Topic Course on Probabilistic Methods (Week 13)
Discrepancy

Linyuan Lu

University of South Carolina
Selected topics

- Linearity of Expectation (2 weeks)
- Alterations (1 week)
- The second moment method (1 week)
- The Local Lemma (1-2 weeks)
- Correlation Inequalities (1 week)
- Large deviation inequalities (3 weeks)
- Poisson Paradigm (1 week)
- Random graphs (2 weeks)
- Discrepancy (1 week)
- Entropy (1 week)
Random graphs

- Discrepancy
- Linear discrepancy
- Hereditary discrepancy
- Lower bound
- The Beck-Fiala Theorem
Discrepancy

- $\Omega$: a finite set.
- $\chi: \Omega \to \{-1, 1\}$.
- For any $A \subset \Omega$, $\chi(A) = \sum_{a \in A} \chi(a)$.
- For $\mathcal{A} \subset 2^\Omega$,

$$
disc(\mathcal{A}, \chi) = \max_{A \in \mathcal{A}} |\chi(A)|;
$$

$$
disc(\mathcal{A}) = \min_{\chi} \text{disc}(\mathcal{A}, \chi).
$$
Discrepancy

- $\Omega$: a finite set.
- $\chi: \Omega \rightarrow \{-1, 1\}$.
- For any $A \subset \Omega$, $\chi(A) = \sum_{a \in A} \chi(a)$.
- For $\mathcal{A} \subset 2^\Omega$,

\[
\text{disc}(\mathcal{A}, \chi) = \max_{A \in \mathcal{A}} |\chi(A)|;
\]
\[
\text{disc}(\mathcal{A}) = \min_{\chi} \text{disc}(\mathcal{A}, \chi).
\]

Geometric meaning: Assume $|\Omega| = m$, $|\mathcal{A}| = n$, and $B = (b_{ij})$ be the $m \times n$ incidence matrix. Let $v_1, v_2, \ldots, v_n$ be the column vector of $B$. Then
\[
\text{disc}(\mathcal{A}) = \min |\pm v_1 \pm v_2 \pm \cdots \pm v_n|_\infty.
\]
Theorem: Let $\mathcal{A}$ be a family of $n$ subsets of an $m$-set $\Omega$. Then

$$\text{disc}(\mathcal{A}) \leq \sqrt{2m \ln(2n)}.$$
**Theorem:** Let $\mathcal{A}$ be a family of $n$ subsets of an $m$-set $\Omega$. Then

$$\text{disc}(\mathcal{A}) \leq \sqrt{2m \ln(2n)}.$$ 

**Proof:** Let $\chi : \Omega \rightarrow \{-1, 1\}$ be random.
Theorem: Let $\mathcal{A}$ be a family of $n$ subsets of an $m$-set $\Omega$. Then

$$\text{disc}(\mathcal{A}) \leq \sqrt{2m \ln(2n)}.$$ 

Proof: Let $\chi: \Omega \to \{-1, 1\}$ be random. Let

$$\lambda = \sqrt{2m \ln(2n)}.$$ 

By Azuma’s inequality, we have

$$\Pr(|\chi(A)| > \lambda) < 2e^{-\lambda^2/(2|A|)} \leq \frac{1}{n}.$$ 

With positive probability, we have $|\chi(A)| \leq \lambda$ holds for every $A \in \mathcal{A}$. Therefore $\text{disc}(A) \leq \lambda$. \qed
Theorem [Spencer (1985)]: Let $\mathcal{A}$ be a family of $n$ subsets of an $n$-element set $\Omega$. Then

$$\text{disc}(\mathcal{A}) < K \sqrt{n}.$$
Theorem [Spencer (1985)]: Let $\mathcal{A}$ be a family of $n$ subsets of an $n$-element set $\Omega$. Then

$$\text{disc}(\mathcal{A}) < K\sqrt{n}.$$ 

- In his paper, $K = 6$ is proved; here we will prove a weaker version with $K = 11$. 
Theorem [Spencer (1985)]: Let $\mathcal{A}$ be a family of $n$ subsets of an $n$-element set $\Omega$. Then

$$\text{disc}(\mathcal{A}) < K \sqrt{n}.$$ 

- In his paper, $K = 6$ is proved; here we will prove a weaker version with $K = 11$.
- If $\mathcal{A}$ consists on $n$ sets on $m$ points and $m \leq n$. Then

$$\text{disc}(\mathcal{A}, \chi) < K \sqrt{m} \sqrt{\ln(n/m)}.$$
Basic entropy

Let $X$ be a random variable taking values in some range $S$. The **binary entropy** of $X$, denoted by $H(X)$ is defined by

$$H(X) = - \sum_{x \in S} \Pr(X = x) \log_2 \Pr(X = x).$$
Let $X$ be a random variable taking values in some range $S$. The **binary entropy** of $X$, denoted by $H(X)$, is defined by

$$H(X) = -\sum_{x \in S} \Pr(X = x) \log_2 \Pr(X = x).$$

**Sub-additive property:**

$$H(X, Y) \leq H(X) + H(Y).$$

Here $(X, Y)$ is the random variable taking values in $S \times T$ (where $T$ is the range of $Y$.)
Proof of entropy inequality

Proof:

\[
H(X) + H(Y) - H(X, Y)
\]

\[
= \sum_{i \in S} \sum_{j \in T} \Pr(X = i, Y = j) \log_2 \frac{\Pr(X = i, Y = j)}{\Pr(X = i) \Pr(Y = j)}
\]

\[
= \sum_{i \in S} \sum_{j \in T} \Pr(X = i) \Pr(Y = j) f(z_{ij}),
\]

where \( f(z) = z \log_2 z \) and \( z_{ij} = \frac{\Pr(X = i, Y = j)}{\Pr(X = i) \Pr(Y = j)} \). By the convexity inequality of \( f(z) \), we have

\[
H(X) + H(Y) - H(X, Y) \geq f(1) = 0. \]

□
A map $\chi : \Omega \rightarrow \{-1, 0, 1\}$ is called a **partial coloring**. When $\chi(a) = 0$ we say $a$ is uncolored.

**Lemma 13.2.2:** Let $A$ be a family of $n$ subsets of an $n$-set $\Omega$. Then there is a partial coloring $\chi$ with at most $10^{-9}n$ points uncolored such that $|\chi(A)| \leq 10\sqrt{n}$ for all $A \in A$. 
A map $\chi : \Omega \rightarrow \{-1, 0, 1\}$ is called a **partial coloring**. When $\chi(a) = 0$ we say $a$ is uncolored.

**Lemma 13.2.2:** Let $A$ be a family of $n$ subsets of an $n$-set $\Omega$. Then there is a partial coloring $\chi$ with at most $10^{-9}n$ points uncolored such that $|\chi(A)| \leq 10\sqrt{n}$ for all $A \in A$.

**Proof:** Let $A := \{A_1, A_2, \ldots, A_n\}$. Consider a random coloring

$$
\chi : \Omega \rightarrow \{-1, 1\}.
$$

For $1 \leq i \leq n$ define

$$
b_i = \text{nearest integer to } \frac{\chi(A_i)}{20\sqrt{n}}.
$$
By Chernoff’s inequality, we have

\[
\begin{align*}
\Pr(b_i = 0) & > 1 - 2e^{-50}, \\
\Pr(b_i = 1) & = \Pr(b_i = -1) < 2^{-50}, \\
\Pr(b_i = 2) & = \Pr(b_i = -2) < 2^{-450}, \\
& \vdots \\
\Pr(b_i = s) & = \Pr(b_i = -s) < 2^{-50(2s-1)^2}.
\end{align*}
\]
By Chernoff’s inequality, we have

\[
\Pr(b_i = 0) > 1 - 2e^{-50},
\]

\[
\Pr(b_i = 1) = \Pr(b_i = -1) < 2^{-50},
\]

\[
\Pr(b_i = 2) = \Pr(b_i = -2) < 2^{-450},
\]

\[
\vdots
\]

\[
\Pr(b_i = s) = \Pr(b_i = -s) < 2^{-50(2s-1)^2}.
\]

Recall the entropy \( H(b_i) \) is defined as

\[
H(b_i) = \sum_{s=\infty}^{s=-\infty} -\Pr(b_i = s) \log_2 \Pr(b_i = s).
\]
By the subadditive property, we have

\[
H(b_1, b_2, \ldots, b_n) \leq \sum_{i=1}^{n} H(b_i) \leq \epsilon n.
\]

If a random variable $Z$ assumes no value with probability greater than $2^{-t}$, then $H(Z) \geq t$. This implies there is a particular $n$-tuple $(s_1, s_2, \ldots, s_n)$ so that

\[
\Pr((b_1, \ldots, b_n) = (s_1, \ldots, s_n)) \geq 2^{-\epsilon n}.
\]
Since every coloring has equal probability $2^{-n}$, there is a set $C$ consisting of at least $2^{(1-\epsilon)n}$ colorings $\chi: \Omega \to \{-1, 1\}$, all having the same value $(b_1, b_2, \ldots, b_n)$. 

continue
Since every coloring has equal probability $2^{-n}$, there is a set $C$ consisting of at least $2^{(1-\epsilon)n}$ colorings $\chi : \Omega \rightarrow \{-1, 1\}$, all having the same value $(b_1, b_2, \ldots, b_n)$.

Kleitman (1966) proved that if $|C| \geq \sum_{i \leq r} \binom{n}{i}$ with $r \leq n/2$ then $C$ has diameter (of Hamming distance) at least $2r$. 
Since every coloring has equal probability $2^{-n}$, there is a set $\mathcal{C}$ consisting of at least $2^{(1-\epsilon)n}$ colorings $\chi : \Omega \rightarrow \{-1, 1\}$, all having the same value $(b_1, b_2, \ldots, b_n)$.

Kleitman (1966) proved that if $|\mathcal{C}| \geq \sum_{i \leq r} \binom{n}{i}$ with $r \leq n/2$ then $\mathcal{C}$ has diameter (of Hamming distance) at least $2r$.

Let $r = \alpha n$ and $2^{H(\alpha)} \leq 2^{1-\epsilon}$. Taylor series expansion gives

$$H\left(\frac{1}{2} - x\right) \sim 1 - \frac{2}{\ln 2}x^2.$$

Thus $\mathcal{C}$ has diameter at least $n(1 - 10^{-9})$. Choose $\chi_1, \chi_2 \in \mathcal{C}$ be at the maximal distance. Let $\chi = \frac{\chi_1 - \chi_2}{2}$. Then the partial coloring $\chi$ satisfying all requirements.
We will iterate the procedure to color the remaining uncolored points.

**Lemma 13.2.3:** Let $\mathcal{A}$ be a family of $n$ subsets of an $m$-set $\Omega$ with at most $10^{-40}m$ points uncolored so that

$$\chi(A) < 10\sqrt{m}\sqrt{\ln(n/m)}$$

for all $A \in \mathcal{A}$.
We will iterate the procedure to color the remaining uncolored points.

Lemma 13.2.3: Let $\mathcal{A}$ be a family of $n$ subsets of an $m$-set $\Omega$ with at most $10^{-40}m$ points uncolored so that

$$\chi(A) < 10 \sqrt{m} \sqrt{\ln(n/m)}$$

for all $A \in \mathcal{A}$.

The proof is similar by define

$$b_i = \text{nearest integer to} \frac{\chi(A_i)}{20 \sqrt{m \ln(n/m)}}.$$

The detail is omitted. \qed
Proof: Apply Lemma 13.2.2 to find a partial coloring $\chi^1$ and then apply Lemma 13.2.3 repeatedly on the remaining uncolored points giving $\chi^2, \chi^3, \ldots$ until all points have been colored. Let $\chi = \sum_{i \geq 1} \chi^i$. 
Proof of Theorem

Proof: Apply Lemma 13.2.2 to find a partial coloring $\chi^1$ and then apply Lemma 13.2.3 repeatedly on the remaining uncolored points giving $\chi^2, \chi^3, \ldots$ until all points have been colored. Let $\chi = \sum_{i \geq 1} \chi^i$. Then

$$|\chi(A)| \leq 10\sqrt{n} + 10\sqrt{10^{-9}n\sqrt{\ln 10^9}}$$

$$+ 10\sqrt{10^{-49}n\sqrt{\ln 10^{49}}} + 10\sqrt{10^{-89}n\sqrt{\ln 10^{89}}}$$

$$\leq 11\sqrt{n}. \quad \square$$
Proof of Theorem

Proof: Apply Lemma 13.2.2 to find a partial coloring $\chi^1$ and then apply Lemma 13.2.3 repeatedly on the remaining uncolored points giving $\chi^2, \chi^3, \ldots$ until all points have been colored. Let $\chi = \sum_{i \geq 1} \chi^i$. Then

$$|\chi(A)| \leq 10\sqrt{n} + 10\sqrt{10^{-9}n \sqrt{\ln 10^9}} + 10\sqrt{10^{-49}n \sqrt{\ln 10^{49}}} + 10\sqrt{10^{-89}n \sqrt{\ln 10^{89}}} \leq 11\sqrt{n}.$$ 

The statement of case $r < n$ can be proved similarly.
More points than sets

Suppose $m > n$, $\mathcal{A} = \{A_1, A_2, \ldots, A_n\}$ and $\Omega = [n]$. The linear discrepancy $\text{lindisc}(\mathcal{A})$ is defined by

$$\text{lindisc}(\mathcal{A}) = \max_{p_1, \ldots, p_m \in [0,1]} \min_{\epsilon_1, \ldots, \epsilon_m \in \{0,1\}} \max_{A \in \mathcal{A}} \left| \sum_{i \in A} (\epsilon_i - p_i) \right|.$$
More points than sets

Suppose $m > n$, $\mathcal{A} = \{A_1, A_2, \ldots, A_n\}$ and $\Omega = [n]$. The \textbf{linear discrepancy} $\text{lindisc}(\mathcal{A})$ is defined by

$$\text{lindisc}(\mathcal{A}) = \max_{p_1, \ldots, p_m \in [0,1]} \min_{\epsilon_1, \ldots, \epsilon_m \in \{0,1\}} \max_{A \in \mathcal{A}} \left| \sum_{i \in A} (\epsilon_i - p_i) \right|.$$ 

Setting all $\epsilon_i = \frac{1}{2}$ and scaling $[0, 1]$ to $[-1, 1]$, we have

$$\text{disc}(A) = \min_{\epsilon'_1, \ldots, \epsilon'_m \in \{-1,1\}} \max_{A \in \mathcal{A}} \left| \sum_{i \in A} \epsilon'_i \right|$$

$$= 2 \min_{\epsilon_1, \ldots, \epsilon_m \in \{0,1\}} \max_{A \in \mathcal{A}} \left| \sum_{i \in A} \epsilon_i - \frac{1}{2} \right|$$

$$\leq 2 \cdot \text{lindisc}(A).$$
Theorem 13.3.1 Let $\mathcal{A}$ be a family of $n$ sets on $m$ points with $m \geq n$. Suppose that $\text{lindisc}(\mathcal{A}|_X) \leq K$ for every subset $X$ of at most $n$ points. Then $\text{lindisc}(\mathcal{A}) \leq K$. 
**Theorem 13.3.1** Let $\mathcal{A}$ be a family of $n$ sets on $m$ points with $m \geq n$. Suppose that $\text{lindisc}(\mathcal{A}|_X) \leq K$ for every subset $X$ of at most $n$ points. Then $\text{lindisc}(\mathcal{A}) \leq K$.

**Proof:** For $p_1, \ldots, p_m \in [0, 1]$, call index $j$ fixed if $p_i = 0$ or $1$ otherwise call it floating, and let $F$ denote the set of floating indices.
A theorem

**Theorem 13.3.1** Let $\mathcal{A}$ be a family of $n$ sets on $m$ points with $m \geq n$. Suppose that $\text{lindisc}(\mathcal{A}|_X) \leq K$ for every subset $X$ of at most $n$ points. Then $\text{lindisc}(\mathcal{A}) \leq K$.

**Proof:** For $p_1, \ldots, p_m \in [0, 1]$, call index $j$ fixed if $p_i = 0$ or 1 otherwise call it floating, and let $F$ denote the set of floating indices.

Our goal is to reduce $p_1, p_2, \ldots, p_m$ so that $|F| < n$. 
Theorem 13.3.1 Let $\mathcal{A}$ be a family of $n$ sets on $m$ points with $m \geq n$. Suppose that $\text{lindisc}(\mathcal{A}|_X) \leq K$ for every subset $X$ of at most $n$ points. Then $\text{lindisc}(\mathcal{A}) \leq K$.

Proof: For $p_1, \ldots, p_m \in [0, 1]$, call index $j$ fixed if $p_i = 0$ or 1 otherwise call it floating, and let $F$ denote the set of floating indices.

Our goal is to reduce $p_1, p_2, \ldots, p_m$ so that $|F| < n$.

Suppose $|F| \geq n$. Let $y_1, \ldots, y_m$ be a nonzero solution to the homogeneous system

$$\sum_{j \in A \cap F} y_j = 0, \quad A \in \mathcal{A}.$$
Consider a line

\[ p'_j = \begin{cases} 
  p_j + \lambda y_j, & j \in F, \\
  p_j, & j \notin F.
\end{cases} \]

The line will hit the boundary of the hypercube \( Q^m \) and the intersection point gives a set of \( p'_1, \ldots, p'_m \) with the smaller floating indices. Critically, for all \( A \in A \).

\[
\sum_{j \in A} p'_j = \sum_{j \in A} p_j + \lambda \sum_{j \in A \cap F} y_j = \sum_{j \in S} p_j.
\]

Iterate this process, we get some \( p^*_1, \ldots, p^*_m \) with the set \( X \) of floating indices satisfying \( |X| < n \).
Since $\text{lindisc}(\mathcal{A}|_X) \leq K$, there exists $\epsilon_j, j \in X$ so that

$$\left| \sum_{j \in A \cap X} p_j^* - \epsilon_j \right| \leq K, \quad A \in \mathcal{A}.$$ 

Extend $\epsilon_j$ to $j \in \bar{X}$ by letting $\epsilon_j = p_j^*$. For any $A \in \mathcal{A}$,

$$\left| \sum_{j \in A} (p_j - \epsilon_j) \right| = \left| \sum_{j \in A} (p_j^* - \epsilon_j) \right|$$

$$= \left| \sum_{j \in A \cap X} (p_j^* - \epsilon_j) \right| \leq K.$$ 

Thus, $\text{lindisc}(\mathcal{A}) \leq K$. \qed
The **hereditary discrepancy** \( \text{herdisc}(\mathcal{A}) \) is defined by

\[
\text{herdisc}(\mathcal{A}) = \max_{X \in \Omega} \text{disc}(\mathcal{A}|_X).
\]
The **hereditary discrepancy** $\text{herdisc}(\mathcal{A})$ is defined by

\[
\text{herdisc}(\mathcal{A}) = \max_{X \in \Omega} \text{disc}(\mathcal{A}|_X).
\]

**Theorem 13.3.2:** $\text{lindisc}(\mathcal{A}) \leq \text{herdisc}(\mathcal{A})$. 
The **hereditary discrepancy** $\text{herdisc}(\mathcal{A})$ is defined by

$$\text{herdisc}(\mathcal{A}) = \max_{X \in \Omega} \text{disc}(\mathcal{A}|_X).$$

**Theorem 13.3.2:** $\text{lindisc}(\mathcal{A}) \leq \text{herdisc}(\mathcal{A})$.

**Proof:** Set $K = \text{herdisc}(\mathcal{A})$. Let $p_1, \ldots, p_m \in [0, 1]$ be given. First assume all $p_i$ have finite expansions in base 2. Let $T$ be the minimal integer so that all $p_i2^T \in \mathbb{Z}$. Let $J$ be the set of $i$ for which $p_i2^T$ is odd. As $\text{disc}(\mathcal{A}|_J) \leq K$, there exists $\epsilon_j \in \{-1, 1\}$, so that

$$\left| \sum_{j \in J \cap A} \epsilon_j \right| \leq K, \quad A \in \mathcal{A}.$$
For $i$ from $T$ to 0, let $p_j = p_j^{(T)}$ and $p_j^{(i-1)}$ be the "roundoffs" of $p_j^i$. For any $A \in \mathcal{A}$,

$$
\left| \sum_{j \in A} (p_j^{(i-1)} - p_j^{(i)}) \right| = \sum_{j \in J^{(i)} \cap A} 2^{-i} \epsilon_j^{(i)} \leq 2^{-i} K.
$$

Thus, for any $A \in \mathcal{A}$,

$$
\left| \sum_{j \in A} p_j^{(0)} - p_j^{(T)} \right| \leq \sum_{i=1}^{T} \left| \sum_{j \in A} (p_j^{(i-1)} - p_j^{(i)}) \right| \leq \sum_{i=1}^{T} 2^{-i} K \leq K.
$$
For $p_1, p_2, \ldots, p_m \in [0, 1]$, consider the function

$$f(p_1, \ldots, p_m) = \min_{\epsilon_1, \ldots, \epsilon_m \in \{0, 1\}} \max_{A \in A} \left| \sum_{i \in A} (\epsilon_i - p_i) \right|.$$ 

Note that $f(p_1, p_2, \ldots, p_m)$ is continuous.
For $p_1, p_2, \ldots, p_m \in [0, 1]$, consider the function

$$f(p_1, \ldots, p_m) = \min_{\epsilon_1, \ldots, \epsilon_m \in \{0, 1\}} \max_{A \in \mathcal{A}} \left| \sum_{i \in A} (\epsilon_i - p_i) \right|.$$ 

Note that $f(p_1, p_2, \ldots, p_m)$ is continuous. We just proved that

$$f(p_1, p_2, \ldots, p_m) \leq K$$

for a dense set of $[0, 1]^m$. Thus it holds for any $(p_1, \ldots, p_m) \in [0, 1]^m$. This implies

$$\text{lindisc}(\mathcal{A}) \leq K.$$
Corollary: 13.3.3: Let $\mathcal{A}$ be a family of $n$ sets on $m$ points. Suppose $\text{disc}(\mathcal{A}|_X) \leq K$ for every subset $X$ with at most $n$ points. Then $\text{disc}(\mathcal{A}) \leq 2K$. 
Corollary: 13.3.3: Let $\mathcal{A}$ be a family of $n$ sets on $m$ points. Suppose $\text{disc}(\mathcal{A}|_X) \leq K$ for every subset $X$ with at most $n$ points. Then $\text{disc}(\mathcal{A}) \leq 2K$.

Proof: By Theorem 13.3.2, $\text{lindisc}(\mathcal{A}|_X) \leq K$ for every $X \subset \Omega$ with $|X| \leq n$. By Theorem 13.3.1, $\text{lindisc}(\mathcal{A}) \leq K$. Thus,

$$\text{disc}(\mathcal{A}) \leq 2 \cdot \text{lindisc}(\mathcal{A}) \leq 2K.$$
Corollary: 13.3.3: Let $\mathcal{A}$ be a family of $n$ sets on $m$ points. Suppose $\text{disc}(\mathcal{A}|_X) \leq K$ for every subset $X$ with at most $n$ points. Then $\text{disc}(\mathcal{A}) \leq 2K$.

Proof: By Theorem 13.3.2, $\text{lindisc}(\mathcal{A}|_X) \leq K$ for every $X \subset \Omega$ with $|X| \leq n$. By Theorem 13.3.1, $\text{lindisc}(\mathcal{A}) \leq K$. Thus,

$$\text{disc}(\mathcal{A}) \leq 2 \cdot \text{lindisc}(\mathcal{A}) \leq 2K.$$  

Corollary 13.3.4: For any family $\mathcal{A}$ of $n$ sets of arbitrary size

$$\text{disc}(\mathcal{A}) \leq 12\sqrt{n}.$$
Lower bounds: $\text{disc}(\mathcal{A}) \geq C \sqrt{n}$. 
**Lower bounds:** $\text{disc}(A) \geq C\sqrt{n}$.

Two methods:
- Using Hadamard matrices.
Lower bounds: \( \text{disc}(\mathcal{A}) \geq C\sqrt{n} \).

Two methods:
- Using Hadamard matrices.
- Using probabilistic method.
A **Hadamard matrix** is a $n \times n$ matrix $H = (h_{ij})$ with all entries $\pm 1$ and row vectors mutually orthogonal (and hence with column vectors mutually orthogonal).
A Hadamard matrix is a $n \times n$ matrix $H = (h_{ij})$ with all entries $\pm 1$ and row vectors mutually orthogonal (and hence with column vectors mutually orthogonal).

- $HH' = nI$.
- If $A$ is an $n \times n$ $(\pm)$-matrix, then $|\det(A)| \leq n^{n/2}$. The equality holds if and only if $A$ is an Hadamard matrix.
- If $H_1$ and $H_2$ are Hadamard matrices, then so is $H_1 \otimes H_2$.
- If $\exists n \times n$ Hadamard matrix, then $n = 1, 2$ or $4|n$. 

Hadamard matrices
A Hadamard matrix is a $n \times n$ matrix $H = (h_{ij})$ with all entries $\pm 1$ and row vectors mutually orthogonal (and hence with column vectors mutually orthogonal).

- $HH' = nI$.
- If $A$ is an $n \times n$ $(\pm)$-matrix, then $|\det(A)| \leq n^{n/2}$. The equality holds if and only if $A$ is an Hadamard matrix.
- If $H_1$ and $H_2$ are Hadamard matrices, then so is $H_1 \otimes H_2$.
- If $\exists n \times n$ Hadamard matrix, then $n = 1, 2$ or $4|n$.

It is conjectured that Hadamard matrix exists for every $n = 1, 2$ and all multiples of 4.
A **Hadamard matrix** is a $n \times n$ matrix $H = (h_{ij})$ with all entries $\pm 1$ and row vectors mutually orthogonal (and hence with column vectors mutually orthogonal).

- $HH' = nI$.
- If $A$ is an $n \times n$ $(\pm)$-matrix, then $|\det(A)| \leq n^{n/2}$. The equality holds if and only if $A$ is an Hadamard matrix.
- If $H_1$ and $H_2$ are Hadamard matrices, then so is $H_1 \otimes H_2$.
- If $\exists n \times n$ Hadamard matrix, then $n = 1, 2$ or $4|n$.

It is **conjectured** that Hadamard matrix exists for every $n = 1, 2$ and all multiples of 4.

**Hall (1986)** For all $\epsilon > 0$ and sufficiently large $n$, there is a Hadamard matrix of order between $n(1 - \epsilon)$ and $n$. 

---

**Hadamard matrices**
Let $H$ be a Hadamard matrix of order $n$ (even) with first row and first column all ones. (Any Hadamard matrix can be so “normalized” by multiplying appropriate rows and columns by $-1$.) Let $J$ be all ones square matrix of order $n$. Let $v = (v_1, \ldots, v_n)'$ be the column vector with each $v_i \in \{-1, 1\}$. Then

$$
\langle (H + J)v, (H + J)v \rangle = n^2 + 2n \left( \sum_{i=1}^{n} v_i \right)v_1 + n \left( \sum_{i=1}^{n} v_i \right)^2 \geq n^2.
$$

Setting $H^* = (H + J)/2$, then,

$$
\|H^*v\|_\infty \geq \sqrt{\|H^*v\|^2/n} \geq \frac{\sqrt{n}}{2}.
$$

Let $\mathcal{A}$ be the family of subsets with incidence matrix $H^*$.  

---

**Construction I**

Let $H$ be a Hadamard matrix of order $n$ (even) with first row and first column all ones. (Any Hadamard matrix can be so “normalized” by multiplying appropriate rows and columns by $-1$.) Let $J$ be all ones square matrix of order $n$. Let $v = (v_1, \ldots, v_n)'$ be the column vector with each $v_i \in \{-1, 1\}$. Then

$$
\langle (H + J)v, (H + J)v \rangle = n^2 + 2n \left( \sum_{i=1}^{n} v_i \right)v_1 + n \left( \sum_{i=1}^{n} v_i \right)^2 \geq n^2.
$$

Setting $H^* = (H + J)/2$, then,

$$
\|H^*v\|_\infty \geq \sqrt{\|H^*v\|^2/n} \geq \frac{\sqrt{n}}{2}.
$$

Let $\mathcal{A}$ be the family of subsets with incidence matrix $H^*$.  

Construction II

- \( M \): a random 0, 1 matrix of order \( n \).
- \( d_i \): \( i \)-th row sum of \( M \), \( d_i = (1 + o(1))n/2 \).
- \( v := (v_1, \ldots, v_n)' \), \( v_i = \pm 1 \), set \( Mv = (L_1, L_2, \ldots, L_n) \).

\[
L_i \sim B(d_i, 1/2) - B(d_i, 1/2) \sim N(0, \sqrt{n}/2).
\]
Construction II

- \( M \): a random 0, 1 matrix of order \( n \).
- \( d_i \): \( i \)-th row sum of \( M \), \( d_i = (1 + o(1))n/2 \).
- \( v := (v_1, \ldots, v_n)' \), \( v_i = \pm 1 \), set \( Mv = (L_1, L_2, \ldots, L_n) \).

\[
L_i \sim B(d_i, 1/2) - B(d_i, 1/2) \sim N(0, \sqrt{n/2}).
\]

Pick \( \lambda \) so that

\[
\int_{-\lambda}^{\lambda} \frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt < \frac{1}{2}.
\]

Then \( \Pr(|L_i| < \lambda\sqrt{n/2}) < \frac{1}{2} \). The expected number of \( v \) for which \( |Mv|_\infty < \lambda\sqrt{n/2} \) is less than 1. \( \exists M \) such that \( |Mv|_\infty \geq \lambda\sqrt{n/2} \) for every \( v \).
Construction II

- $M$: a random 0, 1 matrix of order $n$.
- $d_i$: $i$-th row sum of $M$, $d_i = (1 + o(1))n/2$.
- $v := (v_1, \ldots, v_n)'$, $v_i = \pm 1$, set $Mv = (L_1, L_2, \ldots, L_n)$.

$$L_i \sim B(d_i, 1/2) - B(d_i, 1/2) \sim N(0, \sqrt{n}/2).$$

Pick $\lambda$ so that

$$\int_{-\lambda}^{\lambda} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt < \frac{1}{2}.$$

Then $\Pr(|L_i| < \lambda \sqrt{n}/2) < \frac{1}{2}$. The expected number of $v$ for which $|Mv|_\infty < \lambda \sqrt{n}/2$ is less than 1. $\exists M$ such that $|Mv|_\infty \geq \lambda \sqrt{n}/2$ for every $v$.

Let $\mathcal{A}$ be the family of sets with incident matrix $M$. Then

$$\text{disc}(\mathcal{A}) \geq \lambda \lambda \sqrt{n}/2.$$
For any $A$, let $\deg(A)$ denote the maximal number of sets containing any particular points.
Beck-Fiala Theorem

For any $A$, let $\deg(A)$ denote the maximal number of sets containing any particular points.

**Theorem [Beck-Fiala 1981]** Let $A$ be a finite family of finite sets. If $\deg(A) \leq t$, then

$$\text{disc}(A) \leq 2t - 1.$$
For any $A$, let $\deg(A)$ denote the maximal number of sets containing any particular points.

**Theorem [Beck-Fiala 1981]** Let $A$ be a finite family of finite sets. If $\deg(A) \leq t$, then

$$\text{disc}(A) \leq 2t - 1.$$  

**Proof:** Assume $A = \{A_1, A_2, \ldots, A_m\}$ where all $A_i \subset [n]$. Let $x = (x_1, \ldots, x_n) \in [-1, 1]^n$. A set $S_i$ has value $\sum_{j \in S_i} x_j$. We say an index $j$ is **fixed** if $x_j = \pm 1$; otherwise we say $j$ is **floating**. A set $S_i$ is **safe** if it has at most $t$ floating points; otherwise it is **active**.

**Fact:** There are fewer active sets than floating points.
Initially all $j$ are floating; i.e. $x$ is the zero vector. We will change $x$ to $x'$ with fewer floating points while keep the values of all sets to 0.

**Iteration:** For each active set, move the fixed points to the right hand side. We get a system of linear equations where the unknown variables are floating points. Since there are fewer active sets than floating points. This is an underdetermined system. The solution contains a line, parametrized

$$x'_j = x_j + \lambda y_j, \quad j \text{ floating},$$

on which the active sets retain value zero. Choose the smallest $\lambda$ on the absolute value so that one of $x'_j = 1$. 

continue
After many iterations, we get a vector $x$ so that every set is safe and has value $0$. For each floating point $j$, setting $x_j = \pm 1$ arbitrarily. For each set, the value may change less than $2t$ and, as it is an integer, it is at most $2t - 1$. □
After many iterations, we get a vector $x$ so that every set is safe and has value 0. For each floating point $j$, setting $x_j = \pm 1$ arbitrarily. For each set, the value may change less than $2t$ and, as it is an integer, it is at most $2t - 1$. □

**Conjecture:** If $\deg(A) \leq t$, then $\text{disc}(A) \leq K\sqrt{t}$, for some absolute constant.