Lecture 18

In the previous lecture, we showed that for $Z = \mathbb{1}_A$,

$$\boxed{\mathrm{Var}(Z)\log\frac{\mathrm{Var}(Z)}{(2p(1-p))^2\,\delta_p(A)\,I_p(A)} \le p(1-p)\,c(p)\,I_p(A),} \qquad (\#)$$

where $\delta_p(A) = \max_{1\le i\le n} I_p^i(A)$.

Proof was almost complete; only the following claim remains to be proved.

<u>Claim</u>: For $Z_j = \mathbb{E}[Z \mid X^j]$ and $\Delta_j = Z_j - Z_{j-1}$,

$$\mathcal{E}(\Delta_j) = \mathcal{E}(Z_j) - \mathcal{E}(Z_{j-1}).$$

<u>Proof</u>. $$\frac{\mathcal{E}(\Delta_j)}{p(1-p)} = \mathbb{E}\left[\left(\Delta_j(X) - \Delta_j(\bar{X}^{(i)})\right)^2\right]$$

Note that

$$\Delta_j(X) - \Delta_j(\bar{X}^{(i)}) = \mathbb{E}_j\Big[\overbrace{f(X) - f(\bar{X}^{(i)})}^{\text{denote by } Y}\Big] - \mathbb{E}_{j-1}\left[f(X) - f(\bar{X}^{(i)})\right] \equiv \mathbb{E}_j[Y] - \mathbb{E}_{j-1}[Y]$$

which is uncorrelated with $\mathbb{E}_{j-1}\left[f(X) - f(\bar{X}^{(i)})\right] \equiv \mathbb{E}_{j-1}[Y]$

Thus, $\mathbb{E}\left[(\Delta_j(X) - \Delta_j(\bar{X}^{(i)}))^2\right]$

$$= \mathbb{E}\left[\left(\mathbb{E}_j[Y] - \mathbb{E}_{j-1}[Y]\right)^2\right]$$

$$= \mathbb{E}\left[\mathbb{E}_j[Y]^2\right] + \mathbb{E}\left[\mathbb{E}_{j-1}[Y]^2\right] - 2\,\overbrace{\mathbb{E}\left[\mathbb{E}_j[Y]\,\mathbb{E}_{j-1}[Y]\right]}^{= \mathbb{E}\left[\mathbb{E}_{j-1}[Y]^2\right]}$$

$$= \mathbb{E}\left[\mathbb{E}_j[Y]^2\right] - \mathbb{E}\left[\mathbb{E}_{j-1}[Y]^2\right]$$

$$= \left(\mathcal{E}(Z_j) - \mathcal{E}(Z_{j-1})\right) / p(1-p). \qquad \blacksquare$$

D Threshold Phenomenon for monotone sets

Theorem Let $A \subseteq \{-1,1\}^n$ be a symmetric monotone set, i.e., a monotone set s.t. $I_p^i(A) = I_p^j(A)$ $\forall i,j$.

Then,

$$\boxed{p_{1-\epsilon} - p_\epsilon \leq \frac{8 \log 1/2\epsilon}{\log \frac{n}{16}}}$$

Some examples:

1 Majority: Let $A = \{x : \sum_{i=1}^{n} x_i > 0\}$.

Thus, $\mathbb{P}_p(A) = \mathbb{P}\left(\sum_{i=1}^{n} x_i - (2p-1)n > (1-2p)n\right)$

$$\leq e^{-2n(1-2p)^2} \quad \text{(by Hoeffding's)}$$

Therefore, $\mathbb{P}_p(A) \leq \epsilon$ if $p \leq \frac{1}{2} - \sqrt{\frac{\log(1/\epsilon)}{8n}}$

and similarly,

$$1 - \mathbb{P}_p(A) = \mathbb{P}_{1-p}(A^c) \leq \epsilon \text{ if } 1-p \leq \frac{1}{2} - \sqrt{\frac{\log(1/\epsilon)}{8n}}.$$

which gives

$$p_\epsilon - p_{1-\epsilon} \leq \sqrt{\frac{\log(1/\epsilon)}{2n}}.$$

Thus, we can do better than the width promised by the theorem above.

2 Dictatorship: Let $A = \{x : x_1 = 1\}$.

Then, $P_p(A) = p \Rightarrow p_\epsilon - p_{1-\epsilon} = 1-2\epsilon$. But $I_p^1(A) = p$ and $I_p^i(A) = 0$ $\forall i > 1$. $\Rightarrow$ Symmetry is essential.

A useful corollary of Russo's Lemma

For a symmetric, monotone set A

$$(\#\#)\quad \delta_p(A) \le \sqrt{\frac{\mathbb{P}_p(A)}{n\,p(1-p)}} \le \frac{1}{\sqrt{n\,p(1-p)}}.$$

Proof.
$$I^p(A) = \frac{d}{dp}\,\mathbb{P}_p(A)$$
$$= \frac{d}{dp}\sum_{x\in A} p^{\|x\|}(1-p)^{n-\|x\|}$$
$$= \sum_{x\in A} \|x\|\, p^{\|x\|-1}(1-p)^{n-\|x\|} - (n-\|x\|)\, p^{\|x\|}(1-p)^{n-\|x\|-1}$$
$$= \frac{1}{p(1-p)}\,\mathbb{E}\left[(\|x\|-np)\,\mathbb{1}_A\right]$$
$$\le \frac{1}{p(1-p)}\sqrt{\mathrm{Var}(\|X\|)\,\mathbb{P}_p(A)}$$
$$= \sqrt{\frac{n\,\mathbb{P}_p(A)}{p(1-p)}}.$$

By symmetry,
$$\delta_p(A) = \frac{I_p(A)}{n} \le \sqrt{\frac{\mathbb{P}_p(A)}{n\,p(1-p)}}. \qquad \blacksquare$$

Proof of Theorem We first need to simplify $(\#)$.

Claim: $I_p(A) \ge \dfrac{\mathrm{Var}(Z)}{2p(1-p)c(p)} \log \dfrac{c(p)}{4p(1-p)\delta_p}$

Pf. For $y = \dfrac{\mathrm{Var}(Z)}{(2p(1-p))^2\,\delta_p\, I}$ and $z = \dfrac{c(p)}{4p(1-p)\delta_p}$,

(#) gives $z \geq y \log y$.

But then $\boxed{\frac{2z}{\log z} \geq y}$ since otherwise

$$y \log y > \frac{2z}{\log z} \log \frac{2z}{\log z} > z,$$ where the last

inequality hold iff $\frac{\log z}{2} > \log \frac{\log z}{2}$, which is always true. ∎

Now, proceeding using Russo's lemma, denoting $f(p) = \mathbb{P}_p(A)$, we get

$$\frac{d f(p)}{dp} \geq \frac{f(p)(1-f(p))}{2p(1-p)c(p)} \cdot \log \frac{c(p)}{4p(1-p)\delta_p}$$

$$\geq \frac{f(p)(1-f(p))}{2p(1-p)c(p)} \log \frac{c(p)}{4} \sqrt{\frac{n}{p(1-p)}}$$

(by (##))

$\boxed{\text{Claim: } 4\sqrt{2p(1-p)} \leq c(p) \leq \frac{1}{2p(1-p)}}$

$$\geq \frac{f(p)(1-f(p))}{2} \log \frac{n}{16}$$

We now integrate. First consider $p \in [p_\epsilon, p_{1/2}]$. In this range, $(1-f(p)) \geq \frac{1}{2}$. Thus,

$$\frac{f'(p)}{f(p)} \geq \frac{1}{4} \log \frac{n}{16} \Rightarrow \log \frac{1}{2} - \log \epsilon \geq \left(\frac{p_{1/2} - p_\epsilon}{4}\right) \log \frac{n}{16}.$$

Similarly,

$$-\frac{d}{dp}\log(1-f(p)) \geq \frac{1}{4}\log\frac{n}{16} \quad \forall \quad p > p_{1/2}.$$

Thus,

$$\log\frac{1}{2} - \log(1-\epsilon) \geq \frac{(p_{1-\epsilon} - p_{1/2})}{4}\log\frac{n}{16}.$$

which upon combining with the previous bound yields

$$\log\frac{1}{4\epsilon(1-\epsilon)} \geq \frac{p_{1-\epsilon} - p_{\epsilon}}{4}\log\frac{n}{16}.$$

But $\log\frac{1}{4\epsilon(1-\epsilon)} \leq \log\frac{1}{4\epsilon^2}$, which yields the claimed bound. ∎

(To see the claim:

$$c(p) = \frac{\log (1-p)/p}{1-2p} = \frac{\log(1-p) - \log p}{(1-p)-p}$$

$$= \frac{\log(1+2x) - \log(1-2x)}{2x}, \quad x = \frac{1}{2} - p$$

$$= 2\cdot\left(1 + \frac{4x^2}{3} + \frac{16x^4}{5} + \cdots\right) \leq \frac{2}{1-4x^2} = \frac{1}{2p(1-p)}$$

and $c(p) \geq 2$.

Thus, $\frac{c(p)}{\sqrt{p(1-p)}} \geq \frac{4}{\sqrt{c(p)p(1-p)}} \geq 4\sqrt{2}$.)