W1 3D1
18/1/15

Prob. Methods in Comb.

Thm. $P \sim \mathrm{Pois}(\mu)$, then (i) $\mathbb{P}(P \leq (1-\varepsilon)\mu) \leq e^{\varepsilon^2 \mu/2}$, $\mathbb{P}(P \geq (1+\varepsilon)\mu) \leq [e^{\varepsilon}(1+\varepsilon)^{-1-\varepsilon}]^{\mu}$.

Recall We construct $R \subseteq \Omega$ at random by $\mathbb{P}(x \in R) = p_x$ indep. of other x. $B_i =$ "$A_i \subseteq R$". $\mu = \sum_{i \in I}^{m} \mathbb{P}(B_i) = \sum_{i=1}^{n} \prod_{x \in A_i} p_x$,

$\Delta = \sum_{\substack{(i,j) \in I^2 \\ i \sim j}} \mathbb{P}(B_i \cap B_j)$ where $i \sim j \Leftrightarrow i \neq j$ and $A_i \cap A_j \neq \emptyset$.

$J \subseteq I$ is a maximal disjoint family $\Longleftrightarrow \left(\bigcap_{i \in J} B_i\right.$ holds and $\forall i \notin J$, $\left. B_i \to i \overset{\text{some } j \in J}{\sim} J\right)$.

For $s \in \mathbb{N}$, $\mu_s = \min\{\sum_{i \not\sim J} \mathbb{P}(B_i) \mid J \subseteq I$ of size $s\}$. $\nu = \max_{j \in I} \sum_{i \sim j} \mathbb{P}(B_i)$. Then observe $\mu_s \geq \mu - s \cdot \nu$.

Lemma $\mathbb{P}(\exists$ max. disj. family J of size $s) \leq \frac{\mu^s}{s!} e^{-\mu_s} e^{\Delta/2}$.

pf. Given $J \subseteq I$ of size s, $j \not\sim i$ for all $i, j \in J$, $\mathbb{P}(J$ is max. disj. family$) = \mathbb{P}\left(\bigcap_{j \in J} B_j \cap \bigcap_{i \not\sim J} \overline{B}_i\right) \underset{\text{ind.}}{=} \prod_{j \in J} \mathbb{P}(B_j) \cdot \mathbb{P}\left(\bigcap_{i \not\sim J} \overline{B}_j\right) \underset{\text{Janson}}{\leq} \prod_{j \in J} \mathbb{P}(B_j) e^{-\mu(J) + \Delta(J)/2} \leq$
$\prod_{j \in J} \mathbb{P}(B_j) e^{-\mu_s + \Delta/2}$. So, $\mathbb{P}(\exists$ such family$) \leq \sum_{\substack{J \subseteq I \\ |J| = s}} \prod_{j \in J} \mathbb{P}(B_j) e^{-\mu_s + \Delta/2} \underset{\text{(like last time)}}{\leq} \frac{\mu^s}{s!} e^{-\mu_s + \Delta/2}$

Note Previously we could prove $\mathbb{P}(\exists$ max. disj. fam. of size $\geq 3\mu) \leq \mathbb{P}(\exists$ disj family of size $3\mu) \leq \left(\frac{e}{3}\right)^{\mu}$. If $\Delta = o(1)$ and $\mu_{3\mu} = \mu - o(1)$ then for $s \leq 3\mu$ we get $\mathbb{P}(\exists$ max. disj. fam. of size $s) = (1+o(1)) \frac{\mu^s}{s!} e^{-\mu} = (1+o(1)) \mathbb{P}(\mathrm{Pois}(\mu) = s)$.

Every vertex of $G(n,p)$ is in a Triangle (Rev.)

Let $N(x) = \#$triangles which contain x. $\mathbb{E}N(x) = \binom{n-1}{2} p^3 = \mu$.

Thm Let p be such that $\mu >> \log n$. Then a.a.s. $N(x) = (1+o(1))\mu$ for every x. More precisely, $\mathbb{P}(|N(x) - \mu| \geq \varepsilon\mu) = o(n^{-1})$ for every $x \in V(G)$, $\varepsilon > 0$ (and the result follows from the union bound).

pf. For simplicity assume $\mu = n^{o(1)}$ (eg. $p = n^{-2/3 + o(1)}$). $\Delta = O(n^3 p^5) = o(1)$, $\nu = O(np^3)$, hence $\mu\nu = o(1)$. $\mu_s = \mu - o(1)$. Define $\tilde{N}(x) =$ max. # edge-disjoint Δs containing x. Clearly $N(x) \geq \tilde{N}(x)$ and
$\mathbb{P}(N(x) \leq (1-\varepsilon)\mu) \leq \mathbb{P}(\tilde{N}(x) \leq (1-\varepsilon)\mu) = (1+o(1)) \mathbb{P}(\mathrm{Pois}(\mu) \leq (1-\varepsilon)\mu) \leq (1+o(1)) e^{-\varepsilon^2\mu/2} \leq o(n^{-2})$ for large enough n. For the upper tail,

$\mathbb{P}(\tilde{N}(x) \geq (1+\varepsilon)\mu) = \sum_{s=(1+\varepsilon)\mu}^{3\mu} \mathbb{P}(\tilde{N}(x)=s) + \mathbb{P}(\tilde{N}(x) > 3\mu) \leq 1/n^2 + (e/3)^{\mu} \leq 1/n^2 + 1/n^2$ for large enough n, since $\mathbb{P}(\mathrm{Pois}(\mu) \geq (1+\varepsilon)\mu) \leq (1-\delta_\varepsilon)^{\mu} \xrightarrow[n\to\infty]{} 0$.

$\mathbb{P}(\exists \quad) \leq n^6 p^{11} = o(n^{-1})$, $\mathbb{P}(\quad z_1', \ldots \quad z_4') \leq n^{12} p^{20} = o(n^{-1})$.

Consider an aux. graph T with vertices $\leftrightarrow$ triangles containing x, edge $\leftrightarrow$ common edge. We proved w. prob $1-o(n^{-1})$: $\alpha(T) \leq (1+\varepsilon)\mu$, $\Delta(T) \leq 3$, max. induced matching ≤ 3. We claim $\alpha(T) \geq V(T) - 27$ so $V(T) \leq \alpha(T) + 27 \leq (1+2\varepsilon)\mu$. Why is that? While $\exists$ edge, delete both endpoints and their neighbours (≤ 6), after 3 iterations we can't have other edges (otherwise we'd have an induced matching of size 4).

X r.v. $\in R$ finite set.

<u>Self-Information</u> is a measure of the information content associated with the outcome of X. $I(S)$ - info. content of $x \in S$. (i) $I(S)$ depends only on $\mathbb{P}(x \in S)$, (ii) If $x \in S_1$ and $x \in S_2$ are indep., $I(S_1 \cap S_2) = I(S_1) + I(S_2)$, $I(S)$ nonzero.

Only choice is $\log_b 1/\mathbb{P}(x\in S)$ for some $b > 1$. $b=2$ "bits", $b=e$ "nats", $b=10$ "bans/dits/hartleys".

<u>Entropy</u> Is the "average" information content of an outcome:

$$H(X) = \mathbb{E}\, I(x) = \sum_{x\in R} \mathbb{P}(X=x) \log_2 \frac{1}{\mathbb{P}(X=x)}.$$

<u>Conditional Entropy</u> If $X, Y: \Omega \to R, S$ r.v. then $(X,Y): \Omega \to R\times S$.

Then $H(X|Y) \overset{\text{def}}{=} H(X,Y) - H(Y)$.

<u>Lemma</u> X, Y, Z r.v. - (i) $0 \leq H(X) \leq \log_2 |\mathrm{range}(X)|$

(ii) $H(X,Y) \geq H(X)$

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

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

<u>pf.</u> (i) $z \mapsto \log_2 z$ is concave, hence $0 \leq \overset{\text{def}}{=} H(x)$ " $\sum_{x\in R} \mathbb{P}(X=x) \log_2 1/\mathbb{P}(X=x) \overset{\text{Jensen}}{\leq}$

$\log_2 \frac{1}{1/|R|} = \log_2 |R| = \log_2 |\mathrm{range}(X)|$. Sharp iff $\mathbb{P}(X=x) = 1/|R|$ $(\forall x \in R)$.

(ii) $H(x,y) = \sum_{x\in R}\sum_{y\in S} \mathbb{P}(X=x, Y=y) \log_2 \frac{1}{\mathbb{P}(X=x,Y=y)} \geq \sum_{x\in R}\sum_{y\in S} \mathbb{P}(X=x, Y=y) \frac{1}{\mathbb{P}(X=x)} = \sum_{x\in R} \mathbb{P}(X=x) \log_2 \frac{1}{\mathbb{P}(X=x)} = H(X)$.

(iii) $H(X) + H(Y) - H(X,Y) = \sum_{\substack{x\in R\\ y\in S}} \mathbb{P}(X=x,Y=y) \log_2 \underbrace{\frac{\mathbb{P}(X=x,Y=y)}{\mathbb{P}(X=x)\mathbb{P}(Y=y)}}_{z_{x,y}} = \sum_{\substack{x\in R\\ y\in S}} \mathbb{P}(X=x)\mathbb{P}(Y=y) f(z_{x,y}) \overset{\text{Jensen}}{\geq}$ $f(z) = z\log z$ convex

$f\left(\sum_{x\in R} \mathbb{P}(X=x, Y=y)\right) = f(1) = 0$.

(iv) will be shown later.

<u>Prop.</u> Let $X=(X_1,\dots,X_n): \Omega \to R_1\times\dots\times R_n$. By applying (iii) repeatedly, $H(X)\leq\sum_{i=1}^{n}H(X_i)$.

<u>cor.</u> Let $\mathcal{F}$ be a family of subsets of $[n]$, p_i = fraction of sets in $\mathcal{F}$ which contain i. Then $|\mathcal{F}|\leq 2^{\sum_{i=1}^{n}H(p_i)}$ where $H(p)=H(\mathrm{Bin}(p))=p\log_2 1/p+(1-p)\log_2 1/(1-p)$.

<u>pf.</u> Choose $X\in\mathcal{F}$ u.a.r. $H(X)=\log_2|\mathcal{F}|\leq H(X_1)+\dots+H(X_n)=\sum_{i=1}^{n}H(p_i)$.

<u>cor.</u> $\forall n\geq 1$, $\alpha\leq 1/2$, $\binom{n}{\alpha n}\leq\sum_{k=0}^{\lfloor\alpha n\rfloor}\binom{n}{k}\leq 2^{nH(\alpha)}$ ← sharp, up to poly(n) factor. $\binom{n}{\alpha n}=(1+o(1))\frac{2^{nH(\alpha)}}{\sqrt{2\pi n\alpha(1-\alpha)}}$

<u>pf.</u> Take $\mathcal{F}$ to be the family of subsets of size $\leq\alpha n$ of $[n]$. $p_i\leq\alpha$ and H is increasing on $[0,1/2]$.

<u>Cor</u> Let $X_1,\dots,X_n\sim U([-1,1])$ i.i.d. Then $\forall c\geq 0$, $\mathbb{P}(|X|\geq c\sqrt{n})\leq 2^{1-c^2/2}$ where $X=\sum_{i=1}^{n}X_i$.

<u>Lemma</u> (Shearer '86) Let $X=(X_1,\dots,X_n)$ r.v. $\mathcal{G}\subseteq\mathcal{P}([n])$, such that any $i\in[n]$ belongs to at least k $G\in\mathcal{G}$. Then $H(X)\leq\frac{1}{k}\sum_{G\in\mathcal{G}}H(X(G))$ where $X(G)=(X_i)_{i\in G}$.

We'll probably only show this next week due to time constraints.

<u>Cor</u> $\mathcal{F}\subseteq S_1\times\dots\times S_n$, $\mathcal{G}=\{G_1,\dots,G_n\}$ such that each $i\in[n]$ belongs to at least k $G\in\mathcal{G}$. Let $\mathcal{F}_i$ be the proj. of $\mathcal{F}$ onto G_i. Then $|\mathcal{F}|^k\leq\prod_{i\in[n]}|\mathcal{F}_i|$.

<u>pf.</u> $X\sim\mathrm{Uni}(\mathcal{F})$, $X_i=X|_{G_i}$ then by Shearer $\log_2|\mathcal{F}|=H(X)\leq\frac{1}{k}\sum_{i\in[n]}\log_2|\mathcal{F}_i|$.

<u>Cor.</u> (Loomis-Whitney '49) Let $B\subseteq\mathbb{R}^n$ measurable, $\forall i$ $B_i=\pi_i(B)$ (The projection to $x_i=0$). Then $\mathrm{Vol}_n(B)^{n-1}\leq\prod_{i=1}^{n}\mathrm{Vol}_{n-1}(B_i)$

<u>pf</u> Approximate B using $\varepsilon\times\dots\times\varepsilon$ cubes, $\mathcal{F}$ = set of cubes, use prev. cor.

<u>cor.</u> Let $\mathcal{F}$ be a family of graphs on $V=[n]$ s.t. $\forall G_1,G_2\in\mathcal{F}$, $K_3\subseteq G_1\cap G_2$. Then $|\mathcal{F}|\leq\frac{1}{4}2^{\binom{n}{2}}$

$|\mathcal{G}|=m$, A ~~complement~~ has $s\leq\frac{1}{2}\binom{n}{2}$ edges $\quad sm=k\binom{n}{2}$

<u>pf.</u> Let $\mathcal{G}=\{A\subseteq[n],\ |A|=\lfloor\frac{n}{2}\rfloor\}$. $\forall G\in\mathcal{G}$, $\pi_G(\mathcal{F})$ intersecting, hence of size $\leq 2^{e(G)-1}$.

~~Claim m~~ Each edge is covered $k\leq m/2$ times by $\mathcal{G}$. By cor. above, $|\mathcal{F}|\leq$

$\left|\prod_{G\in\mathcal{G}}2^{e(G)-1}\right|^{1/k}=(2^{s-1})^{m/k}<2^{\binom{n}{2}-2}=\frac{1}{4}2^{\binom{n}{2}}$.