PROBABILISTIC METHOD

Colouring Problem

Theorem 1

Let $A_1, A_2, \ldots, A_n$ be subsets of A and $|A_i| = k$ for $1 \leq i \leq n$. If $n < 2^{k-1}$ then there exists a partition $A = R \cup B$ such that

$$A_i \cap R \neq \emptyset \text{ and } A_i \cap B \neq \emptyset \qquad 1 \leq i \leq n.$$

[R = Red elements and B= Blue elements.]

Proof Randomly colour A.
$\Omega = \{R, B\}^A = \{f : A \to \{R, B\}\}$, uniform distribution.

$$BAD = \{\exists i : \; A_i \subseteq R \text{ or } A_i \subseteq B\}.$$

Claim: $\mathbb{P}(BAD) < 1$.
Thus $\Omega \setminus BAD \neq \emptyset$ and this proves the theorem.

$$BAD(i) = \{A_i \subseteq R \text{ or } A_i \subseteq B\} \text{ and } BAD = \bigcup_{i=1}^{n} BAD(i).$$

Boole's Inequality: if $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_N$ are a collection of events, then

$$\mathbb{P}\left(\bigcup_{i=1}^{N} \mathcal{A}_i\right) \leq \sum_{i=1}^{N} \mathbb{P}(\mathcal{A}_i).$$

This easily proved by induction on N. When $N = 2$ we use

$$\mathbb{P}(\mathcal{A}_1 \cup \mathcal{A}_2) = \mathbb{P}(\mathcal{A}_1) + \mathbb{P}(\mathcal{A}_2) - \mathbb{P}(\mathcal{A}_1 \cap \mathcal{A}_2) \leq \mathbb{P}(\mathcal{A}_1 \cup \mathcal{A}_2).$$

In general,

$$\mathbb{P}\left(\bigcup_{i=1}^{N} \mathcal{A}_i\right) \leq \mathbb{P}\left(\bigcup_{i=1}^{N-1} \mathcal{A}_i\right) + \mathbb{P}(\mathcal{A}_N) \leq \sum_{i=1}^{N-1} \mathbb{P}(\mathcal{A}_i) + \mathbb{P}(\mathcal{A}_N).$$

The first inequality is the two event case and the second is by induction on N.

So,

$$\begin{aligned}
\mathbb{P}(BAD) &\leq \sum_{i=1}^{n} \mathbb{P}(BAD(i)) \\
&= \sum_{i=1}^{n} \left(\frac{1}{2}\right)^{k-1} \\
&= n/2^{k-1} \\
&< 1.
\end{aligned}$$

Example of system which is not 2-colorable.

Let $n = \binom{2k-1}{k}$ and $A = [2k-1]$ and

$$\{A_1, A_2, \ldots, A_n\} = \binom{[2k-1]}{k}.$$

Then in any 2-coloring of $A_1, A_2, \ldots, A_n$ there is a set A_i all of whose elements are of one color.

Suppose A is partitioned into 2 sets R, B. At least one of these two sets is of size at least k (since $(k-1)+(k-1) < 2k-1$). Suppose then that $R \geq k$ and let S be any k-subset of R. Then there exists i such that $A_i = S \subseteq R$.

De-randomising the coloring procedure.

We describe how we can **deterministically** color the elements of A one at a time so that we end up with a coloring satisfying $A_i \cap R \neq \emptyset$ and $A_i \cap B \neq \emptyset$, $1 \leq i \leq n$.

We need some notation: Suppose that we have only colored a subset C of A and $C = R \cup B$ defines the colors of the elemtns in C. (Abusing notation, R, B now refer to a partial coloring of A).

Let $Z(R, B)$ be the number of sets among the A_i that will be mono-colored if we randomly color the remaining elements in $A \setminus (R \cup B)$.

$$Z(R,B)) = \sum_{i=1}^{n} Z_i(R,B)$$

where

$$\mathbb{E}(Z_i(R,B)) = \begin{cases} 1 & A_i \subseteq R \text{ or } A_i \subseteq B \\ 0 & A_i \cap R \neq \emptyset \text{ and } A_i \cap B \neq \emptyset \\ 2^{1-k} & A_i \cap C = \emptyset \\ 2^{-|A_i \setminus C|} & A_i \cap R \neq \emptyset \text{ and } A_i \cap B = \emptyset \\ 2^{-|A_i \setminus C|} & A_i \cap R = \emptyset \text{ and } A_i \cap B \neq \emptyset \end{cases}$$

Initially we have $\mathbb{E}(Z(\emptyset,\emptyset)) < 1$.

Not also that we can compute $\mathbb{E}(Z(R,B))$ in $O(n|A|)$ time.

Suppose now that we have managed to color some of the elements of A and $\mathbb{E}(Z(R,B)) < 1$.

Suppose that x is an arbitrary element of $A \setminus C$. Then if we consider the random color c for x then

$$1 > \mathbb{E}(Z(R,B)) =$$
$$\mathbb{E}(Z(R,B) \mid c = Red)\mathbb{P}(c = Red) +$$
$$\mathbb{E}(Z(R,B) \mid c = Blue)\mathbb{P}(c = Blue) =$$
$$\frac{\mathbb{E}(Z(R \cup \{x\}, B)) + \mathbb{E}(Z(R, B \cup \{x\}))}{2}$$

It follows that at least one of $\mathbb{E}(Z(R \cup \{x\}, B)), \mathbb{E}(Z(R, B \cup \{x\})$ is less than 1.

If $\mathbb{E}(Z(R \cup \{x\}, B)) < 1$ then we color x Red, otherwise we color it Blue.

We continue in this way until we find R, B such that

$$R \cup B = A \text{ and } \mathbb{E}(Z(R, B)) < 1.$$

Now if $R \cup B = A$ then there are no more random choices and $\mathbb{E}(Z(R, B)) = Z(R, B)$ is the number of mono-colored sets.

Since $Z(R, B) < 1$, this number is zero.

Theorem 2

Let $A_1, A_2, \ldots, A_n$ be subsets of A and $|A_i| = k \geq 2$ for $1 \leq i \leq n$. If $n < 2^{k-1}k^{1/4}/3$ then there exists a partition $A = R \cup B$ such that

$$A_i \cap R \neq \emptyset \text{ and } A_i \cap B \neq \emptyset \qquad 1 \leq i \leq n.$$

[R = Red elements and B= Blue elements.]

Randomly order the elements of A as $a_1, a_2, \ldots, a_N$.

Assume that we have colored $a_1, a_2, \ldots, a_{i-1}$. Then we color a_i Red, unless there is an edge A_i for which $A_i \setminus \{a_i\}$ is all Red. In which case, we color a_i Blue.

We now argue that with a positive probability, this algorithm colors A so that no set is mono-colored.

If this fails then there exists j such that A_j is all Blue, by construction. Let v be the first element of A_j to be colored.

Then there exists A_i such that (i) $A_i \cap A_j = \{v\}$ and (ii) v is the last element of A_i to be colored.

Because v is Blue, it is the last element of A_i to be colored. Also (i) holds because all other elements of A_i are Red.

Supppose that $n = 2^{k-1}\ell$. Then the probability of (i), (ii) is at most

$$(2^{k-1}\ell)^2 \cdot \frac{1}{2k-1} \cdot \frac{1}{\binom{2k-2}{k-1}}.$$

For such a pair A_i, A_j we have $|A_i \cup A_j| = 2k-1$. The probability that v is the middle element selected is $1/(2k-1)$ and given this the probability that the first $k-1$ elements of $A_i \cup A_j$ are $A_i \setminus \{v\}$ is $1/\binom{2k-2}{k-1}$.

$(2^{k-1}\ell)^2$ bounds the number of choices for i, j.

Using Stirling's formula $N! \sim (2\pi N)(N/e)^N$ we see that $\binom{2M}{M} \geq 2^M/(3M^{1/2})$ for all M.

It follows that the probability of failure is bounded by

$$2^{2k-2}\ell^2 \cdot \frac{3(k-1)^{1/2}}{2^{2k-2}(k-1)} = \frac{3\ell^2}{(k-1)^{1/2}} < 1,$$

if $\ell \leq k^{1/4}/3$.

Tournaments

n players in a tournament each play each other i.e. there are $\binom{n}{2}$ games.

Fix some k. Is it possible that for every set S of k players there is a person w_S who beats everyone in S?

Suppose that the results of the tournament are decided by a random coin toss.

Fix S, $|S| = k$ and let $\mathcal{E}_S$ be the event that nobody beats everyone in S.

The event

$$\mathcal{E} = \bigcup_S \mathcal{E}_S$$

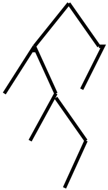
is that there is a set S for which w_S does not exist.

We only have to show that $\mathbb{P}(\mathcal{E}) < 1$.

$$
\begin{aligned}
\mathbb{P}(\mathcal{E}) &\leq \sum_{|S|=k} \mathbb{P}(\mathcal{E}_S) \\
&= \binom{n}{k}(1-2^{-k})^{n-k} \\
&< n^k e^{-(n-k)2^{-k}} \\
&= \exp\{k \ln n - (n-k)2^{-k}\} \\
&\to 0
\end{aligned}
$$

since we are assuming here that k is fixed independent of n.

Random Binary Search Trees



A binary tree consists of a set of *nodes*, one of which is the *root*. Each node is connected to 0,1 or 2 nodes below it and every node other than the root is connected to exactly one node above it. The root is the highest node.

The depth of a node is the number of edges in its path to the root.

The depth of a tree is the maximum over the depths of its nodes.

Starting with a tree T_0 consisting of a single root r, we grow a tree T_n as follows:

The n'th *particle* starts at r and flips a fair coin. It goes left (L) with probability 1/2 and right (R) with probability 1/2.

It tries to move along the tree in the chosen direction. If there is a node below it in this direction then it goes there and continues its random moves. Otherwise it creates a new node where it wanted to move and stops.

Let D_n be the depth of this tree.
Claim: for any $t \geq 0$,

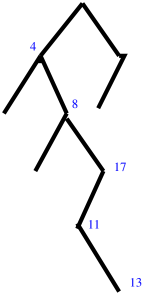
$$\mathbb{P}(D_n \geq t) \leq (n2^{-(t-1)/2})^t.$$

Proof The process requires at most n^2 coin flips and so we let $\Omega = \{L, R\}^{n^2}$ – most coin flips will not be needed most of the time.

$$DEEP = \{D_n \geq t\}.$$

For $P \in \{L, R\}^t$ and $S \subseteq [n]$, $|S| = t$ let $DEEP(P, S)$ = {the particles $S = \{s_1, s_2, \ldots, s_t\}$ follow P in the tree i.e. the first i moves of s_i are along P, $1 \leq i \leq t$}.

$$DEEP = \bigcup_P \bigcup_S DEEP(P, S).$$



S={4,8,11,13,17}

t=5 and DEEP(P,S) occurs if

4 goes L...

8 goes LR...

17 goes LRR...

11 goes LRRL...

13 goes LRRLR...

$$
\begin{aligned}
\mathbb{P}(DEEP) &\leq \sum_P \sum_S \mathbb{P}(DEEP(P,S)) \\
&= \sum_P \sum_S 2^{-(1+2+\cdots+t)} \\
&= \sum_P \sum_S 2^{-t(t+1)/2} \\
&= 2^t \binom{n}{t} 2^{-t(t+1)/2} \\
&\leq 2^t n^t 2^{-t(t+1)/2} \\
&= (n2^{-(t-1)/2})^t.
\end{aligned}
$$

So if we put $t = A\log_2 n$ then

$$
\mathbb{P}(D_n \geq A\log_2 n) \leq (2n^{1-A/2})^{A\log_2 n}
$$

which is very small, for $A > 2$.

Secretary Problem

There are n applicants for a secretarial position and CEO Pat will interview them in random order. The rule is that Pat must decide on the spot whether to hire the current applicant or interview the next one. Pat is an excellent judge of quality, but she does not know the set of applicants a priori. She wants to give herself a good chance of hiring the best.

Here is her strategy: She chooses a number $m < n$, interviews the first m and then hires the first person in $m+1, \ldots, n$ who is the best so far. (There is a chance that she will not hire anyone).

Let S be the event that Pat chooses the best person and let P_i be the event that the best person is the ith applicant. Then

$$\mathbb{P}(S) = \sum_{i=1}^{n} \mathbb{P}(S \mid P_i)\mathbb{P}(P_i) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{P}(S \mid P_i).$$

Now Pat's strategy implies that $\mathbb{P}(S \mid P_i) = 0$ for $1 \le i \le m$. If P_i occurs for $i > m$ then Pat will succeed iff the best of the first $i-1$ applicants (j say) is one of the first m, otherwise Pat will mistakenly hire j. Thus, for $i > m$, $\mathbb{P}(S \mid P_i) = \frac{m}{i-1}$ and hence

$$\mathbb{P}(S) = \frac{m}{n}\sum_{i=m+1}^{n} \frac{1}{i-1}.$$

Now assume that n is large and that $m = \alpha n$. Then

$$\mathbb{P}(S) \sim \alpha(\ln n - \ln \alpha n) = \alpha \ln 1/\alpha.$$

Pat will want to choose the value of α that maximises $f(\alpha) = \alpha \ln 1/\alpha$. But $f'(\alpha) = \ln 1/\alpha - 1$ and so the optimum choice for α is $1/e$. In which case,

$$\mathbb{P}(S) \sim e^{-1}.$$

A problem with hats

There are n people standing a circle. They are blind-folded and someone places a hat on each person's head. The hat has been randomly colored Red or Blue.

They take off their blind-folds and everyone can see everyone else's hat. Each person then simultaneously declares (i) my hat is red or (ii) my hat is blue or (iii) or I pass.

They win a big prize if the people who opt for (i) or (ii) are all correct. They pay a big penalty if there is a person who incorrectly guesses the color of their hat.

Is there a strategy which means they will win with probability better than 1/2?

Suppose that we partition $Q_n = \{0,1\}^n$ into 2 sets W, L which have the property that L is a cover i.e. if $x = x_1 x_2 \cdots x_n \in W = Q_n \setminus L$ then there is $y_1 y_2 \cdots y_n \in L$ such that $h(x, y) = 1$ where

$$h(x, y) = |\{j : x_j \neq y_j\}|.$$

Hamming distance between x and y.

Assume that $0 \equiv Red$ and $1 \equiv Blue$. Person i knows x_j for $j \neq i$ (color of hat j) and if there is a unique value ξ of x_i which places x in W then person i will declare that their hat has color ξ.

The people assume that $x \in W$ and if indeed $x \in W$ then there is at least one person who will be in this situation and any such person will guess correctly.

Is there a small cover L?

Let $p = \frac{\ln n}{n}$. Choose L_1 randomly by placing $y \in Q_n$ into L_1 with probability p.

Then let L_2 be those $z \in Q_n$ which are not at Hamming distance ≤ 1 from some member of L_1.

Clearly $L = L_1 \cup L_2$ is a cover and $\mathbb{E}(|L|) = 2^n p + 2^n (1-p)^{n+1} \leq 2^n (p + e^{-np}) \leq 2^n \frac{2 \ln n}{n}$.

So there must exist a cover of size at most $2^n \frac{2 \ln n}{n}$ and the players can win with probability at least $1 - \frac{2 \ln n}{n}$.

Markov Inequality: if X is a non-negative random variable with mean μ then for any $x \geq 0$

$$\mathbb{P}(X \geq x) \leq \frac{\mu}{x}.$$

First moment method: if X is a random variable taking values in $\{0, 1, \ldots\}$ then

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

We just apply the Markov inequality with $x = 1$.

$$\begin{aligned} \mathbb{E}(X) &= \mathbb{E}(X \mid X = 0)\mathbb{P}(X = 0) + \mathbb{E}(X \mid X \geq 1)\mathbb{P}(X \geq 1) \\ &\geq \mathbb{P}(X \geq 1). \end{aligned}$$

Intersection Safe Families

Let $\mathcal{A}$ be a family of sub-sets of $[n]$. We say that $\mathcal{A}$ is *Intersection Safe* if for distinct $A, B, C \in \mathcal{A}$ we have $C \not\supseteq A \cap B$. We use the probabilistic method to show the existence of an Intersection Safe family of exponential size.
Suppose that $\mathcal{A}$ consists of p randomly and independently chosen sets $X_1, X_2, \ldots, X_p$. Let Z denote the number of 3-tples i, j, k such that $X_i \cap X_j \subseteq X_k$. Then

$$\mathbb{E}(Z) = p(p-1)(p-2)\mathbb{P}(X_i \cap X_j \subseteq X_k) = p(p-1)(p-2)\left(\frac{7}{8}\right)^n.$$

(Observe that $\mathbb{P}(x \in (X_i \cap X_j) \setminus X_k = 1/8.$)

So if $p \leq (8/7)^{n/3}$ then

$$\mathbb{P}(Z \geq 1) \leq \mathbb{E}(Z) < p^3 \left(\frac{7}{8}\right)^n \leq 1$$

implying that there exists a union free family of size p.

There is a small problem here in that we might have repetitions $X_i = X_j$ for $i \neq j$. Then our set will not be of size p.

But if Z_1 denotes the number of pairs i, j such that $X_i = X_j$ then

$$\mathbb{P}(Z_1 \neq 0) \leq \mathbb{E}(Z_1) = \binom{p}{2} 2^{-n}$$

and so we should really choose p so that
$\mathbb{P}(Z + Z_1 \neq 0) \leq \mathbb{E}(Z) + \mathbb{E}(Z_1) < p^3 \left(\frac{7}{8}\right)^n + p^2 \left(\frac{1}{2}\right)^n \leq 1.$

Application: Suppose that we have a central storage containing n keys $\{k_1, k_2, \ldots, k_n\}$.

We must distribute sets of keys to p people. Person i will get the set $K_i = \{k_j : j \in X_i\}$. The sets $X_1, X_2, \ldots, X_p$ are public knowledge.

If person r wishes to communicate with person s then he/she will send them $\{k_j : j \in X_r \cap X_s\}$ as a means of proving their identity.

If the sets $X_1, X_2, \ldots, X_p$ are intersection safe, then person t cannot pretend to be person r.

It is possible therefore to have a “secure“ system with p people that requires each person to get $O(\ln p)$ keys.

Graph Crossing Number

The crossing number of a graph G is the minimum number of edge crossings of a drawing of G in the plane.

Euler's forula implies that a planar graph with n vertices has at most $3n$ edges.

This implies that a graph $G = (V, E)$ requires at least $|E| - 3|V|$ crossings.

Theorem 3

If $|E| > 4|V|$ then G has crossing number $\Omega(|E|^3/|V|^2)$.

If $|E| \approx |V|^{3/2}$ then this gives $\Omega(|V|^{5/2})$ whereas $|E| - 3|V| = O(|V|^{3/2})$.

Proof

Suppose that G has a drawing with k crossings and let $0 < p < 1$.

Let $G_p = (V_p, E_p)$ denote the subgraph of G obtained by including each vertex in V_p independently with probability p.

E_p is then the set of edges $\{x, y\}$ such that $x, y \in V_p$.

$$\mathbb{E}(|V_p|) = p|V| \text{ and } \mathbb{E}(|E_p|) = p^2|E|.$$

Also,

$$\mathbb{E}(\text{number of crossings in the drawing of } G_p) = p^4 k.$$

So,

$$p^4 k \geq \mathbb{E}(|E_p| - 3|V_p|) = p^2|E| - 3p|V|.$$

So

$$k \geq \frac{p^2|E| - 3p|V|}{p^4}.$$

Maximising the RHS over $p \leq 1$ gives $p = 4|V|/|E|$ and

$$k \geq \frac{|E|^3}{64|V|^2}.$$

$\square$

Connectivity of a Random Graph

The random graph $G_{n,p}$ is defined as follows: each edge of the complete graph K_n is included independently with probability p.

Thus if $G = (V, E)$ has vertex set $V = [n]$ then

$$\mathbb{P}(G_{n,p} = G) = p^{|E|}(1-p)^{\binom{n}{2}-|E|}.$$

Theorem 4

Let $p = \frac{\log n + c_n}{n}$. *Then*

$$\lim_{n\to\infty} \mathbb{P}(G_{n,p} \text{ is connected}) = \begin{cases} 0 & c_n \to -\infty. \\ e^{-e^{-c}} & c_n \to c. \\ 1 & c_n \to \infty. \end{cases}$$

Proof

A vertex of $G_{n,p}$ is said to isolated if it has degree zero. If $G_{n,p}$ has isolated vertices then it is not connected. Let X_1 denote the number of isolated vertices in $G_{n,p}$. Therefore

$$\mathbb{P}(X_1 > 0) \leq \mathbb{P}(G_{n,p} \text{ is not connected})$$
$$\leq \mathbb{P}(X_1 > 0) + \mathbb{P}\left(G_{n,p} \text{ has component with } 2 \leq k \leq \frac{n}{2} \text{ vertices}\right).$$

It will suffice to prove that if $c_n \to c$ then

$$\lim_{n\to\infty} \mathbb{P}(G_{n,p} \text{ is connected}) = e^{-e^{-c}}.$$

This is because if $p' \geq p$ then

$$\mathbb{P}(G_{n,p} \text{ is connected}) \leq \mathbb{P}(G_{n,p'} \text{ is connected}).$$

$e^{-e^{-c_n}} \to 0$ if $c_n \to -\infty$ and $e^{-e^{-c_n}} \to 1$ if $c_n \to \infty$.

Let X_k denote the number of components in $G_{n,p}$ with k vertices. Then

$$\mathbb{P}\left(G_{n,p} \text{ has component with } 2 \le k \le \frac{n}{2} \text{ vertices}\right) = \mathbb{P}\left(\exists 2 \le k \le \frac{n}{2} : X_k > 0\right).$$

So, using Boole's inequality and the first moment method,

$$\mathbb{P}\left(\exists 2 \le k \le \frac{n}{2} : X_k > 0\right) \le \sum_{k=2}^{n/2} \mathbb{P}(X_k > 0) \le \sum_{k=2}^{n/2} \mathbb{E}(X_k)$$
$$\le \sum_{k=2}^{n/2} \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)} = \sum_{k=2}^{n/2} u_k.$$

The factor k^{k-2} in the second line is the number of spanning trees of K_k. In effect we are bounding the number of connected components by the number of spanning trees in such components.

Now, for $2 \le k \le 10$,

$$u_k \le e^k n^k \left(\frac{\log n + c}{n}\right)^{k-1} e^{-k(n-10)(\log n + c)/n}$$
$$\le (1+o(1)) e^{k(1-c)} \left(\frac{\log n}{n}\right)^{k-1},$$

and for $k > 10$

$$u_k \le \left(\frac{ne}{k}\right)^k k^{k-2} \left(\frac{\log n + c}{n}\right)^{k-1} e^{-k(\log n + c)/2}$$
$$\le n \left(\frac{e^{1-c/2+o(1)} \log n}{n^{1/2}}\right)^k.$$

So

$$\sum_{k=2}^{n/2} u_k \le (1+o(1)) \frac{e^{2(1-c)} \log n}{n} + \sum_{k=10}^{n/2} n^{1+o(1)-k/2}$$
$$= O\left(n^{o(1)-1}\right).$$

We have so far proved that if $p = (\log n + c)/n$ then

$\mathbb{P}(G_{n,p} \text{ is connected}) = (1+o(1))\mathbb{P}(G_{n,p} \text{ has no isolated vertices.})$

We use inclusion-exclusion to estimate the RHS probability.

Let $\mathcal{A}$ denote the set of graphs with vertex set $[n]$. For a graph $G = (V, E) \in \mathcal{A}_i$ we let $w_G = p^{|E|}(1-p)^{\binom{n}{2}-|E|}$.

For $i \in [n]$, we let $\mathcal{A}_i = \{G : i \text{ isolated in } G\}$.

Then

$$\mathbb{P}(G_{n,p} \text{ has no isolated vertices}) = w\left(\bigcap_{i=1}^{n} \bar{\mathcal{A}}_i\right).$$

If $S \subseteq [n]$ then

$$w(\mathcal{A}_S) = \mathbb{P}(\text{ the vertices in } S \text{ are isolated}) = (1-p)^{|S|(n-|S|)+\binom{|S|}{2}}.$$

So, by inclusion-exclusion,

$$\begin{aligned} w\left(\bigcap_{i=1}^{n} \bar{\mathcal{A}}_i\right) &= \sum_{S \subseteq [n]} (-1)^{|S|}(1-p)^{|S|(n-|S|)+\binom{|S|}{2}} \\ &= \sum_{s=0}^{n} \binom{n}{s} (-1)^s (1-p)^{s(n-s)+\binom{s}{2}}. \end{aligned}$$

Now fix an arbitrary positive integer, independent of n. The Bonferroni inequalities imply that

$$
\begin{aligned}
w\left(\bigcap_{i=1}^{n} \bar{A}_i\right) &\leq \sum_{s=0}^{2k} \binom{n}{s} (-1)^s (1-p)^{s(n-s)+\binom{s}{2}} \\
&= \left(1 + O\left(\frac{k^2 \log n}{n}\right)\right) \sum_{s=0}^{2k} \frac{n^s}{s!} (-1)^s e^{-snp} \\
&= \left(1 + O\left(\frac{k^2 \log n}{n}\right)\right) \sum_{s=0}^{2k} \frac{n^s}{s!} (-1)^s e^{-s(\log n + c)} \\
&= \left(1 + O\left(\frac{k^2 \log n}{n}\right)\right) \sum_{s=0}^{2k} \frac{(-1)^s e^{-cs}}{s!}.
\end{aligned}
$$

So, for all positive integers k,

$$
\limsup_{n\to\infty} w\left(\bigcap_{i=1}^{n} \bar{A}_i\right) \leq \sum_{s=0}^{2k} \frac{(-1)^s e^{-cs}}{s!}.
$$

Similarly, for all positive integers k,

$$\liminf_{n\to\infty} w\left(\bigcap_{i=1}^{n} \bar{\mathcal{A}}_i\right) \geq \sum_{s=0}^{2k-1} \frac{(-1)^s e^{-cs}}{s!}.$$

But,

$$\lim_{k\to\infty} \sum_{s=0}^{2k} \frac{(-1)^s e^{-cs}}{s!} = \lim_{k\to\infty} \sum_{s=0}^{2k-1} \frac{(-1)^s e^{-cs}}{s!} = e^{-e^{-c}}.$$

And so

$$\lim_{n\to\infty} w\left(\bigcap_{i=1}^{n} \bar{\mathcal{A}}_i\right) = e^{-e^{-c}}.$$

It is instructive to (i) use the first moment method to show that if $c_n \to \infty$ then $\mathbb{P}(X_1 > 0) = o(1)$ and (ii) use the second moment method to prove that if $c_n \to -\infty$ then $\mathbb{P}(X_1 > 0) = 1 - o(1)$.
(i) Suppose that $c_n = \omega \to \infty\,, \omega = O(\log n)$:

$$\mathbb{P}(X_1 > 0) \leq \mathbb{E}(X_1) = n(1-p)^{n-1} \leq ne^{-(n-1)(\log n+\omega)/n}$$
$$= n\exp\left\{-\log n - \omega + \frac{\log n + \omega}{n}\right\} \leq (1+o(1))e^{-\omega} \to 0.$$

(ii) Second Moment Method (Paley-Zygmund).

Suppose that X is a random variable taking values in $\{0,1,2,\ldots\}$. Then

$$\mathbb{P}(X \geq 1) \geq \frac{\mathbb{E}(X)^2}{\mathbb{E}(X^2)}.$$

Let $Y = 1_{X \geq 1}$. The $Y^2 = Y$ and $XY = X$. Applying the Cauchy-Schwartz inequality

$$\mathbb{E}(X)^2 = \mathbb{E}(XY)^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2) = \mathbb{E}(X^2)\mathbb{E}(Y) = \mathbb{E}(X^2)\mathbb{P}(Y = 1) = \mathbb{E}(X^2)\mathbb{P}(X \geq 1).$$

We apply this to X_1 – the number of isolated vertices.

$$\mathbb{E}(X_1) = n(1-p)^{n-1} = n \exp\left\{-(n-1)\sum_{k=1}^{\infty} \frac{p^k}{k}\right\} \geq n \exp\left\{-(n-1)(p+p^2)\right\} = (1-o(1))e^{\omega} \to \infty.$$

$$\mathbb{E}(X_1^2) = \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{P}(i \text{ and } j \text{ are both isolated}) = \mathbb{E}(X_1) + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n} \mathbb{P}(i \text{ and } j \text{ are both isolated})$$

If $i \neq j$ then

$$\mathbb{P}(i \text{ and } j \text{ are both isolated}) = (1-p)^{2n-3}.$$

So,

$$\begin{aligned}\frac{\mathbb{E}(X_1^2)}{\mathbb{E}(X_1)^2} &= \frac{\mathbb{E}(X_1) + n(n-1)(1-p)^{2n-3}}{n^2(1-p)^{2n-2}} \\ &= \frac{1}{\mathbb{E}(X)} + \frac{(n-1)(1-p)}{n} \\ &\to 1.\end{aligned}$$

So, of course,

$$\frac{\mathbb{E}(X_1)^2}{\mathbb{E}(X_1^2)} \to 1.$$

Average case of Quicksort

Quicksort is an algorithm for sorting numbers. Given distinct $x_1, x_2, \ldots, x_n$ we

1. Randomly choose an integer p between 1 and and n – the pivot.
2. Divide the remaining numbers into 2 sets L, R where $L = \{x_j : \ x_j < x_p\}$ and $R = \{x_j : \ x_j > x_p\}$.
3. Recursively sort L, R.

Let T_n be the expected number of comparisons taken by Quicksort.

We have $T_0 = 0$ and for $n \geq 1$

$$T_n =$$

$$\sum_{i=1}^{n} \mathbb{E}(\textit{No. comparisons} \mid p \textit{ is } i'\textit{th largest})\mathbb{P}(p \textit{ is } i'\textit{th largest}) =$$

$$\sum_{i=1}^{n} (n - 1 + T_{i-1} + T_{n-i}) \times \frac{1}{n}$$

$$= n - 1 + \frac{2}{n} \sum_{i=0}^{n-1} T_i$$

or

$$nT_n = n(n-1) + 2 \sum_{i=0}^{n-1} T_i.$$

Let $T(x) = \sum_{n=0}^{\infty} T_n x^n$ be the generating function for T_n.

We note that

$$\sum_{n=1}^{\infty} n T_n x^n = xT'(x).$$

$$\sum_{n=1}^{\infty} n(n-1)x^n = \frac{2x^2}{(1-x)^3}.$$

$$\sum_{n=1}^{\infty} \left(\sum_{i=0}^{n-1} T_i \right) x^n = \frac{xT(x)}{1-x}.$$

Thus,

$$T'(x) = \frac{2x}{(1-x)^3} + \frac{2T(x)}{1-x}$$

or

$$(1-x)^2 T'(x) - 2(1-x)T(x) = \frac{2x}{1-x}$$

or

$$\frac{d}{dx}((1-x)^2 T(x)) = \frac{2x}{1-x}$$

and so

$$(1-x)^2 T(x) = C - 2x - 2\ln(1-x).$$

$$(1-x)^2T(x) = C - 2x - 2\ln(1-x).$$

Now $T(0) = 0$ implies that $C = 0$ and so

$$\begin{aligned} T(x) &= -\frac{2x}{(1-x)^2} - \frac{2\ln(1-x)}{(1-x)^2} \\ &= -2\sum_{n=0}^{\infty} nx^n + 2\sum_{n=0}^{\infty}\left(\sum_{k=1}^{n}\frac{n-k+1}{k}\right)x^n \end{aligned}$$

So

$$\begin{aligned} T_n &= -4n + 2(n+1)\sum_{k=1}^{n}\frac{1}{k} \\ &\approx 2n\ln n. \end{aligned}$$

Hashing

Let $U = \{0, 1, \ldots, N-1\}$ and $H = \{0, 1, \ldots, n-1\}$ where n divides N and $N \gg n$. $f : U \to H$, $f(u) = u \mod n$.
(H is a hash table and U is the universe of objects from which a subset is to be stored in the table.)
Suppose $u_1, u_2, \ldots, u_m,\ m = \alpha n$, are a random subset of U. A copy of u_i is stored in "cell" $f(u_i)$ and u_i's that "hash" to the same cell are stored as a linked list.
Questions: u is chosen uniformly from U.
(i) What is the expected time T_1 to determine whether or not u is in the table?
(ii) If it is given that u is in the table, what is the expected time T_2 to find where it is placed?
Time = The number of comparisons between elements of U needed.

Let $M = N/n$, the average number of $u's$ that map to a cell. Let X_k denote the number of u_i for which $f(u_i) = k$. Then

$$\begin{aligned}
\mathbb{E}(T_1) &= \sum_{k=1}^{n} \mathbb{E}(T_1 \mid f(u) = k)\mathbb{P}(f(u) = k) \\
&= \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}(T_1 \mid f(u) = k) \\
&\leq \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}(X_k) \\
&= \frac{1}{n} \mathbb{E}\left(\sum_{k=1}^{n} X_k\right) \\
&= \alpha.
\end{aligned}$$

Let X denote $X_1, X_2, \ldots, X_n$ and let $\mathcal{X}$ denote the set of possible values for X. Then

$$\begin{aligned}
\mathbb{E}(T_2) &= \sum_{X \in \mathcal{X}} \mathbb{E}(T_2 \mid X)\mathbb{P}(X) \\
&= \sum_{X \in \mathcal{X}} \sum_{k=1}^{n} \mathbb{E}(T_2 \mid f(u) = k, X)\mathbb{P}(f(u) = k)\mathbb{P}(X) \\
&= \sum_{X \in \mathcal{X}} \sum_{k=1}^{n} \mathbb{E}(T_2 \mid f(u) = k, X)\frac{X_k}{m}\mathbb{P}(X) \\
&= \sum_{X \in \mathcal{X}} \sum_{k=1}^{n} \left(\frac{1 + X_k}{2}\right) \frac{X_k}{m}\mathbb{P}(X) \\
&= \frac{1}{2m} \sum_{X \in \mathcal{X}} \sum_{k=1}^{n} X_k(1 + X_k)\mathbb{P}(X)
\end{aligned}$$

$$
\begin{aligned}
\mathbb{E}(T_2) &= \frac{1}{2} + \frac{1}{2M}\mathbb{E}(X_1^2 + \cdots + X_n^2) \\
&= \frac{1}{2} + \frac{1}{2\alpha}\mathbb{E}(X_1^2) \\
&= \frac{1}{2} + \frac{1}{2\alpha}\sum_{t=1}^{m} t^2 \frac{\binom{M}{t}\binom{N-M}{m-t}}{\binom{N}{m}}.
\end{aligned}
$$

If α is small and t is small then we can write

$$\begin{aligned}\frac{\binom{M}{t}\binom{N-M}{m-t}}{\binom{N}{m}} &\approx \frac{M^t}{t!}\frac{(N-M)^{m-t}}{(m-t)!}\frac{m!}{N^m} \\ &\approx \left(1-\frac{1}{n}\right)^m \frac{m^t}{t!n^t} \\ &\approx \frac{\alpha^t e^{-\alpha}}{t!}.\end{aligned}$$

Then we can further write

$$\mathbb{E}(T_2) \approx \frac{1}{2} + \frac{1}{2\alpha}\sum_{t=1}^{\infty} t^2 \frac{\alpha^t e^{-\alpha}}{t!} = 1 + \frac{\alpha}{2}$$

Finding Minimum

Consider the following program which computes the minimum of the n numbers $x_1, x_2, \ldots, x_n$.

begin
$min := \infty$;
for $i = 1$ **to** n **do**
begin
if $x_i < min$ **then** $min := x_i$
end
output min
end

If the x_i are all different and in random order, what is the expected number of times that that the statement $min := x_i$ is executed?

Ω = {permutations of $1, 2, \ldots, n$} – uniform distribution.
Let X be the number of executions of statement $min := x_i$. Let

$$X_i = \begin{cases} 1 & \text{statement executed at } i. \\ 0 & \text{otherwise} \end{cases}$$

Then $X_i = 1$ iff $x_i = \min\{x_1, x_2, \ldots, x_i\}$ and so

$$\mathbb{P}(X_i = 1) = \frac{(i-1)!}{i!} = \frac{1}{i}.$$

[The number of permutations of $\{x_1, x_2, \ldots, x_i\}$ in which x_i is the largest is $(i-1)!$.]

So

$$\begin{aligned}
\mathbb{E}(X) &= \mathbb{E}\left(\sum_{i=1}^{n} X_i\right) \\
&= \sum_{i=1}^{n} \mathbb{E}(X_i) \\
&= \sum_{i=1}^{n} \frac{1}{i} \qquad (= \; H_n) \\
&\approx \log_e n.
\end{aligned}$$

Inequalities

Markov Inequality: let $X : \Omega \to \{0, 1, 2, \ldots, \}$ be a random variable. For any $t \geq 1$

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}(X)}{t}.$$

Proof

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{\infty} k\mathbb{P}(X = k) \\
&\geq \sum_{k=t}^{\infty} k\mathbb{P}(X = k) \\
&\geq \sum_{k=t}^{\infty} t\mathbb{P}(X = k) \\
&= t\mathbb{P}(X \geq t).
\end{aligned}$$

In particular, if $t = 1$ then $\mathbb{P}(X \neq 0) \leq \mathbb{E}(X)$.

Chebycheff Inequality

Now let $\sigma = \sqrt{\textbf{Var}(Z)}$.

$$\begin{aligned}\mathbb{P}(|Z-\mu| \geq t\sigma) &= \mathbb{P}((Z-\mu)^2 \geq t^2\sigma^2) \\ &\leq \frac{\mathbb{E}((Z-\mu)^2)}{t^2\sigma^2} \\ &= \frac{1}{t^2}.\end{aligned}$$

(1) comes from the Markov inequality applied to the random variable $(Z-\mu)^2$.

Back to Binomial: $\sigma = \sqrt{np(1-p)}$.

$$\mathbb{P}(|B_{n,p} - np| \geq t\sqrt{np(1-p)}) \leq \frac{1}{t^2}$$

which implies

$$\mathbb{P}(|B_{n,p} - np| \geq \epsilon np) \leq \frac{1}{\epsilon^2 np}$$

[Law of large numbers.]

Hoeffding's Inequality – I

Let $X_1, X_2, \ldots, X_n$ be independent random variables taking values such that $\mathbb{P}(X_i = 1) = 1/2 = \mathbb{P}(X_i = -1)$ for $i = 1, 2, \ldots, n$. Let $X = X_1 + X_2 + \cdots + X_n$. Then for any $t \geq 0$

$$\mathbb{P}(|X| \geq t) < 2e^{-t^2/2n}.$$

Proof: For any $\lambda > 0$ we have

$$\begin{aligned} \mathbb{P}(X \geq t) &= \mathbb{P}(e^{\lambda X} \geq e^{\lambda t}) \\ &\leq e^{-\lambda t}\mathbb{E}(e^{\lambda X}). \end{aligned}$$

Now for $i = 1, 2, \ldots, n$ we have

$$\mathbb{E}(e^{\lambda X_i}) = \frac{e^{-\lambda} + e^{\lambda}}{2} = 1 + \frac{\lambda^2}{2!} + \frac{\lambda^4}{4!} + \cdots < e^{\lambda^2/2}.$$

So, by independence,

$$\mathbb{E}(e^{\lambda X}) = \mathbb{E}\left(\prod_{i=1}^{n} e^{\lambda X_i}\right) = \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i}) \le e^{\lambda^2 n/2}.$$

Hence,

$$\mathbb{P}(X \ge t) \le e^{-\lambda t + \lambda^2 n/2}.$$

We choose $\lambda = t/n$ to minimise $-\lambda t + \lambda^2 n/2$. This yields

$$\mathbb{P}(X \ge t) \le e^{-t^2/2n}.$$

Similarly,

$$\begin{aligned} \mathbb{P}(X \le -t) &= \mathbb{P}(e^{-\lambda X} \ge e^{\lambda t}) \\ &\le e^{-\lambda t}\mathbb{E}(e^{-\lambda X}) \\ &\le e^{-\lambda t + \lambda^2 n/2}. \end{aligned}$$

Discrepancy

Suppose that $|X| = n$ and $\mathcal{F} \subseteq \mathcal{P}(X)$. If we color the elements of X with Red and Blue i.e. partition X in $R \cup B$ then the discrepancy $disc(\mathcal{F}, R, B)$ of this coloring is defined

$$disc(\mathcal{F}, R, B) = \max_{F \in \mathcal{F}} disc(F, R, B)$$

where $disc(F, R, B) = ||R \cap F| - |B \cap F||$ i.e. the absolute difference between the number of elements of F that are colored Red and the number that are colored Blue.

Claim:

If $|\mathcal{F}| = m$ then there exists a coloring R, B such that $disc(\mathcal{F}, R, B) \leq (2n\log_e(2m))^{1/2}$.

Proof Fix $F \in \mathcal{F}$ and let $s = |F|$. If we color X randomly and let $Z = |R \cap F| - |B \cap F|$ then Z is the sum of s independent ± 1 random variables.

So, by the Hoeffding inequality,

$$\mathbb{P}(|Z| \geq (2n\log_e(2m))^{1/2}) < 2e^{-n\log_e(2m)/s} \leq \frac{1}{m}.$$

Switching Game:

We are given an $n \times n$ matrix A where $A(i,j) = \pm 1$. We interpret $A(i,j) = 1$ as the light at i,j is on.

Now suppose that $x, y \in \{\pm 1\}^n$ are switches. The light at i,j is on if $A(i,j)x_i y_j = 1$ and off otherwise.

Let $\sigma(A) = \max_{x,y} \left| \sum_{i,j} A(i,j)x_i y_j \right|$ be the maximum absolute difference between the number of lights which are on and those that are off, obtaianble by switching.

Claim: There exists A such that $\sigma(A) \leq cn^{3/2}$ where $c = 2(\ln 2)^{1/2}$.

Fix $x, y \in \{\pm 1\}^n$ and let A be a random ± 1 matrix. Consider the random variable

$$Z_{x,y} = \sum_{i,j} A(i,j) x_i y_j.$$

This is the sum of n^2 independent random variables ($A(i,j)x_i y_j$) taking values in ± 1.

It follows from the Hoeffding inequality that

$$|Z_{x,y}| \geq cn^{3/2} < 2e^{-(cn^{3/2})^2/2n^2} = 2^{-2n}$$

So

$$\mathbb{P}(\max_{x,y} |Z_{x,y}| \geq cn^{3/2}) < 2^n \times 2^n \times = 2^{-2n} = 1.$$

Hence there exists A such that $\sigma(A) \leq cn^{3/2}$.

Hoeffding's Inequality – II

Now let $S_n = X_1 + X_2 + \cdots + X_n$ where $X_i, i = 1, \ldots, n$ are independent random variables where $0 \leq X_i \leq 1$ and $\mathbb{E}X_i = \mu_i$ for $i = 1, 2, \ldots, n$. Let $\mu = \mu_1 + \mu_2 + \cdots + \mu_n$. Then for $\lambda \geq 0$

$$\mathbb{P}(S_n \geq \mu + t) \leq e^{-\lambda(\mu+t)} \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i}) \tag{2}$$

and for $\lambda \leq 0$

$$\mathbb{P}(S_n \leq \mu - t) \leq e^{-\lambda(\mu-t)} \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i}). \tag{3}$$

Now the convexity of e^x and $0 \leq X_i \leq 1$ implies that

$$e^{\lambda X_i} \leq 1 - X_i + X_i e^{\lambda}.$$

Taking expectations we get

$$\mathbb{E}(e^{\lambda X_i}) \leq 1 - \mu_i + \mu_i e^{\lambda}.$$

Equation (2) becomes, for $\lambda \geq 0$,

$$\mathbb{P}(S_n \geq \mu + t) \leq e^{-\lambda(\mu+t)} \prod_{i=1}^{n} (1 - \mu_i + \mu_i e^{\lambda})$$
$$\leq e^{-\lambda(\mu+t)} \left(\frac{n - \mu + \mu e^{\lambda}}{n} \right)^n . \quad (4)$$

The second inequality follows from the fact that the geometric mean is at most the arithmetic mean i.e.

$$(x_1 x_2 \cdots x_n)^{1/n} \leq \frac{x_1 + x_2 + \cdots + x_n}{n}$$

for non-negative $x_1, x_2, \ldots, x_n$.
The right hand side of (4) attains its minimum, as a function of λ, at

$$e^{\lambda} = \frac{(\mu + t)(n - \mu)}{(n - \mu - t)\mu}. \quad (5)$$

Hence, by (4) and (5), assuming that $\mu + t < n$,

$$\mathbb{P}(S_n \geq \mu + t) \leq \left(\frac{\mu}{\mu + t}\right)^{\mu+t} \left(\frac{n-\mu}{n-\mu-t}\right)^{n-\mu-t},$$

while for $t > n - \mu$ this probability is zero.
Now let

$$0 \leq \phi(x) = (1+x)\log(1+x) - x, \quad x \geq -1,$$

and let $\phi(x) = \infty$ for $x < -1$. Now, for $0 \leq t \leq n - \mu$, we can rewrite the bound (68) as

$$\mathbb{P}(S_n \geq \mu + t) \leq \exp\left\{-\mu\phi\left(\frac{t}{\mu}\right) - (n-\mu)\phi\frac{-t}{n-\mu}\right\}.$$

Since $\phi(x) \geq 0$ for every $x \geq -1$, we get

$$\mathbb{P}(S_n \geq \mu + t) \leq e^{-\mu\phi(t/\mu)}. \tag{6}$$

Similarly, putting $n - S_n$ for S_n, or by an analogous argument, using (3), we get for $0 \leq t \leq \mu$,

$$\mathbb{P}(S_n \leq \mu - t) \leq \exp\left\{-\mu\phi\left(\frac{-t}{\mu}\right) - (n-\mu)\phi\left(\frac{t}{n-\mu}\right)\right\}.$$

Hence,

$$\mathbb{P}(S_n \leq \mu - t) \leq e^{-\mu\phi(-t/\mu)}. \tag{7}$$

We can simplify the expressions (6) and (7) by observing that

$$\phi(x) \geq \frac{x^2}{2(1+\frac{x}{3})}. \tag{8}$$

To see this observe that for $|x| \leq 1$ we have

$$\phi(x) - \frac{x^2}{2(1+\frac{x}{3})} = \sum_{k=2}^{\infty} (-1)^k \left(\frac{1}{k(k-1)} - \frac{1}{2 \cdot 3^{k-2}} \right) x^k.$$

Equation (8) for $|x| \leq 1$ follows from $\frac{1}{k(k-1)} - \frac{1}{2 \cdot 3^{k-2}} \geq 0$ for $k \geq 2$.

For $x \geq 1$ we let $f(x) = \phi(x) - \frac{x^2}{2(1+\frac{x}{3})}$ and then check that $f'(1) \geq 0$ and

$$f'(x) = \log(1+x) - 3 + \frac{9}{3+x} + \frac{3x^2}{(3+x)^2} \geq \log 2 - 0.75 + \frac{3x^2}{(3+x)^2} \geq 0$$

for $x \geq 1$.

Taking this into account we arrive at the following theorem,

Theorem 5

Suppose that $S_n = X_1 + X_2 + \cdots + X_n$ *where (i)* $0 \le X_i \le 1$ *and* $\mathbb{E}X_i = \mu_i$ *for* $i = 1, 2, \ldots, n$, *(ii)* $X_1, X_2, \ldots, X_n$ *are independent. Let* $\mu = \mu_1 + \mu_2 + \cdots + \mu_n$. *Then for* $t \ge 0$,

$$\mathbb{P}(S_n \ge \mu + t) \le \exp\left\{ -\frac{t^2}{2(\mu + \frac{t}{3})} \right\}$$

and for $t \le \mu$,

$$\mathbb{P}(S_n \le \mu - t) \le \exp\left\{ -\frac{t^2}{2(\mu - \frac{t}{3})} \right\}. \tag{9}$$

Putting $t = \epsilon\mu$, for $0 < \epsilon < 1$, one can immediately obtain the following bounds.

Corollary 6

Let $0 < \epsilon < 1$, *then*

$$\mathbb{P}(S_n \geq (1+\epsilon)\mu) \leq \left(\frac{e^{\epsilon}}{(1+\epsilon)^{1+\epsilon}}\right)^{\mu} \leq \exp\left\{-\frac{\mu\epsilon^2}{3}\right\},$$

while

$$\mathbb{P}(S_n \leq (1-\epsilon)\mu) \leq \exp\left\{-\frac{\mu\epsilon^2}{2}\right\}$$

The formula (6) follows directly from (9) and (6) follows from (68).

For large deviations we have the following result.

Corollary 7

If $c > 1$ *then*

$$\mathbb{P}(S_n \geq c\mu) \leq \left\{ \frac{e}{ce^{1/c}} \right\}^{c\mu} \leq \left(\frac{e}{c} \right)^{c\mu}. \tag{10}$$

Put $t = (c - 1)\mu$ into (6).

Valiant-Brebner routing algorithm:

Let $Q_n = (V_n = \{0,1\}^n, E_n)$ be the n-cube where $(\mathbf{x}, \mathbf{y}) \in E_n$ iff

$$h(\mathbf{x}, \mathbf{y}) = |\{j : x_j \neq y_j\}| = 1.$$

Given a permutation $\pi : V_n \to V_n$ we wish to synchronously send a packet $p_{\mathbf{x}}$ from $\mathbf{x}$ to $\pi(\mathbf{x})$ along a path of Q_n for all $\mathbf{x} \in V_n$.

At most one packet can cross any edge in one time step. Packets form a queue waiting to cross.

Bit Fixing Path: Given $\mathbf{x}, \mathbf{y}$ let $\mathbf{z}_i = (y_1, y_2, \ldots, y_i, x_{i+1}, \ldots, x_n)$. Let $BFP(\mathbf{x}, \mathbf{y})$ be the path $(\mathbf{x}, \mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_n = y)$.

- Step For each $\mathbf{x} \in V_n$ choose $\delta(\mathbf{x})$ independently and randomly from V_n.
- Step Send packet $p_{\mathbf{x}}$ to $\delta(\mathbf{x})$ along the path $P(\mathbf{x}) = BFP(\mathbf{x}, \delta(\mathbf{x}))$.
- Step Send packet $p_{\mathbf{x}}$ to $\pi(\mathbf{x})$ along the path $Q(\mathbf{x}) = BFP(\delta(\mathbf{x}), \pi(\mathbf{x}))$.

Note that for a given $\mathbf{x}$, $P(\mathbf{x}) = (\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n = \delta(\mathbf{x}))$ where $\mathbf{x}_{i+1}$ is obtained from $\mathbf{x}_i$ by flipping the $i+1$th bit of $\mathbf{x}_i$ with probability 1/2. Note also that the length of $P(\mathbf{x})$ is equal to the number of i such that $\mathbf{x}_{i+1} \neq \mathbf{x}_i$.

Let $D(\mathbf{x})$ be the time spent by $p_{\mathbf{x}}$ waiting in a queue. Let

$$S(\mathbf{x}) = \{\mathbf{y} \neq \mathbf{x} : P(x) \cap P(\mathbf{y}) \neq \emptyset\}.$$

Observation: $D(\mathbf{x}) \leq |S(\mathbf{x})|$. This follows from $|P(\mathbf{x}) \cap P(\mathbf{y})| \leq 1$ for all $\mathbf{x}, \mathbf{y}$. This follows from the fact that paths can meet, continue together for a while and diverge. They cannot meet again once they diverge.

We claim that

$$\mathbb{P}(|S(\mathbf{x})| \geq 3n) \leq 3^{-n}. \tag{11}$$

It follows from this that

$$\mathbb{P}(\exists \mathbf{x} : \text{Step 2 takes more than 4n time}) \leq 2^n \times 3^{-n} = o(1).$$

Step 3 can be analysed similarly.

Proof of (11)**:** We write

$$|S(\mathbf{x})| = \sum_{\mathbf{y}\in V_n\setminus\{\mathbf{x}\}} Z_{\mathbf{y}}$$

where

$$Z_{\mathbf{y}} = \mathbf{1}_{P_{\mathbf{y}}\cap P_{\mathbf{x}}\neq\emptyset}.$$

Observe next that if $\mathbf{u},\mathbf{v}$ are chosen randomly from V_n then

$$\sum_{\mathbf{x},\mathbf{y}\in V_n} \mathbb{P}(P(\mathbf{x})\cap P(\mathbf{y})\neq\emptyset) = 2^{2n}\mathbb{P}(P(\mathbf{v})\cap P(\mathbf{u})\neq\emptyset)$$

$$\leq 2^{2n}\sum_{k=1}^{n}\frac{1}{2^{k+1}}\cdot\frac{1}{2^{n-k}}$$

$$= 2^{n-1}n.$$

Here $P(\mathbf{v})\cap P(\mathbf{u})\neq\emptyset$ only if there is some k such that $\mathbf{u},\mathbf{v}$ agree on the last $n-k$ components and $P(\mathbf{u})$ amends the first $k+1$ components of $\mathbf{u}$ so that they agree with the first $k+1$ components of the $k+1$st vertex of $P(\mathbf{v})$.

On the other hand, symmetry yields that for any fixed $\mathbf{x} \in V_n$,

$$\sum_{\mathbf{x},\mathbf{y} \in V_n} \mathbb{P}(P(\mathbf{x}) \cap P(\mathbf{y}) \neq \emptyset) = 2^n \sum_{\mathbf{y} \in V_n} \mathbb{P}(P(\mathbf{x}) \cap P(\mathbf{y}) \neq \emptyset)$$
$$\geq 2^n \mathbb{E}(|S(\mathbf{x})|).$$

It follows that

$$\mathbb{E}(|S(\mathbf{x})|) \leq \frac{n}{2}.$$

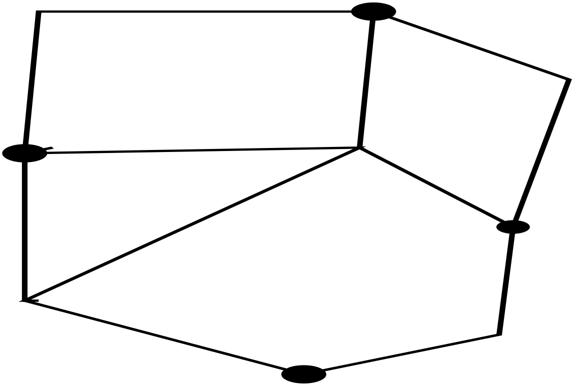
Applying the Chernoff bound (10) we see that

$$\mathbb{P}(|S(\mathbf{x})| \geq \alpha n) \leq \left(\frac{e}{2\alpha}\right)^{\alpha n}$$

and we obtain (11) by putting $\alpha = 3$.

Independent sets and cliques

$S \subseteq V$ is *independent* if no edge of G has both of its endpoints in S.



$\alpha(G)$=maximum size of an independent set of G.

Theorem 8

If graph G has n vertices and m edges then

$$\alpha(G) \geq \frac{n^2}{2m+n}.$$

Note that this says that $\alpha(G)$ is at least $\frac{n}{d+1}$ where d is the average degree of G.

Proof Let $\pi(1), \pi(2), \ldots, \pi(\nu)$ be an arbitrary permutation of V. Let $N(v)$ denote the set of neighbours of vertex v and let

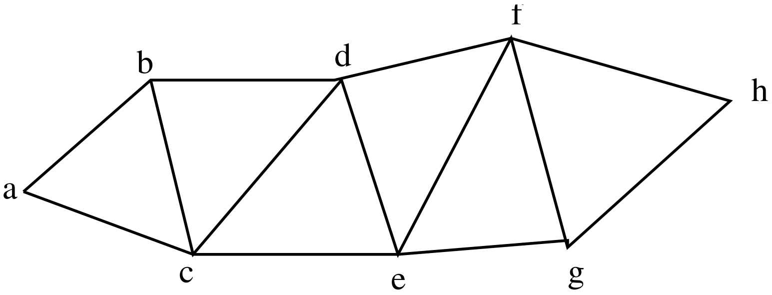
$$I(\pi) = \{v : \ \pi(w) > \pi(v) \text{ for all } w \in N(v)\}.$$

Claim 1

I is an independent set.

Proof of Claim 1

Suppose $w_1, w_2 \in I(\pi)$ and $w_1 w_2 \in E$. Suppose $\pi(w_1) < \pi(w_2)$. Then $w_2 \notin I(\pi)$ — contradiction. □



	a	b	c	d	e	f	g	h	l
π_1	c	b	f	h	a	g	e	d	$\{c, f\}$
π_2	g	f	h	d	e	a	b	c	$\{g, a\}$

Claim 2

If π is a random permutation then

$$\mathbb{E}(|I|) = \sum_{v \in V} \frac{1}{d(v)+1}.$$

Proof: Let $\delta(v) = \begin{cases} 1 & v \in I \\ 0 & v \notin I \end{cases}$

Thus

$$\begin{aligned} |I| &= \sum_{v \in V} \delta(v) \\ \mathbb{E}(|I|) &= \sum_{v \in V} \mathbb{E}(\delta(v)) \\ &= \sum_{v \in V} \mathbb{P}(\delta(v) = 1). \end{aligned}$$

Now $\delta(v) = 1$ iff v comes before all of its neighbours in the order π. Thus

$$\mathbb{P}(\delta(v) = 1) = \frac{1}{d(v) + 1}$$

and the claim follows. $\square$

Thus there exists a π such that

$$|I(\pi)| \geq \sum_{v \in V} \frac{1}{d(v) + 1}$$

and so

$$\alpha(G) \geq \sum_{v \in V} \frac{1}{d(v) + 1}.$$

We finish the proof of the theorem by showing that

$$\sum_{v \in V} \frac{1}{d(v)+1} \geq \frac{n^2}{2m+n}.$$

This follows from the following claim by putting $x_v = d(v)+1$ for $v \in V$.

Claim 3

If $x_1, x_2, \ldots x_k > 0$ *then*

$$\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_k} \geq \frac{k^2}{x_1 + x_2 + \cdots + x_k}.$$

Proof

Multiplying (3) by $x_1 + x_2 + \cdots + x_k$ and subtracting k from both sides we see that (3) is equivalent to

$$\sum_{1 \le i < j \le k} \left(\frac{x_i}{x_j} + \frac{x_j}{x_i} \right) \ge k(k-1).$$

But for all $x, y > 0$

$$\frac{x}{y} + \frac{y}{x} \ge 2$$

and (86) follows. $\square$

Corollary 9

If G contains no clique of size k then

$$m \leq \frac{(k-2)n^2}{2(k-1)}$$

For example, if G contains no triangle then $m \leq n^2/4$.

Proof Let $\bar{G}$ be the *complement* of G i.e. $G + \bar{G} = K_n$.

By assumption

$$k - 1 \geq \alpha(\bar{G}) \geq \frac{n^2}{n(n-1) - 2m + n}.$$

$\square$

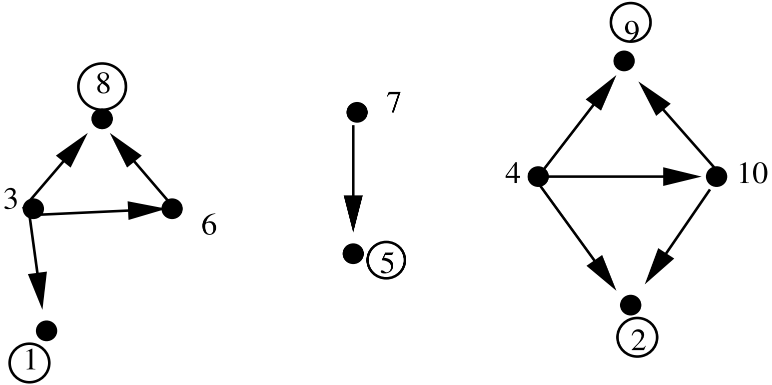
Parallel searching for the maximum – Valiant

We have n processors and n numbers $x_1, x_2, \ldots, x_n$. In each round we choose n pairs i, j and compare the values of x_i, x_j. The set of pairs chosen in a round can depend on the results of previous comparisons.

Aim: find i^* such that $x_{i^*} = \max_i x_i$.

Claim 4

For any algorithm there exists an input which requires at least $\frac{1}{2}\log_2\log_2 n$ rounds.



Suppose that the first round of comparisons involves comparing x_i, x_j for edge ij of the above graph and that the arrows point to the larger of the two values. Consider the independent set $\{1, 2, 5, 8, 9\}$. These are the indices of the 5 largest elements, but their relative order can be arbitrary since there is no implied relation between their values.

Let $C(a,b)$ be the maximum number of rounds needed for a processors to compute the maximum of b values in this way.

Lemma 10

$$C(a,b) \geq 1 + C\left(a, \left\lceil \frac{b^2}{2a+b} \right\rceil\right).$$

Proof The set of b comparisons defines a b-edge graph G on a vertices where comparison of x_i, x_j produces an edge ij of G. Now,

$$\alpha(G) \geq \left\lceil \frac{b}{\frac{2a}{b}+1} \right\rceil = \left\lceil \frac{b^2}{2a+b} \right\rceil.$$

For any independent set I it is always possible to define values for $x_1, x_2, \ldots, x_a$ such I is the index set of the $|I|$ largest values and so that the comparisons do not yield any information about the ordering of the elements $x_i, i \in I$.
Thus after one round one has the problem of finding the maximum among $\alpha(G)$ elements. □
Now define the sequence $c_0, c_1, \ldots$ by $c_0 = n$ and

$$c_{i+1} = \left\lceil \frac{c_i^2}{2n + c_i} \right\rceil .$$

It follows from the previous lemma that

$$c_k \geq 2 \text{ implies } C(n, n) \geq k + 1.$$

Claim 4 now follows from

Claim 5

$$c_i \geq \frac{n}{3^{2^i-1}}.$$

By induction on i. Trivial for $i = 0$. Then

$$\begin{aligned}
c_{i+1} &\geq \frac{n^2}{3^{2^{i+1}-2}} \times \frac{1}{2n + \frac{n}{3^{2^i-1}}} \\
&= \frac{n}{3^{2^{i+1}-1}} \times \frac{3}{2 + \frac{1}{3^{2^i-1}}} \\
&\geq \frac{n}{3^{2^{i+1}-1}}.
\end{aligned}$$

□

The Local Lemma

We go back to the coloring problem at the beginning of these slides. We now place a different restriction on the sets involved.

Theorem 11

Let $A_1, A_2, \ldots, A_n$ be subsets of A where $|A_i| \geq k$ for $1 \leq i \leq n$. If each A_i intersects at most 2^{k-3} other sets then there exists a partition $A = R \cup B$ such that

$$A_i \cap R \neq \emptyset \text{ and } A_i \cap B \neq \emptyset \qquad 1 \leq i \leq n.$$

Symmetric Local Lemma: We consider the following situation. $X = \{x_1, x_2, \ldots, x_N\}$ is a collection of independent random variables. Suppose that we have events $\mathcal{E}_i, i = 1, 2, \ldots, m$ where $\mathcal{E}_i$ depends only on the set $X_i \subseteq X$. Thus if $X_i \cap X_j = \emptyset$ then $\mathcal{E}_i$ and $\mathcal{E}_j$ are independent.

The dependency graph Γ has vertex set $[m]$ and an edge (i, j) iff $X_i \cap X_j \neq \emptyset$.

Theorem 12

Let

$$p = \max_i \mathbb{P}(\mathcal{E}_i) \text{ and let } d \text{ be the maximum degree of } \Gamma.$$

$$4dp \leq 1 \text{ implies that } \mathbb{P}\left(\bigcap_{i=1}^{m} \bar{\mathcal{E}}_i\right) \geq (1 - 2p)^m > 0.$$

Proof: We prove by induction on $|S|$ that for any i,

$$\mathbb{P}\left(\mathcal{E}_i \middle| \bigcap_{j \in S} \bar{\mathcal{E}}_j\right) \leq 2p. \tag{12}$$

Notice that this suffices, since

$$\mathbb{P}\left(\bigcap_{i=1}^{m} \bar{\mathcal{E}}_i\right) = \prod_{i=1}^{m} \mathbb{P}\left(\bar{\mathcal{E}}_i \middle| \bigcap_{j=1}^{i-1} \bar{\mathcal{E}}_j\right)$$

The base case $|S| = 0$ for (12) is trivial.

Inductive Step: Renumber for convenience so that $i = n, S = [s]$ and $(i, x) \notin \Gamma$ for $x > d$. Now

$$\mathbb{P}\left(\mathcal{E}_n \middle| \bigcap_{i=1}^{s} \bar{\mathcal{E}}_i\right) = \frac{\mathbb{P}\left(\mathcal{E}_n \cap \bigcap_{i=1}^{d} \bar{\mathcal{E}}_i \middle| \bigcap_{i=d+1}^{s} \bar{\mathcal{E}}_i\right)}{\mathbb{P}\left(\bigcap_{i=1}^{d} \bar{\mathcal{E}}_i \middle| \bigcap_{i=d+1}^{s} \bar{\mathcal{E}}_i\right)},$$

$$\leq \frac{\mathbb{P}\left(\mathcal{E}_n \middle| \bigcap_{i=d+1}^{s} \bar{\mathcal{E}}_i\right)}{\mathbb{P}\left(\bigcap_{i=1}^{d} \bar{\mathcal{E}}_i \middle| \bigcap_{i=d+1}^{s} \bar{\mathcal{E}}_i\right)},$$

$$\leq \frac{\mathbb{P}\left(\mathcal{E}_n \middle| \bigcap_{i=d+1}^{s} \bar{\mathcal{E}}_i\right)}{1 - \sum_{i=1}^{d} \mathbb{P}\left(\mathcal{E}_i \middle| \bigcap_{i=d+1}^{s} \bar{\mathcal{E}}_i\right)}. \qquad (13)$$

Now

$$\mathbb{P}\left(\mathcal{E}_n \middle| \bigcap_{i=d+1}^{s} \bar{\mathcal{E}}_i\right) = \mathbb{P}(\mathcal{E}_n) \leq p, \tag{14}$$

since $\mathcal{E}_n$ is independent of $\mathcal{E}_{d+1}, \ldots, \mathcal{E}_s$.

Furthermore, we can assume that $d > 0$, else the events $\mathcal{E}_1, \ldots, \mathcal{E}_m$ are independent and the result is trivial. So, by induction, we have that

$$1 - \sum_{i=1}^{d} \mathbb{P}\left(\mathcal{E}_i \middle| \bigcap_{i=d+1}^{s} \bar{\mathcal{E}}_i\right) \geq 1 - 2dp \geq \frac{1}{2}. \tag{15}$$

The induction is now completed by using (14) and (15) in (13).

Proof of Theorem 11: We randomly color the elements of A Red and Blue. Let $\mathcal{E}_i$ be the event that A_i is mono-colored. Clearly, $\mathbb{P}(\mathcal{E}_i) \leq 2^{-(k-1)}$. Thus,

$$p \leq 2^{-(k-1)}.$$

The degree of vertex i of Γ is the number of j such that $A_i \cap A_j \neq \emptyset$. So, by assumption,

$$d \leq 2^{k-3}.$$

Theorem 12 implies that $\mathbb{P}\left(\bigcap_{i=1}^{n} \bar{\mathcal{E}}_i\right) > 0$ and so the required coloring exists.

Theorem 13

Let $G = (V, E)$ be an r-regular graph. If r is sufficiently large, then E can be partitioned into E_1, E_2 so that if $G_i = (V, E_i), i = 1, 2$ then

$$\frac{r}{2} - (20r \log r)^{1/2} \leq \delta(G_i) \leq \Delta(G_i) \leq \frac{r}{2} + (20r \log r)^{1/2}.$$

Proof: We randomly partition the edges of G by independently placing e into E_1 E_1 with probability 1/2. For $v \in V$, we let $\mathcal{E}_v$ be the event that the degree $d_1(v)$ in G_1 satisfies

$$d_1(v) \notin \left[\frac{r}{2} - (3r \log r)^{1/2}, \frac{r}{2} + (3r \log r)^{1/2}\right].$$

It follows from Hoeffding's Inequality - I with $t = (3r \log r)^{1/2}$ that

$$\mathbb{P}(\mathcal{E}_v) \le 2e^{-t^2/2r} = 2r^{-3/2}.$$

Furthermore, $\mathcal{E}_v$ is independent of the events $\mathcal{E}_w$ for vertices w at distance 2 or more from v in G. Thus,

$$d \le r.$$

Clearly, $4 \cdot 2r^{-3/2} \cdot r \le 1$ for r large and the result follows from Theorem 12. I.e. $\mathbb{P}\left(\bigcap_{v \in V} \bar{\mathcal{E}}_v\right) > 0$ which imples that there exists a partition where none of the events $\mathcal{E}_v, v \in V$ occur.

For the next application, let $D = (V, E)$ be a k-regular digraph. By this we mean that each vertex has exactly k in-neighbors and k out-neighbors.

Theorem 14

Every k-regular digraph has a collection of $\lfloor k/(4 \log k) \rfloor$ vertex disjoint cycles.

Proof: Let $r = \lfloor k/(4 \log k) \rfloor$ and color the vertices of D with colors $[r]$. For $v \in V$, let $\mathcal{E}_v$ be the event that there is a color missing at the out-neighbors of v. We will show that $\mathbb{P}\left(\bigcap_{v \in V} \bar{\mathcal{E}}_v\right) > 0$.

Suppose then that none of the events $\mathcal{E}_v, v \in V$ occur. Consider the graph D_j induced by a single color $j \in [r]$. Note that D_j is not the empty graph. Let $P_j = (v_1, v_2, \ldots, v_m)$ be a longest directed path in D_j. Let w be an out-neighbor of v_m of color j. We must have $w \in \{v_1, \ldots, v_m\}$, else P_j is not a longest path in D_j. Thus each $D_j, j \in [r]$ contains a cycle and these cycles are vertex disjoint.

We first estimate

$$\mathbb{P}(\mathcal{E}_v) \leq r\left(1 - \frac{1}{r}\right)^k \leq ke^{-k/r} \leq ke^{-4\log k} = k^{-3}.$$

On the other hand, if $N^+(v)$ denotes the out-neighbors of v plus v then $\mathcal{E}_v$ is independent of all events $\mathcal{E}_w$ for which $N^+(v) \cap N^+(w) = \emptyset$. It follows that

$$d \leq k^2.$$

To apply Theorem 12 we need to have $4k^{-3}k^2 \leq 1$. This is true for $k \geq 4$. For $k \leq 3$ we have $r = 1$ and the local lemma is not needed.

Constructive version: Moser/Tardos

Suppose now that in the context of Theorem 12 we have

$$p\frac{d^d}{(d-1)^{d-1}} \leq 1. \tag{16}$$

(Notice that this is a weaker assumption than $4dp \leq 1$.)

Algorithm MT:

- Assign values to $x_1, x_2, \ldots, x_N$.
- **While** $\exists j : \mathcal{E}_j$ holds **do**
- Pick smallest j such that $\mathcal{E}_j$ occurs.
- Randomly re-set $\{x \in X_j\}$.
- **od**

Theorem 15

Assuming (16) *holds, algorithm MT that finds an assignment of values to* $x \in X$ *such that* $\bigcap_{i=1}^{m} \bar{\mathcal{E}}_i$ *holds in* $O(N)$ *expected number of iterations.*

Let j_t be the value of j in Step 3 at the tth execution of Steps 2–5.

The LOG of the execution is the sequence $Y_t = X_{j_t}, t \geq 1$.

For $j \in M$ we let $COUNT(j)$ denote the number of times that $j_t = j$.

Theorem 15 follows from

Lemma 16

$$\mathbb{E}(COUNT(j)) \leq \frac{1}{d-1} \text{ for } j \in [m].$$

It follows that the expected length of LOG is at most $\frac{m}{d-1}$. Then $m \leq Nd$ because each $x \in X$ can be in at most d of the X_i's.

Proof of Lemma 16: Given an execution of MT of length at least t we define a rooted tree $TREE(t)$ with vertices from $X_1, X_2, \ldots, X_m$ as follows: Its root is Y_t. Now for $i = t-1, t-2, \ldots, 1$ we see if there exists k such that $i < k \leq t$ and $Y_i \cap Y_k \neq \emptyset$. If so, choose Y_k furthest from the root and make Y_i a child of Y_k. Otherwise do nothing.

We observe the following:

- P1 $s \neq t$ implies that $TREE(s) \neq TREE(t)$.
  Reason: Suppose not and $Y = Y_s = Y_t$. Then $TREE(t)$ will have at least one more appearance of Y (its root) than $TREE(s)$, contradiction.

- P2 If Y', Y'' are children of Y in $TREE(t)$ then $Y' \cap Y'' = \emptyset$.
  Reason: Otherwise, Y'' is a child of Y' or vice-versa.

- P3 Let $x \in X$ and let $Y_{i_1}, Y_{i_2}, \ldots, Y_{i_s}$ be those Y_i that contain x, ordered by their appearance in LOG. Then each value for x comes from a different independent sample.
  Reason: Because Y_{i_a} appears in LOG, we re-sample the random variables in $Y_{i_a} \cap Y_{i_{a+1}}$ before $Y_{i_{a+1}}$ appears.

Next let T be a rooted tree with vertices labelled by $X_1, X_2, \ldots, X_m$ such that if Y' is a child of Y then $Y \cap Y' \neq \emptyset$. Let $OCCUR(T)$ be the event that $TREE(t) = T$ for some t.

$$\mathbb{P}(OCCUR(T)) \leq p^{|T|}. \tag{17}$$

Let the vertices of $T = TREE(t)$ be $X_{i_1}, X_{i_2}, \ldots, X_{i_s}$ ordered by their appearence in LOG. Then

$$\mathbb{P}(OCCUR(T)) \leq \prod_{j=1}^{s} \mathbb{P}\left(\mathcal{E}_{i_j} \mid \mathcal{E}_{i_1}, \ldots \mathcal{E}_{i_{j-1}}\right) = \prod_{j=1}^{s} \mathbb{P}\left(\mathcal{E}_{i_j}\right) \leq p^{|T|}.$$

The appearance of X_{i_j} implies that $\mathcal{E}_{i_j}$ has occurred and P3 means that the values of $x \in X_{i_j}$ are fresh with respect to $\mathcal{E}_{i_1}, \ldots \mathcal{E}_{i_{j-1}}$.

Next let $\mathcal{T}$ denote an infinite tree with branching factor d rooted at a vertex ρ. Then, for each j,

$$\mathbb{E}(Count(j)) \leq y = \sum_{T} p^{|T|}$$

where T ranges over subtrees of $\mathcal{T}$ that are rooted at ρ.
This is because each ocurrence of Y_j corresponds to a distinct $T \subseteq \mathcal{T}$.
We complete the proof of Theorem 15 by proving

$$\xi \leq \frac{1}{d-1}. \tag{18}$$

Let $\mathcal{T}_s$ denote the set of sub-trees of $\mathcal{T}$ that are rooted at ρ. Let $y_s = \sum_{T \in \mathcal{T}_s} p^{|T|}$ so that $y = \lim_{s \to \infty} y_s$. Then $y_0 \leq p$ and

$$y_{s+1} \leq p \sum_{i=0}^{d} \binom{d}{i} y_s^i. \tag{19}$$

Note that (19) follows from the fact that if $T \in \mathcal{T}_{s+1}$ then it will contain ρ and $0 \leq i \leq d$ subtrees, each of which is isomorphic to a tree in $\mathcal{T}_s$. The factor y_s^i follows from (17).
All that remains to prove (18) is to prove that $y_s \leq \frac{1}{d-1}$ for $s \geq 0$. We do this by induction on s.

Now for the base case we have

$$y_0 \leq p \leq \frac{(d-1)^{d-1}}{d^d} \leq \frac{1}{d-1}.$$

For the inductive step we have

$$y_{s+1} \leq p \sum_{i=0}^{d} \binom{d}{i} \frac{1}{(d-1)^i} = p \left(1 + \frac{1}{d-1}\right)^d \leq$$

$$\frac{(d-1)^{d-1}}{d^d} \cdot \left(\frac{d}{d-1}\right)^d = \frac{1}{d-1}.$$

Note that y_s is monotone increasing and bounded and so has a limit.