Chernoff Bounds

Let $X_1, ..., X_n$ be independent 0-1 random variables with

$$Pr(X_i = 1) = p_i \qquad\qquad Pr(X_i = 0) = 1 - p_i.$$

Let $X = \sum_{i=1}^{n} X_i$,

$$\mu = E[X] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p_i$$

We want a bound on

$$Pr(|X - \mu| > \delta\mu).$$

$$Var[X] = npq$$

If we use Chebyshev's Inequality we get

$$Pr(|X - \mu| > \delta\mu) \leq \frac{npq}{\delta^2 n^2 p^2} = \frac{q}{\delta^2 \mu}$$

Chernoff bound will give

$$Pr(|X - \mu| > \delta\mu) \leq 2e^{-\mu\delta^2/3}.$$

The Basic Idea

Using Markov inequality we have:
For any $t > 0$,

$$Pr(X \geq a) = Pr(e^{tX} \geq e^{ta}) \leq \frac{E[e^{tX}]}{e^{ta}}.$$

Similarly, for any $t < 0$

$$Pr(X \leq a) = Pr(e^{tX} \geq e^{ta}) \leq \frac{E[e^{tX}]}{e^{ta}}.$$

$$Pr(X \geq a) \leq \min_{t>0} \frac{E[e^{tX}]}{e^{ta}}.$$

$$Pr(X \leq a) \leq \min_{t<0} \frac{E[e^{tX}]}{e^{ta}}.$$

Moment Generating Function

Definition

The moment generating function of a random variable X is defined for any real value t as

$$M_X(t) = E[e^{tX}].$$

Theorem

Let X be a random variable with moment generating function $M_X(t)$. Assuming that exchanging the expectation and differentiation operands is legitimate, then for all $n \geq 1$

$$E[x^n] = M_X^{(n)}(0),$$

where $M_X^{(n)}(0)$ is the n-th derivative of $M_X(t)$ evaluated at $t = 0$.

Proof.

$$M_X^{(n)}(t) = E[X^n e^{tX}].$$

Computed at $t = 0$ we get

$$M_X^{(n)}(0) = E[X^n].$$

□

Theorem

Let X and Y be two random variables. If

$$M_X(t) = M_Y(t)$$

for all $t \in (-\delta, \delta)$ for some $\delta > 0$, then X and Y have the same distribution.

Theorem

If X and Y are independent random variables then

$$M_{X+Y}(t) = M_X(t)M_Y(t).$$

Proof.

$$M_{X+Y}(t) = E[e^{t(X+Y)}] = E[e^{tx}]E[e^{tY}] = M_X(t)M_Y(t).$$

$\square$

Chernoff Bound for Sum of Bernoulli Trials

Let $X_1, \ldots, X_n$ be a sequence of independent Bernoulli trials with $Pr(X_i = 1) = p_i$. Let $X = \sum_{i=1}^{n} X_i$, and let

$$\mu = E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} p_i.$$

$$\begin{aligned} M_{X_i}(t) &= E[e^{tX_i}] \\ &= p_i e^t + (1 - p_i) \\ &= 1 + p_i(e^t - 1) \\ &\leq e^{p_i(e^t - 1)}. \end{aligned}$$

Taking the product of the n generating functions we get

$$\begin{aligned} M_X(t) &= \prod_{i=1}^{n} M_{X_i}(t) \\ &\le \prod_{i=1}^{n} e^{p_i(e^t-1)} \\ &= e^{\sum_{i=1}^{n} p_i(e^t-1)} \\ &= e^{(e^t-1)\mu} \end{aligned}$$

Theorem

Let $X_1, \ldots, X_n$ be independent Bernoulli random variables such that $Pr(X_i = 1) = p_i$.

1. *For any $\delta > 0$,*

$$Pr(X \geq (1+\delta)\mu) < \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}. \tag{1}$$

2. *For $0 < \delta < 1$,*

$$Pr(X \geq (1+\delta)\mu) \leq e^{-\mu\delta^2/3}. \tag{2}$$

3. *For $R \geq 6\mu$,*

$$Pr(X \geq R) \leq 2^{-R}. \tag{3}$$

Applying Markov's inequality we have for any $t > 0$

$$\begin{aligned} Pr(X \geq (1+\delta)\mu) &= Pr(e^{tX} \geq e^{t(1+\delta)\mu}) \\ &\leq \frac{E[e^{tX}]}{e^{t(1+\delta)\mu}} \\ &< \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu}} \end{aligned} \tag{4}$$

For any $\delta > 0$, we can set $t = \ln(1+\delta) > 0$ to get:

$$Pr(X \geq (1+\delta)\mu) \leq \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu.$$

We show that for $0 < \delta < 1$,

$$\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}} \leq e^{-\delta^2/3}$$

or that

$$f(\delta) = \delta - (1+\delta)\ln(1+\delta) + \delta^2/3 \leq 0$$

in that interval. Computing the derivatives of $f(\delta)$ we get

$$\begin{aligned} f'(\delta) &= 1 - \frac{1+\delta}{1+\delta} - \ln(1+\delta) + \frac{2}{3}\delta && (5)\\ &= -\ln(1+\delta) + \frac{2}{3}\delta, && (6)\\ f''(\delta) &= -\frac{1}{1+\delta} + \frac{2}{3}. && (7) \end{aligned}$$

$f''(\delta) < 0$ for $0 \leq \delta < 1/2$, and $f''(\delta) > 0$ for $\delta > 1/2$.
$f'(\delta)$ first decreases and then increases over the interval $[0, 1]$.
Since $f'(0) = 0$ and $f'(1) < 0$, $f'(\delta) \leq 0$ in the interval $[0, 1]$.
Since $f(0) = 0$, we have that $f(\delta) \leq 0$ in that interval.

For $R \geq 6\mu$, $\delta \geq 5$.

$$\begin{aligned} Pr(X \geq (1+\delta)\mu) &\leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu} \\ &\leq \left(\frac{e}{6}\right)^{R} \\ &\leq 2^{-R}. \end{aligned}$$

Theorem

*Let $X_1, \ldots, X_n$ be independent Bernoulli random variables such that $Pr(X_i = 1) = p_i$. Let $X = \sum_{i=1}^{n} X_i$ and $\mu = E[X]$.
For $0 < \delta < 1$,*

$$Pr(X \leq (1 - \delta)\mu) \leq e^{-\mu\delta^2/2} \tag{8}$$

Using Markov's inequality, for any $t < 0$,

$$\begin{aligned} Pr(X \le (1-\delta)\mu) &= Pr(e^{tX} \ge e^{(1-\delta)t\mu}) \\ &\le \frac{E[e^{tX}]}{e^{t(1-\delta)\mu}} \\ &\le \frac{e^{(e^t-1)\mu}}{e^{t(1-\delta)\mu}} \end{aligned}$$

For $0 < \delta < 1$, we set $t = \ln(1-\delta) < 0$ to get:

$$Pr(X \le (1-\delta)\mu) \le \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu}$$

We need to show:

$$f(\delta) = -\delta - (1-\delta)\ln(1-\delta) + \frac{1}{2}\delta^2 \le 0. \tag{9}$$

We need to show:

$$f(\delta) = -\delta - (1-\delta)\ln(1-\delta) + \frac{1}{2}\delta^2 \leq 0. \tag{10}$$

Differentiating $f(\delta)$ we get

$$\begin{aligned} f'(\delta) &= \ln(1-\delta) + \delta, \\ f''(\delta) &= -\frac{1}{1-\delta} + 1. \end{aligned}$$

$f(0) = 0$, and since $f'(\delta) \leq 0$ in the range $[0,1)$, $f(\delta)$ is monotonically decreasing in that interval.

Example: Coin flips

Let X be the number of heads in a sequence of n independent fair coin flips.

$$Pr\left(|X - \frac{n}{2}| \geq \frac{1}{2}\sqrt{4n \ln n}\right)$$

$$= Pr\left(X \geq \frac{n}{2}\left(1 + \sqrt{\frac{4 \ln n}{n}}\right)\right)$$

$$+ Pr\left(X \leq \frac{n}{2}\left(1 - \sqrt{\frac{4 \ln n}{n}}\right)\right)$$

$$\leq e^{-\frac{1}{3}\frac{n}{2}\frac{4 \ln n}{n}} + e^{-\frac{1}{2}\frac{n}{2}\frac{4 \ln n}{n}} \leq \frac{2}{n}.$$

Using the Chebyshev's bound we had:

$$Pr\left(|X - \frac{n}{2}| \geq \frac{n}{4}\right) \leq \frac{4}{n}.$$

Using the Chernoff bound in this case, we obtain

$$\begin{aligned} Pr\left(|X - \frac{n}{2}| \geq \frac{n}{4}\right) &= Pr\left(X \geq \frac{n}{2}\left(1 + \frac{1}{2}\right)\right) \\ &+ Pr\left(X \leq \frac{n}{2}\left(1 - \frac{1}{2}\right)\right) \\ &\leq e^{-\frac{1}{3}\frac{n}{2}\frac{1}{4}} + e^{-\frac{1}{2}\frac{n}{2}\frac{1}{4}} \\ &\leq 2e^{-\frac{n}{24}}. \end{aligned}$$

Example: Estimating a Parameter

- Evaluating the probability that a particular gene mutation occurs in the population.
- Given a DNA sample, a lab test can determine if it carries the mutation.
- The test is expensive and we would like to obtain a relatively reliable estimate from a minimum number of samples.
- $p =$ the unknown value;
- $n =$ number of samples, $\tilde{p}n$ had the mutation.
- Given sufficient number of samples we expect the value p to be in the neighborhood of sampled value $\tilde{p}$, but we cannot predict any single value with high confidence.

Confidence Interval

Instead of predicting a single value for the parameter we give an interval that is likely to contain the parameter.

Definition

A $1-q$ **confidence interval** *for a parameter* T *is an interval* $[\tilde{p}-\delta, \tilde{p}+\delta]$ *such that*

$$Pr(T \in [\tilde{p}-\delta, \tilde{p}+\delta]) \geq 1-q.$$

We want to minimize 2δ and q, with minimum n.
Using $\tilde{p}n$ as our estimate for pn, we need to compute δ and q such that

$$Pr(p \in [\tilde{p}-\delta, \tilde{p}+\delta]) = Pr(np \in [n(\tilde{p}-\delta), n(\tilde{p}+\delta)]) \geq 1-q.$$

- The random variable here is the interval $[\tilde{p} - \delta, \tilde{p} + \delta]$ (or the value $\tilde{p}$), while p is a fixed (unknown) value.
- $n\tilde{p}$ has a binomial distribution with parameters n and p, and $E[\tilde{p}] = p$. If $p \notin [\tilde{p} - \delta, \tilde{p} + \delta]$ then we have one of the following two events:
  1. If $p < \tilde{p} - \delta$, then $n\tilde{p} \geq n(p + \delta) = np(1 + \frac{\delta}{p})$, or $n\tilde{p}$ is larger than its expectation by a $\frac{\delta}{p}$ factor.
  2. If $p > \tilde{p} + \delta$, then $n\tilde{p} \leq n(p - \delta) = np(1 - \frac{\delta}{p})$, and $n\tilde{p}$ is smaller than its expectation by a $\frac{\delta}{p}$ factor.

$$
\begin{aligned}
& \Pr(p \notin [\tilde{p} - \delta, \tilde{p} + \delta]) \\
= \ & \Pr(n\tilde{p} \leq np(1 - \frac{\delta}{p})) + \Pr(n\tilde{p} \geq np(1 + \frac{\delta}{p})) \\
\leq \ & e^{-\frac{1}{2} np (\frac{\delta}{p})^2} + e^{-\frac{1}{3} np (\frac{\delta}{p})^2} \\
= \ & e^{-\frac{n\delta^2}{2p}} + e^{-\frac{n\delta^2}{3p}}.
\end{aligned}
$$

But the value of p is unknown, A simple solution is to use the fact that $p \leq 1$ to prove

$$
\Pr(p \notin [\tilde{p} - \delta, \tilde{p} + \delta]) = e^{-\frac{n\delta^2}{2}} + e^{-\frac{n\delta^2}{3}}.
$$

Setting $q = e^{-\frac{n\delta^2}{2}} + e^{-\frac{n\delta^2}{3}}$, we obtain a tradeoff between δ, n and the error probability q.

Better Bound

The binomial probabilities are monotone increasing up to the expectation, and then monotone decreasing.

$$
\begin{aligned}
& \Pr(p \notin [\tilde{p} - \delta, \tilde{p} + \delta]) \\
\leq \ & \Pr(n\tilde{p} \leq np(1 - \frac{\delta}{p})) + \Pr(n\tilde{p} \geq np(1 + \frac{\delta}{p})) \\
\leq \ & \max_{p \leq \tilde{p} - \delta} e^{-np(\frac{\tilde{p}-p}{p})^2/2} + \max_{p \geq \tilde{p} + \delta} e^{-np(\frac{p-\tilde{p}}{p})^2/3} \\
\leq \ & e^{-\frac{n\delta^2}{2(\tilde{p}-\delta)}} + e^{-\frac{n\delta^2}{3(\tilde{p}+\delta)}},
\end{aligned}
$$

Setting

$$q = e^{-\frac{n\delta^2}{2(\tilde{p}-\delta)}} + e^{-\frac{n\delta^2}{3(\tilde{p}+\delta)}}$$

gives a tighter tradeoff between δ, n and q.

Application: Set Balancing

Given an $n \times n$ matrix $\mathcal{A}$ with entries in $\{0, 1\}$, let

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ \dots \\ b_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \dots \\ \dots \\ c_n \end{pmatrix}.$$

Find a vector $\bar{b}$ with entries in $\{-1, 1\}$ that minimizes

$$||\mathcal{A}\bar{b}||_\infty = \max_{i=1,\dots,n} |c_i|.$$

Theorem

For a random vector $\bar{b}$*, with entries chosen independently and with equal probability from the set* $\{-1, 1\}$*,*

$$Pr(||\mathcal{A}\bar{b}||_\infty \geq \sqrt{12n \ln n}) \leq \frac{4}{n}.$$

- Consider the i-th row $\bar{a}_i = a_{i,1}, \ldots, a_{i,n}$. Let k be the number of 1's in that row.
- If $k \le \sqrt{12n \ln n}$ clearly $|\bar{a}_i \cdot \bar{b}| \le \sqrt{12n \ln n}$.
- If $k > \sqrt{12n \ln n}$, let

$$X_i = |\{j \mid a_{i,j} = 1 \text{ and } b_j = 1\}|$$

and

$$Y_i = |\{j \mid a_{i,j} = 1 \text{ and } b_j = -1\}|.$$

- Thus, X_i counts the number of $+1$'s in the sum $\sum_{j=1}^{n} a_{i,j} b_j$,
- Y_i counts the number of -1's
- $X_i + Y_i = k$.

if $|X_i - Y_i| \leq \sqrt{12n \log n}$ then $|X_i - (k - X_i)| \leq \sqrt{12n \log n}$ which implies

$$\frac{k}{2}(1 - \frac{\sqrt{12n \log n}}{k}) \leq X_i \leq \frac{k}{2}(1 + \frac{\sqrt{12n \log n}}{k}).$$

Using Chernoff bounds,

$$Pr\left(X_i \geq \frac{k}{2}\left(1+\sqrt{\frac{12n\ln n}{k^2}}\right)\right) \leq e^{-(\frac{k}{2})(\frac{1}{3})(\frac{12n\ln n}{k^2})} \leq e^{-2\ln n}$$

$$Pr\left(X_i \leq \frac{k}{2}\left(1-\sqrt{\frac{12n\ln n}{k^2}}\right)\right) \leq e^{-(\frac{k}{2})(\frac{1}{2})(\frac{12n\ln n}{k^2})} \leq e^{-3\ln n}$$

Hence, for a given row,

$$Pr(|X_i - Y_i| \geq \sqrt{12n\ln n}) \leq \frac{2}{n^2}$$

Since there are n rows, the probability that any row exceeds that bound is bounded by $\frac{2}{n}$.

Chernoff Bound for Sum of $\{-1, +1\}$ Random Variables

Theorem

Let $X_1, ..., X_n$ be independent random variables with

$$Pr(X_i = 1) = Pr(X_i = -1) = \frac{1}{2}.$$

Let $X = \sum_1^n X_i$. For any $a > 0$,

$$Pr(X \geq a) \leq e^{-a^2/2n}$$

For any $t > 0$,

$$E[e^{tX_i}] = \frac{1}{2}e^t + \frac{1}{2}e^{-t}.$$

$$e^t = 1 + t + \frac{t^2}{2!} + \cdots + \frac{t^i}{i!} + \ldots$$

and

$$e^{-t} = 1 - t + \frac{t^2}{2!} + \cdots + (-1)^i \frac{t^i}{i!} + \ldots$$

Thus,

$$\begin{aligned} E[e^{tX_i}] &= \frac{1}{2}e^t + \frac{1}{2}e^{-t} = \sum_{i \geq 0} \frac{t^{2i}}{(2i)!} \\ &\leq \sum_{i \geq 0} \frac{\left(\frac{t^2}{2}\right)^i}{i!} = e^{t^2/2} \end{aligned}$$

$$E[e^{tX}] = \prod_{i=1}^{n} E[e^{tX_i}] \leq e^{nt^2/2},$$

$$Pr(X \geq a) = Pr(e^{tX} > e^{ta}) \leq \frac{E[e^{tX}]}{e^{ta}} \leq e^{t^2 n/2 - ta}.$$

Setting $t = a/n$ yields

$$Pr(X \geq a) \leq e^{-a^2/2n}.$$

By symmetry we also have

Corollary

Let $X_1, ..., X_n$ be independent random variables with

$$Pr(X_i = 1) = Pr(X_i = -1) = \frac{1}{2}.$$

Let $X = \sum_{i=1}^{n} X_i$. Then for any $a > 0$,

$$Pr(|X| > a) \leq 2e^{-a^2/2n}.$$

Application: Set Balancing Revisited

Theorem

For a random vector $\bar{b}$, with entries chosen independently and with equal probability from the set $\{-1, 1\}$,

$$Pr(||\mathcal{A}\bar{b}||_\infty \geq \sqrt{4n \ln n}) \leq \frac{2}{n} \tag{11}$$

- Consider the i-th row $\bar{a}_i = a_{i,1}, ..., a_{i,n}$.
- Let k be the number of 1's in that row.
- $Z_i = \sum_{j=1}^{k} a_{i,i_j} b_{i_j} = \sum_{j=1}^{k} b_{i_j}$.
- If $k \leq \sqrt{4n \ln n}$ then clearly Z_i satisfies the bound.

If $k > \sqrt{4n \log n}$, the k non-zero terms in the sum Z_i are independent random variables, each with probability $1/2$ of being either $+1$ or -1.
Using the Chernoff bound:

$$Pr\left\{|Z_i| > \sqrt{4n \log n}\right\} \leq 2e^{-4n \log n/2k} \leq \frac{2}{n^2},$$

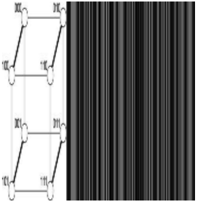
where we use the fact that $n \geq k$.

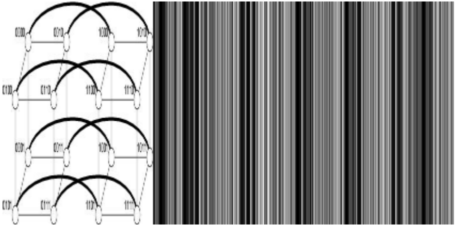
Packet Routing on Parallel Computer

Communication network:

- Nodes - processors, switching nodes.
- edges - communication links.

The n-cube:
$N = 2^n$ nodes.
Let $\bar{x} = (x_1, ..., x_n)$ be the number of node x in binary.
Nodes x and y are connected by an edge iff their binary representations differ in exactly one bit.
Bit-wise routing: correct bit i in the i-th transition - route has length $\leq n$.

The 3-cube: 

The 4-cube: 

A permutation communication request: each node is the source and destination of exactly one packet.
Up to one packet can cross an edge per step, each packet can cross up to one edge per step.
What is the time to route an arbitrary permutation on the n-cube?

Two phase routing algorithm:

1. Send packet to a randomly chosen destination.
2. Send packet from random place to real destination.

Path: Correct the bits, starting at x_0 to x_{n-1}.
Any greedy queuing method - if some packet can traverse an edge one does.

Theorem

The two phase routing algorithm routes an arbitrary permutation on the n-cube in $O(\log N) = O(n)$ parallel steps with high probability.

- We focus first on phase 1. We bound the routing time of a given packet M.
- Let $e_1, \ldots, e_m$ be the $m \leq n$ edges traversed by a given packet M is phase 1.
- Let $X(e)$ be the total number of packets that traverse edge e at that phase.
- Let $T(M)$ be the number of steps till M finished phase 1.

Lemma

$$T(M) \leq \sum_{i=1}^{m} X(e_i).$$

- We call any path $P = (e_1, e_2, \ldots, e_m)$ of $m \leq n$ edges that follows the bit fixing algorithm a *possible packet path*.
- We denote the corresponding nodes $v_0, v_1, \ldots, v_m$, with $e_i = (v_{i-1}, v_i)$.
- For any possible packet path P, let $T(P) = \sum_{i=1}^{m} X(e_i)$.

- If phase I takes more than T steps then for some possible packet path P,

$$T(P) \geq T$$

- There are at most $2^n \cdot 2^n = 2^{2n}$ possible packet paths.
- Assume that e_k connects $(a_1, ..., a_i, ..., a_n)$ to $(a_1, .., \bar{a}_i, ..., a_n)$.
- Only packets that started in address

$$(*, ..., *, a_i, ..., a_n)$$

can traverse edge e_k, and only if their destination addresses are

$$(a_1, ..., a_{i-1}, \bar{a}_i, *, ..., *)$$

.

- There are 2^{i-1} possible packets, each has probability 2^{-i} to traverse e_i.

- There are 2^{i-1} possible packets, each has probability 2^{-i} to traverse e_i.
- $$E[X(e_k)] \leq 2^{i-1} \cdot 2^{-i} = \frac{1}{2}.$$
- $$E[T(P)] \leq \sum_{i=1}^{m} E[X(e_i)] \leq \frac{1}{2} \cdot m \leq n.$$
- **Problem:** The $X(e_i)$'s are not independent.

- A packet is active with respect to possible packet path P if it ever use an edge of P.
- For $k = 1, \ldots, N$, let $H_k = 1$ if the packet starting at node k is active, and $H_k = 0$ otherwise.
- The H_k are independent, since each H_k depends only on the choice of the intermediate destination of the packet starting at node k, and these choices are independent for all packets.
- Let $H = \sum_{k=1}^{N} H_k$ be the total number of active packets.
- $$E[H] \le E[T(P)] \le n$$
- Since H is the sum of independent $0-1$ random variables we can apply the Chernoff bound

$$\Pr(H \ge 6n \ge 6E[H]) \le 2^{-6n}.$$

For a given possible packet path P,

$$\begin{aligned} & \Pr(T(P) \geq 36n) \\ \leq \; & \Pr(H \geq 6n) + \Pr(T(P) \geq 36n \mid H < 6n) \\ \leq \; & 2^{-6n} + \Pr(T(P) \geq 36n \mid H < 6n). \end{aligned}$$

Lemma

If a packet leaves a path (of another packet) it cannot return to that path in the same phase.

Proof.

Leaving a path at the i-th transition implies different i-th bit, this bit cannot be changed again in that phase. □

Lemma

The number of transitions that a packet takes on a given path is distributed $G(\frac{1}{2})$.

Proof.

The packet has probability $1/2$ of leaving the path in each transition. □

The probability that the active packets cross edges of P more than $36n$ times is less than the probability that a fair coin flipped $36n$ times comes up heads less than $6n$ times.
Letting Z be the number of heads in $36n$ fair coin flips, we now apply the Chernoff bound:

$$
\begin{aligned}
& \Pr(T(P) \geq 36n \mid H \leq 6n) \leq \Pr(Z \leq 6n) \\
\leq \quad & e^{-18n(2/3)^2/2} = e^{-4n} \leq 2^{-3n-1}.
\end{aligned}
$$

$$
\begin{aligned}
\Pr(T(P) \geq 36n) \quad & \leq \quad \Pr(H \geq 6n) \\
& + \quad \Pr(T(P) \geq 36n \mid H \leq 6n) \\
& \leq \quad 2^{-6n} + 2^{-3n-1} \leq 2^{-3n}
\end{aligned}
$$

As there are at most 2^{2n} possible packet paths in the hypercube, the probability that there is *any* possible packet path for which $T(P) \geq 36n$ is bounded by

$$2^{2n}2^{-3n} = 2^{-n} = O(N^{-1}).$$

- The proof of phase 2 is by symmetry:
- The proof of phase 1 argued about the number of packets crossing a given path, no "timing" considerations.
- The path from "one packet per node" to random locations is similar to random locations to "one packet per node" in reverse order.
- Thus, the distribution of the number of packets that crosses a path of a given packet is the same.

Oblivious Routing

Definition

A routing algorithm is **oblivious** if the path taken by one packet is independent of the source and destinations of any other packets in the system.

Theorem

Given an N-node network with maximum degree d the routing time of any deterministic oblivious routing scheme is

$$\Omega(\sqrt{\frac{N}{d^3}}).$$