Discrepancy I

Problem: Let $\mathcal{F}$ be a system of subsets of X, where $|X|=n$. $c: X \longrightarrow \{1,-1\}$

$$\mathrm{disc}(\mathcal{F}, c) = \max_{S\in\mathcal{F}} |c(S)|, \text{ where } c(S) = \sum_{x\in S} c(x).$$

$$\mathrm{disc}(\mathcal{F}) = \min_{c} \max_{S\in\mathcal{F}} |c(S)|.$$

Fact: If $\mathcal{F} = 2^X$ i.e. all sets then $\mathrm{disc}(\mathcal{F}) = \frac{n}{2}$

Pf: $\mathcal{F} = 2^X \Rightarrow \mathrm{disc}(\mathcal{F}) \geq \frac{n}{2}$.

Consider $c: X \longrightarrow \{1,-1\}$ with $|c^{-1}(-1)| = |c^{-1}(1)| = \frac{n}{2}$ ← w.l.o.g assume n is even

Note that $c(S) \leq \frac{n}{2}$ for each $S \in \mathcal{F}$, so $\max_{S\in\mathcal{F}} |c(S)| \leq \frac{n}{2}$. **QED**

Discrepancy II

Thm Let $|X|=n$ and $|\mathcal{F}|=m$. Then $\operatorname{disc}(\mathcal{F}) \leq \sqrt{2n\ln(2m)}$.

pf: Let $c: X \longrightarrow \{1,-1\}$ be a random coloring.

$$\Pr\Big(\bigcup_{S\in\mathcal{F}} \{c(S)>t\}\Big) \leq \sum_{S\in\mathcal{F}} P(c(S)>t) < \sum_{S\in\mathcal{F}} e^{-\frac{t^2}{2|S|}}$$

Chernoff's ineq.

$$\leq \sum_{S\in\mathcal{F}} e^{-\frac{t^2}{2n}} = m e^{-\frac{t^2}{2n}}.$$

$$m e^{-\frac{t^2}{2n}} \leq 1 \Longleftrightarrow e^{-\frac{t^2}{2n}} \leq \frac{1}{m} \Longleftrightarrow t \geq \sqrt{2n\ln(2m)}$$