Lecture 15

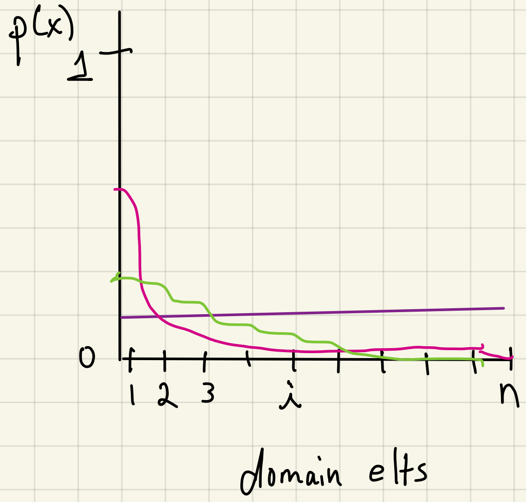
Learning & Testing Distributions:

Monotonicity

Monotone distributions (over totally ordered domain)

total order

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



Monotonicity tester:

with prob $\geq 1-\delta$

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

h.w. l.b. $\Omega(\sqrt{n})$ samples

Useful Tool: Birgé Decomposition + Flattening

← different than decomposition in H.W

OBLIVIOUS

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

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

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

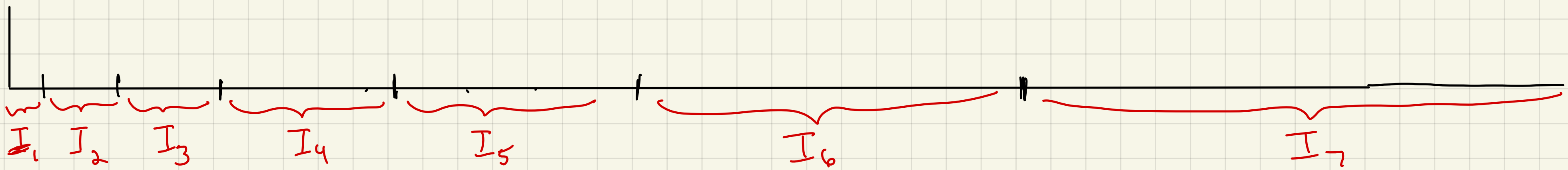
← will drop ε from notation since ε is fixed by algorithm

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

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

$\vdots$

but then at some point the exponential "takes off"



Birge Decomposition

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

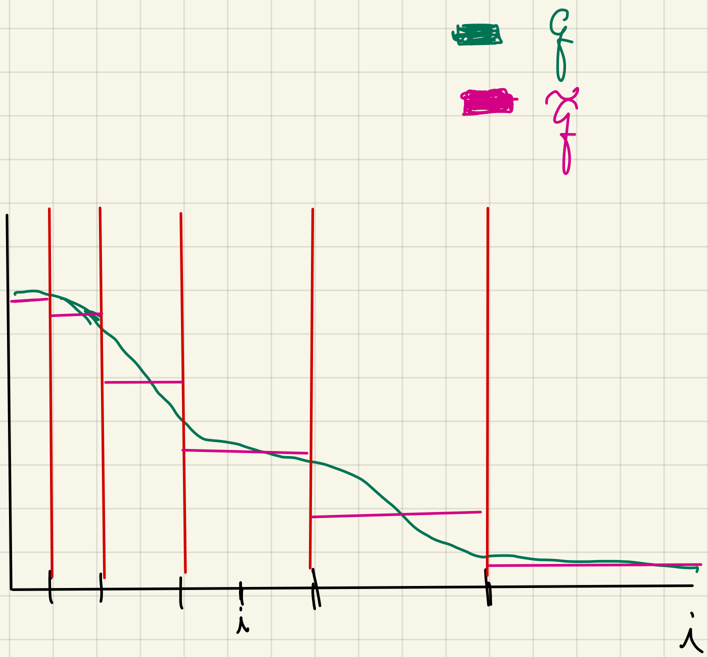
Def. "flattened distribution": given q

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

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

$\leftarrow$ total wt of interval

$\leftarrow$ # of domain elts in interval



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

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

Corr: " " " ε-close to " " " "

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:

- Take m samples S' of q. how many?
- For each Birge partition I_j: parameter $\frac{\varepsilon}{c}$!

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

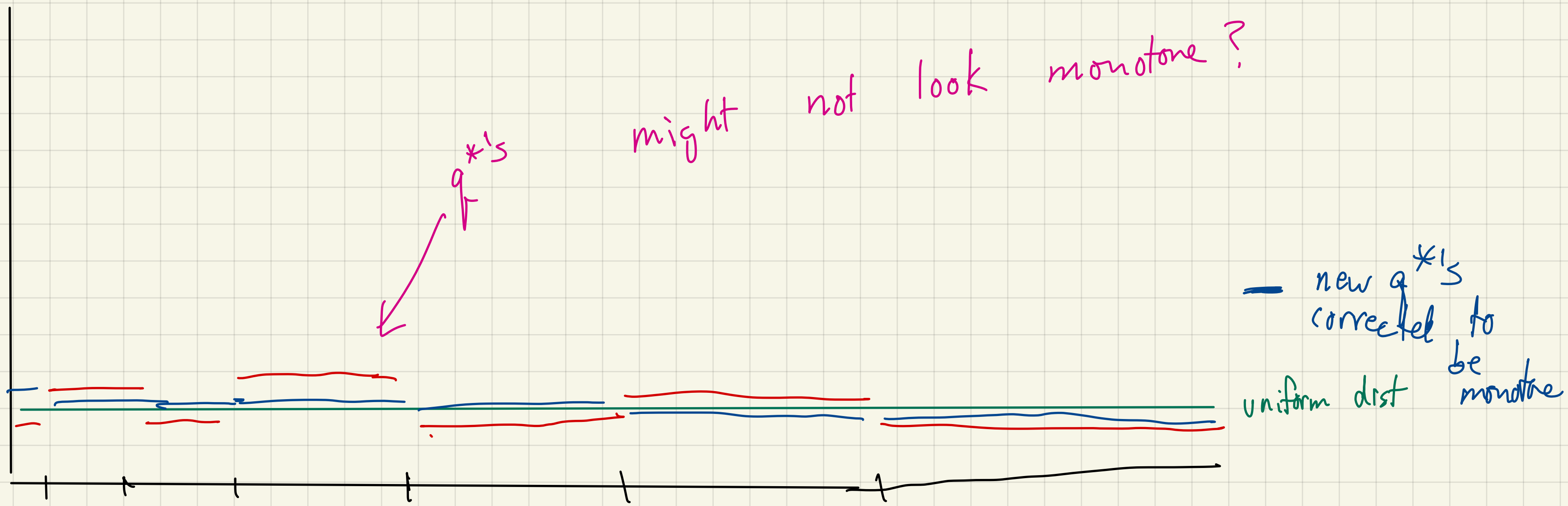
$$\hat{w}_j \leftarrow \frac{|S_j|}{m}$$ estimate of $q(I_j)$

- Define q^*: $\forall\ i \in I_j,\ q^*(i) = \frac{\hat{w}_j}{|I_j|}$
- Use LP on $\hat{w}_j$'s to verify that q^* is $\frac{\varepsilon}{c}$ close to monotone
  if no, Fail + halt

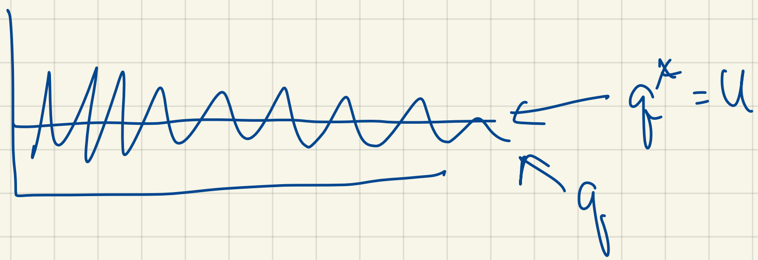
  no new samples needed
  this is LP in $O(\log n)$ vars

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

even if q monotone q + q^* are only close
how do we pass all "good" (monotone) q?
previous algorithms are not tolerant



Another issue: what if q not 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) (q monotone $\Rightarrow$ ~~tester~~ tester passes whp)

- if q monotone then $\tilde{q}$ monotone

  + Birgé $\Rightarrow \|q - \tilde{q}\|_1 < \frac{\varepsilon}{c'}$

- Since $\hat{w}_j$'s are close to $q(I_j)$ $\leftarrow$ via Chernoff bnd argument

  $\Rightarrow \|\tilde{q} - q^*\|_1 < \frac{\varepsilon}{c'}$

- 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 $q = q^*$ from $\|q - q^*\|_1 > \varepsilon$ in $O(\sqrt{n})$ samples

here we need to distinguish $\|q - q^*\|_1 < \varepsilon'$ from $\|q - q^*\|_1 > \varepsilon$ (in $O(\sqrt{n}$ | samples) ?

if q arbitrary, not possible. But q is monotone so we can do it.

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

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

equivalent

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

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

$\Rightarrow q$ is $\frac{2\varepsilon}{c}$-close to monotone via $\Delta \neq$ $\blacksquare$

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:

- 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 $\|\tilde{q} - q\|_1 < O(\varepsilon)$

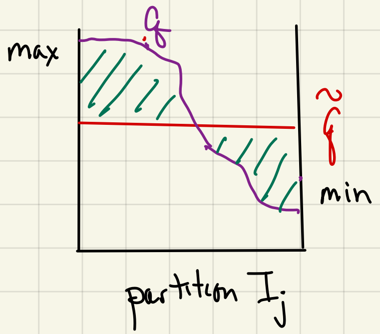
Birge Flattening

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

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

Proof of Birge's Thm

error in partition:



gross upper bnd on error:

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

Type of Intervals:

- Size 1 intervals $|I_j| = 1$ — no error on these ← if have any short intervals then there are $\ge \frac{1}{\varepsilon}$ size 1 intervals
- Short intervals $|I_j| < 1/\varepsilon$ — if have any of these, max prob $\le \varepsilon$ — why?
- Long intervals $|I_j| \ge 1/\varepsilon$

- # size 1 intervals $\ge \frac{1}{\varepsilon}$ (by partitioning)
- last size 1 interval has prob $\le \varepsilon$ (min wt)

why? if last size 1 interval has wt $> \varepsilon$ then all previous size intervals have wt $> \varepsilon$

$\Rightarrow$ total wt of size 1 intervals $> \frac{1}{\varepsilon} \cdot \varepsilon > 1$

contradiction

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

$$= \sum_{\substack{\text{size 1} \\ \text{intervals}}} 1 \cdot 0 + \sum_{\substack{\text{short} \\ \text{intervals}}} |I_j| (\max - \min) + \underbrace{\sum_{\substack{\text{long} \\ \text{intervals}}} |I_j| (\max - \min)}_{\text{will bound now}}$$

omitted but similar to long (need to group same size short intervals)

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

green rectangles upper bnd error

$$l_i = \lfloor (1+\varepsilon)^i \rfloor$$

$$(1+\varepsilon)^{i+1} - (1+\varepsilon)^i$$

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

$$\le h_i l_i + h_{i+1}(l_{i+1} - l_i) + h_{i+2} \underbrace{(l_{i+2} - l_{i+1})}_{\approx\, \varepsilon \cdot l_{i+1}} + \ldots$$

all h_i's $\le \varepsilon$

$\approx \varepsilon \cdot l_{i+1}$ by the way we partitioned

$$\le \varepsilon \left[l_i + \underbrace{\sum h_i l_{i-1}}_{} \right]$$

get rid of when bound short intervals.

area of red rectangles which is upper bounded by q so sum ≤ 1

[Daskalakis Diakonikolas Servedio] + [Daskalakis et al]

Slight change of perspective:

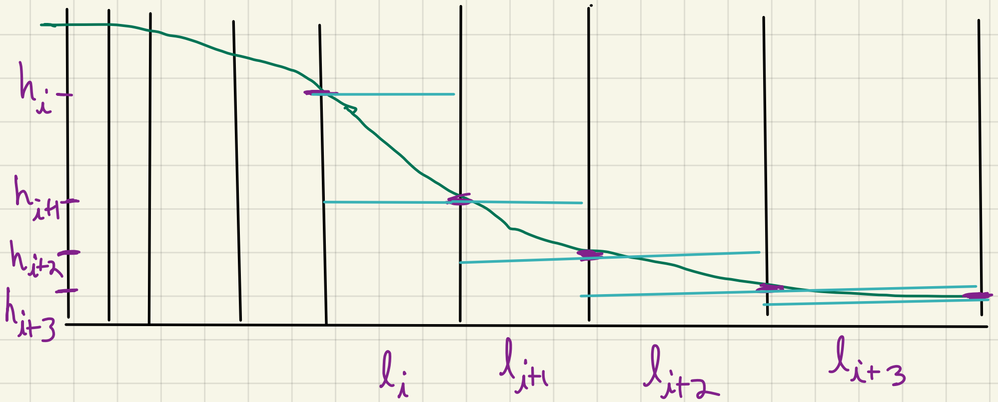
if we know q is monotone, can we learn it?

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

Birgé's thm $\Rightarrow$ can learn monotone distributions

to w/in ε L_1-error

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


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

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 $< \frac{\varepsilon}{c}$