Lecture 17

<u>Review</u>:
* monotone properties
* Influence $I_p(A)$
* Russo's Lemma: $\boxed{\dfrac{d\,\mathbb{P}_p(A)}{dp} = I_p(A)}$

[A] Proof of Russo's Lemma (contd.)

<u>Proof</u>. Let $U_1, \dots, U_n$ be iid unif $[0,1]$. Given $\underline{p} = (p_1, \dots, p_n)$ and

$\underline{\hat{p}} = (p_1, \dots, p_{i-1}, \hat{p}_i, p_{i+1}, \dots, p_n)$, let

$$X_j = 2\mathbb{1}\{U_j \le p_j\} - 1 \quad \text{and} \quad \hat{X}_j = 2\mathbb{1}\{U_j \le \hat{p}_j\} - 1, \quad 1 \le j \le n.$$

Consider the case $\hat{p}_i > p_i$. Then, since A is monotone,

$X \in A \Rightarrow \hat{X} \in A$.

Thus,
$$\begin{aligned}
&\mathbb{P}(\hat{X} \in A) - \mathbb{P}(X \in A) \\
&= \mathbb{P}(\hat{X} \in A, X \notin A) \\
&= \mathbb{P}\left(U_i \in (p_i, \hat{p}_i), X_i^+ \in A, X_i^- \notin A\right) \\
&= (\hat{p}_i - p_i) . \mathbb{P}(X_i^+ \in A, X_i^- \notin A).
\end{aligned}$$

For $p_j = p$ for every $j \ne i$, we get
$$\mathbb{P}(\hat{X} \in A) - \mathbb{P}(X \in A) = (\hat{p}_i - p_i)\, I_p^i(A)$$

and similarly, when $\hat{p}_i < p_i$,
$$\mathbb{P}(\hat{X} \in A) - \mathbb{P}(X \in A) = -(\hat{p}_i - p_i)\, I_p^i(A).$$

Thus, denoting by $Q_{\underline{p}}(A)$ the prob. $X \in A$ for $\underline{p} = (p, \dots, p)$,

$$\frac{\partial}{\partial p_i} \mathbb{Q}_{\underline{p}}(A) = I_p^i(A)$$

which gives

$$\frac{d\,\mathbb{P}_p(A)}{dp} = \sum_{i=1}^{n} \frac{\partial\, \mathbb{Q}_{\underline{p}}(A)}{\partial p_i} = \sum_{i=1}^{n} I_p^i(A) = I(A)$$

■

$\boxed{B}$ Formalising the threshold phenomenon

We have already noting that for a monotone set $A \neq \phi$ or $\{-1,1\}^n$, $\mathbb{P}_0(A) = 0$, $\mathbb{P}_1(A) = 1$. Also, for such a set A, $I_p(A) > 0$. Indeed, since $(-1,-1,\dots,-1) \in A^c$ and $(1,1,\dots,1) \in A$, $\exists\, i$ and a sequence x s.t. $x_i^- \in A^c$ and $x_i^+ \in A$. Thus, $I_p^i(A) > 0 \Rightarrow I_p(A) > 0$. We have established the following.

Corollary. For every monotone set $A \neq \phi$ or $\{-1,1\}^n$, $f(p) \overset{def}{=} I_p(A)$ is strictly increasing and $f(0)=0$, $f(1)=1$.

For $0 < \epsilon < 1$, let $p_\epsilon \overset{def}{=} \inf\{p : \mathbb{P}_p(A) \geq \epsilon\}$. Then, in view of the corollary above, $p_0 = 0$, $p_1 = 1$. We are interested in bounding $p_\epsilon - p_{1-\epsilon}$ for $\epsilon < 1/2$.

While we can use a version of log-Sobolev inequality for arbitrary p to lower bound $I_p(A)$, this will not do it.

$\boxed{C}$ A strengthening of log-Sobolev inequality

Theorem (General BLSI)

Let $X_1, \dots, X_n$ be iid with $P_{X_i}(1) = 1 - P_{X_i}(-1) = p$.

Then, for every $f : \{-1,1\}^n \to \mathbb{R}$ and $c(p) = \frac{1}{1-2p} \log \frac{1-p}{p}$,

$$\mathrm{Ent}(f^2) \le 2c(p)\,\mathcal{E}(f)$$

where
$$\mathcal{E}(f) = \frac{1}{2}\sum_{i=1}^n \mathbb{E}\left[\left(f(X) - f(X_i')\right)^2\right]$$
$$= p(1-p)\,\underbrace{I_p(f)}_{\hookrightarrow \sum_{i=1}^n \mathbb{E}\left[\left(f(X) - f(\bar{X}^{(i)})\right)^2\right]}$$

Proof. Just as in the proof of BLSI for the uniform case, it suffices to show the inequality for $n=1$. This case reduces to showing

$$pa^2 \log a^2 + (1-p)b^2 \log b^2 - (pa^2 + (1-p)b^2)\log(pa^2 + (1-p)b^2)$$
$$\le p(1-p)\,c(p)\,(|a| - |b|)^2. \quad (HW!)$$ ∎

Check:
$$\mathbb{E}\left[\left(f(X) - f(X_i')\right)^2\right] = 2p(1-p)\,\mathbb{E}\left[\left(f(X_i^+) - f(X_i^-)\right)^2\right]$$
$$\mathbb{E}\left[\left(f(X) - f(\bar{X}^{(i)})\right)^2\right] = \mathbb{E}\left[\left(f(X_i^+) - f(X_i^-)\right)^2\right].$$

A direct proof for the case $f = \mathbb{1}_A$ is simple.

Theorem ("Stregthened" general BLSI and Efron-Stein)

For $Z = \mathbb{1}_A$ and $\delta_p(A) = \max_{1 \le i \le n} I_p^i(A)$,

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

Remark. BLSI: Bounds $P(A)\log\frac{1}{P(A)}$

BESI: Bounds $P(A)(1-P(A))$ $\left(\le P(A)\log\frac{1}{P(A)}\right)$

Since there is no dependence on n in these inequalities, we can't expect them to enable a proof of the threshold phenomenon.

Proof. Step 1: Subadditivity and tensorization

$Z_i = \mathbb{E}[Z|X^i]$, $\Delta_i = Z_i - Z_{i-1}$

$\left(\mathrm{Var}(\Delta_i) = \mathbb{E}\Delta_i^2 \le \mathbb{E}\,\mathrm{Var}^{(i)}(Z)\right)$: ES proof

$$\mathcal{E}(\Delta_j) = \frac{1}{2}\sum_{i=1}^{n}\mathbb{E}\left[(\Delta_j(X)-\Delta_j(X_i'))^2\right]$$
$$= p(1-p)\sum_{i=1}^{n}\mathbb{E}\left[(\Delta_j(X)-\Delta_j(\bar{X}^{(i)}))^2\right]$$

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

Proof: $\Delta_j(X) - \Delta_j(\bar{X}^{(i)})$

$$= \mathbb{E}_j[f(X)-f(\bar{X}^{(i)})] - \mathbb{E}_{j-1}[f(X)-f(\bar{X}^{(i)})]$$

Also, denoting $\bar{Z}_{j,i} = \mathbb{E}_j[f(\bar{X}^{(i)})]$,

$$Z_{j-1} - \bar{Z}_{j-1,i} = \mathbb{E}_{j-1}[f(X) - f(\bar{X}^{(i)})].$$

Since $\mathbb{E}_j[Y] - \mathbb{E}_{j-1}[Y] \perp \mathbb{E}_{j-1}[Y]$ for any rv Y,

$(\Delta_j - \Delta_j(\bar{X}^{(i)}))$ is uncorrelated with $Z_{j-1} - \bar{Z}_{j-1,i}$.

Thus,

$$\mathcal{E}(\Delta_j) = p(1-p)\sum_{i=1}^{n}\mathbb{E}\left[\mathbb{E}_j[(f(X)-f(\bar{X}^{(i)}))]^2 - \mathbb{E}_{j-1}[f(X)-f(\bar{X}^{(i)})]^2\right]$$
$$= \mathcal{E}(Z_j) - \mathcal{E}(Z_{j-1})$$

$$\Rightarrow \mathcal{E}(f) = \sum_{j=1}^{n} \mathcal{E}(\Delta_i).$$

Step 2: log-Sobolev and other inequalities

(a) $\mathcal{E}(f) \geq \mathcal{E}(|f|)$

Pf. $(a-b)^2 \geq (|a|-|b|)^2$

(b) $\mathcal{E}(|f|) \geq \frac{1}{2C(p)} \mathrm{Ent}(f^2)$

(c) For $f : \{-1,1\}^n \to \mathbb{R}_+$,

$$\mathrm{Ent}(f^2) \geq \mathbb{E}[f^2] \log \frac{\mathbb{E}[f^2]}{\mathbb{E}[f]^2}.$$

Pf. $\mathrm{Ent}(f^2) = \mathbb{E}[f^2 \log(f^2/\mathbb{E}[f^2])]$

$$= \mathbb{E}\left[f^2 \log \frac{f^2}{\mathbb{E}[f^2]^2} \cdot \mathbb{E}[f]^2\right]$$

$$+ \mathbb{E}[f^2] \log \mathbb{E}[f^2] / (\mathbb{E}[f])^2$$

The first term on the right-side equals,

$$2\,\mathbb{E}\left[f^2 \log \frac{f}{\mathbb{E}[f^2]} \cdot \mathbb{E}[f]\right] \geq 2\mathbb{E}\left[f^2\left(1 - \frac{\mathbb{E}[f^2]}{f \cdot \mathbb{E}[f]}\right)\right]$$

$$= 2\,\mathbb{E}[f^2] - 2\,\mathbb{E}[f^2] = 0,$$

which completes the proof. ∎

(d) (log-sum inequality) $\sum_i a_i \log \frac{a_i}{b_i} \geq \sum_i a_i \log \frac{\sum_i a_i}{\sum_i b_i}$

Step 3: Finally, combining everything

$$\mathcal{E}(f) = \sum_{i=1}^{n} \mathcal{E}(\Delta_i) \geq \sum_{i=1}^{n} \mathcal{E}(|\Delta_i|) \geq \frac{1}{2c(p)} \sum_{i=1}^{n} \mathrm{Ent}(\Delta_i^2)$$

$$\geq \frac{1}{2c(p)} \sum_{i=1}^{n} \mathbb{E}[\Delta_i^2] \log \frac{\mathbb{E}[\Delta_i^2]}{(\mathbb{E}[|\Delta_i|])^2}$$

$$\geq \frac{1}{2c(p)} \sum_{i=1}^{n} \mathbb{E}[\Delta_i^2] \log \frac{\sum_{i=1}^{n} \mathbb{E}[\Delta_i^2]}{\sum_{i=1}^{n} (\mathbb{E}[|\Delta_i|])^2}.$$

$$= \frac{1}{2c(p)} \mathrm{Var}(f) \log \frac{\mathrm{Var}(f)}{\sum_{i=1}^{n} \mathbb{E}[|\Delta_i|]^2}$$

Finally, note for $f = \mathbb{1}_A$ at

$$\mathbb{E}[|\Delta_i|] = \mathbb{E}[|Z_i - Z_{i-1}|] \leq \mathbb{E}[|f(X) - \mathbb{E}^{(i)} f|]$$

$$= 2p(1-p)\, I_p^i(A).$$

Thus, $$\mathcal{E}(f) \geq \frac{1}{2c(p)} \mathrm{Var}(\mathbb{1}_A) \log \frac{\mathrm{Var}(\mathbb{1}_A)}{4p^2(1-p)^2 \sum_{i=1}^{n} I_p^i(A)^2}$$

$$\geq \frac{1}{2c(p)} \mathrm{Var}(\mathbb{1}_A) \log \frac{\mathrm{Var}(\mathbb{1}_A)}{4p^2(1-p)^2 \delta_p(A) I_p(A)}$$

∎