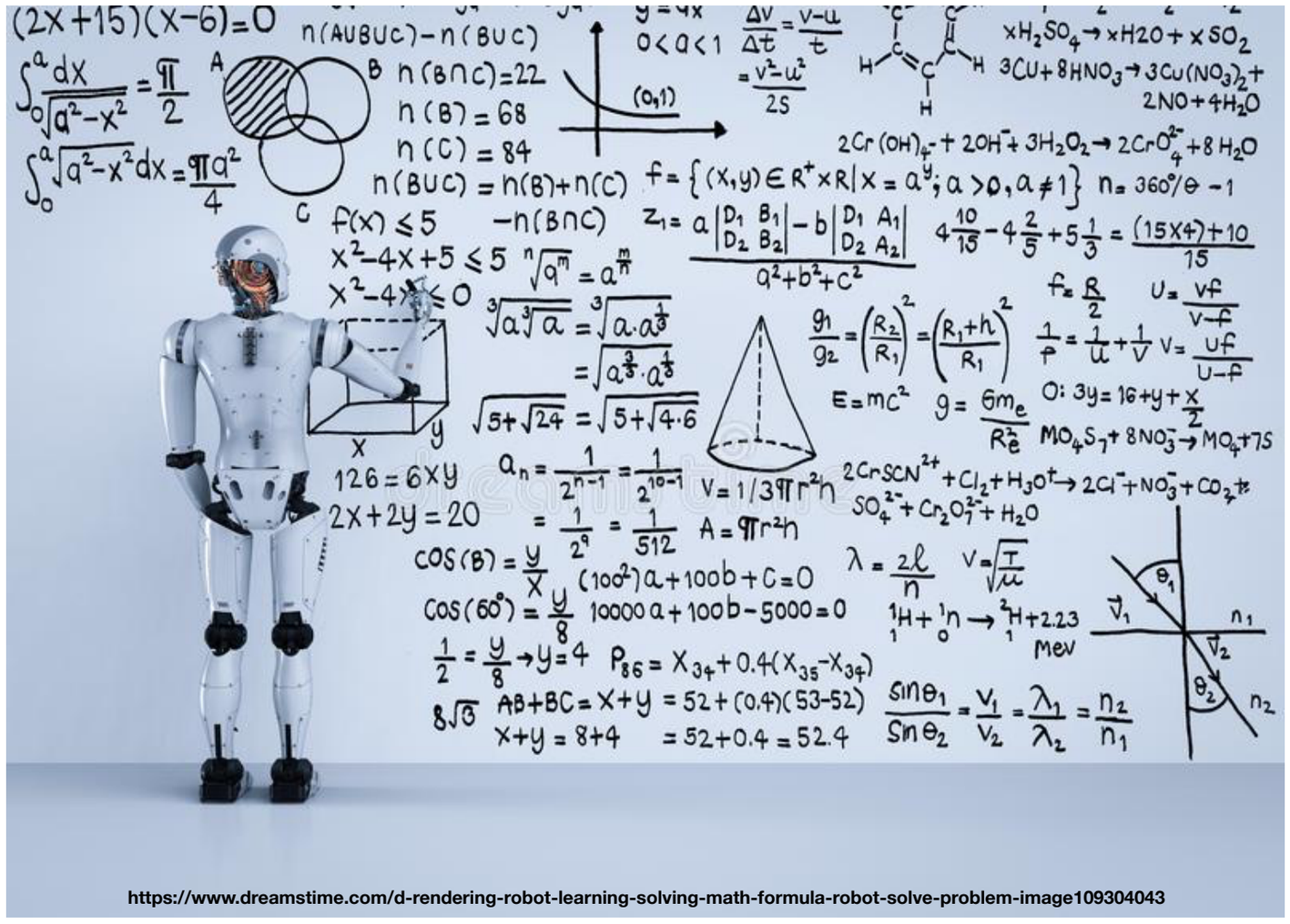
Introduction to Theory of Active Machine Learning



https://www.dreamstime.com/d-rendering-robot-learning-solving-math-formula-robot-solve-problem-image109304043

IFDS @ UW-Madison

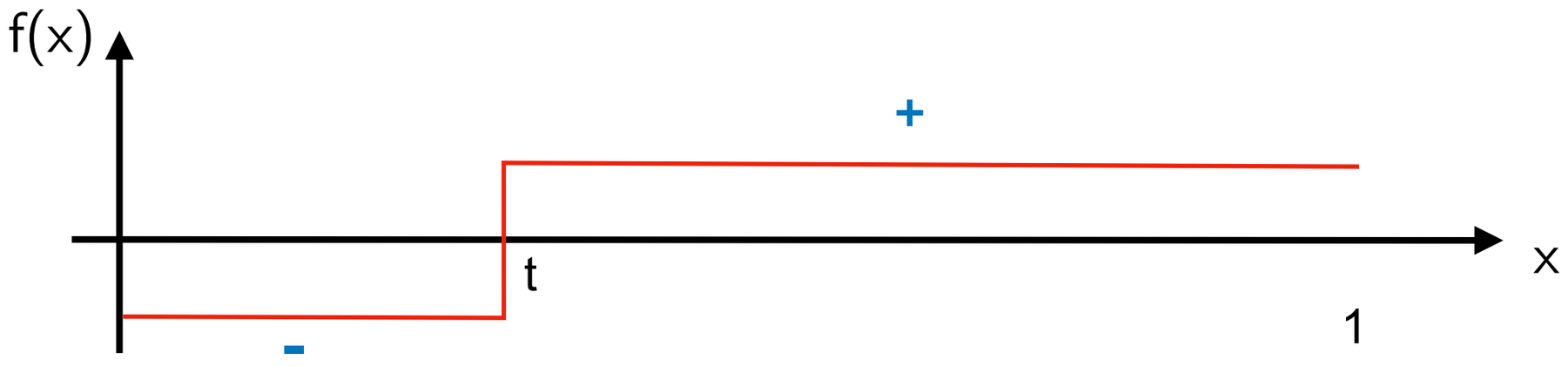
Summer School
July 25, 2018

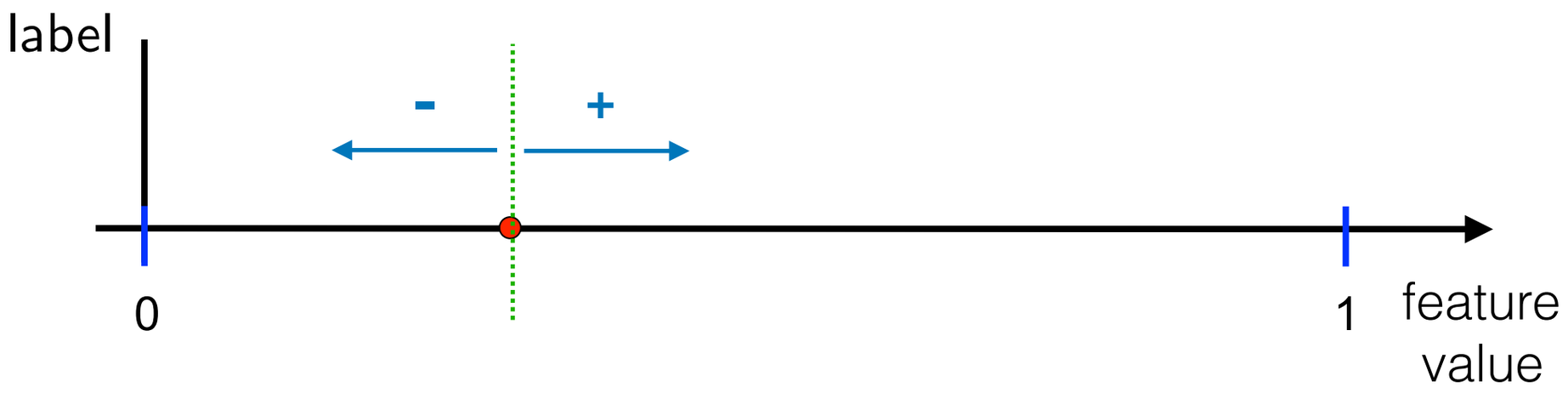
Robert Nowak
UW-Madison
rdnowak@wisc.edu

Outline

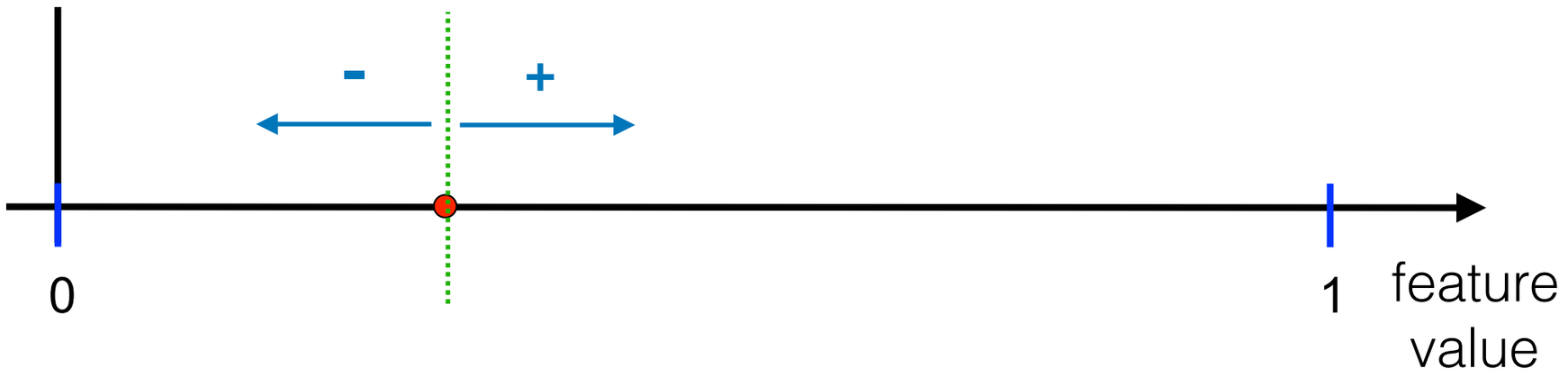
1. Binary Search
2. Coping with Noise: Confidence Intervals
3. Noisy Binary Search
4. Machine Learning: Empirical Risk Minimization
5. Active Machine Learning
6. Multi-Armed Bandits

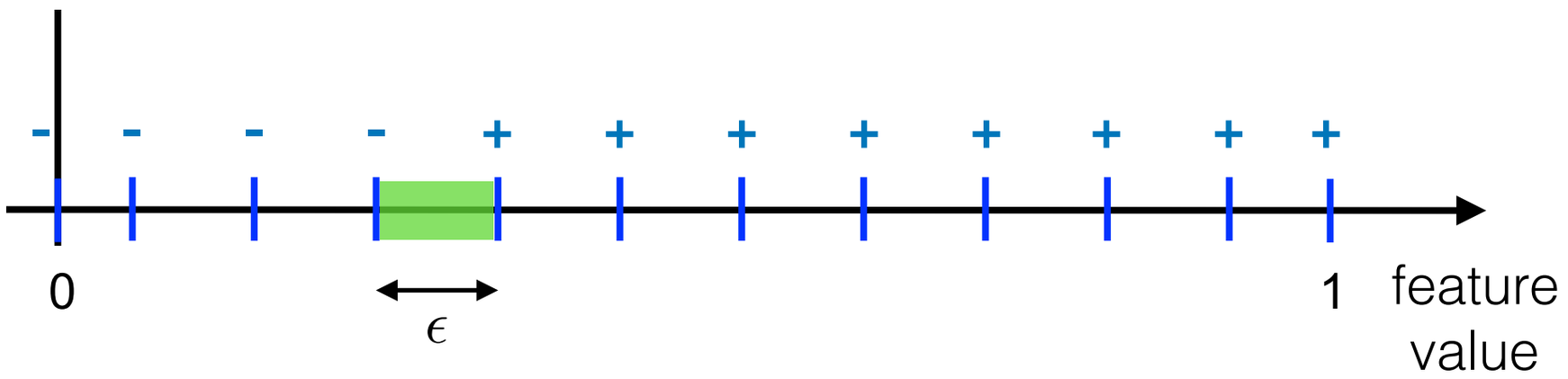
Learning a 1-D Threshold Function (classifier)





Learning a 1-d Linear Classifier





$n = 1/\epsilon$ evaluations

Binary Search

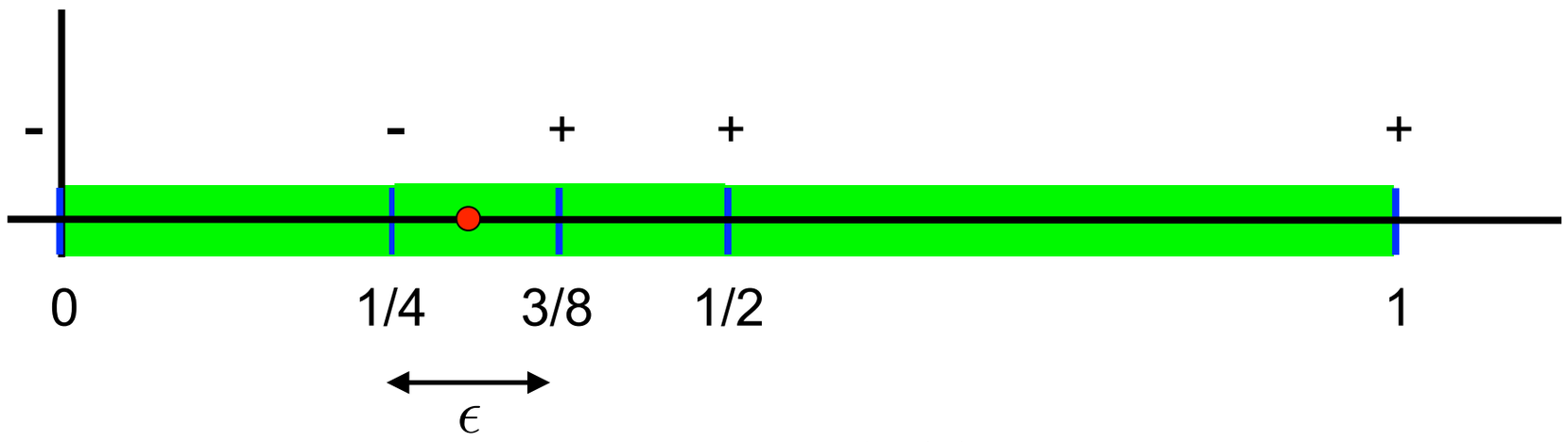
Binary Search

initialize: $\mathcal{H} = [0, 1]$

while (*stopping-criterion*) not met

1. **sample** at midpoint of $\mathcal{H}$
2. **label** midpoint
3. **reduce** $\mathcal{H}$ half of interval

output: $\mathcal{H}$, a small subset of $[0, 1]$



$n = \log(1/\epsilon)$ evaluations

Noisy Binary Search

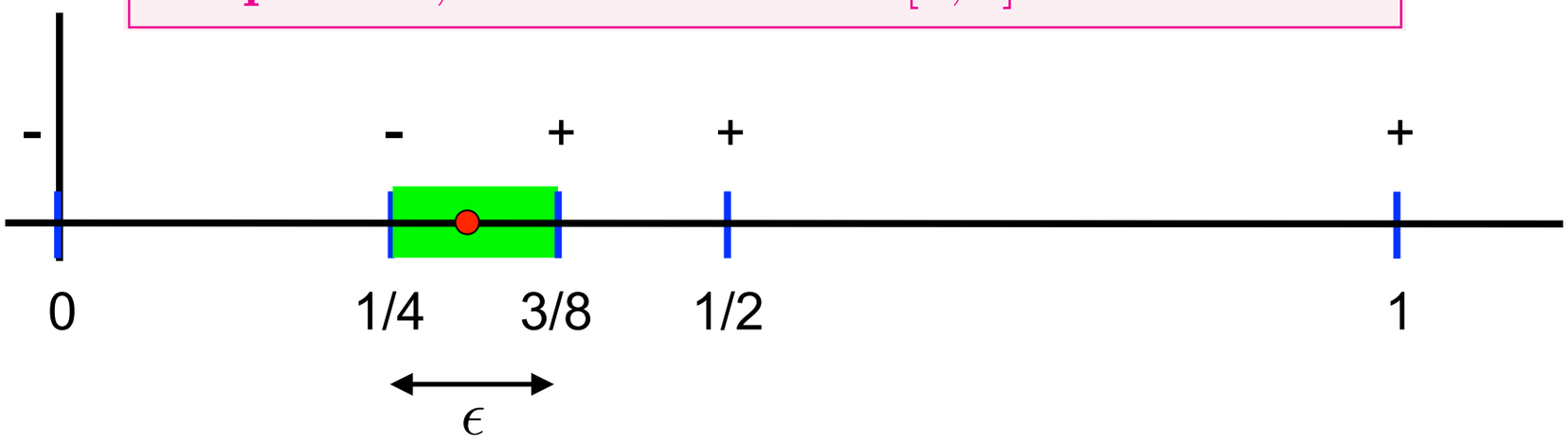
Noisy Binary Search

initialize: $\mathcal{H} = [0, 1]$

while (*stopping-criterion*) not met

1. **sample** m-times (iid) at midpoint of $\mathcal{H}$
2. **label** midpoint by majority vote
3. **reduce** $\mathcal{H}$ half of interval

output: $\mathcal{H}$, a small subset of $[0, 1]$

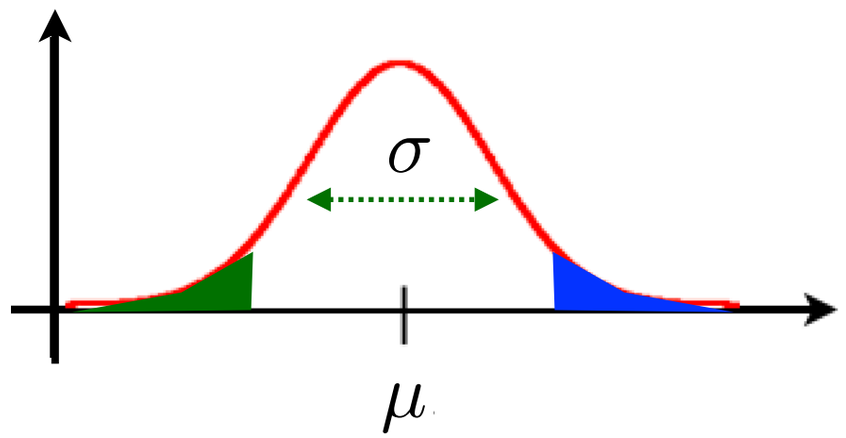


if $n = \log(1/\epsilon) \underbrace{\log\left(\frac{\log(1/\epsilon)}{\delta}\right)}_{m}$ samples, then correct with probability $\geq 1 - \delta$

Confidence Intervals

Gaussian Tails

$$y \sim \mathcal{N}(\mu, \sigma^2)$$



$$\mathbb{P}(y - \mu \geq t) \;\leq\; \frac{1}{2}\exp(-\frac{t^2}{2\sigma^2})$$

$$\mathbb{P}(\mu - y \geq t) \;\leq\; \frac{1}{2}\exp(-\frac{t^2}{2\sigma^2})$$

$$y_1, \ldots, y_m \overset{\text{iid}}{\sim} \mathcal{N}(\mu, 1) \qquad \widehat{\mu}_m \sim \mathcal{N}(\mu, 1/m)$$

$$\mathbb{P}(\widehat{\mu}_m - \mu \geq t) \;\leq\; \frac{1}{2}\exp(-\frac{mt^2}{2})$$

$$\mathbb{P}(\mu - \widehat{\mu}_m \geq t) \;\leq\; \frac{1}{2}\exp(-\frac{mt^2}{2})$$

Chernoff's Bound

The same sort of bounds hold for averages of independent random variables distributed with tails that decay at least as fast as Gaussians

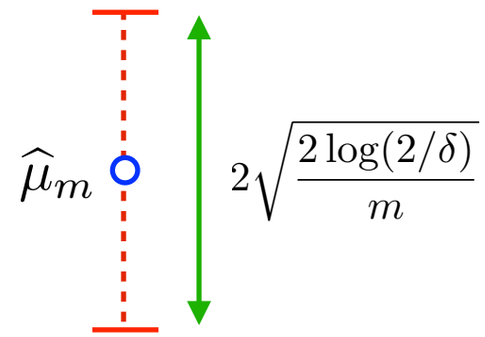
If $y_1, \ldots, y_m$ are iid bounded in $[-1, 1]$, then we have Chernoff's bound

$$\mathbb{P}(\widehat{\mu}_m - \mu \geq t) \quad \leq \quad \exp(-\frac{mt^2}{2})$$

$$\mathbb{P}(\mu - \widehat{\mu}_m \geq t) \quad \leq \quad \exp(-\frac{mt^2}{2})$$

set $\quad \delta = 2\exp(-\frac{mt^2}{2})$

$$\widehat{\mu}_m - \sqrt{\frac{2\log(2/\delta)}{m}} \;\leq\; \mu \;\leq\; \widehat{\mu}_m + \sqrt{\frac{2\log(2/\delta)}{m}}$$



true mean value is in this interval
with probability at least $1 - \delta$

Noisy Binary Search

Noisy Binary Search

Noisy Binary Search

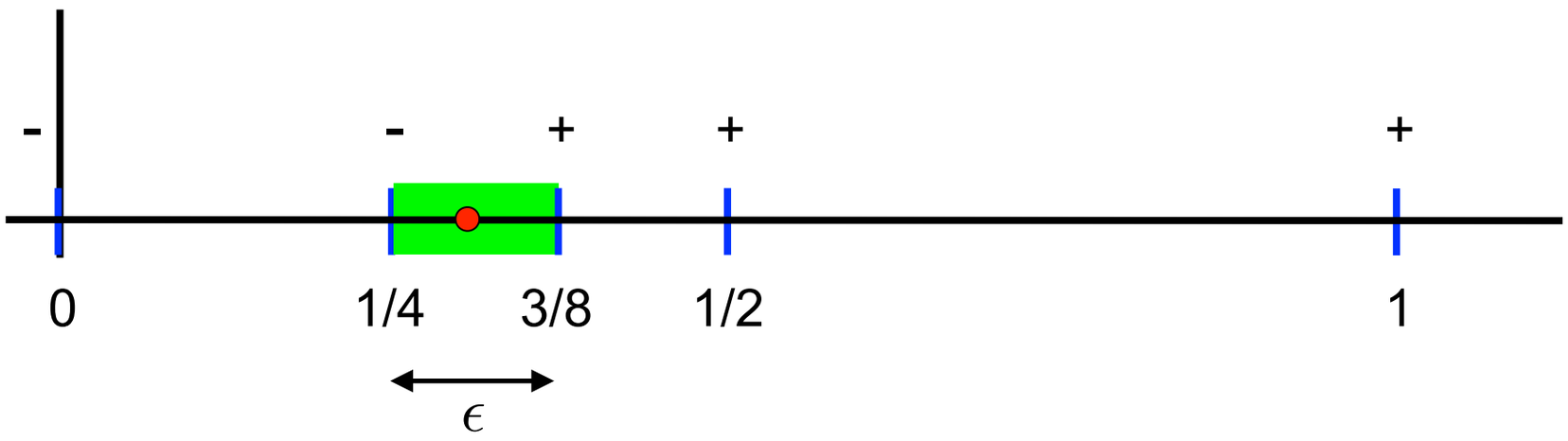
initialize: $\mathcal{H} = [0, 1]$

while (*stopping-criterion*) not met

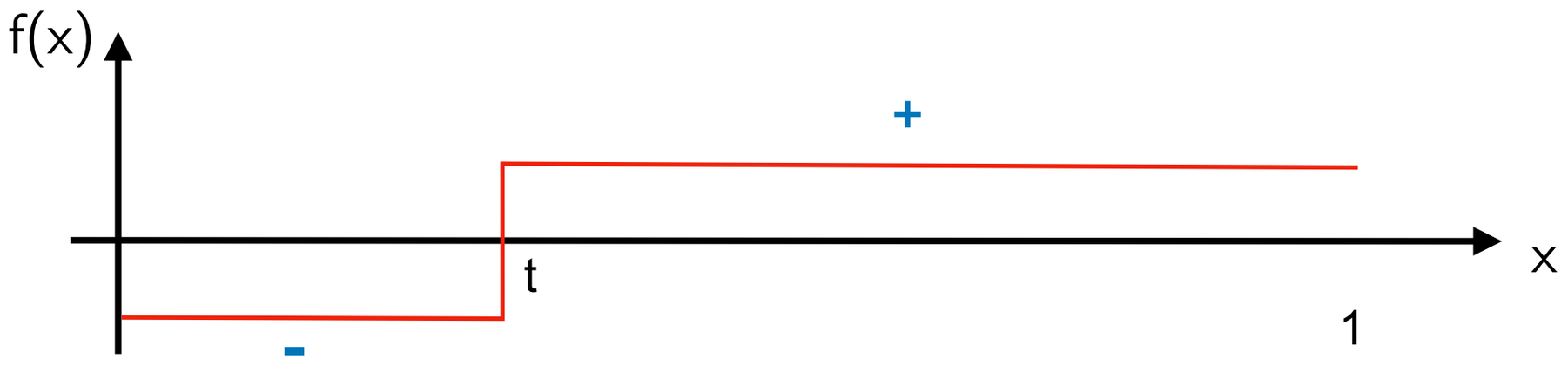
1. **sample** m-times (iid) at midpoint of $\mathcal{H}$
2. **label** midpoint by majority vote
3. **reduce** $\mathcal{H}$ half of interval

output: $\mathcal{H}$, a small subset of $[0, 1]$

majority vote =
sign of empirical mean



Coping with Noise



noise model $\quad y(x) \;=\; \begin{cases} f(x) & \text{w.p. } (1+\Delta)/2 \\ -f(x) & \text{w.p. } (1-\Delta)/2 \end{cases} \qquad \Delta > 0$

collect iid $y_1(x), \ldots, y_m(x)$ and let $\widehat{\mu}_m(x)$ be the average

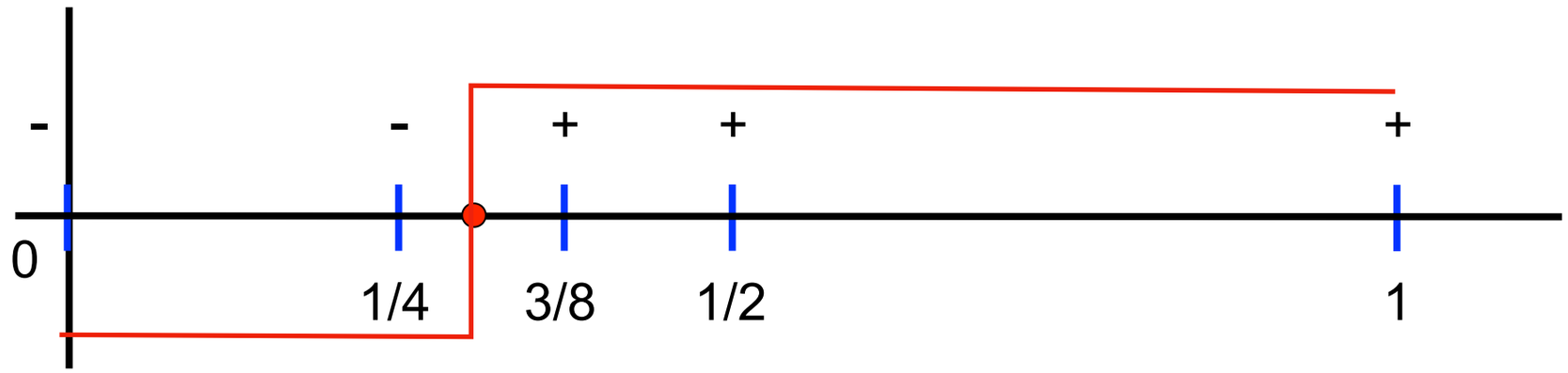
$$\widehat{\mu}_m(x) + \sqrt{\frac{2\log(2/\delta)}{m}} \geq \mu(x) = \Delta \qquad \text{(assuming } f(x) = +1\text{)}$$

the sign of $\widehat{\mu}_m(x)$ is correct with probability at least $1-\delta$ if

$$\Delta - \sqrt{\frac{2\log(2/\delta)}{m}} \geq 0 \qquad \text{or equivalently} \qquad m \;\geq\; \frac{2\log(2/\delta)}{\Delta^2}$$

Noisy Binary Search

Test k points in binary search



probability that one or more tests fail is less than $k\delta$ so substitute

$$\delta \to \frac{\delta}{k}$$

union bound:
$\mathbb{P}(A \text{ or } B) \leq \mathbb{P}(A) + \mathbb{P}(B)$

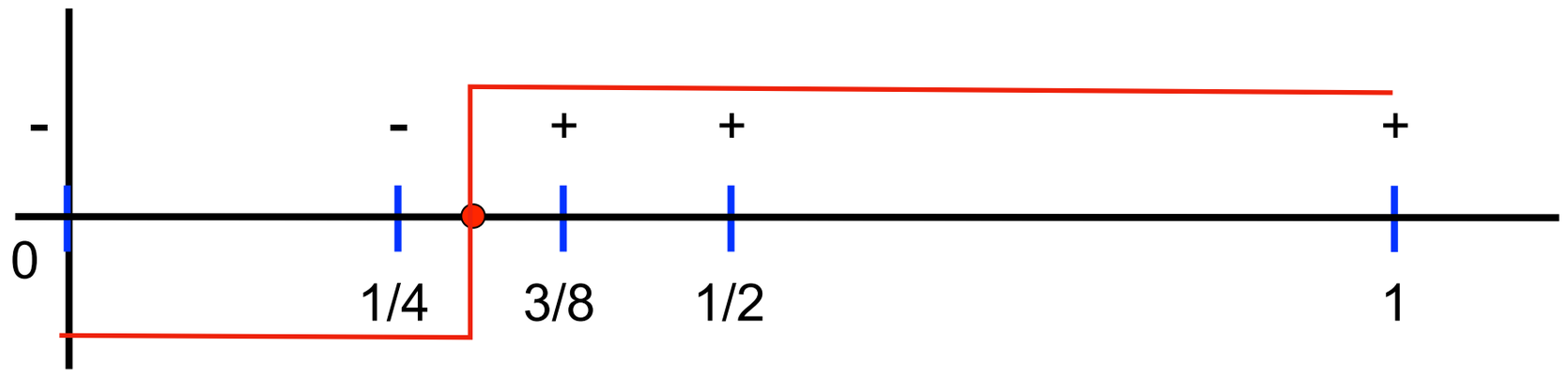
$\Rightarrow$ prob one or more tests fails $\leq \delta$

sufficient number of samples per test:

$$m \;=\; O\left(\frac{2\log(2k/\delta)}{\Delta^2}\right)$$

Noisy Binary Search

Test k points in binary search



sufficient number of samples per test:

$$m = O\left(\frac{2\log(2k/\delta)}{\Delta^2}\right)$$

For ϵ-accurate solution take $k = O(1/\epsilon)$

total number of samples:

$$O\left(\log(1/\epsilon)\underbrace{\frac{2\log(2\log(1/\epsilon)/\delta)}{\Delta^2}}_{\text{price paid due to noise}}\right)$$

Machine Learning Perspective

Classification: Machine Learning Perspective

Given training data $\{(x_j, y_j)\}_{j=1}^m$, learn a function f to predict y from x

Consider a finite set of hypotheses $f_1, f_2, \ldots, f_k$ and define risk (error rate) of each:

$$R(f_i) \;:=\; \mathbb{P}(f_i(x) \neq y)$$

error rate on training data:

$$\widehat{R}(f_i) \;=\; \frac{1}{m}\sum_{i=1}^{M} \mathbb{1}(f(x_i) \neq y_i)$$

"empirical risk"

Chernoff's bound:

$$|R(f_i) - \widehat{R}(f_i)| \;\leq\; \sqrt{\frac{\log(2/\delta)}{2m}}$$

all k empirical risks close to true risks w.p. $\geq 1-\delta$:

$$|R(f_i) - \widehat{R}(f_i)| \;\leq\; \sqrt{\frac{\log(2k/\delta)}{2m}} \text{ for } i = 1, \ldots, k$$

union bound:
$\mathbb{P}(A \text{ or } B) \leq \mathbb{P}(A) + \mathbb{P}(B)$

Empirical Risk Minimization (ERM)

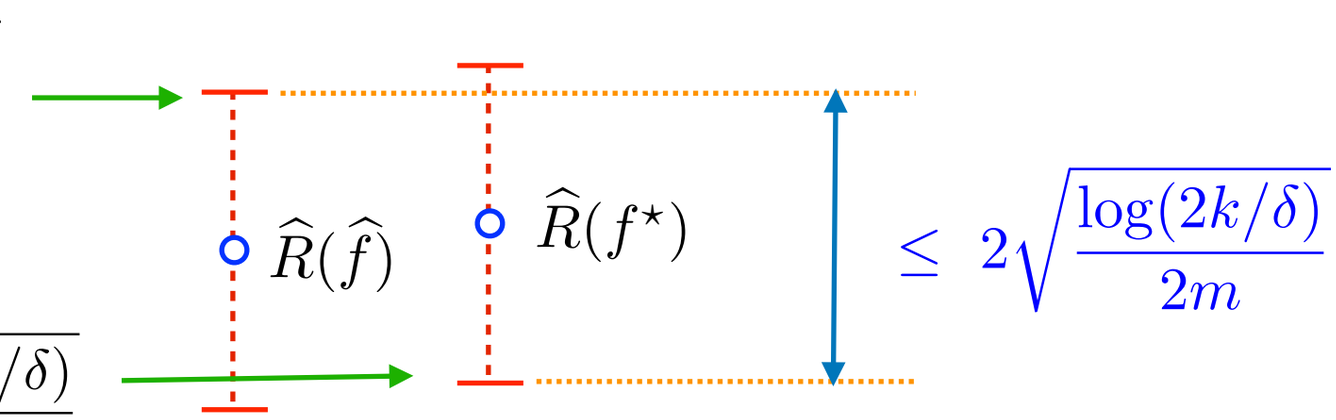
Goal: select hypothesis with true error rate within $\gamma > 0$ of $\min_i R(f_i)$

$$f^\star = \arg\min_{f \in \{f_i\}} R(f) \quad \text{true risk minimizer}$$

$$\widehat{f} = \arg\min_{f \in \{f_i\}} \widehat{R}(f) \quad \text{empirical risk minimizer}$$

$$R(\widehat{f}) \le \widehat{R}(\widehat{f}) + \sqrt{\frac{\log(2k/\delta)}{2m}}$$

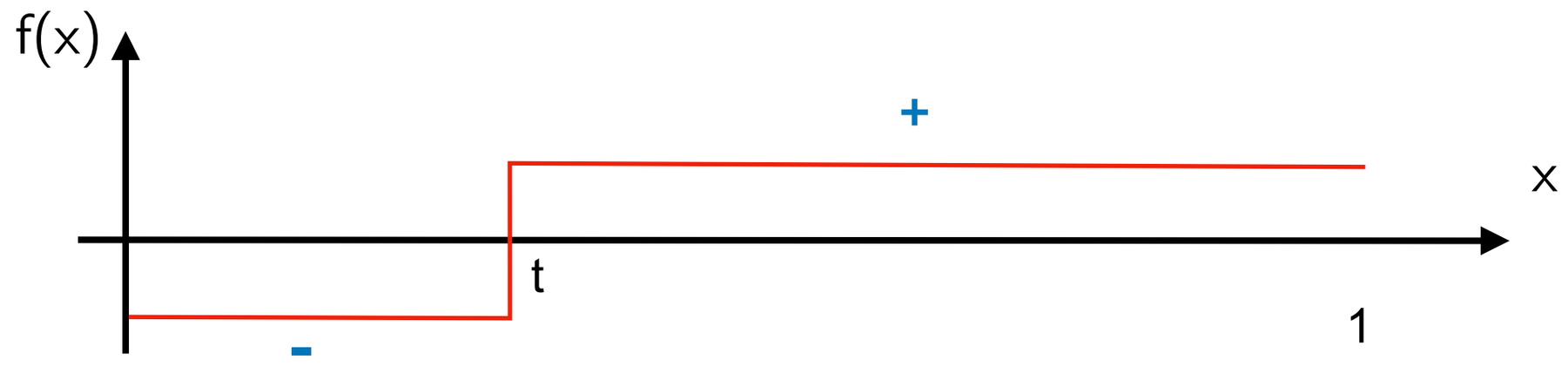
$$R(f^\star) \ge \widehat{R}(f^\star) - \sqrt{\frac{\log(2k/\delta)}{2m}}$$



sufficient number of training examples:

$$2\sqrt{\frac{\log(2k/\delta)}{2m}} \le \gamma \text{ or } m \ge \frac{2\log(2k/\delta)}{\gamma^2}$$

ERM to Learn 1-D Binary Classifier

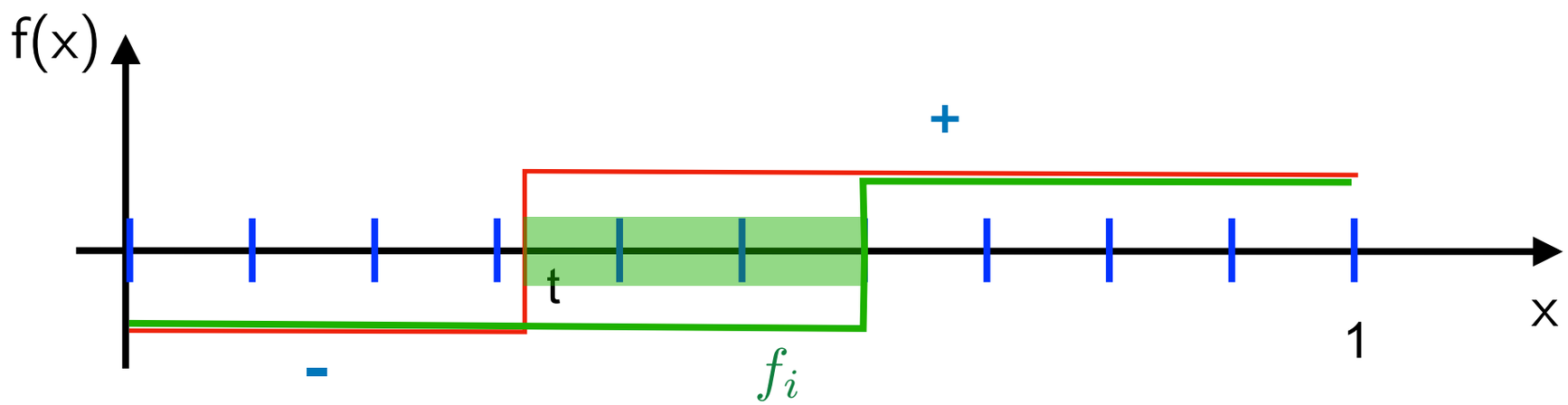


Analyze ERM under following data assumptions:

$$x \sim \text{uniform}(0,1)$$

$$y|x = \begin{cases} f(x) & \text{w.p. } (1+\Delta)/2 \\ -f(x) & \text{w.p. } (1-\Delta)/2 \end{cases}$$

Hypothesis Class



$$f_i(x) = \begin{cases} -1 & \text{if } x < i/k \\ \\ +1 & \text{if } x \geq i/k \end{cases} \qquad i = 1, \ldots, k$$

error rates: let I_i denote subinterval where $f_i \neq f$

$$\begin{aligned} R(f_i) &= \mathbb{P}(f_i(x) \neq y) \\ &= \mathbb{P}(x \in I_i)\,\mathbb{P}(f_i(x) \neq y | x \in I_i) \; + \; (1 - \mathbb{P}(x \in I_i))\mathbb{P}(f_i(x) \neq y | x \notin I_i) \\ &= |t - i/k|(1 + \Delta)/2 \; + \; (1 - |t - i/k|)(1 - \Delta)/2 \\ &= |t - i/k|\Delta + (1 - \Delta)/2 \end{aligned}$$

Near-Optimal Classifier Selection

$$R(f_i) \quad = \quad |t - i/k|\Delta + (1 - \Delta)/2$$

$$R(f_i) - R(f) = |t - i/k|\Delta \qquad\qquad \min_i R(f_i) - R(f) \leq \Delta/k$$

to select hypothesis with error within γ of best:

$$m \geq \frac{2\log(2k/\delta)}{\gamma^2}$$

to find best in class set $\gamma = \Delta/k$

$$m \geq \frac{2k^2\log(2k/\delta)}{\Delta^2}$$

to locate optimal threshold to within ϵ set $k = 1/\epsilon$

$$m = O\left(\frac{1}{\epsilon^2}\frac{2\log(2/\epsilon\delta)}{\Delta^2}\right)$$ scales quadratically in ϵ^{-1}

Active Learning

Active Binary Classification

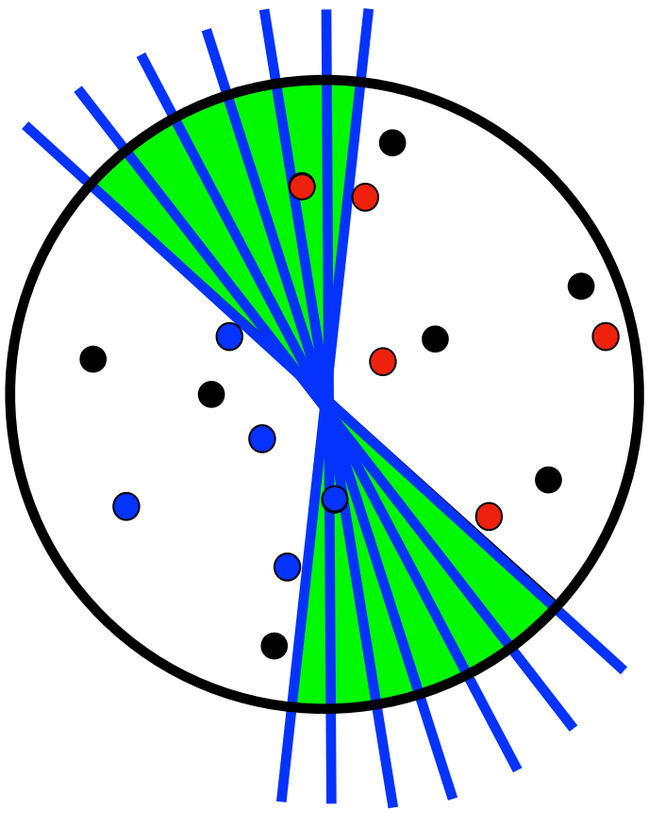
Active Learning for Classifiers

initialize: $\mathcal{H}$ = all linear classifiers

while (*stopping-criterion*) not met

1. **sample** at random from available dataset
2. **label** only those samples in *region of disagreement* of $\mathcal{H}$
3. **reduce** $\mathcal{H}$ by removing all probably suboptimal classifiers

output: a classifier in final $\mathcal{H}$



ϵ-optimal classifier requires

passive: $O(1/\epsilon^2)$ labeled examples

acrtive: $O(\log(1/\epsilon))$ labeled examples

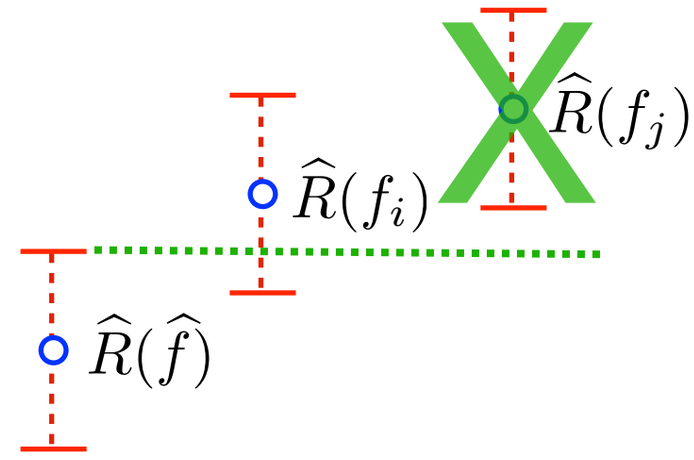
(Balcan & Long '13)

Active Binary Classification

key idea: successive elimination of hypotheses

use initial training data to remove all hypotheses
that are probably worse than the best

$$\widehat{R}(f_j) - \sqrt{\frac{\log(2k/\delta)}{2m}} \geq \widehat{R}(\widehat{f}) + \sqrt{\frac{\log(2k/\delta)}{2m}}$$



Active Learning for Classifiers

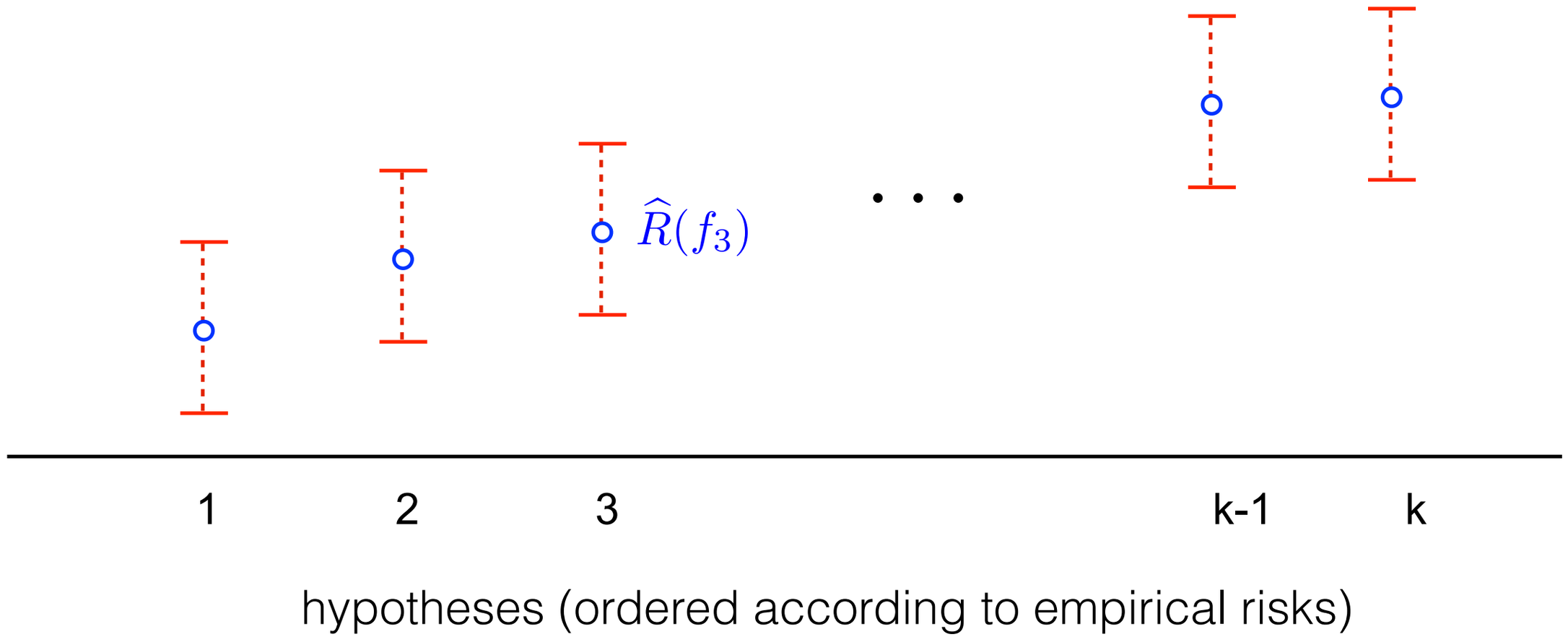
initialize: $\mathcal{H} =$ all linear classifiers, assume $|\mathcal{H}| = k$

while (*stopping-criterion*) not met

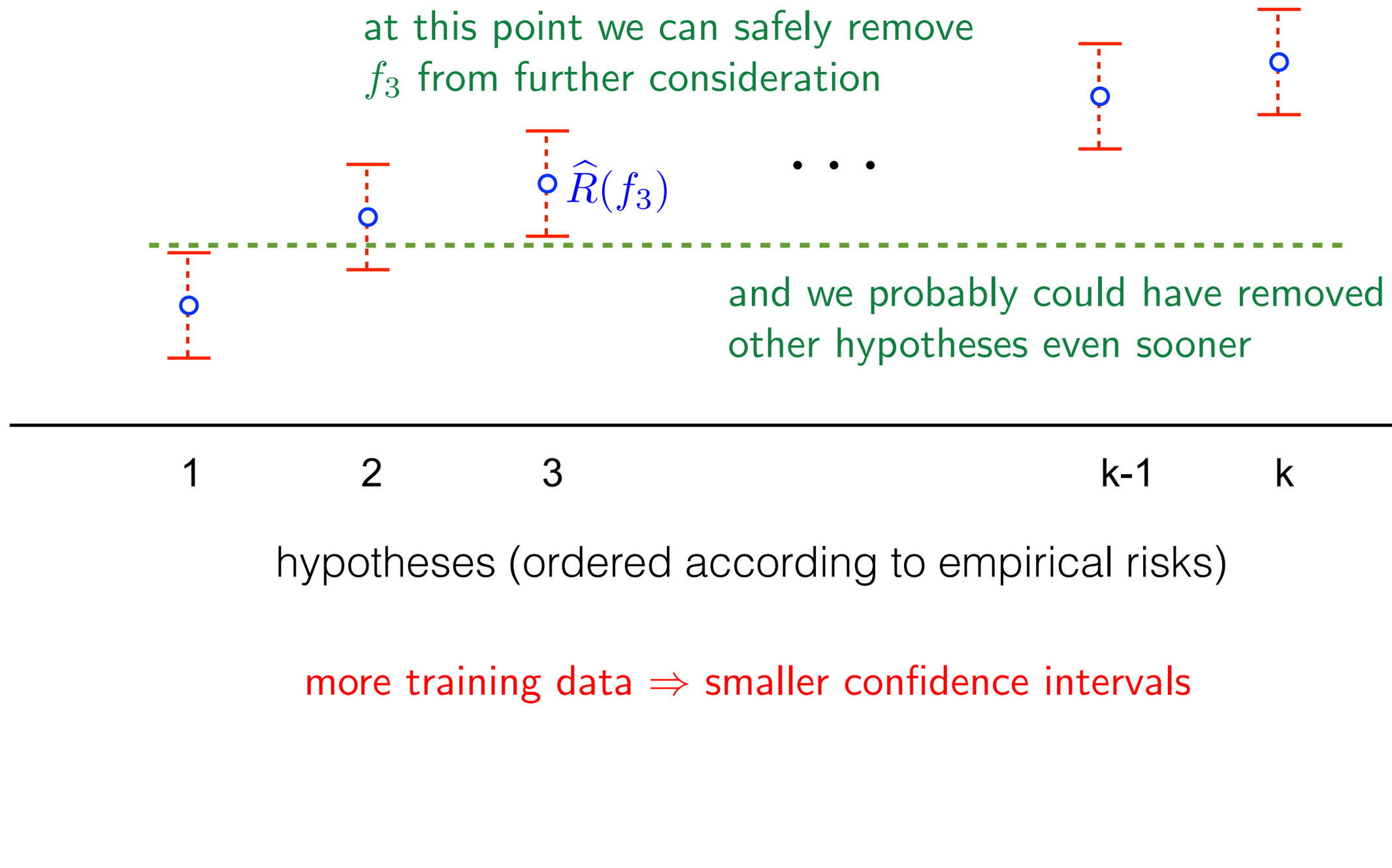
1. **sample** at random from available dataset
2. **label** only m samples in *region of disagreement* of $\mathcal{H}$
3. **reduce** remove all hypotheses f_j that satisfy
   $\widehat{R}(f_j) - \sqrt{\log(2k/\delta)/2m)} \;\geq\; \widehat{R}(\widehat{f}) + \sqrt{\log(2k/\delta)/2m)}$

output: a classifier in final $\mathcal{H}$

Empirical Risks and Confidence Intervals



Empirical Risks and Confidence Intervals



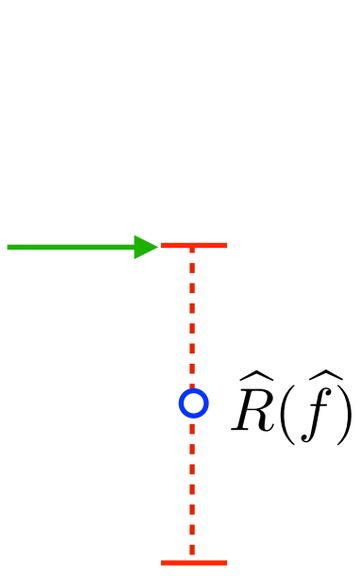
more training data $\Rightarrow$ smaller confidence intervals

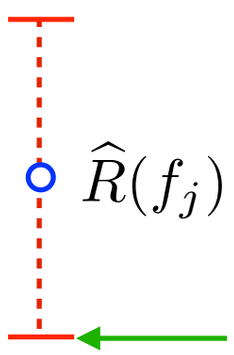
Chipping Away at Error Rate

suppose we want to remove all hypotheses with error rates more than γ larger than that of best classifier

how large should m be to ensure that no hypotheses with $R(f_j) \geq \min_i R(f_i) + \gamma$ remain?

$\min_i R(f_i)$ must be less than this





must be at least $\min_i R(f_i) + \gamma - 2\sqrt{\frac{\log(2k/\delta)}{2m}}$

so m must be large enough to satisfy

$$2\sqrt{\frac{\log(2k/\delta)}{2m}} \leq \gamma \quad \text{or equivalently} \quad m \geq \frac{2\log(2k/\delta)}{\gamma^2}$$

Active Binary Classification

Active Learning for Classifiers

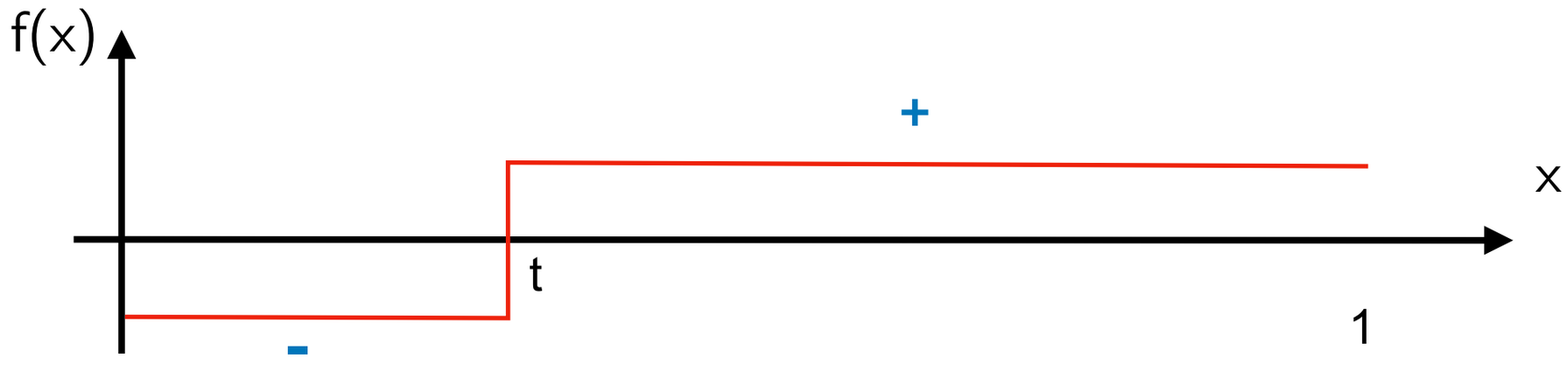
initialize: $\mathcal{H} =$ all linear classifiers, assume $|\mathcal{H}| = k$

while (*stopping-criterion*) not met

1. **sample** at random from available dataset
2. **label** $m = \lceil \frac{2\log(2k/\delta)}{\gamma^2} \rceil$ samples in *disagreement region*
3. **reduce** remove all hypotheses f_j that satisfy
$$\widehat{R}(f_j) - \sqrt{\log(2k/\delta)/2m)} \;\geq\; \widehat{R}(\widehat{f}) + \sqrt{\log(2k/\delta)/2m)}$$

output: a classifier in final $\mathcal{H}$

Active Learning of 1-D Classifier



analyze active learning under following data assumptions:

$$x \sim \text{uniform}(0,1)$$

$$y|x = \begin{cases} f(x) & \text{w.p. } (1+\Delta)/2 \\ -f(x) & \text{w.p. } (1-\Delta)/2 \end{cases}$$

assume we have access to a very large (unlimited) pool of unlabeled data

hypotheses: $$f_i(x) = \begin{cases} -1 & \text{if } x < i/k \\ +1 & \text{if } x \geq i/k \end{cases} \qquad i = 1, \ldots, k$$

Active Learning of 1-D Classifier

Active Learning for Classifiers

initialize: $\mathcal{H} = \{f_1, \ldots, f_k\}$

while $|\mathcal{H}| > 1$

1. **sample** at random from available dataset
2. **label** m samples in *region of disagreement* of $\mathcal{H}$
3. **reduce** remove all hypotheses f_j that satisfy

$$\widehat{R}(f_j) - \sqrt{\log(2k/\delta)/2m)} \;\geq\; \widehat{R}(\widehat{f}) + \sqrt{\log(2k/\delta)/2m)}$$

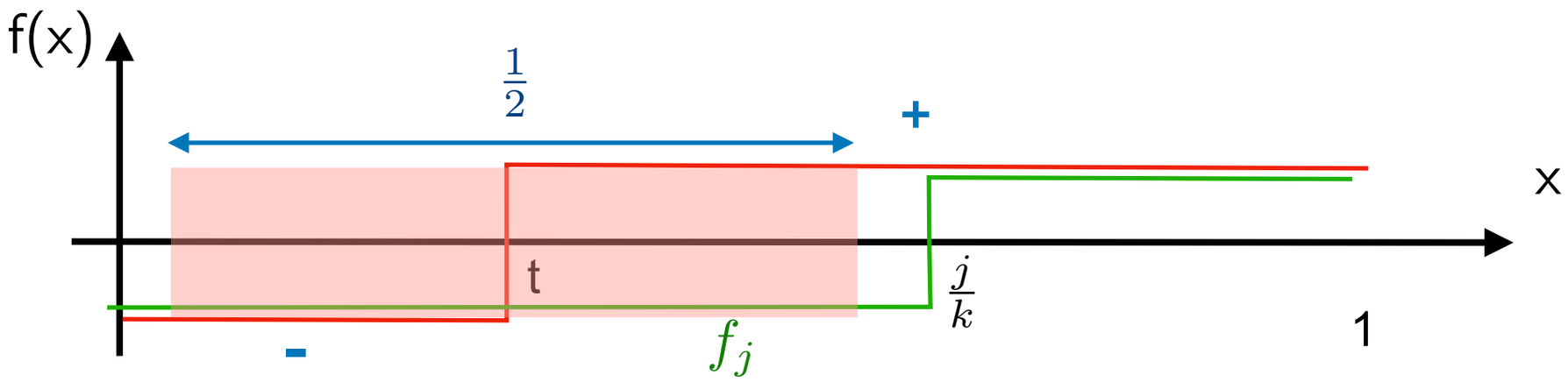
output: classifier in final $\mathcal{H}$

consider special case of successive elimination: halving
(remove half of remaining hypotheses at each step)

this will require $O(\log k)$ stages (just like binary search)

Stage 1 $\mathcal{H} = \{f_1, \ldots, f_k\}$

to remove at least $1/2$ of the hypotheses it suffices
to remove all f_j for which $|\frac{j}{k} - t| \geq 1/4$



the error rates of such classifiers satisfy

$$R(f_j) - \min R(f_i) \geq \Delta/4$$

setting $\gamma = \Delta/4$ yields requirement $m \geq \frac{2\log(2k/\delta)}{(\Delta/4)^2}$

Subsequent Stages

$\leq k/2$ hypotheses will remain after stage 1,

$\leq k/4$ after stage 2,

and so on

we only require confidence intervals for the hypotheses remaining after each stage, so the total number is at most $k + k/2 + k/4 + \cdots \leq 2k$

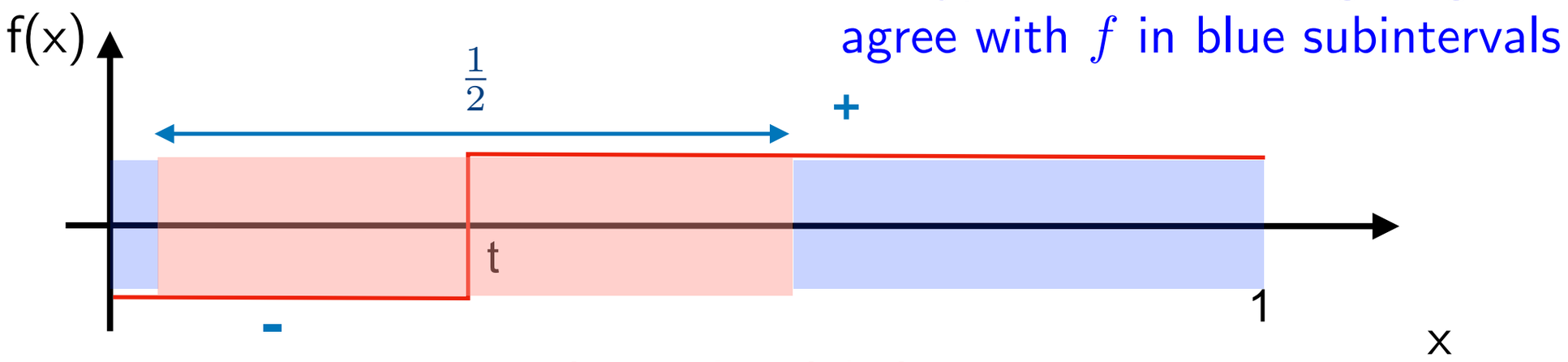
so rather than just k confidence intervals, over the entire active learning process we will consider up to $2k$ intervals: modify the sample size to account for this

$$m \geq \frac{2\log(2k/\delta)}{(\Delta/4)^2} \longrightarrow m \geq \frac{2\log(4k/\delta)}{(\Delta/4)^2}$$

union bound:
$\mathbb{P}(A \text{ or } B) \leq \mathbb{P}(A) + \mathbb{P}(B)$

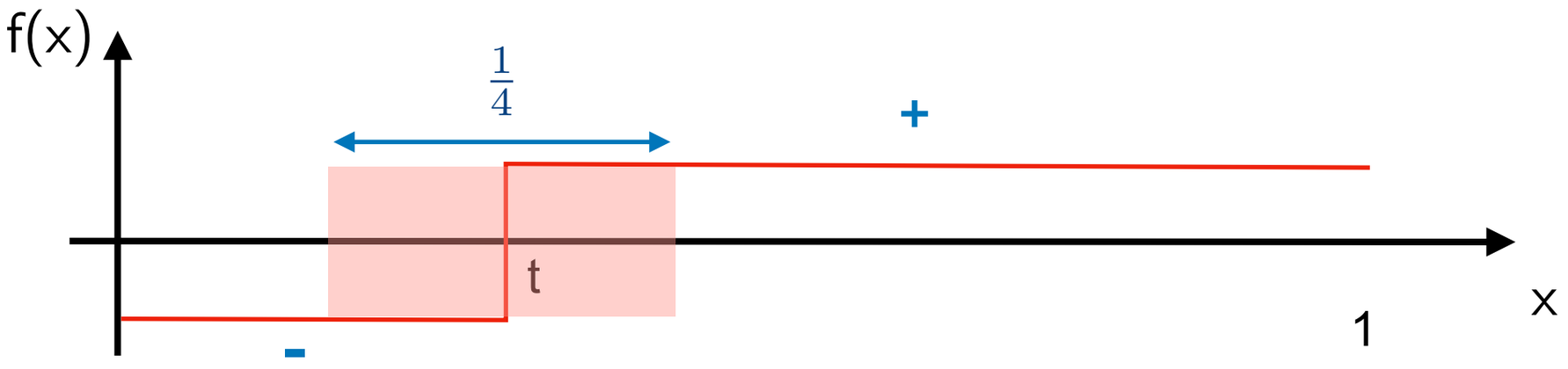
Stage 2

we now want to remove all f_j with $|\frac{j}{k} - t| \leq 1/8$

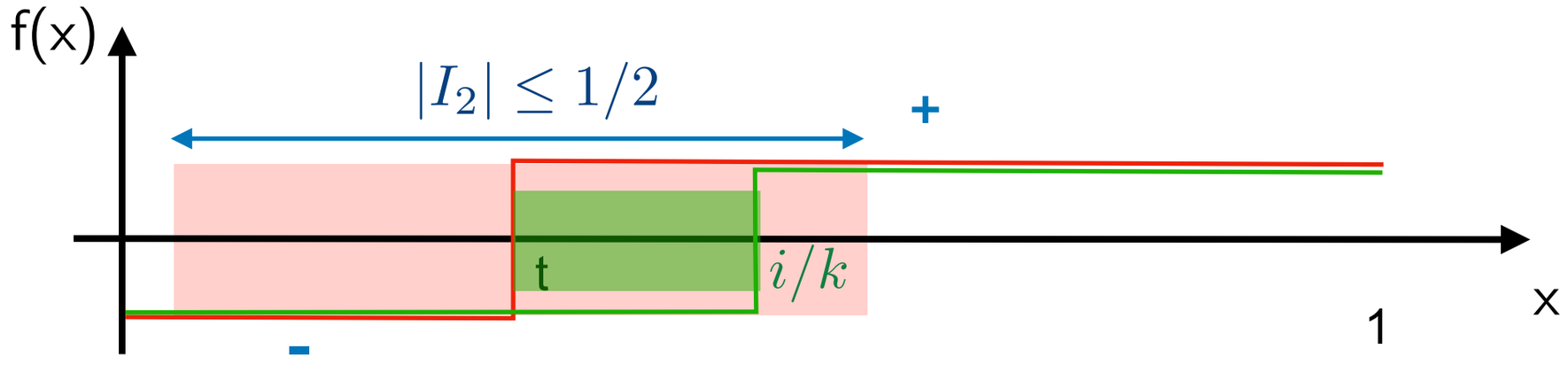


only need to label
examples in red interval

after stage 2



Each Stage is Carbon Copy of the First



since labeled examples only taken in interval of width $|I_2|$,
probability of example in green disagreement region $= \frac{|t-i/k|}{|I_2|}$

$$
\begin{aligned}
R_2(f_i) &= \mathbb{P}(f_i(x) \neq y | x \in I_2) \\
&= \frac{1}{|I_2|} |t - i/k| \, (1+\Delta)/2 \; + \; \frac{1}{|I_2|} (|I_2| - |t - i/k|) \, (1-\Delta)/2 \\
&= \frac{1}{|I_2|} |t - i/k| \, \Delta \; + \; \frac{1-\Delta}{2}
\end{aligned}
$$

so for f_j with $|t - j/k| \geq 1/8$

$$
R_2(f_j) - \min_i R_2(f_i) \; \geq \; \frac{1}{|I_2|} \frac{1}{8} \Delta \; \geq \; \Delta/4
$$

exactly the same as in stage 1

Subsequent Stages

sufficient number of labeled examples per stage:

$$m \geq \frac{2\log(4k/\delta)}{(\Delta/4)^2}$$

total number of labeled examples in $O(\log k)$ stages:

$$m = O\left(\frac{\log k \log(4k/\delta)}{\Delta^2}\right)$$

guarantee: $f^\star = \arg\min_{f \in \{f_i\}_{i=1}^k} R(f)$ selected with probability at least $1-\delta$

to obtain classifier with boundary within ϵ of optimal threshold t set $k = 1/\epsilon$:

$$m = O\left(\frac{\log(\frac{1}{\epsilon})\log(\frac{4}{\epsilon\delta})}{\Delta^2}\right)$$

Active vs. Passive

to obtain classifier with boundary within ϵ of optimal threshold t

active: $$m = O\left(\frac{\log(\frac{1}{\epsilon})\log(\frac{4}{\epsilon\delta})}{\Delta^2}\right)$$

passive: $$m = O\left(\frac{1}{\epsilon^2}\frac{2\log(2/\epsilon\delta)}{\Delta^2}\right)$$

noisy binary search: $$O\left(\log(1/\epsilon)\underbrace{\frac{2\log(2\log(1/\epsilon)/\delta)}{\Delta^2}}_{\text{price paid due to noise}}\right)$$

Dealing with Unknown Noise Levels

the analysis above assumed that noise level Δ was known
(used to set sample size m in each stage)

we can automatically adapt to *unknown*
levels of noise with the sample doubling trick

Active Learning for Classifiers

initialize: $\mathcal{H} = \{f_1, \ldots, f_k\}$

while $|\mathcal{H}| > 1$

1. **sample** at random from available dataset
2. **label** m samples in *region of disagreement* of $\mathcal{H}$
3. **reduce** remove all hypotheses f_j that satisfy
   $\widehat{R}(f_j) - \sqrt{\log(2k/\delta)/2m)} \;\geq\; \widehat{R}(\widehat{f}) + \sqrt{\log(2k/\delta)/2m)}$

   if more than $|\mathcal{H}|/2$ remain, return to label step and double m

output: classifier in final $\mathcal{H}$

Dealing with Unknown Noise Levels

Active Learning for Classifiers

initialize: $\mathcal{H} = \{f_1, \ldots, f_k\}$

while $|\mathcal{H}| > 1$

1. **sample** at random from available dataset
2. **label** m samples in *region of disagreement* of $\mathcal{H}$
3. **reduce** remove all hypotheses f_j that satisfy
   $\widehat{R}(f_j) - \sqrt{\log(2k/\delta)/2m)} \;\geq\; \widehat{R}(\widehat{f}) + \sqrt{\log(2k/\delta)/2m)}$

   if more than $|\mathcal{H}|/2$ remain, return to label step and double m

output: classifier in final $\mathcal{H}$

starting with $m = 2$ and doubling until set of remaining hypotheses is cut in half requires another union bound, resulting in the following bound on the sample complexity per step (automatically adaptive to unknown Δ):

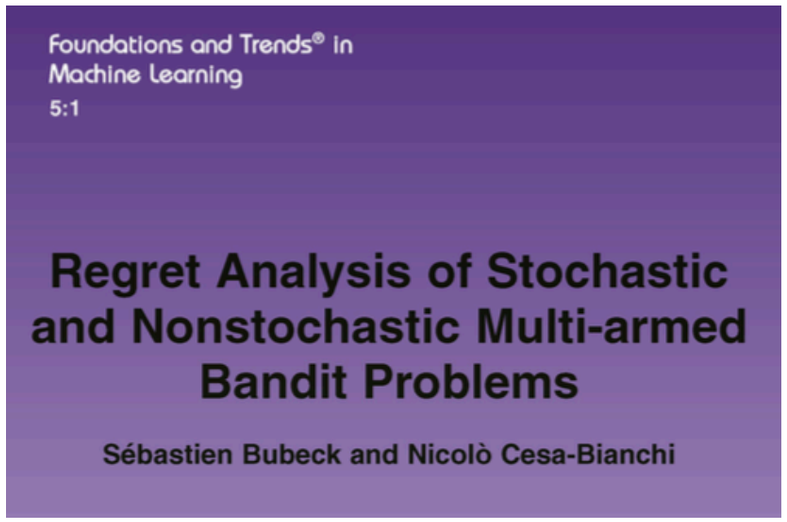
$$m = O\left(\frac{\log(2k/\delta)}{\Delta^2}\log(4\log(2k/\delta)/(\Delta/4)^2)\right)$$

same as before up to small log factor

Agnostic **active learning**
MF Balcan, A Beygelzimer, J **Langford** - … on Machine **learning**, 2006 - dl.acm.org

Multi-Armed Bandits

good starting points for comprehensive reviews



Sébastien Bubeck, from Microsoft Research, talked about the multi bandit problem more

Slides:

- Slides #1
- Slides #2

THE NEW YORKER Cartoon Caption Contest

www.newyorker.com/cartoons/vote

Please rank the entries for this Cartoon Caption Contest image, then click the "Done" button. You can rank as many or as few captions as you like, but five is too few and five thousand is way too many.

"They brought him in from the outside."

UNFUNNY | SOMEWHAT FUNNY | FUNNY

DONE

ROBERT MANKOFF

EMMA ALLEN

- $n \approx 5000$ captions/week
- goal: identify funniest
- $\approx 10K$ raters / week
- 100+ weeks of experiments

Multi-Armed Bandit Problem

Stochastic Bandit Problem

known: $n = \#$arms

unknown: arm "reward" distributions $\{p_1, \ldots, p_n\}$

while (*stopping-criterion*) not met

1. **select** arm $i_t \in \{1, \ldots, n\}$ at step t
2. **draw** reward $x_{i_t,t} \sim p_{i_t}$ independently from past

Goal:

identify arm with largest expected reward (mean)

mean = average rating

classic work: Thompson (1933), Robbins (1952), Bechhofer (1958), Paulson (1964)

UCB algorithm: Lai & Robbins (1985)

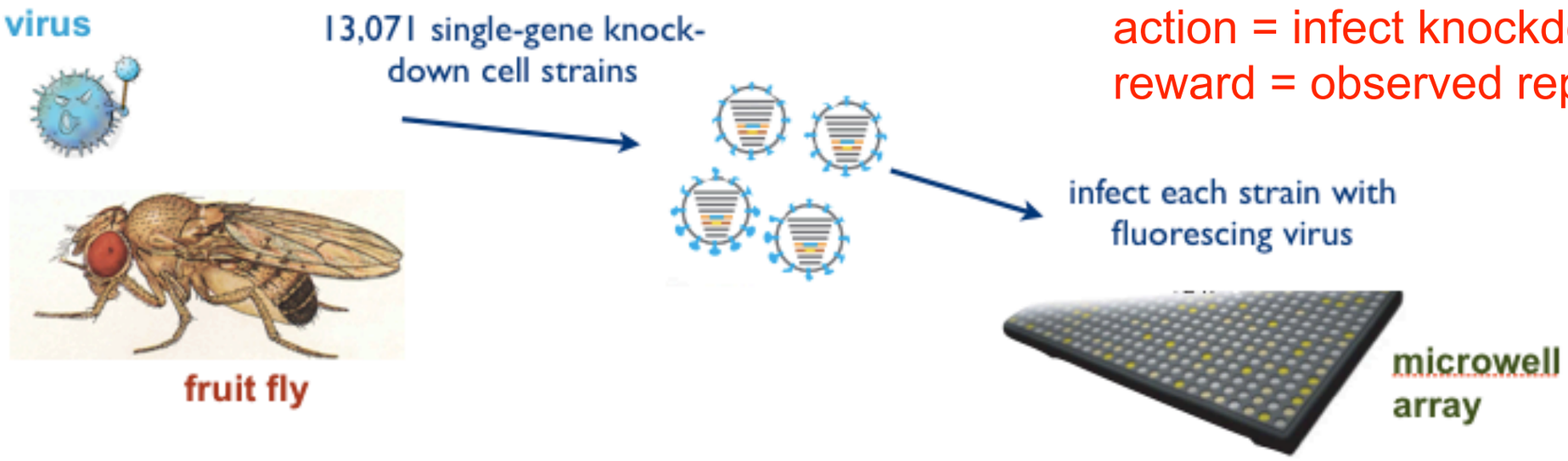
successive elimination algorithm:
Even-Dar, Mannor & Mansour (2006)

optimal algorithms:
Karnin, Koren & Somekh (2013)
Malloy and N (2013)
Jamieson, Malloy, N & Bubeck (2013, 2014)
Tanczos, Mankoff & N (2017)

Applications of the Stochastic Bandit Problem

adaptively select the best ad to display on a webpage

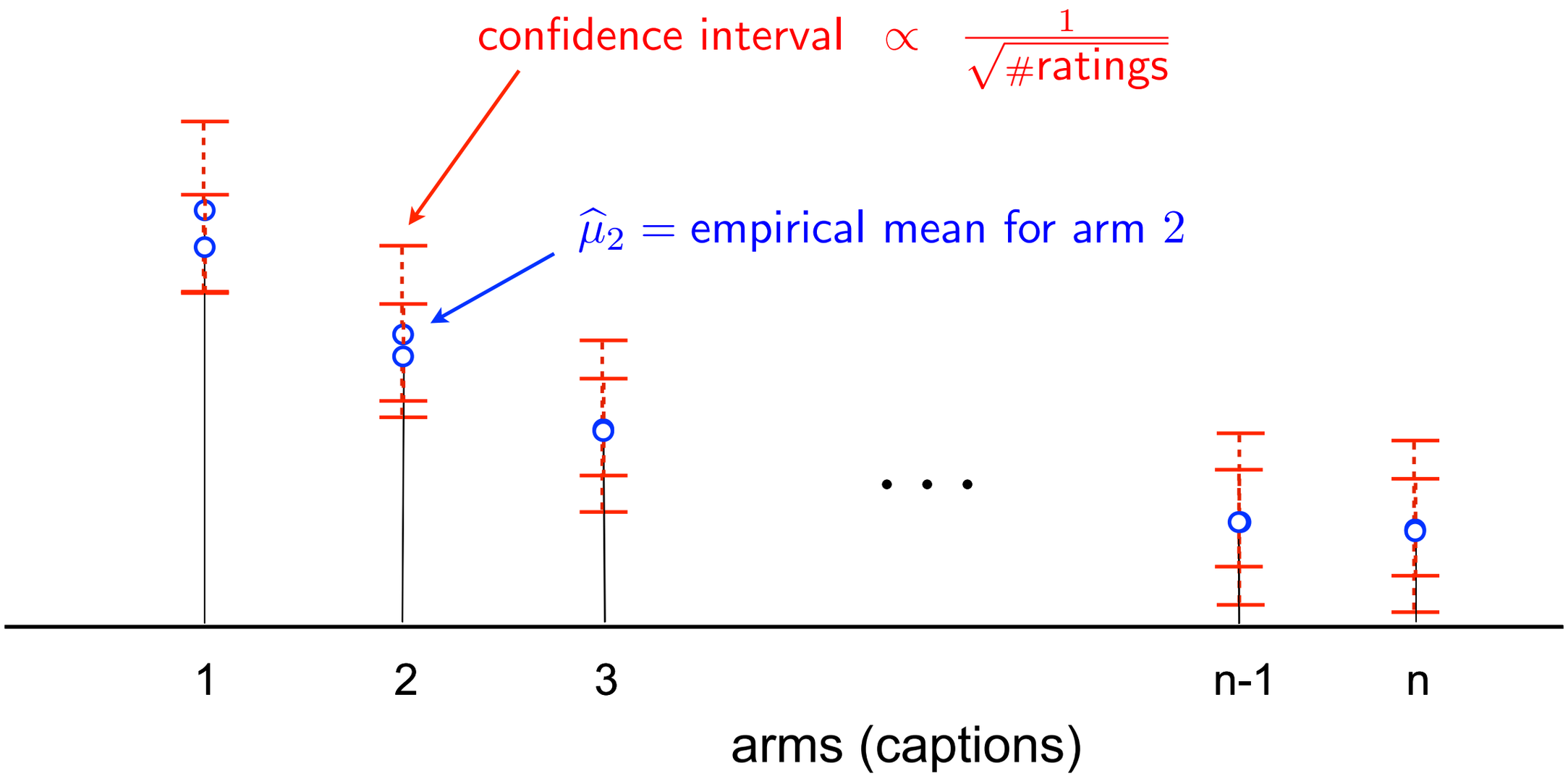
arms = ads/creatives
action = display ad
reward = click



help scientists adaptively select experiments to determine which genes are the most involved in disease

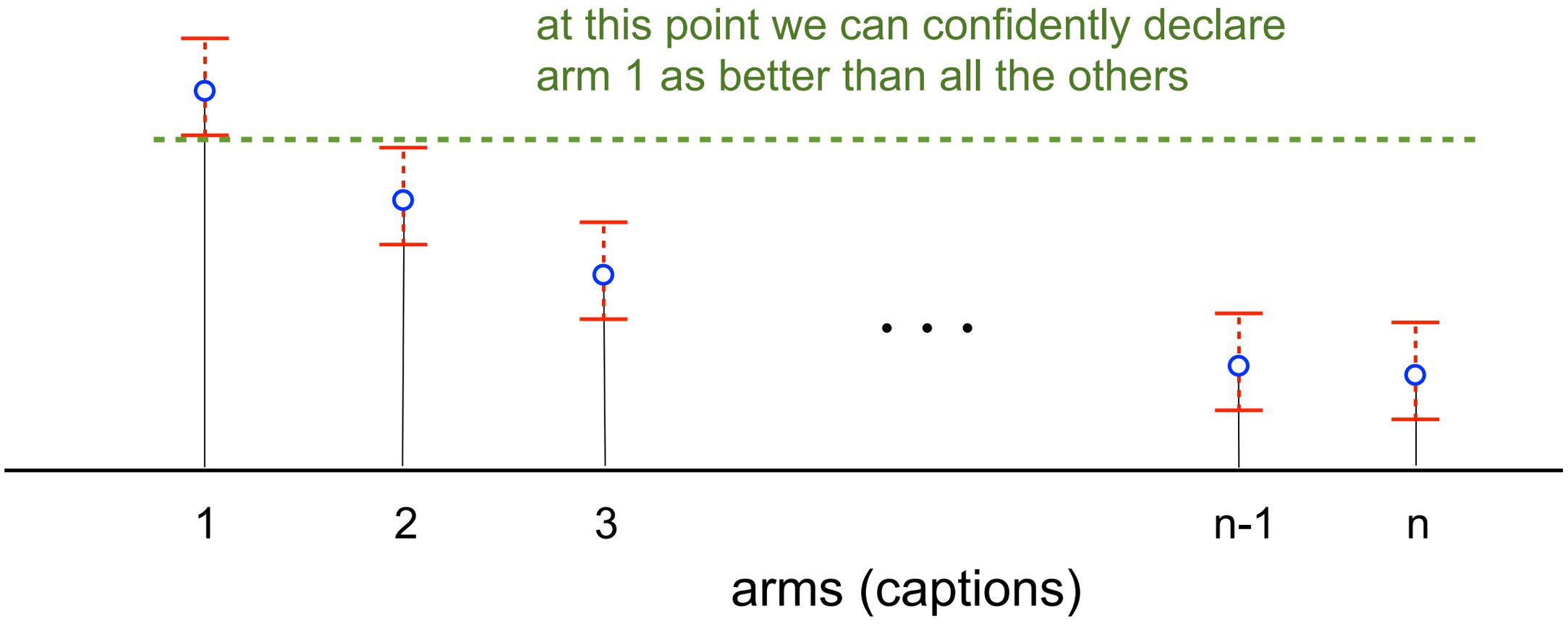
arms = genes/proteins
action = infect knockdown strain
reward = observed replication rate

Empirical Means and Confidence Intervals



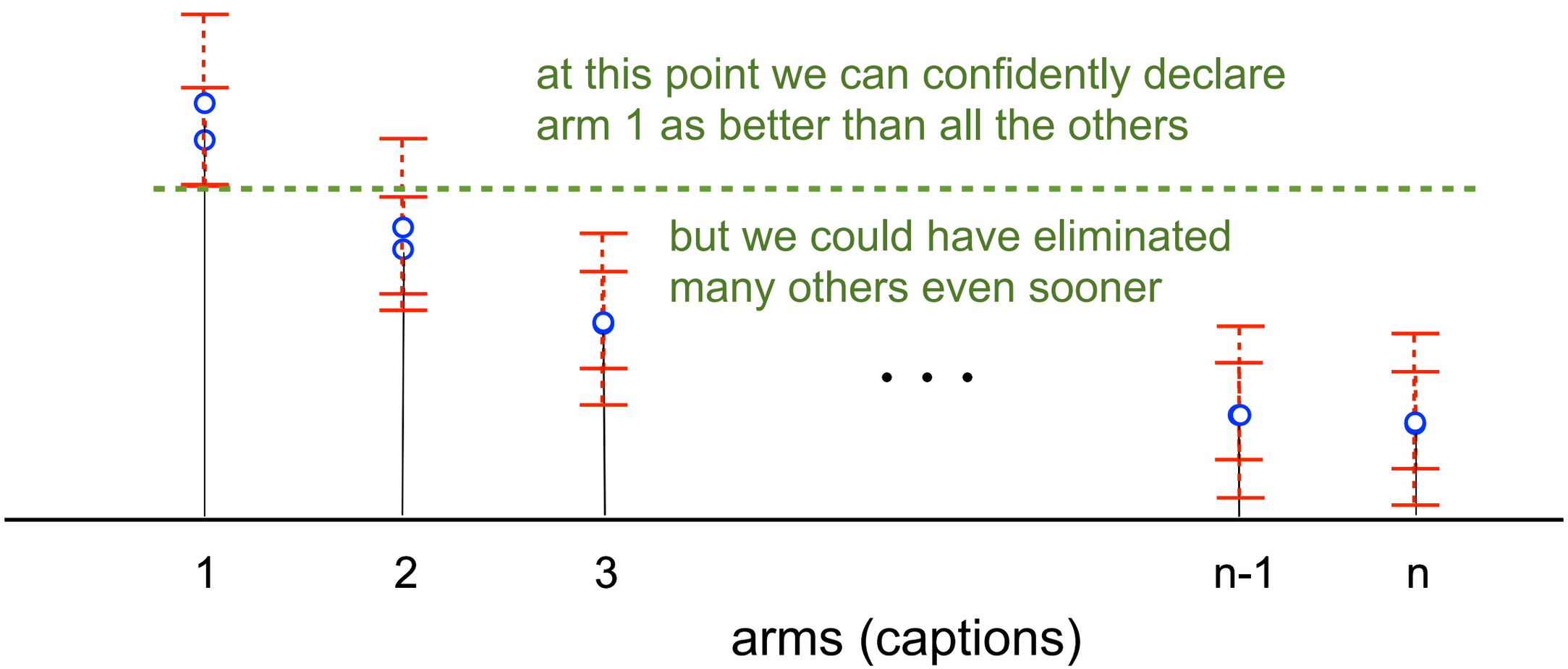
more ratings $\Rightarrow$ smaller intervals

Empirical Means and Confidence Intervals



more ratings $\Rightarrow$ smaller intervals

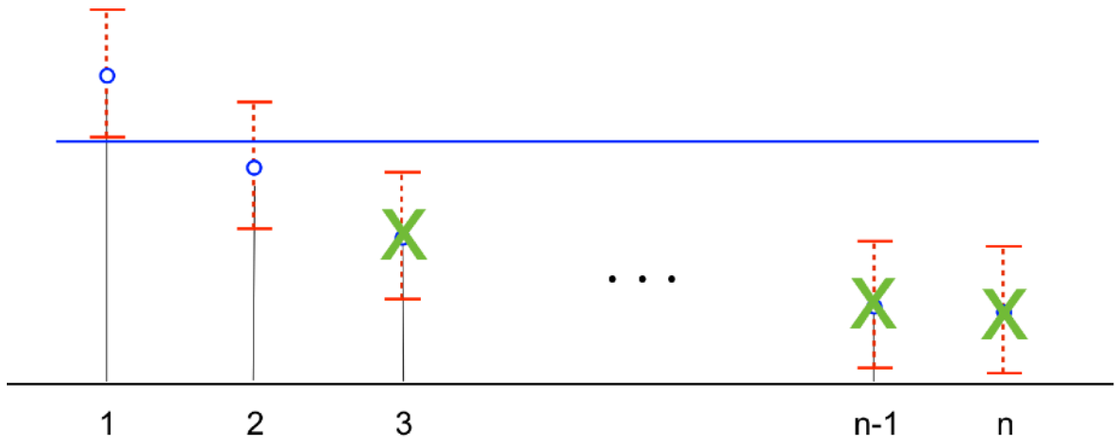
Empirical Means and Confidence Intervals



more samples $\Rightarrow$ smaller intervals

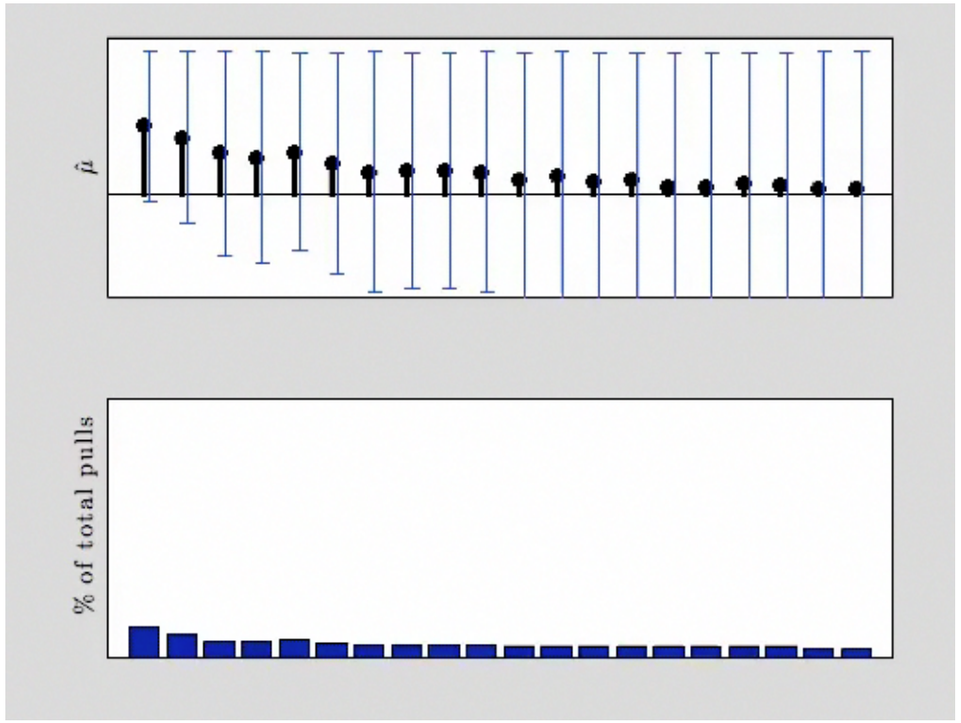
Two Strategies

successive elimination algorithm



stop sampling arms with
upper confidence bounds (UCBs) <
largest lower confidence bound

UCB algorithm



sample arm with largest UCB

both algorithms enjoy optimal sample complexity (up to constant factors)

…but harsh keep-or-kill nature of successive elimination leads to excessively large constant factors, making it ineffective in practice

Notation

n arms

$\mu_1 > \mu_2 \geq \cdots \geq \mu_n$, expected reward of each arm (unknown)

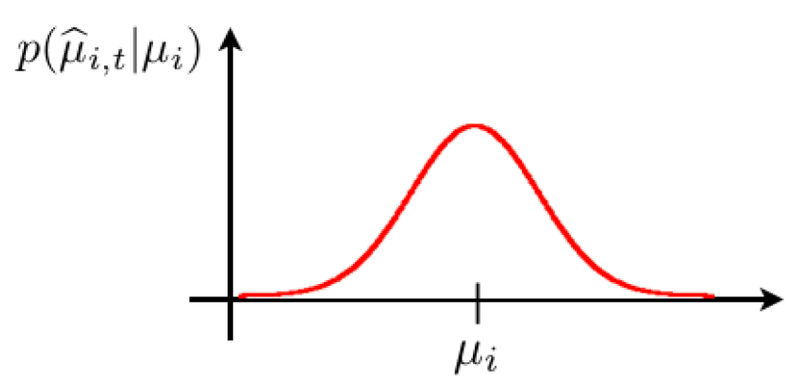
$\Delta_i = \mu_1 - \mu_i$, $i > 1$, "gaps"

$x_{it} \sim p_i$, independent random reward from arm i at time t

$\widehat{\mu}_{i,t_i} = \frac{1}{t_i} \sum_{j=1}^{t_i} x_{ij}$, empirical mean from t_i rewards

bounded/subGaussian rewards:

$$\mathbb{P}(|\widehat{\mu}_{i,t} - \mu_i| \geq \epsilon) \;\leq\; 2e^{-2t\epsilon^2}$$



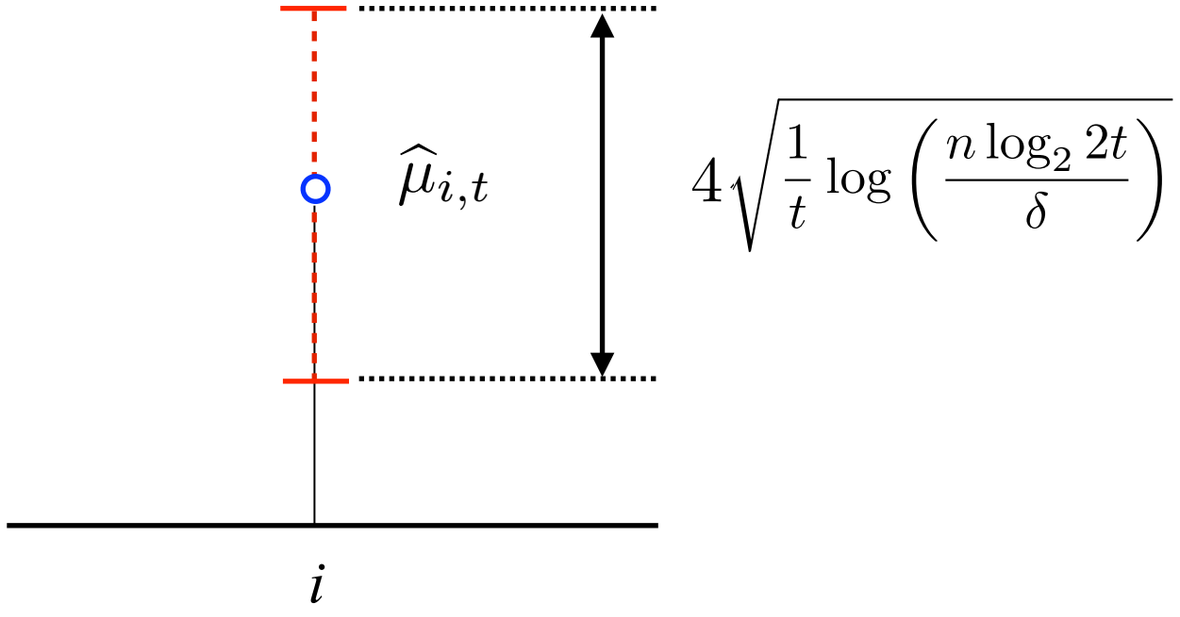
for all i and all t

$$\widehat{\mu}_{i,t} - 2\sqrt{\frac{1}{t}\log\left(\frac{n\log_2 2t}{\delta}\right)} \;\leq\; \mu_i \;\leq\; \widehat{\mu}_{i,t} + 2\sqrt{\frac{1}{t}\log\left(\frac{n\log_2 2t}{\delta}\right)}$$

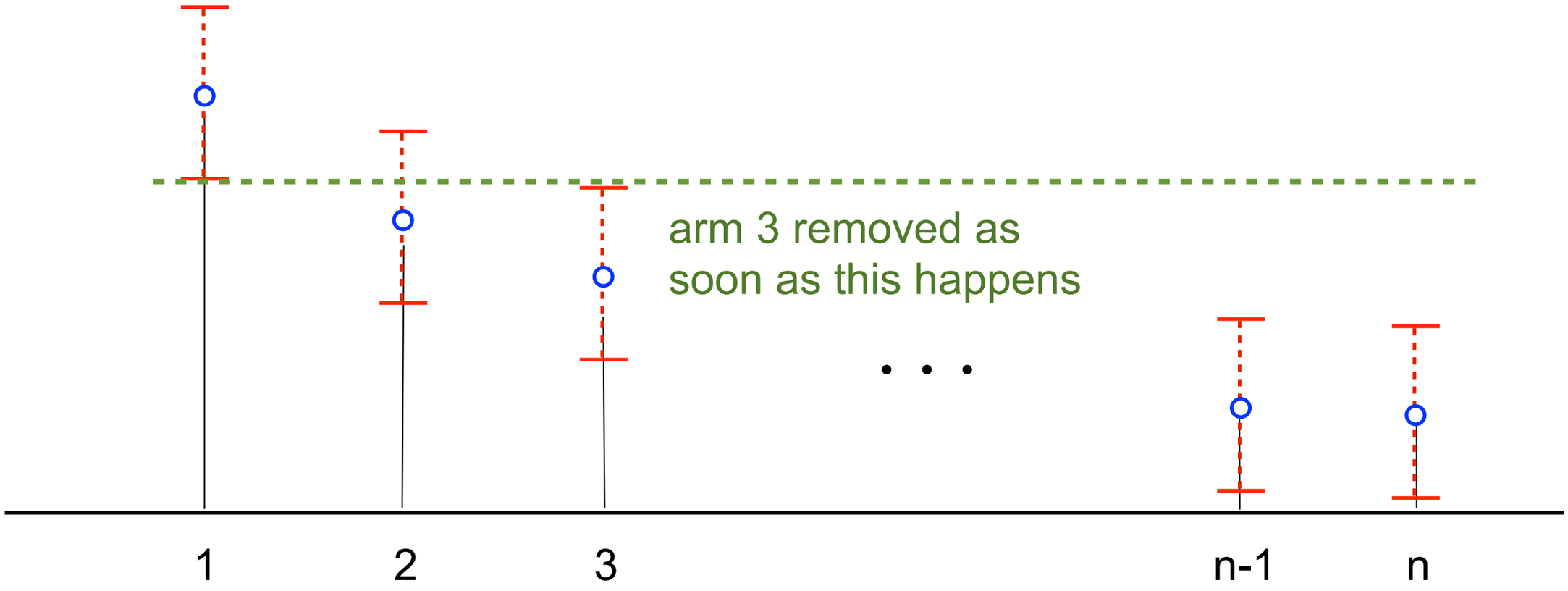
Anytime Confidence Bounds

for all i and all t

$$\widehat{\mu}_{i,t} - 2\sqrt{\frac{1}{t}\log\left(\frac{n\log_2 2t}{\delta}\right)} \;\leq\; \mu_i \;\leq\; \widehat{\mu}_{i,t} + 2\sqrt{\frac{1}{t}\log\left(\frac{n\log_2 2t}{\delta}\right)}$$



Analysis of Successive Elimination

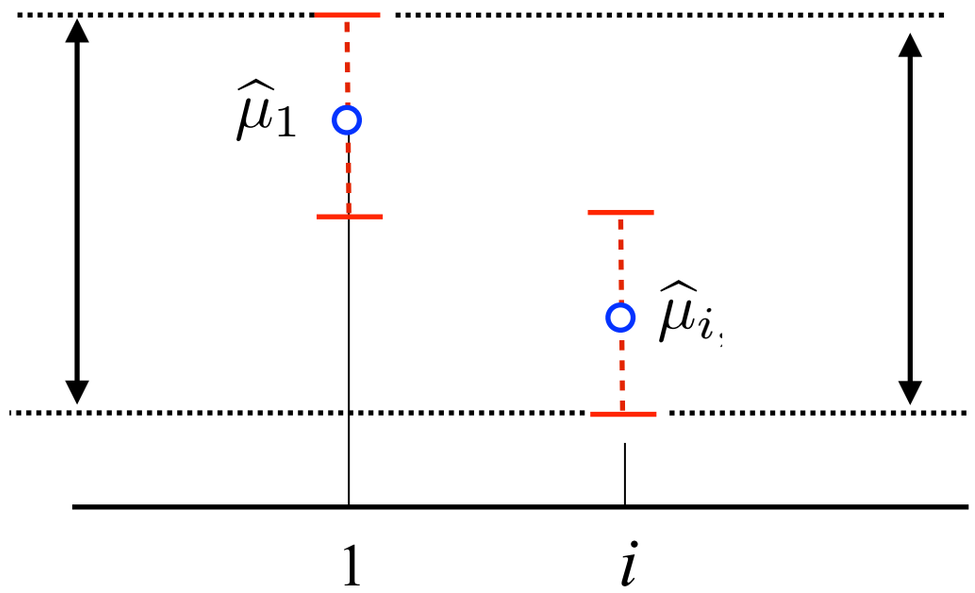


how many times will arm 3 be sampled?

Arm Elimination

if arm i is sampled

$$\Delta_i := \mu_1 - \mu_i \;\leq$$



$$\leq \; 8\sqrt{\frac{1}{t}\log\left(\frac{n\log_2 2t}{\delta}\right)}$$

(otherwise nonoverlapping)

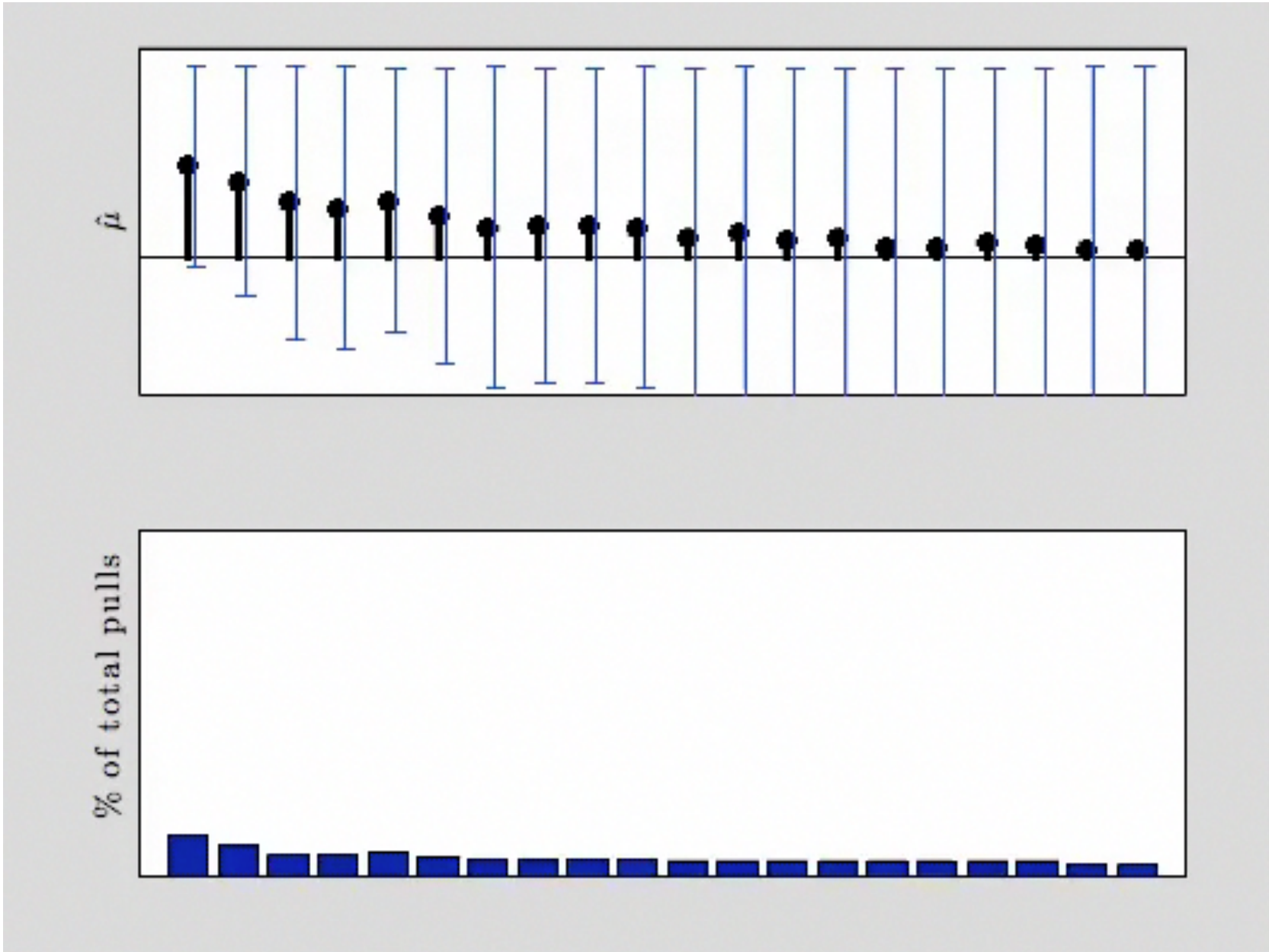
arm i is sampled $\Rightarrow$ $\Delta_i \;\leq\; 8\sqrt{\frac{1}{t}\log\left(\frac{n\log_2 2t}{\delta}\right)}$

$\Rightarrow$ arm i sampled $O(\Delta_i^{-2}\log(n\log(\Delta_i^{-2})/\delta)$ times

total # of samples $\quad O\left(\sum_{i\geq 2}\Delta_i^{-2}\log\left(\frac{n\log\Delta_i^{-2}}{\delta}\right)\right)$

Analysis of UCB

optimism: sample arm with largest upper confidence bound (UCB)

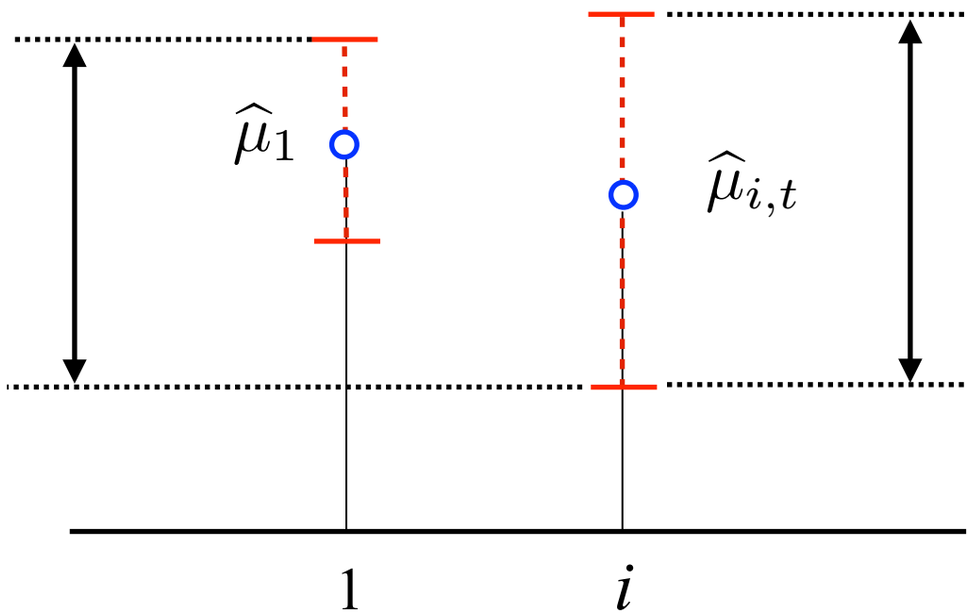


Jamieson, Malloy, N & Bubeck (2014)

UCB Samples Suboptimal Arms Finitely Often

if arm i is sampled

$$\Delta_i := \mu_1 - \mu_i \;\leq$$



$$4\sqrt{\frac{\log\log t/\delta}{t}}$$

arm i is sampled $\Rightarrow$ $\Delta_i \;\leq\; 4\sqrt{\dfrac{\log\log t/\delta}{t}}$

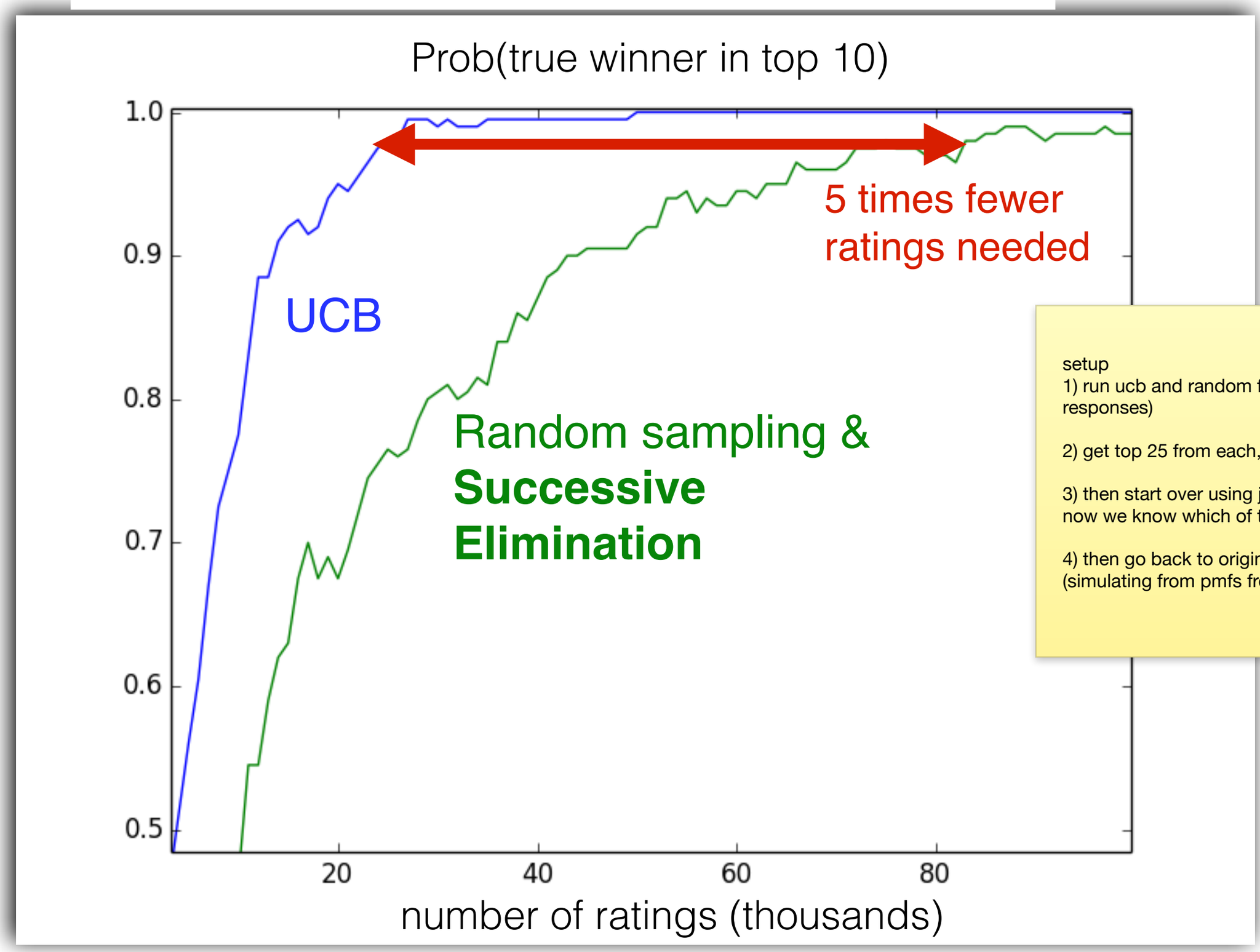
$\Rightarrow$ arm i only sampled $T_i = O(\Delta_i^{-2} \log\log \Delta_i^{-2})$ times

sub-exponential concentration inequality $\Rightarrow$

total # of samples from subopt arms $\quad \displaystyle\sum_{i\geq 2} T_i \;\lesssim\; \sum_{i\geq 2} \Delta_i^{-2} \log\log \Delta_i^{-2}$

Jamieson, Malloy, N & Bubeck (2014)

THE NEW YORKER Ranking Accuracy



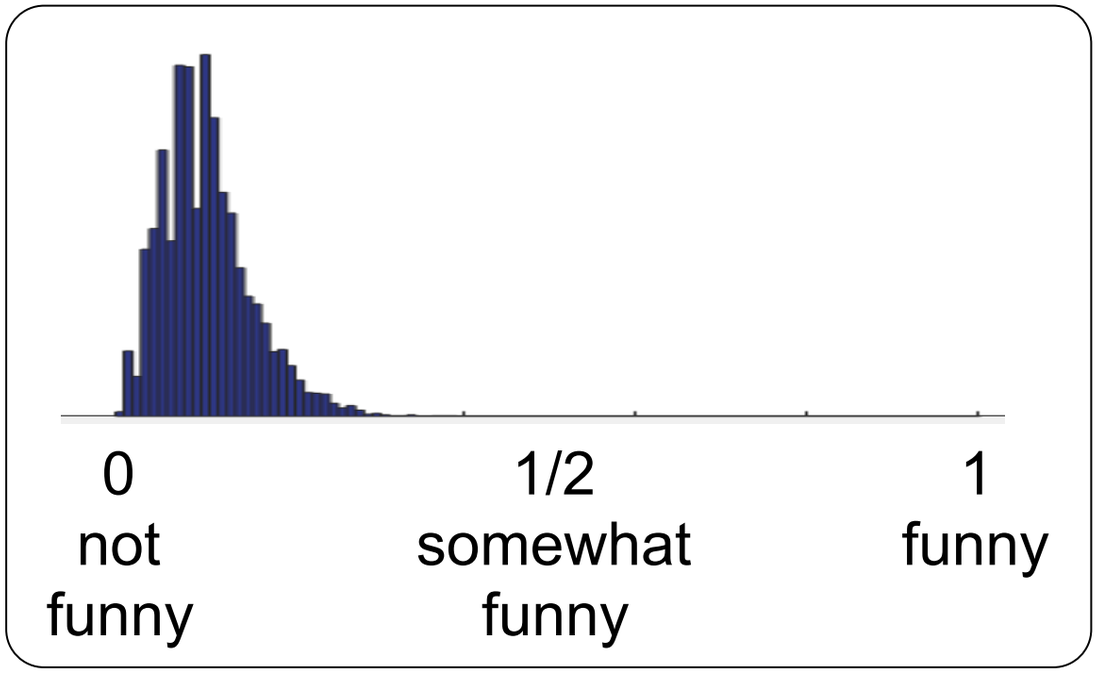
setup
1) run ucb and random for first 3 days (few h
responses)

2) get top 25 from each, take union <= 50

3) then start over using just union and rando
now we know which of these was best

4) then go back to original data... where is w
(simulating from pmfs from original data)

What We Learned: Most Captions Are Not Funny



histogram of ratings
for typical contest

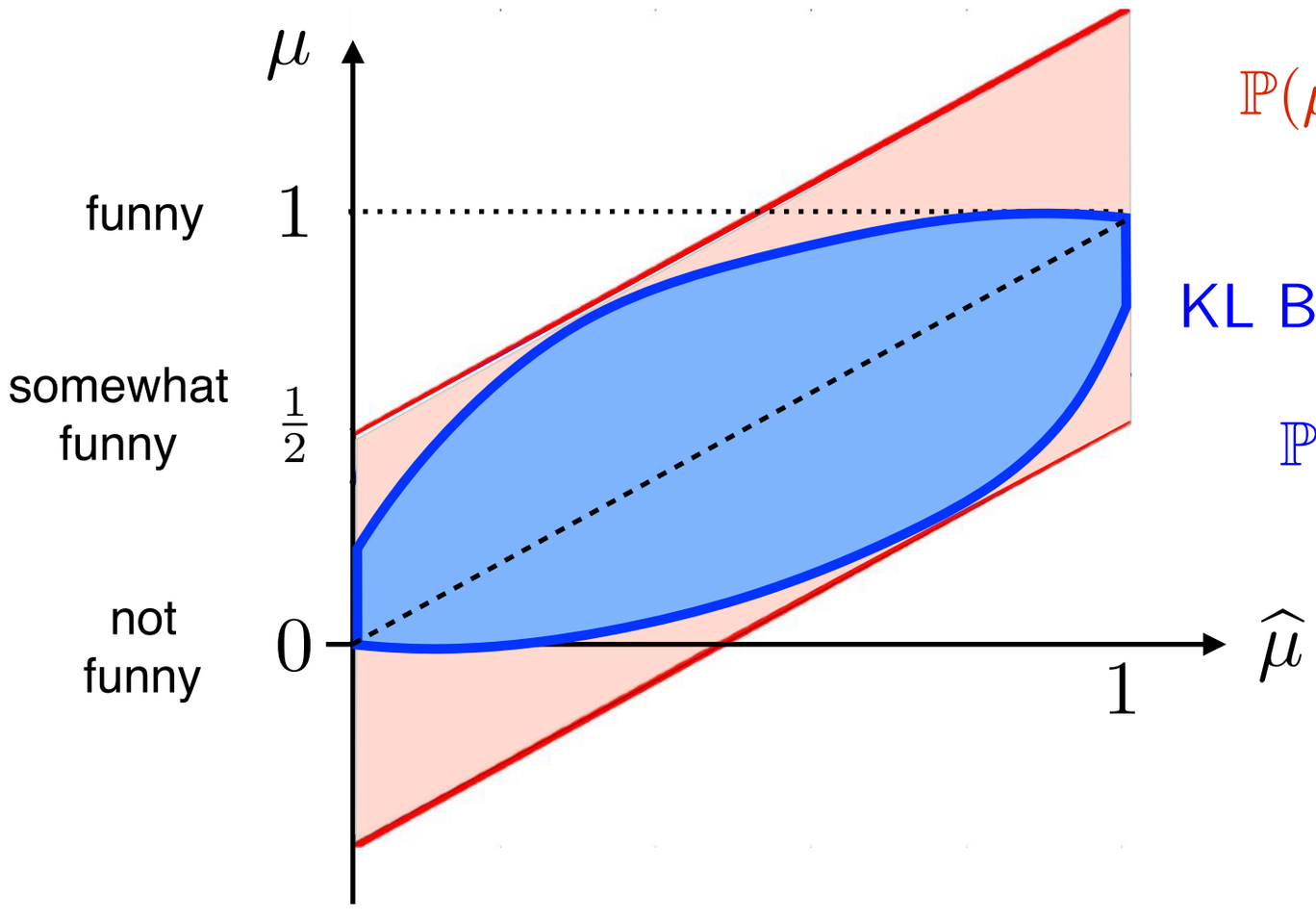
most captions have low average ratings

→ rating variance typically much less than 1/4

sharper confidence bounds can exploit this fact

KL-Based Confidence Intervals

if $x_1, x_2, \ldots$ are independent $[0,1]$-bounded, mean μ rvs,
then $\widehat{\mu}_t = \frac{1}{t}\sum_{i=1}^{t} x_i$ satisfies



subGaussian Bound:

$$\mathbb{P}(\mu - \widehat{\mu}_t \geq \epsilon) \;\leq\; e^{-2t\epsilon^2}$$

KL Bound:

$$\mathbb{P}(\mu - \widehat{\mu}_t > \epsilon) \;\leq\; \exp(-t\,\underbrace{\mathsf{KL}(\mu-\epsilon, \mu)}_{\text{Bernoulli KL}})$$

KL-Based Confidence Intervals

$$\mathbb{P}(\mu - \widehat{\mu}_t \geq \epsilon) \leq e^{-2t\epsilon^2} \quad \Rightarrow \quad \sum_{i \geq 2} T_i \lesssim \sum_{i \geq 2} \Delta_i^{-2} \log\log \Delta_i^{-2}$$

main challenge with KL is deriving finite sample LIL-style confidence intervals

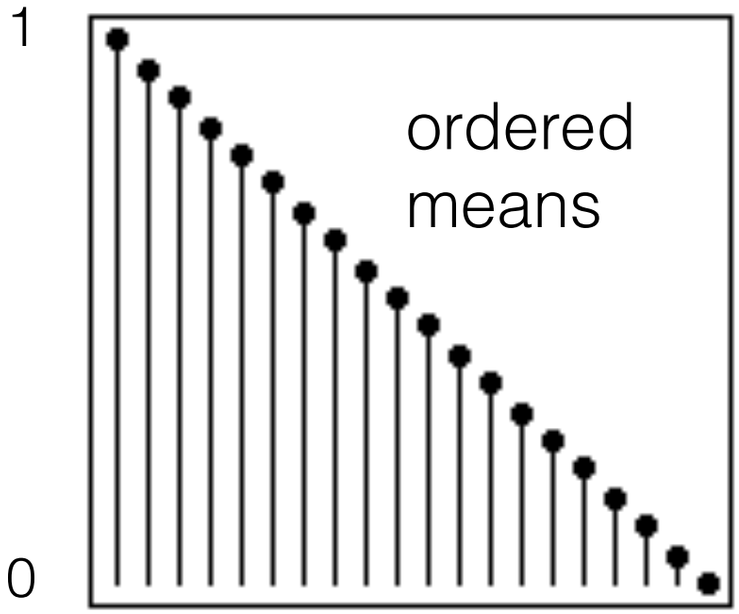
$$\mathbb{P}(\mu - \widehat{\mu}_t > \epsilon) \leq e^{-t\,\mathsf{KL}(\mu-\epsilon,\mu)} \Rightarrow \sum_{i \geq 2} T_i \lesssim \sum_{i \geq 2} C_i^{-1} \log\log C_i^{-1}$$

where $C_i = C(\mu_1, \mu_i) \geq \Delta_i^2$ is the *Chernoff Information* between Bernoulli distributions with means μ_1 and μ_i

if $\mu_1 \approx 1$, then $C_i \leq 2\Delta_i$ $\Rightarrow$ $\sum_{i \geq 2} T_i \lesssim \sum_{i \geq 2} \Delta_i^{-1} \log\log \Delta_i^{-1}$

Sample Complexity: SG vs. KL Intervals

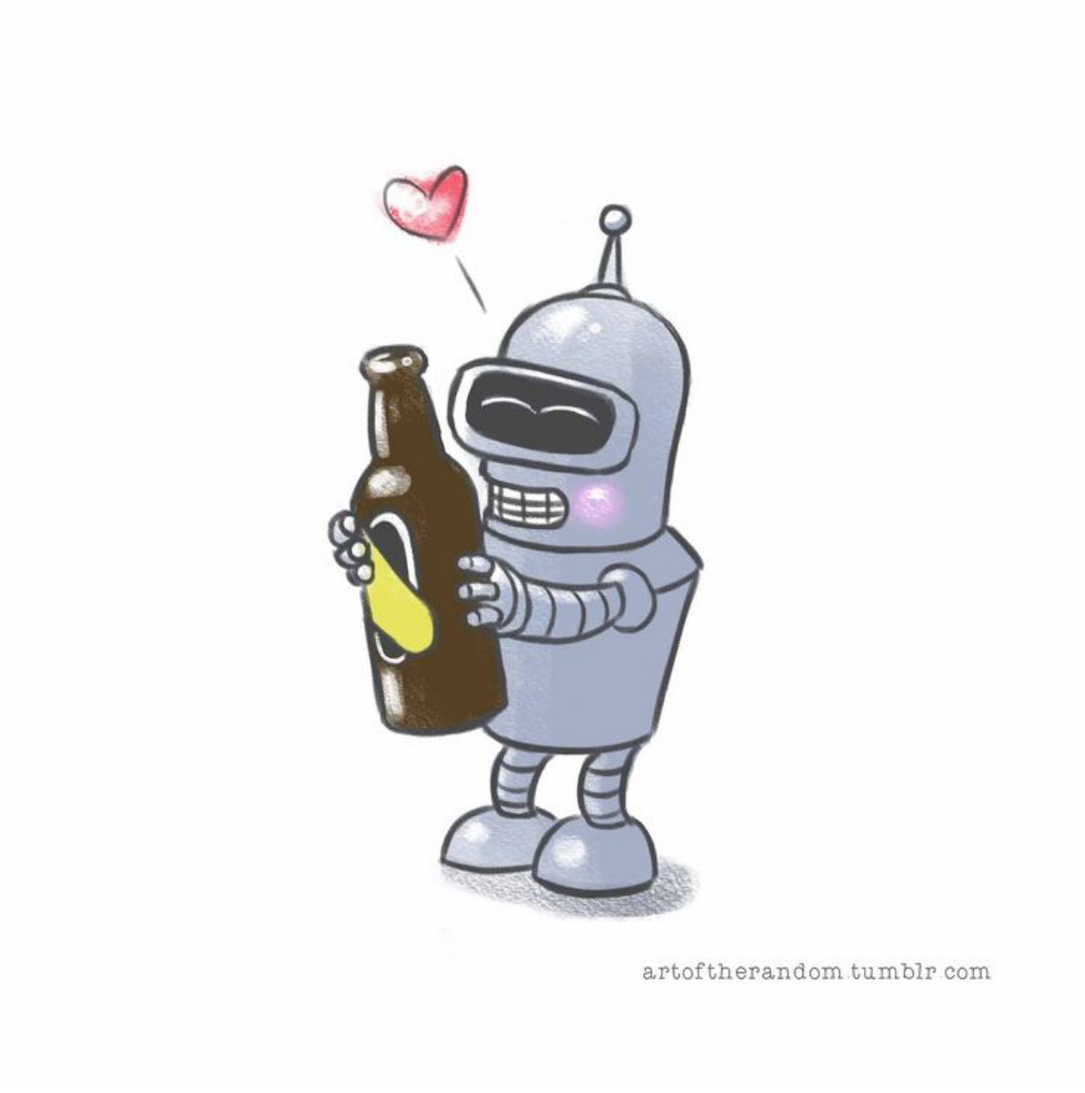
$$\sum_{i \geq 2} \Delta_i^{-2} \log\log \Delta_i^{-2} \quad \text{vs.} \quad \sum_{i \geq 2} \Delta_i^{-1} \log\log \Delta_i^{-1}$$



	$\sum_{i=2}^{n} T_i$
non-adaptive	n^3
SG bound	n^2
KL bound	$n \log n$

Tanczos, Mankoff, N (2017)

Thanks!



lecture notes online:

http://nowak.ece.wisc.edu/intro2ActiveML.pdf