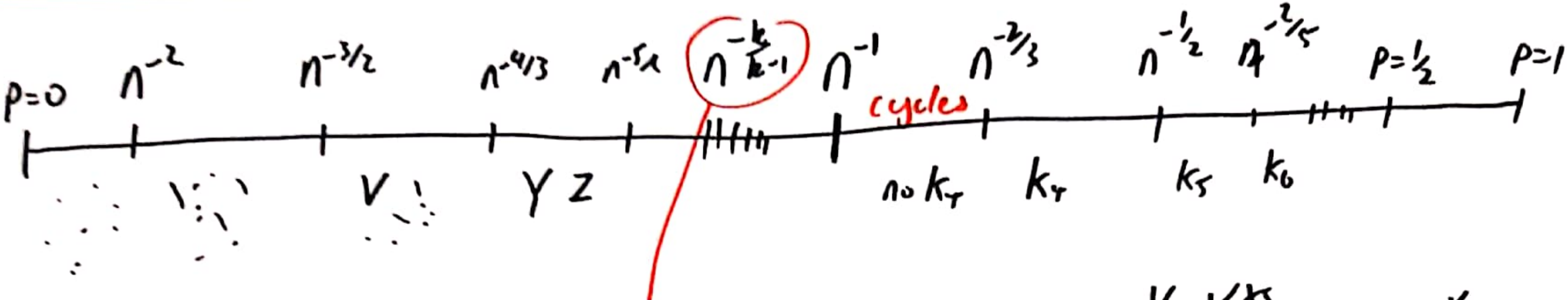


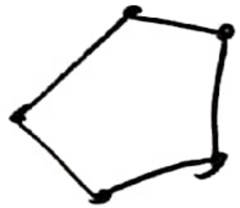
Trees size k appear

v vts

e edges

$n^{-\frac{v}{e}}$

$H =$ 

 n^{-1}

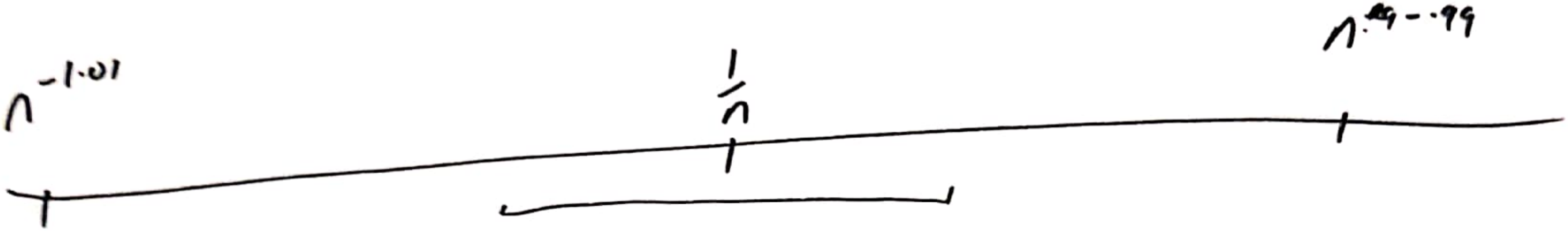
$e(G(n,p)) \sim Bi(\binom{n}{2}, p)$

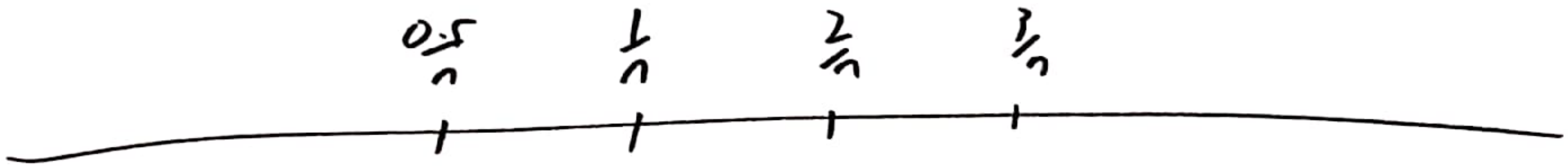
$\frac{n^2 p}{2}$

3 vx

2 edges

  $n^{-4/3}$





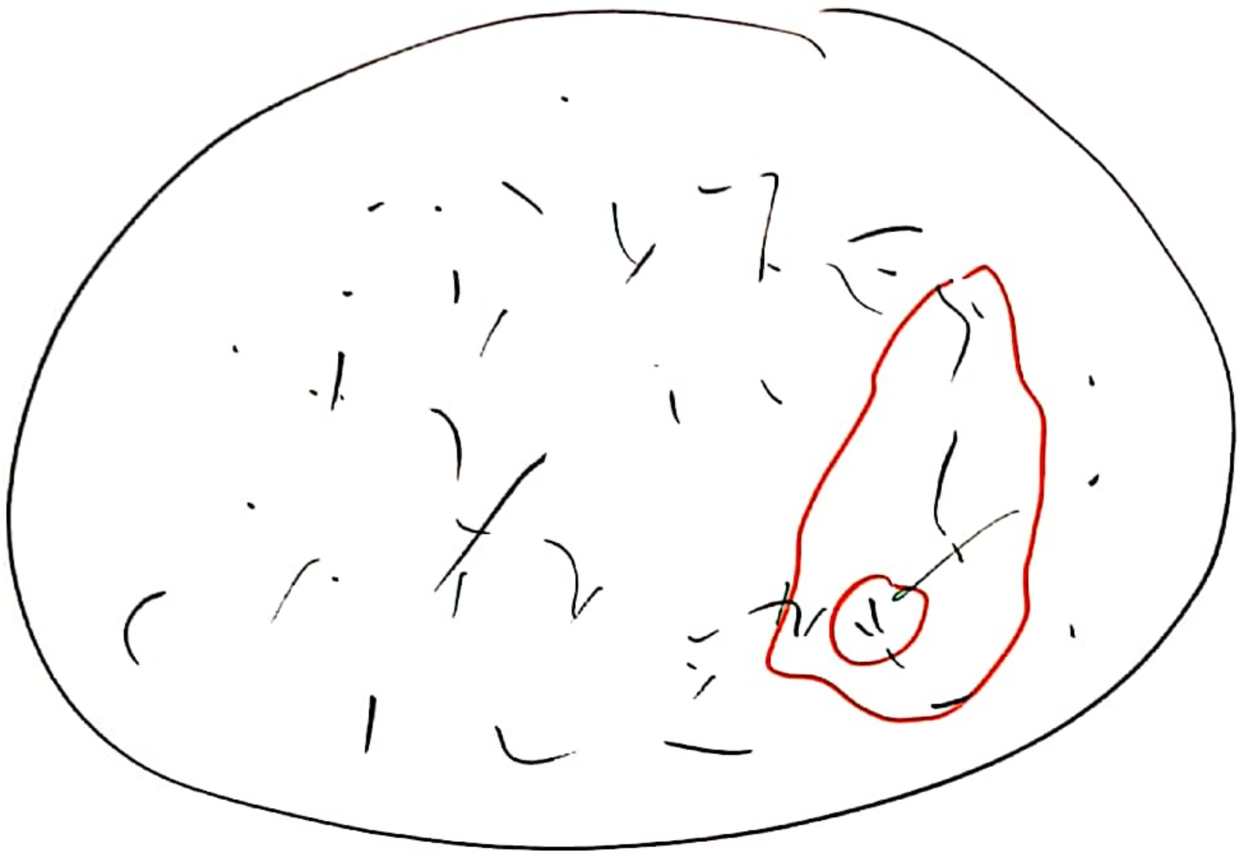
Erdős - Rényi 1960

If $p = \frac{c}{n}$

for $c<1$ typically all components are small.

largest $\sim A_c \log n$

for $c>1$ $\exists!$ giant component size $\sim g_c n$



$p = \frac{0.9}{n}$ Hedges

$\approx 0.45n$

$p = \frac{1.1}{n}$ $0.55n$

Z = a distribution on $\mathbb{Z}_{\geq 0}$ (dist. of a counting RV)

gen 0 $X_t = 1$

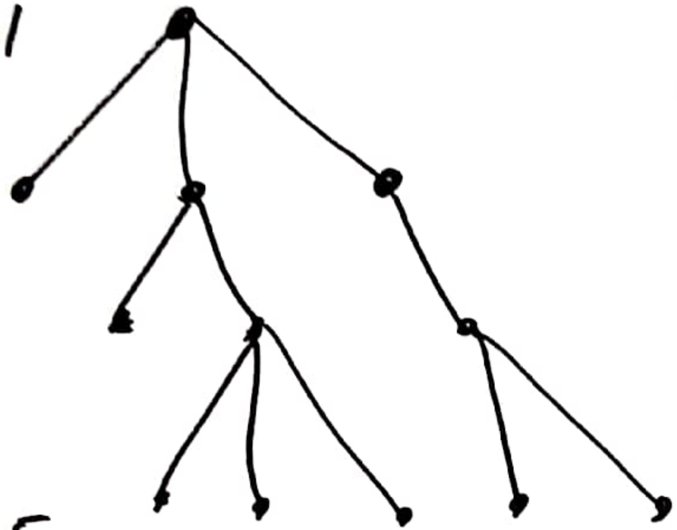
1 $X_1 = 3$

2 $X_2 = 3$

3 $X_3 = 5$

4 $X_4 = 0$

$\vdots$ $X_5 = 0$



$\frac{1}{8}$

$\times \frac{1}{4} \times \frac{1}{8} \times \frac{1}{2}$

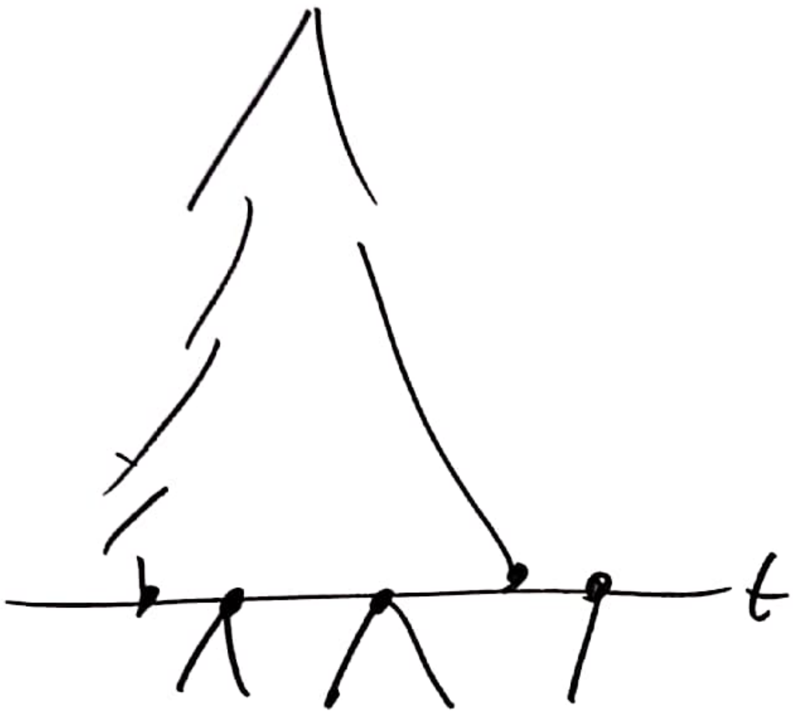
$p_0 = \mathbb{P}(Z=0) = \frac{1}{4}$

$p_1 = \mathbb{P}(Z=1) = \frac{1}{2}$

$\mathbb{P}(Z=2) = \frac{1}{8}$

$\mathbb{P}(Z=3) = \frac{1}{8}$

$\underset{\sim}{X} = (X_t)_{t \geq 0}$ is called the Galton-Watson branching process with offspring distribution Z.



Distribution of X_{t+1}

given $X_0 = 1, X_1 = k_1, X_2 = k_2 \dots$

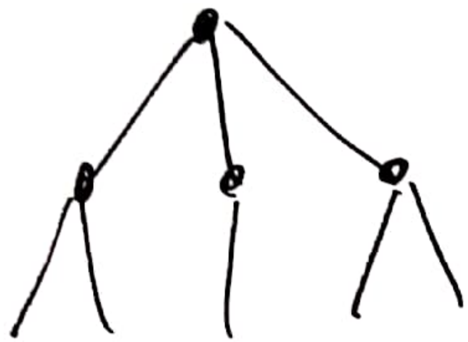
$X_t = k_t$

is sum of k_t indep. copies of Z.

We say the BP goes extint / dies out

if $\exists t$ st. $X_t = 0$ ($\Leftrightarrow$ tree finite)

Survives if not



Francis Galton 1873

Henry Watson

Let $\lambda = \mathbb{E}[Z]$ $\bullet$

Given that $X_t = k$

$$\mathbb{E}[X_{t+1} \mid X_t = k] = k\,\mathbb{E}[Z] = k\lambda$$

$$\mathbb{E}[X_{t+1}] = \sum_k \mathbb{P}(X_t = k)\,\mathbb{E}[\qquad] = \sum_k k\lambda\,\mathbb{P}(X_t = k)$$

$$= \lambda\,\mathbb{E}[X_t]$$

$\mathbb{E}[X_0] = 1 \qquad \therefore \quad \mathbb{E}[X_t] = \lambda^t$

Suppose $\lambda < 1$

$$\mathbb{P}(\underset{\sim}{X} \text{ survives}) \le \mathbb{P}(X_t > 0) = \mathbb{P}(X_t \geq 1) \le \mathbb{E}[X_t] = \lambda^t$$

Applies $\forall t$, since $\lambda \to 1 \Rightarrow \mathbb{P}(\underset{\sim}{X} \text{ survives}) = 0$

{If $\mathbb{P}(Z=0)>0$ then for $\lambda>1$ possible Z dies out

If Z is a counting RV its pgf $f_Z : [0,1] \to \mathbb{R}$

is the function defined by $f_Z(x) = \mathbb{E}[x^Z]$

$(\mathbb{E}Z<\infty)$

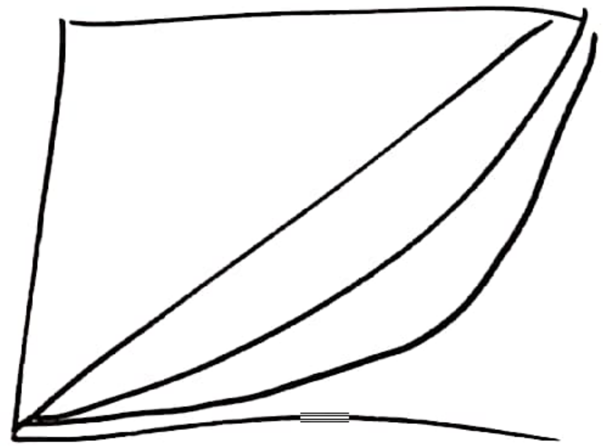
$$= \sum_k x^k \, \mathbb{P}(Z=k)$$

$f_Z(0) = \mathbb{P}(Z=0) \geq 0$

$f_Z(1) = 1$

$f_Z(x)$ is continuous, increasing

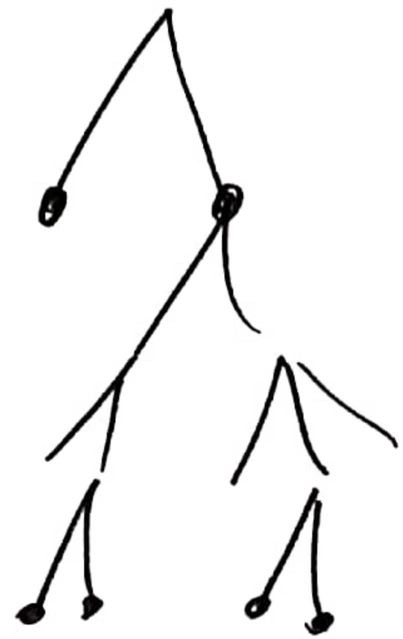
$f_Z'(1) = \mathbb{E}(Z)$



If $\mathbb{P}(Z\geq 2)>0$, $f_Z''(x)>0$ on $(0,1)$.

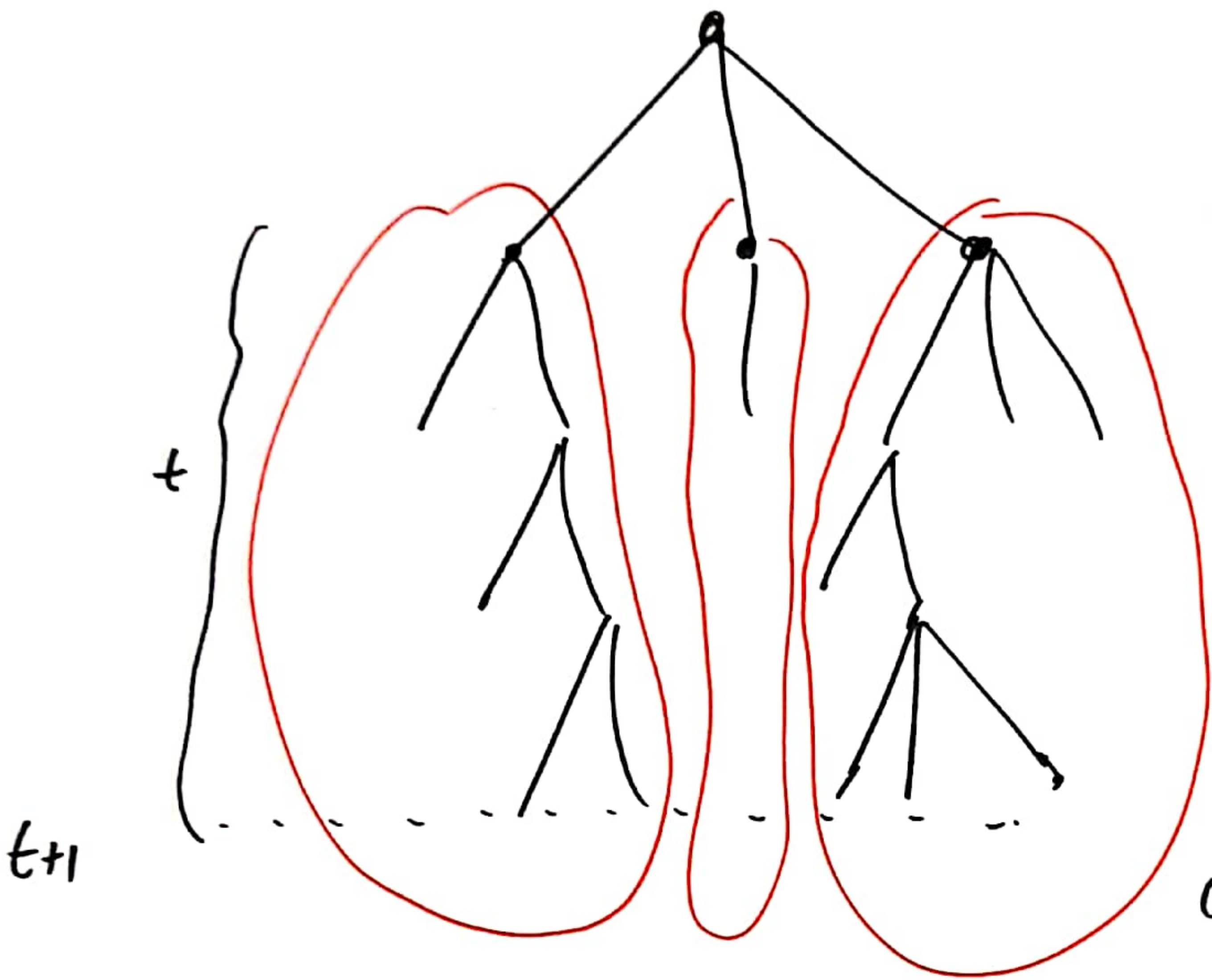
Let $\eta_t = \mathbb{P}(X_t = 0)$

$t+1 = 1+t$



t

$t+1$

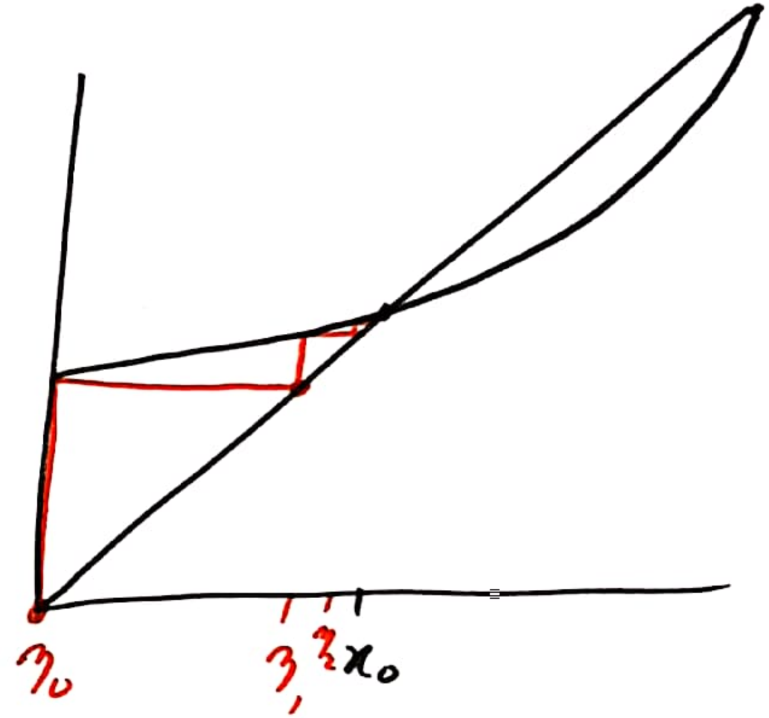


Conditional on $X_1 = k$ the k trees $T_1, \dots, T_k$ giving the descendants of the k individuals in gen. 1 are INDEP copies of $\underset{\sim}{X}$

$$\mathbb{P}(X_{t+1} = 0 \mid X_t = k) = \eta_t^k$$

$$\eta_{t+1} : \mathbb{P}(X_{t+1} = 0) = \sum_k \mathbb{P}(Z = k)\, \eta_t^k = f_Z(\eta_t)$$

Thm 5.1 For any prob. dist Z on $\mathbb{Z}_{\geq 0}$, $\eta(Z)$, the prob. that $\underset{\sim}{X}_Z$ dies out, is the smallest solution in $[0,1]$ to $f_Z(x) = x$.



Proof $f_Z(x) = x$ has a solution $x = 1$.

Has a smallest solution x_0 by continuity

Recall $\eta_{t+1} = f_Z(\eta_t)$, $\eta_0 = 0$

The events $\{X_1 = 0\} \subseteq \{X_2 = 0\} \subseteq \{X_3 = 0\} \cdot \subseteq$ are nested, & $\{\underset{\sim}{X} \text{ dies out}\}$ is their union.

$\therefore$ (η_t) is increasing & converges to $\eta = \eta(Z)$

$$\eta_{t+1} = f_2(\eta_t)$$
$$\downarrow \qquad \downarrow$$
$$\eta \qquad f_2(\eta)$$

$\therefore \eta = f_2(\eta)$ ie. η is <u>A</u> solution to this equation. $\therefore \eta \geq x_0$

Now $\eta_0 = 0 \leq x_0$

$\eta_t \leq x_0 \Rightarrow f_2(\eta_t) \leq f_2(x_0)$

ie. $\eta_{t+1} \leq x_0$

so by induction $\eta_t \leq x_0 \quad \forall t$

$$\downarrow$$
$$\eta \leq x_0$$

$\square$

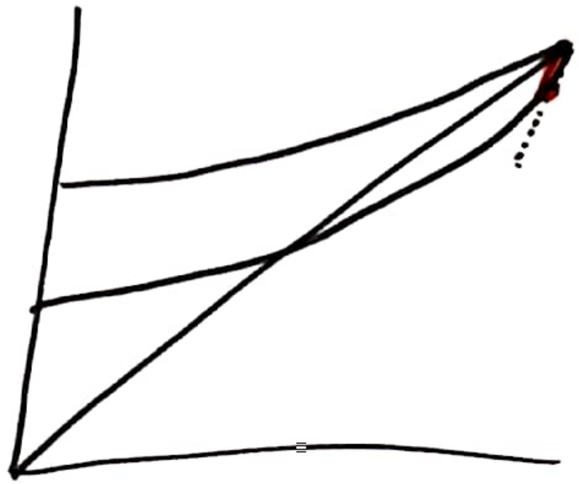
Cor 5.2 Suppose $\mathbb{E}[Z]>1$ then $\eta(Z)<1$.

Q is $\mathbb{P}(\text{survives})>0$?

$\eta(Z)<1$

If $\mathbb{E}[Z]<1$ then $\eta(Z)=1$

If $\mathbb{E}[Z]=1$ & $\mathbb{P}(Z=1)<1$ then $\eta(Z)=1$



Proof By Thm 5.1 $\eta(Z)<1 \iff$

$\exists$ a solution to $f_Z(x)=x$ in $[0,1)$

- if $\lambda=\mathbb{E}Z>1$ $\quad f_Z'(1)>1$

  $\Rightarrow f_Z(1-\varepsilon)<1-\varepsilon$ for small ε

Since $f_Z(0)\geq 0$ by IVT $\exists\, x\in[0,1-\varepsilon]$

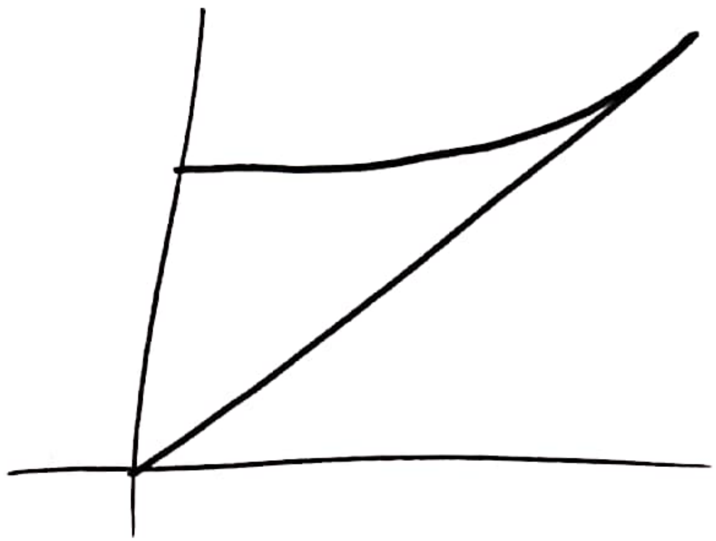
s.t. $f_Z(x)=x$.

$\mathbb{E}Z < 1$ done already

$\mathbb{E}Z = 1$ & $\mathbb{P}(Z=1) < 1$

Then $\mathbb{P}(Z \geq 2) > 0$

$\therefore f_Z$ strictly convex

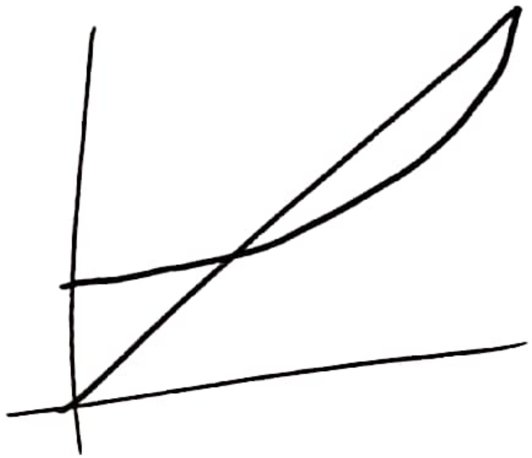


$f_Z''(x) > 0$ on $[0,1)$

$\Rightarrow f_Z'(x) < f_Z'(1) = 1$ on $(0,1)$

$\Rightarrow f_Z(x) > x$ on $[0,1)$

$\square$



Defn For $c>0$ a RV Z has the Poisson distribution with parameter (mean) c if $\mathbb{P}(Z=k) = \frac{c^k}{k!}e^{-c}$ $k=0,1,2,\dots$

written $Z \sim Po(c)$

Lemma 5.3 Suppose $n\to\infty$, $p\to 0$ with $np\to c$, $c>0$ is constant. let Z_n have binomial distribution $Bi(n,p)$. Then $Z_n \xrightarrow{d} Z \sim Po(c)$, i.e. for any fixed k $\mathbb{P}(Z_n=k) \to \mathbb{P}(Z=k)$

$(1-p)^k \sim 1$

Proof Fix k

$1-p = e^{-\underbrace{(-\log(1-p))}_{p+O(p^2)}}$

$$\mathbb{P}(Z_n=k) = \binom{n}{k}p^k(1-p)^{n-k} \sim \frac{n^k}{k!}p^k(1-p)^n$$

$$= \frac{(np)^k}{k!}e^{-np+O(np^2)} \to \frac{c^k}{k!}e^{-c} \quad \text{as } np\to c$$

$np^2 = (np)\cdot p \to c\cdot 0$
$=0$ □

$\eta(c)$ $\eta_{Po(c)}$

<u>Theorem 5.4</u> For $c > 0$ the extinction probability $\eta := \eta(c)$ of the $Po(c)$ - Galton - Watson BP satisfies

$$\eta = e^{-c(1-\eta)}$$

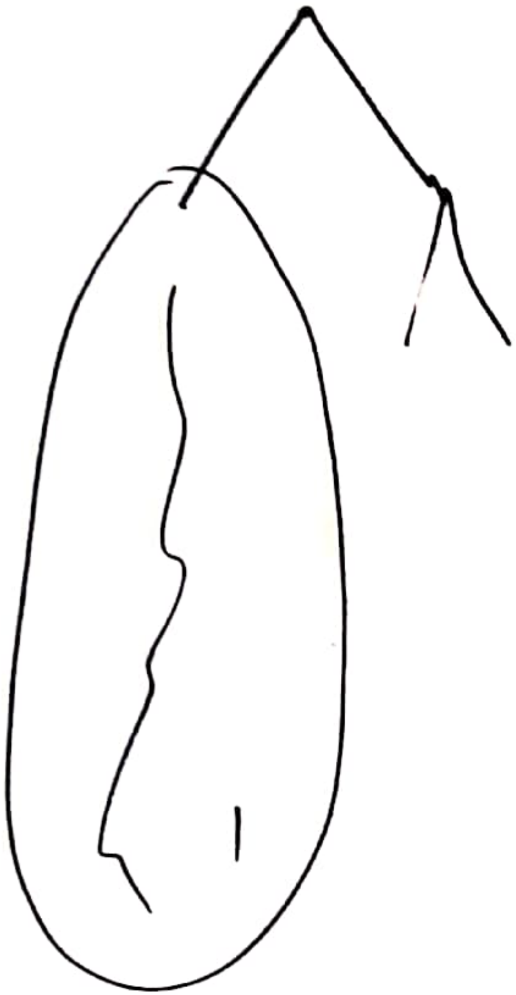
Moreover $\eta < 1 \Leftrightarrow c > 1$.

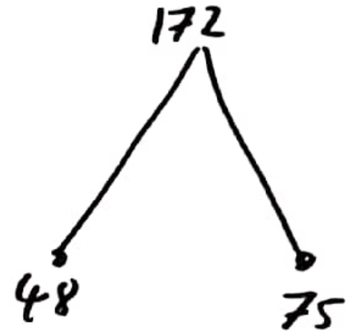
<u>Pf</u> Standard 1st year maths & if $Z \sim Po(c)$

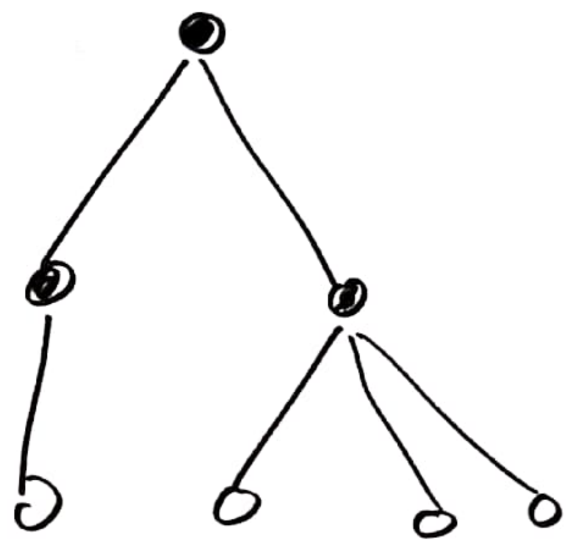
$$f_Z(x) = e^{-c(1-x)}.$$

By Cor 5.2 $\eta = \eta(c)$ is a solution to this equation

& $\eta < 1 \Leftrightarrow c = \mathbb{E}[Z] > 1$ $\square$

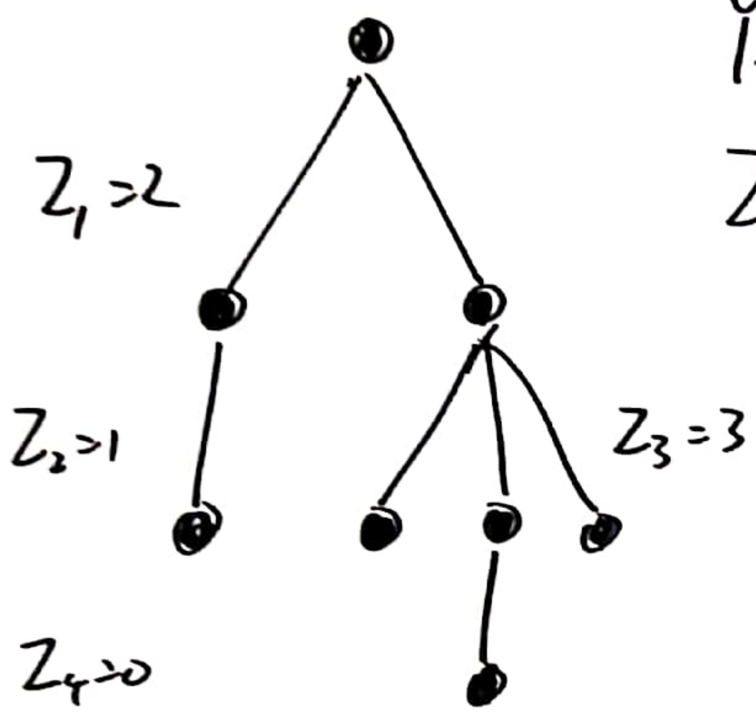



172
48
75



Y_t^{bp} = # live children after t steps

Z_t = # born in step t



t	0	1	2	3	4	5	6	7	8
Z_t		2	1	3	0	0	1	0	0
Y_t^{bp}	1	2	2	4	3	2	2	1	0

$$Y_t^{bp} = \begin{cases} Y_{t-1}^{bp} + Z_t - 1 & \text{if } Y_{t-1}^{bp} > 0 \\ 0, & \text{if } Y_{t-1}^{bp} = 0 \end{cases}$$

○ = live

● = dead

so (Y_t^{Gp}) is just a random walk with iid

increments $Z_t - 1$, except that if hit zero then stop.

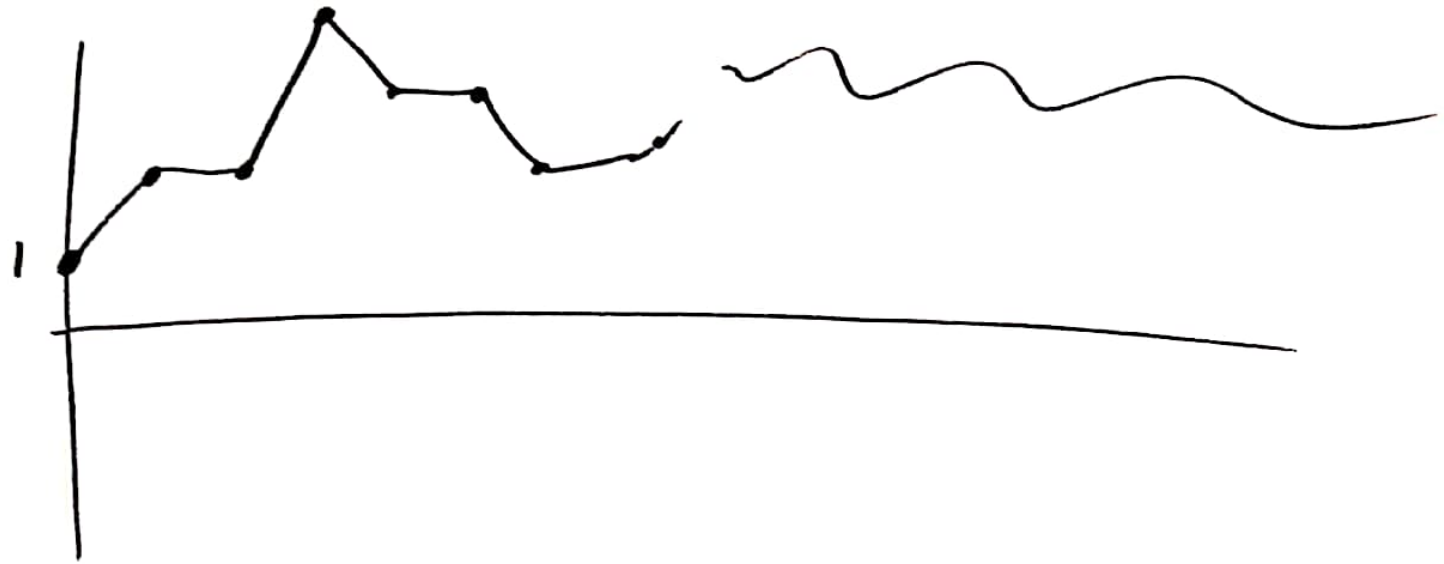
If hit zero at step m then at end there are m

dead individuals

$\therefore$ total size of $\underset{\sim}{X}$ $\left(= \sum_i x_i\right) = m$

If dont hit zero $\underset{\sim}{X}$ is ∞

so $|\underset{\sim}{X}| = \inf\{t : Y_t^{Gp} = 0\}$.



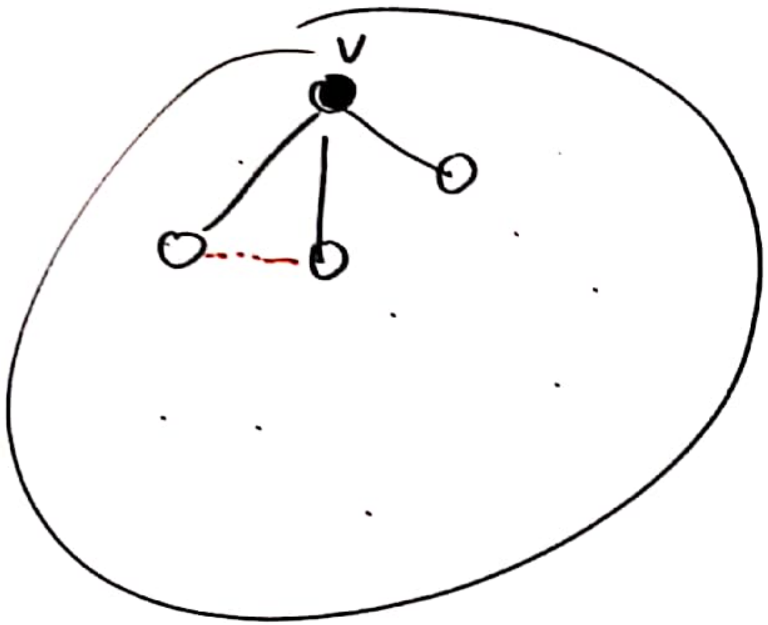
'drift' of RW $= \mathbb{E}[\text{step size}] = \mathbb{E}(Z) - 1$

$= \lambda - 1$

If $\lambda \leq 1$ RW has -ve / 0 drift. $\Rightarrow \mathbb{P}(\text{hit } 0) = 1$

$\lambda > 1$ RW has +ve drift

$\mathbb{P}(\text{hit } 0) < 1$



o = live

• = (dead) = processed

· = unreached

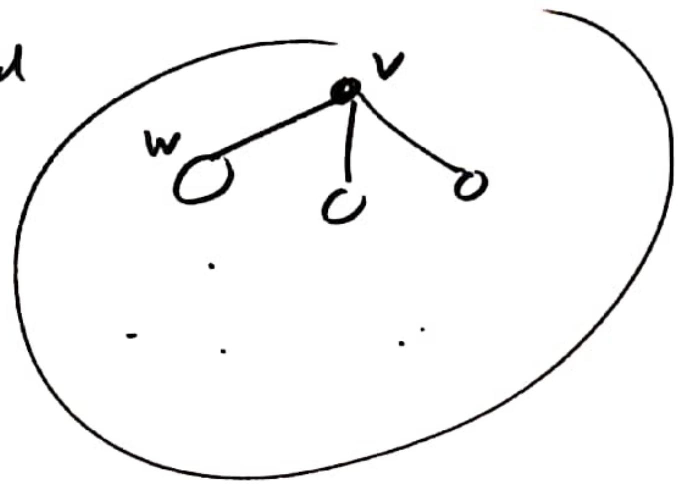
Start at time $t=0$ with v live all other vertices unreached

After t step let Y_t = # live vxs

Note there are then t processed

$U_t = n - t - Y_t$ unreached vertices

In step t: pick a live vx w if $\exists$ (o/w stop)

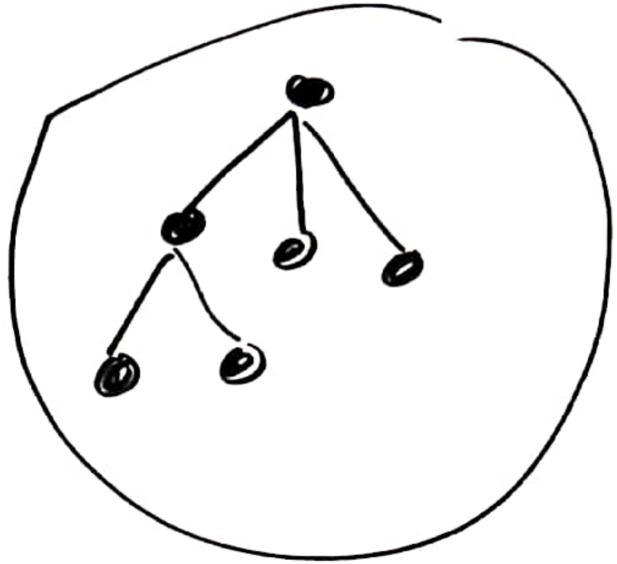


- For each w' that is unreached (after $t-1$ steps) 'test' is $ww' \in E(G)$. If so, w' becomes live
- w becomes processed.

Let R_t = # vxs reached in step t

Note $Y_t = \begin{cases} Y_{t-1} + R_t - 1 & \text{if } Y_{t-1} > 0 \\ 0 & \text{o/w} \end{cases}$

t	0	1	2	3	4	5	6
R_t		3	2	0	0	0	0
Y_t	1	3	4	3	2	1	0



Key observation status of a given changes only

$$\text{unreached} \xrightarrow{?} \text{live} \xrightarrow{\checkmark} \text{processed.}$$

At end, no live vxs

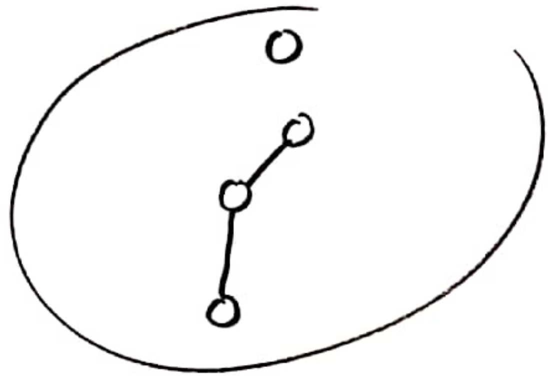
$S = \{$processed vxs at end$\} = \{w : \exists\, vw \text{ path}\}$

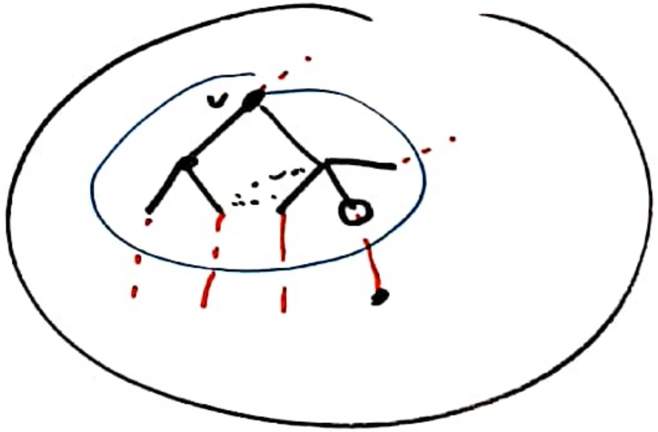
$= V(\text{cpt of } G \text{ containing } v)$

Easy check (ind on t)

w becomes live at step t

$\Rightarrow$ $\exists\, vw$ path





unreached

↓

live

↓

processed

Key fact: algorithm never tests an edge twice:

or ⊘

when we test $e = ww'$ 

state before testing is 

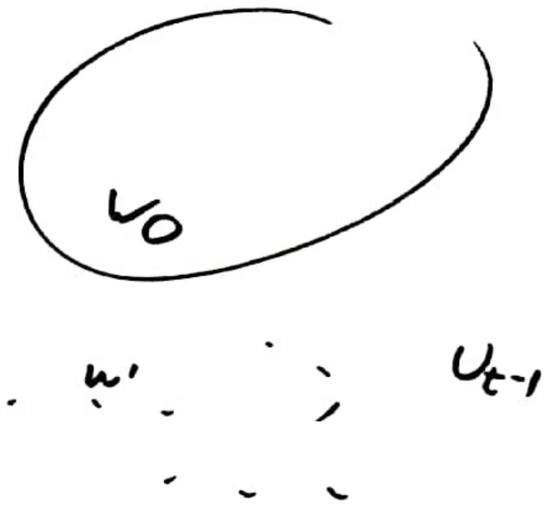
after

from here never test again as one end processed.

Consequence: given history (conditional on $Y_0 = y_0$, $Y_1 = y_1, \dots Y_{t-1} = y_{t-1}$) conditional distribution of R_t is $\mathrm{Bin}(u_{t-1}, p)$

where $u_{t-1} = n - (t-1) - y_{t-1}$

$=$ # unreached vxs at start of step t.



$\underset{\sim}{X}_{Po(c)}$ $\quad |\underset{\sim}{X}_{Po(c)}| = \sum_{t \geq 0} X_t \leq \infty$

For $k \geq 1$ let $\rho_k(c) = \mathbb{P}(|\underset{\sim}{X}_{Po(c)}| = k)$

C_v = cpt containing v

Lemma 5.5

$(p = \frac{c}{n})$

$np \to c$

Suppose $p = p(n)$ satisfies $np \to c$ as $n \to \infty$ where $c > 0$ is constant. Then for any fixed k

$$\mathbb{P}(|C_v| = k) \to \rho_k(c) \text{ as } n \to \infty.$$

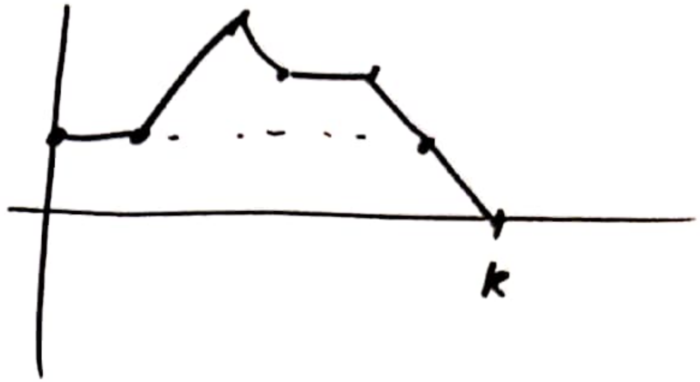
Proof Consider $(Y_t)_{t \geq 0}$ & $(Y_t^{op})_{t \geq 0}$ as just defined.

Note $|C_v| = k \iff Y_t$ first hits 0 at $t = k$

$|\underset{\sim}{X}_{Po(c)}| = k \iff Y_t^{op}$ " " "

Recall $Y_t = Y_{t-1} + R_t - 1$

$$Y_t^{bp} = Y_{t-1}^{bp} + Z_t - 1$$

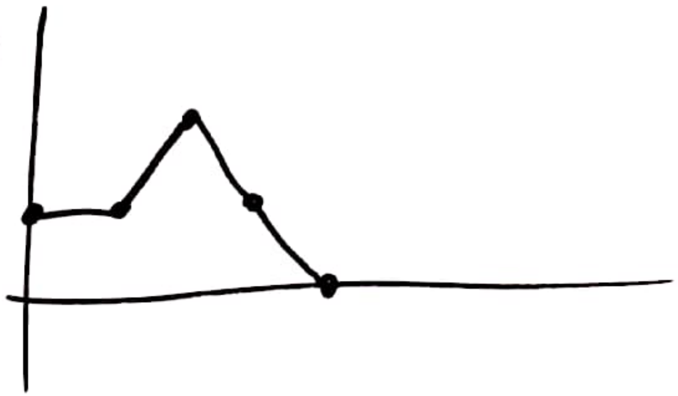


Fix k. Set S_k of possible trajectories first hitting 0 at $t=k$ is FINITE.

Fix one. STP $\mathbb{P}(Y_t \text{ follows this})$

$$\to \mathbb{P}(Y_t^{bp} \ \ldots) \quad \text{as } n \to \infty$$

So (Y_t) follows this specific trajectory ⊛ eg $k=4$



$\Rightarrow R_1=1,\ R_2=2,\ R_3=0,\ R_4=0$

$\mathbb{P}(R_t = r_t \mid \text{history})$

$$= \mathbb{P}(\text{Bin}(\underbrace{n-(t-1)-y_{t-1}}_{U_{t-1}}, p) = r_t)$$

t	0	1	2	3	4
y_t	1	1	2	1	0
r_t		1	2	0	0
U_t	$n-1$	$n-2$	$n-4$	$n-4$	$n-4$

$\therefore$

$$⊛ = \mathbb{P}(\text{Bin}(n-1,p)=1)\,\mathbb{P}(\text{Bin}(n-2,p)=2)$$
$$\cdot\,\mathbb{P}(\text{Bin}(n-4,p)=0)\,\mathbb{P}(\text{Bin}(n-4,p)=0)$$

But $(n-1)p,\ (n-2)p,\ (n-4)p \sim np \to c$ as $n \to \infty$

$\therefore$ by Lemma 5.3 $\quad ⊛ \to \mathbb{P}(\text{Po}(c)=1)\,\mathbb{P}(\text{Po}(c)=2)\,\mathbb{P}(\text{Po}(c)=0)\,\mathbb{P}(\text{Po}(c)=0)$

$= \mathbb{P}(Y_t^{\text{up}} \text{ follows trajectory})$ $\square$

Notn Let $N_k(G) = \#$ vertices of G in k-vertex cpts

$(\#\ v \text{ s.t. } |C_v| = k)$

$= k \times \#$ cpts with k vertices

Cor 5.6 Suppose $p = p(n)$ s.t. $np \to c > 0$ constant.

For any fixed k $\quad \mathbb{E}[N_k(G(n,p))] \sim n \rho_k(c)$

Proof $N_k = \sum_{v \in V} \mathbb{1}_{\{|C_v| = k\}}$

$\therefore \mathbb{E}[N_k] = \sum_{v \in V} \mathbb{P}(|C_v| = k) = n\,\mathbb{P}(|C_v| = k)$ $\square$