PROBABILISTIC COMBINATORICS

Lecture 1

1. Ramsey numbers, cliques and chromatic numbers

1. Binomial random graphs

$0 < p = p(n) < 1$; $\mathcal{G}(n,p)$ = space of random graphs on $[n] = \{1, \ldots, n\}$ obtained by joining the vertices with probability p independently of each other. $G_{n,p} \leftarrow$ sample/RV

$\mathcal{G}(n,p)$: if H is any graph on $[n]$ with ℓ edges then $\mathbb{P}_{n,p}(G_{n,p} = H) = p^{\ell}(1-p)^{\binom{n}{2}-\ell}$.

This is the BINOMIAL MODEL or Erdős–Rényi model, originally they used uniform

A property (of graphs) is said to hold whp (with high probability) if it holds with probability $\to 1$.

F fixed graph, k vertices, ℓ edges, a elements in automorphism group

$X_F = X_F(G_{n,p})$ = # F-subgraphs of $G_{n,p}$ X counting

$$\mathbb{E}_{n,p}(X_F) = \underbrace{\binom{n}{k}\frac{k!}{a}}_{\#F\text{-graphs on }[n]} p^{\ell} = \frac{(n)_k}{a} p^{\ell} \qquad (n)_k = n(n-1)\ldots(n-k+1) \quad \text{FALLING FACTORIAL}$$

$\mathbb{1}_{X \geq 1} \leq X$

If $\mathbb{E}_{n,p}(X_F) \to 0$ then whp $G_{n,p}$ contains no F-subgraph. Indeed $\mathbb{P}(\exists F) = \mathbb{P}(X_F \geq 1) \leq \mathbb{E}(X_F)$

e.g. if $p(n) = o(\frac{1}{n})$ then $\forall$ fixed k whp there are no C_k in $G_{n,p}$. $\mathbb{P}(\exists C_k) \leq \frac{(n)_k}{2k} p^k \leq \frac{(np)^k}{2k}$

If p is fixed then $\mathcal{G}(\mathbb{N},p)$ couples all $G(n,p)$ $n = 1, 2, \ldots$ by taking $G(n,p)$ to be the initial segment of length n of $\mathcal{G}(\mathbb{N},p)$. RANDOM GRAPH

2. Ramsey numbers

$G_{n,p} \supseteq K_s \vee \overline{G_{n,p}} \supseteq K_t$

$s, t \geq 2$ $R(s,t)$ minimal n st ... $R(s,t) = R(t,s)$ $R(2,t) = t$ $R(3,3) = 6$

Ramsey's theorem (1930): $R(s,t) < \infty$

Erdős–Szekeres (1935)

Theorem 1 For $s, t \geq 2$ $R(s,t) \leq R(s-1,t) + R(s,t-1)$

OFF-DIAGONAL $R(s,t) \leq \binom{s+t-2}{s-1}$ DIAGONAL $R(s) = R(s,s) \leq \binom{2s-2}{s-1} < \frac{c}{\sqrt{s}} 4^s$

Erdős (1947)

Theorem 2 If $\binom{n}{s} 2^{-\binom{s}{2}+1} < 1$ then $R(s) \geq n+1$

Proof Consider $\mathcal{G}(n,\frac{1}{2})$. $X_s = X_s(G_{n,\frac{1}{2}}) = \# K_s \subseteq G_{n,\frac{1}{2}}$ $\overline{X_s} = \# K_s \subseteq \overline{G_{n,\frac{1}{2}}}$ = # of sets of s independent vertices

$$\mathbb{P}(X_s + \overline{X_s} \geq 1) \leq \mathbb{E}(X_s + \overline{X_s}) = 2\mathbb{E}(X_s) = 2\binom{n}{s} 2^{-\binom{s}{2}} < 1$$

Hence $\exists$ $G_{n,\frac{1}{2}}$ with $X_s(G_{n,\frac{1}{2}}) = \overline{X_s}(G_{n,\frac{1}{2}}) = 0$. Thus $R(s) \geq n+1$

Corollary 3 $R(s) \geq (\frac{s}{e}) 2^{\frac{(s-1)}{2}} > \frac{s}{e} 2^{\frac{s}{2}}$

Upper bound: Yackel (1972) (any); Rödl & Thomason (1980s) (small); Conlon (2007) (wide current)

$\exists c, s_0$ $R(s)$ ... $\frac{\log s}{\log\log s} 4^s$ best known

... lower bound presently

3. Cliques

$cl(G)$ = maximal ... for a complete subgraph

$\alpha(G)$ ... # ... $cl(G_{n,p}) = \max\{r : X_r(G_{n,p}) \geq 1\}$ $\mu_r = \binom{n}{r} p^{\binom{r}{2}}$

$b = \frac{1}{p}$ $2\log_b n$

Things to know: Expected number of F-subgraphs in $G_{n,p}$

$R(s,t) \leq \binom{s+t-2}{2}$ $\qquad$ $R(s) \leq \frac{c}{\sqrt{s}} 4^s$

$R(s) > \frac{s}{5} 2^{\frac{s}{2}}$

$b=\frac{1}{p}$; if $r \geq 2\log_b n$ then $np^{\frac{r}{2}} \leq 1$

$$\mu_r \leq \frac{(np^{\frac{r}{2}})^r p^{-\frac{r}{2}}}{r!} \leq \frac{p^{-\frac{r}{2}}}{r!} \to 0$$

If $r \leq (2-\varepsilon)\log_b n$ then $np^{\frac{r}{2}} \geq n^{\frac{\varepsilon}{2}}$

$$\mu_r \geq c\frac{n^r}{r!}p^{\binom{r}{2}} \geq c\left(\frac{np^{\frac{r}{2}}}{r}\right)^r \geq c\left(\frac{n^{\frac{\varepsilon}{2}}}{r}\right)^r \to \infty \text{ as } n\to\infty$$

$\times p^{\frac{1}{2}r}$

$\mathbb{P}(X\geq 1)\leq \mathbb{E}(X)$

Theorem 4 Let $0<p<1$ be fixed. Then if $\mu_r \to 0$ then whp $X_r(G_{n,p})=0$; if $\mu_r\to\infty$ then whp $X_r(G_{n,p})\geq 1$.

Proof $b=\frac{1}{p}$ First ✓: $\mathbb{P}(X_r\geq 1)\leq \mathbb{E}(X_r)=\mu_r\to 0$. SEE NEXT PAGE

$\binom{r+1}{2}=\binom{r}{2}+\binom{r}{1}$

$$\frac{\mu_{r+1}}{\mu_r}=\frac{n-r}{r+1}p^{(+)r}=n^{-1+o(1)} \text{ if } r=(2+o(1))\log_b n$$

wlog $\mu_r \leq n^{\frac{3}{2}}$; $r=(2+o(1))\log_b n$

Write σ_r^2 for the variance of X_r it suffices to prove that $\frac{\sigma_r}{\mu_r}\to 0$ $\left[\mathbb{P}(X=0)\leq\frac{\sigma_r^2}{\mu_r^2}\right]$

$$\mathbb{E}(X_r^2)=\sum_{l=0}^{r}\binom{n}{r}\binom{r}{l}\binom{n-r}{r-l}p^{2\binom{r}{2}-\binom{l}{2}} \qquad \mathbb{E}(X_r)^2=\left(\binom{n}{r}p^{\binom{r}{2}}\right)^2=\sum_{l=0}^{r}\binom{n}{r}\binom{r}{l}\binom{n-r}{r-l}p^{2\binom{r}{2}}$$

$$\frac{\sigma_r^2}{\mu_r^2}=\sum_{l=2}^{r}\frac{\binom{n}{r}\binom{r}{l}\binom{n-r}{r-l}p^{2\binom{r}{2}}\left(p^{-\binom{l}{2}}-1\right)}{\binom{n}{r}^2p^{2\binom{r}{2}}} \leq \sum_{l=2}^{r}\frac{\binom{r}{l}\binom{n-r}{r-l}}{\binom{n}{r}}p^{-\binom{l}{2}}=\sum_{l=2}^{r}\varepsilon_l$$

$$\varepsilon_2=O\left(\frac{(\log_b n)^4}{n}\right) \quad ; \quad \varepsilon_r=\frac{1}{\mu_r}$$

Also, $\frac{\varepsilon_{l+1}}{\varepsilon_l}=\frac{(r-l)^2}{l+1}\frac{\cancel{n-2r+l+2}}{n-2r+l+1}p^{-l}$ $\therefore \sum_{2}^{r}\varepsilon_l\leq 2(\varepsilon_2+\varepsilon_r)=o(1)$ □

$\left(\mathcal{G}(\mathbb{N},p) \quad \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \quad G_{\mathbb{N},p}\right)$ n $\mathcal{G}(n,p)$

Theorem 5 $\exists\ r_0=r_0(n)$, $r_1=r_1(n)$ st whp the clique number of $G_{n,p}$ satisfies $r_0\leq cl(G_{n,p})\leq r_1$, $r_1\leq r_0+1$

In fact, for most values of n, $r_0=r_1$ (the density of n with $r_0\neq r_1$ is 0).

Proof $\frac{\mu_{r+1}}{\mu_r}=n^{-1+o(1)}$ $(r\sim 2\log_b n)$. Hence, if n is large, $\exists$ at most one value of r st $\frac{1}{r}\leq\mu_r\leq r$.

If $r_0(n)$ is the maximal r with $\mu_r\geq r$ and $r_1(n)$ is the minimal r with $\mu_{r+1}\leq\frac{1}{r}$ then whp $cl(G_{n,p})\in[r_0,r_1]$. Also, $r_1\leq r_0+1$ since $\frac{\mu_{r+1}}{\mu_r}=n^{-1+o(1)}$. □

Note, with $\mu_{n,r}=\mathbb{E}_{n,p}(X_r)$, $\frac{\mu_{n+1,r}}{\mu_{n,r}}=\frac{n+1}{n-r+1}=1+O\left(\frac{\log n}{n}\right)$

Theorem 6 Given p, $\exists\ (n_r)$, (n_r') st with probability 1 $G_{\mathbb{N},p}\in\mathcal{G}(\mathbb{N},p)$ is st for n large enough satisfying $n_r\leq n\leq n_r'$, $cl(G_{\mathbb{N},p}[n])=r$. Also, the density of $\bigcup[n_r',n_{r+1}]$ is 0.

$$n_r<\underbrace{n_r'<n_{r+1}}_{\text{small}}<\underbrace{n_{r+1}'<\ldots}_{\text{small}}$$

Proof As before, prove Borel-Cantelli. □

If $\frac{1}{2}\log_b n\leq l\leq\frac{3}{2}\log_b n$ then $\varepsilon_l\leq\left(\frac{r^2p^{-\frac{(l-1)}{2}}}{n}\right)^l\leq n^{-\frac{l}{5}}<n^{-5}$ $l\to\infty$

See $\rho_l=\frac{\varepsilon_{l+1}}{\varepsilon_l}=\frac{(r-l)^2p^{-l}}{(l+1)(n-2r+l+1)}$

If $2\leq l\leq\frac{1}{2}\log_b n$ then $\rho_l\leq n^{-\frac{1}{3}}$

If $\frac{3}{2}\log_b n\leq l\leq r$ then $\rho_l\geq n^{\frac{1}{5}}$

Hence $\sum_{2}^{r}\varepsilon_l\leq\frac{3}{2}(\varepsilon_2+\varepsilon_r)+n^{-4}\leq 2\varepsilon_r=\frac{2}{\mu_r}\to 0$

~~$\mathbb{P}(X_r=0)\leq\frac{\sigma^2}{\mu_r^2}\to 0$ $\Rightarrow$ $X_r\Rightarrow$ whp~~

□

$\binom{n}{r} \geq \frac{cn^r}{r!} \iff c \leq \frac{n(n-1)\dots(n-r+1)}{n^r}$

If $r \leq \alpha \log n$, $\alpha > 0$ then $\frac{n(n-1)\dots(n-r+1)}{n^r} = 1.(1-\frac{1}{n}).\quad\dots.(1-\frac{r-1}{n})$

$\geq \left(1-\frac{\alpha \log n}{n}\right)^{\alpha \log n}$

$\geq e^{-\frac{\alpha^2 \beta (\log n)^2}{n}}$ any $\beta > 1$

continuous, positive, $\to 1$ so bounded below

Theorem 4 $0<p=\frac{1}{b}<1$ fixed, $X_r = \#(r\text{-cliques in } G_{n,p})$. Write $\mu_r = \mathbb{E}(X_r) = \binom{n}{r} p^{\binom{r}{2}}$. L2a

If $r \geq 2\log_b n$ then $\mu_r \to 0$ and whp $X_r = 0$.

If $r \leq (2-\varepsilon)\log_b n$ then $\mu_r \to \infty$ and whp $X_r \neq 0$.

Proof. $r \geq 2\log_b n \Rightarrow p^{-r} \geq n^2 \Rightarrow np^{\frac{r}{2}} \leq 1$.

$$\mu_r \leq \frac{n^r}{r!} p^{\binom{r}{2}} = \frac{(np^{\frac{r}{2}})^r p^{-\frac{r}{2}}}{r!} \leq \frac{p^{-\frac{r}{2}}}{r!} \to 0 \text{ as } n\to\infty \ (\Rightarrow r\to\infty)$$

$\mathbb{P}(X_r \geq 1) \leq \mathbb{E}(X_r)$ gives $X_r = 0$ whp.

$r \leq (2-\varepsilon)\log_b n \Rightarrow p^{-r} \leq n^{2-\varepsilon} \Rightarrow np^{\frac{r}{2}} \geq n^{\frac{\varepsilon}{2}}$.

Claim For these r, $\binom{n}{r} \geq c\frac{n^r}{r!}$

Proof Equivalently, $c \leq \frac{n(n-1)\cdots(n-r+1)}{n^r} = 1\cdot(1-\frac{1}{n})\cdots(1-\frac{r-1}{n}) \geq (1-\frac{r}{n})^r$

so the expression is bounded below $\geq e^{-\frac{\alpha r^2}{n}} \to 1$ any $\alpha > 1$

$$\text{So } \mu_r \geq \frac{cn^r}{r^r} p^{\frac{r^2}{2}} = c\left(\frac{np^{\frac{r}{2}}}{r}\right)^r \geq c\left(\frac{n^{\frac{\varepsilon}{2}}}{r}\right)^r \geq c\left(\frac{n^{\frac{\varepsilon}{2}}}{2\log_b n}\right) \to \infty \text{ as } n\to\infty$$

$$\mathbb{E}(X_r^2) = \mathbb{E}(\#(\text{pairs of } r\text{-cliques})) = \sum_{l=0}^{r} \binom{n}{r}\binom{r}{l}\binom{n-r}{r-l} p^{2\binom{r}{2}-\binom{l}{2}}$$

$$\mathbb{E}(X_r)^2 = \binom{n}{r}^2 p^{2\binom{r}{2}} = \sum_{l=0}^{r} \underbrace{\binom{n}{r}}_{}\underbrace{\binom{r}{l}}_{\text{first set}}\underbrace{\binom{n-r}{r-l}}_{\text{second set}} p^{2\binom{r}{2}}$$

0 for $l=0,1$

$$\Rightarrow \frac{\mathrm{Var}(X_r)}{\mu_r^2} = \frac{\sum_{l=0}^{r} \binom{n}{r}\binom{r}{l}\binom{n-r}{r-l} p^{2\binom{r}{2}}\left(p^{-\binom{l}{2}}-1\right)}{\binom{n}{r}^2 p^{2\binom{r}{2}}} \leq \sum_{l=2}^{r} \underbrace{\frac{\binom{r}{l}\binom{n-r}{r-l} p^{-\binom{l}{2}}}{\binom{n}{r}}}_{\varepsilon_l}$$

$$\frac{\varepsilon_{l+1}}{\varepsilon_l} = \frac{\binom{r}{l+1}\binom{n-r}{r-l-1} p^{-\binom{l+1}{2}}}{\binom{r}{l}\binom{n-r}{r-l} p^{-\binom{l}{2}}} = \frac{r-l}{l+1}\,\frac{r-l}{n-2r+l+1}\,p^{-l} = \frac{(r-l)^2 p^{-l}}{(l+1)(n-2r+l+1)}$$

$$2 \leq l \leq \tfrac{1}{2}\log_b n \Rightarrow p^{-l} \leq n^{\frac{1}{2}} \Rightarrow \frac{\varepsilon_{l+1}}{\varepsilon_l} \leq \frac{(2\log_b n)^2 n^{\frac{1}{2}}}{3(n-4\log_b n)} \leq n^{-\frac{1}{3}} \quad \left(\text{just slower than } n^{-\frac{1}{2}}\right)$$

$$\tfrac{3}{2}\log_b n \leq l \leq 2\log_b n \Rightarrow p^{-l} \geq n^{\frac{3}{2}} \Rightarrow \frac{\varepsilon_{l+1}}{\varepsilon_l} \geq \frac{n^{3/2}}{2\log_b n\,(n)} = \frac{n^{\frac{1}{2}}}{2\log_b n} \geq n^{\frac{1}{3}}$$

$$\tfrac{1}{2}\log_b n \leq l \leq \tfrac{3}{2}\log_b n \Rightarrow \varepsilon_l = \frac{r(r-1)\cdots(r-l+1)}{l!}\,\frac{(n-r)(n-r-1)\cdots(n-2r+l+1)}{(r-l)!}\cdot\frac{r!}{n(n-1)\cdots(n-r+1)}\,p^{-\binom{l}{2}}$$

$$\leq \frac{r^{2l}}{2^l(\frac{n}{2})^l} p^{-\binom{l}{2}} = \frac{r^{2l}p^{-\frac{1}{2}l(l-1)}}{n^l} = \left(\frac{r^2 p^{-\frac{(l-1)}{2}}}{n}\right)^l$$

$$\leq \left(\frac{r^2 n^{\frac{3}{4}}}{n}\right)^l \leq n^{-\frac{l}{5}} \underset{l\text{ large}}{\leq} n^{-1}$$

n^{-100} true too

$$\text{Then } \sum_{l=2}^{r} \cdots \leq \frac{\varepsilon_2}{1-n^{-\frac{1}{3}}} + \frac{\varepsilon_r}{1-n^{-\frac{1}{3}}} + \frac{\log_b n}{n^{\frac{1}{5}\log_b n}} \to 0 \text{ as } n\to\infty$$

$$\varepsilon_r = \frac{1}{\binom{n}{r}p^{\binom{r}{2}}} = \frac{1}{\mu_r} \to 0 \qquad \to 0$$

$$\mathbb{P}(X_r = 0) \leq \frac{\mathrm{Var}(X_r)}{\mu_r^2} \to 0$$

$$\varepsilon_2 = \frac{\binom{r}{2}\binom{n-r}{r-2}p^{-\binom{2}{2}}}{\binom{n}{r}} \leq \frac{\frac{1}{2}r^4p^{-1}}{n(n-1)} \to 0 \text{ as } n\to\infty$$

$$= O\left(\frac{(\log n)^4}{n^2}\right)$$

$\Rightarrow X_r \geq 1$ whp

4. The Chromatic Number

A k-COLOURING of G is a partition of $V(G)$ into k independent sets.
Equivalently, $\phi: V(G) \to [k]$ st $uv \in E(G) \Rightarrow \phi(u) \neq \phi(v)$.
Vertices of the same colour are not adjacent.
The CHROMATIC NUMBER of G, $\chi(G)$ is the minimum k st G is k-colourable.

$\chi(G) = 2 \quad \equiv \quad G$ is bipartite

$\chi(G) \geq 3 \quad \equiv \quad \exists$ odd cycle

Question: $\chi(G_{n,p})$?

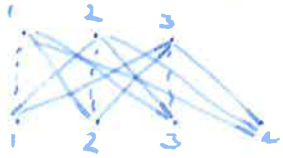
Always: $\chi(G_n) \geq \frac{n}{\mathrm{ind}(G_n)}$ independence number - largest independent set

$\chi(G_n) \geq \alpha(G_n)$ (clique number of complement)

$\overline{G_{n,p}} \sim G_{n,q}$ $q = 1-p$; $d = \frac{1}{q}$

$\mathrm{ind}(G_{n,p}) \leq 2\log_d n$ $\therefore$ whp $\chi(G_{n,p}) \geq \frac{n}{2\log_d n}$ weaker $(\chi(G_{n,p}) \geq (2+o(1))\log_b n)$

Greedy colouring



$\chi_g(G_{n,p}) = \#$ colours used by greedy on $G_{n,p}$

Check: greedy works for some ordering

$\chi_g(G_{n,p}) \geq \chi(G_{n,p})$ $\leftarrow$ worst-case behaviour

Theorem 7 Let $0<p<1$ be fixed, $d = \frac{1}{q} = \frac{1}{1-p}$. Then whp $\chi_g(G_{n,p}) = (1+o(1))\frac{n}{\log_d n}$
$\left(\forall \varepsilon > 0, \text{ whp } (1-\varepsilon)\frac{n}{\log_d n} \leq \chi_g(G_{n,p}) \leq (1+\varepsilon)\frac{n}{\log_d n}\right)$

Proof Aim: given $\varepsilon > 0$, whp $(1-\varepsilon)\frac{n}{\log_d n} \leq \chi_g(G_{n,p}) \leq (1+\varepsilon)\frac{n}{\log_d n}$

"I don't like it when people ask things in lectures, overall for me good lesson"

i) Set $r = \lfloor \log_d n \rfloor$, $r_0 = \lfloor \varepsilon \log_d n \rfloor$; $\frac{n}{r+r_0} \geq \frac{(1-\varepsilon)n}{\log_d n}$. We'll prove $\chi_g(G_{n,p}) \geq \frac{n}{r+r_0}$

Claim $\forall i$, $1 \leq i \leq n$, the probability that the ith colour in greedy has more than $r+r_0$ vertices is $\leq n^{-2}$.

Proof $P(j_1, \dots, j_{r_0})$ for the probability that the $(r+i)^{\text{th}}$ vertex of the ith colour class is exactly j_i for $1 \leq i \leq r_0$. $P(j_1,\dots,j_{r_0}) \leq q^r q^{r+1} \cdots q^{r+r_0-1} = q^{rr_0 + \binom{r_0}{2}} \leq q^{\varepsilon(\log_d n)^2} = n^{-\varepsilon \log_d n}$

$\therefore \mathbb{P}(i^{\text{th}}$ class has $\geq r+r_0$ vertices$) \leq \binom{n}{r_0} n^{-\varepsilon\log_d n} \leq \frac{1}{r_0!} \leq n^{-2}$

Lower-bound follows: with probability $\geq 1 - \frac{1}{n}$ every colour class has $\leq r+r_0$ vertices $(1-\frac{1}{n^2})^n$

$\therefore \chi_g(G_{n,p}) \geq \frac{n}{r+r_0} > (1-\varepsilon)\frac{n}{\log_d n}$.

ii) $k = \left\lfloor \frac{(1+\varepsilon)n}{\log_d n} \right\rfloor$ $\frac{n}{k} < (1-\frac{\varepsilon}{2})\log_d n$

$A = A^{k+1}$ we need $\geq k+1$ colours

$B_{j+1} = B_{j+1}^{k+1}$: $j+1$ is the first vertex getting colour $k+1$

$C(j_1, \dots, j_k)$: up to j we have j_i of colour i $\quad j = \sum j_i, \; j_i \geq 1$

$A = \bigcup_{j \geq k}^{n-1} B_{j+1}$ disjoint $\quad B_{j+1} \subseteq \bigcup_{(j_i)} C(j_1, \dots, j_k)$ smaller vertex, use smaller colours $\quad \sum j_k = j \quad j_i \geq 1$

Take a $\chi(G)$-colouring in which no vertex can be coloured with a lower number without changing the colour of any other vertex. Then a vertex colour j is connected to a vertex colour i $\forall i<j$. Enumerate the vertices as x_i so that $i<j \Rightarrow c(x_i) \leq c(x_j)$: then greedy reproduces the original colouring

$s = \langle \log_d n \rangle$ $\quad s_0 s + s r_0 + s_0 r \leq 1 + r + r_0 = O(\log_d n) \leq \binom{r_0}{2}$ for large n

$s_0 = \langle \varepsilon \log_d n \rangle$ $\quad \Rightarrow \varepsilon (\log_d n)^2 \leq r r_0 + \binom{r_0}{2}$

$\lfloor \varepsilon \log n \rfloor ! \geq \left(\frac{\varepsilon}{2} \log n\right)^{\frac{\varepsilon}{2}\log n} \geq \left(e^{\frac{10}{\varepsilon}}\right)^{\frac{\varepsilon}{2}\log n} \geq n^5 \geq n^1$ for large n (or $\varepsilon \to 0$)

careful

$$\mathbb{P}(B_{j+1} \mid C(j_1,\dots,j_k)) = \prod_{i=1}^{k} (1-q^{j_i}) \le (1-q^{j/k})^k \le e^{-kq^{\frac{j}{k}}} = P_{k,j} \quad [\log(1-q^x) \text{ is } \cancel{\text{convex}} \text{ concave}]$$

(k concave; $1-x \le e^{-x}$; $\uparrow$ in j)

$$\mathbb{P}(B_{j+1}) = \sum_{j_i} \mathbb{P}(B \mid C(j_1,\dots,j_k))\,\mathbb{P}(j_1,\dots,j_k) \le P_{k,j} \underbrace{\sum \mathbb{P}(C(\dots))}_{\le 1} \le P_{k,j}.$$

Also, $P_{k,j}$ is an increasing function of j.

Hence $\mathbb{P}(A) \le \sum_{j=k}^{n-1} \mathbb{P}(B_{j+1}) \le n P_{k,n} = n\exp(-kq^{\frac{n}{k}}) \le n\exp(-n^{\frac{\varepsilon}{3}}) \to 0$

$kq^{\frac{n}{k}} > n^{\frac{\varepsilon}{3}}$?

$kq^{\frac{n}{k}} > kn^{-(1-\frac{\varepsilon}{2})}$

Better? Of course: pick a largest independent set $\to 1$;

—————— " —————— from the remainder $\to \frac{2}{3}$

For $W \subseteq V(G_{n,p})$, $|W| = m$, $G_{n,p}[W] \sim G_{m,p}$

$2\log_d m \sim 2\log_d n$ if $m > \frac{n}{(\log n)^2}$ RUBBISH clearly...

Uses $\sim \frac{n}{2\log n} + \frac{n}{(\log n)^2} \sim \frac{n}{2\log_d n}$ colours

II Concentration of probability

1. Multiplicatively independent r.v.s and martingales

A sequence of r.v.s $(Y_k)_1^n$ is MULTIPLICATIVELY INDEPENDENT if $\forall a_k, b_k$,

$$\mathbb{E}\Big(\prod_1^n (a_k + b_k Y_k)\Big) = \prod_1^n a_k$$

Example (Y_k) independent, $\mathbb{E}(Y_k) = 0 \ \forall k$

Prime example: (Y_k) is a MARTINGALE DIFFERENCE SEQUENCE

$(X_k)_0^n$ martingales $Y_k = X_k - X_{k-1}$

2. Basic inequalities

The SPREAD of a r.v. X is $\min\{b-a : \mathbb{P}(X \in [a,b]) = 1\}$.

First: (Y_k), all with same spread. $=1$

Theorem 1 Let $(Y_k)_1^n$ be mult. indep, $-p_k \le Y_k \le q_k = 1-p_k \ \forall k$, $p_k, q_k > 0$ (and <1)

Set $X = \sum Y_k$; $p = \frac{1}{n}\sum p_k$, $q = \frac{1}{n}\sum q_k$.

Then for $h \in \mathbb{R}$, $\mathbb{E}(e^{hX}) \le e^{-nph}\prod_1^n (q_k + p_k e^h) \le e^{-nph}(q+pe^h)^n \le e^{nh^2/8}$

Proof e^{ht} is a convex function of t, so if $-p_k \le t \le q_k$ then $e^{ht} \le (t+p_k)e^{q_k h} + (q_k - t)e^{-p_k h}$

$e^{hY_k} \le (Y_k + p_k)e^{q_k h} + (q_k - Y_k)e^{-p_k h}$ a.s.

$$= p_k e^{q_k h} + q_k e^{-p_k h} + Y_k(e^{q_k h} - e^{-p_k h})$$

$$\therefore \mathbb{E}(e^{hX}) \le \mathbb{E}\prod(\quad\downarrow\quad)$$

$$= \prod(q_k e^{-p_k h} + p_k e^{q_k h}) = e^{-h\sum p_k}\prod(q_k + p_k e^h)$$

AM-GM $\Rightarrow \le e^{-nph}(q+e^h)^n$

$q + pe^h \le \dots$? IA Ex.

$(q+pe^h)^n \le e^{pnh} e^{ch^2 n}$ c as small as possible

omitted

$$k2^{\frac{n}{k}} > \frac{n}{2\log_d n}\, 2^{\frac{\log_d n}{1+\frac{\varepsilon}{2}}} = \frac{n^{1-(1+\frac{\varepsilon}{2})^{-1}}}{2\log_d n} > \frac{n^{1-(1-\frac{\varepsilon}{4})}}{2\log_d n} > n^{\frac{\varepsilon}{5}}$$

smaller k

$-p \qquad t \qquad q \qquad p+q=1$

$$t = q(p+t) - p(q-t) \Rightarrow e^{ht} \le (p+t)e^{qh} + (q-t)e^{-ph} = pe^{qh} + qe^{-ph} + t(e^{qh} - e^{-ph})$$

$$\Rightarrow \mathbb{E}(e^{hX}) \le \mathbb{E}\left(\prod\left((p_k e^{q_k h} + q_k e^{-p_k h}) + (e^{q_k h} - e^{-p_k h})Y_k\right)\right) = \prod(p_k e^{q_k h} + q_k e^{-p_k h})$$

$$= e^{-nph}\prod(q_k + p_k e^{h}) \le e^{-nph}(q+pe^{h})^n \text{ by AM-GM}$$

Write $\phi(t) = \log(q+pe^t)$. $\phi'(t) = \frac{pe^t}{q+pe^t} = \frac{p}{qe^{-t}+p}$ $\qquad \phi(0)=0 \qquad \phi'(0)=p$

$$\phi''(t) = \frac{-p(-qe^{-t})}{(qe^{-t}+p)^2} \le \frac{1}{4}$$

By Taylor's theorem $\phi(h) \le ph + \frac{h^2}{8} \Rightarrow q+pe^h \le e^{ph+\frac{h^2}{8}}$

$$\Rightarrow e^{-nph}(q+pe^h)^n \le e^{nh^2/8}$$

Remains: $e^{-ph}(q+pe^h) \le e^{h^2/8}$; $q+pe^h \le e^{ph+h^2/8}$

$\phi(t) = \log(q+pe^t)$ $\quad\phi(0)=0\quad$ Apply Taylor's theorem with error term:

$\phi'(t) = \frac{pe^t}{q+pe^t} = \frac{p}{qe^{-t}+p}$ $\quad\phi'(0)=p\quad$ $\phi(h) = ph + \frac{\phi''(t)h^2}{2}$ for some $t\in[0,h]$ or $[h,0]$

$\phi''(t) = \frac{pqe^{-t}}{(qe^{-t}+p)^2}$ $\quad\phi''(t)\le \frac14\quad$ $\le ph + \frac{h^2}{8}$ □

cf. T; Chernoff's inequality

Sup binomial: $\mathbb{P}(S_{n,p} \ge (p+t)n) \le \left[\left(\frac{p}{p+t}\right)^{p+t}\left(\frac{q}{q-t}\right)^{q-t}\right]^n$ $\quad 0<t<q=1-p$

Theorem 2 Let (Y_k), etc., be as in Theorem 1.

Then $\mathbb{P}\left(\sum_1^n Y_k \ge tn\right) \le \left[\left(\frac{p}{p+t}\right)^{p+t}\left(\frac{q}{q-t}\right)^{q-t}\right]^n$ for $0<t<q$

Proof $X=\sum_1^n Y_k$; $h>0$

$\mathbb{P}(X\ge tn)e^{htn} \le \mathbb{E}(e^{hX}) \le e^{-nph}(q+pe^{h})^n$

$\mathbb{P}(X\ge tn) \le e^{-nh(p+t)}(q+pe^h)^n$

$\phi(h) = e^{-h(p+t)}(q+pe^h)$

$\phi'(h)e^{h(p+t)} = -(p+t)(q+pe^h)+pe^h$

To minimise $\phi(h)$ we'll choose h st $pe^h = (p+t)(q+pe^h)$ — OK if we miss minimum

$e^h = \frac{q(p+t)}{p(q-t)} > 1$ so $\exists h$

$q+pe^h = q\left[\frac{p+t}{q-t}+1\right] = \frac{q}{q-t}$

With that h, $\phi(h) = \left(\frac{p(q-t)}{q(p+t)}\right)^{p+t}\left(\frac{q}{q-t}\right) = \left(\frac{p}{p+t}\right)^{p+t}\left(\frac{q}{q-t}\right)^{q-t}$. □

Lemma 3 For $0<t<q$, $\left(\frac{p}{p+t}\right)^{p+t}\left(\frac{q}{q-t}\right)^{q-t} \le e^{-2t^2}$

Proof Set $\phi(t) = \log(\mathrm{LHS}(t))$. Need $\phi(t)\le -2t^2$. $\quad\mathbb{P}(X\ge tn)\le e^{-2nt^2}$

$\phi(0)=0$; $\phi'(0)=0$; $\phi''(t)\le -4$. Done by Taylor's theorem

Theorem 4 Let (Y_k) be a mult. ind. seq. with each Y_k having spread at most 1. Then for $t>0$

$$\mathbb{P}\left(\sum Y_k \ge tn\right) \le e^{-2t^2 n}$$

Proof Theorem 2 and Lemma 3.

$u>0$: $\mathbb{P}\left(\sum Y_k\ge u\right) \le e^{-2u^2/n}$

Also from Theorem 1 using the last bound. $X=\sum Y_k$

$\mathbb{P}(X\ge tn)e^{tnh} \le \mathbb{E}(e^{hX}) \le e^{nh^2/8}$

$\mathbb{P}(X\ge tn) \le e^{-tnh+\frac{nh^2}{8}} = e^{-nh(t-\frac h8)} = e^{-4nt(t-\frac t2)} = e^{-2t^2n}$ $\quad t=h/4$; $h=4t$

BACK TO 1. Martingales and, ...

Martingales: $(\Omega,\mathcal F,\mathbb P)$; filtration $\overset{=\{\emptyset,\Omega\}}{\mathcal F_0}\subset \mathcal F_1\subset\dots\subset\mathcal F_n=\mathcal F$

X r.v. on Ω ($\mathcal F$-measurable)

The MARTINGALE defined by the filtration $(\mathcal F_k)_0^n$ and X is $(X_k)_0^n$ where $X_k = \mathbb E(X|\mathcal F_k)$ is the orthogonal projection of X into the space of $\mathcal F_k$-measurable functions (all in L_2).

$$\mathbb E(UX) = \mathbb E(U\,\mathbb E(X|\mathcal F_k))\ \forall\, U\ \mathcal F_k \text{ measurable}$$

$\mathcal F$ on Ω finite; $\mathcal F \longleftrightarrow \{\text{atoms}\}$ = partition $\mathcal P$

$\mathcal F_0\subset\dots\subset\mathcal F_n = \mathcal F = \mathcal P(\Omega)$

$\mathcal P_n = \{\text{singletons}\}$

Exercise: what is X_k?

$0 < u < q$

$\phi(u) = (p+u)(\log p - \log(p+u)) + (q-u)(\log q - \log(q-u))$ $\quad \phi(0) = 0$

$\phi'(u) = \log p - \log(p+u) - \log q + \log(q-u) - 1 + 1$ $\quad \phi'(0) = 0$

$\phi''(u) = -\frac{1}{p+u} - \frac{1}{q-u} = \frac{-1}{(p+u)(q-u)} = \frac{-1}{(q-u)(1-(q-u))} \leq -4$

so $\phi(t) \leq -2t^2$

$\Rightarrow \mathbb{P}(\sum Y_k \geq tn) \leq e^{-2t^2 n}$

$\mathbb{P}(\sum Y_k \geq u) \leq e^{-\frac{2u^2}{n}}$

Conditional expectation $\mathcal{F}' \subset \mathcal{F}$ X $\mathcal{F}$-measurable

$L_2(\mathcal{F}) = L_2(\Omega, \mathcal{F}, \mathbb{P})$; $L_2(\mathcal{F}')$ $L_2(\mathcal{F}') \subset L_2(\mathcal{F})$

For $X \in L_2(\mathcal{F})$, $\mathbb{E}(X|\mathcal{F}')$ is the orthogonal projection of X into $L_2(\mathcal{F}')$

Equivalently: $x \in H \supset H_0$ $P_{H_0}(x) = y$ if $y \in H_0$ and $\forall z \in H_0$ $\langle x, z\rangle = \langle y, z\rangle$.

Do not need all of H_0, only a set S with $\overline{\lim S} = H_0$

Thus for $\mathcal{F}$-measurable X, $\mathbb{E}(X|\mathcal{F}')$ is the unique $\mathcal{F}'$-measurable function st

$$\mathbb{E}(UX) = \mathbb{E}(U\,\mathbb{E}(X|\mathcal{F}')) \ \forall\ \mathcal{F}'\text{-measurable } U, \text{ provided LHS exists}$$

suffices to take $\{\mathbb{1}_S : S \in \mathcal{F}'\}$. "$S =$"

If Ω is finite then everything is elem. to-each. $(\mathcal{F}_n)_0^n \longleftrightarrow (\mathcal{P}_n)_0^n$ $\mathcal{P}_n$ partn of Ω into atoms of $\mathcal{F}_n$

$\mathcal{F}_n$-measurable $\Leftrightarrow$ constant on sets in $\mathcal{P}_n$

$\mathbb{E}(X|\mathcal{F}_n) = X_n$: $X_n(\omega) = \mathbb{E}(X \mathbb{1}_{P_i}) / \mathbb{P}(P_i)$

where $\omega \in P_i \in \mathcal{P}_n$.

We'll take $\mathcal{F}_n = \mathbb{P}(\Omega)$; $\mathcal{P}_n = \{\text{singletons}\}$.

$X = (X_n)_0^n$ martingale difference sequence $(Y_n)_1^n$ $Y_n = X_n - X_{n-1}$ $\Leftrightarrow$

Then Y_n is $\mathcal{F}_n$-measurable and $\mathbb{E}(Y_n | \mathcal{F}_{n-1}) = 0$ $= X_{n-1} - X_{n-1}$ $\Rightarrow$ MI as con.... polynomial

Also, $(Y_n)_1^n$ is mult. ind. Indeed, if $1 \le k_1 < k_2 < \dots < k_\ell \le n$ then $\mathbb{E}\left(\prod_1^\ell Y_{k_i}\right) = 0$. Why?

$$\mathbb{E}\left(\prod_1^\ell Y_{k_i}\right) = \cancel{\mathbb{E}\left(\prod_1^{\ell-1} Y_{k_i}\,\mathbb{E}(Y_{k_\ell}\right)} \ \mathbb{E}\left(\mathbb{E}\left(\prod_1^\ell Y_{k_i} \,\middle|\, \mathcal{F}_{k_\ell - 1}\right)\right) = \mathbb{E}\left(\prod_1^{\ell-1} Y_{k_i}\, \mathbb{E}(Y_{k_\ell} | \mathcal{F}_{k_\ell - 1})\right) = 0$$

all $k_i - 1$ $\mathcal{S}$ measurable

Conversely, from $(Y_n)_1^n$ we obtain a martingale. $\Leftrightarrow$

Recall: Theorem 4 $(Y_n)_1^n$ is mult ind^n, each Y_n has spread $\le \delta$, then for $u > 0$

$$\mathbb{P}\left(\sum_1^n Y_n > u\right) \le e^{-2u^2/n\delta^2}$$

SPECIAL CASE Corollary 5 $X = (X_n)_0^n$ is a martingale with difference sequence bounded by δ in modulus

then $\mathbb{P}(X_n - \mathbb{E}(X) \ge u) \le e^{-u^2/2n\delta^2}$

3. Martingale inequality theory

$X \leftrightarrow X_n$

Aim: Improve Theorem 4 and Corollary 5. We shall not demand uniformity.

Theorem 6 Let $(Y_n)_1^n$ be MI with $|Y_n| \le s_n$ $S = S_n = \sum_1^n s_n^2$. Then for $u > 0$, $\mathbb{P}\left(\sum_1^n Y_n \ge u\right) \le e^{-u^2/2S}$

Proof e^{tx} is convex on $\mathbb{R}$ (as a function of x); in particular, for $-1 \le t \le 1$, $e^{tx} \le \frac{1+t}{2}e^{x} + \frac{1-t}{2}e^{-x}$

$e^{tY_n} = e^{(ts_n)(Y_n/s_n)} \le \cosh(ts_n) + \frac{Y_n}{s_n}\sinh(ts_n)$ $= \cosh x + t \sinh x$

$$\mathbb{E}(e^{t\sum Y_n}) \le \mathbb{E}\left(\prod \left(\cosh(ts_n) + \frac{Y_n}{s_n}\sinh(ts_n)\right)\right) = \prod_1^n \cosh(ts_n)$$

$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots \le e^{x^2/2}$ $\le e^{\frac{t^2 S}{2}}$

$$e^{tu}\,\mathbb{P}\left(\sum Y_n \ge u\right) \le \mathbb{E}(e^{t\sum Y_n}) \le e^{t^2\frac{S}{2}}$$

$$\mathbb{P}\left(\sum Y_n \ge u\right) \le e^{-t(u - \frac{tS}{2})} \qquad u = tS \qquad t = u/S$$

$$\mathbb{P}\left(\sum Y_n \ge u\right) \le e^{-\frac{u}{S}\cdot\frac{u}{2}} = e^{-u^2/2S} \quad \square$$

Special case:

Corollary 7 Let $X = (X_n)_0^n$ be a martingale with difference sequence $(Y_n)_1^n$ $|Y_n| \le s_n$ $\forall n$

Then $\mathbb{P}(X_n - \mathbb{E}(X_n) \ge u) \le e^{-u^2/2S}$ $S = \sum_1^n s_n^2$

$$\mathbb{P}(|X_n - \mathbb{E}X| \ge u) \le 2e^{-u^2/2S}$$

Classical Hoeffding–Azuma inequality

$U=1$ gives $\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|\mathcal{F}_n))$

Y_n $\mathcal{F}_n$-measurable, $\mathbb{E}(Y_n|\mathcal{F}_{n-1}) = 0$ then $X_n = \sum_1^n Y_n + c$ is a martingale $X_0 = c$

Let $X=(X_n)_0^n$ be a martingale on a filtration $(\mathcal{F}_n)_0^n$; $(Y_n)_1^n$ difference sequence.

The spread of Y_n on $\mathcal{F}_{n-1}$, $s(Y_n,\mathcal{F}_{n-1})$, is the minimal s s.t. $\exists$ $\mathcal{F}_{n-1}$-measurable function f_{n-1} st $f_{n-1}\le Y_n \le f_{n-1}+s$ (equivalently st $f_{n-1}\le X_n \le f_{n-1}+s$) – this is the maximum of the spread on the atoms of $\mathcal{F}_{n-1}$.

Applying Theorem 1 on atoms gives $\mathbb{E}(e^{hY_n}\mid\mathcal{F}_{n-1})(\omega)=\mathbb{E}_P(e^{hY_n}) \le e^{(hs_n)^2/8}$ where $\mathbb{P}_P(A)=\mathbb{P}(A)/\mathbb{P}(P)$ for $A\subseteq P$, $\omega\in P\in\mathcal{P}_{n-1}$, $s_k=s(Y_k,\mathcal{F}_{k-1})$

<u>Lemma 8</u> Take $S_n=\sum_1^n s_k^2$, $X_0=0$ $(=\mathbb{E}(X))$

Then $\mathbb{E}(e^{hX_n})\le e^{h^2 S_n/8}$

<u>Proof</u> By induction on n. $n=0$ ✓ (we also know for $n=1$). Induction step from $n-1$ to n:

$$\mathbb{E}(e^{hX_n})=\mathbb{E}(e^{hX_{n-1}}e^{hY_n})=\mathbb{E}(e^{hX_{n-1}}\mathbb{E}(e^{hY_n}\mid\mathcal{F}_{n-1}))\le\mathbb{E}(e^{hX_{n-1}})e^{h^2s_n^2/8}\le e^{h^2S_n/8}$$ by the induction hypothesis. □

We call S_n the <u>SIZE</u> of X.

<u>Theorem 9</u> Let $(X_k)_0^n$ be a martingale of size $\le S_n$. Then $\forall u>0$ $\mathbb{P}(X_n-\mathbb{E}(X_n)\ge u)\le e^{-2u^2/S_n}$.

2's in location of use $|Y_i|\le c_i$

<u>Proof</u> wlog $\mathbb{E}(X_n)=X_0=0$ (this doesn't change the spreads). For $u>0$,

$$\mathbb{P}(X_n\ge u)e^{hu}\le\mathbb{E}(e^{hX_n})\le e^{h^2S_n/8} \Rightarrow \mathbb{P}(X_n\ge u)\le e^{-h(u-\frac{hS_n}{8})}$$ put $u=\frac{hS_n}{4}$, $h=\frac{4u}{S_n}$

$$= e^{-\frac{4u}{S_n}\cdot\frac{u}{2}} = e^{-\frac{2u^2}{S_n}}.\ \square$$

<u>Theorem 10</u> Let $Z_1,\dots,Z_n$ be independent r.v.s taking values in finite sets $\Omega_1,\dots,\Omega_n$ respectively. Let $f:\Omega=\Omega_1\times\dots\times\Omega_n\to\mathbb{R}$ be st $|f(\omega)-f(\omega')|\le c_k$ whenever ω,ω' differ only in the k^{th} coordinate. Then $Z=f(Z_1,\dots,Z_n)$ is concentrated about its mean: $\mathbb{P}(Z\ge\mathbb{E}(Z)+u)\le e^{-2u^2/S_n}$, where $S_n=\sum_1^n c_k^2$.

<u>Proof</u> Take $\mathcal{P}_k = \cancel{\{A\times\Omega_{k+1}\times\dots\times\Omega_n : A\in\Omega_1\times\dots\times\Omega_k\}} = \{\{a_1\}\times\dots\times\{a_k\}\times\Omega_{k+1}\times\dots\times\Omega_n : a_j\in\Omega_j\}$,

$X_k=\mathbb{E}(Z\mid\mathcal{F}_k)$. $s(X_k,\mathcal{F}_{k-1})\le c_k$, so the result follows from Theorem 9. □ Independence relevant?

<u>III Applications of Martingale Theory</u>

<u>i. Symmetric group</u>

Maurey (1984)

S_n symmetric group: $|S_n|=n!$

Probability space: $\mathbb{P}(\{\pi\})=\frac{1}{n!}$

HAMMING DISTANCE

$d(\pi,\rho)=|\{i:\pi(i)\ne\rho(i)\}|$

$f:S_n\to\mathbb{R}$ is LIPSCHITZ with constant c if

$|f(\pi)-f(\rho)|\le c\,d(\pi,\rho)$ all functions Lipschitz

<u>Theorem 1</u>

Let $f:S_n\to\mathbb{R}$ be Lipschitz with constant c. Then $\mathbb{P}(f-\mathbb{E}(f)\ge u)\le e^{\frac{-u^2}{2nc^2}}$

<u>Proof</u> wlog $c=1$ (c and f move a scale together). Take $X=f$ and define $\equiv_k$ on S_n by $\pi\equiv_k\rho$ if $\pi(i)=\rho(i)$ for $i\le k$. This gives a filtration $(\mathcal{F}_k)_0^n$ (with atoms equivalence classes) and corresponding $(X_k)_0^n$.

Need $A,B\in\mathcal{P}_k$, $A\cup B\subseteq C\in\mathcal{P}_{k-1}$ wlog $A=\{\pi:\pi(i)=i,\ i\le k-1;\ \pi(k)=l\}$

$B=\{\pi:\pi(i)=i,\ i\le k-1;\ \pi(k)=m\}$

$C=\{\pi:\pi(i)=i,\ i\le k-1\}$

e^{-2u^2/nc^2} with d^*

$\mathcal{P}_k = \{\{x_1\} \times \cdots \times \{x_k\} \times \Omega_{k+1} \times \cdots \times \Omega_n : x_i \in \Omega_i\}$

Take $A, B \in \mathcal{P}_k$, $A \cup B \subseteq C \in \mathcal{P}_{k-1}$

$A = \{x_1\} \times \cdots \times \{x_{k-1}\} \times \{a\} \times \Omega_{k+1} \times \cdots \times \Omega_n$

$B = \quad '' \quad\quad '' \quad \{b\} \quad '' \quad\quad ''$

$C = \{x_1\} \times \cdots \times \{x_{k-1}\} \times \Omega_k \times \cdots \times \Omega_n$

$X_k = \mathbb{E}(X \mid \mathcal{G}_k)$

independence

For $\alpha \in A$, $X_k(\alpha) = \dfrac{\mathbb{E}\, X(\alpha) \mathbb{1}_A}{\mathbb{P}(A)} = \displaystyle\sum_{y_{k+1}, \dots, y_n} \frac{f(x_1, \dots, x_{k-1}, a, y_{k+1}, \dots, y_n) \mathbb{P}(\{x_1\}) \cdots \mathbb{P}(\{x_{k-1}\}) \mathbb{P}(\{a\}) \mathbb{P}(\{y_{k+1}\}) \cdots \mathbb{P}(\{y_n\})}{\mathbb{P}(\{x_1\}) \cdots \mathbb{P}(\{x_{k-1}\}) \mathbb{P}(\{a\})}$

$= \displaystyle\sum_{y_{k+1}, \dots, y_n} f(x_1, \dots, x_{k-1}, a, y_{k+1}, \dots, y_n) \mathbb{P}(\{y_{k+1}\}) \cdots \mathbb{P}(\{y_n\})$

$\Rightarrow |X_k(\alpha) - X_k(\beta)| \leq \displaystyle\sum_{y_{k+1}, \dots, y_n} \mathbb{P}(\{y_{k+1}\}) \cdots \mathbb{P}(\{y_n\}) \left| f(x_1, \dots, x_{k-1}, a, y_{k+1}, \dots, y_n) - f(x_1, \dots, x_{k-1}, b, y_{k+1}, \dots, y_n) \right|$

$\leq \displaystyle\sum_{y_{k+1}, \dots, y_n} \mathbb{P}(\{y_{k+1}\}) \cdots \mathbb{P}(\{y_n\})\, c_k = c_k$

Proof $\pi \equiv_k \rho$ if $\pi(i)=\rho(i)\ \forall i \le k$ $(\equiv_k) \leftrightarrow (\mathcal{P}_k)_0^{n-1} \leftrightarrow (\mathcal{F}_k)$; $X=f$; (X_k)

$A, B \in \mathcal{P}_k$, $A \cup B \subseteq C \in \mathcal{P}_{k-1}$ wlog $A=\{\pi : \pi(i)=i \ \ i \le k-1, \pi(k)=l\}$

Define $\phi : A \to B$ by $\phi(\pi)=\tau\pi$ $\qquad B=\{\pi : \pi(i)=i \ \ i\le k-1,\ \pi(k)=m\}$

$\tau=(l\ m)$ $\qquad C=\{\pi : \pi(i)=i,\ i \le k-1\}$

$$\phi(\pi)(i)=\begin{cases} m & \pi(i)=l \\ l & \pi(i)=m \\ \pi(i) & \text{otherwise}\end{cases}$$

$d^*=1$

Then $d(\pi,\phi(\pi))=2$, so our martingale has size $\le 4(n-1)$ $\quad (*)\ \le n$

Hence $\mathbb{P}(X-\mathbb{E}X \ge u) \le e^{-2u^2/4(n-1)} \le e^{-u^2/2n}$ $\square$

$A \subseteq S_n$; t-NEIGHBOURHOOD of A: $A_t=\{x : d(x,A) \le t\}$; $d(x,A)=\inf\{d(x,y) : y \in A\}$

sometimes $<$ fine. Borel measurability

Theorem 2 Let $A \subseteq S_n$; $\mathbb{P}(A)=\alpha>0$. If $t=(x+1)(2n\log(\frac{1}{\alpha}))^{\frac12}$, $x\ge 0$, then $\mathbb{P}(A_t) \ge 1-\alpha^{x^2}$

$d(\pi,A) \le d(\pi,\rho)+d(\rho,A)$ and vice versa

Proof $f(\pi)=d(\pi,A)$, Lipschitz with constant 1 $\quad \mu=\mathbb{E}f$

$\alpha=\mathbb{P}(f\overset{\le}{=}0)=\mathbb{P}(-f+\mu\ge\mu) \le e^{-\mu^2/2n}$ hence $\mu \le (2n\log(\frac{1}{\alpha}))^{\frac12}$

Therefore $\mathbb{P}(A_t^c)=\mathbb{P}(f-\mu>t-\mu) < e^{-2(t-\mu)^2/4n} \le \alpha^{x^2}$. $\square$

Corollary 3 $A, B \subseteq S_n$; $\mathbb{P}(A)=\alpha$, $\mathbb{P}(B)=\beta$, $d=d(A,B)$ $\ \inf_{a\in A} d(a,B)$

$\Rightarrow d(A,B) \le \left(\log(\frac{1}{\alpha})^{\frac12}+\log(\frac{1}{\beta})^{\frac12}\right)\sqrt{2n}$

$\Rightarrow \min\{\alpha,\beta\} \le e^{-d^2/8n}$ $\ \le e^{-d^{*2}/8n}$

Another natural metric on S_n: $d^*(\pi,\rho)=$ minimal # of transpositions τ_i st $\rho=\pi\tau_1\ldots\tau_k$.

$d^*(\pi,\rho)+1 \le d(\pi,\rho) \le 2d^*(\pi,\rho)$ all touched once 2

First lives pairs touched 2 times, rest ≥ 1

Theorem 4 ... for d^*

Proof (S_n, d^*)

2. The chromatic number of random graphs

Theorem 5 Let $0<p=p(n)<1$; set $X_n=\chi(G_{n,p})$.

Then $\mathbb{P}(|X_n-\mathbb{E}(X_n)|\ge t) \le 2e^{-2t^2/n}$

In particular, whp X_n is within $\omega(n)\sqrt{n}$ of its mean, if $\omega(n)\to\infty$

Proof Take the vertex-revealing martingale. $G \equiv_k H$ if $G[k]=H[k]$. $X_n \to (X_k)_0^n$

Spread of Y_k wrt $\mathcal{P}_{k-1}^{\le 1}$, size $\le n$ $\square$

It is easily checked that size $\le n$ ← warming phase (official)

$\updownarrow$ G' preserving bijection

χ differs by ≤ 1

Expectation works by independence, which is in essence Proof 2

$$t > \left(\log\left(\tfrac{1}{\alpha}\right)^{\frac{1}{2}} + \log\left(\tfrac{1}{\beta}\right)^{\frac{1}{2}}\right)\sqrt{2n} = \left(2n\log\left(\tfrac{1}{\alpha}\right)\right)^{\frac{1}{2}}\left(1 + \frac{\log\left(\frac{1}{\beta}\right)^{\frac{1}{2}}}{\log\left(\frac{1}{\alpha}\right)^{\frac{1}{2}}}\right)$$

$$\Rightarrow \mathbb{P}(A_t) > 1 - \alpha^{\frac{\log\left(\frac{1}{\beta}\right)}{\log\left(\frac{1}{\alpha}\right)}} = 1 - e^{-\log\frac{1}{\beta}} = 1 - \beta$$

$$\Rightarrow \mathbb{P}(A_t) + \mathbb{P}(B) > 1$$

$$\Rightarrow A_t \cap B \neq \phi$$

$$\Rightarrow d \leq t$$

Take the infimum over all t to get the result

$$d \leq \left(2n\log\left(\tfrac{1}{\alpha}\right)\right)^{\frac{1}{2}}\left(1 + \frac{\log\left(\frac{1}{\beta}\right)^{\frac{1}{2}}}{\log\left(\frac{1}{\alpha}\right)^{\frac{1}{2}}}\right) \leq 2\left(2n\log\left(\tfrac{1}{\alpha}\right)\right)^{\frac{1}{2}} \qquad \alpha \leq \beta$$

$$\Rightarrow e^{-d^2/8n} \geq \alpha$$

and similarly if $\beta \leq \alpha$

2nd proof Ω_k = collection of sets of edges ending in vertex $k+1$; $|\Omega_k| = 2^k$

$z_1, \dots, z_{n-1}$ independent r.v.s with Z_k taking values in Ω_k; $\mathbb{P}(Z_k$ is a certain set of l edges$) = p^l(1-p)^{k-l}$

$\Omega = \Omega_1 \times \dots \times \Omega_{n-1}$; $\omega = (\omega_i)_1^{n-1}$. Then $\omega \Leftrightarrow G_\omega$ (in the obvious way). Then $G_Z \sim G_{n,p}$ where $Z = (z_n)_1^{n-1}$.

Define $X(\omega) = \chi(G_\omega)$: if ω, ω' differ in one coordinate then $|X(\omega) - X(\omega')| \le 1$. Claim follows by result in Ch II.3. $\square$

Same proof

There are two purposes of applying martingale concentration inequalities.

1. Can show that a r.v. is concentrated on a very short interval
2. ——— " ——— is highly concentrated on a not-too-long interval

In particular, if $\omega(n) \to \infty$ then whp $\chi(G_{n,p})$ is within $\omega(n)\sqrt{n}$ of $\mathbb{E}(X)$. $e^{-\omega(n)}$ of failure

$\chi(G_{n,p})$ is ?? whp. Know: for $0<p<1$ fixed, whp $\frac{n}{2\log_d n} \le \chi(G_{n,p}) \le (1+o(1))\frac{n}{\log_d n}$

where $d = \frac{1}{q} = \frac{1}{1-p}$

could happen: $n = 2^{2^{2k}}$ $\chi \sim$ lower bound

$n = 2^{2^{2k+1}}$ $\chi \sim$ upper bound

Edge-revealing martingale

$e_1, \dots, e_N$; $N = \binom{n}{2}$; $G \equiv_k H$ if they agree up to e_k : for $i \le k$ $e_i \in E(G) \Leftrightarrow e_i \in E(H)$

X r.v. $\to$ martingale $(X_k)_0^N$ 'finest'

$\Omega = \{0,1\}^N = \{0,1\}^{E(n)^{(2)}}$; $\omega \in \Omega \Leftrightarrow G_\omega$: $e_i \in E(G_\omega) \Leftrightarrow \omega_i = 1$

$(Z_i)_1^N$ i.i.d. $Be(p)$; $G_Z \sim G_{n,p}$

If f is a graph parameter st $|f(G_\omega) - f(G_{\omega'})| \le c_k$ whenever $\omega_i = \omega_i'$ $\forall i \ne k$, then

$$\mathbb{P}(|f - \mathbb{E}f| \ge t) \le 2e^{-2t^2/S} \qquad S = \sum_1^N c_k^2$$

If f is integer-valued then $S \ge N \sim \frac{n^2}{2}$ (if any edge can change it)

To pin down $\mathbb{E}(\chi(G_{n,p}))$ or $\chi(G_{n,p})$, we'll run the simple algorithm we already discussed

$G_1 = G_{n,p}$ Find maximal independent set I_1, colour it 1 , $G_2 = G_1 - I_1$

G_2 in G_2, I_2 2 $G_3 = G_2 - I_2$

Hope: G_k has an independent set $\sim$ as the corresponding binomial random graph $\sim$ about as large $\sim$ as # independent vertices of $G_{n,p}$; $2\log_d n$

$\sim 2\log_d n$

$$m = \frac{n}{(\log_d n)^2}$$

$$\chi \le \frac{n}{r} + \frac{n}{(\log_d n)^2} \sim \frac{n}{2\log_d n}$$

It would work if for every set of m vertices our graph contains r independent vertices.

i.e. $\exists r = r(n) \sim \frac{n}{2\log_d n}$ st whp every set of $m \sim \frac{n}{(\log_d n)^2}$ vertices contains r independent vertices

Lemma 6. Let $0<p<1$; $X_r(G) = \#$ r-cliques; suppose $\mu_r = \mathbb{E}_{n,p}(X_r) \le n^{9/5}$.
Then $\mathbb{P}(X_r = 0) \le e^{-3\mu_r^2/n^2}$ $r \sim 2\log_b n$

Proof. We'll apply the edge-revealing martingale; more precisely, its variant: $Z_1,\dots,Z_N$ independent $Be(p)$, G_ω, $G_Z \sim G_{n,p}$. Would like $f(\omega) = f(G_\omega)$ st if ω and ω' differ in one coordinate then $|f(\omega)-f(\omega')|$ is small.

$f(\omega) = X_r(\omega)$ $\binom{n-2}{r-2}$ ※   Let $\tilde{X}_r(G)$ be the number of K_rs in G none of which shares an edge with another K_r. NOT max #-disjoint

Now $f(\omega) = \tilde{X}_r(G_\omega)$ does change by $\le 1^*$. Hence, with $\nu_r = \mathbb{E}_{n,p}(\tilde{X}_r)$ *Could change by $\binom{r}{2}$

$$\mathbb{P}(X_r=0) \le \mathbb{P}(\tilde{X}_r = 0) \le e^{-2\nu_r^2/N} \le e^{-4\nu_r^2/n^2} \qquad \left[N = \binom{n}{2}\right]$$

Prove ν_r is about μ_r. Let $W = [r]$; $E = \{G_{n,p}[W] = K_r\}$, $F = \{\exists\, K_r$ with at least 2 and at most $r-1$ vertices in $W\}$.

Then $\nu_r = \binom{n}{r}\mathbb{P}(E\setminus F) = \binom{n}{r}(\mathbb{P}(E) - \mathbb{P}(E\cap F)) = \binom{n}{r}\mathbb{P}(E)(1-\mathbb{P}(F|E)) = \mu_r(1-\mathbb{P}(F|E))$

$$\mathbb{P}(F|E) \le \sum_{\ell=2}^{r-1}\binom{r}{\ell}\binom{n-r}{r-\ell}p^{\binom{r}{2}-\binom{\ell}{2}} = \mu_r \sum_2 \varepsilon_\ell$$

But we know from ch. I.3 that this is $\le n^{-\frac{1}{6}}$. In particular, $\nu_r \ge \frac{\sqrt{3}}{2}\mu_r$. □   Result true, but don't use this method

Theorem 7. Let $0<p<1$ be fixed. Then whp $\chi(G_{n,p}) = (1+o(1))\frac{n}{2\log_d n}$, $d = \frac{1}{q} = \frac{1}{1-p}$.

Proof $\ge$ ✓ Need: $\le$. Let $m = \left\lfloor \frac{n}{(\log n)^2}\right\rfloor$ and let $r = r(n)$ be maximal st $\mathbb{E}_{m,p}(Y_r) \ge m^{9/5}$ where $Y_r(G) = \#$ independent r-sets. Then $r \sim 2\log_d n$. We claim that whp $\chi(G_{n,p}) \le \frac{n}{r} + m$. (*)

(*) holds if every set of m vertices contains an independent r-set.

wlog $\mathbb{E}_{m,p}(Y_r) = m^{9/5}$. By Lemma 6, ~~[illegible]~~

$\mathbb{P}$(a fixed m-set does not contain r independent vertices) $\le e^{-3n^{18/5}/n} \le e^{-n^{\frac{3}{2}}}$

$\mathbb{P}(\text{error}) \le \binom{n}{m} e^{-n^{3/2}} \le 2^n e^{-n^{\frac{3}{2}}} = o(1)$

IV. Talagrand's inequality and its applications

1. The inequality

$(\Omega, \mathbb{P})$ probability space $\Omega^n = \Omega \times \dots \times \Omega$ multiple conflation

$x, y \in \Omega^n$ $\{i : x_i \neq y_i\}$ similar $\sigma_x(y) = \mathbb{1}_{x_i \neq y_i}$ $\sigma_x(y) \in \{0,1\}^n$ lifted

$x \in \Omega^n$ $\emptyset \neq A \subseteq \Omega^n$ $U(x,A)$ = up-set (in cube $\{0,1\}^n$) generated by $\sigma_x(A)$

$\sigma_x(x) = \underline{0}$ $\underline{1} \in U(x,A)$ $[0,1]^n$

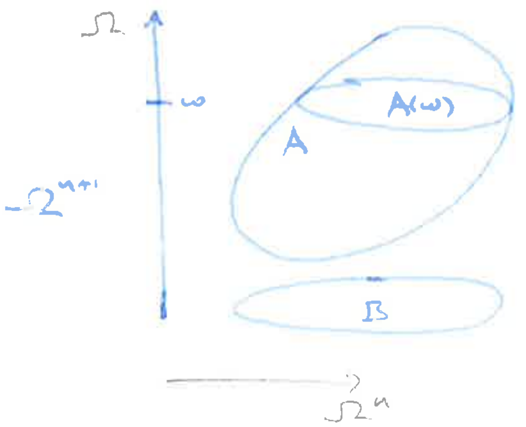
$T(x,A)$ = convex hull of $U(x,A)$ in $\mathbb{R}^n$

The TALAGRAND DISTANCE $d_T(x,A)$ is the euclidean distance from $\underline{0}$ to $T(x,A)$ $= \min\{\|w\| : w \in T(x,A)\}$

$A \subseteq \Omega^{n+1}$; $\omega \in \Omega$, the SECTION of A at height ω is

$$A(\omega) = \{x \in \Omega^n : (x,\omega) \in A\}$$

The PROJECTION of A into Ω^n is $B = \{x \in \Omega^n : (x,\omega) \in A \text{ for some } \omega\} = \bigcup_\omega A(\omega)$



$A = \bigcup_\omega (A(\omega), \omega)$

$d_H(0, \sigma_x(A)) = h$ (Hamming) $\quad \frac{h}{\sqrt{h}} = \|(\frac{1}{h}, \dots, \frac{1}{h})\| \le d_T(x, A) \le \sqrt{h}$

$A \subseteq \Omega^{n+1}$; $\omega \in \Omega$; $A(\omega)$ section; B projection; $x \in \Omega^n$

$\sigma_{(x,\omega)}(A) = \sigma_{(x,\omega)}(A(\omega), \omega) \cup \bigcup_{\alpha \ne \omega} \sigma_{(x,\omega)}(A(\alpha), \alpha) = (\sigma_x(A(\omega)), 0) \cup \bigcup_{\alpha \ne \omega} (\sigma_x(A(\alpha)), 1)$

Claim $T((x,\omega), A) \supseteq (T(x, A(\omega)), 0) \cup (T(x, B), 1)$ $\quad$ abuse of notation

Proof $T((x,\omega), A) \supseteq \sigma_{(x,\omega)}(A(\omega), \omega) = (\sigma_x(A(\omega)), 0)$

$\therefore$ —— " —— $\supseteq$ $(T(x, A(\omega)), 0)$

$\sigma_{(x,\omega)}(A) = (\sigma_x(A(\omega)), 0) \cup \bigcup_{\alpha \ne \omega} (\sigma_x(A(\alpha)), 1)$

$\therefore U((x,\omega), A) \supseteq \bigcup_\alpha (\sigma_x(A(\alpha)), 1) = (\sigma_x(B), 1)$

$T((x,\omega), A) \supseteq (T(x, B), 1)$

Lemma 1 $\emptyset \ne A \subseteq \Omega^{n+1}$; $\omega \in \Omega$, $0 \le \lambda \le 1$, $x \in \Omega^n$

$d_T^2((x,\omega), A) \le \lambda^2 + \lambda d_T^2(x, B) + (1-\lambda) d_T^2(x, A(\omega))$

Proof Let $w_\omega \in T(x, A(\omega))$ and $w_B \in T(x, B)$ with $\|w_\omega\| = d_T(x, A(\omega))$, $\|w_B\| = d_T(x, B)$.

Then $\lambda(w_B, 1) + (1-\lambda)(w_\omega, 0) \in T((x,\omega), A)$

$\therefore d_T^2((x,\omega), A) \le \lambda^2 + \|\lambda w_B + (1-\lambda) w_\omega\|^2$

$\le \lambda^2 + (\lambda\|w_B\| + (1-\lambda)\|w_\omega\|)^2$ $\quad \Delta$-inequality

$\le \lambda^2 + \lambda\|w_B\|^2 + (1-\lambda)\|w_\omega\|^2$ $\quad \square$ convexity

Lemma 2 Let $0 \le r \le 1$. Then $\inf_{0 \le \lambda \le 1} e^{\lambda^2/4} r^{\lambda - 1} \le 2 - r$ $\quad \square$ $\quad$ true but ugly

Theorem 3 (Talagrand's inequality)

$\emptyset \ne A \subseteq \Omega^n$. Then $\int_{\Omega^n} \exp\left(\frac{1}{4} d_T^2(x, A)\right) d\mathbb{P}(x) \le \frac{1}{\mathbb{P}(A)}$

Proof Induction on n. $n=1$: $d_T(x, A) = \begin{cases} 0 & x \in A \\ 1 & x \notin A \end{cases}$

$\int_\Omega = \mathbb{P}(A) + (1 - \mathbb{P}(A)) e^{\frac{1}{4}} \overset{?}{\le} \frac{1}{\mathbb{P}(A)}$ $\qquad p = \mathbb{P}(A)$

$(1-p) e^{\frac{1}{4}} \overset{?}{\le} \frac{1}{p} - p \qquad e^{\frac{1}{4}} \overset{?}{\le} \frac{\frac{1}{p} - p}{1-p} = \frac{1-p^2}{p(1-p)} = 1 + \frac{1}{p} \ge 2$ ✓

Inductive step. $n \ge 1$ $\quad n+1$? Let $\omega \in \Omega$, $E_n(A, \omega) = \int_{\Omega^n} \exp\left(\frac{1}{4} d_T^2((x,\omega), A)\right) d\mathbb{P}(x)$

Task: $\int_\Omega E_n(A, \omega) d\mathbb{P}(\omega) \le \frac{1}{\mathbb{P}(A)}$ $\qquad$ equivalent

$E_n(A, \omega) \overset{\text{Lemma 1}}{\le} e^{\lambda^2/4} \int_{\Omega^n} \exp\left(\frac{1}{4} d_T^2(x, B)\right)^\lambda \exp\left(\frac{1}{4} d_T^2(x, A(\omega))\right)^{(1-\lambda)} d\mathbb{P}(x)$

$\le e^{\lambda^2/4} \left(\int \exp\left(\frac{1}{4} d_T^2(x, B)\right)\right)^\lambda \left(\int \exp\left(\frac{1}{4} d_T^2(x, A(\omega))\right)\right)^{1-\lambda}$ $\quad$ Hölder

inductive hypothesis $\le e^{\lambda^2/4} \left(\frac{1}{\mathbb{P}(B)}\right)^\lambda \left(\frac{1}{\mathbb{P}(A(\omega))}\right)^{1-\lambda} = e^{\frac{\lambda^2}{4}} \left(\frac{\mathbb{P}(A(\omega))}{\mathbb{P}(B)}\right)^{\lambda - 1} \frac{1}{\mathbb{P}(B)}$

This holds $\forall\, 0 \le \lambda \le 1$ $\quad E_n(A, \omega) \le \left(2 - \frac{\mathbb{P}(A(\omega))}{\mathbb{P}(B)}\right) \frac{1}{\mathbb{P}(B)}$ $\quad$ take inf and use Lemma 2

$\therefore \int_\Omega E_n(A, \omega) d\mathbb{P}(\omega) \le \left(2 - \frac{\mathbb{P}(A)}{\mathbb{P}(B)}\right) \frac{1}{\mathbb{P}(B)}$

$= \left(2 - \frac{\mathbb{P}(A)}{\mathbb{P}(B)}\right) \frac{\mathbb{P}(A)}{\mathbb{P}(B)} \frac{1}{\mathbb{P}(A)} \le \frac{1}{\mathbb{P}(A)}$

$x(2-x)$ 

The Talagrand distance of two sets $A, B \subseteq \Omega^n$ is $d_T(A,B) = \inf\{d_T(x,B) : x \in A\} = \min\{\cdots\}$

Corollary 4 $\emptyset \neq A, B \subseteq \Omega^n$, $d_T(A,B) \geq t$. Then $\mathbb{P}(A)\mathbb{P}(B) \leq e^{-t^2/4}$ Another form of T's inequality

Proof $\mathbb{P}(A) e^{t^2/4} \leq \int_{\Omega^n} \exp\left(\frac{1}{4} d_T^2(x,B)\right) d\mathbb{P}(x) \overset{T}{\leq} \frac{1}{\mathbb{P}(B)}$ $\square$ $\quad \mathbb{1}_A t \leq d_T(\cdot, B)$

Lower bounds for $d_T(x, A)$, $d_T(A,B)$.

T = convex hull of S $\quad d(0,T) = \inf\{\|x\| : x \in T\}$

$\|\phi\| = 1$ $\quad \phi(x) \geq d \quad \forall x \in S$ then $d(0,T) \geq d$

$d_T(x,A) = \sup_{\|\phi\|=1} \inf_{y \in A} \phi(\sigma_x(y))$

LINEAR FUNCTIONALS

$= \sup\{d : \exists \phi \in \mathbb{R}^n, \ \phi(\sigma_x(y)) = \sum \phi_i \sigma_i(x,y)$ $\geq$

$= \sum_{i: x_i \neq y_i} \phi_i \geq d \|\phi\| \quad \forall y \in A\}$

Corollary 5 $\emptyset \neq A, B \subseteq \Omega^n$, $d > 0$; $\forall x \in A \ \exists \phi_x = (\phi_i(x))_1^n \in \mathbb{R}^n$ st $\phi_x(\sigma_x(y)) = \sum_{i: x_i \neq y_i} \phi_i(x) \geq d \|\phi_x\| \ \forall y \in B$

Then $d_T(A,B) \geq d$ and $\mathbb{P}(A)\mathbb{P}(B) \leq e^{-d^2/4}$. $\square$

Let $f: \Omega^n \to \mathbb{R}$ be a r.v., $c > 0$. We say that f is of class $\mathrm{Tal}(c)$ on $A \subseteq \Omega^n$ if $\forall x \in A \ \exists \phi_x = (\phi_i(x))_1^n \in \mathbb{R}^n$

st $\|\phi_x\| \leq c$ and $f(x) - f(y) \leq \phi_x(\sigma_x(y)) = \sum_{i: x_i \neq y_i} \phi_i(x) \ \forall y \in \Omega^n$ ϕ_x: bounding function $\quad$ ~Lipschitz

Corollary 6 Let f be a r.v. on Ω^n; $a > b$; suppose f is $\mathrm{Tal}(c)$ on $A = \{f \geq a\}$. Then

$$\mathbb{P}(\underset{A}{f \geq a})\,\mathbb{P}(\underset{B}{f \leq b}) \leq e^{-(a-b)^2/4c^2}$$

Proof Let $x \in A$, $y \in B$; ϕ_x = functional belonging to x (bounding functional), $\|\phi_x\| \leq c$. Then

$a - b \leq f(x) - f(y) \leq \phi_x(\sigma_x(y))$. Hence $d_T(A,B) \geq \frac{a-b}{c}$. (normalise ϕ_x) $\square$

2. Increasing subsequences

S_n; $n!$; $\frac{1}{n!}$ $\qquad \pi \in S_n \quad \pi = a_1 \ldots a_n \qquad I(\pi) = \max\{k : a_{i_1} < a_{i_2} < \cdots < a_{i_k}\}$

$I(2\,5\,8\,1\,4\,3\,7\,6\,9) = 4 \qquad I^*(\quad) = \text{decreasing} = 3$

I on S_n: I_n. Ulam (1961): I_n? $\quad \mu_n = \mathbb{E}(I_n) = ? \quad O(\sqrt{n})$

Hammersley: $\mu_n = (c + o(1))\sqrt{n}$; $\quad \frac{\pi}{2} \leq c \leq e$

Kingman $\qquad\qquad\qquad\qquad (8/\pi)^{\frac{1}{2}}$

Logan–Shepp $\quad c \leq 2$

Vershik–Kerov $\quad c = 2 \qquad$ Proper combinatorial proof

Erdős–Szekeres: $\mathbb{P}(\{I_n \geq \sqrt{n}\} \cup \{I_n^* \geq \sqrt{n}\}) = 1$ $\quad$ largest increasing × decreasing ... $\quad \Rightarrow n \leq kl$

symmetry

$\mathbb{P}(I_n \geq \sqrt{n}) \geq \frac{1}{2} \quad \therefore \mu_n \geq \frac{\sqrt{n}}{2}$; median $m_n \geq \sqrt{n}$

$\mathbb{P}(\underbrace{I_n \geq k}_{=\{X_{n,k} \geq 1\}}) \leq \mathbb{E}(X_{n,k}) = \frac{1}{k!}\binom{n}{k} \leq \frac{1}{2\pi k}\left(\frac{e^2 n}{k^2}\right)^k \leq \frac{1}{2\pi k}\left(\frac{e}{c}\right)^{2k} \qquad k = c\sqrt{n}$

$X_{n,k}$ = # increasing subsequences of length k $\qquad\qquad \times \frac{1}{k}$ throughout $\quad$ probably not

Hence $\mu_n \leq (e + o(1))\sqrt{n}$ and $m_n \leq e\sqrt{n}$

$\frac{\sqrt{n}}{2} \leq \mu_n \leq (e + o(1))\sqrt{n} \qquad \sqrt{n} \leq m_n \leq e\sqrt{n}$

I_n; $\frac{\sqrt{n}}{2} \leq \mu_n \leq (e+o(1))\sqrt{n}$; $\sqrt{n} \leq m_n \leq e\sqrt{n}$

$I_n(\pi)$, $\pi \in S_n$

$[0,1]^n$ - probability space Ω^n; $x \in \Omega^n$, $x = (x_i)_1^n$, $\Omega = [0,1]$

$I_n(x) = \max\{k : \exists\, i_1 < \dots < i_k \text{ st } x_{i_1} \leq \dots \leq x_{i_k}\}$

Then $I_n(\pi)$ and $I_n(x)$ have the same distribution choose points, then label randomly

If $I_n(x) \geq k$ then $\exists$ $K_x = \{i_1, \dots, i_k\}$, $i_1 < \dots < i_k$ st $x_{i_1} \leq \dots \leq x_{i_k}$. We call K_x a CERTIFICATE of $I(x) \geq k$

$y \in \Omega^n$. Then $k - I(y) \leq \sum_{i \in K_x} \sigma_i(x,y)$ (because if $I(y) < k$ to drop down to $I(y)$ must change coordinates)

<u>Lemma 7</u> Let $a > b > 0$, $A = \{x : I(x) \geq a\}$, $B = \{y : I(y) \leq b\}$. Then $d_T(A,B) \geq \frac{a-b}{\sqrt{a}}$ and $\mathbb{P}(A)\mathbb{P}(B) \leq e^{-\frac{(a-b)^2}{4a}}$

<u>Proof</u> $x \in A$. Let K_x be a certificate of the event $\{I(x) \geq a\}$; $|K_x| = a$.

Let $\phi_x \in \mathbb{R}^n$, $\phi_x = (\phi_i(x))_1^n = \mathbb{1}_{K_x}$. Then $\|\phi_x\| = \sqrt{a}$. Also, for $y \in B$,

$$a - b \leq a - I(y) \leq \sum_i \phi_i(x)\sigma_i(x,y)$$

$\therefore d_T(A,B) \geq \frac{a-b}{\|\phi_x\|} = \frac{a-b}{\sqrt{a}}$. The second part follows from Corollary 4. $\square$

<u>Theorem 8</u> m = median of I. Then $\forall t > 0$, $\mathbb{P}(I \geq m+t) \leq 2e^{-t^2/4(m+t)}$

<u>Proof</u> First, take $a = m+t$, $b = m$. $\qquad \mathbb{P}(I \leq m-t) \leq 2e^{-t^2/4m}$

Then $\mathbb{P}(A)\frac{1}{2} \leq e^{-t^2/4(m+t)}$. Second, set $a = m$, $b = m-t$; then $\frac{1}{2}\mathbb{P}(B) \leq e^{-t^2/4m}$ $\square$

$\mathbb{P}(I \leq m-t) \leq 2e^{-t^2/4e\sqrt{n}}$ (good concentration) m_n concentrated in interval length $c\,\epsilon(n)\sqrt[4]{n}$

<u>Note</u> $|m_n - \mu_n| = O(n^{\frac{1}{4}})$

Indeed, $m_n - \mu_n \leq 2\int_0^\infty e^{-t^2/4e\sqrt{n}}\,dt = O(n^{\frac{1}{4}})$ $\quad = c n^{\frac{1}{4}}$ $\quad m_n - \mu_n \leq 2\int t\, d(e^{-t^2/4m_n})$ $\quad (4m_n = 4e\sqrt{n})$

<u>3. Steiner trees, minimal spanning trees and the travelling salesman problem</u>

$[0,1]^2 \subseteq \mathbb{R}^2$ $\quad X \subset [0,1]^2$ $\quad 3 \leq |X| < \infty$

$MST(X)$ = minimal length of spanning tree

$TS(X)$ = minimal length of tour of X closed walk

$ST(X)$ = minimal length of tree whose vertex set contains X

We call a graph $G = (V,E)$ a GEOMETRIC GRAPH if $V \subset [0,1]^2$ and all the edges are straight-line segments (that may cross)

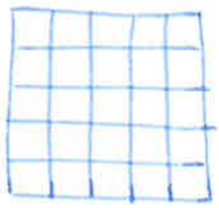
$p > 0$ $\quad G = (V,E)$ geometric

$\ell_p(G) = \ell_p(E) = \sum_{xy \in E} \|x-y\|^p$

$\ell_1 = \ell$ = length

ℓ_2 = square length

<u>Theorem 1</u> If $X \subset [0,1]^2$, $|X| = n$, then $TS(X) \leq 2\sqrt{n}$

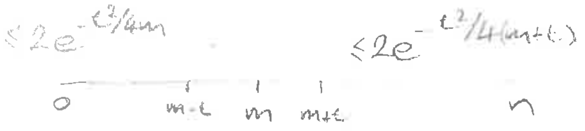


$n = (m+1)^2$

steps length $\geq \frac{1}{m}$

$\frac{(m+1)^2}{m} \sim \sqrt{n}$

$4 = 2\sqrt{4}$



$\mu - m \leq t + 2ne^{-t^2/4(e\sqrt{n}+t)}$ not good enough for $O(\sqrt[4]{n})$

$\mu - m \geq -t(1-2e^{-t^2/4m})$

$m - \mu \leq t(1-2e^{-t^2/4m}) \leq t$ provided $e^{-t^2/4m} < \frac{1}{2}$

$t^2/4m > \log 2$

$t^2 > 4m\log 2$

$\Leftarrow t^2 > 4e\log 2\,\sqrt{n} \approx 7.537\sqrt{n}$

$\approx 2.745^2\sqrt{n}$

$m - \mu < 3\sqrt[4]{n}$

$$m-\mu = \mathbb{E}(m-I) = \sum_{k=1}^{\infty}\mathbb{P}(m-I \geq k) = \sum_{k=1}^{\infty}\mathbb{P}(I \leq m-k)$$

$$\leq \sum_{k=1}^{\infty} 2e^{-k^2/4m} \leq 2\int_0^{\infty} e^{-t^2/4m}\,dt$$

$$= \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{t}{\sqrt{2m}}\right)^2} dt = \sqrt{2\pi}\sqrt{2m} \leq 2\sqrt{e\pi}\,n^{\frac{1}{4}}$$

$$\mu - m = \mathbb{E}(I-m) \leq \sum_{k=1}^{\infty}\mathbb{P}(I-m \geq k) = \sum_{k=1}^{\infty}\mathbb{P}(I \geq m+k) \leq \sum_{k=1}^{\infty} 2e^{-k^2/4(m+k)}$$

$$\leq 2\int_0^{\infty} e^{-t^2/4(m+t)}\,dt \leq 2\int_0^{m} e^{-t^2/8m}\,dt + 2\int_m^{\infty} e^{-t/8}\,dt$$

$$\leq \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{t}{2\sqrt{m}}\right)^2} dt + \left[-16e^{-t/8}\right]_m^{\infty}$$

$$= \sqrt{2\pi}\,2\sqrt{m} + 16e^{-m/8}$$

$$= O(n^{\frac{1}{4}})$$

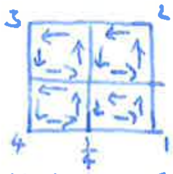
$ST \le MST \le TS \le 2ST$

Theorem 9 There is a continuous surjection $f:[0,4]\to[0,1]^2$ st for $0\le u<t\le 4$,

$\|f(t)-f(u)\|^2 \le |t-u|$ $\quad f(0)=f(4)=(0,0)$

Proof (Sketch)

Γ_1



Corollary 10 $\forall X\subset[0,1]^2$, $|X|=n\ge 3$ $\exists$ circuit C, $V(C)=X$, of square length ≤ 4

Proof $X=\{x_1,\dots,x_n\}$: wlog $\exists\ 0\le t_1<t_2<\dots<t_n\le 4$, $f(t_i)=x_i$. Then

$\ell_2(C)=\sum_1^n\|f(t_{i+1})-f(t_i)\|^2\le|t_{n+1}-t_1|\le 4$ □

Corollary 11 $\forall X\subseteq[0,1]^2$, $|X|=n$, $TS(X)\le 2\sqrt{n}$

Proof Take a tour with lengths $l_1,\dots,l_n$, $\sum_1^n l_i^2\le 4$. Then $\sum_1^n l_i\cdot 1\le\left(\sum l_i^2\right)^{\frac12}n^{\frac12}\le 2\sqrt{n}$ □

$\mathbb{P}\left(d(x_1,x_i)\ge\frac{1}{\sqrt{\pi n}}\ \forall i\ge 2\right)\ge\left(1-\frac1n\right)^{n-1}>\frac1e$

Hence $\mathbb{E}(\min(\text{d of edge leaving } x_i))\ge\frac{1}{e\sqrt{\pi}\sqrt{n}}$

$\therefore\ \mathbb{E}(TS)\ge\frac{\sqrt{n}}{e\sqrt{\pi}}>\frac{\sqrt{n}}{5}$; $\quad\mathbb{E}(ST)\ge\frac{\sqrt{n}}{10}$

$X\leftrightarrow x\in([0,1]^2)^n$; $\quad f(x)=ST(X)$

Theorem 12 For $0<b<a$, $\mathbb{P}(f\ge a)\mathbb{P}(f\le b)\le e^{-(a-b)^2/64}$

Proof Claim: $f(x)$ is of class Tal(4). $A=\{f\ge a\}$

Let $x\in([0,1]^2)^n$. Let H_x be a tour of X with square length ≤ 4

$\phi_i(x)=$ sum of the lengths of the two edges of H_x incident with x_i

Set $\phi_x=(\phi_i(x))_1^n$ $\|\phi_x\|^2=\sum(l_i+l_{i+1})^2\le 2\sum(l_i^2+l_{i+1}^2)=4\sum l_i^2\le 16$ $\quad\therefore\|\phi_x\|\le 4$

Let $y\in([0,1]^2)^n$; $y\leftrightarrow Y$. Let S_y be a Steiner tree for Y. Let $H^*_{x,y}$ be the set of edges of H_x incident with $x_i\ne y_i$.

Then $S_y\cup H^*_{x,y}$ is a connected graph containing X. Hence ($X\cap Y\ne\emptyset$ or in component)

$$f(X)\le L(S_y\cup H^*_{x,y})\le \underbrace{f(y)}_{l(S_y)}+\underbrace{\sum_{x_i\ne y_i}\phi_i(x)}_{l(H^*_{x,y})}$$

$\therefore f$ is of class Tal(4). □

Corollary 13 If m_n is the median of f_n then $\mathbb{P}(|f_n-m_n|\ge t)\le 4e^{-t^2/64}$. □

MST(X) and TS(X) have similar concentration

V Entropy inequalities and applications

1. Introduction

Shannon entropy (discrete, combinatorial)

$X, Y, \dots, X_i, \dots$ discrete, finitely many values   only sum over $p_i \neq 0$

$X \sim p = p(x) = (p_i)_1^n$ don't care what values are

$H(X) = \sum p_i \log\left(\frac{1}{p_i}\right) = \mathbb{E} \log \frac{1}{p}$ base 2

Conditional entropy: X, Y $H(Y|X) = \mathbb{E}_X H(Y|X=x_i)$

$\mathbb{P}(X=x_i) = p_i$; $H(Y|X) = \sum p_i H(Y|X=x_i)$

$(p_i), (r_{ij})$; X and (X,Y) $H(Y|X) = \sum_i p_i \sum_j \frac{r_{ij}}{p_i} \log \frac{p_i}{r_{ij}} = \sum_{i,j} r_{ij} \log p_i + \sum r_{ij} \log \frac{1}{r_{ij}} = H(X,Y) - H(X)$

Theorem 1 X, Y, Z r.v.s

(i) $H(X) \geq 0$, equality $\iff$ X is constant (a.e.)

(ii) $H(Y|X) \geq 0$, equality $\iff$ Y is a function of X

(iii) $H(X,Y) \geq H(X)$, equality $\iff$ Y is a function of X

(iv) $H(Y|X) = H(X,Y) - H(X)$

(v) $H(X,Y|Y,Z) = H(X|Y,Z)$

(vi) If X takes $\leq n$ values then $H(X) \leq \log n$. Equality $\iff$ X is uniformly distributed on a set with n elements ($p_i = \frac{1}{n}\ \forall i$)

(vii) $H(X|Y) \leq H(X)$

Proof (i) ✓ (ii) $p = p' + p''$ $p', p'' > 0$: $p \log \frac{1}{p} \overset{?}{\leq} p' \log \frac{1}{p'} + p'' \log \frac{1}{p''}$ (iv) ✓ (v) ✓ (iii) follows from (ii) and (iv)

$p' \log \frac{p}{p'} + p'' \log \frac{p}{p''} \overset{?}{\geq} 0$ ✓

(vi) $H(X) = \sum_1^n p_i \log \frac{1}{p_i}$

$\leq \log\left(\sum_1^n p_i \left(\frac{1}{p_i}\right)\right) = \log n$ [$\log x$ is strictly concave on $(0,\infty)$]

Theorem 2 X, Y, Z ; $X_0, X_1, \dots, X_n$ r.v.s

(i) $H(X|Y,Z) \leq H(X|Z) \leq H(X,Y|Z)$ ← gets bigger

(ii) $H(X_1, \dots, X_n | X_0) = H(X_1|X_0) + H(X_2|X_0, X_1) + \dots + H(X_n|X_0, \dots, X_{n-1})$ (Shannon's chain rule)

(iii) $H(X_1, \dots, X_n) \leq H(X_1) + \dots + H(X_n)$ take X_0 constant   (Shannon's subadditivity)

Proof (i) second ✓ First $(p_{ijk}), (q_{ik}), (r_{jk}), (s_k)$: $(X,Y,Z), (X,Z), (Y,Z), Z$

$\text{LHS} - \text{RHS} = \sum p_{ijk} \log \frac{r_{jk}}{p_{ijk}} - \sum\sum q_{ik} \log \frac{s_k}{q_{ik}} = \sum p_{ijk} \log \frac{r_{jk}\, q_{ik}}{p_{ijk}\, s_k}$

$\leq \log\left(\sum p_{ijk} \frac{r_{jk} q_{ik}}{p_{ijk} s_k}\right) = \log 1 = 0$

Check: $\sum_{j,k} \sum_i \frac{r_{jk} q_{ik}}{s_k} = \sum_{j,k} r_{j,k} = 1$

(ii) Telescopic sum

(iii) $H(X|Y) \leq H(X)$ + (ii)

2. Inequalities

$h: \mathcal{P}(n) \to \mathbb{R}$ is SUB~~ADDITIVE~~MODULAR if $h(A \cup B) + h(A \cap B) \leq h(A) + h(B)$ $\forall$ $A, B \subseteq [n]$

$X = (X_1, \dots, X_n)$; $X_A = (X_i)_{i \in A}$

Theorem 3 The map $A \mapsto H(X_A)$ is sub~~additive~~modular

Proof $H(X_{A\cup B}) + H(X_{A \cap B}) \overset{?}{\leq} H(X_A) + H(X_B)$

$H(X_{A\cup B}|X_A) = H(X_{B\setminus A}|X_A) \leq H(X_{B \setminus A}|X_{A \cap B}) = H(X_B|X_{A\cap B})$ □

(i) $H(X) = \sum p_i \log \frac{1}{p_i} \geq \log \sum p_i \frac{1}{p_i} = \log 1 = 0$ equality $\Leftrightarrow$ X constant

(iv) $H(Y|X) = \sum_i p_i \sum_j \frac{q_{ij}}{p_i} \log \frac{p_i}{q_{ij}} = \sum_{ij} q_{ij} \log \frac{1}{q_{ij}} - \sum_i p_i \log \frac{1}{p_i} = H(X,Y) - H(X)$

(v) $H(X,Y|Y,Z) = H(X,Y,Y,Z) - H(Y,Z) = H(X,Y,Z) - H(Y,Z) = H(X|Y,Z)$

(iii) Suppose Y takes 2 values for same value of X. Then $p \log \frac{1}{p}$ is replaced by $p' \log \frac{1}{p'} + p'' \log \frac{1}{p''}$ $\quad p' + p'' = p$

$$p' \log \frac{1}{p'} + p'' \log \frac{1}{p''} - p \log \frac{1}{p} = p' \log \frac{p}{p'} + p'' \log \frac{p}{p''} > 0 \qquad \boxed{H(Y,X) \geq H(X) \text{ eq.} \Leftrightarrow Y = f(X)}$$

Get ≥ 2 finitely many values by induction. If $Y = f(X)$ each $Y|X=x_i$ is constant so has entropy 0

(ii) $H(Y|X) = H(X,Y) - H(X) \geq 0$ eq. $\Leftrightarrow Y = f(X)$ by (iii)

(vi) $H(X) = \sum p_i \log \frac{1}{p_i} \leq \log \left(\sum p_i \frac{1}{p_i}\right) = \log n$ eq. $\Leftrightarrow \frac{1}{p_i}$ constant

(vii) $H(X|Y) - H(X) = \sum_j q_j \sum_i \frac{r_{ij}}{q_j} \log \frac{q_j}{r_{ij}} - \sum_i p_i \log \frac{1}{p_i} = \sum r_{ij} \log \frac{p_i q_j}{r_{ij}} \leq \log \left(\sum r_{ij} \frac{p_i q_j}{r_{ij}} \right) = \log 1 = 0$

$h : \mathcal{P}(n) \to \mathbb{R}$ is SUBMODULAR if $h(A\cup B) + h(A\cap B) \le h(A) + h(B)$

$X = (X_1, \dots, X_n)$ $X_A = (X_i)_{i\in A}$ $A \subseteq [n]$

Corollary 3 $A \mapsto H(X_A)$ is submodular □

$\mathcal{A} \subseteq \mathcal{P}(n)$; in fact, we'll take multisets of subsets of $[n]$

$M_{n,m}$ = collection of multisets of subsets of $[n]$ with total number of elements $= m$

An ELEMENTARY COMPRESSION of $\mathcal{A} \in M_{n,m}$ is a family $\mathcal{A}'$ obtained from $\mathcal{A}$ by replacing some two sets $A, B \in \mathcal{A}$ by $A\cup B$ and $A \cap B$. In case $A\cap B = \phi$ we do not take it.

A COMPRESSION of ~~[illegible]~~ $\mathcal{A}$ is the application of a sequence of elementary compressions.

$\mathcal{A} > \mathcal{B}$. Then '$>$' turns $M_{n,m}$ into a poset: indeed, $\sum_{A\in\mathcal{A}} |A|^2 < \sum_{B\in\mathcal{B}} |B|^2$. convexity

Minimal elements of $M_{n,m}$ are nested families. Also, for $\mathcal{A} \in M_{n,m}$ $\exists!$ minimal element $\mathcal{A}^\#$ dominated by $\mathcal{A}$. $A_j = \{i : i$ is in at least j sets of $\mathcal{A}\}$ $\mathcal{A}^\# = \{A_1, A_2, \dots\}$.

$\mathcal{A}$ is a k-UNIFORM COVER of $[n]$ if every $i \in [n]$ is in exactly k sets in $\mathcal{A}$. [$\mathcal{A}^\# = [n]$ k times]

Theorem 4 If $\mathcal{A} > \mathcal{B}$ then $\sum_{A\in\mathcal{A}} H(X_A) \ge \sum_{B\in\mathcal{B}} H(X_B)$

Proof $\mathcal{A} = \{A_1, \dots, A_m\}$ $\mathcal{A}_{(ij)} = \{A_1, \dots, \hat{A}_i, \dots, \hat{A}_j, \dots, A_m, A_i\cup A_j, A_i \cap A_j\}$

$$\Rightarrow \sum_{A\in\mathcal{A}_{(ij)}} H(X_A) - \sum_{A\in\mathcal{A}} H(X_A) = H(X_{A\cup B}) + H(X_{A\cap B}) - H(X_A) - H(X_B) \le 0. \ \square$$

Corollary 5 For any multiset $\mathcal{A}$ $\sum_{A\in\mathcal{A}^\#} H(X_A) \le \sum_{A\in\mathcal{A}} H(X_A)$. □

Corollary 6 (Shearer's Inequality)

If $\mathcal{A}$ is a k-uniform cover of $[n]$ then $H(X) \le \frac{1}{k}\sum_{A\in\mathcal{A}} H(X_A)$ □

3. Applications

Given G: # independent sets of vertices?

Kahn & Lovasz

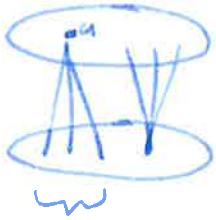
Theorem 7 Let G be an r-regular $n\times n$ bipartite graph. Then $i(G) \le (2^{r+1}-1)^{n/r}$

Proof $\mathcal{I}(G)$ = collection of independent sets; $i(G) = |\mathcal{I}(G)|$. $\mathcal{I}(G)$ = probability space, all outcomes equiprobable $X = (X_v)_{v\in V(G)}$ $X_v = \begin{cases} 1 & \text{if } v \in I \\ 0 & \text{otherwise}\end{cases}$ $I \in \mathcal{I}(G) \mapsto X = \mathbb{1}_I$

$H(X) = H(\mathcal{I}) = \log i(G)$

subadditivity and forget some conditions

G:



U

W

$|\Gamma(u)| = r$

$$H(X) = H(X_U, X_W) = H(X_U | X_W) + H(X_W)$$

Shearer

$$\le \sum_{u\in U} H(X_u | X_{\Gamma(u)}) + \frac{1}{r}\sum_{u\in U} H(X_{\Gamma(u)})$$

Set $p_u = \mathbb{P}(X_{\Gamma(u)} = \underline{0}_{\Gamma(u)})$. Then $H(X_u | X_{\Gamma(u)}) \le p_u \log_2 2 = p_u$

$$H(X_{\Gamma(u)}) \le p_u \log \frac{1}{p_u} + (1-p_u)\log\frac{2^r - 1}{1-p_u}$$

$$\therefore r H(X_u | X_{\Gamma(u)}) + H(X_{\Gamma(u)}) \le r p_u + p_u \log\frac{1}{p_u} + (1-p_u)\log\frac{2^r-1}{1-p_u}$$

$$= \log(2^r - 1) + (1-p_u)\log\frac{1}{1-p_u} + p_u \log\left(\frac{2^r}{(2^r-1)p_u}\right)$$

concavity $\le \log(2^r-1) + \log\left(1 + \frac{2^r}{2^r-1}\right) = \log(2^{r+1}-1)$

$$\therefore H(X) \le \frac{n}{r}\log(2^{r+1}-1)$$

Zhao: Theorem True if G just r-regular

MARTINGALES

$(Y_k)_1^n$ MULTIPLICATIVELY INDEPENDENT if $\mathbb{E}\left(\prod_1^n (a_k + b_k Y_k)\right) = \prod_1^n a_k$ $\left(\mathbb{E}(\prod_1^n Y_k) = 0\right)$

Correspond to MARTINGALE $(X_k)_0^n$ $Y_k = X_k - X_{k-1}$ (MARTINGALE DIFFERENCE SEQUENCE)

Have nested σ-algebras $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_n = \mathcal{P}(\Omega)$, corresponding partitions $\mathcal{P}_n$ FILTRATION

X r.v. on Ω $X_k = \mathbb{E}(X \mid \mathcal{F}_k)$ = expected value of X given which part of $\mathcal{P}_k$ ω is in

$$\mathbb{E}(X \mid \mathcal{F}_k)(\omega) = \frac{\mathbb{E}(X \mathbb{1}_{P_i})}{\mathbb{P}(P_i)} \qquad \omega \in P_i \in \mathcal{P}_k.$$

X_k is the orthogonal projection of X onto the space of $\mathcal{F}_k$-measurable functions:

$$\mathbb{E}(UX) = \mathbb{E}(U \mathbb{E}(X \mid \mathcal{F}_k)) \text{ for all } \mathcal{F}_k\text{-measurable } U.$$

Theorem 1 $(Y_k)_1^n$ MI $-p_k \le Y_k \le q_k = 1 - p_k$ $p = \frac{1}{n}\sum p_k$ $q = \frac{1}{n}\sum q_k$

$h > 0$: $\mathbb{E}(e^{hX}) \le e^{-nph} \prod (q_k + p_k e^h) \le e^{-nph}(q + pe^h)^n \le e^{\frac{nh^2}{8}}$ uniformity

Theorem 2 $\mathbb{P}(X \ge tn) \le \left[\left(\frac{p}{p+t}\right)^{p+t}\left(\frac{q}{q-t}\right)^{q-t}\right]^n$ $0 < t < q$ CHERNOFF

Lemma 3 $\left(\frac{p}{p+t}\right)^{p+t}\left(\frac{q}{q-t}\right)^{q-t} \le e^{-2t^2}$

Theorem 4 $\mathbb{P}(X \ge tn) \le e^{-2nt^2}$ $\mathbb{P}(X \ge u) \le e^{-\frac{2u^2}{n}}$

If the Y_k have spread s: $\mathbb{P}(X \ge u) \le e^{-\frac{2u^2}{ns^2}}$

Corollary 5 $X = (X_k)_0^n$ difference sequence bounded by s in modulus

$$\mathbb{P}(X - \mathbb{E}(X) \ge u) \le e^{-\frac{u^2}{2ns^2}}$$

Theorem 6 $(Y_k)_1^n$ MI $|Y_k| \le s_k$ $S = \sum_1^n s_k^2$ $\mathbb{P}(X \ge u) \le e^{-\frac{u^2}{2S}}$ symmetry

Corollary 7 $\mathbb{P}(X - \mathbb{E}(X) \ge u) \le e^{-\frac{u^2}{2S}}$ HOEFFDING–AZUMA

$s(Y_k, \mathcal{F}_{k-1})$ least s st. $f_{k-1} \le Y_k \le f_{k-1} + s$ some $\mathcal{F}_{k-1}$-measurable f_{k-1}

(or $f_{k-1} \le X_k \le f_{k-1} + s$)

$S_n = \sum_1^n s_k^2$ SIZE of X

Lemma 8 $X_0 = 0$ $\mathbb{E}(e^{hX_n}) \le e^{\frac{h^2 S_n}{8}}$

Theorem 9 $\mathbb{P}(X - \mathbb{E}(X) \ge u) \le e^{-\frac{2u^2}{S_n}}$

Theorem 10 $Z_1, \dots, Z_n$ independent taking values in $\Omega_1, \dots, \Omega_n$

$f: \Omega_1 \times \cdots \times \Omega_n \to \mathbb{R}$ $|f(x) - f(y)| \le c_k$ when x, y differ only in kth coordinate

Then $\mathbb{P}(f - \mathbb{E}(f) \ge u) \le e^{-\frac{2u^2}{S_n}}$ $S_n = \sum_1^n c_k^2$

SYMMETRIC GROUP

$d(\pi,\rho) = \#(i : \pi(i) \neq \rho(i))$ $\quad d^*(\pi,\rho) = \min k : \pi = \rho\, \tau_1 \cdots \tau_k$ $\quad \tau_i$ transpositions

$d \leq 2d^*$ $\quad d^* \leq d-1$

__Theorem__ $f : S_n \to \mathbb{R}$ Lipschitz with constant c. Then $\mathbb{P}(f - \mathbb{E}(f) \geq u) \leq e^{-\frac{u^2}{2nc^2}}$

__Theorem 2__ $A \subseteq S_n$ $\cancel{\mathbb{P}(x) > 0}$ $\mathbb{P}(A) = \alpha > 0$ $\quad A_t = \{x : d(x,A) \leq t\}$

If $t = (x+1)\left(2n \log \frac{1}{\alpha}\right)^{\frac{1}{2}}$, $x \geq 0$ then $\mathbb{P}(A_t) \geq 1 - \alpha^{x^2}$

__Theorem 3__ $A, B \subseteq S_n$ $\quad \mathbb{P}(A) = \alpha$ $\quad \mathbb{P}(B) = \beta$ $\quad d = d(A,B)$

Then $d \leq \left(\log\left(\frac{1}{\alpha}\right)^{\frac{1}{2}} + \log\left(\frac{1}{\beta}\right)^{\frac{1}{2}} \right) \sqrt{2n}$

$$\min(\alpha,\beta) \leq e^{-d^2/8n}$$

Variants for d^*

CHROMATIC NUMBER

__Theorem 5__ $0 < p < 1$ fixed $\quad X_n = \chi(G_{n,p})$ $\quad \mathbb{P}(|X_n - \mathbb{E}(X_n)| \geq t) \leq 2e^{-\frac{2t^2}{n}}$

(VERTEX-REVEALING MARTINGALE has size n)

TALAGRAND'S INEQUALITY

For $x, y \in \Omega^n$, $\sigma_x(y) = \mathbb{1}_{x_i \neq y_i} \in \{0,1\}^n$

For $\phi \neq A \subseteq \Omega^n$ let $U(x, A)$ be the up-set in $[0,1]^n$ generated by $\sigma_x(A)$

$T(x, A)$ = convex hull of $U(x, A)$

TALAGRAND DISTANCE $d_T(x, A) = d(0, T(x, A))$

$A \subseteq \Omega^{n+1}$, $\omega \in \Omega$ SECTION $A(\omega) = \{x \in \Omega^n : (x, \omega) \in A\}$

PROJECTION $B = \bigcup_{\omega \in \Omega} A(\omega)$

$T((x,\omega), A) \supseteq (T(x, A(\omega)), 0) \cup (T(x, B), 1)$

Lemma 1 $\phi \neq A \subseteq \Omega^{n+1}$ $\omega \in \Omega$ $0 \leq \lambda \leq 1$ $x \in \Omega^n$

$$\Rightarrow d_T((x,\omega), A)^2 \leq \lambda^2 + \lambda d_T{}^2(x, B) + (1-\lambda) d_T(x, A(\omega))^2$$

Lemma 2 $0 < r \leq 1 \Rightarrow \inf_{0 \leq \lambda \leq 1} e^{\lambda^2/4} r^{\lambda - 1} \leq 2 - r$

Theorem 3 (Talagrand's inequality)

$$\phi \neq A \subseteq \Omega^n \Rightarrow \int_{\Omega^n} e^{\frac{1}{4} d_T(x, A)^2} \, d\mathbb{P}(x) \leq \frac{1}{\mathbb{P}(A)}$$

Corollary 4 $\mathbb{P}(A)\mathbb{P}(B) \leq e^{-t^2/4}$, $t = d_T(A, B) = \inf d_T(a, B)$

Corollary 5 $\phi \neq A, B \subseteq \Omega^n$ $d > 0$

If $\forall x \in A \ \exists \ \phi_x = (\phi_i(x))_1^n$ s.t. $\forall y \in B$ $\sum_{x_i \neq y_i} \phi_i(x) \geq d \|\phi_x\|$ then $\mathbb{P}(A)\mathbb{P}(B) \leq e^{-d^2/4}$

$f : \Omega^n \to \mathbb{R}$, $c > 0$

f OF CLASS $\mathrm{Tal}(c)$ on A if $\forall x \in A$ $\exists$ bounding functions ϕ_x $\|\phi_x\| \leq c$

$$f(x) - f(y) \leq \phi_x(\sigma_x(y)) \quad \forall y \in \Omega^n$$

Corollary 6 $a > b$ f $\mathrm{Tal}(c)$ on $\{f \geq a\}$

$$\Rightarrow \mathbb{P}(f \geq a)\mathbb{P}(f \leq b) \leq e^{-\frac{(a-b)^2}{4c^2}}$$

INCREASING SUBSEQUENCES

$x = (x_1, \ldots, x_n) \in S_n$ $I(x)$ length of longest increasing subsequence

$I^*(x)$ " " " decreasing "

$\frac{1}{2}\sqrt{n} \leq \mu_n \leq (e + o(1))\sqrt{n}$

$\sqrt{n} \leq M_n \leq e\sqrt{n}$ Equivalently choose $x \in [0,1]^n$ (choose points then labels)

If $I(x) \geq h$ $\exists$ CERTIFICATE $K_x = \{i_1, \ldots, i_h\}$ $x_{i_1} \leq \cdots \leq x_{i_h}$

$$h - I(y) \leq \sum_{i \in K_x} \sigma_i(x, y)$$

Corollary 7 $\mathbb{P}(I \geq a)\mathbb{P}(I \leq b) \leq e^{-\frac{(a-b)^2}{4a}}$

Theorem 8 $\mathbb{P}(I \geq M + t) \leq 2e^{-\frac{t^2}{4(M+t)}}$ M concentrated in interval

$\mathbb{P}(I \leq M - t) \leq 2e^{-t^2/4M}$ length $\omega(n)\sqrt{n}$

TRAVELLING SALESMAN

Minimal Spanning Tree
Steiner Tree
Travelling Salesman

of $X \subseteq [0,1]^2$

$3 \leq |X| < \infty$

$\ell_p(G) = \sum_{xy \in E} \|x-y\|^p$

Theorem 9 $\exists$ continuous surjection $f : [0,1] \to [0,1]^2$ $\|f(u)-f(t)\|^2 \leq |u-t|$

Corollary 10 ~~$TS(X) \leq 4$~~ $\exists$ TS-tour of square length ≤ 4

Corollary 11 $TS(X) \leq 2\sqrt{n}$

$\mathbb{E}(TS) \geq \frac{\sqrt{n}}{5}$ all have means $\Theta(\sqrt{n})$

$f(x) = ST(x)$

Theorem 12 $a > b > 0$ $\mathbb{P}(f \geq a)\mathbb{P}(f \leq b) \leq e^{-\frac{(a-b)^2}{64}}$

Corollary 13 $\mathbb{P}(|f_n - m_n| \geq t) \leq 4e^{-t^2/64}$

ENTROPY

SHANNON ENTROPY $H(X) = \sum p_i \log \frac{1}{p_i} = \mathbb{E} \log \frac{1}{p}$ (base 2)

CONDITIONAL ENTROPY $H(Y|X) = \mathbb{E}_X H(Y|X=x_i)$

<u>Basic properties</u> $H(X) \geq 0$ with equality if and only if X constant

$H(Y|X) = H(X,Y) - H(X)$

If X takes n values then $H(X) \leq \log n$ with equality if and only if X is uniformly distributed

$H(X,Y) \geq H(X)$ with equality if and only if Y is a function of X

$H(Y|X) \geq 0$ with equality if and only if Y is a function of X

$H(X,Y|Y,Z) = H(X|Y,Z)$

<u>More properties</u>

$H(X|Y,Z) \leq H(X|Z) \leq H(X,Y|Z)$

$H(X_1,\dots,X_n|X_0) = H(X_n|X_{n-1},\dots,X_0) + H(X_{n-1}|X_{n-2},\dots,X_0) + \cdots + H(X_1|X_0)$ SHANNON'S CHAIN RULE

$H(X_1,\dots,X_n) \leq H(X_1) + \cdots + H(X_n)$ SHANNON'S SUBADDITIVITY

$h: \mathcal{P}([n]) \to \mathbb{R}$ SUBMODULAR if $h(A\cup B) + h(A\cap B) \leq h(A) + h(B)$ $\forall A,B \subseteq [n]$

$X = (X_1,\dots,X_n)$ $X_A = (X_i)_{i\in A}$

<u>Theorem 3</u> $A \mapsto H(X_A)$ submodular

$\mathcal{M}_{n,m}$ collection of multisets of subsets of $\mathcal{P}([n])$ with m elements in total

ELEMENTARY COMPRESSION of $\mathcal{A} \in \mathcal{M}_{n,m}$ is $\mathcal{A}'$ obtained by replacing A and B by $A\cup B$ and $A\cap B$ (throw away empty sets)

$\mathcal{A} > \mathcal{B}$ if $\mathcal{B}$ is a COMPRESSION of $\mathcal{A}$ – partial order as $\sum |A|^2$ strictly increasing

$\exists$ unique minimum compression $\mathcal{A}^\#$

$\mathcal{A}$ a k-UNIFORM COVER of $[n]$ if each element used k times ($\mathcal{A}^\# = k[n]$)

<u>Theorem 4</u> $\mathcal{A} > \mathcal{B} \Rightarrow \sum_{A\in\mathcal{A}} H(X_A) \geq \sum_{B\in\mathcal{B}} H(X_B)$

<u>Corollary 5</u> $\sum_{A\in\mathcal{A}^\#} H(X_A) \leq \sum_{A\in\mathcal{A}} H(X_A)$

<u>Corollary 6</u> $H(X) \leq \frac{1}{k} \sum_{A\in\mathcal{A}} H(X_A)$ if $\mathcal{A}$ k-uniform cover (SHEARER'S INEQUALITY)

<u>Theorem 7</u> G r-regular $n\times n$-bipartite $\Rightarrow$ #(independent sets) $\leq (2^{r+1}-1)^{\frac{n}{r}}$