Lecture 15

Learning & Testing Distributions

Monotonicity
Monotone distributions (over totally ordered domain)

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

Monotonicity tester:

- if \( p \) monotone decreasing, output PASS
- if \( p \) \( \epsilon \) far in \( L_1 \) from any mon dec dist \( q \), output FAIL

h.w. \( i.i.d. \ L(\sqrt{n}) \) samples
Useful Tool: Birge Decomposition & Flattening

Given $\epsilon$, decompose domain $D = 1..n$ into $l = \Theta(\frac{\log n}{\epsilon})$ intervals

$I_1^\epsilon$, $I_2^\epsilon$, ..., $I_l^\epsilon$ st.

$|I_k^\epsilon| = L(1+\epsilon)^k$

will drop $\epsilon$ from notation

since $\epsilon$ is fixed by algorithm

Note that

$|I_1^\epsilon| = |I_2^\epsilon| = ... = 1 \quad \Theta(\frac{1}{\epsilon})$ intervals

$|I_a^\epsilon| = |I_{a+1}^\epsilon| = ... = 2$

but then at some point the exponential "takes off"
Def. "flattened distribution": given \( q \),

\[
\forall \text{ intervals } 1 \leq j \leq l, \quad \forall i \in I_j
\]

\[
\tilde{q}(i) = \frac{q(I_j)}{|I_j|} \quad \text{total \ wt \ of \ interval} \quad \# \text{ of \ domain \ elts \ in \ interval}
\]

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

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

Corr: \( \varepsilon \)-close to
Birge’s Thm: \( \| \tilde{q} - q \|_1 \leq O(\varepsilon) \)

Testing algorithm:

- Take \( m \) samples \( S \) of \( q \).
- For each Birge partition \( I_j \):
  \( S_j \leftarrow S \cap I_j \)
  \( w_j \leftarrow \frac{|S_j|}{m} \)
  \( \hat{q}_j \leftarrow \text{estimate of } q(I_j) \)
- Define \( q^* \): \( \forall i \in I_j \), \( q^*(i) = \frac{\hat{q}_j}{|I_j|} \)
- Use LP on \( \hat{w}_j \)’s to verify that \( q^* \) is \( \varepsilon \)-close to monotone
  - if no, Fail + halt
- Test that \( L_1 \)-dist of \( q + q^* \) is \( \leq \frac{\varepsilon}{2} \)
  - if no, Fail + halt
  - else accept

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

Birge Flattening

\( I_{\text{init}} = \{(0, \varepsilon)^k\} \)
Another issue: what if $q$ not monotone?

Another issue: what if $q^*$ not monotone?

New $q^*$s corrected to be uniform dist monotone.
Birge's Thm: If \( q \) is \( \varepsilon \)-close to monotone decreasing then \( \| \hat{q} - q \|_1 < O(\varepsilon) \)

Correctness (high level) (\( q \) monotone \( \Rightarrow \) test passes whp)

- If \( q \) monotone then \( \hat{q} \) monotone
- Birge \( \Rightarrow \) \( \| \hat{q} - q \|_1 < \frac{\varepsilon}{2} \)
- Since \( \hat{w}_j \)'s are close to \( q(I_j) \) \( \Rightarrow \) \( \| \hat{q} - q^* \|_1 < \frac{\varepsilon}{2} \)
- So \( q^* \) is \( \frac{\varepsilon}{C} \)-close to monotone
- \( \| q - q^* \|_1 < 2 \cdot \frac{\varepsilon}{C} \), by \( \Delta^t \)

**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 \( O(n) \) samples)

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

Birge Flattening

\[ \| I_{k+1} \| = \left( \frac{1}{1+\varepsilon} \right)^k \]

\[ \forall i \in I_{j}, \hat{q}(i) = \frac{q(I_j)}{|I_j|} \]

Testing algorithm:

- Take \( m \) samples \( S' \) of \( q \).
- For each Birge partition \( I_j \):
  - \( S_j = S' \cap I_j \)
  - \( \hat{w}_j = \frac{|S_j|}{m} \)
- Define \( \hat{q}^* \) \( \forall i \in I_j, \hat{q}^*(i) = \frac{\hat{w}_j}{|I_j|} \)
- verify that \( \hat{q}^* \) is \( \varepsilon \)-close to monotone (no sample)
- Test that \( L_1 \)-dist of \( q, q^* \) is \( < \frac{\varepsilon}{C} \)
Birge's Thm: If $q$ is $(\varepsilon, \text{close})$-monotone decreasing then $\| \tilde{q} - q \|_1 < O(\varepsilon)$

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

Show contrapositive: tester passes whp $\Rightarrow q$ is $\varepsilon$-close to monotone

Testing algorithm:

1. Take $m$ samples $S'$ of $q$.
2. For each Birge partition $I_j$:
   - $S_j = S' \cap I_j$
   - $\hat{w}_j = |S_j| / m$
3. Define $\tilde{q}^*$:
   $\tilde{q}^*(i) = \hat{w}_j / |I_j|$
4. Verify that $\tilde{q}^*$ is $\varepsilon$-close to monotone (no sample)
5. Test that $L_1$-dist of $q$ and $q^*$ is $\frac{\varepsilon}{2}$
Birge's Thm: If \( \tilde{q} \) is (\( \varepsilon \)-close) monotone decreasing, then \( \| \tilde{q} - \hat{q} \| < O(\varepsilon) \)

Proof of Birge's Thm

error in partition:

![Diagram of error in partition]

gross upper bound on error:

\[ \leq (\text{max} - \text{min}) \cdot \text{partition length} \]

Type of Intervals:

- Size 1 intervals: \( |I_j| = 1 \) no error on these \( \leftarrow \) if have any short intervals then there are \( \geq \frac{1}{\varepsilon} \) size 1 intervals
- Short intervals: \( |I_j| < \frac{1}{\varepsilon} \) (why?)
- Long intervals: \( |I_j| \geq \frac{1}{\varepsilon} \)

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

\[ = \sum_{\text{size intervals}} |I_j| \cdot (\text{max-min}) + \sum_{\text{long intervals}} |I_j| \cdot (\text{max-min}) \]

\[ \geq \frac{1}{3} \cdot \varepsilon \cdot \varepsilon > 1 \text{ (contradiction)} \]
Bounding \( \sum |I_j| \) (max-min) of long intervals:

- Green rectangles: upper bound error

Error \( \leq (h_i - h_{i+1}) l_i + (h_{i+1} - h_{i+2}) l_{i+1} + (h_{i+2} - h_{i+3}) l_{i+2} + \ldots \)

\( \leq h_i l_i + h_{i+1} (l_i + l_{i+1}) + h_{i+2} (l_{i+1} + l_{i+2}) + \ldots \)

All \( h_i \)'s \( \leq 3 \)

Get rid of short intervals.

Area of red rectangles, which is upper bounded by 9.

By the way, we partitioned

[Dasgulakos Oranikolos Servedau] \& [Dasgulakos et al]
Slight change of perspective:

if we know \( q \) is monotone, can we learn it?

Yes! Use sampling to estimate \( \hat{q}(I_j) \)’s.

Birge’s then \( \Rightarrow \) can learn monotone distributions

to within \( \varepsilon \) \( L_1 \) error

in \( O \left( \frac{1}{\varepsilon^2 \log n} \right) \) samples.
Testing algorithm:

- Take \( m \) samples \( S' \) of \( q \).

- For each Birge partition \( I_j \):
  
  \[ S_j = S' \cap I_j; \quad n_j = |S_j|; \quad \hat{w}_j = \frac{|S_j|}{m} \]

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

- Verify that \( q^* \) is \( \varepsilon \)-close to monotone.

- Test that \( L_1 \)-dist of \( q + q^* \) is \( < \frac{3\varepsilon}{2} \).