3 Methods of conditional events

3.1 Property B & Recoloring

Recall that

a hypergraph $H = (V, E)$ has Property B, or 2-colorable,
if $\exists$ a 2-coloring of V
s.t. no edge in H is monochromatic.

Have proved: Any k-uniform hypergraph with $|H| < 2^{k-1}$ have property B.

Theorem (Beck '78) If a k-uniform hypergraph H has at most

$$ck^{1/3}(\ln k)^{-1/2}2^{k-1}$$

edges for a constant $c < \sqrt{3}$, then H has property B.

Remark:

Radhakrishnan & Srinivasan (2000)

If

$$|H| \leq 0.7\left(\frac{k}{\ln k}\right)^{1/2} 2^k$$

then then H has property B.

On the other hand,

$$\exists \text{ non-2-colorable } k\text{-uniform}$$
$$\text{hypergraph with } ck^2 2^k \text{ edges}$$

Proof. Ex. (Take random $m = ck^2 2^k$ k-subsets of $\{1, ..., k^2\}$ and use the first moment method.)

Let

$$f(k) = \min\{m : \exists \text{ non-2-colorable } k\text{-uniform hypergraph with } m \text{ edges }\}.$$

Then

$$0.7\left(\frac{k}{\ln k}\right)^{1/2} 2^k \le f(k) \le ck^2 2^k.$$

Proof. Let $p = \ln k/(3k)$ and $t = ck^{1/3}(\ln k)^{-1/2}$, where c is a constant $< \sqrt{3}$. Then

$$2te^{-pk} + t^2 pe^{pk} < 1.$$

We need to show that if

$$|H| \le t2^{k-1}$$

then H has property B.

Color each vertex, randomly and independently, either B (blue) or R (red) with equal probability (as before).

Call this "first coloring"

Let W be the set all vertices which belong to at lease one monochromatic edge.

Independently change the color of each $v \in W$ with probability p

Call this "recoloring"

An edge e after recoloring is monochromatic by two reasons:

1. Event A_e : e is monochromatic in both the first coloring and the recoloring

$$\begin{aligned}
\Pr[A_e] &= \Pr[e \text{ monoch. in first col. }] \\
&\quad \cdot \Big(\Pr[\text{ no color in } e \text{ is changed (in the recol.)}] \\
&\qquad + \Pr[\text{ all colors in } e \text{ is changed (in the recol.)}] \Big) \\
&= 2^{1-k}((1-p)^k + p^k) \\
&\leq 2^{1-k}(2(1-p)^k) \leq 2^{2-k}e^{-pk}
\end{aligned}$$

and

$$\Pr[\exists \text{ such } e] \leq |H| 2^{2-k} e^{-pk} \leq 2te^{-pk}.$$

2. Event B_e : $\exists \emptyset \neq U_0 \subseteq e$ s.t.

$$\text{in the first col, } e \setminus U_0 \text{ was red and } U_0 \text{ was blue,}$$

and

$$\text{in the recol. all col. in } e \setminus U_0 \text{ remains the same and}$$
$$\text{all col. in } U_0 \text{ are changed,}$$

or the same with red/blue reversed.

Notice that this case also requires at least one edge, say f, with

$$f \cap e \neq \emptyset, \; f \cap e \subseteq U_0, \text{ and } f \text{ is blue monoch.}$$

Let $U = U_0 \setminus f$. Then the case implies that $\exists f, U$(possibly $\emptyset$) with $e \cap f \neq \emptyset$ and $U \subseteq e \setminus f$ s.t. the event B_{efU} that,

$$\text{in the first col., } e \setminus (U \cup f) \text{ was red, } U \cup f \text{ was blue}$$

and

$$\begin{gathered}\text{in the recol., all col. in } e \setminus (U \cup f) \text{ remains the same} \\ \text{and all col. in } U \cup (e \cap f) \text{ are changed,}\end{gathered}$$

or the same with red/blue reversed, occurs.

Since

$$\Pr[B_{efU}] \leq 2^{1-2k+|e\cap f|} p^{|e\cap f|+|U|},$$

we have

$$\Pr[\cup_{efU} B_{efU}] \leq \sum_{e} \sum_{f:e\cap f\neq\emptyset} \sum_{U:U\subseteq e\setminus f} \Pr[B_{efU}].$$

It follows that

$$\begin{aligned}
\sum_{U:U\subseteq e\setminus f} \Pr[B_{efU}] &\leq \sum_{u=0}^{k-|e\cap f|} \binom{k-|e\cap f|}{u} \cdot 2^{1-2k+|e\cap f|} p^{|e\cap f|+u} \\
&= 2^{1-2k+|e\cap f|} p^{|e\cap f|} (1+p)^{k-|e\cap f|} \\
&\leq 2^{1-2k} e^{pk} \left(\frac{2p}{1+p}\right)^{|e\cap f|}.
\end{aligned}$$

Thus

$$\sum_{U:U\subseteq e\setminus f} \Pr[B_{efU}] \le 2^{1-2k} e^{pk} \left(\frac{2p}{1+p}\right)^{|e\cap f|} \le 2^{2-2k} e^{pk} p$$

yields

$$\begin{aligned} \Pr[\cup_e B_e] \le \Pr[\cup_{efU} B_{efU}] &\le \sum_e \sum_{f:e\cap f\neq\emptyset} 2^{2-2k} p e^{pk} \\ &\le |H|^2 2^{2-2k} p e^{pk} \\ &\le t^2 p e^{pk}, \end{aligned}$$

and hence

$$\begin{aligned} \Pr[\cup_e A_e \cup \cup_{e,f,U} B_{efu}] &\le \Pr[\cup_e A_e] + \Pr[\cup_{e,f,U} B_{efu}] \\ &\le 2te^{-pk} + t^2 p e^{pk} < 1. \qquad \square \end{aligned}$$

3.2 Independence numbers of sparse graphs

Theorem (Ajtai, Komlós and Szemerédi ('81))
For a triangle-free graph G,

$$E[|I|] \geq \frac{cn \log t}{t}$$

($t = t(G)$: the average degree).

Ex. This implies that

$$R(3,t) \leq \frac{c' t^2}{\log t}.$$

For a graph G, let $\mathcal{I}$ denote the set of all independent sets of G. Take an independent set I in $\mathcal{I}$ uniformly at random so that

$$\Pr[I = J] = \frac{1}{|\mathcal{I}|} \quad \forall J \in \mathcal{I}.$$

I : (uniformly) random independent set

Theorem (Shearer) For a K_r-free graph G,

$$E[|I|] \geq \frac{c_r n \log t}{t \log \log t} \ .$$

Proof of AKS Theorem. (Alon's version of Shearer's proof) Assume $maxdeg(G) \leq 2t$ (WHY?). For $v \in V(G)$, let

$$X_v = \begin{cases} t & \text{if } v \in I \\ |N(v) \cap I| & \text{otherwise,} \end{cases}$$

or equivalently,

$$X_v := t1(v \in I) + \sum_{w \sim v} 1(w \in I) \ .$$

Note that

$$\begin{aligned} \sum_{v \in V(G)} X_v &= t|I| + \sum_{v \in V(G)} \sum_{w \sim v} 1(w \in I) \\ &\leq t|I| + 2t|I| = 3t|I| \end{aligned}$$

Enough to show that

$$E[X_v] \geq c \log t,$$

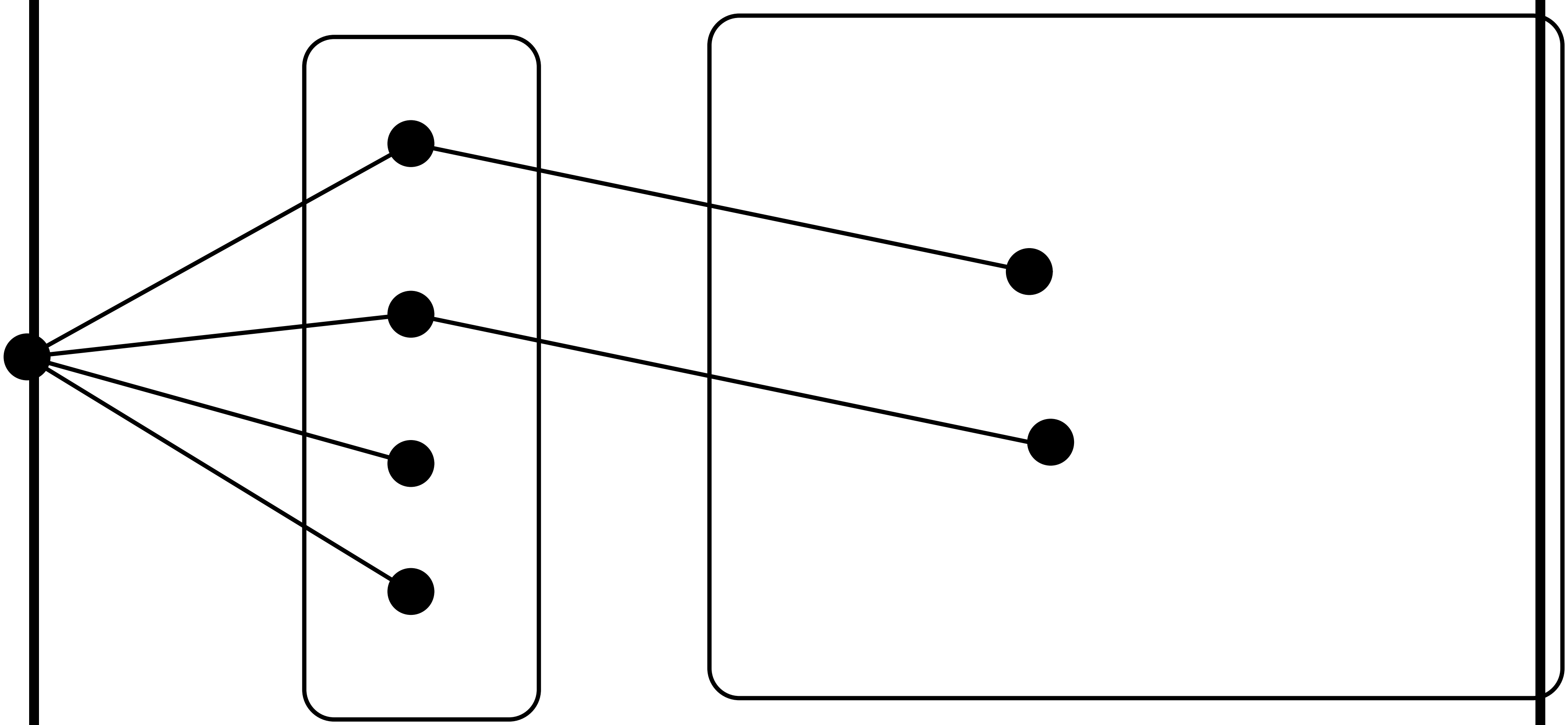
for this implies

$$3tE[|I|] \geq E[\sum_v X_v] \geq cn \log t.$$

Let $J = I \setminus (\{v\} \cup N(v))$, we will show that

$$E[X_v | J] \geq c \log t \ .$$

Let r be the number of vertices in $N(v)$
to which no vertex in J is adjacent.



$$E[X_v|J] = \frac{t}{1+2^r} + \frac{(r/2)2^r}{1+2^r}\ .$$

If $2^r \le t/\log t - 1$, then

$$\frac{t}{1+2^r} \ge \log t \ .$$

If $2^r \ge t/\log t - 1$, then

$$\frac{(r/2)2^r}{1+2^r} \sim r/2 \ge c\log t - \log\log t \ .$$

$\square$

3.3 Covering hypercube

Recall

n-cube Q_n :

$$\{-1,1\}^n = \{(x_i) : x_i = 1 \text{ or } -1, i = 1, ..., n\}$$

Let $X_1, ..., X_m$ be (mutually) independent uniform random vectors in Q_n, in particular,

$$\Pr[X_j = u] = 2^{-n} \quad \text{for any} \quad u \in Q_n.$$

Theorem If $m = (1+\varepsilon)n$ for $\varepsilon > 0$, then

$$\Pr[\exists\, w \in Q_n \text{ with } w \cdot X_j > 0 \ \ \forall j = 1, ..., m] \le 2^{-\varepsilon n} \longrightarrow 0,$$

as $n \to \infty$.

- **Motivation: Neural Networks**

Consider voice recognition (of the brain)

How the brain recognizes

Voice 1, Voice 2, ..., Voice m.

WANT to store)

Voice 1, Voice 2, ..., Voice m so that ... .

Voice 1 = $X^{(1)}$, Voice 2 = $X^{(2)}$, ...

$$X^{(1)} = (1, -1, ..., 1)$$

$$X^{(2)} = (-1, -1, ..., -1)$$

$$\cdots\cdots\cdots$$

To store = To find a weight matrix J (between neurons)

which satisfies certain properties

The weight matrix $J = (J_{ij})$ is a (matrix-valued) function of $X^{(1)}, X^{(2)}, ..., X^{(m)}$.

(We always assume $J_{ii} = 0$.)

A1. Neural network models, or simply neural networks, are interconnected systems of neurons with binary activity. That is,

$$X^{(i)} \in Q_n = \{-1, 1\}^n.$$

A2. A neural network evolves using certain weights, called synaptic weights (or learning rules), between neurons.

A3. Interconnections among the neurons collectively encode information. That is, $J = J(X^{(1)}, ..., X^{(m)})$.

HOW to find the closest $X^{(i)}$???

1. Compute all the Hamming distances

$$d_H(Y, X^{(1)}),\ d_H(Y, X^{(2)}), ...,\ d_H(Y, X^{(m)})$$

2. Evolution

$$\begin{aligned} Y(0) &:= Y, \\ Y(t+1) &:= F_J(Y(t)) \quad for \quad t = 1, 2, ... \end{aligned}$$

O.K. if $Y(t)$ converges to $X^{(i)}$ for some i.

Suppose the weight matrix J is given.

Define $F_J = (f_1, ..., f_n) : Q_n \longrightarrow Q_n$ s.t.

$$f_i(Y) := \text{sgn}(\sum_{j=1}^{n} J_{ij} Y_j) \ ,$$

where

$$\text{sgn}(z) := \begin{cases} 1 & \text{if } z \geq 0 \\ -1 & \text{if } z < 0. \end{cases}$$

HOW to find the closest $X^{(i)}$?

1. Compute all the Hamming distances

$$d_H(Y, X^{(1)}), d_H(Y, X^{(2)}), ..., d_H(Y, X^{(m)}).$$

2. Evolution

$$\begin{aligned} Y(0) &:= Y, \\ Y(t+1) &:= F_J(Y(t)) \quad for \quad t = 1, 2, ... \end{aligned}$$

O.K. if $Y(t)$ converges to $X^{(i)}$ for some i.

• Evolution

$$\begin{aligned} Y(0) &:= Y, \\ Y(t+1) &:= F_J(Y(t)) \quad for \quad t = 1, 2, ... \end{aligned}$$

O.K. if $Y(t)$ converges to $X^{(i)}$ for some i.

That is,

$$Y(t) = Y(t+1) = X^{(i)} \ ,$$

for some t.

Necessary and Sufficient Conditions:

Fixed Point Property:

All $X^{(i)}$ have to be fixed points of F_J, that is,

$$F_J(X^{(i)}) = X^{(i)} \ .$$

Attracting Property:

All $X^{(i)}$ have to be attractive.

In ideal cases,

we need only $O(\log\log n)$ evolutions.

Note that

$$n \sim 10^{11} \ .$$

So

$$\log\log n \sim 3.23 \ .$$

For Fixed Point Property

WANT a weight matrix J such that

$$F_J(X^{(r)}) = X^{(r)} \quad \text{for all} \ \ r = 1, ..., m \ .$$

That is, for all $r = 1, ..., m$, $i = 1, ..., n$,

$$X_i^{(r)} = \mathrm{sgn}(\sum_{j=1}^{n} J_{ij} X_j^{(r)}) \ ,$$

or equivalently,

$$\sum_{j=1}^{n} J_{ij} X_j^{(r)} X_i^{(r)} \geq 0 \quad \text{for all } r = 1, ..., m. \tag{1}$$

($J_{ii} = 0$.)

Let $i = 1$ and $X_1^{(r)} = 1$. Then (1) becomes

$$\sum_{j=2}^{n} J_{1j} X_j^{(r)} \geq 0 \quad \text{for all } r = 1, ..., m.$$

Problem:

$$\text{Is there } w \in Q_n \text{ with } w \cdot X_j > 0 \;\; \forall j = 1, ..., m?$$

Let

$$P_{b,b}(n,m) = \Pr[\exists\, w \in Q_n \text{ with } w \cdot X_j > 0 \;\; \forall j = 1, ..., m,$$

for i.i.d uniform random vector $X_1, ..., X_m$ in Q_n. One may similarly define

$$P_{s,s}(n,m) = \Pr[\exists\, w \in S_{n-1} \text{ with } w \cdot X_j > 0 \;\; \forall j = 1, ..., m,$$

for i.i.d uniform random vector $X_1, ..., X_m$ in S_{n-1}, and $P_{b,s}$, $P_{s,b}$.

Wendel ('62)

$$\mathrm{P}_{s,s}(n,k) = 2^{-k+1}\sum_{i=0}^{n-1}\binom{k-1}{i}\ .$$

Füredi('86)

$$\mathrm{P}_{b,s}(n,k) = 2^{-k+1}\sum_{i=0}^{n-1}\binom{k-1}{i} + O(n^{-1/2})\ .$$

Kahn, Komlós and Szemerédi('93)

$$\begin{aligned}\mathrm{P}_{b,s}(n,k) &= 2^{-k+1}\sum_{i=0}^{n-1}\binom{k-1}{i} \\ &\quad + o((0.99910)^n n^2)\ .\end{aligned}$$

Tao & Vu (2005)

$$\mathrm{P}_{b,s}(n,k) \quad = \quad 2^{-k+1}\sum_{i=0}^{n-1}\binom{k-1}{i} \\ + \ o((3/4)^n n^2)\ .$$

In particular,

$$\begin{aligned} \lim_{n\to\infty} P_{b,s}(n,(2-\varepsilon)n) &= 1 \\ \lim_{n\to\infty} P_{s,s}(n,(2+\varepsilon)n) &= 0 \end{aligned}$$

and

$$\begin{aligned} \lim_{n\to\infty} P_{s,s}(n,(2-\varepsilon)n) &= 1 \\ \lim_{n\to\infty} P_{b,s}(n,(2+\varepsilon)n) &= 0. \end{aligned}$$

(K & Roche '98) For $\varepsilon = 0.0037$,

$$\lim_{n\to\infty} P_{b,b}(n, (1-\varepsilon)n) = 0 \ ,$$

and, for $\rho = 0.005$,

$$\lim_{n\to\infty} P_{b,b}(n, \rho n) = 1 \ .$$

For the proof of

$$\lim_{n\to\infty} P_{b,b}(n,(1-\varepsilon)n) = 0,$$

let $W \in Q_n$ be fixed and assume

$$W \cdot X^{(i)} = \sum_{j=1}^{n} W_j X_{ij} \geq 0, \quad \forall \;\; i = 1, ..., k,$$

where $X_{ij} = X_j^{(i)}$. Then, for

$$U_j := \sum_{i=1}^{k} X_{ij} \quad \text{and} \quad U := (U_j) \; ,$$

we have

$$
\begin{aligned}
\sum_{i=1}^{k} \left| \sum_{j=1}^{n} W_j X_{ij} \right| &= \sum_{i=1}^{k} \sum_{j=1}^{n} W_j X_{ij} \\
&= \sum_{j=1}^{n} \sum_{i=1}^{k} W_j X_{ij} \\
&= \sum_{j=1}^{n} W_j \sum_{i=1}^{k} X_{ij} \\
&= \sum_{j=1}^{n} W_j U_j \\
&= W \cdot U
\end{aligned}
$$

On the other hand, if ε is small enough,

then

$$\sum_{i=1}^{k} |\sum_{j=1}^{n} W_j X_{ij}| \approx \sum_{j} |U_j| \ .$$

Together with

$$\sum_{i=1}^{k} |\sum_{j=1}^{n} W_j X_{ij}| = \sum_{j=1}^{n} W_j U_j \ ,$$

this gives

$$W \approx (\mathrm{sgn}(U_j)) =: \mathrm{sgn} U \ .$$

Thus

$$P_{b,b}(n,(1-\varepsilon)n) \lesssim \sum_{\substack{w \in Q_n \\ w \approx \mathrm{sgn} U}} \Pr[w \cdot X_j \ \ \forall\, j = 1, ..., (1-\varepsilon)n]$$

$$P_{b,b}(n,(1-\varepsilon)n)$$

$$\lesssim \sum_{u\in\mathcal{U}} \Pr[U=u] \sum_{\substack{w\in Q_n \\ w\approx \operatorname{sgn} u}} \Pr(w\cdot X_j \quad \forall\, j=1,\ldots,(1-\varepsilon)n|U=u)$$

Open problems:

Conjecture. There $c > 0$ such that

$$\begin{aligned} \lim_{n\to\infty} P_{b,b}(n,(c-\delta)n) &= 1 \\ \lim_{n\to\infty} P_{b,b}(n,(c+\delta)n) &= 0 \ . \end{aligned}$$

$$c = ??? \quad \text{if exists}$$

Krauth and Opper ('89)

A simulation up to $n \leq 25$ predicts $c \approx .82$.

Krauth and Mézard ('89)

The replica method with so-called symmetry breaking, which is not rigorous, gives $c \approx .83$.

4 Second Moment Method

4.1 Two examples

Let X be a nonnegative integral valued RV.

$$\Pr[X = 0] =???$$

E.g.

$$X = \# \text{ of 2-colorings satisfying certain properties}$$

If $E[X] \to 0$, then

$$0 \leq \Pr[X > 0] \leq E[X] \to 0$$

yields

$$\Pr[X = 0] \to 1.$$

If

$$E[X] = 100 \quad ????$$

For

$$\Pr[X = 99] = \Pr[X = 101] = 1/2,$$

we have

$$\Pr[X = 0] = 0.$$

IF

$$\Pr[Y = 0] = 0.999, \Pr[Y = 10^6] = 10^{-4}$$

then

$$\Pr[Y = 0] = 0.999.$$

Notice that $E[X] = E[Y] = 100$ but

$$\sigma^2(X) = 1, \quad \sigma^2(Y) = 10^{12} \cdot 10^{-4} - (100)^2 = 10^8 - 10^4 \approx 10^8.$$

4.2 Chebyschev's Inequality:

For any positive λ and any RV X,

$$\Pr[|X - \mu| \geq \lambda\sigma] \leq \frac{1}{\lambda^2}$$

Proof. Note that

$$\Pr[|X - \mu| \geq \lambda\sigma] = \Pr[|X - \mu|^2 \geq \lambda^2\sigma^2]$$

Markov Ineq. implies that

$$\Pr[|X - \mu|^2 \geq \lambda^2\sigma^2] \leq \frac{E[|X - \mu|^2]}{\lambda^2\sigma^2} = \frac{1}{\lambda^2}. \qquad \square$$

Corollary For a nonnegative integral valued RV X,

$$\Pr[X = 0] \leq \frac{\sigma^2(X)}{E[X]^2}.$$

Proof. Taking $\lambda = \mu/\sigma$,

$$\Pr[X = 0] \leq \Pr[|X - \mu| \geq \lambda\sigma] \leq 1/\lambda^2 = \sigma^2/\mu^2.$$

$\square$

Corollary If

$$\sigma^2(X) \ll E[X]^2, \quad \text{or equivalently} \quad \frac{\sigma^2(X)}{E[X]^2} \to 0$$

then

$$\Pr[X > 0] \to 1.$$

4.3 Arithmetic Progression

Let $A = A(n,p)$ be a random subset of $\{1, ..., n\}$ such that each $i \in \{1, ..., n\}$ independently belongs to A with probability p, that is,

$$\Pr[i \in A] = p.$$

Theorem For fixed positive integer $k \geq 2$,

$$\Pr[A \text{ contains a } k\text{-term AP}] \to \begin{cases} 0 & \text{if } p \ll n^{-2/k} \\ 1 & \text{if } p \gg n^{-2/k} \end{cases} .$$

The property that A contains a k-term AP has a threshold function $n^{-2/k}$.

Proof. Let $\phi(n,k)$ be the number of k-term AP's in $\{1, ..., n\}$. Then

$$\phi(n,k) = \Theta(n^2).$$

For the set $\{S_1, ..., S_{\phi(n,k)}\}$ of k-term AP's with a certain order, we define

$$X_i = \begin{cases} 1 & \text{if all ele. of } S_i \text{ are in } A \\ 0 & \text{otherwise} \end{cases}$$

and the number of k-term AP's in A

$$X = \sum_i X_i.$$

Then

$$E[X] = \sum_i E[X_i] = \phi(n,k)p^k = \Theta(n^2 p^k).$$

If $p \ll n^{-2/k}$,

$$\Pr[A \text{ contains a } k\text{-term AP}] = \Pr[X > 0] \leq E[X] \longrightarrow 0.$$

If $p \gg n^{-2/k}$,

$$E[X] \longrightarrow \infty,$$

NEED variance: It is enough to show that

$$\sigma^2(X) \ll E[X]^2, \quad \text{or} \quad \frac{\sigma^2(X)}{E[X]^2} \longrightarrow 0.$$

Notice that

$$\begin{aligned} E[X^2] - E[X]^2 &= E[\sum_{i,j} X_i X_j] - \sum_{i,j} E[X_i]E[X_j] \\ &= \sum_{i,j} E[X_i X_j] - E[X_i]E[X_j]. \end{aligned}$$

If $S_i \cap S_j = \emptyset$, then

$$\mathrm{cov}(X_i, X_j) := E[X_i X_j] - E[X_i]E[X_j] = p^{2k} - p^k p^k = 0.$$

If $|S_i \cap S_j| = 1$, then

$$\mathrm{cov}(X_i, X_j) = E[X_i X_j] - E[X_i][X_j] \leq p^{2k-1}.$$

If $|S_i \cap S_j| > 1$, then we use a trivial bound

$$\mathrm{cov}(X_i, X_j) = E[X_i X_j] - E[X_i]E[X_j] \leq E[X_i] = p^k.$$

Using

$$\sum_{i,j} \mathrm{cov}(X_i, X_j) = \sum_i \Big(\sum_{j:S_i \cap S_j = \emptyset} \mathrm{cov}(X_i, X_j)$$

$$+ \sum_{j:|S_i \cap S_j|=1} \mathrm{cov}(X_i, X_j) + \sum_{j:|S_i \cap S_j|>1} \mathrm{cov}(X_i, X_j)\Big),$$

and, for fixed i,

$$|\{j : |S_i \cap S_j| = 1\}| = O(n),$$

$$|\{j : |S_i \cap S_j| > 1\}| = O(1),$$

we have

$$\sum_{i,j} \mathrm{cov}(X_i, X_j) = O(n^3 p^{2k-1} + n^2 p^k).$$

Thus,

$$\sigma^2(X) = O(n^2 p^k + n^3 p^{2k-1}),$$

and $p \gg n^{-2/k}$ yields

$$\frac{\sigma^2(X)}{E[X]^2} = O\Big(\frac{1}{n^2 p^k} + \frac{1}{np}\Big) = o(1).$$

$\square$

4.4 Random graph $G(n,p)$

Each of $\binom{n}{2}$ edges is independently in $G(n,p)$ with pr. p.

For a fixed graph G with m edges,

$$\Pr[G(n,p) = G] = p^m (1-p)^{\binom{n}{2}-m}.$$

Theorem For $G = G(n,p)$,

$$\Pr[\omega(G) \geq 4] \to \begin{cases} 0 & \text{if } p \ll n^{-2/3} \\ 1 & \text{if } p \gg n^{-2/3}. \end{cases}$$

(The property $w(G) \geq 4$ has a threshold function $n^{-2/3}$.)

Proof. Ex.

4.5 Randomized Selection

S: a set of n distinct elements
Select the kth smallest element in S.

Note that there are sorting Algorithms with running time $O(n \log n)$.

Notation: For $t \in S$,

$$r_S(t) = \text{the rank of } t$$
$$S(i) = \text{the element } t \in S \text{ with } r_S(t) = i.$$

LazySelect Algorithm:
Input: A set of n distinct elements.
Output: The kth smallest element in S, $S(k)$.

Algorithm for $k = n/2$:

1. Choose $n^{3/4}$ elements from S uniformly at random with *replacement*. Denoted by T is the (multi)set of the elements.

2. Sort T in $O(n^{3/4} \log n)$ steps using any optimal sorting algorithms.

3. Let $x = kn^{-1/4} = n^{3/4}/2$. For $\ell = \lfloor x - \sqrt{n} \rfloor$ and $h = \lceil x + \sqrt{n} \rceil$, choose $a = T(\ell)$ and $b = T(h)$.
   By comparing a and b with all elements of S, determine $r_S(a)$ and $r_S(b)$.

4. If
$$\frac{1n - 3n^{3/4}}{2} \leq r_S(a) \leq n/2 \leq r_S(b) \leq \frac{n + 3n^{3/4}}{2}, \qquad (2)$$
   then set $P = \{y \in S \mid a \leq y \leq b\}$ and sort P in $O(|P| \log |P|)$ steps to identify $P(n/2 - r_S(a) + 1)$, which is $S(n/2)$.
   If not, repeat Steps 1-3 until (2) holds.

Theorem With probability $1 - O(n^{-1/4})$, the **LazySelect** finds $S(n/2)$ on the first pass through step 1-5. Thus, it performs only $(2 + o(1))n$ comparisons.

Proof. (Probability of failure) .

It is enough to show that

$$\Pr[|r_S(a) - (n/2 - n^{3/4})| \geq n^{3/4}/2] = O(n^{-1/4})$$

and $$\Pr[|r_S(b) - (n/2 + n^{3/4})| \geq n^{3/4}/2] = O(n^{-1/4})$$

(WHY?). We prove only

$$\Pr\left[r_S(a) \leq \frac{n - 3n^{3/4}}{2}\right] = O(n^{-1/4}).$$

Other inequalities may be obtained by similar arguments.

Clearly, $r_S(a) \leq n/2 - 3n^{3/4}/2$ implies that T contains at least $n^{3/4}/2 - n^{1/2}$ elements less than or equal to the $(n/2 - 3n^{3/4}/2)^{\text{th}}$ element in S.

Let $X_i = 1$ if the i^{th} sample of T is less than or equal to $S(n/2 - 3n^{3/4}/2)$, and 0 otherwise. Then X_i's are i.i.d with

$$\Pr[X_i = 1] = \frac{1 - 3n^{-1/4}}{2} \quad \text{and} \quad \Pr[X_i = 0] = \frac{1 + 3n^{-1/4}}{2}.$$

Then, for $X = \sum_{i=1}^{n^{3/4}} X_i$,

$$\Pr\left[r_S(a) \le \frac{n - 3n^{3/4}}{2}\right] \le \Pr\left[X \ge n^{3/4}/2 - n^{1/2}\right].$$

As $E[X] = \frac{n^{3/4} - 3n^{1/2}}{2}$ and

$$\begin{aligned} E[X^2] - E[X]^2 &= \sum_{i,j=1}^{n^{3/4}} E[X_i X_j] - E[X_i]E[X_j] \\ &= \sum_{i}^{n^{3/4}} \frac{1 - 3n^{-1/4}}{2} - \left(\frac{1 - 3n^{-1/4}}{2}\right)^2 \le \frac{n^{3/4}}{3}, \end{aligned}$$

Chebyschev's Inequality yields

$$\Pr\left[X \geq n^{3/4}/2 - n^{1/2}\right] \leq \frac{n^{3/4}/3}{(n^{1/2}/2)^2} = O(n^{-1/4}).$$

$\square$

4.6 The coupon Collector's Problem

There are n types of coupons and at each trial a coupon is chosen uniformly at random.

How many trials are needed to get all coupons?

General example of waiting for combinations of events to happen. Expected case analysis:

Elementary Analysis: For any $0 \leq i \leq n-1$,

$$X_i = \text{number of trials to get } (i+1)^{\text{th}} \text{ new coupon after getting } i \text{ coupons.}$$

Then, $X = \sum_{i=0}^{n-1} X_i$ is the random variable representing the number of trials needed to get all coupons.

- Distribution of X_i:

$$\Pr[X_i = \ell] = p_i (1 - p_i)^{\ell - 1},$$

where the success probability $p_i = \frac{n-i}{n}$.

That is, X_i is geometrically distributed with parameter p_i.

In particular,

$$E[X_i] = \sum_{\ell=1}^{\infty} \ell p_i (1-p_i)^{\ell-1} = 1/p_i = \frac{n}{n-i},$$

and hence

$$E[X] = \sum_{i=0}^{n} \frac{n}{n-i} = n \sum_{i=1}^{n} \frac{1}{i} = nH_n = n \log n + O(n).$$

For the variance of X, notice that X_i are independent. thus

$$\sigma^2(X) = \sum_{i=1}^{n} \sigma^2(X_i).$$

As

$$\sigma^2(X_i) = \sum_{\ell=1}^{\infty} \ell^2 p_i (1-p_i)^{\ell-1} - \frac{1}{p_i^2} = \frac{1-p_i}{p_i^2} = \frac{in}{(n-i)^2},$$

we have

$$\begin{aligned}
\sigma^2(X) &= \sum_{i=0}^{n-1} \frac{in}{(n-i)^2} \sum_{i=1}^{n} \frac{n(n-i)}{i^2} \\
&= n^2 \sum_{i=1}^{n} \frac{1}{i^2} - nH_n \\
&= (1+o(1))n^2 \sum_{i=1}^{n} \frac{1}{i^2}.
\end{aligned}$$

Therefore, the Chebyschev inequality gives

$$X = n \ln n + O(n)$$

with high probability.

4.7 The Coupon Collector's Problem vs. The Occupancy Problem

Occupancy Problems:

Insert each of m balls to n
distinct bins uniformly at random.

Theorem 2 *If $m = n \ln n + cn$, then*

$$\Pr[\ \exists \textit{ empty bin }] \to 1 - e^{-e^{-c}},$$

(as $n \to \infty$).

Corollary 3 *For the number of trials X for the coupon collector's problem and $m = n \ln n + cn$,*

$$\Pr[X > m] \to 1 - e^{-e^{-c}}.$$

Poisson Approximation

- Properties of Poisson random variables

Property 1: If X, Y are independent $\mathrm{Poi}(\lambda)$ and $\mathrm{Poi}(\mu)$, respectively, then $X+Y$ is a $\mathrm{Poi}(\lambda+\mu)$.

Pf.

$$
\begin{aligned}
&\Pr[X+Y=j] \\
&= \sum_{i=0}^{k} \Pr[X=i, Y=j-i] \\
&= \sum_{i=0}^{j} e^{-\lambda}\frac{\lambda^i}{i!} e^{-\mu}\frac{\mu^{j-i}}{(j-i)!} \\
&= e^{-(\lambda+\mu)}\frac{(\lambda+\mu)^j}{j!}\sum_{i=0}^{j}\frac{j!}{i!(j-i)!}\Big(\frac{\lambda}{\lambda+\mu}\Big)^i\Big(\frac{\mu}{\lambda+\mu}\Big)^{j-i} \\
&= e^{-(\lambda+\mu)}\frac{(\lambda+\mu)^j}{j!}\Big(\frac{\lambda}{\lambda+\mu}+\frac{\mu}{\lambda+\mu}\Big)^j.
\end{aligned}
$$

Property 2: The ‘converse’ is also true.

Property 2: The 'converse' is also true.

Let W be a $\mathrm{Poi}(\rho)$. We take W balls and color each ball red with probability p and blue with probability $1-p$, independently of the others. Let $\lambda = p\rho$ and $\mu = (1-p)\rho$. Then the numbers X, Y of red and blue balls, respectively, are independent $\mathrm{Poi}(\lambda)$ and $\mathrm{Poi}(\mu)$, respectively.

Proof. We need to show that

$$\Pr[X = i, Y = j] = e^{-\lambda} \frac{\lambda^i}{i!} e^{-\mu} \frac{\mu^i}{j!}.$$

$$\begin{aligned}
\Pr[X = i, Y = j] &= \Pr[W = i + j] \binom{i+j}{i} p^i (1-p)^j \\
&= e^{-\rho} \frac{\rho^{i+j}}{(i+j)!} \binom{i+j}{i} p^i (1-p)^j \\
&= e^{-\rho} \frac{(p\rho)^i}{i!} \frac{((1-p)\rho)^j}{j!}.
\end{aligned}$$

Using $\rho = \lambda + \mu$ and $p\rho = \lambda, (1-p)\rho = \mu$, we have that

$$\Pr[X = i, Y = j] = e^{-\lambda-\mu} \frac{\lambda^i}{i!} \frac{\mu^j}{j!}$$

Generally,

$$\text{if } X_i\text{'s are independent Poi}(\lambda_i)\text{'s}$$
$$\text{then } \sum X_i \text{ is a Poi}(\sum \lambda_i).$$

Conversely,

Let W be a Poi(ρ).
Take W balls and color each ball i with probability p_i, $\sum p_i = 1$, independently of the others. Denote X_i to be the numbers of balls colored i.
Then X_i's are independent Poi(λ_i)'s, where $\lambda_i = p_i \rho$,

Proof of Theorem 2 Take a Poisson random variable M_1 with mean $m_1 = m - n^{1/2}\ln^2 n$. Notice that

$$\Pr[M_1 \geq m] \to 0.$$

We choose M_1 balls and insert each of them to the n bins uniformly at random. Then, the numbers Y_i of balls in the i^{th} bins are i.i.d Poisson $\lambda_1 := m_1/n = \ln n + c + o(1)$ random variables. Thus

$$\begin{aligned} \Pr[\ \exists \text{ empty bin }\ ] &\leq \Pr_1[\ \exists \text{ empty bin} | M_1 \leq m] \\ &\leq (1 + o(1))(1 - \Pr[Y_i > 0 \ \ \forall\ i]) \end{aligned}$$

For

$$\Pr[Y_i > 0 \ \ \forall\ i] = \Pr[Y_1 > 0]^n = (1 - e^{-\lambda_1})^n,$$

and

$$e^{-\lambda_1} = e^{-\ln n - c + o(1)} = \frac{(1 + o(1))e^{-c}}{n},$$

we have

$$(1 - e^{-\lambda_1})^n = (1 + o(1))e^{-e^{-c}}.$$

Therefore,

$$\Pr[\ \exists \text{ empty bin }\ ] \le (1+o(1))(1-e^{e^{-c}}).$$

Similarly, we may take a Poisson random variable M_2 with mean $m_2 = m + n^{1/2}\ln^2 n$, and choose M_2 balls to obtain

$$\begin{aligned}\Pr[\ \exists \text{ empty bin }\ ] &\ge \Pr_2[\ \exists \text{ empty bin} | M_2 \ge m] \\ &\ge 1 - \Pr[Z_i > 0 \ \ \forall\ i] + o(1),\end{aligned}$$

for the numbers Z_i of balls in the i^{th} bins, which are i.i.d. Poisson random variables with mean $\lambda_2 = m_2/n = \ln n + c + o(1)$.

$\square$

4.8 Perfect matchings in random uniform hypergraphs

List of Papers

Z. Furedi, Random Polytopes in the d-Dimensional Cube, Discrete Comput. Geom. 1: 315-319 (1986).

J. Kim and J. Roche, *Covering Cubes by Random Half Cubes with Applications to Binary Neural Networks*, J. Comput. Syst. Sci. 56(2): 223-252 (1998).

Radhakrishnan & Srinivasan, *Improved Bounds and Algorithms for Hypergraph 2-coloring*, Random Structures & Algorithms 16, 4-32, (2000).