pm \frac{u_\alpha\left(\sum_{i=1}^{t}(X_i - \widehat{X}_i)^2\right)}{t} \qquad \text{where } \widehat{X}_i = \widehat{Y}_i(1) - \widehat{Y}_i(0).$$

- $\sum_{i=1}^{t}(X_i - \widehat{X}_i)^2$ is an "online" estimate of $\mathrm{Var}(S_t)$.
- Estimation precision depends on prediction accuracy.
- $u_\alpha(v) = \mathcal{O}(\sqrt{v \log v})$, so $u_\alpha(v)$ is like $z_{1-\alpha}\sqrt{v}$, but the "z-factor" grows over time (slowly).

The uniform boundary grows only slightly faster than $\mathcal{O}(\sqrt{n})$




Upper Confidence sequence for ATE$_t$
0.3
0.2
0.1
0.0
$y^{max} = 1$
$y^{max} = 5$
$y^{max} = 20$
ATE$_t$
Point estimate
10^2
10^3
10^4
10^5
t (log scale)

Confidence sequences are instrumental to many best-arm identification algorithms.

Multi-armed bandit: set of distributions indexed by $k = 1, \ldots, K$. Distribution k has mean μ_k.

Confidence sequences are instrumental to many best-arm identification algorithms.

Multi-armed bandit: set of distributions indexed by $k = 1, \ldots, K$. Distribution k has mean μ_k.

ϵ-best-arm identification with fixed confidence $1 - \delta$: choose an arm $k_\star$ such that, with probability at least $1 - \delta$, we have $\mu_{k_\star} \geq \max_k \mu_k - \epsilon$. [Even-Dar et al. 2002]

Confidence sequences are instrumental to many best-arm identification algorithms.

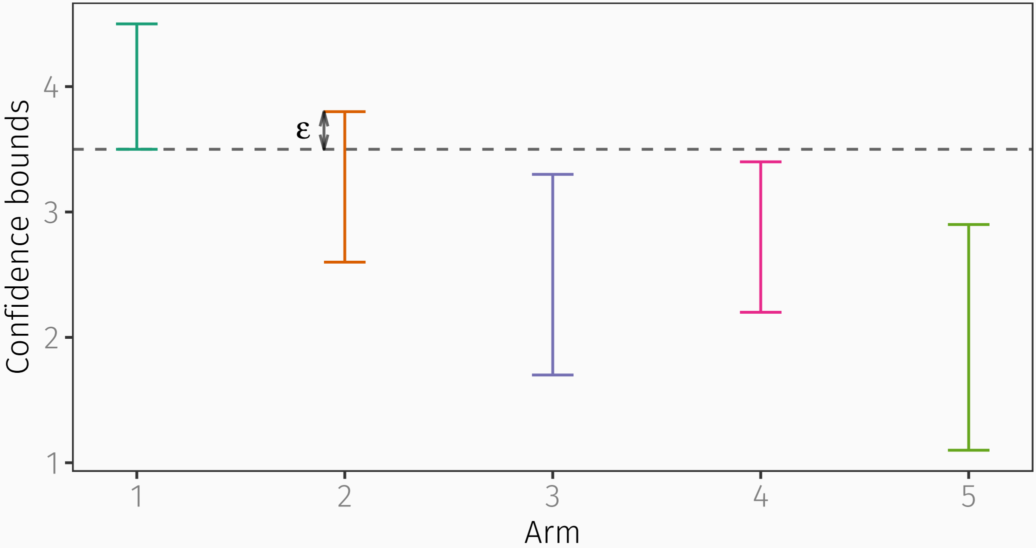
Multi-armed bandit: set of distributions indexed by $k = 1, \ldots, K$. Distribution k has mean μ_k.

ϵ-best-arm identification with fixed confidence $1 - \delta$: choose an arm $k_\star$ such that, with probability at least $1 - \delta$, we have $\mu_{k_\star} \geq \max_k \mu_k - \epsilon$. [Even-Dar et al. 2002]

Common strategy:

- Construct confidence sequence $(L_{kt}, U_{kt})_{t=1}^{\infty}$ for each arm k, so that $L_{kt} \leq \mu_k \leq U_{kt}$ for all k, t with probability at least $1 - \delta$.
- Stop the first time there exists some $k_\star$ such that

$$L_{k_\star t} \geq U_{kt} - \epsilon \text{ for all } k \neq k_\star.$$



Confidence bounds
4
3
2
1
ε
1
2
3
4
5
Arm

Arm differences can be examined directly.

Comparing confidence intervals for *individual arm means* is wasteful.

Better to compute confidence intervals on *pairwise differences of arm means* directly.

- Related to asymptotically optimal stopping rule of Garivier and Kaufmann (2016).

Arm differences can be examined directly.

Comparing confidence intervals for *individual arm means* is wasteful.

Better to compute confidence intervals on *pairwise differences of arm means* directly.

- Related to asymptotically optimal stopping rule of Garivier and Kaufmann (2016).

Can be done in our framework, allowing more efficient stopping in nonparametric settings.

- Complete theory is work in progress

Quantile best-arm identification

Let $Q_k(p)$ denote the p^{th} quantile of arm k.

Quantile ϵ-best-arm identification with fixed confidence $1 - \delta$: choose an arm $k_\star$ such that, with probability at least $1 - \delta$, we have $Q_{k_\star}(p + \epsilon) \geq \max_k Q_k(p)$. [Szörényi et al. 2015]

Quantile best-arm identification

Let $Q_k(p)$ denote the p^{th} quantile of arm k.

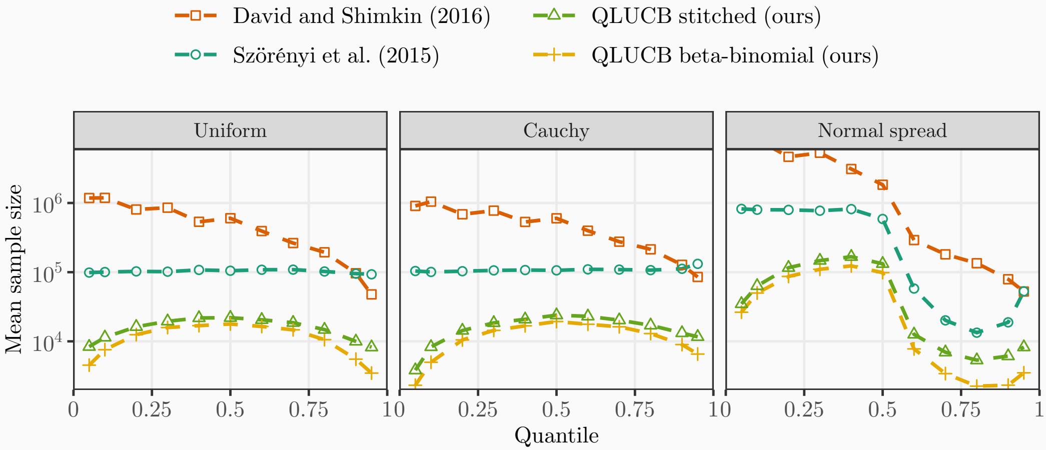
Quantile ϵ-best-arm identification with fixed confidence $1-\delta$: choose an arm $k_\star$ such that, with probability at least $1-\delta$, we have $Q_{k_\star}(p+\epsilon) \geq \max_k Q_k(p)$. [Szörényi et al. 2015]

QLUCB algorithm [H. and Ramdas 2019]: at each round,

- sample arm h_t with highest LCB for $Q_k(p+\epsilon)$;
- sample arm with highest UCB for $Q_k(p)$, excluding h_t; and
- stop when LCB for $Q_k(p+\epsilon)$ is above UCB for $Q_j(p)$ for all $j \neq k$, for some k.

[cf. Kalyanakrishnan et al. 2012]

Quantile best-arm simulations



A taste of the underlying framework

Reminder: sub-Gaussianity and Hoeffding bound

A random variable X is *sub-Gaussian* with variance parameter σ^2 if

$$\log \mathbb{E} e^{\lambda X} \leq \frac{\lambda^2 \sigma^2}{2} \quad \text{for all } \lambda \in \mathbb{R}. \tag{6}$$

Reminder: sub-Gaussianity and Hoeffding bound

A random variable X is *sub-Gaussian* with variance parameter σ^2 if

$$\log \mathbb{E} e^{\lambda X} \leq \frac{\lambda^2 \sigma^2}{2} \quad \text{for all } \lambda \in \mathbb{R}. \tag{6}$$

Hoeffding bound (1963): if $X_i \in [0, 1]$ independent, $i = 1, \ldots, t$, then

$$\mathbb{P}\left(\sum_{i=1}^{t} (X_i - \mathbb{E} X_i) \geq \sqrt{\frac{t \log \alpha^{-1}}{2}} \right) \leq \alpha. \tag{7}$$

Reminder: sub-Gaussianity and Hoeffding bound

A random variable X is *sub-Gaussian* with variance parameter σ^2 if

$$\log \mathbb{E} e^{\lambda X} \leq \frac{\lambda^2 \sigma^2}{2} \quad \text{for all } \lambda \in \mathbb{R}. \tag{6}$$

Hoeffding bound (1963): if $X_i \in [0, 1]$ independent, $i = 1, \ldots, t$, then

$$\mathbb{P}\left(\sum_{i=1}^{t} (X_i - \mathbb{E} X_i) \geq \sqrt{\frac{t \log \alpha^{-1}}{2}} \right) \leq \alpha. \tag{7}$$

Proof involves two main pieces:

1. Show X_i is sub-Gaussian with variance parameter 1/4, and
2. Use Cramér-Chernoff method to obtain (7) from (6).

Reminder: sub-Gaussianity and Hoeffding bound

A random variable X is *sub-Gaussian* with variance parameter σ^2 if

$$\log \mathbb{E} e^{\lambda X} \leq \psi(\lambda)\sigma^2 \quad \text{for all } \lambda \in \mathbb{R}. \qquad \text{Here } \psi(\lambda) = \frac{\lambda^2}{2} \tag{6}$$

Hoeffding bound (1963): if $X_i \in [0, 1]$ independent, $i = 1, \ldots, t$, then

$$\mathbb{P}\left(\sum_{i=1}^{t}(X_i - \mathbb{E}X_i) \geq u_\psi(t\sigma^2)\right) \leq \alpha. \qquad \text{Here } u_\psi(v) = \sqrt{2v \log \alpha^{-1}}, \ \ \sigma^2 = \frac{1}{4} \tag{7}$$

Proof involves two main pieces:

1. Show X_i is sub-Gaussian with variance parameter 1/4, and
2. Use Cramér-Chernoff method to obtain (7) from (6).

Our recipe for confidence sequences

1. Assume a null hypothesis $H_{\theta_0} : \theta = \theta_0$.

Our recipe for confidence sequences

1. Assume a null hypothesis $H_{\theta_0} : \theta = \theta_0$.
2. Show process $S_t^{\theta_0}$ is *sub-ψ with variance process* $V_t^{\theta_0}$ under H_{θ_0}.

Our recipe for confidence sequences

1. Assume a null hypothesis $H_{\theta_0} : \theta = \theta_0$.
2. Show process $S_t^{\theta_0}$ is *sub-ψ with variance process* $V_t^{\theta_0}$ under H_{θ_0}.
3. Choose any *sub-ψ uniform boundary* $u_\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$. Then, under H_{θ_0},

$$\mathbb{P}\left(S_t^{\theta_0} \geq u_\alpha(V_t^{\theta_0}) \text{ for some } t \in \mathbb{N}\right) \leq \alpha.$$

Our recipe for confidence sequences

1. Assume a null hypothesis $H_{\theta_0} : \theta = \theta_0$.
2. Show process $S_t^{\theta_0}$ is *sub-ψ with variance process* $V_t^{\theta_0}$ under H_{θ_0}.
3. Choose any *sub-ψ uniform boundary* $u_\alpha : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$. Then, under H_{θ_0},

$$\mathbb{P}\left(S_t^{\theta_0} \geq u_\alpha(V_t^{\theta_0}) \text{ for some } t \in \mathbb{N}\right) \leq \alpha.$$

4. At time t, a confidence set for θ is

$$\mathrm{CI}_t = \left\{\theta_0 \in \mathbb{R} : S_t^{\theta_0} < u_\alpha(V_t^{\theta_0})\right\}.$$

Example: testing the bias of a coin

Suppose a coin has bias p. Let S_t be the number of heads observed up to time t.

Example: testing the bias of a coin

Suppose a coin has bias p. Let S_t be the number of heads observed up to time t.

The process $S_t - tp$ is sub-Gaussian with variance process $V_t = t/4$.

Example: testing the bias of a coin

Suppose a coin has bias p. Let S_t be the number of heads observed up to time t.

The process $S_t - tp$ is sub-Gaussian with variance process $V_t = t/4$.

Then, for any $\lambda > 0$,

$$\mathbb{P}\left(S_t - tp \geq \underbrace{\frac{\log \alpha^{-1}}{\lambda} + \frac{\lambda}{2} \cdot \frac{t}{4}}_{\text{A linear boundary}} \text{ for some } t \in \mathbb{N} \right) \leq \alpha.$$

Example: testing the bias of a coin

Suppose a coin has bias p. Let S_t be the number of heads observed up to time t.

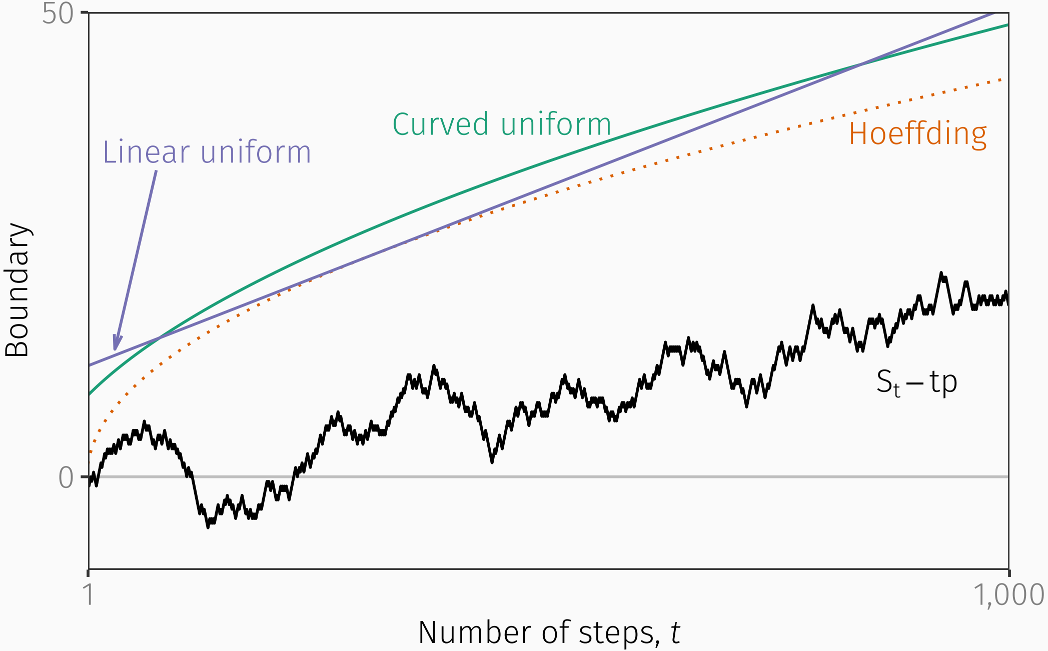
The process $S_t - tp$ is sub-Gaussian with variance process $V_t = t/4$.

Then, for any $\lambda > 0$,

$$\mathbb{P}\left(S_t - tp \geq \underbrace{\frac{\log \alpha^{-1}}{\lambda} + \frac{\lambda}{2} \cdot \frac{t}{4}}_{\text{A linear boundary}} \text{ for some } t \in \mathbb{N} \right) \leq \alpha.$$

This yields the confidence sequence

$$\left| \frac{S_t}{t} - p \right| < \frac{\log \alpha^{-1}}{\lambda t} + \frac{\lambda}{8}, \quad \text{for all } t\text{, with probability at least } 1 - 2\alpha.$$



50
0
Boundary
Curved uniform
Linear uniform
Hoeffding
$S_t - tp$
1
1,000
Number of steps, t

Our framework includes self-normalized bounds.

Suppose X_i has a symmetric distribution conditional on $X_1, \ldots, X_{i-1}$.

Our framework includes self-normalized bounds.

Suppose X_i has a symmetric distribution conditional on $X_1, \ldots, X_{i-1}$.

Then $S_t = \sum_{i=1}^t X_i$ is sub-Gaussian with variance process $V_t = \sum_{i=1}^t X_i^2$.

[de la Peña 1999]

Our framework includes self-normalized bounds.

Suppose X_i has a symmetric distribution conditional on $X_1, \ldots, X_{i-1}$.

Then $S_t = \sum_{i=1}^t X_i$ is sub-Gaussian with variance process $V_t = \sum_{i=1}^t X_i^2$.

[de la Peña 1999]

So

$$\mathbb{P}\left(\sum_{i=1}^t X_i \geq u_\alpha\left(\sum_{i=1}^t X_i^2\right) \text{ for some } t\right) \leq \alpha.$$

Our framework includes matrix bounds.

Suppose X_i is a Hermitian, $d \times d$ random matrix with a conditionally symmetric distribution, $X_i \sim -X_i \mid X_1, \ldots, X_{i-1}$.

Our framework includes matrix bounds.

Suppose X_i is a Hermitian, $d \times d$ random matrix with a conditionally symmetric distribution, $X_i \sim -X_i \mid X_1, \ldots, X_{i-1}$.

Then $S_t = \gamma_{\max}\left(\sum_{i=1}^t X_i\right)$ is sub-Gaussian with variance process $V_t = \gamma_{\max}\left(\sum_{i=1}^t X_i^2\right)$. [de la Peña 1999, Tropp 2011]

Our framework includes matrix bounds.

Suppose X_i is a Hermitian, $d \times d$ random matrix with a conditionally symmetric distribution, $X_i \sim -X_i \mid X_1, \ldots, X_{i-1}$.

Then $S_t = \gamma_{\max}\left(\sum_{i=1}^t X_i\right)$ is sub-Gaussian with variance process $V_t = \gamma_{\max}\left(\sum_{i=1}^t X_i^2\right)$. [de la Peña 1999, Tropp 2011]

So

$$\mathbb{P}\left(\gamma_{\max}\left(\sum_{i=1}^t X_i\right) \geq u_{\alpha,d}\left(\gamma_{\max}\left(\sum_{i=1}^t X_i^2\right)\right) \text{ for some } t\right) \leq \alpha.$$