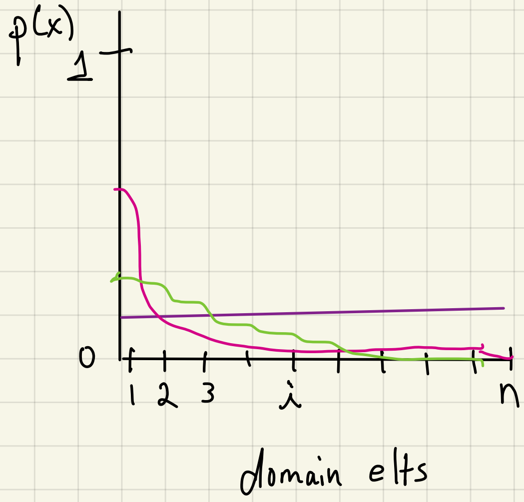
Lecture 15

Learning & Testing Distributions:

Monotonicity

Monotone distributions (over totally ordered domain)

Def. p over domain $[n]$ is "monotone decreasing" if $\forall i \in [n-1]$ $p(i) \geq p(i+1)$



Monotonicity tester:

- if p monotone decreasing, output PASS
- if p ε-far in L_1 from any mon dec dist q, output FAIL

with probability $\geq 1-\delta$

Useful Tool: Birge Decomposition + Flattening

← different than decomposition in pset OBLIVIOUS!

Given ε, decompose domain $D = 1..n$ into $\ell = \Theta\left(\frac{\log n}{\varepsilon}\right)$ intervals

$$I_1^{\varepsilon}, I_2^{\varepsilon}, \ldots, I_\ell^{\varepsilon} \quad \text{s.t.}$$

$$|I_{k+1}^{\varepsilon}| = \lfloor (1+\varepsilon)^k \rfloor$$

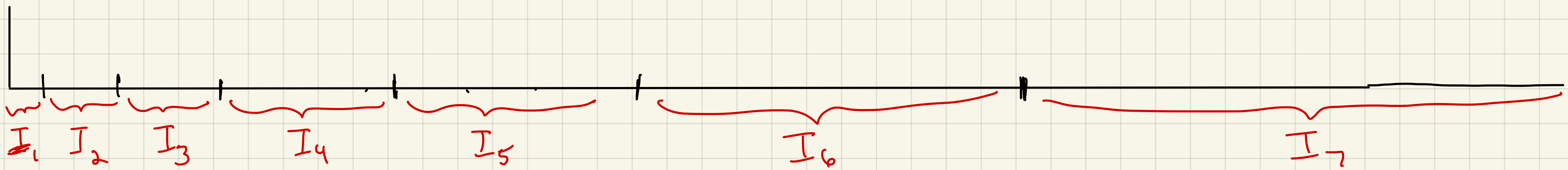
← will drop ε in notation since ε is fixed by algorithm

Note that $|I_1^{\varepsilon}| = |I_2^{\varepsilon}| = \ldots = 1$ ← $\Theta\left(\frac{1}{\varepsilon}\right)$ intervals

$|I_a^{\varepsilon}| = |I_{a+1}^{\varepsilon}| = \ldots = 2$

$\vdots$

but then at some point the exponential "takes off"



Birge Decomposition

$|I_{k+1}| = \lfloor (1+\varepsilon)^k \rfloor$

Def. "flattened distribution":

$\forall$ intervals $1 \le j \le \ell$, $\forall\, i \in I_j$

$$\tilde{q}(i) = \frac{q(I_j)}{|I_j|}$$

$\leftarrow$ total weight of interval

$\leftarrow$ # elts in interval

all elts in interval assigned same weight

Note: $q(I_j) = \tilde{q}(I_j)$

Birge's Thm: If q is monotone decreasing then $\|\tilde{q} - q\|_1 < \varepsilon$

Corr: " " " ε-close to " " " $\|\tilde{q} - q\|_1$ is $O(\varepsilon)$

Birge's Thm: If q is (ε-close to) monotone decreasing then $\|\tilde{q} - q\|_1 < O(\varepsilon)$

Birge Flattening

$$|I_{k+1}| = \lfloor (1+\varepsilon)^k \rfloor$$

$$\forall\ i \in I_j,\ \tilde{q}(i) = \frac{q(I_j)}{|I_j|}$$

Testing algorithm:

how many?

- Take m samples S' of q.

parameter $\frac{\varepsilon}{c}$!

- For each Birge partition I_j:

$$S_j \leftarrow S' \cap I_j$$

estimates of $q(I_j)$

$$\hat{w}_j \leftarrow \frac{n_j}{m}$$

- Define q^*: $\forall\ i \in I_j,\ q^*(i) = \dfrac{\hat{w}_j}{|I_j|}$

no new samples needed

this is LP in $O(\log n)$ vars

- Use LP on $\hat{w}_j$'s to verify that q^* is $\frac{\varepsilon}{c}$-close to monotone
  if no, Fail + halt

- Test that L_1-dist of q + q^* is $< \frac{\varepsilon}{c}$
  if no, Fail + halt
  else accept

q + q^* are only close

how will we pass "good" q?

Birge's Thm: If q is ε-close to monotone decreasing then $\|\tilde{q} - q\|_1 < O(\varepsilon)$

Birge Flattening

$$|I_{k+1}| = \lfloor (1+\varepsilon)^k \rfloor$$

$$\forall\, i \in I_j,\ \tilde{q}(i) = \frac{q(I_j)}{|I_j|}$$

Correctness (high level)

- If q monotone then $\tilde{q}$ is monotone
  $\left(+\ \|q - \tilde{q}\|_1 < \frac{\varepsilon}{c}\right)$
- Since $\hat{w}_j$'s are close to $q(I_j)$ $\leftarrow$ via Chernoff argument
  $\|\tilde{q} - q^*\|_1 < \frac{\varepsilon}{c}$
- so q^* is $\frac{\varepsilon}{c}$-close to monotone.
- $\|q - q^*\|_1 < 2 \cdot \frac{\varepsilon}{c}$ by $\Delta \neq$

Testing algorithm:

- Take m samples S' of q.
- For each Birge partition I_j:
  $S_j \leftarrow S' \cap I_j$
  $\hat{w}_j \leftarrow \frac{|S_j|}{m}$
- Define q^*: $\forall\, i \in I_j,\ q^*(i) = \frac{\hat{w}_j}{|I_j|}$
- verify that q^* is $\frac{\varepsilon}{c}$ close to monotone (no samples)
- Test that L_1-dist of q & q^* is $< \frac{\varepsilon}{c}$

difficulty:
we can distinguish $\|q - q^*\|_1 = 0$ from $\|q - q^*\|_1 > \varepsilon$ in $O(\sqrt{n})$ samples

but we ~~don't~~ can't in general distinguish $\|q - q^*\|_1 < \frac{\varepsilon}{c'}$ from $\|q - q^*\|_1 > \varepsilon$ in $O(\sqrt{n})$ samples

Luckily: this is a special case since we know q is monotone!

Birge's Thm: If q is (ε-close to) monotone decreasing then $\|\tilde{q} - q\|_1 < O(\varepsilon)$

Birge Flattening

$|I_{k+1}| = \lfloor (1+\varepsilon)^k \rfloor$

$\forall\, i \in I_j,\ \tilde{q}(i) = \frac{q(I_j)}{|I_j|}$

Correctness (high level) to show: q ε-far from monotone $\Rightarrow$ tester fails whp

show contrapositive: tester passes whp $\Rightarrow$ q ε-close to monotone

- Tester passes $\Rightarrow$ q^* $\frac{\varepsilon}{c}$-close to monotone
- Tester passes $\Rightarrow$ $\|q - q^*\|_1 < \frac{\varepsilon}{c}$
  $\Rightarrow$ q is $\frac{2\varepsilon}{c}$-close to monotone

Testing algorithm:

- Take m samples S' of q.
- For each Birge partition I_j:
  $S_j \leftarrow S' \cap I_j$
  $\hat{w}_j \leftarrow \frac{|S_j|}{m}$
- Define q^*: $\forall\, i \in I_j,\ q^*(i) = \frac{\hat{w}_j}{|I_j|}$
- verify that q^* is $\frac{\varepsilon}{c}$ close to monotone (no samples)
- Test that L_1-dist of q & q^* is $< \frac{\varepsilon}{c}$

Birge's Thm: If q is [ε-close to] monotone decreasing then $\|\tilde{q} - q\|_1 < O(\varepsilon)$

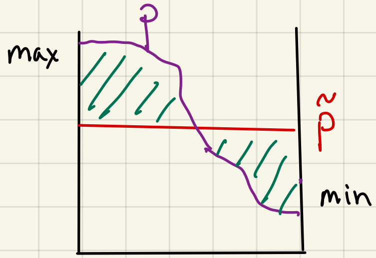
Birge Flattening

$|I_{k+1}| = \lfloor (1+\varepsilon)^k \rfloor$

$\forall\, i \in I_j,\ \tilde{q}(i) = \frac{q(I_j)}{|I_j|}$

Proof of Birge's Thm

error in partition:



gross upper bnd on error:

$\leq$ (max - min) $\cdot$ partition length

Type of Intervals:

- Size 1 intervals $|I_j| = 1$ no error on these ← if have any short intervals, there are $\geq 1/\varepsilon$ size 1 intervals
- Short intervals $|I_j| < 1/\varepsilon$
- Long intervals $|I_j| \geq 1/\varepsilon$

if we have these then max prob $\leq \varepsilon$ } why?

- # size 1 intervals is $\geq \frac{1}{\varepsilon}$
- last (min) size 1 interval has prob $\leq \varepsilon$

why? if last size 1 interval has prob α, then all previous size 1 intervals have prob $> \alpha$

$\Rightarrow$ total wt $\frac{1}{\varepsilon} \cdot \alpha < 1$

$\Rightarrow \alpha < \varepsilon$

$$\text{Total error} \leq \sum_{j=1}^{\ell} |I_j| \left(\text{max prob in } I_j - \text{min prob in } I_j\right)$$

$$= \sum_{\substack{\text{size 1}\\ \text{intervals}}} 1 \cdot 0 + \underbrace{\sum_{\substack{\text{short}\\ \text{intervals}}} |I_j| \cdot (\text{max} - \text{min})}_{\substack{\text{omitted but similar}\\ \text{to long intervals (need to group intervals}\\ \text{of same size)}}} + \sum_{\substack{\text{long}\\ \text{intervals}}} |I_j| (\text{max} - \text{min})$$

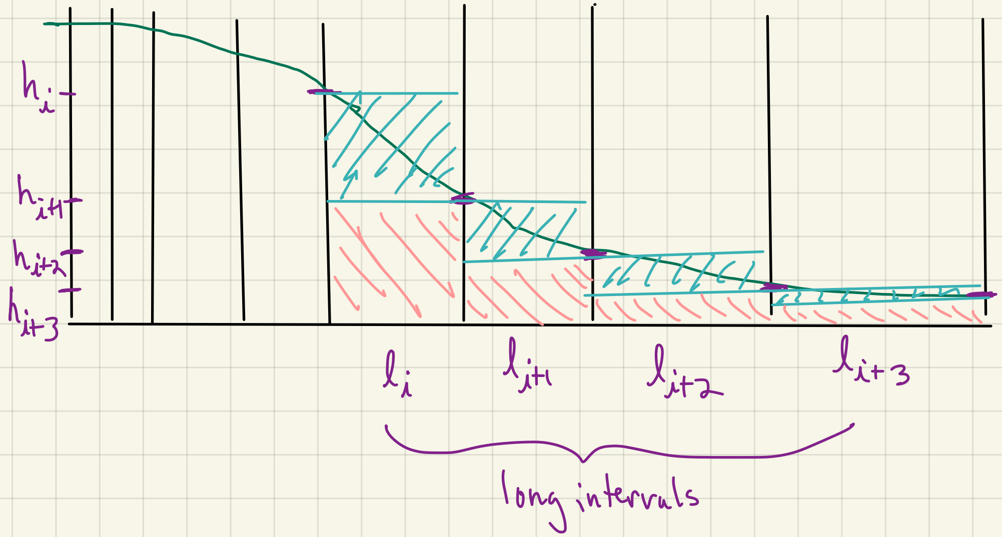
Birge Flattening

$$|I_{k+1}| = \lfloor (1+\varepsilon)^k \rfloor$$

$$\forall\, i \in I_j,\ \tilde{q}(i) = \frac{q(I_j)}{|I_j|}$$

Bounding $\sum_{\substack{\text{long} \\ \text{intervals}}} |I_j| (\max - \min)$:

green rectangles = upper bnd on error

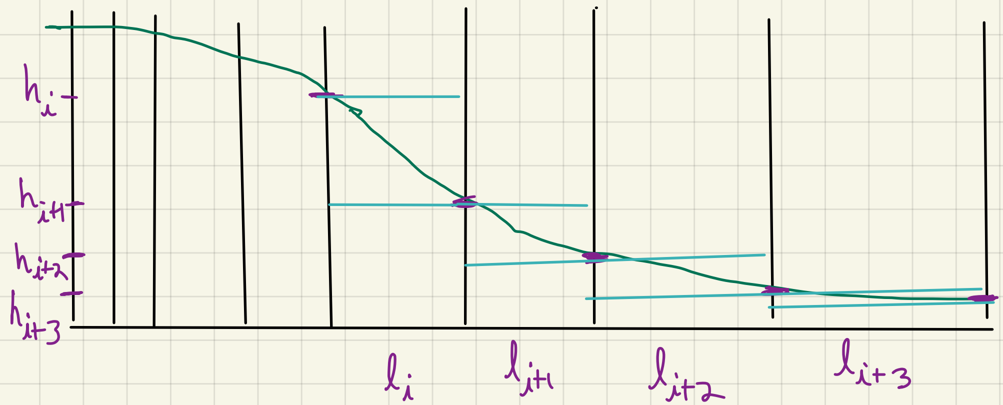


$$\text{error} \le (h_i - h_{i+1})\, l_i + (h_{i+1} - h_{i+2})\, l_{i+1} + (h_{i+3} - h_{i+2})\, l_{i+2} + \dots$$

$$\le \underbrace{h_i l_i}_{\text{all } h_j\text{'s in this area are } < \varepsilon!} + h_{i+1} \underbrace{(l_{i+1} - l_i)} + h_{i+2} \underbrace{(l_{i+2} - l_{i+1})}_{\text{positive } \& \approx \varepsilon \cdot l_{i+1} \text{ by the way we partitioned!}} + \dots$$

$$\le \varepsilon \big[\, l_i + \underbrace{\textstyle\sum h_i l_{i-1}}_{\text{area of red rectangles which is upper bnded by } p\text{, so sum} \le 1} \big]$$

get rid of when bounding short intervals


h_i
h_{i+1}
h_{i+2}
h_{i+3}
l_i
l_{i+1}
l_{i+2}
l_{i+3}

Slight change of perspective:

if we know q is monotone, can we learn it?

Yes! use sampling to estimate $\tilde{q}(I_j)$'s

Birgé's Thm $\Rightarrow$ Can learn monotone distributions

to w/in $\leq \varepsilon$ L_1-error

in $\Theta\left(\frac{1}{\varepsilon^2}\log n\right)$ samples.

Testing algorithm:

- Take m samples S' of q.
- For each Birge partition I_j:

$$S_j \leftarrow S' \cap I_j$$

$$n_j \leftarrow |S'_j| \quad \& \quad \hat{w}_j \leftarrow \frac{|S_j|}{m}$$

- Define q^*: $\forall\, i \in I_j,\ q^*(i) = \frac{\hat{w}_j}{|I_j|}$
- verify that q^* is $\frac{\varepsilon}{c}$ close to monotone
- Test that L_1-dist of q & q^* is $< \frac{\varepsilon}{c}$