STATS 325 Exam Solutions Semester 2, 2014

1a) First-step analysis equations: $d_x = \mathbb{P}(\text{ever reach state } 1 \mid \text{start in state } x)$, $x = 2, 3, 4$.

$\Rightarrow \; d_2 = \frac{3}{5} + \frac{2}{5} d_3$ ①

$d_3 = \frac{3}{5} d_2$ ②

Substitute ② in ①: $d_2 = \frac{3}{5} + \frac{2}{5} \cdot \frac{3}{5} d_2$

$\Rightarrow 19 d_2 = 15$

$\Rightarrow \; d_2 = \frac{15}{19}$ as stated.

b) Let $h_x = \mathbb{P}(\text{ever reach state } 4 \mid \text{start at state } x)$, $x = 0, \dots, 4$.

Then: $h_0 = 0$, $h_4 = 1$

and: $h_1 = \frac{2}{5} h_2$

$h_2 = \frac{3}{5} h_1 + \frac{2}{5} h_3$

$h_3 = \frac{3}{5} h_2 + \frac{2}{5}$

c) Consider $\mathbb{P}(D \mid S_2 \cap W) = \mathbb{P}_2(D \mid W)$ where $\mathbb{P}_2(\cdot) = \mathbb{P}(\cdot \mid S_2)$ for shorthand.

$$= \frac{\mathbb{P}_2(W \mid D)\,\mathbb{P}_2(D)}{\mathbb{P}_2(W)} \quad \circledast \text{ Bayes Thm.}$$

Now $\mathbb{P}_2(D) = \mathbb{P}(D \mid S_2) = d_2 = \frac{15}{19}$ from definitions and (a).

Also $\mathbb{P}_2(W) = \mathbb{P}(W \mid S_2) = h_2 = \frac{20}{65}$ (given).

Finally, $\mathbb{P}_2(W \mid D) = \mathbb{P}(W \mid S_1)$ because trajectory starts in state 2, and later reaches state 1 (event D), so only the most recent state (state 1) is relevant.

1c cont.) So $\mathbb{P}_2(W \mid D) = \mathbb{P}(W \mid S_1) = h_1 = \frac{8}{65}$ (given.)

Substituting all in $\circledast$:

$$\mathbb{P}(D \mid S_2 \cap W) = \frac{\frac{8}{65} * \frac{15}{19}}{\frac{20}{65}} \quad \text{ie.} \quad \frac{h_1 * d_2}{h_2}$$

$$\mathbb{P}(D \mid S_2 \cap W) = \frac{6}{19} \quad (0.316).$$

So $\mathbb{P}(D \mid S_2) = d_2 = \frac{15}{19}$ is larger than $\mathbb{P}(D \mid S_2 \cap W) = \frac{6}{19}$.

This is expected because $\mathbb{P}(D \mid S_2)$ includes trajectories where the game is lost, all of which must go through state 1; whereas $\mathbb{P}(D \mid S_2 \cap W)$ includes only trajectories where the game is won, many of which will avoid state 1 altogether (e.g. trajectory $2 \to 3 \to 4$). *See also comment below

1d) Communicating classes: $\{0\}$ closed,

$\{1, 2, 3\}$ not closed,

$\{4\}$ closed.

1e) The Markov chain does not converge to equilibrium independent of start state. For example, if the start state is state 0, then $X_t = 0$ for all t with probability 1; whereas if the start state is 4, then $X_t = 4$ for all t w.p. 1.

*Further comment on Q1c: it is NOT correct to say that $\mathbb{P}(D \mid S_2)$ is larger because $S_2 \cap W$ is a smaller set than S_2, or because there are fewer paths from state $2 \to$ state $1 \to$ win than from state $2 \to 1$. We condition on $S_2 \cap W$ so its size or probability relative to that of S_2 is irrelevant. The question is whether we are more likely to visit state 1 if we do or don't condition on W as well as on S_2.
Note $\mathbb{P}_2(D) = \mathbb{P}_2(D \mid W)\mathbb{P}_2(W) + \mathbb{P}_2(D \mid \overline{W})\mathbb{P}_2(\overline{W}) > \mathbb{P}_2(D \mid W)\{\mathbb{P}_2(W) + \mathbb{P}_2(\overline{W})\} = \mathbb{P}_2(D \mid W)$

2a) $G(s) = (1+\mu-\mu s)^{-1}$

$\Rightarrow G'(s) = -(1+\mu-\mu s)^{-2} * (-\mu)$

$\Rightarrow G'(1) = \dfrac{\mu}{(1+\mu-\mu)^2} = \mu.$

So $\mathbb{E}(Y) = G'(1) = \mu$ as stated.

b) γ is the smallest non-negative solution of $G(s)=s$:

$G(s)=s \Rightarrow \dfrac{1}{1+\mu-\mu s} = s$

$1 = s + s\mu - \mu s^2$

$\Rightarrow \mu s^2 - (\mu+1)s + 1 = 0$

Factorising: we know $(s-1)$ must be a factor.

$\Rightarrow (s-1)(\mu s-1) = 0$

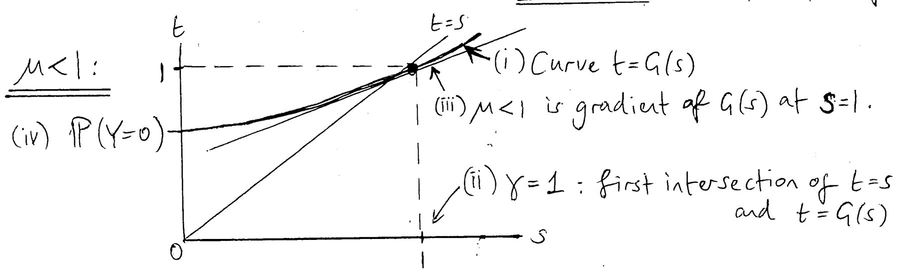
Solutions: $s=1,\ s=\frac{1}{\mu}$. The smaller of these is γ.

Case (i): $\mu<1 \Rightarrow \frac{1}{\mu}>1 \Rightarrow \gamma=1$ (definite extinction).

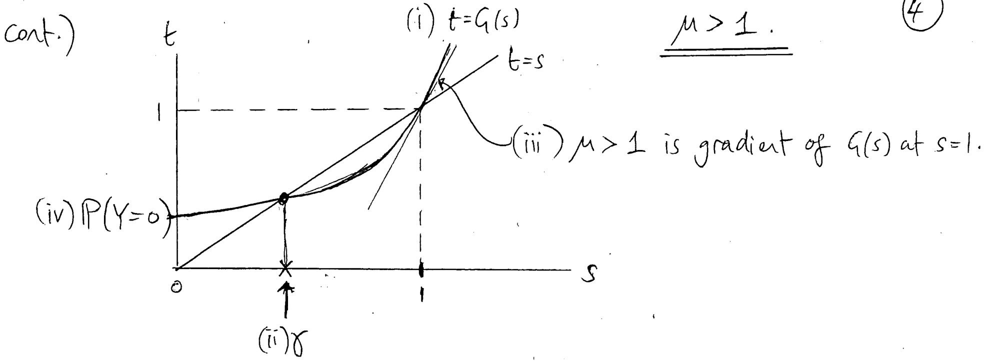
Case (ii): $\mu=1 \Rightarrow \frac{1}{\mu}=1 \Rightarrow \gamma=1$ (definite extinction)

Case (iii): $\mu>1 \Rightarrow \frac{1}{\mu}<1 \Rightarrow \gamma=\frac{1}{\mu}$ (extinction possible but not definite)

c) $\mu<1$:



2c cont.) $\mu>1$.



2d) $\mu=1 \Rightarrow G(s) = \dfrac{1}{1+\mu-\mu s} = \dfrac{1}{2-s}.$

Then $\mathbb{P}(Z_1=0) = G(0) = \frac{1}{2}$

and $\mathbb{P}(Z_2=0) = G(G(0)) = \dfrac{1}{2-\frac{1}{2}} = \dfrac{2}{3}.$

2e) RTP: $G_n(0) = \dfrac{n}{n+1}$ for $n=1,2,3,\dots$ (*).

Base case: $n=1$.

LHS (*) $= G_1(0) = G(0) = \frac{1}{2}$ from part (d).

RHS (*) $= \frac{1}{1+1} = \frac{1}{2} =$ LHS (*).

So (*) is proved for base case $n=1$.

General case: Suppose (*) is true for $n=1,2,\dots,x$.

So we may assume $G_x(0) = \dfrac{x}{x+1}$. (a)

RTP (*) holds for $n=x+1$

$\Rightarrow$ RTP $G_{x+1}(0) = \dfrac{x+1}{x+2}$ (**)

2e cont.) $\text{LHS}(**) = G_{x+1}(0)$

$= G(G_x(0))$ by hint in Question

$= G\left(\frac{x}{x+1}\right)$ using allowed (a)

$= \dfrac{1}{2 - \frac{x}{x+1}}$ using $G(s)$ from part (d)

$= \dfrac{x+1}{2(x+1) - x}$

$= \dfrac{x+1}{x+2}$

$= \text{RHS}(**)$

So if (*) holds for $n = x$, we have proved (*) holds for $n = x+1$.
We proved (*) holds for $n=1$, thus (*) is proved for all $n = 1, 2, 3, \ldots$ □

2f) $\mathbb{P}(\text{first goes extinct at generation } 10) = \mathbb{P}(Z_{10} = 0 \cap Z_9 > 0)$

$= \mathbb{P}(Z_{10} = 0) - \mathbb{P}(Z_9 = 0)$

$= G_{10}(0) - G_9(0)$

$= \frac{10}{11} - \frac{9}{10}$ using $G_n(0) = \frac{n}{n+1}$ from part (e)

$= 0.0091$

Comment about Q6d: NOT required for marks.
In the random walk (RW), a path ABAAB... is actually two possible paths, RLRRL... or LRLLR... So the probability of either of 2 paths of length $2t$, each of which has probability $\left(\frac{1}{2}\right)^{2t}$, is $2 * \left(\frac{1}{2}\right)^{2t} = \left(\frac{1}{2}\right)^{2t-1}$.
In the branching process (BP), with $Y \sim \text{Geo}(\frac{1}{2})$, a Geometric "success" is required for the conclusion of every family (prob $= \frac{1}{2}$); and every individual except the first arises from a Geometric "failure" (prob $= \frac{1}{2}$). If $T = t$, there are t families, so we need t successes and $t-1$ failures. So each graph with $T = t$ has prob $\left(\frac{1}{2}\right)^{2t-1}$.

3a) $P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}$

b) $\mathbb{P}(C_2 \cap S_3 \mid C_1) = \frac{3}{4} * \frac{1}{4} = \frac{3}{16}$ (0.188).

c) Require $\underset{\sim}{\pi}^T P = \underset{\sim}{\pi}^T$ (1) and $\pi_1 + \pi_2 = 1$. (2)
Sufficient to substitute the $\underset{\sim}{\pi}$ given:

$\text{LHS}(1) = \left(\frac{1}{3} \;\; \frac{2}{3}\right) \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix} = \left(\frac{1}{6} + \frac{2}{12} \;\;\; \frac{1}{6} + \frac{6}{12}\right) = \left(\frac{1}{3} \;\; \frac{2}{3}\right) = \text{RHS}(1)$

So $\underset{\sim}{\pi}^T = \left(\frac{1}{3} \;\; \frac{2}{3}\right)$ satisfies (1), and satisfies (2) clearly,
so $\underset{\sim}{\pi}^T = \left(\frac{1}{3} \;\; \frac{2}{3}\right)$ is an equilibrium distribution for $\{Y_t\}$.

d) The chain is irreducible and aperiodic, and an equilibrium distn exists, so the chain does converge to this distn $\underset{\sim}{\pi}$ as $t \to \infty$.

e) $X_t = \sum_{i=1}^{n} Y_{it}$ so $\mathbb{E}(X_t) = \sum_{i=1}^{n} \mathbb{E}(Y_{it})$

$= n * \mathbb{E}(Y_{it})$

So $\lim_{t \to \infty} \mathbb{E}(X_t) = \frac{2}{3} n$

because as $t \to \infty$, each $Y_{it} = \begin{cases} 1 & \text{with probability } \pi_1 = \frac{2}{3} \\ 0 & \text{" \quad " \quad } \pi_0 = \frac{1}{3} \end{cases}$

and we established in part (d) that this is true because each chain $\{Y_{it}\}_{t \in \mathbb{N}}$ converges to equilibrium $\underset{\sim}{\pi}$.

f) X_t is the sum of n Bernoulli random variables $Y_{1t}, \ldots, Y_{nt}$.
As $t \to \infty$, each Y_{it} converges to equilibrium with $\mathbb{P}(Y_{it} = 1) = \frac{2}{3}$.

3f cont.) So X_t must converge to the equilibrium distribution constituting the sum of these independent Bernoulli r.v.s.

So the equilibrium distn for $\{X_t\}$ is $\underline{\underline{\text{Binomial}(n, \frac{2}{3})}}$

and it <u>does</u> converge to this distribution as $t \to \infty$.

3g) W_{t+1} is the number of the $n - X_t$ whales that were at sea on day t, that return to the coast on day $t+1$.

Clearly, $\underline{\underline{U_{t+1} \mid X_t \sim \text{Binomial}(X_t, \frac{3}{4})}}$

and $\underline{\underline{W_{t+1} \mid X_t \sim \text{Binomial}(n - X_t, \frac{1}{2})}}$.

3h) <u>Option B is correct.</u>

We need $Q_{rs} = \mathbb{P}(X_{t+1} = s \mid X_t = r)$.

Option A just specifies the equilibrium distribution for X_{t+1}, which is not the same as the one-step conditional distn of $X_{t+1} \mid X_t$.

Option B partitions correctly over the values of U_{t+1}, $u = 1, \dots, r$:

$$q_{rs} = \mathbb{P}(X_{t+1} = s \mid X_t = r) = \sum_{u=0}^{r} \mathbb{P}(U_{t+1} = u \mid X_t = r)\,\mathbb{P}(W_{t+1} = s - u \mid X_t = r).$$

---

4) A: Voter Process. Q1: yes, gives $\mathbb{P}$(concensus for one opinion).
Q2: yes, e.g. in gene spread, $\mathbb{E}$(time to fixation).
Q3: no: the presence of two absorbing states implies no convergence to equilibrium. Any distn of the form $(\pi_0, 0, 0, \dots, 0, 1-\pi_0)$ is an equilibrium distn so we learn nothing about long-term behaviour by solving equilibrium eqns.

B: Two-Armed Bandit. Q1: no, hitting probability is 1 for both states.
Q2: no, there is no absorbing state.
Q3: yes, interest in long-run prob of success.

C: Random Walk. Q1: yes, interested in whether we will definitely reach particular states, studied via defective r.v.s and PGFs.
Q2: no, there is no absorbing state.
Q3: no, equilibrium distns won't exist.

5a)
$$
\begin{aligned}
G_Y(t) = \mathbb{E}(t^Y) &= \sum_{y=0}^{\infty} t^y\, \mathbb{P}(Y = y) \\
&= \sum_{y=0}^{\infty} t^y p (1-p)^y \\
&= p \sum_{y=0}^{\infty} \{t(1-p)\}^y \\
&= \frac{p}{1-(1-p)t} \quad \text{for } |t(1-p)| < 1 : \text{sum of Geometric Series}
\end{aligned}
$$

So $\underline{\underline{G_Y(t) = \dfrac{p}{1-(1-p)t} \quad \text{for} \quad -\dfrac{1}{1-p} < t < \dfrac{1}{1-p}}}$.

b) $N = Y_1 + \dots + Y_k$

so
$$
\begin{aligned}
G_N(s) = \mathbb{E}(s^N) &= \mathbb{E}(s^{Y_1 + Y_2 + \dots + Y_k}) \\
&= \mathbb{E}(s^{Y_1})\,\mathbb{E}(s^{Y_2}) \dots \mathbb{E}(s^{Y_k}) \quad \text{by independence} \\
&= G_Y(s)\, G_Y(s) \dots G_Y(s) \quad \text{by definition} \\
&= \{G_Y(s)\}^k
\end{aligned}
$$

$\underline{\underline{\Rightarrow G_N(s) = \left\{ \dfrac{p}{1-(1-p)s} \right\}^k \quad \text{for} \quad \dfrac{-1}{1-p} < s < \dfrac{1}{1-p}}}$.

c)
$$
\begin{aligned}
G_X(s) &= \mathbb{E}(s^X) \\
&= \mathbb{E}_Y\{\mathbb{E}(s^X \mid Y)\} \quad \text{Law of Total Expectation} \\
&= \mathbb{E}_Y\left\{ \left(\frac{2/3}{1 - s/3} \right)^{Y+1} \right\} \quad \text{because } X \mid Y \sim \text{NegBin}(Y+1, \tfrac{2}{3}) \text{ so } \mathbb{E}(s^X \mid Y) = \left(\frac{2/3}{1-s/3}\right)^{Y+1} \\
&= \frac{2}{3-s} \cdot \mathbb{E}_Y\left\{ \left(\frac{2}{3-s}\right)^Y \right\} \\
&= \frac{2}{3-s} \cdot G_Y(t) \quad \text{where} \quad t = \frac{2}{3-s}
\end{aligned}
$$

5c cont.) So $G_X(s) = \dfrac{2}{3-s} \cdot \dfrac{2/5}{1-3t/5}$ for $Y \sim \text{Geometric}\left(\frac{2}{5}\right)$: $G_Y(t) = \dfrac{2/5}{1-3t/5}$ by (a)

$$= \frac{2}{3-s} \cdot \frac{2}{5-3t}$$

$$= \frac{2}{3-s} \cdot \frac{2}{5 - \frac{3*2}{3-s}} \quad \text{replacing } t = \frac{2}{3-s} \text{ from above}$$

$$= \frac{4}{5(3-s)-6}$$

$$\Rightarrow \underline{\underline{G_X(s) = \frac{4}{9-5s}}} \quad \text{as stated.}$$

Thus $G_X(s) = \dfrac{4/9}{1-5s/9} = \dfrac{p}{1-(1-p)s}$ where $p = \frac{4}{9}$: PGF of Geo$\left(\frac{4}{9}\right)$.

$$\Rightarrow \underline{\underline{X \sim \text{Geometric}\left(\tfrac{4}{9}\right).}}$$

5d) Consider $\mathbb{E}(s^Z) = \mathbb{E}(s^{X+Y}) = \mathbb{E}_Y\{\mathbb{E}(s^{X+Y} | Y)\}$

$$\Rightarrow \mathbb{E}(s^Z) = \mathbb{E}_Y\{s^Y \mathbb{E}(s^X | Y)\}$$

$$= \mathbb{E}_Y\left\{s^Y \cdot \left(\frac{2}{3-s}\right)^{Y+1}\right\} \quad \text{using } \mathbb{E}(s^X | Y) \text{ from above}$$

$$= \frac{2}{3-s} \mathbb{E}_Y\left\{\left(\frac{2s}{3-s}\right)^Y\right\}$$

$$= \frac{2}{3-s} \cdot G_Y(t) \quad \text{where } t = \frac{2s}{3-s}$$

$$= \frac{2}{3-s} \cdot \frac{2}{5 - \frac{3*2s}{3-s}} \quad \text{using } G_Y(t) = \frac{2}{5-3t} \text{ and } t = \frac{2s}{3-s}$$

$$= \frac{4}{15-5s-6s}$$

$$\Rightarrow \underline{\underline{\mathbb{E}(s^Z) = \frac{4/15}{1-(11/15)s}}} \quad \text{So } \underline{\underline{Z \sim \text{Geometric}\left(\tfrac{4}{15}\right)}} \text{ as stated.}$$

6a) $T^{(n)} = 1 + T_1^{(n-1)} + \dots + T_Y^{(n-1)}$

$$\Rightarrow H_n(s) = \mathbb{E}(s^{T^{(n)}})$$

$$= \mathbb{E}\left(s^{1+T_1^{(n-1)} + \dots + T_Y^{(n-1)}}\right)$$

$$= s\, \mathbb{E}\left(s^{T_1^{(n-1)}} \cdots\cdots s^{T_Y^{(n-1)}}\right)$$

$$= s\, \mathbb{E}_Y\left\{\mathbb{E}\left(s^{T_1^{(n-1)} + \dots + T_Y^{(n-1)}} | Y\right)\right\} \quad \text{LoTE}$$

$$= s\, \mathbb{E}_Y\left\{\mathbb{E}\left(s^{T_1^{(n-1)}} \cdots s^{T_Y^{(n-1)}}\right)\right\}$$

where Y is considered constant inside $\{\ \}$: valid because $T_1^{(n-1)}, T_2^{(n-1)}, \dots$ are independent of Y.

$$= s\, \mathbb{E}_Y\left\{\mathbb{E}\left(s^{T_1^{(n-1)}}\right) \cdots \mathbb{E}\left(s^{T_Y^{(n-1)}}\right)\right\}$$

because $T_1^{(n-1)}, T_2^{(n-1)}, \dots$ are indept of each other

$$= s\, \mathbb{E}_Y\left\{H_{n-1}(s)^Y\right\} \quad \text{by defn of } \mathbb{E}\left(s^{T_i^{(n-1)}}\right) = H_{n-1}(s)$$

$$= s\, G_Y\left(H_{n-1}(s)\right)$$

$$\Rightarrow \underline{\underline{H_n(s) = \frac{s}{2 - H_{n-1}(s)}}} \quad (*) \quad \text{using } G_Y(t) = \frac{1}{2-t} \text{ given.}$$

– as stated.

6b) Take limit of both sides of (*) as $n \to \infty$: $\lim_{n\to\infty} H_n(s) = \lim_{n\to\infty} H_{n-1}(s) = H(s)$ given

$$\Rightarrow H(s) = \frac{s}{2 - H(s)}$$

$$\Rightarrow 2H(s) - \{H(s)\}^2 = s$$

$$\Rightarrow \{H(s)\}^2 - 2H(s) + s = 0$$

6b cont.) Solving quadratic: $H(s) = \dfrac{2 \pm \sqrt{4-4s}}{2}$

$$\Rightarrow \quad \underline{H(s) = 1 \pm \sqrt{\tfrac{4-4s}{4}} = 1 \pm \sqrt{1-s}}$$

Now $H(0) = \mathbb{P}(T=0) = 0$ is known, because $T = 1 + Z_1 + \dots \geq 1$.

Check (+) root: $1 + \sqrt{1-0} = 2 \neq 0$, so (+) root is wrong.

Check (−) root: $1 - \sqrt{1-0} = 0$; correct.

So $\underline{\underline{H(s) = 1 - \sqrt{1-s}}}$ as stated.

6c) Let V = #steps to move one place to the right.

Note: V has the same distn as #steps to move one place to the left, by symmetry.

Let $K(s) = \mathbb{E}(s^V)$:

now $V = \begin{cases} 1 & \text{w.p. } 1/2 \\ 1 + V' + V'' & \text{w.p. } 1/2 \end{cases}$ where $V' \sim V'' \sim V$, V', V'' are indept.

So $\mathbb{E}(s^V) = K(s) = \frac{1}{2}s + \frac{1}{2}s\,K(s)^2$ by indep of V', V'' and because $V' \sim V'' \sim V$.

So $s\,K(s)^2 - 2K(s) + s = 0$

$$\Rightarrow \quad \underline{K(s) = \frac{2 \pm \sqrt{4-4s^2}}{2s} = \frac{1}{s}\left\{1 \pm \sqrt{1-s^2}\right\}}.$$

Now $\mathbb{P}(V=0) = 0 \Rightarrow K(0) = 0$, only possible with (−) root.

So $\underline{K(s) = \mathbb{E}(s^V) = \frac{1}{s}\left\{1 - \sqrt{1-s^2}\right\}}$.

We require $W = \begin{cases} 1 + V^* & \text{w.p. } 1/2 \\ 1 + V^{**} & \text{w.p. } 1/2 \end{cases}$ where $V^* \sim V^{**} \sim V$

6c cont.) So $J(s) = \mathbb{E}(s^W) = s\,\mathbb{E}(s^V) = s\,K(s)$

$$\Rightarrow \quad \underline{\underline{J(s) = 1 - \sqrt{1-s^2}}} \quad \text{as stated.}$$

6d) We have: $H(s) = \mathbb{E}(s^T) = 1 - \sqrt{1-s}$ for T = total progeny

$J(s) = \mathbb{E}(s^W) = 1 - \sqrt{1-s^2}$ for W = return time

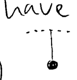
So $J(s) = H(s^2)$

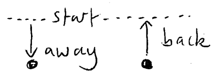
ie. $\underline{\mathbb{E}(s^W) = \mathbb{E}((s^2)^T) = \mathbb{E}(s^{2T})}$

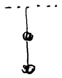
Thus W and $2T$ have the same PGF, so $\underline{\underline{W \sim 2T.}}$

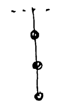
(relationship between W and T).

This suggests that, for every completed graph of a branching process, containing T individuals, we should be able to construct a unique path from 0 to 0 in the random walk consisting of $2T$ steps, and that (assuming this correspondence does hold) the path should have the same probability as the graph does in the branching process.

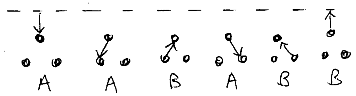
i) 

$T=1$, $W=2T=2$ steps. Path is AB : 

ii)  $T=2$, $W=2T=4$ steps. Path is AABB : the only way of first return to 0 taking 4 steps.

iii)  $T=3$, $W=2T=6$ steps. Logical extension of (ii) is path AAABBB.

iv)  $T=3$, $W=2T=6$ steps again. Only other way of a 6-step path for first return to 0 is AABABB.



Note: (i)–(iv) describe an algorithm for matching graphs of size T to paths of length $2T$. Not required, but for proof of same probability, see Box on Page 5.