Sequential (Interactive) Testing in High Dimensions

Matthew Malloy

University of Wisconsin-Madison
Department of Electrical
and Computer Engineering

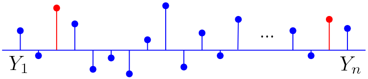
http://homepages.cae.wisc.edu/ mmalloy/

BIRS 2012 - Interactive Information Theory
Joint work with Rob Nowak

Problem Setup

- Consider a support set $\mathcal{S} \subset \{1, ..., n\}$ with $|\mathcal{S}| = s \ll n$ and the random variables

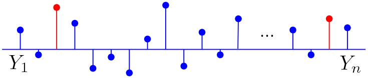
$$Y_i \sim \begin{cases} P_0 & i \notin \mathcal{S} \\ P_1 & i \in \mathcal{S} \end{cases} \qquad i = 1, ..., n$$



Problem Setup

- Consider a support set $\mathcal{S} \subset \{1, ..., n\}$ with $|\mathcal{S}| = s \ll n$ and the random variables

$$Y_i \sim \begin{cases} P_0 & i \notin \mathcal{S} \\ P_1 & i \in \mathcal{S} \end{cases} \qquad i = 1, ..., n$$

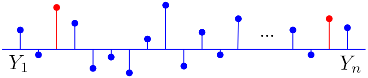


- **Goal:** exact recovery of $\mathcal{S}$ from multiple i.i.d. samples of each random variable Y_i, $i = 1, ..., n$.

Problem Setup

- Consider a support set $\mathcal{S} \subset \{1, ..., n\}$ with $|\mathcal{S}| = s \ll n$ and the random variables

$$Y_i \sim \begin{cases} P_0 & i \notin \mathcal{S} \\ P_1 & i \in \mathcal{S} \end{cases} \qquad i = 1, ..., n$$

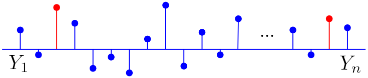


- **Goal:** exact recovery of $\mathcal{S}$ from multiple i.i.d. samples of each random variable Y_i, $i = 1, ..., n$.
  - Conventional non-sequential approach: take a fixed number of samples of each index i

Problem Setup

- Consider a support set $\mathcal{S} \subset \{1, \ldots, n\}$ with $|\mathcal{S}| = s \ll n$ and the random variables

$$Y_i \sim \begin{cases} P_0 & i \notin \mathcal{S} \\ P_1 & i \in \mathcal{S} \end{cases} \qquad i = 1, \ldots, n$$

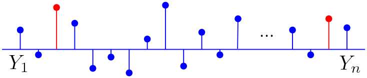


- **Goal:** exact recovery of $\mathcal{S}$ from multiple i.i.d. samples of each random variable Y_i, $i = 1, \ldots, n$.
  - Conventional non-sequential approach: take a fixed number of samples of each index i
  - Sequential approach: decision to re-sample Y_i based on prior observation

Problem Setup

- Consider a support set $\mathcal{S} \subset \{1, ..., n\}$ with $|\mathcal{S}| = s \ll n$ and the random variables

$$Y_i \sim \begin{cases} P_0 & i \notin \mathcal{S} \\ P_1 & i \in \mathcal{S} \end{cases} \qquad i = 1, ..., n$$



- **Goal:** exact recovery of $\mathcal{S}$ from multiple i.i.d. samples of each random variable Y_i, $i = 1, ..., n$.
  - Conventional non-sequential approach: take a fixed number of samples of each index i
  - Sequential approach: decision to re-sample Y_i based on prior observation
- **Definition:** take on average m samples of each index:

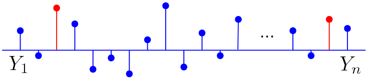
$$m := \mathbb{E}\left[\sum_{i=1}^{n} J_i\right] / n$$

where J_i is a r.v. representing number of times index i is sampled.

Problem Setup

- Consider a support set $\mathcal{S} \subset \{1, \ldots, n\}$ with $|\mathcal{S}| = s \ll n$ and the random variables

$$Y_i \sim \begin{cases} P_0 & i \notin \mathcal{S} \\ P_1 & i \in \mathcal{S} \end{cases} \qquad i = 1, \ldots, n$$



- **Goal:** exact recovery of $\mathcal{S}$ from multiple i.i.d. samples of each random variable Y_i, $i = 1, \ldots, n$.
  - Conventional non-sequential approach: take a fixed number of samples of each index i
  - Sequential approach: decision to re-sample Y_i based on prior observation
- **Definition:** take on average m samples of each index:

$$m := \mathbb{E}\left[\sum_{i=1}^{n} J_i\right] / n$$

where J_i is a r.v. representing number of times index i is sampled.

- **This talk:** find relationship between (n, s, m) such that $\mathbb{P}(\hat{\mathcal{S}} = \mathcal{S}) \to 1$.

Related work

A. Wald and J. Wolfowitz
Optimum character of the sequential probability ratio test. 1948. 1-dimensional simple binary hypothesis test

E. Posner
Optimal Search Procedures. 1963. AWGN, $s = 1$, SPRT

J. Haupt, R. Castro, and R. Nowak
Distilled Sensing: Adaptive Sampling for Sparse Detection and Estimation. 2010. AWGN, $Y_{i,j} = x_i + \gamma_{i,j}^{-1/2} W_{i,j}$, FDP/NDP

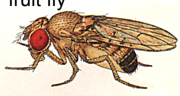
L. Lai, H. Vincent Poor, Y. Xin, and G. Georgiadis
Quickest Search Over Multiple Sequences. Trans. on Info Theory. 2011. Find one element in $\mathcal{S}$

E. Bashan, G. Newstadt, and A. Hero
Two-Stage Multi-Scale Search for Sparse Targets. 2011. AWGN, two stage procedure

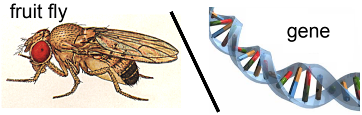
A. Tajer, R. Castro
Adaptive Spectrum Sensing for Agile Cognitive Radios. 2010.
Spectrum Sensing

Motivation – Gene knock-out studies in biology

fruit fly

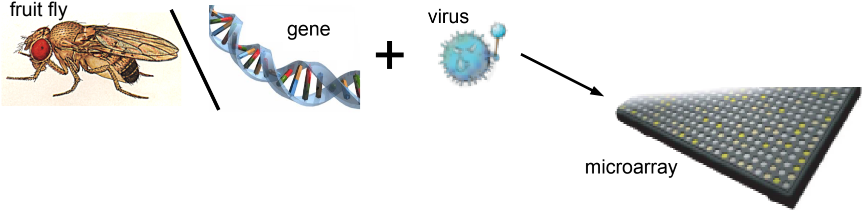


Motivation – Gene knock-out studies in biology

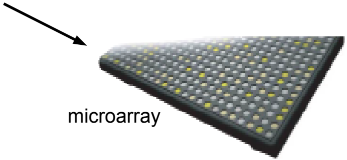


Motivation – Gene knock-out studies in biology

Motivation – Gene knock-out studies in biology

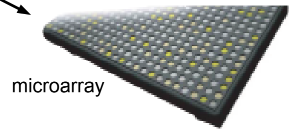


Motivation – Gene knock-out studies in biology



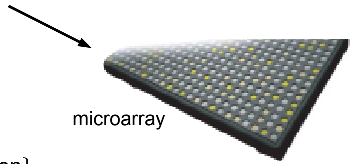
- $n = 13,071$ single gene knock-outs

Motivation – Gene knock-out studies in biology



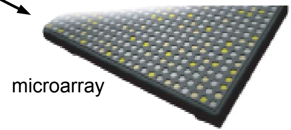
- $n = 13,071$ single gene knock-outs
- $\mathcal{S} = \{\sim$ 100 genes vital to influenza replication$\}$

Motivation – Gene knock-out studies in biology



- $n = 13,071$ single gene knock-outs
- $\mathcal{S} = \{\sim 100$ genes vital to influenza replication$\}$
- $Y_i =$ noisy observed fluorescence levels of gene knockout i

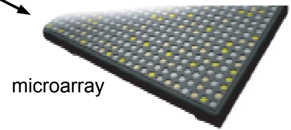
Motivation – Gene knock-out studies in biology



- $n = 13,071$ single gene knock-outs
- $\mathcal{S} = \{\sim 100$ genes vital to influenza replication$\}$
- $Y_i =$ noisy observed fluorescence levels of gene knockout i

Goal: find genes vital to virus replication (recover $\mathcal{S}$)

Motivation – Gene knock-out studies in biology

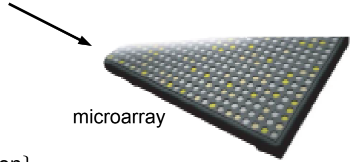


- $n = 13,071$ single gene knock-outs
- $\mathcal{S} = \{\sim 100$ genes vital to influenza replication$\}$
- $Y_i =$ noisy observed fluorescence levels of gene knockout i

Goal: find genes vital to virus replication (recover $\mathcal{S}$)

Challenges: n large, observations noisy, distributions of fluorescence unknown

Motivation – Gene knock-out studies in biology



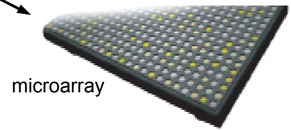
- $n = 13,071$ single gene knock-outs
- $\mathcal{S} = \{\sim 100$ genes vital to influenza replication$\}$
- $Y_i =$ noisy observed fluorescence levels of gene knockout i

Goal: find genes vital to virus replication (recover $\mathcal{S}$)

Challenges: n large, observations noisy, distributions of fluorescence unknown

1. *Non-sequential*

Motivation – Gene knock-out studies in biology



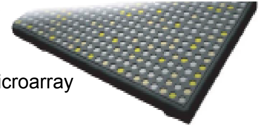
- $n = 13,071$ single gene knock-outs
- $\mathcal{S} = \{\sim 100$ genes vital to influenza replication$\}$
- $Y_i =$ noisy observed fluorescence levels of gene knockout i

Goal: find genes vital to virus replication (recover $\mathcal{S}$)

Challenges: n large, observations noisy, distributions of fluorescence unknown

1. *Non-sequential*
   - Measure each gene knock-out a fixed number of times.

Motivation – Gene knock-out studies in biology



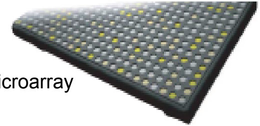
- $n = 13,071$ single gene knock-outs
- $\mathcal{S} = \{\sim 100$ genes vital to influenza replication$\}$
- $Y_i =$ noisy observed fluorescence levels of gene knockout i

Goal: find genes vital to virus replication (recover $\mathcal{S}$)

Challenges: n large, observations noisy, distributions of fluorescence unknown

1. *Non-sequential*
   - Measure each gene knock-out a fixed number of times.
2. *Sequential* [L. Hao et.al. Drosophilia RNAi screen identifies host genes important to influenza virus replication, *Nature 08*]

Motivation – Gene knock-out studies in biology



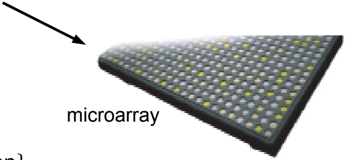
- $n = 13,071$ single gene knock-outs
- $\mathcal{S} = \{\sim 100$ genes vital to influenza replication$\}$
- $Y_i =$ noisy observed fluorescence levels of gene knockout i

Goal: find genes vital to virus replication (recover $\mathcal{S}$)

Challenges: n large, observations noisy, distributions of fluorescence unknown

1. *Non-sequential*
   - Measure each gene knock-out a fixed number of times.
2. *Sequential* [L. Hao et.al. Drosophilia RNAi screen identifies host genes important to influenza virus replication, *Nature 08*]
   - Round 1: observe all 13k knock-outs, twice. Retain the approximately 1k with highest fluorescence (13k - 1k).

Motivation – Gene knock-out studies in biology



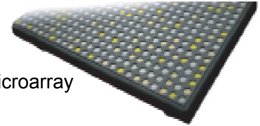
- $n = 13,071$ single gene knock-outs
- $\mathcal{S} = \{\sim 100$ genes vital to influenza replication$\}$
- $Y_i =$ noisy observed fluorescence levels of gene knockout i

Goal: find genes vital to virus replication (recover $\mathcal{S}$)

Challenges: n large, observations noisy, distributions of fluorescence unknown

1. *Non-sequential*
   - Measure each gene knock-out a fixed number of times.
2. *Sequential* [L. Hao et.al. Drosophilia RNAi screen identifies host genes important to influenza virus replication, *Nature 08*]
   - Round 1: observe all 13k knock-outs, twice. Retain the approximately 1k with highest fluorescence (13k - 1k).
   - Round 2: measure remaining strains 10 times each. Retain only those with 'significant' fluorescence.

Motivation – Gene knock-out studies in biology



- $n = 13,071$ single gene knock-outs
- $\mathcal{S} = \{\sim 100$ genes vital to influenza replication$\}$
- $Y_i =$ noisy observed fluorescence levels of gene knockout i

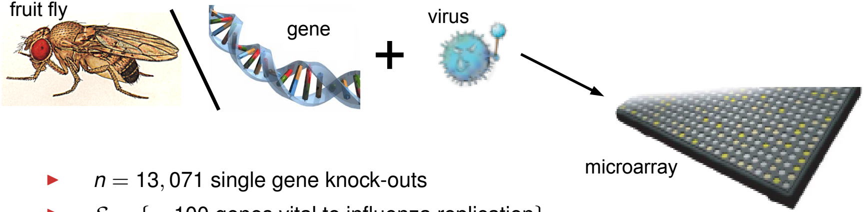
Goal: find genes vital to virus replication (recover $\mathcal{S}$)

Challenges: n large, observations noisy, distributions of fluorescence unknown

1. *Non-sequential*
   - Measure each gene knock-out a fixed number of times.
2. *Sequential* [L. Hao et.al. Drosophilia RNAi screen identifies host genes important to influenza virus replication, *Nature 08*]
   - Round 1: observe all 13k knock-outs, twice. Retain the approximately 1k with highest fluorescence (13k - 1k).
   - Round 2: measure remaining strains 10 times each. Retain only those with 'significant' fluorescence.

How much can we gain from sequential testing?

Motivation – Gene knock-out studies in biology



- $n = 13,071$ single gene knock-outs
- $\mathcal{S} = \{\sim 100$ genes vital to influenza replication$\}$
- $Y_i =$ noisy observed fluorescence levels of gene knockout i

Goal: find genes vital to virus replication (recover $\mathcal{S}$)

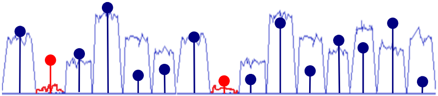
Challenges: n large, observations noisy, distributions of fluorescence unknown

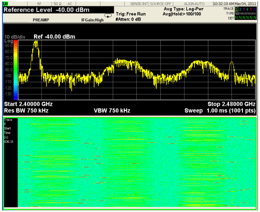
1. *Non-sequential*
   - Measure each gene knock-out a fixed number of times.
2. *Sequential* [L. Hao et.al. Drosophilia RNAi screen identifies host genes important to influenza virus replication, *Nature 08*]
   - Round 1: observe all 13k knock-outs, twice. Retain the approximately 1k with highest fluorescence (13k - 1k).
   - Round 2: measure remaining strains 10 times each. Retain only those with 'significant' fluorescence.

How much can we gain from sequential testing?

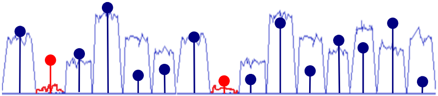
What can we do with limited knowledge of distributions?

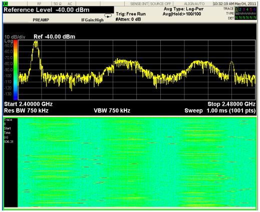
Motivation – Spectrum Sensing





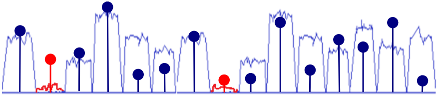
Motivation – Spectrum Sensing

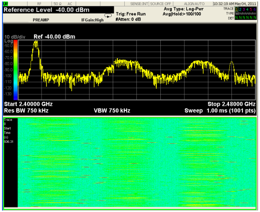




- n = number of channels under consideration

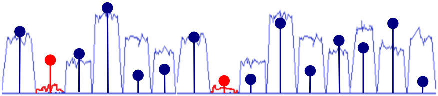
Motivation – Spectrum Sensing

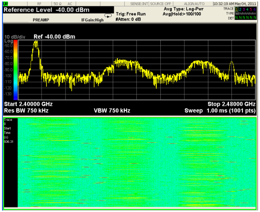




- $n =$ number of channels under consideration
- $\mathcal{S} = \{$vacant channels$\}$

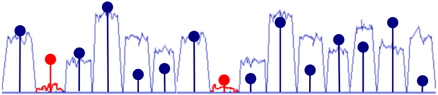
Motivation – Spectrum Sensing

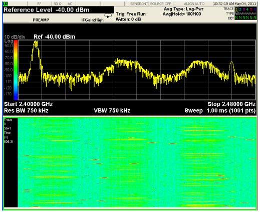




- $n =$ number of channels under consideration
- $\mathcal{S} = \{$vacant channels$\}$
- $Y_i \sim \mathcal{CN}(0, \mathrm{SNR}_i + 1)$

Motivation – Spectrum Sensing

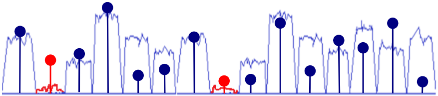


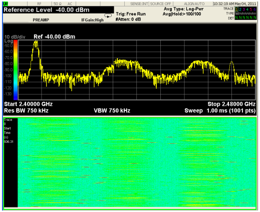


- $n =$ number of channels under consideration
- $\mathcal{S} = \{\text{vacant channels}\}$
- $Y_i \sim \mathcal{CN}(0, \mathrm{SNR}_i + 1)$

Goal: map spectral occupation (recover $\mathcal{S}$)

Motivation – Spectrum Sensing



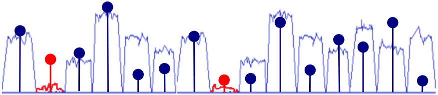


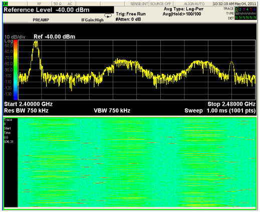
- $n =$ number of channels under consideration
- $\mathcal{S} = \{$vacant channels$\}$
- $Y_i \sim \mathcal{CN}(0, \mathrm{SNR}_i + 1)$

Goal: map spectral occupation (recover $\mathcal{S}$)

Challenges: noise floor is unknown (P_1 not fully known)

Motivation – Spectrum Sensing





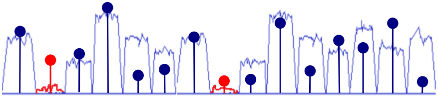
- $n =$ number of channels under consideration
- $\mathcal{S} = \{\text{vacant channels}\}$
- $Y_i \sim \mathcal{CN}(0, \mathrm{SNR}_i + 1)$

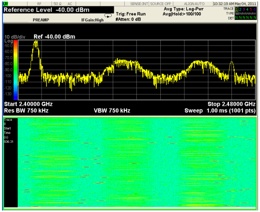
Goal: map spectral occupation (recover $\mathcal{S}$)

Challenges: noise floor is unknown (P_1 not fully known)

1. *Non-sequential*
    - Measure each channel a fixed number of times

Motivation – Spectrum Sensing





- $n =$ number of channels under consideration
- $\mathcal{S} = \{\text{vacant channels}\}$
- $Y_i \sim \mathcal{CN}(0, \mathrm{SNR}_i + 1)$

Goal: map spectral occupation (recover $\mathcal{S}$)

Challenges: noise floor is unknown (P_1 not fully known)

1. *Non-sequential*
   - Measure each channel a fixed number of times
2. *Sequential*
   - sequential thresholding or coordinate-wise SPRT [A. Tajer 2010, W. Zhang 2010, M. Malloy 2011]

Limitations of non-sequential testing

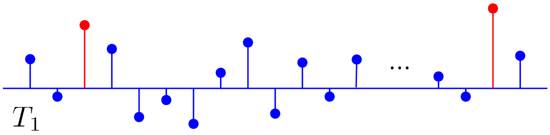
$$T_i^{(m)} := \frac{1}{m} \sum_{j=1}^{m} \log \frac{p_1(Y_{i,j})}{p_0(Y_{i,j})}$$

Limitations of non-sequential testing

$$T_i^{(m)} := \frac{1}{m}\sum_{j=1}^{m} \log \frac{p_1(Y_{i,j})}{p_0(Y_{i,j})} \quad \xrightarrow{a.s.} \quad \begin{cases} -D(P_0||P_1) & i \notin \mathcal{S} \\ D(P_1||P_0) & i \in \mathcal{S} \end{cases}$$

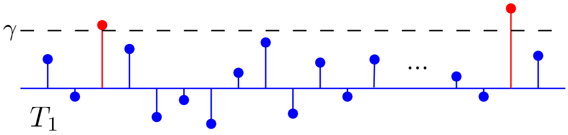
Limitations of non-sequential testing

$$T_i^{(m)} := \frac{1}{m}\sum_{j=1}^{m} \log \frac{p_1(Y_{i,j})}{p_0(Y_{i,j})} \xrightarrow{a.s.} \begin{cases} -D(P_0||P_1) & i \notin \mathcal{S} \\ D(P_1||P_0) & i \in \mathcal{S} \end{cases}$$



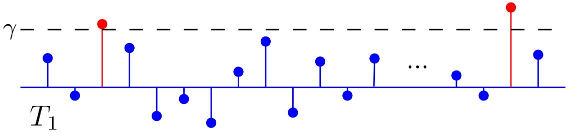
Limitations of non-sequential testing

$$T_i^{(m)} := \frac{1}{m}\sum_{j=1}^{m} \log \frac{p_1(Y_{i,j})}{p_0(Y_{i,j})} \quad \xrightarrow{a.s.} \quad \begin{cases} -D(P_0||P_1) & i \notin \mathcal{S} \\ D(P_1||P_0) & i \in \mathcal{S} \end{cases}$$



Limitations of non-sequential testing

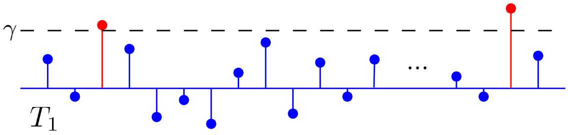
$$T_i^{(m)} := \frac{1}{m}\sum_{j=1}^{m} \log \frac{p_1(Y_{i,j})}{p_0(Y_{i,j})} \quad \xrightarrow{a.s.} \quad \begin{cases} -D(P_0||P_1) & i \notin \mathcal{S} \\ D(P_1||P_0) & i \in \mathcal{S} \end{cases}$$



$$\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) \quad \leq \quad (n-s)\,\mathbb{P}_0\left(T^{(m)} > \gamma\right) + s\,\mathbb{P}_1\left(T^{(m)} \leq \gamma\right)$$

Limitations of non-sequential testing

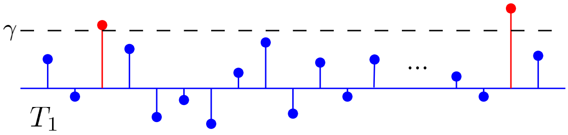
$$T_i^{(m)} := \frac{1}{m} \sum_{j=1}^{m} \log \frac{p_1(Y_{i,j})}{p_0(Y_{i,j})} \quad \xrightarrow{a.s.} \quad \begin{cases} -D(P_0||P_1) & i \notin \mathcal{S} \\ D(P_1||P_0) & i \in \mathcal{S} \end{cases}$$



$$\begin{aligned} \mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) \quad & \leq \quad (n-s)\,\mathbb{P}_0\left(T^{(m)} > \gamma\right) + s\,\mathbb{P}_1\left(T^{(m)} \leq \gamma\right) \\ & \lesssim \quad n\,e^{-mD(P_1||P_0)} \end{aligned}$$

Limitations of non-sequential testing

$$T_i^{(m)} := \frac{1}{m}\sum_{j=1}^{m} \log \frac{p_1(Y_{i,j})}{p_0(Y_{i,j})} \quad \xrightarrow{a.s.} \quad \begin{cases} -D(P_0||P_1) & i \notin \mathcal{S} \\ D(P_1||P_0) & i \in \mathcal{S} \end{cases}$$

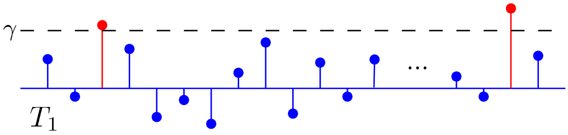


$$\begin{aligned} \mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) &\leq (n-s)\,\mathbb{P}_0\left(T^{(m)} > \gamma\right) + s\,\mathbb{P}_1\left(T^{(m)} \leq \gamma\right) \\ &\lesssim n\, e^{-mD(P_1||P_0)} \end{aligned}$$

- How fast does m have to grow with n for exact recovery of $\mathcal{S}$?

Limitations of non-sequential testing

$$T_i^{(m)} := \frac{1}{m}\sum_{j=1}^{m} \log \frac{p_1(Y_{i,j})}{p_0(Y_{i,j})} \xrightarrow{a.s.} \begin{cases} -D(P_0||P_1) & i \notin \mathcal{S} \\ D(P_1||P_0) & i \in \mathcal{S} \end{cases}$$



$$\begin{aligned} \mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) &\leq (n-s)\,\mathbb{P}_0\left(T^{(m)} > \gamma\right) + s\,\mathbb{P}_1\left(T^{(m)} \leq \gamma\right) \\ &\lesssim n\, e^{-mD(P_1||P_0)} \end{aligned}$$

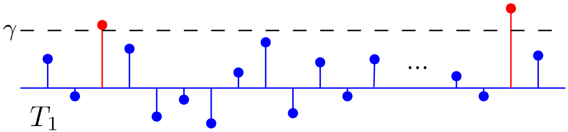
- How fast does m have to grow with n for exact recovery of $\mathcal{S}$?

$$m > \frac{\log n}{D(P_1||P_0)}$$

is a necessary condition for exact recovery of $\mathcal{S}$.

Limitations of non-sequential testing

$$T_i^{(m)} := \frac{1}{m}\sum_{j=1}^{m} \log \frac{p_1(Y_{i,j})}{p_0(Y_{i,j})} \quad \xrightarrow{a.s.} \quad \begin{cases} -D(P_0||P_1) & i \notin \mathcal{S} \\ D(P_1||P_0) & i \in \mathcal{S} \end{cases}$$



$$\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) \leq (n-s)\,\mathbb{P}_0\left(T^{(m)} > \gamma\right) + s\,\mathbb{P}_1\left(T^{(m)} \leq \gamma\right)$$

$$\lesssim n\, e^{-mD(P_1||P_0)}$$

- How fast does m have to grow with n for exact recovery of $\mathcal{S}$?

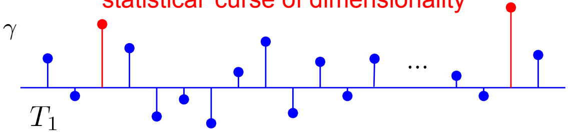
$$m > \frac{\log n}{D(P_1||P_0)} \iff \lim_{n\to\infty} \frac{m}{\log n} > \frac{1}{D(P_1||P_0)}$$

is a necessary condition for exact recovery of $\mathcal{S}$.

Limitations of non-sequential testing

$$T_i^{(m)} := \frac{1}{m}\sum_{j=1}^{m} \log \frac{p_1(Y_{i,j})}{p_0(Y_{i,j})} \xrightarrow{a.s.} \begin{cases} -D(P_0||P_1) & i \notin \mathcal{S} \\ D(P_1||P_0) & i \in \mathcal{S} \end{cases}$$

Non-sequential methods cannot overcome
statistical 'curse of dimensionality'



$$\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) \leq (n-s)\,\mathbb{P}_0\left(T^{(m)} > \gamma\right) + s\,\mathbb{P}_1\left(T^{(m)} \leq \gamma\right)$$
$$\lesssim n\, e^{-mD(P_1||P_0)}$$

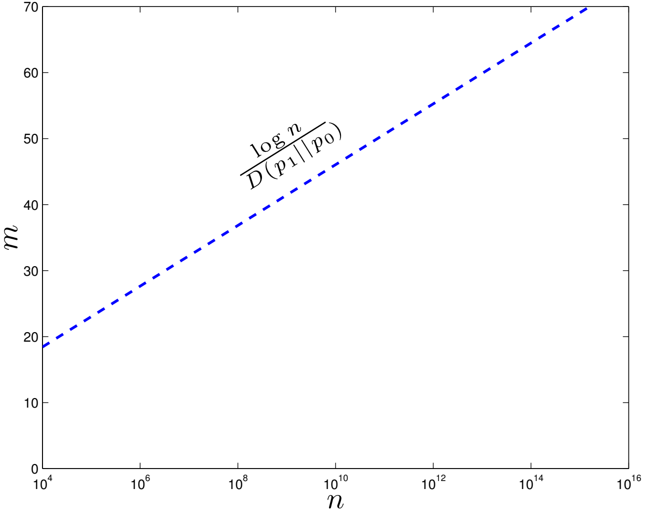
- How fast does m have to grow with n for exact recovery of $\mathcal{S}$?

$$m > \frac{\log n}{D(P_1||P_0)} \iff \lim_{n\to\infty} \frac{m}{\log n} > \frac{1}{D(P_1||P_0)}$$

is a necessary condition for exact recovery of $\mathcal{S}$.

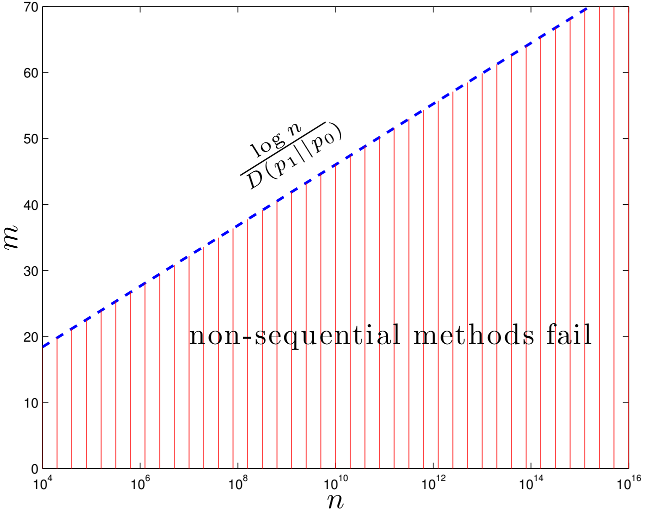
Results

$s = n^{\frac{1}{4}}$ $\qquad D(P_0||P_1) = \frac{1}{2}.$



Results

$s = n^{\frac{1}{4}}$ $\qquad D(P_0||P_1) = \frac{1}{2}.$



The Sequential Probability Ratio Test

How much better can sequential methods do? [Wald, Wolfowitz, Optimum Char. of the SPRT, 1948]

On a coordinate-wise basis, take additional measurement of index i if:

$$A < \prod_{j=1}^{j'} \frac{p_1(Y_j)}{p_0(Y_j)} < B$$

The Sequential Probability Ratio Test

How much better can sequential methods do? [Wald, Wolfowitz, Optimum Char. of the SPRT, 1948]

On a coordinate-wise basis, take additional measurement of index i if:

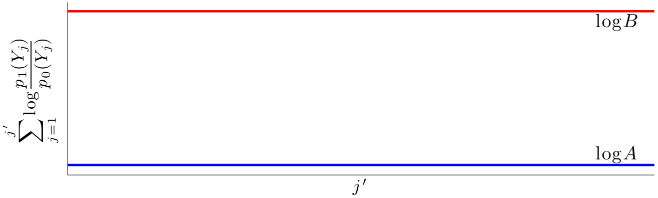
$$A < \prod_{j=1}^{j'} \frac{p_1(Y_j)}{p_0(Y_j)} < B$$

The Sequential Probability Ratio Test

How much better can sequential methods do? [Wald, Wolfowitz, Optimum Char. of the SPRT, 1948]

On a coordinate-wise basis, take additional measurement of index *i* if:

$$A < \prod_{j=1}^{j'} \frac{p_1(Y_j)}{p_0(Y_j)} < B$$

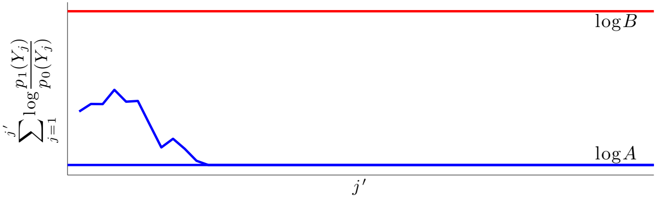


The Sequential Probability Ratio Test

How much better can sequential methods do? [Wald, Wolfowitz, Optimum Char. of the SPRT, 1948]

On a coordinate-wise basis, take additional measurement of index *i* if:

$$A < \prod_{j=1}^{j'} \frac{p_1(Y_j)}{p_0(Y_j)} < B$$

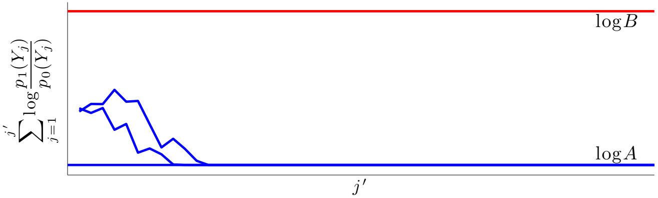


The Sequential Probability Ratio Test

How much better can sequential methods do? [Wald, Wolfowitz, Optimum Char. of the SPRT, 1948]

On a coordinate-wise basis, take additional measurement of index *i* if:

$$A < \prod_{j=1}^{j'} \frac{p_1(Y_j)}{p_0(Y_j)} < B$$

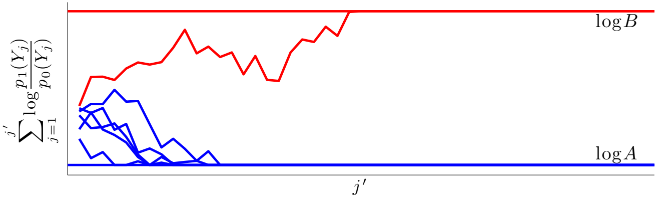


The Sequential Probability Ratio Test

How much better can sequential methods do? [Wald, Wolfowitz, Optimum Char. of the SPRT, 1948]

On a coordinate-wise basis, take additional measurement of index *i* if:

$$A < \prod_{j=1}^{j'} \frac{p_1(Y_j)}{p_0(Y_j)} < B$$

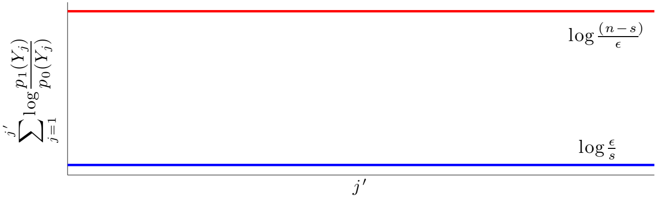


The Sequential Probability Ratio Test

How much better can sequential methods do? [Wald, Wolfowitz, Optimum Char. of the SPRT, 1948]

On a coordinate-wise basis, take additional measurement of index i if:

$$A < \prod_{j=1}^{j'} \frac{p_1(Y_j)}{p_0(Y_j)} < B$$

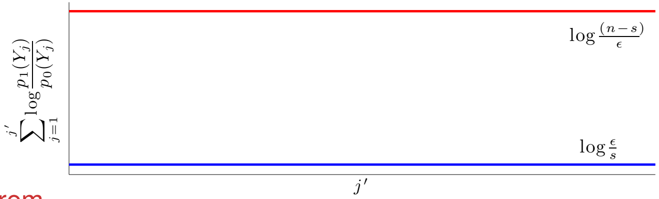


The Sequential Probability Ratio Test

How much better can sequential methods do? [Wald, Wolfowitz, Optimum Char. of the SPRT, 1948]

On a coordinate-wise basis, take additional measurement of index *i* if:

$$A < \prod_{j=1}^{j'} \frac{p_1(Y_j)}{p_0(Y_j)} < B$$



Theorem

Set $A = \frac{\epsilon}{2s}$, $B = \frac{2(n-s)}{\epsilon}$. The SPRT recovers $\mathcal{S}$ with probability

$$\mathbb{P}(\hat{\mathcal{S}} = \mathcal{S}) \geq 1 - \epsilon$$

and requires fewer than

$$m \leq (1 + \epsilon_0) \frac{\log s + \log \epsilon^{-1}}{D(P_0 || P_1)} \quad (1)$$

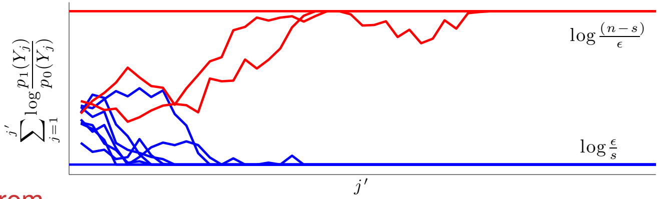
samples per dimension in expectation for any $\epsilon_0 > 0$, n sufficiently large.

The Sequential Probability Ratio Test

How much better can sequential methods do? [Wald, Wolfowitz, Optimum Char. of the SPRT, 1948]

On a coordinate-wise basis, take additional measurement of index i if:

$$A < \prod_{j=1}^{j'} \frac{p_1(Y_j)}{p_0(Y_j)} < B$$



Theorem

Set $A = \frac{\epsilon}{2s}$, $B = \frac{2(n-s)}{\epsilon}$. The SPRT recovers $\mathcal{S}$ with probability

$$\mathbb{P}(\hat{\mathcal{S}} = \mathcal{S}) \geq 1 - \epsilon$$

and requires fewer than

$$m \leq (1 + \epsilon_0) \frac{\log s + \log \epsilon^{-1}}{D(P_0 || P_1)} \quad (1)$$

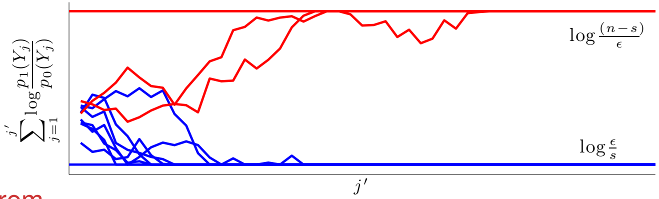
samples per dimension in expectation for any $\epsilon_0 > 0$, n sufficiently large.

The Sequential Probability Ratio Test

How much better can sequential methods do? [Wald, Wolfowitz, Optimum Char. of the SPRT, 1948]

On a coordinate-wise basis, take additional measurement of index i if:

$$A < \prod_{j=1}^{j'} \frac{p_1(Y_j)}{p_0(Y_j)} < B$$



Theorem

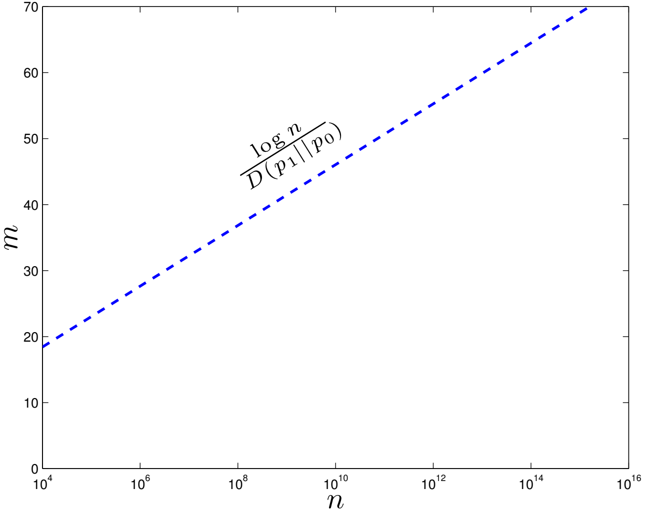
There exist thresholds A and B such that the SPRT recovers $\mathcal{S}$ exactly if

$$m > \frac{\log s}{D(P_0||P_1)} + \frac{\log \epsilon_n^{-1}}{D(P_0||P_1)}$$

for any $\epsilon_n \to 0$.

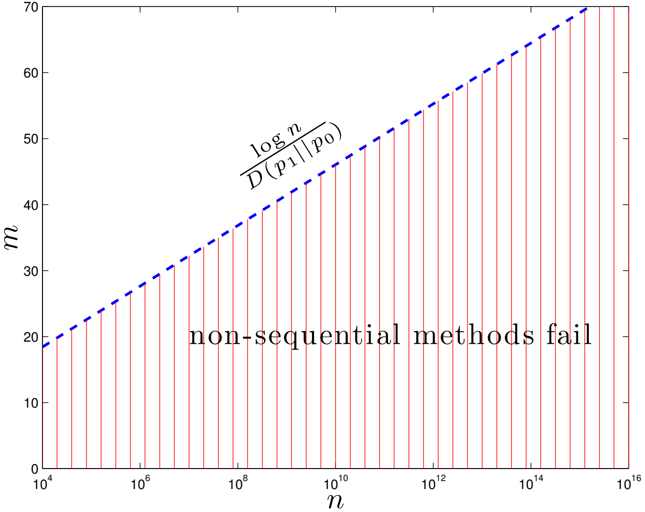
Results

$s = n^{\frac{1}{4}}$ $\qquad D(P_0||P_1) = \frac{1}{2}.$



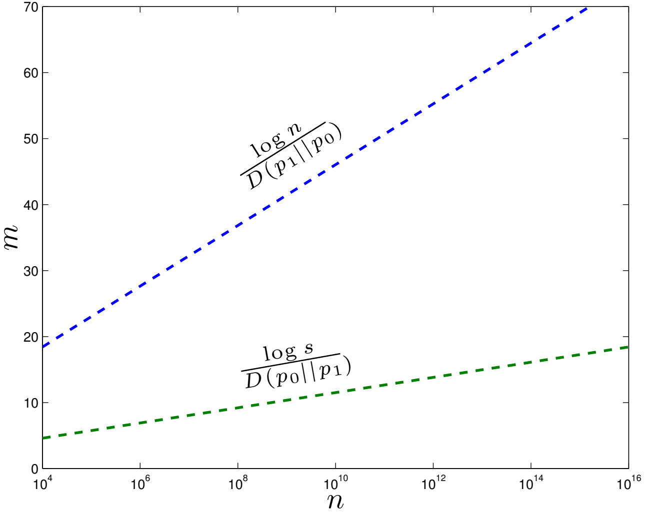
Results

$s = n^{\frac{1}{4}}$ $\qquad$ $D(P_0||P_1) = \frac{1}{2}$.



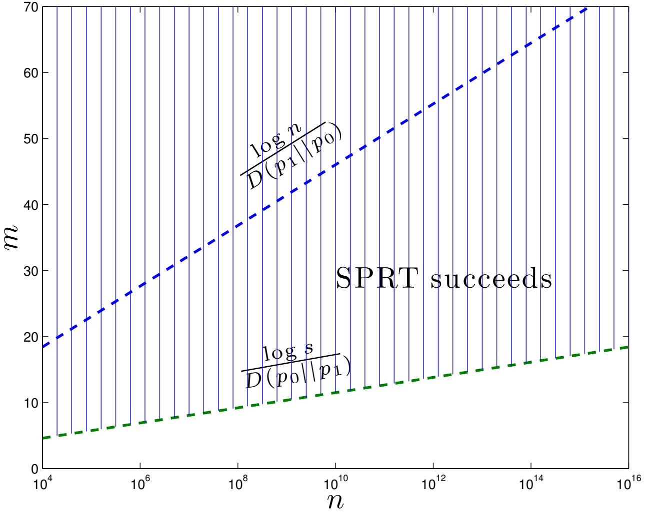
Results

$s = n^{\frac{1}{4}}$ $\qquad D(P_0||P_1) = \frac{1}{2}.$



Results

$s = n^{\frac{1}{4}}$ $\qquad D(P_0||P_1) = \frac{1}{2}.$



Sketch of Lower Bound

Sketch of coordinate-wise lower bound

Sketch of Lower Bound

Sketch of coordinate-wise lower bound

$$m \approx \mathbb{E}_0[J]$$

- *1)* for $s \ll n$

Sketch of Lower Bound

Sketch of coordinate-wise lower bound

$$m \approx \mathbb{E}_0[J] = \frac{\mathbb{E}_0\left[\sum_{j=1}^{J} \log \frac{p_0(y_j)}{p_1(y_j)}\right]}{D(P_0||P_1)}$$

- *1)* for $s \ll n$
- *2)* Wald's identity

Sketch of Lower Bound

Sketch of coordinate-wise lower bound

$$m \approx \mathbb{E}_0[J] = \frac{\mathbb{E}_0\left[\sum_{j=1}^{J} \log \frac{p_0(y_j)}{p_1(y_j)}\right]}{D(P_0||P_1)} \geq \frac{\alpha \log \frac{\alpha}{1-\beta} + (1-\alpha)\log \frac{1-\alpha}{\beta}}{D(P_0||P_1)}$$

- *1)* for $s \ll n$
- *2)* Wald's identity
- *3)* log-sum inequality for any test with false negative β, false positive α

Sketch of Lower Bound

Sketch of coordinate-wise lower bound

$$m \approx \mathbb{E}_0[J] = \frac{\mathbb{E}_0\left[\sum_{j=1}^{J} \log \frac{p_0(y_j)}{p_1(y_j)}\right]}{D(P_0||P_1)} \geq \frac{\alpha \log \frac{\alpha}{1-\beta} + (1-\alpha) \log \frac{1-\alpha}{\beta}}{D(P_0||P_1)}$$

- *1)* for $s \ll n$
- *2)* Wald's identity
- *3)* log-sum inequality for any test with false negative β, false positive α
- *4)* assume $\alpha \leq 1/(n-s)$ and $\beta \leq 1/s$

Sketch of Lower Bound

Sketch of coordinate-wise lower bound

$$m \approx \mathbb{E}_0[J] = \frac{\mathbb{E}_0\left[\sum_{j=1}^{J} \log \frac{p_0(y_j)}{p_1(y_j)}\right]}{D(P_0||P_1)} \geq \frac{\alpha \log \frac{\alpha}{1-\beta} + (1-\alpha) \log \frac{1-\alpha}{\beta}}{D(P_0||P_1)}$$

- *1)* for $s \ll n$
- *2)* Wald's identity
- *3)* log-sum inequality for any test with false negative β, false positive α
- *4)* assume $\alpha \leq 1/(n-s)$ and $\beta \leq 1/s$
- *5)* imposing condition

$$m < \frac{\log s}{D(P_0||P_1)}$$

Sketch of Lower Bound

Sketch of coordinate-wise lower bound

$$m \approx \mathbb{E}_0[J] = \frac{\mathbb{E}_0\left[\sum_{j=1}^{J} \log \frac{p_0(y_j)}{p_1(y_j)}\right]}{D(P_0||P_1)} \geq \frac{\alpha \log \frac{\alpha}{1-\beta} + (1-\alpha)\log \frac{1-\alpha}{\beta}}{D(P_0||P_1)}$$

- *1)* for $s \ll n$
- *2)* Wald's identity
- *3)* log-sum inequality for any test with false negative β, false positive α
- *4)* assume $\alpha \leq 1/(n-s)$ and $\beta \leq 1/s$
- *5)* imposing condition

$$m < \frac{\log s}{D(P_0||P_1)}$$

and taking limits as $n \to \infty$ implies contradiction, and thus

$$\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) \to 1$$

Sketch of Lower Bound

Sketch of coordinate-wise lower bound

$$m \approx \mathbb{E}_0[J] = \frac{\mathbb{E}_0\left[\sum_{j=1}^{J} \log \frac{p_0(y_j)}{p_1(y_j)}\right]}{D(P_0||P_1)} \geq \frac{\alpha \log \frac{\alpha}{1-\beta} + (1-\alpha)\log\frac{1-\alpha}{\beta}}{D(P_0||P_1)}$$

- *1)* for $s \ll n$
- *2)* Wald's identity
- *3)* log-sum inequality for any test with false negative β, false positive α
- *4)* assume $\alpha \leq 1/(n-s)$ and $\beta \leq 1/s$
- *5)* imposing condition

$$m < \frac{\log s}{D(P_0||P_1)}$$

and taking limits as $n \to \infty$ implies contradiction, and thus

$$\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) \to 1$$

Theorem

Any coordinate-wise sequential method fails to recover $\mathcal{S}$ provided

$$m < \frac{\log s}{D(P_0||P_1)}$$

Sketch of Lower Bound

Sketch of coordinate-wise lower bound

$$m \approx \mathbb{E}_0[J] = \frac{\mathbb{E}_0\left[\sum_{j=1}^{J} \log \frac{p_0(y_j)}{p_1(y_j)}\right]}{D(P_0||P_1)} \geq \frac{\alpha \log \frac{\alpha}{1-\beta} + (1-\alpha)\log\frac{1-\alpha}{\beta}}{D(P_0||P_1)}$$

- *1)* for $s \ll n$
- *2)* Wald's identity
- *3)* log-sum inequality for any test with false negative β, false positive α
- *4)* assume $\alpha \leq 1/(n-s)$ and $\beta \leq 1/s$
- *5)* imposing condition

$$m < \frac{\log s}{D(P_0||P_1)}$$

and taking limits as $n \to \infty$ implies contradiction, and thus

$$\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) \to 1$$

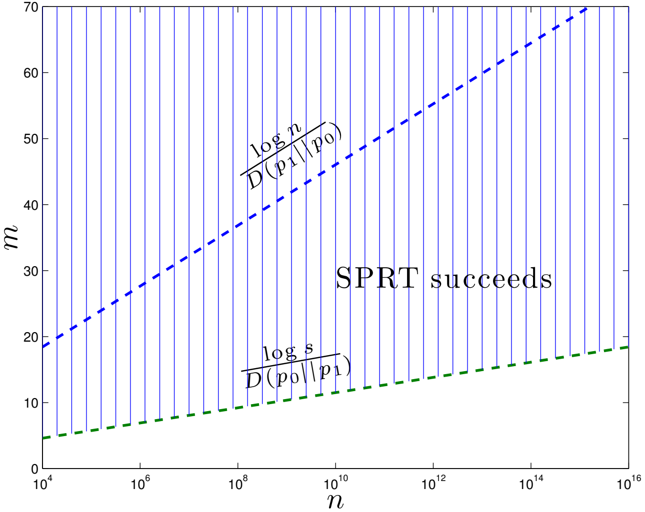
Theorem

Any ~~coordinate-wise~~ sequential method fails to recover $\mathcal{S}$ provided

$$m < \frac{\log s}{D(P_0||P_1)}$$

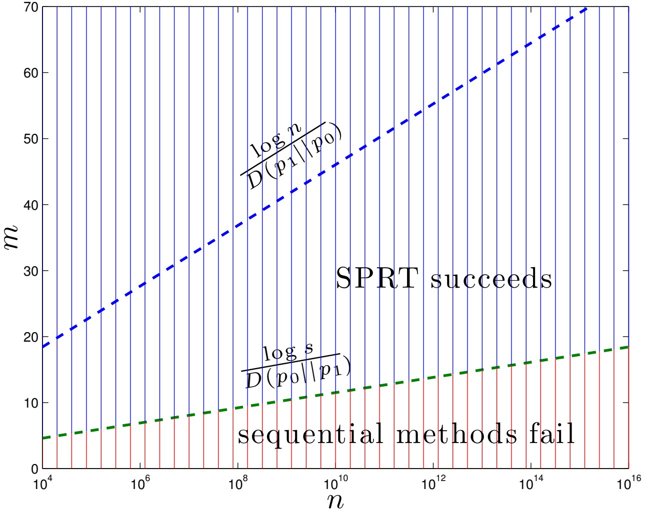
Results

$s = n^{\frac{1}{4}}$ $\quad D(P_0||P_1) = \frac{1}{2}.$



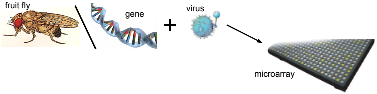
Results

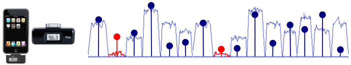
$s = n^{\frac{1}{4}}$ $\qquad D(P_0||P_1) = \frac{1}{2}.$



SPRT Implementation issues

Practical concerns:

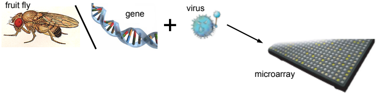


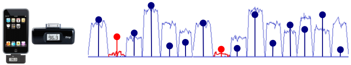


SPRT Implementation issues

Practical concerns:

- level of sparsity is often unknown

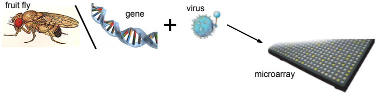


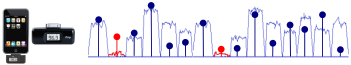


SPRT Implementation issues

Practical concerns:

- level of sparsity is often unknown
- full knowledge of P_0 and P_1 is restrictive

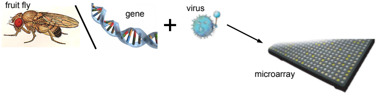


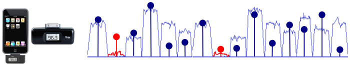


SPRT Implementation issues

Practical concerns:

- level of sparsity is often unknown
- full knowledge of P_0 and P_1 is restrictive
    - often some parameter of P_1 unknown

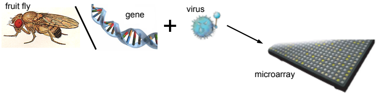


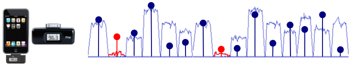


SPRT Implementation issues

Practical concerns:

- level of sparsity is often unknown
- full knowledge of P_0 and P_1 is restrictive
  - often some parameter of P_1 unknown

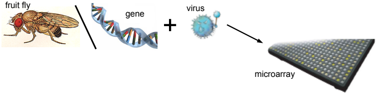




- SPRT not optimal for composite tests: consider $P_0 \sim \mathcal{N}(0, 1)$ and $P_1 \sim \mathcal{N}(\mu, 1)$, $\mu > 0$ unknown

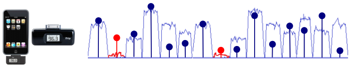
$$A < \prod_{j=1}^{j'} \frac{p_1(Y_j)}{p_0(Y_j)} < B$$

SPRT Implementation issues



Practical concerns:

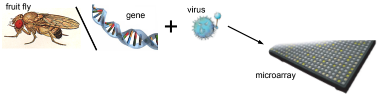
- level of sparsity is often unknown
- full knowledge of P_0 and P_1 is restrictive
  - often some parameter of P_1 unknown



- SPRT not optimal for composite tests: consider $P_0 \sim \mathcal{N}(0,1)$ and $P_1 \sim \mathcal{N}(\mu,1)$, $\mu > 0$ unknown

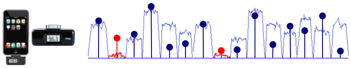
$$A < \prod_{j=1}^{j'} \frac{p_1(Y_j)}{p_0(Y_j)} < B \qquad \Longrightarrow \qquad \frac{\log A}{\mu} + \frac{j'\mu}{2} < \sum_{j=1}^{j'} Y_j < \frac{\log B}{\mu} + \frac{j'\mu}{2}$$

SPRT Implementation issues



Practical concerns:

- level of sparsity is often unknown
- full knowledge of P_0 and P_1 is restrictive
  - often some parameter of P_1 unknown

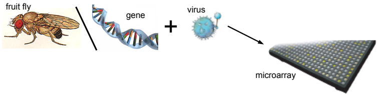


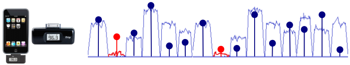
- SPRT not optimal for composite tests: consider $P_0 \sim \mathcal{N}(0, 1)$ and $P_1 \sim \mathcal{N}(\mu, 1)$, $\mu > 0$ unknown

$$A < \prod_{j=1}^{j'} \frac{p_1(Y_j)}{p_0(Y_j)} < B \qquad \Longrightarrow \qquad \frac{\log A}{\mu} + \frac{j'\mu}{2} < \sum_{j=1}^{j'} Y_j < \frac{\log B}{\mu} + \frac{j'\mu}{2}$$

- thresholds depend on μ, number of samples

SPRT Implementation issues





Practical concerns:

- level of sparsity is often unknown
- full knowledge of P_0 and P_1 is restrictive
  - often some parameter of P_1 unknown

- SPRT not optimal for composite tests: consider $P_0 \sim \mathcal{N}(0,1)$ and $P_1 \sim \mathcal{N}(\mu,1)$, $\mu > 0$ unknown

$$A < \prod_{j=1}^{j'} \frac{p_1(Y_j)}{p_0(Y_j)} < B \qquad \Longrightarrow \qquad \frac{\log A}{\mu} + \frac{j'\mu}{2} < \sum_{j=1}^{j'} Y_j < \frac{\log B}{\mu} + \frac{j'\mu}{2}$$

- thresholds depend on μ, number of samples
- What can we do without knowledge of P_1 or s?

Sequential Thresholding

Input:

- $K \approx \log n$ measurement passes
- threshold $\gamma : \mathbb{P}_0 \left(T^{(m/2)} \leq \gamma \right) = \frac{1}{2}$

Sequential Thresholding

Input:

- $K \approx \log n$ measurement passes (depends only on n)
- threshold $\gamma : \mathbb{P}_0\left(T^{(m/2)} \leq \gamma\right) = \frac{1}{2}$

Sequential Thresholding

Input:

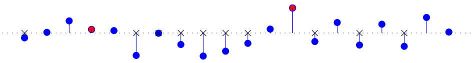
- $K \approx \log n$ measurement passes
- threshold $\gamma : \mathbb{P}_0 \left(T^{(m/2)} \leq \gamma \right) = \frac{1}{2}$ (depends only on P_0)

Sequential Thresholding

Input:

- $K \approx \log n$ measurement passes
- threshold $\gamma : \mathbb{P}_0 \left(T^{(m/2)} \leq \gamma \right) = \frac{1}{2}$

$k = 1$



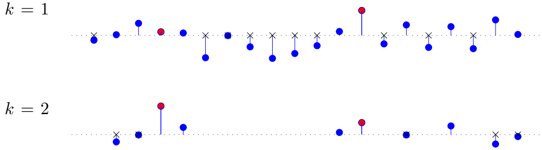
1) sample each index $\frac{m}{2}$ times

$$\mathbb{E}\left[\sum_{i=1}^{n} J_i\right] \approx \frac{mn}{2}$$

Sequential Thresholding

Input:

- $K \approx \log n$ measurement passes
- threshold $\gamma : \mathbb{P}_0 \left(T^{(m/2)} \leq \gamma \right) = \frac{1}{2}$



1) sample each index $\frac{m}{2}$ times

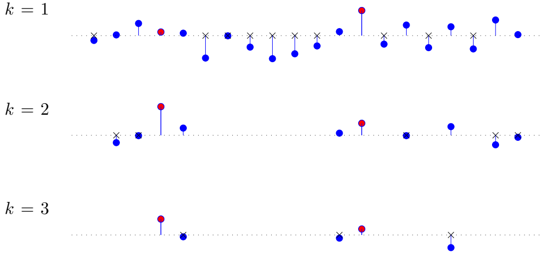
2) re-measure only indices above threshold

$$\mathbb{E}\left[\sum_{i=1}^{n} J_i\right] \approx \frac{mn}{2} + \frac{mn}{4}$$

Sequential Thresholding

Input:

- $K \approx \log n$ measurement passes
- threshold $\gamma : \mathbb{P}_0 \left(T^{(m/2)} \le \gamma \right) = \frac{1}{2}$



1) sample each index $\frac{m}{2}$ times

2) re-measure only indices above threshold

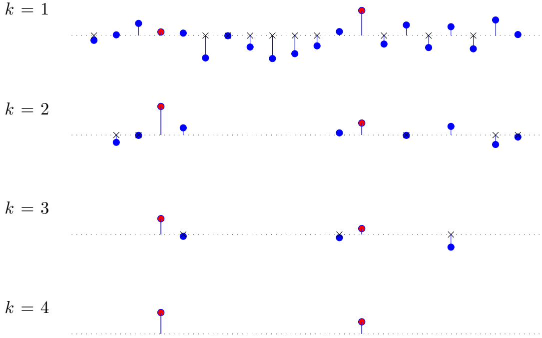
3) repeat

$$\mathbb{E}\left[\sum_{i=1}^{n} J_i\right] \approx \frac{mn}{2} + \frac{mn}{4} + \frac{mn}{8}$$

Sequential Thresholding

Input:

- $K \approx \log n$ measurement passes
- threshold $\gamma : \mathbb{P}_0 \left(T^{(m/2)} \leq \gamma \right) = \frac{1}{2}$



1) sample each index $\frac{m}{2}$ times

2) re-measure only indices above threshold

3) repeat

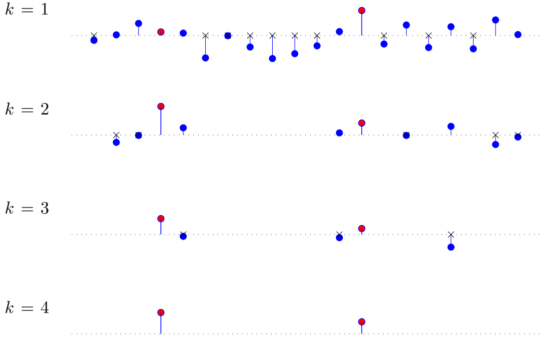
4) ...

$$\mathbb{E}\left[\sum_{i=1}^{n} J_i\right] \approx \frac{mn}{2} + \frac{mn}{4} + \frac{mn}{8} + \ldots$$

Sequential Thresholding

Input:

- $K \approx \log n$ measurement passes
- threshold $\gamma : \mathbb{P}_0\left(T^{(m/2)} \leq \gamma\right) = \frac{1}{2}$



1) sample each index $\frac{m}{2}$ times

2) re-measure only indices above threshold

3) repeat

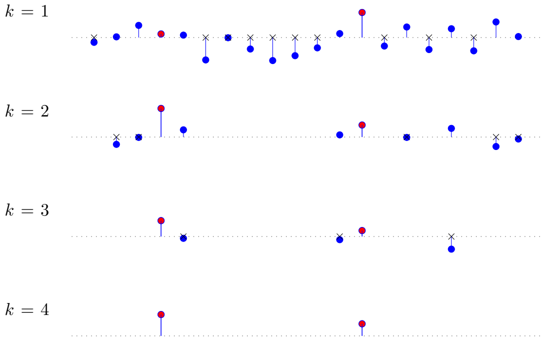
4) ...

$$\mathbb{E}\left[\sum_{i=1}^{n} J_i\right] \approx \frac{mn}{2} + \frac{mn}{4} + \frac{mn}{8} + \ldots \leq mn$$

Sequential Thresholding

Input:

- $K \approx \log n$ measurement passes
- threshold $\gamma : \mathbb{P}_0 \left(T^{(m/2)} \leq \gamma \right) = \frac{1}{2}$



1) sample each index $\frac{m}{2}$ times

2) re-measure only indices above threshold

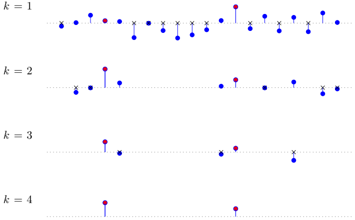
3) repeat

4) ...

$$\mathbb{E}\left[\sum_{i=1}^{n} J_i\right] \approx \frac{mn}{2} + \frac{mn}{4} + \frac{mn}{8} + \ldots \leq mn$$

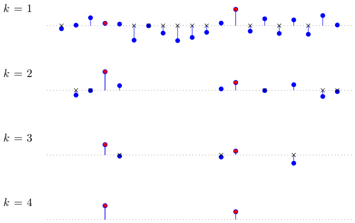
After $K \approx \log n$ passes, return remaining indices!

Controlling Family Wise Error Rates



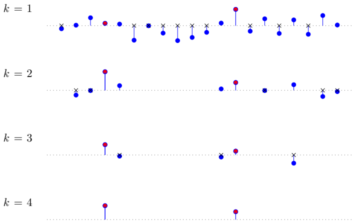
Controlling Family Wise Error Rates

$$\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) \quad = \quad \mathbb{P}\left(\left\{\bigcup_{i \notin \mathcal{S}} \bigcap_{k=1}^{K} T_{i,k} \geq \gamma\right\} \cup \left\{\bigcup_{i \in \mathcal{S}} \bigcup_{k=1}^{K} T_{i,k} \leq \gamma\right\}\right)$$



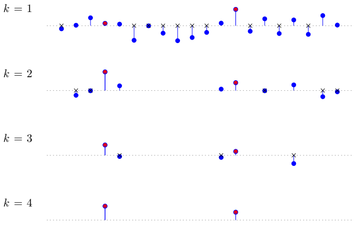
Controlling Family Wise Error Rates

$$\begin{aligned}\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) &= \mathbb{P}\left(\left\{\bigcup_{i \notin \mathcal{S}} \bigcap_{k=1}^{K} T_{i,k} \geq \gamma\right\} \cup \left\{\bigcup_{i \in \mathcal{S}} \bigcup_{k=1}^{K} T_{i,k} \leq \gamma\right\}\right) \\ &\leq (n-s)\frac{1}{2^K} + sK\mathbb{P}_1\left(T_i^{(m/2)} < \gamma\right)\end{aligned}$$



Controlling Family Wise Error Rates

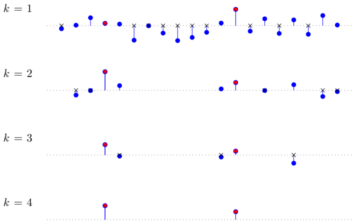
$$\begin{aligned}
\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) &= \mathbb{P}\left(\left\{\bigcup_{i \notin \mathcal{S}} \bigcap_{k=1}^{K} T_{i,k} \geq \gamma\right\} \cup \left\{\bigcup_{i \in \mathcal{S}} \bigcup_{k=1}^{K} T_{i,k} \leq \gamma\right\}\right) \\
&\leq (n-s)\frac{1}{2^K} + sK\mathbb{P}_1\left(T_i^{(m/2)} < \gamma\right) \\
&\lesssim (n-s)\frac{1}{2^K} + sKe^{-\frac{m}{2}D(P_0||P_1)}
\end{aligned}$$



Controlling Family Wise Error Rates

$$
\begin{aligned}
\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) &= \mathbb{P}\left(\left\{\bigcup_{i \notin \mathcal{S}} \bigcap_{k=1}^{K} T_{i,k} \geq \gamma\right\} \cup \left\{\bigcup_{i \in \mathcal{S}} \bigcup_{k=1}^{K} T_{i,k} \leq \gamma\right\}\right) \\
&\leq (n-s)\frac{1}{2^K} + sK\mathbb{P}_1\left(T_i^{(m/2)} < \gamma\right) \\
&\lesssim (n-s)\frac{1}{2^K} + sKe^{-\frac{m}{2}D(P_0||P_1)} \\
&\rightarrow 0
\end{aligned}
$$

provided ...

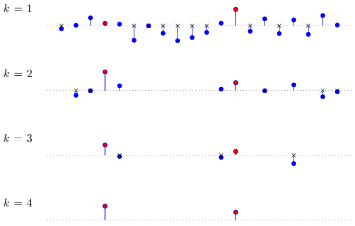


Controlling Family Wise Error Rates

$$
\begin{aligned}
\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) &= \mathbb{P}\left(\left\{\bigcup_{i \notin \mathcal{S}} \bigcap_{k=1}^{K} T_{i,k} \geq \gamma\right\} \cup \left\{\bigcup_{i \in \mathcal{S}} \bigcup_{k=1}^{K} T_{i,k} \leq \gamma\right\}\right) \\
&\leq (n-s)\frac{1}{2^K} + sK\mathbb{P}_1\left(T_i^{(m/2)} < \gamma\right) \\
&\lesssim (n-s)\frac{1}{2^K} + sKe^{-\frac{m}{2}D(P_0||P_1)} \\
&\to 0
\end{aligned}
$$

provided ...

1. $K = (1+\epsilon)\log n$

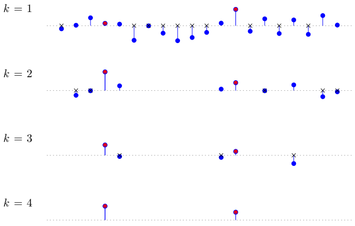


Controlling Family Wise Error Rates

$$
\begin{aligned}
\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) &= \mathbb{P}\left(\left\{\bigcup_{i \notin \mathcal{S}} \bigcap_{k=1}^{K} T_{i,k} \geq \gamma\right\} \cup \left\{\bigcup_{i \in \mathcal{S}} \bigcup_{k=1}^{K} T_{i,k} \leq \gamma\right\}\right) \\
&\leq (n-s)\frac{1}{2^K} + sK\mathbb{P}_1\left(T_i^{(m/2)} < \gamma\right) \\
&\lesssim (n-s)\frac{1}{2^K} + sKe^{-\frac{m}{2}D(P_0||P_1)} \\
&\to 0
\end{aligned}
$$

provided ...

1. $K = (1+\epsilon)\log n$
2. $m > \frac{2\log(sK)}{D(P_0||P_1)}$



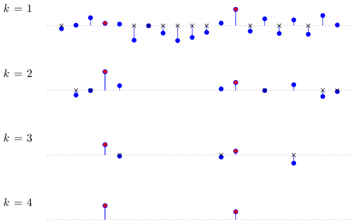
Controlling Family Wise Error Rates

$$
\begin{aligned}
\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) &= \mathbb{P}\left(\left\{\bigcup_{i\notin\mathcal{S}}\bigcap_{k=1}^{K} T_{i,k} \geq \gamma\right\} \cup \left\{\bigcup_{i\in\mathcal{S}}\bigcup_{k=1}^{K} T_{i,k} \leq \gamma\right\}\right) \\
&\leq (n-s)\frac{1}{2^K} + sK\mathbb{P}_1\left(T_i^{(m/2)} < \gamma\right) \\
&\lesssim (n-s)\frac{1}{2^K} + sKe^{-\frac{m}{2}D(P_0||P_1)} \\
&\to 0
\end{aligned}
$$

provided ...

1. $K = (1+\epsilon)\log n$
2. $m > \frac{2\log(sK)}{D(P_0||P_1)}$

Adjust definition of γ removes 2 . . .



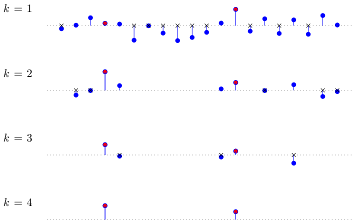
Controlling Family Wise Error Rates

$$\begin{aligned}\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) &= \mathbb{P}\left(\left\{\bigcup_{i \notin \mathcal{S}} \bigcap_{k=1}^{K} T_{i,k} \geq \gamma\right\} \cup \left\{\bigcup_{i \in \mathcal{S}} \bigcup_{k=1}^{K} T_{i,k} \leq \gamma\right\}\right)\\
&\leq (n-s)\frac{1}{2^K} + sK\mathbb{P}_1\left(T_i^{(m/2)} < \gamma\right)\\
&\lesssim (n-s)\frac{1}{2^K} + sKe^{-\frac{m}{2}D(P_0||P_1)}\\
&\to 0\end{aligned}$$

provided ...

1. $K = (1+\epsilon)\log n$
2. $m > \frac{2\log(sK)}{D(P_0||P_1)}$

Adjust definition of γ removes 2 . . .



Theorem

Sequential Thresholding succeeds in exactly recovery of $\mathcal{S}$ if

$$m > \frac{\log s}{D(P_0||P_1)} + \frac{\log\log n}{D(P_0||P_1)}$$

Controlling Family Wise Error Rates

$$
\begin{aligned}
\mathbb{P}(\hat{\mathcal{S}} \neq \mathcal{S}) &= \mathbb{P}\left(\left\{\bigcup_{i \notin \mathcal{S}} \bigcap_{k=1}^{K} T_{i,k} \geq \gamma\right\} \cup \left\{\bigcup_{i \in \mathcal{S}} \bigcup_{k=1}^{K} T_{i,k} \leq \gamma\right\}\right) \\
&\leq (n-s)\frac{1}{2^K} + sK\mathbb{P}_1\left(T_i^{(m/2)} < \gamma\right) \\
&\lesssim (n-s)\frac{1}{2^K} + sKe^{-\frac{m}{2}D(P_0||P_1)} \\
&\rightarrow 0
\end{aligned}
$$

provided ...

1. $K = (1+\epsilon)\log n$
2. $m > \frac{2\log(sK)}{D(P_0||P_1)}$

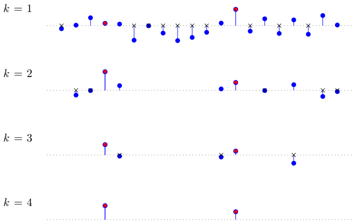
Adjust definition of γ removes 2 . . .

Theorem

Sequential Thresholding succeeds in exactly recovery of $\mathcal{S}$ if

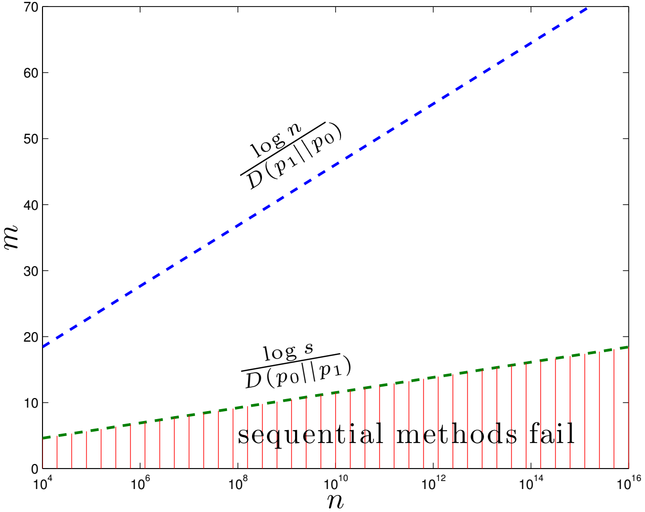
$$
m > \frac{\log s}{D(P_0||P_1)} + \frac{\log\log n}{D(P_0||P_1)}
$$

Note: For certain levels of sparsity, ST is asymptotically optimal!



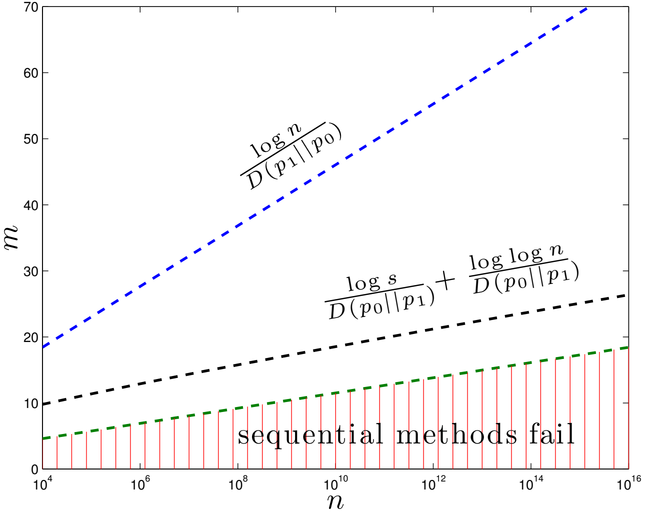
Results

$s = n^{\frac{1}{4}}$ $\qquad$ $D(P_0||P_1) = \frac{1}{2}$.



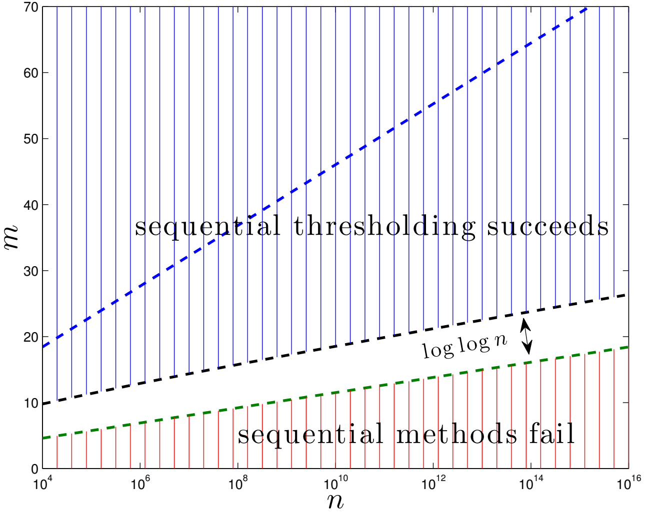
Results

$s = n^{\frac{1}{4}}$ $\qquad$ $D(P_0||P_1) = \frac{1}{2}.$



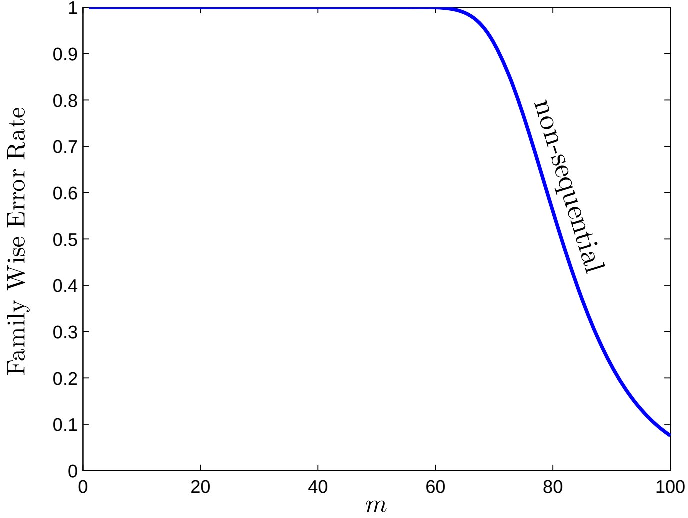
Results

$s = n^{\frac{1}{4}}$ $\qquad D(P_0||P_1) = \frac{1}{2}$.



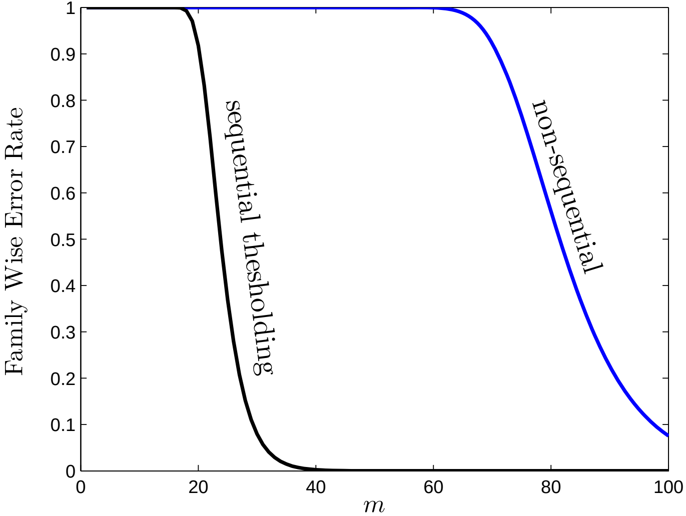
Results

$n = 10^9$ $\qquad s = n^{\frac{1}{4}}$ $\qquad D(P_0||P_1) = \frac{1}{2}.$



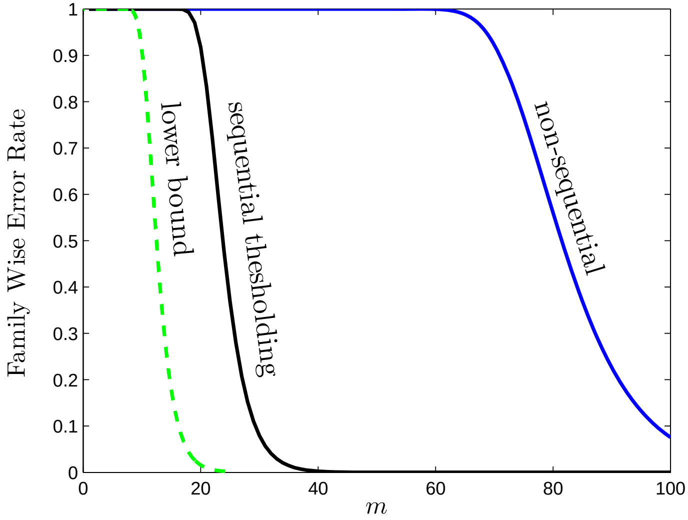
Results

$n = 10^9$ $s = n^{\frac{1}{4}}$ $D(P_0||P_1) = \frac{1}{2}$.



Results

$n = 10^9 \qquad s = n^{\frac{1}{4}} \qquad D(P_0||P_1) = \frac{1}{2}.$



Conclusion

Remaining questions: can procedures remove doubly logarithmic gap without full knowledge of distributions?

For further reading:

M. Malloy, R. Nowak
Sequential Analysis in High Dimensional Multiple Testing and Sparse Recovery.
ISIT 2011.

M. Malloy, R. Nowak
On the limits of Sequential Testing in High Dimensions.
Asilomar 2011.