Universität Zürich[UZH]

UNIVERSITY OF ZURICH
INSTITUTE OF MATHEMATICS

MASTER THESIS

Trees, Clusters and Cycles in Random Graphs

Junzhuo Ren

Supervisor: Prof. Dr. Jean Bertoin

Date: May 2026

Abstract

This thesis provides an introduction to several classical models of random discrete structures and the probabilistic techniques used in their analysis.

After a brief review of Markov chains and their fundamental properties, we study one-dimensional random walks, with particular emphasis on skip-free random walks. We investigate their asymptotic behaviour and review several classical results, including Kemperman's formula and the probability of remaining positive forever.

We then turn to Bienaymé–Galton–Watson (BGW) trees. After introducing the space of plane trees and defining random trees as random variables on this space, we study the construction of BGW trees and their encoding by Łukasiewicz walks. These tools are applied to the study of extinction probabilities and to various combinatorial counting results.

The thesis next considers the Erdős–Rényi random graph model. We begin with a coupling construction and investigate local structural properties of random graphs. Particular attention is devoted to the phase transition associated with the emergence of a giant connected component. We show that the neighbourhood of a given vertex can be approximated by a Poisson(c)-BGW tree, thereby establishing a direct connection with the framework developed earlier.

Finally, we discuss two classical couplings of random permutations: the Feller coupling and the Chinese restaurant coupling. Using the Feller coupling together with Poissonization techniques, we derive limiting distributions for cycle sizes and cycle counts. The Chinese restaurant coupling is then used to study the growth of random recursive trees. We conclude with a brief discussion of the Yule tree and its connection to the random recursive tree.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Prof. Dr. Jean Bertoin, for suggesting this fascinating thesis topic, for his patience in answering my questions, and for the many insightful discussions we had throughout this project. I have learned a great deal from his guidance, both in mathematics and in the way of approaching research.

I am also grateful to Zheng Fang and Dr. Alejandro Rosales Ortiz for their guidance during my studies. In particular, the Probability II course taught by Alejandro introduced me to many useful probabilistic tools that proved invaluable in the preparation of this thesis.

My deepest thanks go to my parents and my family for their unwavering support, encouragement, which made it possible for me to pursue my studies in Switzerland.

Finally, I would like to thank myself—for persevering through the challenges, staying curious, and completing this journey.

Contents

Abstract **3**

Acknowledgements **5**

1 Preliminaries on Markov Chains **8**
1.1 Canonical Markov chain . . . 8
1.2 Strong Markov property . . . 8
1.3 Classification of states . . . 9
1.4 A Useful Lemma on Stopping Times . . . 9

2 Random Walks: The Core Tool **10**
2.1 One-dimensional random walk on $\mathbb{Z}$ . . . 10
2.1.1 Irreducibility, Aperiodicity and SMP . . . 10
2.1.2 Asymptotic behavior . . . 12
2.1.3 Walks with finite mean . . . 13
2.2 Skip-free random walk . . . 14
2.2.1 Duality . . . 15
2.2.2 Cycle Lemma . . . 16
2.2.3 Kemperman's formula . . . 17
2.2.4 Staying positive forever . . . 17

3 Bienaymé–Galton–Watson Trees **21**
3.1 Bienaymé-Galton-Watson trees . . . 21
3.1.1 Plane Tree . . . 21
3.1.2 Tree space . . . 23
3.1.3 Bienaymé–Galton–Watson trees . . . 25
3.2 Łukasiewicz walk and direct applications . . . 27
3.2.1 Łukasiewicz Walk . . . 28
3.2.2 Tree-Walk correspondence . . . 29
3.3 Probabilistic counting of trees . . . 31
3.3.1 Uniform geometric BGW plane trees . . . 31
3.3.2 Cayley and Poisson BGW trees . . . 32

4 Erdős–Rényi Random Graph **34**
4.1 Erdős–Rényi Random Graph . . . 34
4.2 Local property . . . 35
4.2.1 Moment methods . . . 35
4.2.2 Isolated vertices . . . 35
4.2.3 Connectivity . . . 36

4.3 Birth of giant . . . 37
4.3.1 Poisson(c)-BGW tree . . . 37
4.3.2 Cut and Sprinkling . . . 38
4.3.3 Phase transition for the giant component . . . 42

5 Random Permutations and Recursive Trees 50
5.1 Random permutation . . . 50
5.1.1 Foata correspondence . . . 50
5.1.2 Feller Coupling . . . 51
5.1.3 Large cycles and Poisson-Dirichlet distribution . . . 53
5.1.4 Poisson count of small cycles . . . 54
5.2 Random recursive tree . . . 59
5.2.1 Chinese Restaurant process . . . 59
5.3 Yule Tree . . . 61
5.3.1 Properties of exponential laws . . . 61
5.3.2 From Yule Tree to RRT . . . 62

A Auxiliary Proofs 64
A.1 Proof that A is a subgroup of $\mathbb{Z}$ . . . 64
A.2 Proof the case $r = 2$ . . . 65
A.3 Auxiliary Convergence Results . . . 66
A.4 Anscombe's theorem . . . 67

References 69

Chapter 1

Preliminaries on Markov Chains

The material of this chapter is largely based on [4, Chapter 2].

1.1 Canonical Markov chain

Throughout this chapter, E is a finite or countable space. We take

$$\mathbf{\Omega} = E^{\mathbb{Z}_+},$$

so that an element of $\mathbf{\Omega}$ is a sequence $\omega = (\omega_0, \omega_1, \ldots)$ with values in E. The coordinate mappings Y_n, $n \in \mathbb{Z}_+$, are defined by

$$Y_n(\omega) := \omega_n.$$

We equip $\mathbf{\Omega}$ with the smallest σ-field for which all coordinate mappings Y_n, $n \in \mathbb{Z}_+$, are measurable. We denote this σ-field by $\mathfrak{F}$. Equivalently, $\mathfrak{F}$ is generated by the cylinder sets

$$\{(\omega_0, \omega_1, \omega_2, \ldots) \in E^{\mathbb{Z}_+} : \omega_0 = x_0, \ldots, \omega_n = x_n\},$$

where $n \in \mathbb{Z}_+$ and $x_0, \ldots, x_n \in E$.

Lemma 1.1. *Let Q be a stochastic matrix on E. For every $\mu \in \mathcal{M}_E$, there exists a unique probability measure $\mathbb{P}_\mu$ on $(\mathbf{\Omega}, \mathfrak{F})$ such that, under $\mathbb{P}_\mu$, the coordinate process $(Y_n)_{n\in\mathbb{Z}_+}$ is a Markov chain with transition matrix Q, and*

$$\mathbb{P}_\mu(Y_0 = x) = \mu(x).$$

1.2 Strong Markov property

We use the notation θ_T to denote the shift operator at stopping time T. We introduce the canonical filtration

$$\mathfrak{F}_n = \sigma(Y_0, Y_1, \ldots, Y_n), \quad n \geq 0.$$

Theorem 1.2 (Strong Markov property)**.** *Let T be a stopping time of $(\mathfrak{F}_n)_{n\in\mathbb{Z}_+}$, and let G be a nonnegative (or integrable) measurable function on $\mathbf{\Omega}$. Then, for every $\mu \in \mathcal{M}_E$,*

$$\mathbb{E}_\mu \left[\mathbf{1}_{\{T<\infty\}}\, G \circ \theta_T \mid \mathfrak{F}_T \right] = \mathbf{1}_{\{T<\infty\}} \mathbb{E}_{Y_T}[G].$$

1.3 Classification of states

Theorem 1.3. *Suppose the chain is irreducible. Then exactly one of the following cases holds:*

- *All states are recurrent: for every $\mu \in \mathcal{M}_E$ and every $x \in E$,*

$$\mathbb{P}_\mu(N_x = \infty) = 1.$$

- *All states are transient: for every $\mu \in \mathcal{M}_E$ and every $x \in E$,*

$$\mathbb{P}_\mu(N_x < \infty) = 1.$$

1.4 A Useful Lemma on Stopping Times

Lemma 1.4. *If T and R are $\mathfrak{F}_n$-stopping times, then*

$$W := T + R \circ \theta_T$$

is also a $\mathfrak{F}_n$-stopping time.

Proof. Let $n \in \mathbb{N}$. We first observe that

$$\begin{aligned}\{W = n\} &= \bigcup_{m=0}^{n} \{T = m\} \cap \{R \circ \theta_T = n - m\} \\ &= \bigcup_{m=0}^{n} \{T = m\} \cap \{R \circ \theta_m = n - m\}.\end{aligned}$$

Since T is a stopping time, we have

$$\{T = m\} \in \mathfrak{F}_m \subset \mathfrak{F}_n.$$

Moreover, since R is a stopping time,

$$\{R = n - m\} \in \mathfrak{F}_{n-m} = \sigma(Y_0, \ldots, Y_{n-m}).$$

Hence, there exists a measurable set $A \subset E^{n-m+1}$ such that

$$\{R = n - m\} = \{(Y_0, \ldots, Y_{n-m}) \in A\}.$$

By the definition of the shift operator, it follows that

$$\{R \circ \theta_m = n - m\} = \{(Y_m, Y_{m+1}, \ldots, Y_n) \in A\} \in \sigma(Y_m, \ldots, Y_n) \subset \mathfrak{F}_n.$$

Therefore, for each $m \leq n$, the set

$$\{T = m\} \cap \{R \circ \theta_m = n - m\}$$

belongs to $\mathfrak{F}_n$, and hence

$$\{W = n\} \in \mathfrak{F}_n.$$

□

Chapter 2

Random Walks: The Core Tool

The following chapters are based primarily on [1]. Several proofs and arguments have been adapted to fit the present exposition.

2.1 One-dimensional random walk on $\mathbb{Z}$

Definition 2.1. Let $\mu = (\mu_k)_{k\in\mathbb{Z}}$ be a probability distribution on $\mathbb{Z}$ such that

$$\mu(\mathbb{Z}_{<0}) > 0 \quad \text{and} \quad \mu(\mathbb{Z}_{>0}) > 0.$$

Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. $\mathbb{Z}$-valued random variables with common distribution μ. Define the process $(S_n)_{n\geq 0}$ by $S_0 = 0$ and, for $n \geq 1$,

$$S_n = X_1 + \cdots + X_n.$$

The process $(S_n)_{n\geq 0}$ is called a one-dimensional random walk with step distribution μ. It is a Markov chain with transition kernel Q given by

$$Q(i, j) = \mu(j - i), \qquad i, j \in \mathbb{Z}.$$

We now introduce the associated canonical Markov chain $(Y_n)_{n\geq 0}$ with state space $\mathbb{Z}$ and transition kernel Q. The process $(S_n)_{n\geq 0}$ under P has the same distribution as $(Y_n)_{n\geq 0}$ under $\mathbb{P}_0$. In the sequel, we will often work with the canonical version when it is more convenient.

2.1.1 Irreducibility, Aperiodicity and SMP

Proposition 2.2 (Parity)**.** *The Markov chain* $(S_n)_{n\geq 0}$

(i) is irreducible if and only if $\operatorname{supp}(\mu)$ *is not contained in* $a\mathbb{Z}$ *for any integer* $a > 1$*.*

(ii) is aperiodic if $\operatorname{supp}(\mu)$ *is not contained in* $a\mathbb{Z} + b$ *for any integers* $a > 1$ *and* $b \in \mathbb{Z}$*.*

Proof. Note $(S_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ have the same parity.

(i) Suppose that $\operatorname{supp}(\mu) \subset a\mathbb{Z}$ for some $a > 1$. Then

$$P(S_n \in a\mathbb{Z}) = 1 \quad \text{for all } n \geq 0.$$

In particular,

$$P(S_n = 1) = \mathbb{P}_0(Y_n = 1) = 0 \quad \text{for all } n \geq 0,$$

so the chain is not irreducible.

Conversely, let A denote the set of states accessible from 0. (See Appendix A.1 for a proof that A is a subgroup of $\mathbb{Z}$.) It is well known that any subgroup of $\mathbb{Z}$ is of the form $g\mathbb{Z}$ for some integer $g \geq 0$. Hence $A = a\mathbb{Z}$ for some $a \geq 0$. Since $\text{supp}(\mu) \subset A$, the assumption implies that $a = 1$. Moreover, $a \neq 0$ since $\mu(0) \neq 1$. Therefore $A = \mathbb{Z}$, and the chain is irreducible.

(ii) By part (i), the chain is irreducible. Let d denote its period, and suppose that $d > 1$. Then there exist $a, b > 0$ such that

$$Q(0, a) > 0 \quad \text{and} \quad Q(0, -b) > 0,$$

which implies that $d \mid (a + b)$.

Let $x \in \text{supp}(\mu)$, with $x \neq 0$. If $x > 0$, then $d \mid (x + b)$, and together with $d \mid (a + b)$ this yields $d \mid (a - x)$. If $x < 0$, then similarly $d \mid (a - x)$.

Thus, for all $x \in \text{supp}(\mu)$, we have $d \mid (a - x)$, and hence

$$\text{supp}(\mu) \subset d\mathbb{Z} + a,$$

with $d > 1$, which contradicts the assumption. Therefore, the chain is aperiodic. □

Proposition 2.3 (Strong Markov property)**.** *Let $(Y_n)_{n\geq 0}$ be the canonical Markov chain with initial distribution $\mu \in \mathcal{M}_E$, and let τ be a $(\mathfrak{F}_n)_{n\geq 0}$-stopping time. Then, on the event $\{\tau < \infty\}$ (implicitly of positive probability), the process*

$$(Y_n^{(\tau)})_{n\geq 0} = (Y_{\tau+n} - Y_\tau)_{n\geq 0}$$

under $\mathbb{P}_\mu$ is independent of $\mathfrak{F}_\tau$ and is distributed as the initial walk $(Y_n)_{n\geq 0}$, i.e.

$$\mathbb{P}_\mu\Big(A \cap (Y_{\tau+n} - Y_\tau)_{n\geq 0} \in B \mid \tau < \infty\Big) = \mathbb{P}_\mu(A \mid \tau < \infty) \cdot \mathbb{P}_0(B),$$

for $A \in \mathfrak{F}_\tau$, $B \in \mathfrak{F}$.

Proof.

$$\begin{aligned}
\mathbb{P}_\mu\Big(A \cap (Y_{\tau+n} - Y_\tau)_{n\geq 0} \in B, \tau < \infty\Big) &= \mathbb{E}_\mu\Big(1_A \, 1_{(Y_{\tau+n} - Y_\tau)_{n\geq 0} \in B} \, 1_{\tau<\infty}\Big) \\
&= \mathbb{E}_\mu\Big(\mathbb{E}_\mu\Big(1_A \, 1_{(Y_{\tau+n} - Y_\tau)_{n\geq 0} \in B} \, 1_{\tau<\infty} \mid \mathfrak{F}_\tau\Big)\Big) \\
&= \mathbb{E}_\mu\Big(1_A \, 1_{\tau<\infty} \, \mathbb{E}_\mu\Big(1_{(Y_n - Y_0)_{n\geq 0} \in B} \mid \mathfrak{F}_\tau\Big)\Big)
\end{aligned}$$

Recall that the sets of the form

$$C_{i_0,\ldots,i_n} = \{\omega \in \mathbb{Z}^{\mathbb{N}} : \omega_0 = i_0, \cdots, \omega_n = i_n\}$$

for $n \geq 0$, $i_k \in \mathbb{Z}$, $0 \leq k \leq n$, generate the sigma-field $\mathfrak{F}$ and form a π-system. By a monotone class argument, it suffices to show that for every such $C_{i_0,\ldots,i_n}$,

$$\mathbb{P}_x\Big((Y_n - Y_0)_{n\geq 0} \in C_{i_0,\ldots,i_n}\Big) = \mathbb{P}_0\Big((Y_n)_{n\geq 0} \in C_{i_0,\ldots,i_n}\Big).$$

The right-hand side is given by

$$\begin{aligned}\mathbb{P}_0((Y_n)_{n\geq 0} \in C_{i_0,\cdots,i_n}) &= \mathbb{P}_0(Y_0 = i_0, Y_1 = i_1, \cdots, Y_n = i_n) \\ &= 1_{i_0=0} Q(0, i_1) Q(i_1, i_2) \cdots Q(i_{n-1}, i_n).\end{aligned}$$

The left-hand side is given by

$$\begin{aligned}\mathbb{P}_x((Y_n - Y_0)_{n\geq 0} \in C_{i_0,\cdots,i_n}) &= \mathbb{P}_x(0 = i_0, Y_1 - Y_0 = i_1, \cdots, Y_n - Y_0 = i_n) \\ &= \mathbb{P}_x(Y_0 = x, Y_1 = i_1 + x, \cdots, Y_n = i_n + x) \\ &= 1_{i_0=0} Q(x, i_1 + x) Q(i_1 + x, i_2 + x) \cdots Q(i_{n-1} + x, i_n + x) \\ &= 1_{i_0=0} Q(0, i_1) Q(i_1, i_2) \cdots Q(i_{n-1}, i_n).\end{aligned}$$

Moreover, for all non-negative (or integrable) measurable F, we have

$$\mathbb{E}_x\Big(F(Y_n - Y_0)_{n\geq 0}\Big) = \mathbb{E}_0(F).$$

□

2.1.2 Asymptotic behavior

Suppose the chain is irreducible. Let us denote

$$\overline{S} := \limsup_{n\to\infty} S_n \quad \text{and} \quad \underline{S} := \liminf_{n\to\infty} S_n.$$

Proposition 2.4. *For any $k \in \mathbb{Z}$, the probability that $\overline{S}$ or $\underline{S}$ equals k is zero.*

Proof. Suppose there exists $k \in \mathbb{Z}$ such that $P(\overline{S} = k) > 0$. Since (S_n) under P and (Y_n) under $\mathbb{P}_0$ have the same distribution, this implies

$$\mathbb{P}_0(\overline{Y} = k) = P(\overline{S} = k) > 0.$$

It follows that $\mathbb{P}_0(N_k = \infty) > 0$. By the classification of states, the chain is recurrent, and hence $\mathbb{P}_0(N_{k+1} = \infty) = 1$, which is a contradiction. □

This motivates the following definition:

Definition 2.5 (Asymptotic categories of a random walk)**.** A one-dimensional random walk $(S_n)_{n\geq 0}$ satisfies almost surely exactly one of the following:

(i) $\overline{S} = \underline{S} = \infty$, i.e. $S_n \to \infty$, in which case (S_n) is said to **drift towards** $+\infty$;

(ii) $\overline{S} = \underline{S} = -\infty$, i.e. $S_n \to -\infty$, in which case (S_n) is said to **drift towards** $-\infty$;

(iii) $\overline{S} = \infty$ and $\underline{S} = -\infty$, in which case (S_n) is said to **oscillate**.

When a random walk drifts (i.e. $P(S_n \to \infty) = 1$ or $P(S_n \to -\infty) = 1$), the walk is transient.

Proposition 2.6. *If the random walk (S_n) drifts towards $+\infty$, then*

$$P(S_i > 0 \text{ for all } i \geq 1) > 0.$$

Proof. Suppose not, then we have

$$P(S_i \le 0 \text{ for some } i \ge 1) = \mathbb{P}_0(Y_i \le 0 \text{ for some } i \ge 1) = 1.$$

Let

$$\tau_1 := \inf\{n \ge 1 : Y_n - Y_0 \le 0\}, \qquad \tau_{k+1} := \tau_k + \tau_1 \circ \theta_{\tau_k}, \quad k \ge 1.$$

Then $\mathbb{P}_0(\tau_1 < \infty) = 1$. By the strong Markov property, we have

$$\begin{aligned}\mathbb{P}_0(\tau_{k+1} < \infty) &= \mathbb{P}_0(\tau_k < \infty,\ \tau_1 \circ \theta_{\tau_k} < \infty) \\ &= \mathbb{E}_0\Big(\mathbf{1}_{\{\tau_k<\infty\}}\, \mathbb{P}_{Y_{\tau_k}}(\tau_1 < \infty)\Big) \\ &= \mathbb{P}_0(\tau_k < \infty).\end{aligned}$$

Hence,

$$\mathbb{P}_0\Big(\bigcap_{k=1}^{\infty}\{\tau_k < \infty\}\Big) = 1,$$

which is a contradiction.

□

2.1.3 Walks with finite mean

Theorem 2.7 (Dichotomy for walks with finite mean)**.** *Suppose* $E(|X_1|) < \infty$. *Then:*

(i) If $E(X_1) \neq 0$, *then* $(S_n)_{n\ge0}$ *drifts, hence is transient.*

(ii) If $E(X_1) = 0$, *then* $(S_n)_{n\ge0}$ *is recurrent.*

Proof. (i) By the strong law of large numbers, $n^{-1}S_n \to E(X_1)$ almost surely. It follows that $(S_n)_{n\ge0}$ drifts towards $+\infty$ or $-\infty$ depending on the sign of $E(X_1)$.

(ii) By the law of large numbers, $S_n/n \to 0$ almost surely as $n \to \infty$. This implies that for any $\varepsilon > 0$,

$$P(|S_n| \le \varepsilon n) \to 1 \quad \text{as } n \to \infty.$$

In particular,

$$E\left(\sum_{i=0}^{\infty} \mathbf{1}_{\{|S_i|\le\varepsilon i\}}\right) \ge \sum_{i=0}^{n} P(|S_i| \le \varepsilon i) \overset{\text{Cesàro}}{\ge} \frac{n}{2}$$

eventually.

If the walk were transient, then for some constant $C > 0$ we would have

$$E\left(\sum_{i=0}^{\infty} \mathbf{1}_{\{S_i=0\}}\right) = \mathbb{E}_0\left(\sum_{i=0}^{\infty} \mathbf{1}_{\{Y_i=0\}}\right) \le C.$$

For $k \in \mathbb{Z}$, applying the strong Markov property at the stopping time $\tau_k = \inf\{i > 0 : Y_i = k\}$ gives

$$\mathbb{E}_0\left(\sum_{i=0}^{\infty} \mathbf{1}_{\{Y_i=k\}}\right) = \mathbb{P}_0(\tau_k < \infty)\, \mathbb{E}_0\left(\sum_{i=0}^{\infty} \mathbf{1}_{\{Y_i=0\}}\right) \le C.$$

Hence,

$$E\left(\sum_{i=0}^{\infty}\mathbf{1}_{\{|S_i|\le\varepsilon n\}}\right)\le\sum_{-\varepsilon n\le k\le\varepsilon n}E\left(\sum_{i=0}^{\infty}\mathbf{1}_{\{S_i=k\}}\right)\le 2\varepsilon nC,$$

which contradicts the earlier inequality for $\varepsilon>0$ small enough. Therefore, the walk cannot be transient. □

Theorem 2.8 (Wald's equality)**.** *Suppose $E(|X_1|)<\infty$. Let τ be a stopping time of the canonical filtration $(\mathcal{F}_n)_{n\ge0}$, where $\mathcal{F}_n=\sigma(X_1,\cdots,X_n)$. If $E(\tau)<\infty$, then $E(|S_\tau|)<\infty$ exists and satisfies:*

$$E(\tau)\cdot E(X_1)=E(S_\tau).$$

Proof. Note that

$$S_n-n\,E(X_1),\quad n\ge0,$$

is an $(\mathcal{F}_n)_{n\ge0}$-martingale.

Take $\tau\wedge n$, which is also a stopping time. Note that $\mathbf{1}_{\{\tau>k\}}$ is $\mathcal{F}_k$-measurable. Firstly,

$$|S_{\tau\wedge n}|\le\sum_{k=0}^{\infty}|S_{k+1}-S_k|\mathbf{1}_{\{k<\tau\wedge n\}}\le\sum_{k=0}^{\infty}|X_{k+1}|\mathbf{1}_{\{\tau>k\}}=:Y.$$

Now

$$\begin{aligned}E(Y)&=\sum_{k=0}^{\infty}E(|X_{k+1}|\mathbf{1}_{\{\tau>k\}})\\&=\sum_{k=0}^{\infty}E\Big(E(|X_{k+1}|\mid\mathcal{F}_k)\mathbf{1}_{\{\tau>k\}}\Big)\\&=\sum_{k=0}^{\infty}E(|X_1|)\,P(\tau>k).\end{aligned}$$

Since τ is integrable,

$$\sum_{k=0}^{\infty}P(\tau>k)=E(\tau)<\infty.$$

Thus $S_{\tau\wedge n}$ is dominated by an integrable random variable. Hence $S_{\tau\wedge n}-(\tau\wedge n)E(X_1)$ is a uniformly integrable martingale, and by the optional stopping theorem,

$$E(S_\tau-\tau E(X_1))=0.$$

□

2.2 Skip-free random walk

In this section, we still consider a one-dimensional random walk $(S_n)_{n\ge0}$ with step distribution μ, under the additional assumption that $\mathrm{Supp}(\mu)\subseteq\{-1,0,1,2,\dots\}$. The expectation

$$m=\sum_{k\ge-1}k\mu_k$$

is always well-defined and belongs to $[-1,\infty]$. By the 2.2, skip-free random walk is always irreducible. Thus, by Theorem 2.7, the walk is recurrent if $m=0$.

2.2.1 Duality

Proposition 2.9. *For each fixed $n > 0$, we have*

$$(S_0, S_1, \ldots, S_n) \stackrel{d}{=} (S_n - S_n, S_n - S_{n-1}, S_n - S_{n-2}, \ldots, S_n - S_0).$$

Proof. For $n > 0$, by the independence and identical distribution of the increments,

$$(X_n, X_{n-1}, \ldots, X_1) \stackrel{d}{=} (X_1, X_2, \ldots, X_n),$$

and applying the mapping $f(x_1, \ldots, x_n) = (0, x_1, x_1 + x_2, \ldots, x_1 + \cdots + x_n)$ gives the stated equality in distribution. □

Proposition 2.10. *We denote by*

$$\overline{S}_n = \sup\{0 \le k \le n : S_k\}, \qquad \underline{S}_n = \inf\{0 \le k \le n : S_k\}$$

the running supremum and infimum processes. We suppose that (S_n) drifts towards ∞ so that

$$\min S = \underline{S}_\infty > -\infty \quad a.s.$$

Then we have

$$\overline{S}_n - S_n \xrightarrow[n\to\infty]{(d)} -\underline{S}_\infty.$$

Proof. First, observe that the infimum process always converges. More precisely, for a sequence $(a_n)_{n\ge0}$, define

$$b_n := \inf\{a_0, a_1, \ldots, a_n\}.$$

Then the sequence $(b_n)_{n\ge0}$ converges, and its limit is given by

$$\lim_{n\to\infty} b_n = \inf_{n\ge0} a_n.$$

By duality, we have

$$-\min(S_0, \ldots, S_n) \stackrel{d}{=} -\min(S_n - S_n, \ldots, S_n - S_0) = \max(S_0, \ldots, S_n) - S_n$$

Moreover, if $(X_n)_{n\ge0}$ and $(Y_n)_{n\ge0}$ are sequences of random variables such that X_n and Y_n are identically distributed for each n, and if Y_n converges almost surely to a random variable Y, then Y_n converges in distribution to Y. Consequently, X_n also converges in distribution to Y. □

One of the main applications of duality is the following interpretation of hitting times of half-spaces. We denote the hitting time of $\mathbb{Z}_{<0}$ by

$$T_{\mathbb{Z}_{<0}} := \inf\{i \ge 0 : S_i < 0\}.$$

For each $n \ge 1$, let

$$A_n := \{Y_n \ge Y_{n-1},\ Y_n \ge Y_{n-2},\ \ldots,\ Y_n \ge Y_0\}.$$

For $n \ge 1$, we have

$$P(T_{\mathbb{Z}_{<0}} > n) = P(S_0 = 0, S_1 \ge 0, \ldots, S_n \ge 0)$$

$$
\begin{aligned}
&= P(S_n - S_{n-1} \geq 0,\, S_n - S_{n-2} \geq 0,\, \ldots,\, S_n - S_0 \geq 0) \\
&= P(S_n \geq S_{n-1},\, S_n \geq S_{n-2},\, \ldots,\, S_n \geq S_0) \\
&= \mathbb{P}_0(A_n),
\end{aligned}
$$

Define the weak ascending record times recursively by

$$\tau_1 := \inf\{n \geq 1 : Y_n \geq Y_k \text{ for all } 0 \leq k \leq n-1\}, \qquad \tau_{k+1} := \tau_k + \tau_1 \circ \theta_{\tau_k}, \quad k \geq 1,$$

Then, by the strong Markov property,

$$\mathbb{P}_0(\tau_k < \infty) = \mathbb{P}_0(\tau_1 < \infty)^k := p^k, \quad \forall k \geq 1.$$

We also have

$$\mathbb{P}_0\Big(\sum_{n\geq 1} \mathbf{1}_{A_n} \geq k\Big) = \mathbb{P}_0(\tau_k < \infty) = p^k,$$

If $p < 1$, we can compute the expectation of the hitting time:

$$
\begin{aligned}
E(T_{\mathbb{Z}_{<0}}) &= 1 + \sum_{n\geq 1} P(T_{\mathbb{Z}_{<0}} > n) \\
&= 1 + \sum_{n\geq 1} \mathbb{P}_0(A_n) \\
&= 1 + \mathbb{E}_0\Big(\sum_{n\geq 1} \mathbf{1}_{A_n}\Big) \\
&= \frac{1}{1-p}.
\end{aligned}
$$

2.2.2 Cycle Lemma

Let $n \geq 1$, $x_1, x_2, \ldots, x_n \in \{-1, 0, 1, \ldots\}$ be integers, (s_k) defined by

$$s_0 = 0, \quad s_1 = x_1, \quad s_2 = x_1 + x_2, \quad \ldots, \quad s_n = x_1 + x_2 + \cdots + x_n.$$

If $\ell \in \{0, 1, 2, \ldots, n-1\}$, we consider $(s_k^{(\ell)})$ the ℓ-th cyclic shift of the walk obtained by cyclically shifting its increments ℓ times, that is $(x_{l+1}, ..., x_n, ..., x_1)$ when $\ell \geq 1$.

Lemma 2.11 (Feller)**.** *Fix $n \geq 1$, suppose that $s_n = -k$ for some $k \geq 1$. Then there are exactly k cyclic shifts $(s^{(\ell)})$, with $\ell \in \{0, 1, 2, \ldots, n-1\}$, for which time n is the hitting time of $-k$ by the walk $(s_i^{(\ell)})$.*

Proof. The case $n = 1$ is obvious. Suppose $n \geq 2$.

We first describe how to construct one such shift. Let

$$m := \min\{s_0, s_1, \ldots, s_n\}, \qquad l := \min\{i : s_i = m\}.$$

If $l = n$, then we are done. Otherwise, if $0 \leq l < n$, then shift at l

$$s_i^l := \begin{cases} s_{l+i} - s_l, & 0 \leq i \leq n-l, \\ s_n - s_l + s_{i-(n-l)}, & n-l < i < n. \end{cases}$$

Since the walk is skip-free, the only possible shifts occur at times when the walk hits $\{0, -1, \ldots, -(k-1)\}$ before time n. Otherwise, the hitting time of $-k$ would occur strictly before n. Therefore, shifting at each such hitting time produces exactly k cyclic shifts for which n is the hitting time of $-k$. □

2.2.3 Kemperman's formula

Theorem 2.12 (Kemperman's formula)**.** *For every $k \geq 1$, let $T_{-k} = \inf\{i \geq 0, S_i = -k\}$. Then for every $n \geq 1$, we have*

$$\frac{1}{n} P(S_n = -k) = \frac{1}{k} P(T_{-k} = n),$$

Proof. By the above lemma, we have

$$\mathbf{1}_{\{S_n=-k\}} = \frac{1}{k} \sum_{\ell=0}^{n-1} \mathbf{1}_{\{T_{-k}(S^{(\ell)})=n\}}.$$

Taking expectations on both sides and using the i.i.d. property of the increments, we have, for all $0 \leq \ell \leq n-1$,

$$(S_j^{(\ell)})_{0\leq j\leq n} \stackrel{d}{=} (S_j)_{0\leq j\leq n}$$

.

□

In the following, we say that $(S)_{n\geq 0}$ is skip-free ascending if $-(S)_{n\geq 0}$ is skip-free.

Lemma 2.13. *Let $(S_n)_{n\geq 0}$ be a skip-free ascending random walk. Then for every $n \geq 1$ and every $k \geq 1$ we have*

$$P(S_i > 0 \ \forall 1 \leq i \leq n \mid S_n = k) = \frac{k}{n}.$$

Proof. Notice that the walk $S'_t = S_{n-t} - S_n$ is skip-free descending. So by Kemperman's formula, we have

$$\begin{aligned} P(S_i > 0 \ \forall 1 \leq i \leq n \text{ and } S_n = k) &= P(T_{-k}(S') = n) \\ &= \frac{k}{n} P(S'_n = -k) \\ &= \frac{k}{n} P(S_n = k). \end{aligned}$$

□

2.2.4 Staying positive forever

We now compute the probability that the random walk remains positive after time 1 in the case of skip-free ascending walks and skip-free walks.

Theorem 2.14 (Ascending case)**.** *Let $(S_n)_{n\geq 0}$ be a skip-free ascending random walk such that $\mathbb{E}(X_1) > 0$. Then*

$$P(S_i > 0 \text{ for all } i \geq 1) = \mathbb{E}(X_1).$$

Proof.

$$P(S_i > 0 \ \forall i \geq 1) = \lim_{n\to\infty} P(S_i > 0 \ \forall 1 \leq i \leq n) = \lim_{n\to\infty} \sum_{k\geq 1} P(S_n = k)\frac{k}{n}$$

$$= \lim_{n\to\infty} E(\frac{S_n}{n}\mathbf{1}_{\{S_n>0\}}) = E(X_1).$$

□

Theorem 2.15. *If $(S_n)_{n\geq 0}$ is skip-free, then*

$$P(S_n \geq 0 \text{ for all } n > 0) = 1 - \alpha,$$

where α is the smallest solution in $\alpha \in [0, 1]$ to the equation

$$\alpha = \sum_{k\geq -1} \mu_k \alpha^{k+1}.$$

Proof. Let

$$\tau = \inf\{n \geq 1 : Y_n < 0\}.$$

Then

$$\mathbb{P}_0(\exists n \geq 1 : Y_n < 0) = \mathbb{P}_0(\tau < \infty) := \alpha.$$

Moreover,

$$\begin{aligned}\mathbb{E}_0\Big(\mathbf{1}_{\{\tau<\infty\}}\Big) &= \mathbb{E}_0\left(\mathbf{1}_{\{\tau<\infty\}} \sum_{k=-1}^{\infty} \mathbf{1}_{\{Y_1=k\}}\right) \\ &= \sum_{k=-1}^{\infty} \mathbb{P}_0(\tau < \infty, Y_1 = k) \\ &= \mu_{-1} + \sum_{k\geq 0} \mathbb{P}_0(\tau < \infty, Y_1 = k) \\ &= \mu_{-1} + \sum_{k\geq 0} \mathbb{P}_k(\tau < \infty)\, \mu_k.\end{aligned}$$

note the last display is by the Markov property,

$$\mathbb{P}_0(\tau < \infty, Y_1 = k) = \mathbb{P}_k(\tau < \infty)\, \mathbb{P}_0(Y_1 = k) = \mathbb{P}_k(\tau < \infty)\, \mu_k.$$

Now we show that, for all $k \geq 0$,

$$\mathbb{P}_k(\tau < \infty) = \Big(\mathbb{P}_0(\tau < \infty)\Big)^{k+1}.$$

Let

$$\tau_1 = \inf\{n \geq 1 : Y_n - Y_0 = -1\}, \quad \tau_{n+1} = \tau_n + \tau_1 \circ \theta_{\tau_n}$$

Note we have

$$\mathbb{P}_k(\tau_1 < \infty) = \mathbb{P}_0(\tau < \infty)$$

by induction

$$\begin{aligned}\mathbb{P}_k(\tau < \infty) &= \mathbb{P}_k(\tau_{k+1} < \infty) \\ &= \mathbb{P}_k(\tau_k < \infty)\, \mathbb{P}_0(\tau < \infty) \\ &= \mathbb{P}_k(\tau_1 < \infty)\, (\mathbb{P}_0(\tau < \infty))^k \\ &= \mathbb{P}_0(\tau < \infty))^{k+1}\end{aligned}$$

Thus, we have shown $\mathbb{P}_0(\tau < \infty)$ is a solution to the equation

$$\alpha = \sum_{k \geq -1} \mu_k \alpha^{k+1}.$$

Recall that

$$m = \sum_{k \geq -1} k\mu_k$$

is the mean of μ, and

$$g(x) = \sum_{k \geq -1} \mu_k x^{k+1}.$$

Now we show it is the smallest solution.

Case 1: $m \leq 0$**.**

If $m < 0$, then by Theorem 2.7,

$$S_n \to -\infty \qquad \text{a.s.}$$

Hence $\alpha = 1$.

If $m = 0$, then by Theorem 2.7, the walk is recurrent. Thus again $\alpha = 1$.

Moreover,

$$g'(1) = \sum_{k \geq -1} (k+1)\mu_k = m + 1 \leq 1.$$

Since g is convex on $[0, 1]$ and $g(1) = 1$, it follows that $x = 1$ is the unique solution of $x = g(x)$ in $[0, 1]$.

Case 2: $m > 0$**.**

In this case, again by Theorem 2.7:

$$\mathbb{P}_0(Y_n \geq 0 \text{ for all } n \geq 0) > 0.$$

Hence

$$\alpha = \mathbb{P}_0(\tau < \infty) < 1.$$

Furthermore,

$$g'(1) = m + 1 > 1.$$

Since $g(0) = \mu_{-1} \geq 0$, $g(1) = 1$, and g is convex, it follows that the equation $x = g(x)$ has another solution in $[0, 1)$. Thus α is the smallest solution.

Case 3: $m = \infty$**.**

Let $X_i^N = \min\{X_i, N\}$ and define the truncated random walk $S_n^N = \sum_{i=1}^n X_i^N$. Then $S_n \geq S_n^N$ for all n, and by the monotone convergence theorem,

$$m^N = \mathbb{E}[X_1^N] \to \infty.$$

Thus there exists N such that $m^N > 0$. By the previous case,

$$P(S_n^N \geq 0 \text{ for all } n) > 0.$$

Therefore,

$$P(S_n \geq 0 \text{ for all } n) \geq P(S_n^N \geq 0 \text{ for all } n) > 0,$$

which implies $\alpha < 1$. As before, α is the smallest solution of $x = g(x)$ in $[0, 1]$.

Combining all cases, we conclude that

$$P(S_n \geq 0 \text{ for all } n > 0) = 1 - \alpha,$$

where α is the smallest solution in $[0, 1]$ of

$$\alpha = \sum_{k \geq -1} \mu_k \alpha^{k+1}.$$

□

Chapter 3

Bienaymé–Galton–Watson Trees

The construction of BGW trees is based on [2].

3.1 Bienaymé-Galton-Watson trees

3.1.1 Plane Tree

Let $\mathbb{Z}_{>0} = \{1, 2, \dots\}$. We define

$$\mathcal{U} = \bigcup_{n=0}^{\infty} (\mathbb{Z}_{>0})^n,$$

with the convention that $(\mathbb{Z}_{>0})^0 = \{\emptyset\}$.

- $|u|$ denotes the **length** of the word u, with $|\emptyset| = 0$.
- For $u, v \in \mathcal{U}$, uv denotes their **concatenation**, with the convention that $\emptyset u = u\emptyset = u$.
- If $v = uj$ with $j \in \mathbb{Z}_{>0}$, then u is the **parent** of v, and v is a **child** of u.

Definition 3.1 (Plane tree)**.** A **plane tree** t is a (finite or infinite) subset of $\mathcal{U}$ such that:

1. $\emptyset \in t$.
2. If $v \in t \setminus \{\emptyset\}$, then the **parent** of v also belongs to t.
3. For every $u \in t$, there exists $k_u(t) \in \mathbb{N}_0 \cup \{\infty\}$ such that

$$\forall j \geq 1,\ uj \in t \iff j \leq k_u(t).$$

Proposition 3.2. *A plane tree, viewed as a graph where an edge links two vertices u and v if and only if one is the parent of the other, is a tree.*

Proof. We prove that the graph is connected and acyclic.

Connectedness: Let $u \in t$. Write $u = i_1 i_2 \cdots i_k$. By repeatedly taking parents, we obtain a path

$$u \to i_1 i_2 \cdots i_{k-1} \to \cdots \to i_1 \to \emptyset.$$

All these vertices belong to t, so every vertex is connected to the root $\emptyset$. Hence the graph is connected.

Acyclicity: Suppose for contradiction that there exists a cycle

$$u_1, u_2, \ldots, u_n, u_1.$$

For each i, either u_i is the parent of u_{i+1} or vice versa.

If u_1 is the parent of u_2.
Then necessarily u_2 is the parent of u_3, u_3 is the parent of u_4, and so on. Hence

$$|u_1| < |u_2| < \cdots < |u_n| < |u_1|,$$

which is impossible.

If u_1 is a child of u_2.
Then u_1 is the parent of u_n, u_n is the parent of u_{n-1}, and so on, so

$$|u_1| < |u_2| < \cdots < |u_n| < |u_1|,$$

which is impossible.

Hence the graph is acyclic. Therefore, the graph is a tree. □

We use $\#t$ to denote the size (i.e., the total number of elements) of the tree. If $\#t < \infty$, then we have

$$\#t - 1 = \sum_{u \in t} k_u(t).$$

The height of the tree t is the maximum length of its elements,

$$\text{Height}(t) = \sup\{|u| : u \in t\} \in \mathbb{N}_0 \cup \{\infty\}.$$

We denote by $\mathcal{T}_\infty$ the set of trees and by $\mathcal{T}$ the set of locally finite trees, that is,

$$\mathcal{T} := \{t \in \mathcal{T}_\infty \mid k_u(t) < \infty,\ \forall u \in t\}.$$

Further, we denote the subset of $\mathcal{T}$ that contains only finite trees by

$$\mathcal{T}_f := \{t \in \mathcal{T} \mid \#t < \infty\}.$$

$\mathcal{T}_f$ is countable since it is a subset of $\bigcup_{n \in \mathbb{N}_0} \{A \subset \mathcal{U} \mid \text{card}(A) = n\}$.

Definition 3.3 (Truncation and Shift Operators)**.** Let $\mathcal{T}$ be the set of trees. We define the *truncation operator* tr_n and the *shift operator* θ_u as follows:

$$\text{tr}_n : \mathcal{T} \longrightarrow \mathcal{T}, \quad t \longmapsto \{u \in t \mid |u| \le n\}, \quad \forall n \ge 0,$$

$$\theta_u : \mathcal{T} \to \mathcal{T}, \quad \theta_u(t) := \begin{cases} \{v \in \mathcal{U} : uv \in t\}, & \text{if } u \in t, \\ \{\emptyset\}, & \text{if } u \notin t. \end{cases}$$

It is not hard to show that these two mappings are well defined. Also note if $\text{tr}_n(t) = \text{tr}_n(t')$, then for $0 \le k \le n$, we also have $\text{tr}_k(t) = \text{tr}_k(t')$. In addition, $t = t'$ iff $\text{tr}_n(t) = \text{tr}_n(t')$ for all $n \ge 0$.

3.1.2 Tree space

The truncation operator can be used to define a metric on the set $\mathcal{T}$. This endows $\mathcal{T}$ with a natural topology under which it becomes a measurable (in fact Polish) space, allowing us to define probability measures on $\mathcal{T}$.

This framework is essential for the rigorous study of the law of Bienaymé–Galton–Watson trees and their branching property.

For $t_1, t_2 \in \mathcal{T}$, define

$$m(t_1, t_2) := \sup\{n \in \mathbb{N} : \mathrm{tr}_n(t_1) = \mathrm{tr}_n(t_2)\} \in \mathbb{N}_0 \cup \{\infty\}.$$

We then define a function $d_{\mathcal{T}} : \mathcal{T} \times \mathcal{T} \to \mathbb{R}$ by

$$d_{\mathcal{T}}(t_1, t_2) := 2^{-m(t_1,t_2)},$$

with the convention that

$$2^{-\infty} := 0.$$

Lemma 3.4. *$d_{\mathcal{T}}$ is a metric on $\mathcal{T}$.*

Proof. Clearly, $d_{\mathcal{T}}(t_1, t_2) \geq 0$. If $t_1 = t_2$, then $\mathrm{tr}_n(t_1) = \mathrm{tr}_n(t_2)$ for all n, so $m(t_1, t_2) = \infty$, and hence $d_{\mathcal{T}}(t_1, t_2) = 0$. Conversely, if $d_{\mathcal{T}}(t_1, t_2) = 0$, then $m(t_1, t_2) = \infty$. By the monotonicity of the truncation operators, $\mathrm{tr}_n(t_1) = \mathrm{tr}_n(t_2)$ for all n, and therefore $t_1 = t_2$. Symmetry is immediate.

For any $t_1, t_2, t_3 \in \mathcal{T}$, we have

$$m(t_1, t_3) \geq \min\{m(t_1, t_2), m(t_2, t_3)\},$$

because for $n \in \mathbb{N}, n \leq \min\{m(t_1, t_2), m(t_2, t_3)\}$,

$$\mathrm{tr}_n(t_1) = \mathrm{tr}_n(t_2) = \mathrm{tr}_n(t_3).$$

Thus,

$$d_{\mathcal{T}}(t_1, t_3) = 2^{-m(t_1,t_3)} \leq 2^{-\min\{m(t_1,t_2),m(t_2,t_3)\}}.$$

it follows that

$$d_{\mathcal{T}}(t_1, t_3) \leq d_{\mathcal{T}}(t_1, t_2) + d_{\mathcal{T}}(t_2, t_3),$$

which proves the triangle inequality. □

We denote by $\mathcal{B}(\mathcal{T})$ the Borel σ-algebra of the metric space $(\mathcal{T}, d_{\mathcal{T}})$.

Definition 3.5. Let $(\Omega, \mathcal{F})$ be a measurable space. A mapping $\tau : \Omega \to \mathcal{T}$ is called a random tree if it is $\mathcal{F}/\mathcal{B}(\mathcal{T})$-measurable.

We denote

$$[t]_n := \{t' \in \mathcal{T} \mid \mathrm{tr}_n(t') = \mathrm{tr}_n(t)\}.$$

Proposition 3.6. *$(\mathcal{T}, d_{\mathcal{T}})$ is second countable.*

Proof. To this end, consider the collection of open balls

$$\{B(t,r) \mid t \in \mathcal{T},\, r > 0\}.$$

It suffices to show that this collection coincides with

$$\{B(t,2^{-n}) \mid t \in \mathcal{T}_f,\, n \in \mathbb{N}_0\}.$$

Indeed, it is immediate that

$$B(t,2^{-n}) = [t]_{n+1}.$$

For any $r > 0$, one can find $n \in \mathbb{N}_0$ such that

$$B(t,r) = B(t,2^{-n}) = [t]_{n+1} = [\mathrm{tr}_{n+1}(t)]_{n+1}.$$

□

There is another way to characterize $\mathcal{B}(\mathcal{T})$, which turns out to be useful as well.

Proposition 3.7. *Let* $A_u := \{t \in \mathcal{T} \mid u \in t\}$. *Then,*

$$\sigma\Big(A_u : u \in \mathcal{U}\Big) = \mathcal{B}(\mathcal{T}).$$

Proof. For $t \in \mathcal{T}$ and $n \in \mathbb{N}$,

$$[t]_n = \bigcap_{u \in \mathrm{tr}_n(t)} A_u \bigcap_{u \notin \mathrm{tr}_n(t),\ |u| \le n} A_u^c,$$

. Since the intersection is countable and each A_u, A_u^c belong to $\sigma(A_u : u \in \mathcal{U})$, it follows that

$$\mathcal{B}(\mathcal{T}) \subset \sigma(A_u : u \in \mathcal{U}).$$

Conversely, for any $u \in \mathcal{U}$, we can write

$$A_u = \bigcup_{t \in \mathcal{T}_f \cap A_u} [t]_{|u|}.$$

Since $\mathcal{T}_f$ is countable, the above union is countable, and thus $A_u \in \mathcal{B}(\mathcal{T})$. This shows that

$$\sigma(A_u : u \in \mathcal{U}) \subset \mathcal{B}(\mathcal{T}).$$

□

Now we can show the truncation operators tr_n and the shift operators θ_u are measurable

Lemma 3.8. *The truncation operator* tr_n *is measurable.*

Proof.

$$\mathrm{tr}_n^{-1}(B) = \bigcup_{s \in \mathcal{T}_f,\, \mathrm{tr}_n(s) \in B} [s]_n.$$

□

Lemma 3.9. *The shift operator θ_u is measurable.*

Proof.

$$\theta_u^{-1}(A_v) = \Big(\{t \in \mathcal{T} : v \in \theta_u(t)\} \cap A_u\Big) \cup \Big(\{t \in \mathcal{T} : v \in \theta_u(t)\} \cap A_u^c\Big).$$

□

Corollary 3.10. *If t_0 is a finite tree, then $\{t_0\}$ is measurable, so is $\mathcal{T}_f$.*

Proof.

$$\{t_0\} = \{t \in \mathcal{T} : \mathrm{tr}_{\mathrm{Height}(t_0)+1}(t) = t_0\}.$$

□

Corollary 3.11. *$t \longrightarrow \#t$ is measurable*

Proof.

$$\{t \in \mathcal{T}, \#t = n\} = \bigcup_{\substack{t' \in \mathcal{T}_f \\ \#t' = n}} \{t'\}.$$

□

3.1.3 Bienaymé–Galton–Watson trees

Definition 3.12 (Bienaymé–Galton–Watson trees)**.** Let μ be a probability measure on $\mathbb{N}_0$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let

$$\tau : \Omega \to \mathcal{T}$$

be a random tree.

We say that τ is a μ- Galton–Watson tree if the following hold:

1. The number of children of the root satisfies
$$\mathbb{P}\Big(k_\emptyset(\tau) = k\Big) = \mu(k), \qquad \text{for all } k \in \mathbb{N}_0.$$
2. For every $k \geq 1$ with $\mu(k) > 0$, conditionally on
$$\{k_\emptyset(\tau) = k\},$$
the shifted trees
$$\theta_1(\tau), \ldots, \theta_k(\tau)$$
are independent and identically distributed with the same law as τ.

Proposition 3.13 (Construction and Existence of BGW tree)**.** *Let $(C_u)_{u \in \mathcal{U}}$ be i.i.d. random variables with the law μ. We define τ consisting of $\emptyset$ and $j_1 \in \tau$ iff $j_1 \leq C_\emptyset$, for $n \geq 2$,*

$$u = j_1 j_2 \cdots j_n \in \tau$$

iff

$$j_1 \leq C_\emptyset, \quad j_i \leq C_{j_1 \cdots j_{i-1}} \quad \textit{for all } 2 \leq i \leq n.$$

. In particular,

$$k_u(\tau) = C_u \quad \textit{for all } u \in \tau.$$

then τ is a BGW tree with offspring distribution μ.

Proof. We first show that, for every fixed $\omega \in \Omega$, one has $\tau(\omega) \in \mathcal{T}$.

By construction, $\emptyset \in \tau$. Let $u \in \tau$, $u \neq \emptyset$.

- If $u = j_1$, then its parent is $\emptyset \in \tau$.
- If $u = j_1 \cdots j_n$ with $n \geq 2$, then by definition,

$$j_1 \leq C_\emptyset, \qquad j_i \leq C_{j_1 \cdots j_{i-1}} \quad \text{for all } 2 \leq i \leq n.$$

  In particular,

$$j_i \leq C_{j_1 \cdots j_{i-1}} \quad \text{for } 2 \leq i \leq n-1,$$

  which implies that the parent $j_1 \cdots j_{n-1}$ belongs to τ.

Let $u \in \tau$ and $j \in \mathbb{N}$. By definition of τ,

$$uj \in \tau \quad \Longleftrightarrow \quad j \leq C_u.$$

Hence

$$k_u(\tau) = C_u \quad \text{for all } u \in \tau.$$

We next prove that τ is measurable. For each $u \in \mathcal{U}$, consider the event

$$\{u \in \tau\} = \tau^{-1}(A_u).$$

- If $u = \emptyset$, then $\{u \in \tau\} = \Omega$.
- If $u = j_1$, then

$$\{u \in \tau\} = \{j_1 \leq C_\emptyset\}.$$

- If $u = j_1 \cdots j_n$ with $n \geq 2$, then

$$\{u \in \tau\} = \{j_1 \leq C_\emptyset\} \cap \bigcap_{i=2}^{n} \{j_i \leq C_{j_1 \cdots j_{i-1}}\}.$$

In all cases, $\{u \in \tau\}$ is measurable. Hence τ is a random tree.

We now prove that τ is a μ-BGW tree.

First, since $k_\emptyset(\tau) = C_\emptyset$ and $C_\emptyset \sim \mu$, we have

$$\mathbb{P}(k_\emptyset(\tau) = k) = \mu(k).$$

Let $k \in \mathbb{N}$ with $\mu(k) > 0$. We show that, under $\mathbb{P}(\cdot \mid k_\emptyset(\tau) = k)$, the subtrees $\theta_1(\tau), \ldots, \theta_k(\tau)$ are independent and identically distributed with the same law as τ.

Step 1: Construction map.

Define the mapping

$$\Phi : \mathbb{N}^{\mathcal{U}} \to \mathcal{T}$$

by

$$\begin{aligned}\Phi((c_u)_{u \in \mathcal{U}}) = \{\emptyset\} &\cup \{j_1 : j_1 \leq c_\emptyset\} \\ &\cup \Big\{ j_1 \cdots j_n : n \geq 2,\ j_1 \leq c_\emptyset, \\ &\qquad j_i \leq c_{j_1 \cdots j_{i-1}}\ \forall\, 2 \leq i \leq n \Big\}.\end{aligned}$$

By the arguments above, Φ is well-defined and measurable with respect to the product σ algebra, and

$$\tau = \Phi\Big((C_u)_{u\in\mathcal{U}}\Big).$$

Step 2: Subtree representation.

On the event $\{j \in \tau\}$, one has

$$\theta_j(\tau) = \Phi\Big((C_{ju})_{u\in\mathcal{U}}\Big).$$

Since $(C_u)_{u\in\mathcal{U}}$ is an i.i.d. family, we have

$$(C_{ju})_{u\in\mathcal{U}} \overset{d}{=} (C_u)_{u\in\mathcal{U}}.$$

Moreover, $(C_{ju})_{u\in\mathcal{U}}$ is independent of $C_\emptyset$.

Therefore, for any measurable $B \subset \mathcal{B}(\mathcal{T})$,

$$\mathbb{P}(\theta_j(\tau) \in B \mid C_\emptyset = k) = \mathbb{P}(\Phi((C_u)_{u\in\mathcal{U}}) \in B) = \mathbb{P}(\tau \in B),$$

which shows that each $\theta_j(\tau)$ has the same distribution as τ.

Finally, for $j_1 \neq j_2$, the families $(C_{j_1u})_{u\in\mathcal{U}}$ and $(C_{j_2u})_{u\in\mathcal{U}}$ are independent, since they correspond to disjoint coordinates of the i.i.d. family $(C_u)_{u\in\mathcal{U}}$. Consequently, the shifted trees $\theta_1(\tau), \dots, \theta_k(\tau)$ are independent.

Hence τ is a μ-BGW tree. □

Therefore, for any μ, there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a random tree $\tau : \Omega \to \mathcal{T}$ such that τ is a μ-Galton–Watson tree.

Using the explicit construction, we can easily compute the following probability.

Proposition 3.14. *Suppose t is a finite tree. Then for a μ-Galton–Watson tree τ,*

$$\mathbb{P}(\tau = t) = \prod_{u\in t} \mu\Big(k_u(t)\Big).$$

Proof.

$$\{\tau = t\} = \bigcap_{u\in t} \{C_u = k_u(t)\}.$$

Since $(C_u)_{u\in\mathcal{U}}$ are i.i.d. with law μ, we obtain

$$\mathbb{P}(\tau = t) = \mathbb{P}\Big(\bigcap_{u\in t} \{C_u = k_u(t)\}\Big) = \prod_{u\in t} \mathbb{P}(C_u = k_u(t)) = \prod_{u\in t} \mu\Big(k_u(t)\Big).$$

□

3.2 Łukasiewicz walk and direct applications

In this section we will encode BGW trees via one-dimensional walks. This will enable us to derive properties of BGW trees from our previous study of one-dimensional random walks.

3.2.1 Łukasiewicz Walk

The lexicographical order $\preceq$ on $\mathcal{U}$ within the same level is defined as usual: for $n \geq 1$, for $u = i_1 \dots i_n$ and $v = j_1 \dots j_n$, we set

$$u \preceq v \quad \text{if and only if} \quad i_n \leq j_n.$$

The breadth-first order $\trianglelefteq$ on $\mathcal{U}$ is defined by

$$u \trianglelefteq v \quad \text{if and only if} \quad |u| < |v| \quad \text{or} \quad \Big(|u| = |v| \text{ and } u \preceq v\Big).$$

Definition 3.15 (Łukasiewicz walk)**.** Let t be a locally finite rooted tree, and let $(u_1, u_2, \dots)$ be its vertices listed in breadth-first order. The *Łukasiewicz walk* $W(t) = (W_n(t))_{n\geq 0}$ associated with t is defined by

$$W_0(t) = 0,$$

and for $1 \leq n \leq \#t$,

$$W_n(t) = W_{n-1}(t) + k_{u_n}(t) - 1,$$

where $k_{u_n}(t)$ denotes the number of children of the vertex u_n.

We record the following basic properties of the Łukasiewicz walk.

Finite trees. If $\#t < \infty$, then

1. $W_i(t) \geq 0$, for all $0 \leq i \leq \#t - 1$,
2. $W_{\#t}(t) = -1$,
3. $W_i(t) - W_{i-1}(t) \geq -1$, for all $1 \leq i \leq \#t$.

Infinite trees. If $\#t = \infty$, then

1. $W_i(t) \geq 0$, for all $i \geq 0$,
2. $W_i(t) - W_{i-1}(t) \geq -1$, for all $i \geq 1$.

Indeed, we can have a better result.

Proposition 3.16. *Recall $\mathcal{T}_f$ be the set of all finite plane trees. Let W be the set of all finite paths*

$$(w_0, w_1, \dots, w_n), \quad n \in \mathbb{N},$$

such that

$$w_0 = 0, \quad w_n = -1$$

and for all $1 \leq i \leq n$,

$$w_i - w_{i-1} \geq -1 \quad \textit{and} \quad w_i \geq 0.$$

Then the Łukasiewicz path defines a bijective between $\mathcal{T}_f$ and W.

Proof. If we have a plane tree $\tau \in \mathcal{T}_f$, then we only show 1, the other is similar.

Suppose that there exists $0 < n < \#\tau$ such that

$$W_n = \sum_{i=1}^{n} k_{u_i} - n < 0.$$

Since $n < \#\tau$, the vertex u_{n+1} exists. Consider $(u_1, \dots, u_n, u_{n+1})$.

By the exploration order of the tree, for each $2 \le i \le n+1$, the vertex u_i is a child of some vertex in $\{u_1, \dots, u_n\}$. Hence the total number of children of $u_1, \dots, u_n$ is at least n, that is,

$$\sum_{i=1}^{n} k_{u_i} \ge n.$$

This contradicts $W_n < 0$. Therefore $W_n \ge 0$ for all $0 < n < \#\tau$.

Conversely, let $(w_0, \dots, w_n) \in W$. We construct a plane tree as follows. Set the root u_1 and define recursively, for $1 \le i \le n$, the number of children of u_i by

$$k_{u_i} := w_i - w_{i-1} + 1 \in \mathbb{N}.$$

Since $w_i - w_{i-1} \ge -1$, this is well-defined.

We then construct the tree in breadth-first order: when visiting u_i, we attach k_{u_i} new vertices as its children and enqueue them in this order. The condition $w_i \ge 0$ for $i < n$ guarantees that at each step there is at least one active vertex to explore, and the condition $w_n = -1$ ensures that the exploration ends exactly after n steps, producing a finite tree with $\#\tau = n$ vertices.

Moreover,

$$\sum_{i=1}^{n} (k_{u_i} - 1) = w_n - w_0 = -1,$$

so the resulting structure is indeed a tree.

Finally, the construction is unique since the sequence (w_i) uniquely determines (k_{u_i}), and the breadth-first ordering fixes the plane structure. Hence this defines an inverse map to the Łukasiewicz coding.

We conclude that the Łukasiewicz path defines a bijection between $\mathcal{T}_f$ and W. □

3.2.2 Tree-Walk correspondence

Proposition 3.17. *Let τ be a μ-Galton–Watson tree with law $\mathbb{P}_\tau$, assume $0 < \mu(0) \le \mu(0) + \mu(1) < 1$, and let $(S_n)_{n\ge 0}$ be a random walk with i.i.d. increments such that*

$$\mathbb{P}(X_1 = k) = \mu_{k+1}, \quad \text{for } k \ge -1.$$

Let T_{-1} be the first hitting time of -1 by the walk S. Then for every $n \ge 0$ and every $(x_0, \dots, x_n) \in \mathbb{Z}^{n+1}$, one has

$$\mathbb{P}_\tau\Big(\#t = n,\ W_0 = x_0, \dots, W_n = x_n\Big) = \mathbb{P}\Big(T_{-1} = n,\ S_0 = x_0, \dots, S_n = x_n\Big).$$

In particular, $\#\tau$ and T_{-1} have the same law.

Proof. Note $S_0 = 0$, $(S_n)_{n\geq 0}$ is skip-free. On the right-hand side, the condition $T_{-1} = n$ restricts to valid paths, i.e.

$$i_0 = 0, \quad i_n = -1, \quad i_k \geq 0 \ (0 \leq k < n), \quad i_k - i_{k-1} \geq -1 \ (1 \leq k \leq n).$$

For such paths,

$$\mathbb{P}(T_{-1} = n, S_0 = i_0, \ldots, S_n = i_n) = \prod_{k=1}^{n} \mu(i_k - i_{k-1} + 1),$$

and it is 0 otherwise.

On the left-hand side, for such paths, by proposition 3.16, it uniquely determines a plane tree t_0.

$$\mathbb{P}_\tau\Big(\#t = n,\ W_0 = i_0, \ldots, W_n = i_n\Big) = \mathbb{P}_\tau(\{t_0\}) = \prod_{k=1}^{n} \mu(i_k - i_{k-1} + 1).$$

and it is 0 otherwise.

Since

$$\mathbb{P}'(\#\tau = n) = \mathbb{P}_\tau(\#t = n)$$

An immediate corollary is that $\#\tau$ and T_{-1} have the same law by summing all possible paths $(x_0, \ldots, x_n) \in \mathbb{Z}^{n+1}$. □

Remark 3.18. An alternative proof using the explicit construction of BGW tree can be found in [3].

Theorem 3.19 (Extinction probability)**.** *With the same notation, and let m denotes the mean of μ. The probability that τ is finite is given by the smallest solution $\alpha \in [0, 1]$ to the equation:*

$$\alpha = \sum_{k\geq 0} \mu(k)\alpha^k$$

In particular, it equals to 1 if $m \leq 1$

Proof. By the previous correspondence,

$$\mathbb{P}'(\#\tau = \infty) = \mathbb{P}(T_{-1} = \infty) = \mathbb{P}(S_n \geq 0, \forall n \geq 0) := 1 - \alpha$$

By Theorem 2.15, we obtain

$$\alpha = \sum_{k\geq 0} \mu(k)\alpha^k.$$

and is the smallest solution to it.

□

Remark 3.20. For the degenerate cases:

- If $\mu(0) = 0$, then every individual has at least one child almost surely. Hence no extinction is possible, and

$$\mathbb{P}(\tau < \infty) = 0.$$

- If $\mu(0) + \mu(1) = 1$ with $\mu(0) = p \in (0, 1]$, then each individual has either 0 or 1 child. Thus

$$\mathbb{P}(\tau < \infty) = 1.$$

3.3 Probabilistic counting of trees

This section illustrates how BGW trees can be used to derive classical counting formulas for trees. In particular, we recover Catalan's counting formula for plane trees and Cayley's formula for labeled trees through probabilistic arguments involving BGW trees and random walks.

3.3.1 Uniform geometric BGW plane trees

We denote by $\mathcal{T}_n$ the set of all plane trees with n edges.

Proposition 3.21. *Let τ be a BGW tree with geometric offspring distribution of parameter $\frac{1}{2}$, i.e. $\mu_k = 2^{-(k+1)}$ for $k \geq 0$. Then the conditional law of τ given that it has n edges is uniform over $\mathcal{T}_n$.*

Proof. Let $t \in \mathcal{T}_n$ be a plane tree with n edges. Then

$$\mathbb{P}(\tau = t) = \prod_{u \in t} \mu_{k_u(t)} = \prod_{u \in t} 2^{-(k_u(t)+1)} = 2^{-\sum_{u \in t}(k_u(t)+1)}.$$

Note that

$$\sum_{u \in t} k_u(t) = \#t - 1.$$

Since t has n edges, $\#t = n + 1$, so

$$\sum_{u \in t} k_u(t) = n.$$

Hence

$$\mathbb{P}(\tau = t) = 2^{-(2n+1)}.$$

In particular, this probability does not depend on the choice of $t \in \mathcal{T}_n$. Therefore, the conditional law of τ given that it has n edges is uniform:

$$\mathbb{P}(\tau = t \mid \tau \in \mathcal{T}_n) = \frac{1}{|\mathcal{T}_n|}, \qquad \forall\, t \in \mathcal{T}_n.$$

□

Corollary 3.22 (Catalan's counting)**.** *For $n \in \{0, 1, 2, \ldots\}$, we have*

$$\#\{\textit{plane trees with } n \textit{ edges}\} = \#\{\textit{plane trees with } n+1 \textit{ vertices}\} = \frac{1}{n+1}\binom{2n}{n}.$$

Proof. By Kemperman's formula and the , we have

$$\mathbb{P}(\#\tau = n + 1) = \mathbb{P}(T_{-1} = n + 1) = \frac{1}{n+1}\,\mathbb{P}(S_{n+1} = -1),$$

where (S_n) is the random walk whose increments are distributed as

$$\mathbb{P}(S_1 = k) = 2^{-k-2}, \quad k \geq -1,$$

or equivalently as $G - 1$, where G is the geometric distribution of parameter $\frac{1}{2}$.

Recall that the sum of n i.i.d. geometric random variables Geo(p) follows a negative binomial distribution Binneg(n, p). Hence

$$\mathbb{P}(S_{n+1} = -1) = \mathbb{P}(\text{Binneg}(n+1, \frac{1}{2}) = n).$$

Therefore,

$$\frac{1}{n+1}\,\mathbb{P}(S_{n+1} = -1) = \frac{1}{n+1}\,\mathbb{P}(\text{Binneg}(n+1, \frac{1}{2}) = n) = \frac{1}{2}\,4^{-n}\,\frac{1}{n+1}\binom{2n}{n}.$$

On the other hand, by the previous proposition (and its proof), we have

$$\mathbb{P}(\#\tau = n+1) = \#\{\text{plane trees with } n+1 \text{ vertices}\} \cdot \frac{1}{2}\,4^{-n}.$$

Comparing the two expressions yields

$$\#\{\text{plane trees with } n+1 \text{ vertices}\} = \frac{1}{n+1}\binom{2n}{n}.$$

□

3.3.2 Cayley and Poisson BGW trees

Theorem 3.23 (Counting rooted labeled trees)**.** *The number of rooted labeled trees on* $[n]$ *is* n^{n-1}*.*

Proof. Let τ be a BGW tree with Poisson(1) offspring distribution, i.e. $\mu_k = \frac{e^{-1}}{k!}$, $k \geq 0$.

Fix a plane tree t with n vertices and denote by $k_u(t)$ the number of children of u. Then

$$\mathbb{P}(\tau = t) = \prod_{u \in t} \frac{e^{-1}}{k_u(t)!} = e^{-n} \prod_{u \in t} \frac{1}{k_u(t)!}.$$

By Kemperman's formula,

$$\mathbb{P}(\#\tau = n) = \frac{1}{n}\mathbb{P}(S_n = n-1) = \frac{1}{n}e^{-n}\frac{n^{n-1}}{(n-1)!},$$

where $S_n \sim \text{Poisson}(n)$.

Hence the conditional distribution of τ given $\#\tau = n$ is

$$\mathbb{P}(\tau = t \mid \#\tau = n) = C_n \prod_{u \in t} \frac{1}{k_u(t)!}, \qquad C_n = \frac{n(n-1)!}{n^{n-1}}.$$

Since,

$$\sum_{t:\#t=n} \mathbb{P}(\tau = t \mid \#\tau = n) = 1,$$

we have

$$\sum_{t:\#t=n} \prod_{u \in t} \frac{1}{k_u(t)!} = \frac{1}{C_n} = \frac{n^{n-1}}{n(n-1)!}.$$

For a fixed plane tree t, the number of rooted labeled trees whose underlying plane shape is t is

$$\frac{n!}{\prod_{u\in t} k_u(t)!}.$$

Indeed, labeling the vertices gives $n!$ possibilities, and permuting the $k_u(t)$ children at each vertex does not change the resulting labeled tree.

Therefore,

$$\sum_{t:\#t=n} \frac{n!}{\prod_{u\in t} k_u(t)!} = n! \sum_{t:\#t=n} \prod_{u\in t} \frac{1}{k_u(t)!} = n! \cdot \frac{n^{n-1}}{n(n-1)!} = n^{n-1}.$$

This sum counts all rooted labeled trees on $[n]$, hence the result follows. □

Corollary 3.24 (Cayley formula)**.** *The number of labeled trees on* $[n]$ *is*

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

Proof. Each unrooted labeled tree corresponds to exactly n rooted labeled trees (by choosing a root). Hence dividing by n yields Cayley's formula. □

Chapter 4

Erdős–Rényi Random Graph

4.1 Erdős–Rényi Random Graph

Definition 4.1 (Erdős–Rényi random graph)**.** Fix a $n \geq 1$, Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space supporting a family of independent random variables

$$\{U_{i,j} : 1 \leq i < j \leq n\},$$

each uniformly distributed on $[0, 1]$. (Such a space exists; for instance, one may take $\Omega = [0, 1]^{\mathbb{N}}$ endowed with the product Lebesgue measure.)

For fixed $p \in [0, 1]$, the *Erdős–Rényi random graph* $G(n, p)$ is defined as the $\mathcal{G}_n$-valued random variable given by

$$\{i, j\} \in E(G(n, p)(\omega)) \quad \Longleftrightarrow \quad U_{i,j}(\omega) \leq p,$$

for each $\omega \in \Omega$, where $\mathcal{G}_n$ denotes the set of all simple graphs on the vertex set $[n] := \{1, \ldots, n\}$.

In this way, for fixed n, the family $\{G(n, p) : p \in [0, 1]\}$ can be realized on a common probability space. Moreover, if $p_1 \leq p_2$, then

$$G(n, p_1)(\omega) \subseteq G(n, p_2)(\omega) \quad \text{for all } \omega \in \Omega.$$

Here, the notation $\subseteq$ means inclusion as graphs.

Remark 4.2. The random graph $G(n, p)$ is a random variable taking values in $\mathcal{G}_n$. Its distribution is given by

$$\mathbb{P}\Big(G(n, p) = \mathfrak{g}\Big) = p^m (1 - p)^{\binom{n}{2} - m},$$

for any $\mathfrak{g} \in \mathcal{G}_n$, where $m = |E(\mathfrak{g})|$.

Definition 4.3 (Graph property)**.** A (graph) property is a subset $A_n \subseteq \mathcal{G}_n$. The property A_n is said to be *increasing* if for any $\mathfrak{g}, \mathfrak{g}' \in \mathcal{G}_n$ with $\mathfrak{g} \subseteq \mathfrak{g}'$, one has

$$\mathfrak{g} \in A_n \implies \mathfrak{g}' \in A_n.$$

We say that the property A_n holds for $G(n, p_n)$ *with high probability* (w.h.p.) if

$$\mathbb{P}\Big(G(n, p_n) \in A_n\Big) \to 1 \quad \text{as } n \to \infty.$$

Definition 4.4 (Sharp threshold)**.** Let $(A_n)_{n \geq 1}$ be a sequence of increasing graph properties with $A_n \subseteq \mathcal{G}_n$. We say that (A_n) exhibits a *sharp threshold* at $p = p_n$ if for every $\varepsilon > 0$,

$$\mathbb{P}\Big(G(n, (1 - \varepsilon)p_n) \in A_n\Big) \to 0 \quad \text{and} \quad \mathbb{P}\Big(G(n, (1 + \varepsilon)p_n) \in A_n\Big) \to 1 \quad \text{as } n \to \infty.$$

4.2 Local property

4.2.1 Moment methods

Lemma 4.5 (First moment method)**.** *Let X be a non-negative integer-valued random variable. Then we have*

$$\mathbb{P}(X > 0) \leq \mathbb{E}[X].$$

Proof. Since X is non-negative, we can apply Markov's inequality:

$$\mathbb{P}(X > 0) = \mathbb{P}(X \geq 1) \leq \frac{\mathbb{E}[X]}{1} = \mathbb{E}[X].$$

□

Lemma 4.6 (Second moment method)**.** *Let $X \in L^2$ be a non-negative random variable and $\mathbb{E}[X^2] \neq 0$. Then*

$$\mathbb{P}(X > 0) \geq \frac{\Big(\mathbb{E}[X]\Big)^2}{\mathbb{E}[X^2]}.$$

Proof. Since $X \geq 0$, we can write

$$\mathbb{E}[X] = \mathbb{E}\Big[X\mathbf{1}_{\{X>0\}}\Big].$$

By the Cauchy–Schwarz inequality,

$$\mathbb{E}[X]^2 = \Big(\mathbb{E}[X\mathbf{1}_{\{X>0\}}]\Big)^2 \leq \mathbb{E}[X^2] \cdot \mathbb{E}[\mathbf{1}^2_{\{X>0\}}].$$

Since $\mathbf{1}^2_{\{X>0\}} = \mathbf{1}_{\{X>0\}}$, this becomes

$$\mathbb{E}[X]^2 \leq \mathbb{E}[X^2] \cdot \mathbb{P}(X > 0).$$

Rearranging gives

$$\mathbb{P}(X > 0) \geq \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X^2]}.$$

□

4.2.2 Isolated vertices

Let

$$\mathrm{NoIsol}_n = \Big\{\mathfrak{g} \in \mathcal{G}_n : \deg_{\mathfrak{g}}(i) > 0 \text{ for all } 1 \leq i \leq n\Big\},$$

$$\mathrm{Connected}_n = \{\mathfrak{g} \in \mathcal{G}_n : \mathfrak{g} \text{ is connected}\}.$$

Note that

$$\mathrm{Connected}_n \subseteq \mathrm{NoIsol}_n \quad \text{for all } n \geq 1.$$

Proposition 4.7. *The sequence $(\mathrm{NoIsol}_n)_{n\geq 1}$ has a sharp threshold at*

$$p_n = \frac{\log n}{n}.$$

Proof. We use the method of first and second moments. Since the degree of any fixed vertex in $G(n,p)$ follows a $\mathrm{Bin}(n-1,p)$ distribution, let $X(n,p)$ denote the number of isolated vertices in $G(n,p)$. Then

$$\mathbb{P}\Big(G(n,p) \text{ has an isolated vertex}\Big) = \mathbb{P}(X(n,p)>0) \le \mathbb{E}[X(n,p)] = n(1-p)^{n-1}.$$

Let $p_n = \frac{\log n}{n}$. For all $\varepsilon > 0$, we have

$$\mathbb{P}\Big(G(n,(1+\varepsilon)p_n) \in \mathrm{NoIsol}_n\Big) = 1-\mathbb{P}(X(n,(1+\varepsilon)p_n)>0) \ge 1-n\Big(1-(1+\varepsilon)p_n\Big)^{n-1},$$

which goes to 1 when n goes to ∞.

$$\begin{aligned}\mathbb{E}[X(n,p)^2] &= \sum_{1\le i,j\le n} \mathbb{P}\Big(i \text{ and } j \text{ are isolated in } G(n,p)\Big)\\ &= n\,\mathbb{P}(1 \text{ is isolated}) + n(n-1)\,\mathbb{P}(1 \text{ and } 2 \text{ are isolated})\\ &= n(1-p)^{n-1} + n(n-1)(1-p)^{2n-3}.\end{aligned}$$

Thus,

$$\begin{aligned}\mathbb{P}\Big(G(n,(1-\varepsilon)p_n) \in \mathrm{NoIsol}_n\Big) &= 1 - \mathbb{P}\Big(X(n,(1-\varepsilon)p_n) > 0\Big)\\ &\le 1 - \frac{\Big(\mathbb{E}[X(n,(1-\varepsilon)p_n)]\Big)^2}{\mathbb{E}[X(n,(1-\varepsilon)p_n)^2]}.\end{aligned}$$

Note that

$$\frac{\Big(\mathbb{E}[X(n,(1-\varepsilon)p_n)]\Big)^2}{\mathbb{E}[X(n,(1-\varepsilon)p_n)^2]} \to 1,$$

Therefore,

$$\mathbb{P}\Big(G(n,(1-\varepsilon)p_n) \in \mathrm{NoIsol}_n\Big) \to 0.$$

□

4.2.3 Connectivity

We say that a graph $\mathfrak{g}$ has the core property if it is made of a connected component of size at least 2 together with isolated vertices. We denote by Core_n the set of graphs associated with $\mathcal{G}_n$ that satisfy the core property.

Lemma 4.8 (Core property)**.** *For any*

$$\frac{2\log n}{3n} \le p_n \le \frac{2\log n}{n},$$

we have

$$G(n,p_n) \in \mathrm{Core}_n \quad \textit{with high probability}$$

Proof. This follows from Lemma 5.3 of [1]. □

Proposition 4.9. *The sequence* $(\mathrm{Connected}_n)_{n>1}$ *has a sharp threshold transition at*

$$p_n = \frac{\log n}{n}.$$

Proof. Let $p_n = \frac{\log n}{n}$. For all $\varepsilon > 0$. Since

$$\text{Connected}_n \subseteq \text{NoIsol}_n,$$

we have

$$\mathbb{P}\Big(G(n,p) \in \text{Connected}_n\Big) \leq \mathbb{P}\Big(G(n,p) \in \text{NoIsol}_n\Big).$$

From the sharp threshold for (NoIsol_n), it follows that

$$\mathbb{P}\Big(G(n,(1-\varepsilon)p_n) \in \text{Connected}_n\Big) \to 0.$$

Let $p = (1+\varepsilon)p_n = (1+\varepsilon)\frac{\log n}{n}$. We use the inclusion

$$\text{Core}_n \cap \text{NoIsol}_n \subseteq \text{Connected}_n.$$

Therefore,

$$\mathbb{P}\Big(G(n,p) \in \text{Connected}_n\Big) \geq \mathbb{P}\Big(G(n,p) \in \text{Core}_n \cap \text{NoIsol}_n\Big).$$

By the union bound,

$$\mathbb{P}\Big(G(n,p) \in \text{Core}_n \cap \text{NoIsol}_n\Big) \geq 1 - \mathbb{P}\Big(G(n,p) \notin \text{Core}_n\Big) - \mathbb{P}\Big(G(n,p) \notin \text{NoIsol}_n\Big).$$

Now, for $0 < \varepsilon \leq 1$, we have $p \leq 2\frac{\log n}{n}$, so by the Core Property Lemma,

$$\mathbb{P}\Big(G(n,p) \notin \text{Core}_n\Big) \to 0.$$

Moreover, from the sharp threshold of (NoIsol_n),

$$\mathbb{P}\Big(G(n,(1+\varepsilon)p_n) \notin \text{NoIsol}_n\Big) \to 0.$$

Thus,

$$\mathbb{P}\Big(G(n,(1+\varepsilon)p_n) \in \text{Connected}_n\Big) \to 1.$$

For $\varepsilon > 1$, the result follows from monotonicity of the property Connected_n in p.

□

4.3 Birth of giant

In this section, we prove a version of the emergence of the giant component using the sprinkling argument together with the Poisson-BGW tree approximation.

4.3.1 Poisson(c)-BGW tree

Recall that for a Poisson(c) Bienaymé–Galton–Watson tree τ, the extinction probability

$$\alpha(c) := \mathbb{P}(\tau < \infty)$$

is given by the smallest solution $\alpha \in [0,1]$ of the fixed-point equation

$$\alpha = \sum_{k \geq 0} \mu(k)\alpha^k,$$

where μ denotes the offspring distribution.

In the Poisson(c) case,

$$\mu(k) = e^{-c}\frac{c^k}{k!},$$

and therefore

$$\alpha(c) = \sum_{k\geq 0} e^{-c}\frac{c^k}{k!}\alpha(c)^k = e^{-c}\sum_{k\geq 0}\frac{(c\alpha(c))^k}{k!} = e^{-c(1-\alpha(c))}.$$

Hence the extinction probability satisfies

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

Note that if $c \leq 1$, then $\alpha(c) = 1$, whereas if $c > 1$, then $0 < \alpha(c) < 1$.

4.3.2 Cut and Sprinkling

Definition 4.10 (ε-cut)**.** Let $\mathfrak{g}$ be a graph on n vertices, and let $\varepsilon > 0$. An *ε-cut* of $\mathfrak{g}$ is a partition

$$V(\mathfrak{g}) = A \sqcup B$$

such that

$$|A| \geq \varepsilon n, \qquad |B| \geq \varepsilon n,$$

and there are no edges between A and B in $\mathfrak{g}$.

Proposition 4.11. *Suppose every connected component of $\mathfrak{g}$ has size strictly smaller than*

$$(1-2\varepsilon)n.$$

Then $\mathfrak{g}$ admits an ε-cut.

Proof. Indeed, let

$$D_1, D_2, \ldots, D_m$$

be the connected components of $\mathfrak{g}$.

Construct a subset A by adding connected components one by one until, for the first time,

$$|A| \geq \varepsilon n.$$

Before adding the last component, we had $|A| < \varepsilon n$. Since every connected component has size strictly smaller than $(1-2\varepsilon)n$, after adding the last component we obtain

$$|A| < \varepsilon n + (1-2\varepsilon)n = (1-\varepsilon)n.$$

Hence

$$\varepsilon n \leq |A| < (1-\varepsilon)n.$$

Set

$$B = V(\mathfrak{g}) \setminus A,$$

then

$$|B| \geq \varepsilon n.$$

Since A is a union of whole connected components, there is no edge between A and B. Thus (A, B) is an ε-cut. □

Let $\mathfrak{g} \in \mathcal{G}_n$ be a graph whose connected components all have size at least M, and let

$$G\left(n, \frac{\delta}{n}\right) \cup \mathfrak{g}$$

denote the graph obtained by superimposing $\mathfrak{g}$ with $G\left(n, \frac{\delta}{n}\right)$ and deleting any multiple edges. The resulting graph is again a random graph.

Lemma 4.12 (Sprinkling)**.** *Fix $\delta, \varepsilon > 0$ and $M \in \{1, 2, \ldots\}$.*

If

$$-\delta\varepsilon^2 + \frac{\log 2}{M} < 0,$$

then, with high probability (as $n \to \infty$), the graph $G\left(n, \frac{\delta}{n}\right) \cup \mathfrak{g}$ contains a giant component of size at least $(1 - 2\varepsilon)n$.

Proof. By the previous proposition, it remains to prove that with high probability there is no ε-cut.

Let

$$X_n = \#\{\varepsilon\text{-cuts in } G\left(n, \frac{\delta}{n}\right) \cup \mathfrak{g}\}.$$

Let K be the number of connected components of $\mathfrak{g}$. Since every component has size at least M, we have

$$K \leq \frac{n}{M}.$$

Moreover, an ε-cut cannot split a connected component of $\mathfrak{g}$, since otherwise there would already be an edge crossing the cut. Hence each connected component of $\mathfrak{g}$ must lie entirely in one side of the cut.

Therefore the number of possible ε-cuts is at most

$$2^K \leq 2^{n/M}.$$

Fix such a cut

$$V = A_0 \sqcup B_0$$

For this partition to remain a cut after adding the random edges of $G\left(n, \frac{\delta}{n}\right)$, there must be no edge between A_0 and B_0 in the sprinkling graph.

The number of potential edges between A_0 and B_0 is

$$|A_0||B_0| \geq \varepsilon^2 n^2.$$

Since each edge is absent with probability $1 - \frac{\delta}{n}$, we obtain

$$\mathbb{P}\Big((A_0, B_0) \text{ remains a cut}\Big) \leq \left(1 - \frac{\delta}{n}\right)^{\varepsilon^2 n^2}.$$

Hence

$$\mathbb{E}[X_n] \leq 2^{n/M} \left(1 - \frac{\delta}{n}\right)^{\varepsilon^2 n^2} \leq \exp\left(\frac{\log 2}{M} n - \delta\varepsilon^2 n\right).$$

By assumption,

$$\delta\varepsilon^2 > \frac{\log 2}{M},$$

so the exponent is negative. Therefore

$$\mathbb{E}[X_n] \to 0.$$

By Markov's inequality,

$$\mathbb{P}(X_n \geq 1) \leq \mathbb{E}[X_n] \to 0.$$

Thus, with high probability there is no ε-cut, and consequently

$$G\left(n, \frac{\delta}{n}\right) \cup \mathfrak{g}$$

contains a connected component of size at least $(1 - 2\varepsilon)n$.

□

Corollary 4.13 (Random sprinkling)**.** *Fix $\delta, \varepsilon, \rho > 0$ and $M \in \{1, 2, \ldots\}$. Assume that*

$$-\delta\rho\varepsilon^2 + \frac{\log 2}{M} < 0.$$

Let $\mathfrak{g}_n$ be a random graph on $\{1, \ldots, n\}$, and let

$$G_1 \sim G\left(n, \frac{\delta}{n}\right)$$

be independent of $\mathfrak{g}_n$.

For a deterministic graph $g \in \mathcal{G}_n$, define

$$S_g(M) = \#\{1 \leq i \leq n : \#C^i(g) > M\}.$$

Assume that

$$\mathbb{P}\Big(S_{\mathfrak{g}_n}(M) \geq \rho n\Big) \longrightarrow 1.$$

Then, with high probability, the graph $\mathfrak{g}_n \cup G_1$ contains a connected component of size at least $\rho(1 - 2\varepsilon)n$.

Proof. Let

$$\mathcal{H}_n = \{g \in \mathcal{G}_n : S_g(M) \geq \rho n\}$$

and

$$H_n = \{\mathfrak{g}_n \in \mathcal{H}_n\}.$$

By assumption,

$$\mathbb{P}(H_n) \to 1.$$

Let

$$E_n = \{\mathfrak{g}_n \cup G_1 \text{ contains a connected component of size at least } \rho(1 - 2\varepsilon)n\}.$$

We prove that $\mathbb{P}(E_n) \to 1$.

By conditioning on the first graph,

$$\mathbb{P}(E_n) \geq \sum_{g \in \mathcal{H}_n} \mathbb{P}(E_n \mid \mathfrak{g}_n = g)\mathbb{P}(\mathfrak{g}_n = g).$$

Fix $g \in \mathcal{H}_n$. Let

$$T = \{1 \le i \le n : \#C^i(g) > M\}, \qquad m = S_g(M).$$

Then

$$m \ge \rho n.$$

Moreover, every connected component of the induced graph $g[T]$ has size larger than M.

By the independence of G_1 and $\mathfrak{g}_n$, and since any connected component contained in the subgraph on T is also a connected component of the whole graph, we have

$$\begin{aligned}\mathbb{P}(E_n \mid \mathfrak{g}_n = g) &= \mathbb{P}\begin{pmatrix} g \cup G_1 \text{ contains a connected} \\ \text{component of size } \ge \rho(1-2\varepsilon)n \end{pmatrix} \\ &\ge \mathbb{P}\begin{pmatrix} g[T] \cup G_1[T] \text{ contains a connected} \\ \text{component of size } \ge \rho(1-2\varepsilon)n \end{pmatrix}.\end{aligned}$$

Since

$$G_1[T] \sim G\left(m, \frac{\delta}{n}\right)$$

and $m \ge \rho n$, we have

$$\frac{\delta\rho}{m} \le \frac{\delta}{n}.$$

Thus we may couple

$$G_1' \sim G\left(m, \frac{\delta\rho}{m}\right)$$

with $G_1[T]$ so that

$$G_1' \subset G_1[T].$$

Therefore

$$\mathbb{P}(E_n \mid \mathfrak{g}_n = g) \ge \mathbb{P}\begin{pmatrix} g[T] \cup G_1' \text{ contains a connected} \\ \text{component of size } \ge \rho(1-2\varepsilon)n \end{pmatrix}.$$

Now apply the deterministic sprinkling lemma to the deterministic graph $g[T]$, which has m vertices and whose connected components all have size at least M, with sprinkling parameter $\delta\rho$.

Let

$$a = -\left(-\delta\rho\varepsilon^2 + \frac{\log 2}{M}\right) > 0.$$

By the proof of the sprinkling lemma,

$$\mathbb{P}\left(\begin{array}{l} g[T] \cup G_1' \text{ contains a connected} \\ \text{component of size } \ge (1-2\varepsilon)m \end{array}\right) \ge 1 - e^{-am}.$$

Since

$$(1-2\varepsilon)m \ge \rho(1-2\varepsilon)n$$

and

$$e^{-am} \le e^{-a\rho n},$$

we obtain, for every $g \in \mathcal{H}_n$,

$$\mathbb{P}(E_n \mid \mathfrak{g}_n = g) \geq 1 - e^{-a\rho n}.$$

Substituting this into the previous lower bound gives

$$\begin{aligned}\mathbb{P}(E_n) &\geq \sum_{g \in \mathcal{H}_n} \mathbb{P}(E_n \mid \mathfrak{g}_n = g)\mathbb{P}(\mathfrak{g}_n = g) \\ &\geq (1 - e^{-a\rho n}) \sum_{g \in \mathcal{H}_n} \mathbb{P}(\mathfrak{g}_n = g) \\ &= (1 - e^{-a\rho n})\mathbb{P}(H_n).\end{aligned}$$

Since

$$e^{-a\rho n} \to 0 \qquad \text{and} \qquad \mathbb{P}(H_n) \to 1,$$

we conclude that

$$\mathbb{P}(E_n) \to 1.$$

□

4.3.3 Phase transition for the giant component

We denote by $C^i(n,p)$ the connected component containing the vertex i in $G(n,p)$. For each connected component, we relabel its vertices in increasing order by $1, 2, \ldots, \#C^i(n,p)$. Let $\#C_1^{\max}(n,p), \#C_2^{\max}(n,p), \ldots$ denote the size of the connected components of $G(n,p)$ arranged in non-increasing order, with ties broken according to the labels of the vertices. when the underlying graph is clear from the context. If there is only 1 connected components, we set $\#C_2^{\max}(n,p) = 0$. The main theorem of this section is the following.

Theorem 4.14. *For $c > 0$, $p_n = \frac{c}{n}$, let $\alpha(c)$ denote the smallest solution in $[0,1]$ of the equation*

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

Then

$$\frac{\#C_1^{\max}(n,p_n)}{n} \xrightarrow[n\to\infty]{\mathbb{P}} 1 - \alpha(c), \qquad \frac{\#C_2^{\max}(n,p_n)}{n} \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

The subcritical and critical cases ($c \leq 1$) follow directly from the following Poisson-BGW tree approximation.

Proposition 4.15. *Fix $c > 0$ and suppose that $p_n = \frac{c}{n}$ Then, for every $k \geq 1$ and every connected labeled graph $\mathfrak{g} \in \mathcal{G}_k$,*

$$\lim_{n\to\infty} \mathbb{P}\Big(C^1(n,p_n) = \mathfrak{g}\Big) = e^{-ck}\frac{c^{k-1}}{(k-1)!}\mathbf{1}_{\{\mathfrak{g} \text{ is a Cayley tree}\}}.$$

Proof. Let $\mathfrak{g}$ be a fixed connected labeled graph with k vertices and ℓ edges. Then

$$\mathbb{P}(C^1(n,p_n) = \mathfrak{g}) = \binom{n-1}{k-1}(1-p_n)^{k(n-k)}(1-p_n)^{\binom{k}{2}-\ell}p_n^\ell.$$

Taking limits as $n \to \infty$, the above expression converges to 0 if $\ell > k-1$, and to $e^{-ck}\frac{c^{k-1}}{(k-1)!}$ if $\ell = k-1$. □

A useful application of the previous proposition is that the connected component of a fixed vertex in $G(n, p_n)$ 'resembles' a Poisson(c)-BGW tree. More precisely, we have

Corollary 4.16 (Local approximation)**.** *For every $i \geq 1$,*

$$\#C^i(n, p_n) \xrightarrow{d} \#\tau,$$

where τ is a Poisson(c)-BGW tree.

Proof. By symmetry, we can assume $i = 1$. For every fixed $k \geq 1$, we decompose according to all connected labeled graphs $\mathfrak{g}$ with k vertices:

$$\mathbb{P}(\#C^1(n, p_n) = k) = \sum_{\mathfrak{g}:\, |\mathfrak{g}|=k} \mathbb{P}(C^1(n, p_n) = \mathfrak{g}).$$

Since there are only finitely many such graphs,

$$\lim_{n\to\infty} \mathbb{P}(\#C^1(n, p_n) = k) = \sum_{\mathfrak{g}:\, |\mathfrak{g}|=k} \lim_{n\to\infty} \mathbb{P}(C^1(n, p_n) = \mathfrak{g}).$$

By the previous proposition and Cayley's formula, we obtain

$$\lim_{n\to\infty} \mathbb{P}(\#C^1(n, p_n) = k) = k^{k-2} e^{-ck} \frac{c^{k-1}}{(k-1)!} = \mathbb{P}(\#\tau = k).$$

□

Remark 4.17. The same proof shows that Corollary 4.16 remain valid under the weaker assumption $q_n \sim p_n = \frac{c}{n}$.

Corollary 4.18 ((Sub)critical regime)**.** *If $c \leq 1$, then $G(n, p_n)$ has no giant component, i.e.*

$$\frac{\#C_1^{\max}(n, p_n)}{n} \xrightarrow{\mathbb{P}} 0.$$

Proof. We show that for every $\varepsilon > 0$,

$$\mathbb{P}(\#C_1^{\max}(n, p_n) \geq \varepsilon n) \longrightarrow 0.$$

Since $c \leq 1$, the Poisson(c)-BGW tree τ is almost surely finite. From the convergence in distribution established above,

$$\#C^1(n, p_n) \xrightarrow{d} \#\tau.$$

In particular (See A.3),

$$\mathbb{P}\Big(\#C^1(n, p_n) \geq \varepsilon n\Big) \longrightarrow 0.$$

By symmetry of the vertices,

$$\mathbb{P}\Big(\#C^1(n, p_n) \geq \varepsilon n\Big) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{P}\Big(\#C^i(n, p_n) \geq \varepsilon n\Big).$$

Now observe that

$$\{\#C_1^{\max}(n,p_n) \geq \varepsilon n,\ i \in C_1^{\max}(n,p_n)\} \subseteq \Big\{\#C^i(n,p_n) \geq \varepsilon n\Big\}.$$

Therefore,

$$\mathbb{P}\Big(\#C^1(n,p_n) \geq \varepsilon n\Big) \geq \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\mathbf{1}_{\{\#C_1^{\max}(n,p_n)\geq \varepsilon n\}}\mathbf{1}_{\{i\in C_1^{\max}(n,p_n)\}}\right].$$

On the event $\{\#C_1^{\max}(n,p_n) \geq \varepsilon n\}$, at least εn vertices belong to $C_1^{\max}(n,p_n)$. Hence

$$\sum_{i=1}^{n} \mathbf{1}_{\{i\in C_1^{\max}(n,p_n)\}} = \#C_1^{\max}(n,p_n) \geq \varepsilon n.$$

Thus,

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\mathbf{1}_{\{\#C_1^{\max}(n,p_n)\geq \varepsilon n\}}\mathbf{1}_{\{i\in C_1^{\max}(n,p_n)\}}\right] \geq \varepsilon\, \mathbb{P}(\#C_1^{\max}(n,p_n) \geq \varepsilon n).$$

Combining the inequalities yields

$$\mathbb{P}\Big(\#C^1(n,p_n) \geq \varepsilon n\Big) \geq \varepsilon\, \mathbb{P}(\#C_1^{\max}(n,p_n) \geq \varepsilon n).$$

Since the left-hand side converges to 0, we conclude that

$$\mathbb{P}(\#C_1^{\max}(n,p_n) \geq \varepsilon n) \longrightarrow 0.$$

Therefore, there is no giant component when $c \leq 1$. □

The supercritical case relies on a sprinkling argument. We first state some useful observations.

Lemma 4.19. *Conditionally on $C^1(n,p)$, the relabeled remaining graph*

$$G(n,p) \setminus C^1(n,p)$$

has distribution

$$G\Big(n - \#C^1(n,p),\ p\Big).$$

Proof. Let $n \geq 1$, $1 \leq k \leq n$, $\mathfrak{h}$ be a connected labeled graph with k vertices, and let $\mathfrak{g}$ be a labeled graph with $n-k$ vertices.

We compute

$$\mathbb{P}\Big(G(n,p) \setminus C^1(n,p) = \mathfrak{g} \mid C^1(n,p) = \mathfrak{h}\Big).$$

The denominator was computed above. We now compute the numerator. In addition to the event

$$\{C^1(n,p) = \mathfrak{h}\},$$

we further require that the remaining graph, after relabeling, is equal to $\mathfrak{g}$. Therefore,

$$\mathbb{P}\Big(C^1(n,p) = \mathfrak{h},\ G(n,p) \setminus C^1(n,p) = \mathfrak{g}\Big)$$

equals

$$\binom{n-1}{k-1} p^{|E(\mathfrak{h})|}(1-p)^{\binom{k}{2}-|E(\mathfrak{h})|}(1-p)^{k(n-k)} \cdot p^{|E(\mathfrak{g})|}(1-p)^{\binom{n-k}{2}-|E(\mathfrak{g})|}.$$

Dividing the two expressions, we obtain

$$\mathbb{P}\Big(G(n,p) \setminus C^1(n,p) = \mathfrak{g} \mid C^1(n,p) = \mathfrak{h}\Big) = p^{|E(\mathfrak{g})|}(1-p)^{\binom{n-k}{2}-|E(\mathfrak{g})|}.$$

But this is exactly the probability that

$$G(n-k,p) = \mathfrak{g}.$$

Hence the relabeled remaining graph has distribution

$$G\Big(n - \#C^1(n,p),\, p\Big).$$

□

We can establish the asymptotic behaviour of the small components.

Corollary 4.20 (Global approximation)**.** *Fix $A \in \{1,2,\ldots\}$ and define*

$$N_n(A) = \#\left\{1 \le i \le n : \#C^i(n,p_n) \le A\right\}.$$

Then

$$\frac{N_n(A)}{n} \xrightarrow{\mathbb{P}} \mathbb{P}(\#\tau \le A) \xrightarrow[A\to\infty]{} \alpha(c) = \mathbb{P}(\#\tau < \infty)$$

where τ is a Poisson(c)-BGW tree.

Proof. We first compute the expectation. By symmetry of the vertices,

$$\mathbb{E}[N_n(A)] = \sum_{i=1}^{n} \mathbb{P}(\#C^i(n,p_n) \le A) = n\,\mathbb{P}(\#C^1(n,p_n) \le A).$$

From Corollary 4.16,

$$\#C^1(n,p_n) \xrightarrow{d} \#\tau,$$

hence

$$\frac{1}{n}\mathbb{E}[N_n(A)] \longrightarrow \mathbb{P}(\#\tau \le A).$$

To prove convergence in probability, it remains to control the fluctuations. We compute the second moment:

$$\mathbb{E}[N_n(A)^2] = \sum_{1\le i,j\le n} \mathbb{P}(\#C^i(n,p_n) \le A,\ \#C^j(n,p_n) \le A).$$

We split the sum according to whether i and j belong to the same connected component:

$$\mathbb{E}[N_n(A)^2] = \sum_{i,j} \mathbb{P}(\#C^i \le A,\ i \sim j) + \sum_{i,j} \mathbb{P}(\#C^i \le A,\ \#C^j \le A,\ i \not\sim j).$$

By symmetry, the second term equals to

$$n(n-1)\mathbb{P}(\#C^1 \le A,\ \#C^2 \le A,\ 1 \not\sim 2).$$

Decomposing according to all connected labeled graphs $\mathfrak{h}$ with $|\mathfrak{h}| \le A$,

$$\begin{aligned}
&\mathbb{P}(\#C^1 \le A,\ \#C^2 \le A,\ 1 \not\sim 2) \\
&= \sum_{|\mathfrak{h}|\le A} \mathbb{P}(C^1(n,p_n) = \mathfrak{h},\ \#C^2 \le A,\ 1 \not\sim 2) \\
&= \sum_{|\mathfrak{h}|\le A} \mathbb{P}(C^1(n,p_n) = \mathfrak{h})\, \mathbb{P}(\#C^2 \le A,\ 1 \not\sim 2 \mid C^1(n,p_n) = \mathfrak{h}).
\end{aligned}$$

Since we only need an upper bound, we may drop the condition $1 \not\sim 2$. By Lemma 4.19, we have

$$\mathbb{P}(\#C^2 \le A,\ 1 \not\sim 2 \mid C^1(n,p_n) = \mathfrak{h}) \le \mathbb{P}(\#C^1(n - |\mathfrak{h}|, p_n) \le A).$$

Hence

$$\begin{aligned}
&\mathbb{P}(\#C^1 \le A,\ \#C^2 \le A,\ 1 \not\sim 2) \\
&\le \sum_{|\mathfrak{h}|\le A} \mathbb{P}(C^1(n,p_n) = \mathfrak{h})\, \mathbb{P}(\#C^1(n - |\mathfrak{h}|, p_n) \le A).
\end{aligned}$$

By the previous remark, for each fixed $k \le A$,

$$\mathbb{P}(\#C^1(n-k,p_n) \le A) \to \mathbb{P}(\#\tau \le A).$$

Therefore,

$$\begin{aligned}
&\sum_{|\mathfrak{h}|\le A} \mathbb{P}(C^1(n,p_n) = \mathfrak{h})\mathbb{P}(\#C^1(n - |\mathfrak{h}|, p_n) \le A) \\
&= \sum_{k=1}^{A} \mathbb{P}(\#C^1(n,p_n) = k)\mathbb{P}(\#C^1(n-k,p_n) \le A).
\end{aligned}$$

Since the sum over k is finite, we may pass to the limit term by term:

$$\begin{aligned}
&\lim_{n\to\infty} \sum_{k=1}^{A} \mathbb{P}(\#C^1(n,p_n) = k)\mathbb{P}(\#C^1(n-k,p_n) \le A) \\
&= \sum_{k=1}^{A} \mathbb{P}(\#\tau = k)\mathbb{P}(\#\tau \le A) \\
&= \mathbb{P}(\#\tau \le A)^2.
\end{aligned}$$

Thus

$$\mathrm{Var}\left(\frac{N_n(A)}{n}\right) \to 0.$$

By Chebyshev's inequality, it follows that

$$\frac{N_n(A)}{n} \xrightarrow{\mathbb{P}} \mathbb{P}(\#\tau \le A).$$

□

Another useful observation is that for $c, \delta > 0$ and n sufficiently large, the superimposition of two independent random graphs $G\left(n, \frac{c}{n}\right)$ and $G\left(n, \frac{\delta}{n}\right)$ is again a random graph and

$$G\left(n, \frac{c}{n}\right) \cup G\left(n, \frac{\delta}{n}\right) \overset{(d)}{=} G\left(n, 1 - \left(1 - \frac{c}{n}\right)\left(1 - \frac{\delta}{n}\right)\right) \subset G\left(n, \frac{c+2\delta}{n}\right)$$

We have now introduced all the ingredients needed to prove the supercritical regime of birth of the giant component.

Theorem 4.21. *For $c > 0$, $p_n = \frac{c}{n}$, let $\alpha(c)$ denote the smallest solution in $[0,1]$ of the equation*

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

Then

$$\frac{\#C_1^{\max}(n,p_n)}{n} \xrightarrow[n\to\infty]{\mathbb{P}} 1-\alpha(c), \qquad \frac{\#C_2^{\max}(n,p_n)}{n} \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Proof. The case $c \le 1$ was proved in Corollary 4.18. We assume from now on that $c > 1$.

Step 1: Upper bound.

We first prove the upper bound. Fix $\varepsilon > 0$. For $A \ge 1$, define

$$S_n(A) = \#\left\{1 \le i \le n : \#C^i(n,p_n) > A\right\}.$$

By Corollary 4.20,

$$\frac{S_n(A)}{n} = 1 - \frac{N_n(A)}{n} \xrightarrow{\mathbb{P}} 1 - \mathbb{P}(\#\tau \le A) = \mathbb{P}(\#\tau > A).$$

Since

$$\mathbb{P}(\#\tau > A) \xrightarrow[A\to\infty]{} \mathbb{P}(\#\tau = \infty) = 1-\alpha(c),$$

we may choose A sufficiently large such that

$$\mathbb{P}(\#\tau > A) < 1 - \alpha(c) + \frac{\varepsilon}{2}.$$

we have

$$\mathbb{P}\left(\frac{S_n(A)}{n} > \mathbb{P}(\#\tau > A) + \frac{\varepsilon}{2}\right) \longrightarrow 0.$$

Therefore

$$\mathbb{P}\left(\frac{S_n(A)}{n} > 1 - \alpha(c) + \varepsilon\right) \longrightarrow 0.$$

Now observe that

$$\#C_1^{\max}(n,p_n) \le S_n(A) + A.$$

Hence, for n large enough so that $A/n < \varepsilon$,

$$\left\{\frac{\#C_1^{\max}(n,p_n)}{n} > 1 - \alpha(c) + 2\varepsilon\right\} \subset \left\{\frac{S_n(A)}{n} > 1 - \alpha(c) + \varepsilon\right\}.$$

It follows that

$$\mathbb{P}\left(\frac{\#C_1^{\max}(n,p_n)}{n} > 1 - \alpha(c) + 2\varepsilon\right) \longrightarrow 0.$$

Since $\varepsilon > 0$ is arbitrary, this proves

$$\mathbb{P}\left(\frac{\#C_1^{\max}(n,p_n)}{n} > 1 - \alpha(c) + \varepsilon\right) \longrightarrow 0$$

for every $\varepsilon > 0$.

Step 2: Lower bound.

For the lower bound, let $\eta > 0$. If $\eta \geq 1 - \alpha(c)$, the statement is trivial. Hence we may assume that $\eta < 1 - \alpha(c)$. Choose $\delta > 0$ sufficiently small such that

$$c_0 := c - 2\delta > 1$$

and

$$1 - \alpha(c_0) > 1 - \alpha(c) - \frac{\eta}{3}.$$

Set

$$\rho := 1 - \alpha(c_0) - \frac{\eta}{3}.$$

Then

$$\rho > 1 - \alpha(c) - \frac{2\eta}{3}.$$

Choose $\gamma > 0$ sufficiently small such that

$$\rho(1 - 2\gamma) > 1 - \alpha(c) - \eta.$$

We may choose M sufficiently large such that

$$-\delta\rho\gamma^2 + \frac{\log 2}{M} < 0.$$

Let

$$G_0 \sim G\left(n, \frac{c_0}{n}\right), \qquad G_1 \sim G\left(n, \frac{\delta}{n}\right)$$

be independent. Define

$$S_n^{(0)}(M) = \#\left\{1 \leq i \leq n : \#C^i(G_0) > M\right\}.$$

By Corollary 4.20, applied with parameter c_0, we have

$$\mathbb{P}\left(S_n^{(0)}(M) \geq \rho n\right) \longrightarrow 1.$$

Therefore Corollary 4.13, applied to the random graph $\mathfrak{g}_n = G_0$ with parameters δ, γ, ρ and M, implies that, with high probability, the graph

$$G_0 \cup G_1$$

contains a connected component of size at least

$$\rho(1 - 2\gamma)n.$$

By the choice of γ,

$$\rho(1 - 2\gamma) > 1 - \alpha(c) - \eta.$$

Hence

$$\mathbb{P}\left(\#C_1^{\max}(G_0 \cup G_1) \geq (1 - \alpha(c) - \eta)n\right) \longrightarrow 1.$$

Moreover,

$$G_0 \cup G_1 \subset G\left(n, \frac{c}{n}\right).$$

Therefore

$$\#C_1^{\max}\left(n, \frac{c}{n}\right) \geq \#C_1^{\max}(G_0 \cup G_1),$$

and consequently

$$\mathbb{P}\left(\frac{\#C_1^{\max}(n,c/n)}{n} < 1-\alpha(c)-\eta\right) \longrightarrow 0.$$

Together with the upper bound already proved, this yields

$$\frac{\#C_1^{\max}(n,p_n)}{n} \xrightarrow{\mathbb{P}} 1-\alpha(c).$$

Step 3: Unique cluster.

It remains to prove that

$$\frac{\#C_2^{\max}(n,p_n)}{n} \xrightarrow{\mathbb{P}} 0.$$

Fix $\eta > 0$ and choose $0 < \varepsilon < \frac{\eta}{3}$. By Corollary 4.20, one can choose A sufficiently large so that

$$\mathbb{P}(\#\tau > A) < 1-\alpha(c)+\varepsilon.$$

and therefore

$$\mathbb{P}(S_n(A) \le (1-\alpha(c)+2\varepsilon)n) \longrightarrow 1.$$

On the other hand, we have already proved that

$$\frac{\#C_1^{\max}(n,p_n)}{n} \xrightarrow{\mathbb{P}} 1-\alpha(c),$$

hence

$$\mathbb{P}(\#C_1^{\max}(n,p_n) \ge (1-\alpha(c)-\varepsilon)n) \longrightarrow 1.$$

For n sufficiently large, we have $\eta n > A$, then

$$\{\#C_2^{\max}(n,p_n) > \eta n\} \subset \{S_n(A) \ge \#C_1^{\max}(n,p_n) + \#C_2^{\max}(n,p_n)\}.$$

Hence,

$$\begin{aligned}\{\#C_2^{\max}(n,p_n) > \eta n\} \subset\ & \{\#C_1^{\max}(n,p_n) < (1-\alpha(c)-\varepsilon)n\} \\ & \cup \{S_n(A) > (1-\alpha(c)+2\varepsilon)n\}.\end{aligned}$$

As $\eta > 0$ was arbitrary,

$$\frac{\#C_2^{\max}(n,p_n)}{n} \xrightarrow{\mathbb{P}} 0.$$

□

Chapter 5

Random Permutations and Recursive Trees

5.1 Random permutation

In this section, we study the law of the cycle decomposition of a random permutation σ_n chosen uniformly in the symmetric group $\mathfrak{S}_n$ over n elements $\{1, 2, \ldots, n\}$.

In particular, we shall establish Poisson statistics for the number of short cycles and the Poisson–Dirichlet limit for the large cycles.

5.1.1 Foata correspondence

A permutation $\sigma_n \in \mathfrak{S}_n$ can be represented by its cycle decomposition. The *Foata encoding* of σ_n is defined by ordering the cycles according to their minimal elements and then removing the parentheses. More precisely, write

$$\sigma_n = (a_1^{(1)}, \ldots, a_{k_1}^{(1)})(a_1^{(2)}, \ldots, a_{k_2}^{(2)}) \cdots (a_1^{(\ell)}, \ldots, a_{k_\ell}^{(\ell)}),$$

where the cycles are disjoint and arranged so that

$$a_{k_1}^{(1)} = \min_{1 \le i \le k_1} a_i^{(1)} < a_{k_2}^{(2)} = \min_{1 \le i \le k_2} a_i^{(2)} < \cdots < a_{k_\ell}^{(\ell)} = \min_{1 \le i \le k_\ell} a_i^{(\ell)}.$$

The Foata encoding of σ_n is then the word

$$\mathrm{Foata}(\sigma_n) = (a_1^{(1)}, \ldots, a_{k_1}^{(1)}, a_1^{(2)}, \ldots, a_{k_2}^{(2)}, \ldots, a_1^{(\ell)}, \ldots, a_{k_\ell}^{(\ell)}).$$

Let

$$\mathcal{W}_n = \{(w_1, \ldots, w_n) \in [n]^n : \ \{w_1, \ldots, w_n\} = [n]\}.$$

Then

$$\mathrm{Foata} : \mathfrak{S}_n \to \mathcal{W}_n$$

defines a bijection.

Moreover, the number of cycles of σ_n coincides with the number of right-to-left minima of $\mathrm{Foata}(\sigma_n)$. Furthermore, the cycle lengths correspond to the spacings between successive right-to-left minima.

A direct application of the Foata correspondence is that the length of the cycle containing 1 (Thus any $1 \le k \le n$) in a uniform permutation is uniformly distributed on $\{1, 2, \ldots, n\}$.

5.1.2 Feller Coupling

Let

$$L^{(n)} = (L_1^{(n)}, L_2^{(n)}, \dots, 0, 0, \dots)$$

be the sequence of cycle lengths of a uniform permutation $\sigma_n \in \mathfrak{S}_n$, ordered according to the Foata construction and completed with zeros.

A very nice property of a uniform random permutation is its self-similarity: after removing the cycle containing 1 and relabelling the remaining elements in increasing order, the resulting permutation is again uniformly distributed on a smaller symmetric group.

More precisely, conditioning on $L_1^{(n)} = k$, for any fixed permutation $\tau \in \mathfrak{S}_{n-k}$, there are

$$\frac{(n-1)!}{(n-k)!}$$

different permutations that induce τ after removing the cycle containing 1 and relabelling the remaining elements.

The preceding self-similarity property allows us to describe the cycle structure of a uniform random permutation in terms of independent Bernoulli random variables.

Let $(B_k)_{k\geq 1}$ be independent Bernoulli random variables with $B_k \sim \mathrm{Ber}(1/k)$. Let

$$1 = I_1 < I_2 < \cdots$$

be the indices $k \geq 1$ such that $B_k = 1$. For $n \geq 1$, write $\ell_n := \max\{\ell : I_\ell \leq n\}$. Define

$$S^{(n)} = \Big((n+1) - I_{\ell_n}, I_{\ell_n} - I_{\ell_n - 1}, \dots, I_2 - I_1, 0, 0, \dots\Big).$$

Theorem 5.1 (Feller's coupling)**.** *We have*

$$L^{(n)} \stackrel{d}{=} S^{(n)}.$$

Proof. We proceed by induction on n. For $n = 1$, the statement is trivial.

Assume now that the result holds for all integers strictly smaller than n.

By the remark above,

$$\mathbb{P}(L_1^{(n)} = k) = \frac{1}{n}, \qquad 1 \leq k \leq n.$$

On the other hand,

$$S_1^{(n)} = (n+1) - I_{\ell_n}.$$

The event $\{S_1^{(n)} = k\}$ means that

$$B_n = \cdots = B_{n-k+2} = 0, \qquad B_{n-k+1} = 1.$$

Therefore

$$\begin{aligned}
\mathbb{P}(S_1^{(n)} = k) &= \prod_{j=n-k+2}^{n} \left(1 - \frac{1}{j}\right) \cdot \frac{1}{n-k+1} \\
&= \frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdots \frac{n-k+1}{n-k+2} \cdot \frac{1}{n-k+1} \\
&= \frac{1}{n}.
\end{aligned}$$

Thus

$$L_1^{(n)} \stackrel{d}{=} S_1^{(n)}.$$

Fix now $k \in \{1, \dots, n\}$.

Conditionally on the event $\{L_1^{(n)} = k\}$, removing the first cycle leaves a uniform permutation on $n - k$ elements. Hence

$$(L_2^{(n)}, L_3^{(n)}, \dots) \mid \{L_1^{(n)} = k\} \stackrel{d}{=} L^{(n-k)}.$$

Moreover, the event $\{S_1^{(n)} = k\}$ depends only on

$$B_n, \dots, B_{n-k+1},$$

while

$$B_{n-k}, \dots, B_1$$

are independent of these variables.

Thus, we have

$$(S_2^{(n)}, S_3^{(n)}, \dots) \mid \{S_1^{(n)} = k\} \stackrel{d}{=} S^{(n-k)}.$$

By the induction hypothesis,

$$L^{(n-k)} \stackrel{d}{=} S^{(n-k)}.$$

Hence the conditional distributions of the remaining coordinates coincide.

Since the first coordinates have the same distribution and the conditional laws of the remaining coordinates are equal, we conclude that

$$L^{(n)} \stackrel{d}{=} S^{(n)}.$$

□

As a consequence, the number of cycles C_n of σ_n satisfies

$$C_n = \#\{i : L_i^{(n)} > 0\} \stackrel{d}{=} \#\{i : S_i^{(n)} > 0\} = \sum_{k=1}^{n} B_k.$$

The expectation and the variance can be easily calculated. In particular,

$$\mathbb{E}[C_n] \sim \log n, \qquad \mathrm{Var}(C_n) \sim \log n.$$

By the Lindeberg's condition for Central limit Theorem, we have

$$\frac{C_n - \log n}{\sqrt{\log n}} \xrightarrow{d} N(0, 1),$$

5.1.3 Large cycles and Poisson-Dirichlet distribution

Let $U_1, U_2, \dots$ be independent identically distributed uniform random variables on $[0,1]$.

We construct a random partition of the interval $[0,1]$ by successive stick-breaking. Define

$$X_1 = U_1, \qquad X_2 = U_2(1-U_1), \qquad X_3 = U_3(1-U_1)(1-U_2),$$

and so on.

After the first n steps, the remaining length of the stick is

$$R_n = \prod_{i=1}^{n}(1-U_i).$$

Moreover,

$$X_1 + \cdots + X_n = 1 - R_n.$$

Take Logarithm of R_n and by the law of large numbers, we have

$$\sum_{i=1}^{\infty} X_i = 1 \qquad \text{a.s.}$$

Definition 5.2 (Poisson–Dirichlet distribution)**.** The law of the random sequence $(X_1, X_2, \dots)$ obtained from the above stick-breaking construction is called the Poisson–Dirichlet distribution.

Theorem 5.3 (Limit of Cycle Lengths)**.** *We have*

$$\left(\frac{L_i^{(n)}}{n}\right)_{i\geq 1} \xrightarrow[d]{n\to\infty} (X_i)_{i\geq 1}$$

in the space

$$\ell_1^{(1)} = \left\{(x_i)_{i\geq 1} : x_i \geq 0,\ \sum_{i=1}^{\infty} x_i = 1\right\},$$

equipped with the ℓ^1-metric.

Furthermore, let $L^{(n),\downarrow} = \left(L_1^{(n),\downarrow}, L_2^{(n),\downarrow}, \dots\right)$ be the decreasing rearrangement of $L^{(n)}$, and let $X^{\downarrow} = \left(X_1^{\downarrow}, X_2^{\downarrow}, \dots\right)$ be the decreasing rearrangement of $(X_i)_{i\geq 1}$.

Then

$$\left(\frac{L_i^{(n),\downarrow}}{n}\right)_{i\geq 1} \xrightarrow[d]{n\to\infty} (X_i^{\downarrow})_{i\geq 1}.$$

Proof. We first prove convergence of finite-dimensional distributions.

Since $S_1^{(n)}$ is uniformly distributed on $\{1, \dots, n\}$,

$$\frac{S_1^{(n)}}{n} \xrightarrow[d]{n\to\infty} U_1,$$

where $U_1 \sim \text{Unif}[0,1]$.

Furthermore, by the recursive property of Feller's coupling,

$$(S_2^{(n)}, S_3^{(n)}, \dots) \mid \{S_1^{(n)} = k\} \overset{d}{=} S^{(n-k)}.$$

Using this relation recursively, an induction on the dimension shows that for every fixed $r \geq 1$,

$$\left(\frac{S_1^{(n)}}{n}, \ldots, \frac{S_r^{(n)}}{n}\right) \xrightarrow[d]{n\to\infty} (X_1, \ldots, X_r),$$

where $(X_i)_{i\geq 1}$ has the Poisson-Dirichlet distribution.

Since $L^{(n)} \stackrel{d}{=} S^{(n)}$, the same convergence holds for $L^{(n)}$. The last convergence follows by the mapping theorem since decreasing rearrangement of a sequence is a continuous operation on ℓ_1 □

Remark 5.4. See Appendix A.2 for a proof when $r = 2$. The well-definedness and continuity of the decreasing rearrangement operator are discussed in [6]. The discussion of finite-dimensional convergence implies convergence in distribution in $\ell_1^{(1)}$ can be found in [5].

5.1.4 Poisson count of small cycles

Recall that $(B_k)_{k\geq 1}$ are independent random variables with $B_k \sim \mathrm{Ber}(1/k)$. Let

$$1 = I_1 < I_2 < \cdots$$

be the indices $k \geq 1$ such that $B_k = 1$.

For $i \geq 1$, define

$$Y_i := \#\left\{k \geq 1 : I_{k+1} - I_k = i\right\}.$$

Proposition 5.5. *Let $N_i(\sigma_n)$ denote the number of cycles of length i in the cycle decomposition of σ_n. Then*

$$\Big(N_i(\sigma_n) : i \geq 1\Big) \xrightarrow[d]{n\to\infty} \Big(Y_i : i \geq 1\Big)$$

in the sense of finite-dimensional distributions.

Proof. Fix $k_0 \geq 1$. For $1 \leq i \leq k_0$, let

$$Y_i^{(n)} := \#\{m \geq 1 : S_m^{(n)} = i\},$$

that is, $Y_i^{(n)}$ is the number of spacings of length i among

$$1 = I_1 < \cdots < I_{\ell_n} < n+1, \qquad \ell_n := \max\{\ell : I_\ell < n+1\}.$$

Since cycle counts are determined by the cycle-length sequence, Feller's coupling yields

$$\Big(N_i(\sigma_n) : 1 \leq i \leq k_0\Big) \stackrel{d}{=} \Big(Y_i^{(n)} : 1 \leq i \leq k_0\Big).$$

It therefore suffices to prove that

$$\Big(Y_i^{(n)} : 1 \leq i \leq k_0\Big) \underset{d}{\to} \Big(Y_i : 1 \leq i \leq k_0\Big).$$

Let

$$A_n = \Big\{\{I_j : j \geq 1\} \cap \{n+1-k_0, \ldots, n\} \neq \varnothing\Big\},$$

and

$$B_n = \Big\{ \exists j \geq 1 : I_j > n,\ I_{j+1} - I_j \leq k_0 \Big\}.$$

The only possible discrepancy between $(Y_i^{(n)} : 1 \leq i \leq k_0)$ and $(Y_i : 1 \leq i \leq k_0)$ comes from the cut at $n+1$. Indeed, all spacings strictly to the left of $n+1$ are counted in both objects, while the genuine spacing

$$I_{\ell_n+1} - I_{\ell_n}$$

is replaced in the truncated configuration by

$$(n+1) - I_{\ell_n}.$$

Consequently,

$$\Big\{ (Y_i^{(n)} : 1 \leq i \leq k_0) \neq (Y_i : 1 \leq i \leq k_0) \Big\} \subseteq A_n \cup B_n.$$

Let $f : \mathbb{R}^{k_0} \to \mathbb{R}$ be bounded and continuous, with $\|f\|_\infty \leq C$. Then

$$\begin{aligned} &\Big| \mathbb{E}\Big[f(Y_1^{(n)}, \ldots, Y_{k_0}^{(n)}) \Big] - \mathbb{E}\Big[f(Y_1, \ldots, Y_{k_0}) \Big] \Big| \\ &\leq 2C\, \mathbb{P}\Big((Y_i^{(n)} : 1 \leq i \leq k_0) \neq (Y_i : 1 \leq i \leq k_0) \Big) \\ &\leq 2C \Big(\mathbb{P}(A_n) + \mathbb{P}(B_n) \Big). \end{aligned}$$

Now

$$\mathbb{P}(A_n) \leq \sum_{m=n+1-k_0}^{n} \frac{1}{m} \longrightarrow 0.$$

Moreover, for every fixed i,

$$\mathbb{E}[Y_i] = \sum_{m \geq 1} \frac{1}{(m+i-1)(m+i)} < \infty,$$

so $Y_i < \infty$ almost surely. Hence

$$Y_1 + \cdots + Y_{k_0} < \infty \qquad \text{a.s.},$$

which implies

$$\mathbb{P}(B_n) \longrightarrow 0.$$

Therefore

$$\mathbb{E}\Big[f(Y_1^{(n)}, \ldots, Y_{k_0}^{(n)}) \Big] \longrightarrow \mathbb{E}\Big[f(Y_1, \ldots, Y_{k_0}) \Big].$$

Since f was arbitrary,

$$(Y_i^{(n)} : 1 \leq i \leq k_0) \underset{d}{\to} (Y_i : 1 \leq i \leq k_0).$$

Combining this with the distributional identity above gives

$$(N_i(\sigma_n) : 1 \leq i \leq k_0) \underset{d}{\to} (Y_i : 1 \leq i \leq k_0).$$

Since k_0 was arbitrary, the result follows. $\square$

It turns out the law of $(Y_i, i \geq 1)$ is quite simple.

Lemma 5.6 (Cauchy formula)**.** *Let σ_n be a uniform random permutation of $\mathfrak{S}_n$ and let $N_i(\sigma_n)$ denote the number of cycles of length i in σ_n. Then for every collection of nonnegative integers $c_1, \ldots, c_n$ satisfying*

$$\sum_{i=1}^{n} ic_i = n,$$

we have

$$\mathbb{P}(N_i(\sigma_n) = c_i,\ 1 \leq i \leq n) = \prod_{i=1}^{n} \frac{(1/i)^{c_i}}{c_i!}.$$

Proof. See Proposition 9.7 of [1] □

Let n_x be a geometric random variable with mean $\frac{x}{1-x}$, i.e.,

$$\mathbb{P}(n_x = n) = (1-x)x^n,$$

Proposition 5.7 (Poissonnization)**.** *If conditionally on $\{n_x = n\}$ σ_{n_x} is uniformly distributed on $\mathfrak{S}_n$, then the cycle counts*

$$\Big(N_i(\sigma_{n_x}) : i \geq 1\Big)$$

has the same law as independent Poisson random variables with means

$$\frac{x^i}{i}, \qquad i \geq 1.$$

Proof. Let $(c_i)_{i\geq 1}$ be a sequence of nonnegative integers which is eventually zero, and put

$$m = \sum_{i\geq 1} ic_i.$$

Since (c_i) is eventually zero, the above sum is finite.

By conditioning on the value of n_x, we get

$$\mathbb{P}\Big(N_i(\sigma_{n_x}) = c_i,\ i \geq 1\Big) = \sum_{n\geq 0} \mathbb{P}(n_x = n)\mathbb{P}\Big(N_i(\sigma_n) = c_i,\ i \geq 1\Big).$$

If the cycle counts are equal to $(c_i)_{i\geq 1}$, then the total number of elements must be

$$\sum_{i\geq 1} ic_i = m.$$

Hence all terms in the sum vanish except the term $n = m$. Therefore

$$\mathbb{P}\Big(N_i(\sigma_{n_x}) = c_i,\ i \geq 1\Big) = (1-x)x^m\mathbb{P}\Big(N_i(\sigma_m) = c_i,\ i \geq 1\Big).$$

By Cauchy's formula,

$$\mathbb{P}\Big(N_i(\sigma_m) = c_i,\ i \geq 1\Big) = \prod_{i\geq 1} \frac{(1/i)^{c_i}}{c_i!}.$$

Thus

$$\mathbb{P}\Big(N_i(\sigma_{n_x}) = c_i,\ i \geq 1\Big) = (1-x)x^m \prod_{i\geq 1} \frac{(1/i)^{c_i}}{c_i!}.$$

Since

$$m = \sum_{i\geq 1} ic_i,$$

we have

$$x^m = \prod_{i\geq 1} x^{ic_i}.$$

Hence

$$\mathbb{P}\Big(N_i(\sigma_{n_x}) = c_i,\ i \geq 1\Big) = (1-x) \prod_{i\geq 1} \frac{(x^i/i)^{c_i}}{c_i!}.$$

we obtain

$$\mathbb{P}\Big(N_i(\sigma_{n_x}) = c_i,\ i \geq 1\Big) = \prod_{i\geq 1} e^{-x^i/i} \frac{(x^i/i)^{c_i}}{c_i!}.$$

But the right-hand side is exactly

$$\prod_{i\geq 1} \mathbb{P}\left(Z_i = c_i\right),$$

where the random variables Z_i are independent and

$$Z_i \sim \text{Poisson}\left(\frac{x^i}{i}\right).$$

Therefore

$$\Big(N_i(\sigma_{n_x}) : i \geq 1\Big) \stackrel{d}{=} \Big(Z_i : i \geq 1\Big),$$

as desired. □

Theorem 5.8 (Goncharov–Kolchin)**.** *For $i \geq 1$, let $N_i(\sigma_n)$ denote the number of cycles of length i in a uniform permutation $\sigma_n \in \mathfrak{S}_n$.*

Then,

$$\Big(N_i(\sigma_n) : i \geq 1\Big) \xrightarrow[d]{n\to\infty} \Big(Z_i : i \geq 1\Big),$$

where the Z_i's are independent and

$$Z_i \sim \text{Poi}\Big(\frac{1}{i}\Big).$$

Proof. By Proposition 9.6, the sequence

$$X_n := \Big(N_i(\sigma_n) : i \geq 1\Big)$$

converges in distribution to some random vector

$$Y = (Y_i : i \geq 1).$$

For $x \in (0,1)$, let n_x be geometric with

$$\mathbb{P}(n_x = n) = (1-x)x^n, \qquad n \geq 0.$$

Then, for every fixed K,

$$\mathbb{P}(n_x \le K) = \sum_{n=0}^{K} (1-x)x^n = 1 - x^{K+1} \longrightarrow 0 \qquad (x \to 1).$$

Hence

$$n_x \to \infty \quad \text{in probability as } x \to 1.$$

Since $X_n \Rightarrow Y$ and $n_x \to \infty$ in probability, the random index theorem (See Appendix A.4) gives

$$X_{n_x} = \Big(N_i(\sigma_{n_x}) : i \ge 1\Big) \xrightarrow[d]{x\to 1} Y.$$

On the other hand, by Lemma 9.8,

$$\Big(N_i(\sigma_{n_x}) : i \ge 1\Big) \stackrel{d}{=} \Big(Z_i^{(x)} : i \ge 1\Big),$$

where the $Z_i^{(x)}$ are independent and

$$Z_i^{(x)} \sim \text{Poisson}\left(\frac{x^i}{i}\right).$$

As $x \to 1$,

$$\frac{x^i}{i} \to \frac{1}{i}.$$

Therefore, for every fixed k,

$$\Big(Z_1^{(x)}, \dots, Z_k^{(x)}\Big) \xrightarrow[d]{x\to 1} \Big(Z_1, \dots, Z_k\Big),$$

where the Z_i are independent and

$$Z_i \sim \text{Poisson}\left(\frac{1}{i}\right).$$

Thus

$$\Big(Z_i^{(x)} : i \ge 1\Big) \xrightarrow[d]{x\to 1} \Big(Z_i : i \ge 1\Big).$$

Hence the same family

$$\Big(N_i(\sigma_{n_x}) : i \ge 1\Big)$$

has two limits in distribution as $x \to 1$, namely Y and $(Z_i : i \ge 1)$. By uniqueness of limits in distribution,

$$Y \stackrel{d}{=} (Z_i : i \ge 1).$$

Since

$$X_n = \Big(N_i(\sigma_n) : i \ge 1\Big) \xrightarrow[d]{n\to\infty} Y,$$

we conclude that

$$\Big(N_i(\sigma_n) : i \ge 1\Big) \xrightarrow[d]{n\to\infty} \Big(Z_i : i \ge 1\Big).$$

□

Remark 5.9 (Derangements)**.** We recover the classical asymptotic formula for the number of derangements (permutations without fixed points). Indeed, by the Goncharov–Kolchin theorem,

$$N_1(\sigma_n) \xrightarrow[d]{n\to\infty} \text{Poi}(1).$$

Hence

$$\frac{\#\{\sigma \in \mathfrak{S}_n : \sigma \text{ has no fixed point}\}}{n!} = \mathbb{P}\Big(N_1(\sigma_n) = 0\Big) \xrightarrow[n\to\infty]{} \mathbb{P}\Big(\text{Poi}(1) = 0\Big) = e^{-1}.$$

5.2 Random recursive tree

Definition 5.10 (Random recursive tree)**.** For $n \geq 1$, let $\mathcal{T}_n$ be the set of labeled recursive trees on the vertex set

$$\{0, 1, \dots, n\}.$$

Let

$$\mathcal{T} = \bigsqcup_{n \geq 1} \mathcal{T}_n.$$

The random recursive tree (RRT) is the Markov chain

$$(T_n)_{n \geq 1}, \qquad T_n \in \mathcal{T},$$

defined as follows.

The initial state is

$$T_1 = 0\text{–}1.$$

For $n \geq 1$, conditionally on T_{n-1}, the tree T_n is obtained by attaching the new vertex n to a uniformly chosen vertex of T_{n-1}.

5.2.1 Chinese Restaurant process

Let $\sigma_n \in \mathfrak{S}_n$. We define

$$\langle \sigma_n \rangle_{n-1} \in \mathfrak{S}_{n-1}$$

by

$$\langle \sigma_n \rangle_{n-1}(k) = \begin{cases} \sigma_n(k), & \text{if } \sigma_n(k) \neq n, \\ \sigma_n(n), & \text{if } \sigma_n(k) = n, \end{cases} \qquad 1 \leq k \leq n-1.$$

If σ_n is uniform on $\mathfrak{S}_n$, then $\langle \sigma_n \rangle_{n-1}$ is uniform on $\mathfrak{S}_{n-1}$. Indeed, fix $\tau \in \mathfrak{S}_{n-1}$. There are exactly n permutations $\sigma \in \mathfrak{S}_n$ such that

$$\langle \sigma \rangle_{n-1} = \tau.$$

They are obtained either by making n a singleton cycle, or by inserting n immediately after one of the elements $1, \dots, n-1$ in the cycle decomposition of τ. Hence

$$\mathbb{P}(\langle \sigma_n \rangle_{n-1} = \tau) = \frac{n}{n!} = \frac{1}{(n-1)!}.$$

Thus $\langle \sigma_n \rangle_{n-1} \sim \mathrm{Unif}(\mathfrak{S}_{n-1})$.

Iterating this operation, for every $k \leq n$ we may define $\langle \sigma_n \rangle_k \in \mathfrak{S}_k$, and if σ_n is uniform on $\mathfrak{S}_n$, then

$$\langle \sigma_n \rangle_k \sim \mathrm{Unif}(\mathfrak{S}_k).$$

We now describe a coupling of the random permutations $(\sigma_n)_{n \geq 1}$ that realizes them as a Markov chain.

Let

$$\mathcal{S} = \bigsqcup_{n \geq 1} \mathfrak{S}_n$$

be the disjoint union of all finite symmetric groups. We define $(\sigma_n^{\mathrm{cr}})_{n \geq 1}$ on $\mathcal{S}$ as follows.

The initial distribution is

$$\mu = \delta_{(1)}.$$

For $\tau \in \mathfrak{S}_{n-1}$ and $\sigma \in \mathfrak{S}_n$, define the transition kernel

$$Q(\tau, \sigma) = \frac{1}{n} \mathbf{1}_{\{\langle \sigma \rangle_{n-1} = \tau\}}.$$

Equivalently, given $\tau \in \mathfrak{S}_{n-1}$, the next permutation is chosen uniformly among the n extensions of τ.

By the general construction of Markov chains, we obtain a process $(\sigma_n^{\mathrm{cr}})_{n \geq 1}$ on $\mathcal{S}$ with initial distribution μ and transition kernel Q.

We now prove that, for every $n \geq 1$,

$$\sigma_n^{\mathrm{cr}} \sim \mathrm{Unif}(\mathfrak{S}_n).$$

The case $n = 1$ is trivial. Assume that

$$\sigma_{n-1}^{\mathrm{cr}} \sim \mathrm{Unif}(\mathfrak{S}_{n-1}).$$

Fix $\sigma \in \mathfrak{S}_n$, and set

$$\tau = \langle \sigma \rangle_{n-1}.$$

Then $\tau \in \mathfrak{S}_{n-1}$, and it is the unique permutation such that

$$Q(\tau, \sigma) > 0.$$

Therefore

$$\begin{aligned} \mathbb{P}(\sigma_n^{\mathrm{cr}} = \sigma) &= \sum_{\eta \in \mathfrak{S}_{n-1}} \mathbb{P}(\sigma_{n-1}^{\mathrm{cr}} = \eta) Q(\eta, \sigma) \\ &= \mathbb{P}(\sigma_{n-1}^{\mathrm{cr}} = \tau) Q(\tau, \sigma) \\ &= \frac{1}{(n-1)!} \cdot \frac{1}{n} \\ &= \frac{1}{n!}. \end{aligned}$$

Thus every $\sigma \in \mathfrak{S}_n$ has probability $1/n!$, and hence

$$\sigma_n^{\mathrm{cr}} \sim \mathrm{Unif}(\mathfrak{S}_n).$$

The evolution of the cycle structure of $(\sigma_n^{\mathrm{cr}})_{n \geq 1}$ admits a particularly simple interpretation known as the Chinese restaurant process.

Customers $1, 2, 3, \ldots$ arrive sequentially in an imaginary restaurant. Customer 1 sits at a new table. Inductively, when customer n arrives, he either starts a new table with probability $1/n$, or sits immediately to the right of one of the $n - 1$ existing customers, each choice having probability $1/n$.

The Chinese restaurant process is naturally associated with the random recursive tree $(T_n)_{n \geq 1}$ by declaring that customer n corresponds to vertex n, and that vertex n is attached in T_n to the vertex corresponding to the customer sitting immediately to its left, or to the root vertex 0 if customer n starts a new table.

As a consequence of this correspondence, the degree of the root 0 in T_n is equal to the number of cycles in the cycle decomposition of σ_n^{cr}, that is,

$$\deg_{T_n}(0) \overset{(d)}{=} \#\mathrm{Cycles}(\sigma_n^{\mathrm{cr}}).$$

5.3 Yule Tree

It turns out that random recursive trees can be constructed via Yule trees. This construction leads to more refined calculations of tree properties. We begin with some basic properties of exponential random variables, which form the crux of the technique.

5.3.1 Properties of exponential laws

In the following, for $\alpha > 0$ we denote by $\mathcal{E}(\alpha)$ the exponential distribution with mean $1/\alpha$, i.e. the probability measure given by

$$\mathcal{E}(\alpha)(dx) = \alpha e^{-\alpha x} \mathbf{1}_{x \geq 0}\, dx.$$

We shall say that α is the rate of the exponential distribution. One of its key properties is the memoryless property: if $X \sim \mathcal{E}(\alpha)$, then for all $s, t \geq 0$,

$$\mathbb{P}(X > s + t \mid X > s) = \mathbb{P}(X > t).$$

Proposition 5.11 (Choosing with clocks)**.** *Let*

$$X_1 \sim \mathcal{E}(\alpha_1), \quad \ldots, \quad X_k \sim \mathcal{E}(\alpha_k)$$

be independent exponential random variables with parameters $\alpha_1, \ldots, \alpha_k > 0$. Define

$$m = \min\{X_1, \ldots, X_k\},$$

and let $J \in \{1, \ldots, k\}$ be the index such that $X_J = m$ (ties broken by the smallest index).

Then:

$$m \sim \mathcal{E}(\alpha_1 + \cdots + \alpha_k), \qquad \mathbb{P}(J = i) = \frac{\alpha_i}{\alpha_1 + \cdots + \alpha_k}.$$

Moreover, conditionally on the event $\{J = i\}$, the random variables

$$Y_j := X_j - m, \quad j \neq i,$$

are independent, and for each $j \neq i$,

$$Y_j \sim \mathcal{E}(\alpha_j).$$

Proof. We first compute the law of m. For $t \geq 0$,

$$\mathbb{P}(m > t) = \mathbb{P}(X_1 > t, \ldots, X_k > t) = \prod_{i=1}^{k} \mathbb{P}(X_i > t) = \prod_{i=1}^{k} e^{-\alpha_i t} = e^{-(\alpha_1 + \cdots + \alpha_k)t}.$$

Hence,

$$m \sim \mathcal{E}(\alpha_1 + \cdots + \alpha_k).$$

Note that for any $i \neq j$,

$$\mathbb{P}(X_i = X_j) = 0.$$

Hence,

$$\mathbb{P}(J = i) = \mathbb{P}(X_i < X_j \ \forall j \neq i).$$

Using the joint density and Fubini's theorem,

$$\mathbb{P}(J = i) = \frac{\alpha_i}{\sum_{j=1}^{k} \alpha_j}.$$

The remaining statement follows from similar Fubini-type calculations, using the independence and memoryless property of exponential random variables.

□

We end this section on extremal statistics of the exponential distribution.

Lemma 5.12. *Let $X_1, \dots, X_n$ be i.i.d. random variables with exponential law* Exp(1)*. Let*

$$M_n = \max\{X_1, \dots, X_n\}.$$

Then

$$M_n - \log n \underset{d}{\to} G,$$

where G has the Gumbel distribution with distribution function

$$\mathbb{P}(G \leq x) = \exp(-e^{-x}), \quad x \in \mathbb{R}.$$

Proof. For any $x \in \mathbb{R}$, we compute

$$\mathbb{P}(M_n \leq \log n + x) = (\mathbb{P}(X_1 \leq \log n + x))^n = \left(1 - \frac{e^{-x}}{n}\right)^n.$$

Taking the limit as $n \to \infty$, we obtain

$$\mathbb{P}(M_n - \log n \leq x) \to \exp(-e^{-x}).$$

This is the distribution function of the Gumbel distribution. □

5.3.2 From Yule Tree to RRT

Definition 5.13 (Yule tree of order k)**.** Fix an integer $k \geq 1$. The Yule tree of order k is the genealogical tree of a branching process started from a single particle, in which each particle lives for an independent exponential time with parameter 1 and, upon death, is replaced by exactly k offspring particles.

For $t \geq 0$, let Y_t denote the number of leaves of the Yule tree at time t.

Remark 5.14. The leaf-counting process $(Y_t)_{t \geq 0}$ admits the following explicit representation.

Let $(E_n)_{n \geq 1}$ be independent random variables such that

$$E_n \sim \mathrm{Exp}\Big(1 + (n-1)(k-1)\Big).$$

Define

$$T_0 := 0, \qquad T_n := \sum_{j=1}^{n} E_j.$$

Then

$$Y_t = 1 + (k-1)\sup\{n \geq 0 : T_n \leq t\}.$$

Remark 5.15. When $k = 1$, the leaf-counting process is trivial:

$$Y_t \equiv 1.$$

Nevertheless, the jump-counting process

$$N_t := \sup\{n \geq 0 : T_n \leq t\}$$

remains non-trivial. By the classical construction of the Poisson process from exponential waiting times, $(N_t)_{t\geq 0}$ is a standard Poisson process of rate 1.

Informally, we have the following construction of random recursive tree from Yule tree.

Proposition 5.16 (Yule construction of the RRT)**.** *Consider a Yule tree of order* 2*. At each branching event, declare that the left child inherits the identity of its parent, while the right child receives a new label according to its order of appearance. Contracting all left edges yields an increasing tree.*

Then, at the successive branching times

$$0 = T_0 < T_1 < T_2 < \cdots,$$

the resulting increasing trees form a random recursive tree process.

Proof. At time T_n, the Yule tree has $n + 1$ leaves. By the memoryless property of exponential clocks, each leaf carries an independent Exp(1) clock. Hence the next leaf to split is uniformly distributed among the $n + 1$ leaves.

After contracting left edges, a splitting event creates exactly one new vertex, attached to the vertex corresponding to the splitting leaf. Therefore the new vertex chooses its parent uniformly among the existing $n + 1$ vertices, which is precisely the growth rule of the random recursive tree. □

Appendix A

Auxiliary Proofs

A.1 Proof that A is a subgroup of $\mathbb{Z}$

Now we prove that A is a subgroup of $\mathbb{Z}$.

Let

$$a := \min\{\alpha \in A : \alpha > 0\}.$$

We will show that $A = a\mathbb{Z}$.
First, let $x \in A$ with $x < 0$ and set

$$g = \gcd(a, x).$$

By Bézout's identity, there exist integers $m, n > 0$ such that

$$g = ma + nx.$$

Therefore

$$\begin{aligned} Q^{m+n}(0, g) &\geq Q^m(0, ma)\, Q^n(ma, ma + nx) \\ &= Q^m(0, ma)\, Q^n(0, nx) \\ &\geq Q(0, a)^m\, Q(0, x)^n \\ &> 0. \end{aligned}$$

Thus $g \in A$, and by the definition of a we must have $g \geq a$.
Since $g \mid a$, it follows that $g = a$.
Hence $x = -ua$ for some $u > 0$.
Now observe that

$$-a = (u - 1)a + x.$$

Therefore

$$\begin{aligned} Q^u(0, -a) &\geq Q^{u-1}(0, (u-1)a)\, Q((u-1)a, (u-1)a + x) \\ &= Q^{u-1}(0, (u-1)a)\, Q(0, x) \\ &\geq Q(0, a)^{u-1}\, Q(0, x) \\ &> 0. \end{aligned}$$

Hence $-a \in A$.
Finally, let $y \in A$ with $y > 0$.

Then $y \geq a$, so we may write

$$y = qa + r,$$

where $q > 0$ and $0 \leq r < a$.
Then

$$\begin{aligned} Q^{q+1}(0, r) &\geq Q(0, y)\, Q^q(y, y - qa) \\ &= Q(0, y)\, Q^q(0, -qa) \\ &\geq Q(0, y)\, Q(0, -a)^q \\ &> 0. \end{aligned}$$

Thus $r \in A$.
By the definition of a, this implies $r = 0$.
Hence $y = qa$.
We have therefore shown that every element of A is an integer multiple of a, i.e.

$$A \subseteq a\mathbb{Z}.$$

Since $a \in A$ and $-a \in A$, we also have

$$a\mathbb{Z} \subseteq A.$$

Therefore

$$A = a\mathbb{Z}.$$

A.2 Proof the case $r = 2$

We prove the case $r = 2$. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be bounded and continuous. Since

$$S_1^{(n)} \sim \text{Unif}\{1, \ldots, n\},$$

we have

$$\mathbb{E}\left[f\left(\frac{S_1^{(n)}}{n}, \frac{S_2^{(n)}}{n}\right)\right] = \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\left[f\left(\frac{k}{n}, \frac{S_2^{(n)}}{n}\right)\middle| S_1^{(n)} = k\right].$$

By the recursive property,

$$S_2^{(n)} \mid \{S_1^{(n)} = k\} \stackrel{d}{=} S_1^{(n-k)}.$$

Thus

$$\mathbb{E}\left[f\left(\frac{S_1^{(n)}}{n}, \frac{S_2^{(n)}}{n}\right)\right] = \frac{1}{n}\sum_{k=1}^{n-1} \mathbb{E}\left[f\left(\frac{k}{n}, \frac{S_1^{(n-k)}}{n}\right)\right] + \frac{1}{n} f(1, 0).$$

Since $S_1^{(m)} \sim \text{Unif}\{1, \ldots, m\}$, we get

$$\mathbb{E}\left[f\left(\frac{S_1^{(n)}}{n}, \frac{S_2^{(n)}}{n}\right)\right] = \frac{1}{n}\sum_{k=1}^{n-1} \frac{1}{n-k} \sum_{j=1}^{n-k} f\left(\frac{k}{n}, \frac{j}{n}\right) + o(1).$$

Equivalently,

$$\frac{j}{n} = \left(1 - \frac{k}{n}\right) \frac{j}{n-k}.$$

Hence the above expression is

$$\frac{1}{n} \sum_{k=1}^{n-1} \frac{1}{n-k} \sum_{j=1}^{n-k} f\left(\frac{k}{n}, \left(1 - \frac{k}{n}\right) \frac{j}{n-k}\right) + o(1).$$

As $n \to \infty$, this double Riemann sum converges to

$$\int_0^1 \int_0^1 f\big(x, (1-x)u\big)\, du\, dx.$$

Since U_1, U_2 are independent uniform random variables on $[0,1]$,

$$\int_0^1 \int_0^1 f\big(x, (1-x)u\big)\, du\, dx = \mathbb{E}\left[f\big(U_1, (1-U_1)U_2\big)\right].$$

Therefore

$$\mathbb{E}\left[f\left(\frac{S_1^{(n)}}{n}, \frac{S_2^{(n)}}{n}\right)\right] \longrightarrow \mathbb{E}\left[f\big(U_1, (1-U_1)U_2\big)\right].$$

Hence

$$\left(\frac{S_1^{(n)}}{n}, \frac{S_2^{(n)}}{n}\right) \xrightarrow[d]{n\to\infty} (U_1, (1-U_1)U_2).$$

A.3 Auxiliary Convergence Results

Proposition A.1. *Let $X_n \xrightarrow{d} X$, and assume that X is finite almost surely. Then*

$$\frac{X_n}{n} \xrightarrow{\mathbb{P}} 0.$$

Proof. Since $X_n \xrightarrow{d} X$ and X is finite almost surely, the sequence (X_n) is tight.

Hence for every $\delta > 0$, there exists $M > 0$ such that for all sufficiently large n,

$$\mathbb{P}(|X_n| > M) < \delta.$$

Let $\varepsilon > 0$.

Choose n large enough such that $\varepsilon n > M$.

Then

$$\mathbb{P}\left(\left|\frac{X_n}{n}\right| > \varepsilon\right) = \mathbb{P}(|X_n| > \varepsilon n) \leq \mathbb{P}(|X_n| > M) < \delta.$$

Since $\delta > 0$ is arbitrary,

$$\frac{X_n}{n} \xrightarrow{\mathbb{P}} 0.$$

□

Proposition A.2. *Suppose that*

$$\mathbb{E}[X_n] \to c$$

and

$$X_n - \mathbb{E}[X_n] \xrightarrow{\mathbb{P}} 0.$$

Then

$$X_n \xrightarrow{\mathbb{P}} c.$$

Proof. Let $\varepsilon > 0$. By the triangle inequality,

$$|X_n - c| \le |X_n - \mathbb{E}[X_n]| + |\mathbb{E}[X_n] - c|.$$

Since $\mathbb{E}[X_n] \to c$, for all sufficiently large n,

$$|\mathbb{E}[X_n] - c| < \frac{\varepsilon}{2}.$$

Hence

$$\{|X_n - c| > \varepsilon\} \subseteq \left\{|X_n - \mathbb{E}[X_n]| > \frac{\varepsilon}{2}\right\}.$$

Therefore,

$$\mathbb{P}(|X_n - c| > \varepsilon) \le \mathbb{P}\left(|X_n - \mathbb{E}[X_n]| > \frac{\varepsilon}{2}\right).$$

Since

$$X_n - \mathbb{E}[X_n] \xrightarrow{\mathbb{P}} 0,$$

the right-hand side converges to 0. Thus

$$\mathbb{P}(|X_n - c| > \varepsilon) \to 0,$$

which proves that

$$X_n \xrightarrow{\mathbb{P}} c.$$

□

A.4 Anscombe's theorem

Proposition A.3 (Anscombe's theorem)**.** *Let N_x be defined by*

$$\mathbb{P}(N_x = n) = (1 - x)x^n, \qquad n = 0, 1, 2, \ldots.$$

Assume that

$$X_n \xrightarrow{d} Y.$$

Then

$$X_{N_x} \xrightarrow[x \to 1]{d} Y.$$

Proof. Let f be bounded and continuous. It suffices to show that

$$\mathbb{E}[f(X_{N_x})] \to \mathbb{E}[f(Y)].$$

By conditioning on N_x,

$$\mathbb{E}[f(X_{N_x})] = \sum_{n=0}^{\infty} \mathbb{P}(N_x = n)\mathbb{E}[f(X_{N_x}) \mid N_x = n].$$

On the event $\{N_x = n\}$, we have $X_{N_x} = X_n$. Hence

$$\mathbb{E}[f(X_{N_x})] = \sum_{n=0}^{\infty}(1-x)x^n\mathbb{E}[f(X_n)].$$

Set

$$a_n := \mathbb{E}[f(X_n)], \qquad a := \mathbb{E}[f(Y)].$$

Since $X_n \xrightarrow{d} Y$, we have $a_n \to a$. Therefore

$$\mathbb{E}[f(X_{N_x})] - a = \sum_{n=0}^{\infty}(1-x)x^n(a_n - a).$$

Fix $\varepsilon > 0$. Choose M such that for all $n \geq M$,

$$|a_n - a| < \varepsilon.$$

Then

$$\begin{aligned}|\mathbb{E}[f(X_{N_x})] - a| &\leq \sum_{n<M}(1-x)x^n|a_n - a| + \sum_{n\geq M}(1-x)x^n|a_n - a| \\ &\leq 2\|f\|_\infty(1-x^M) + \varepsilon.\end{aligned}$$

Since $1 - x^M \to 0$ as $x \to 1$, we get

$$\limsup_{x\to 1} |\mathbb{E}[f(X_{N_x})] - a| \leq \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary,

$$\mathbb{E}[f(X_{N_x})] \to \mathbb{E}[f(Y)].$$

Hence

$$X_{N_x} \xrightarrow[x\to 1]{d} Y.$$

□

Bibliography

[1] Nicolas Curien. *A Random Walk among Random Graphs*, volume 31 of *Cours Spécialisés*. Société Mathématique de France, 2025.

[2] Thomas Duquesne. Processus markoviens 1: Chaînes de markov, 2024. Lecture notes, not published, 2023–2024.

[3] Jean-François Le Gall. Random trees and applications. *Probability Surveys*, 2:245–311, 2005.

[4] Jean-François Le Gall. *Measure Theory, Probability, and Stochastic Processes*. Springer, 2022.

[5] Andrew Ren. Convergence in law in $\ell_1^{(1)}$. Mathematics Stack Exchange, 2026. Available at: https://math.stackexchange.com/q/5138990, accessed June 2026.

[6] Andrew Ren. Is the rearranging operator continuous? Mathematics Stack Exchange, 2026. Available at: https://math.stackexchange.com/q/5138988, accessed June 2026.

Declaration on the Use of Generative AI. ChatGPT (OpenAI, GPT-5.5) was used as an auxiliary tool to improve the clarity of the writing and to assist in the formulation and presentation of certain proofs. The author takes full responsibility for all mathematical content, including all proofs, results, and conclusions presented in this thesis.