Shortest Paths

Let the arcs of the complete digraph D_n on $[n]$ be given independent lengths X_e, $e \in [n]^2$.

Here X_e is exponential with mean 1

i.e.

$$\Pr(X_e \geq t) = e^{-t}$$

for all $t \geq 0$.

Theorem

Let X_{ij} = distance from i to j. Then

(i) For any fixed i, j,

$$\Pr\left(\left|\frac{X_{i,j}}{\log n / n} - 1\right| \geq \epsilon\right) \to 0, \quad \forall \epsilon > 0$$

(ii) For any fixed i,

$$\Pr\left(\left|\frac{Z_i}{2\log n / n} - 1\right| \geq \epsilon\right) \to 0, \quad \forall \epsilon > 0$$

Here $Z_i = \max_j X_{ij}$.

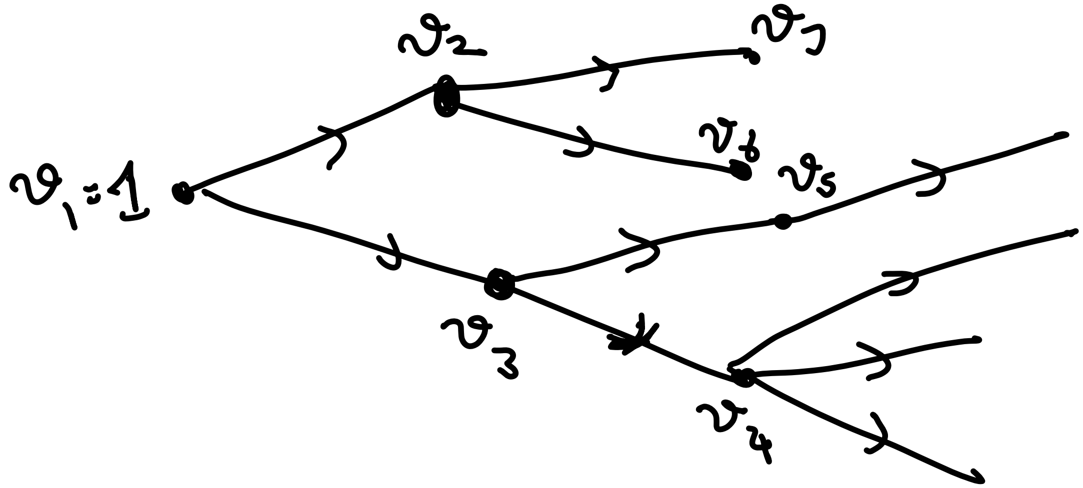
Proof

Two main properties of exponential X:

(P1) $\Pr(X > \alpha + \beta \mid X > \alpha) = \Pr(X > \beta)$.

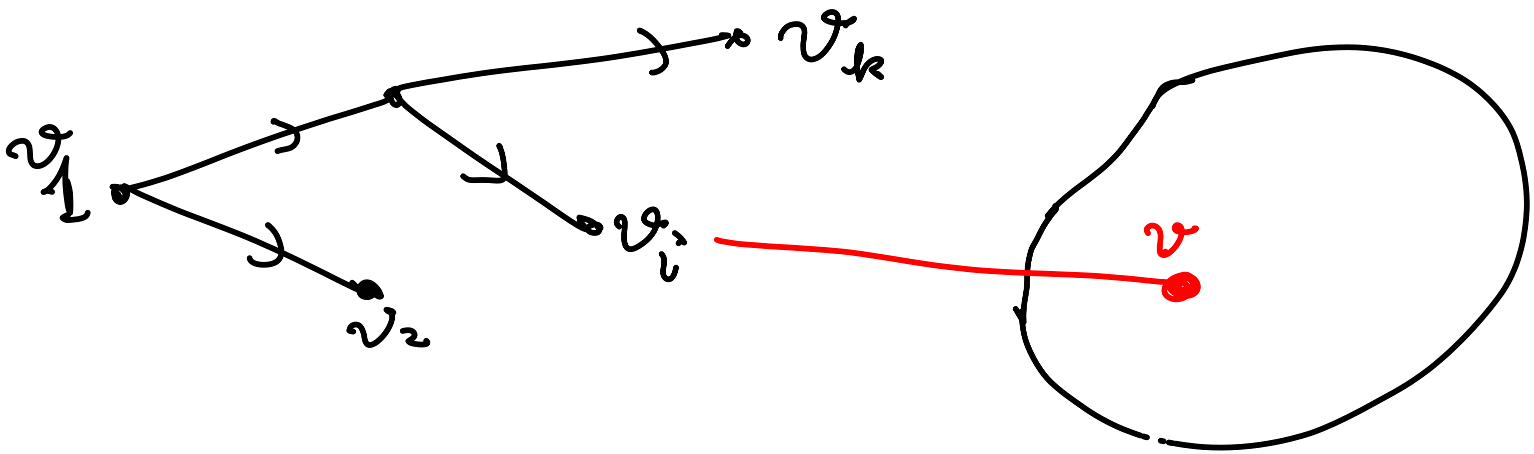
(P2) If $X_1, X_2, \dots, X_m$ are independent exponentials then $\min\{X_1, X_2, \dots, X_m\}$ is an exponential with mean $1/m$.

Fix $i = 1$ and consider Dijkstra's shortest path algorithm. This produces a tree



Suppose that vertices are added to the tree in the order $v_1, v_2, \dots, v_n$ and that $\mathrm{dist}(v_1, v_i) = Y_j$.

It follows from P1 (p3) that



$$Y_{k+1} = \min_{\substack{i=1,\dots,k \\ v}} \Big[\underbrace{Y_i + X_{(v_i, v)}}_{\geq Y_k} \Big]$$

$$\overset{d}{=} Y_k + \text{Exponential}$$

So $Y_{k+1} = Y_k + E_k$ where E_k is exponential with mean $\frac{1}{k(n-k)}$ and is independent of Y_k.

So

$$E(Y_n) = \sum_{k=1}^{n-1} \frac{1}{k(n-k)}$$

$$= \frac{1}{n} \sum_{k=1}^{n-1} \left(\frac{1}{k} + \frac{1}{n-k} \right)$$

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

$$\sim \frac{2 \log_e n}{n}.$$

Also

$$Var(Y_n) = \sum_{k=1}^{n-1} Var(E_k) = \sum_{k=1}^{n-1} \left(\frac{1}{k(n-k)} \right)^2$$

$$\leq 2 \sum_{k=1}^{n/2} \left(\frac{1}{k(n-k)} \right)^2 \leq \frac{8}{n^2} \sum_{k=1}^{n/2} \frac{1}{k^2} = O(n^{-2})$$

and we can use Chebyshev to prove (ii).

Now fix $j=2$. Then if i is defined by $\sigma_i = 2$, we see that i is uniform over $\{2, 3, \ldots, n\}$.

So.

$$E(X_{1,2}) = \frac{1}{n-1} \sum_{i=2}^{n} \sum_{k=1}^{i} \frac{1}{k(n-k)}$$

$$= \frac{1}{n-1} \sum_{k=1}^{n-1} \frac{n-k}{k(n-k)}$$

$$\approx \frac{\log_e n}{n}.$$

For variance we have

$$X_{1,2} = \delta_2 Y_2 + \delta_3 Y_3 + \cdots + \delta_n Y_n$$

where

$\delta_i \in \{0,1\}$; $\delta_2 + \cdots + \delta_n = 1$; $Pr(\delta_i = 1) = \frac{1}{n-1}$.

$$Var(X_{1,2}) = \sum_{i=2}^{n} Var(\delta_i Y_i)$$

$$+ \sum_{i \neq j} Covar(\delta_i Y_i \, \delta_j Y_j)$$

$$\leq \sum_{i=2}^{n} Var(\delta_i Y_i)$$

$$Covar(\delta_i Y_i \, \delta_j Y_j) = E(\underbrace{\delta_i Y_i \, \delta_j Y_j}_{0}) - E(\delta_i Y_i) E(\delta_j Y_j) \leq 0$$

$$\mathrm{Var}(X_{1,2}) \leq \sum_{i=2}^{n} \mathrm{Var}(\delta_i Y_i)$$

$$\leq \sum_{i=2}^{n} \frac{1}{n-1} \sum_{k=1}^{i-1} \left(\frac{1}{k(n-k)} \right)^2$$

$$= O\left(1/n^2 \right).$$

Split sum at $n/2$

We can now use Chebyshev.