Scaling limits of random graphs: exercises

Please send any comments or corrections to goldschm@stats.ox.ac.uk.

1. (Aldous-Broder algorithm) The purpose of this exercise is to convince you that the Aldous-Broder algorithm generates a uniform random tree. We saw in lectures that the tree it generates is stationary, so we just need to prove that the stationary distribution is uniform. Recall that \( T_n \) denotes the set of rooted trees labelled by \( \{1, 2, \ldots, n\} \). Consider the time-reversed chain (which must have the same stationary distribution). For \( \tau, \tau' \in T_n^\ast \), write \( q(\tau, \tau') \) for the transition probability from \( \tau \) to \( \tau' \) for the time-reversed chain.

   (a) Argue that the chain is irreducible on \( T_n^\ast \).
   (b) Show that for fixed \( \tau \), \( q(\tau, \tau') = 0 \) or \( \frac{1}{n-1} \).
   (c) Show that for fixed \( \tau' \), \( q(\tau, \tau') = 0 \) or \( \frac{1}{n-1} \).
   (d) It follows that \( Q = (q(\tau, \tau'))_{\tau, \tau' \in T_n^\ast} \) is a doubly stochastic matrix. Deduce that the stationary distribution must be uniform.

2. (Uniform random trees) Let \( T \) be a Galton-Watson tree with Poisson(1) offspring distribution and total progeny \( N \).

   (a) Fix a particular rooted ordered tree \( t \) with \( n \) vertices having numbers of children \( c_v, v \in t \). What is \( \Pr(T = t) \)?
   (b) Condition on the event \( \{N = n\} \). Assign the vertices of \( T \) a uniformly random labelling by \( [n] \), and let \( \tilde{T} \) be the labelled tree obtained by forgetting the ordering and the root. Show that \( \tilde{T} \) has the same distribution as \( T_n \), a uniform random tree on \( n \) vertices.

   Hint: it suffices to show that the probability of obtaining a particular tree \( t \) is a function of \( n \) only.

3. (Other combinatorial trees) Let \( T \) be a Galton-Watson tree with offspring distribution \( (p_k)_{k \geq 0} \) and total progeny \( N \).

   (a) Show that if \( p_k = 2^{-k-1}, k \geq 0 \) then, conditional on \( N = n \), \( T \) is uniform on the set of ordered rooted trees with \( n \) vertices.
   (b) Show that if \( p_0 = 1/2 \) and \( p_2 = 1/2 \) then, conditional on \( N = n \) (for \( n \) odd), \( T \) is uniform on the set of complete binary trees.

4. (Contour function and convergence to the Brownian CRT) In Lecture 1, we discussed the depth-first walk and the height function of a tree. A third encoding which is often used is the so-called contour function \( (C(i), 0 \leq i \leq 2(n-1)) \). For a rooted ordered tree \( t \), we imagine a particle tracing the outline of the tree from left to right at speed 1. (The picture
below is for a labelled tree, with a planar embedding given by the labels.) Notice that we visit every vertex a number of times given by its degree.

Let $T_n$ be a Galton-Watson tree with offspring distribution $p(k) = \left(\frac{1}{2}\right)^{k+1}, k \geq 0$, conditioned to have total progeny $N = n$, as in Question 3(a). Let $(C^n(i), 0 \leq i \leq 2(n - 1))$ be its contour function. It will be convenient to define a somewhat shifted version: let $\tilde{C}^n(0) = 0$, $\tilde{C}^n(2n) = 0$ and, for $1 \leq i \leq 2n - 1$, $\tilde{C}^n(i) = 1 + C(i - 1)$.

(a) Show that $(\tilde{C}^n(i), 0 \leq i \leq 2n))$ has the same distribution as a simple symmetric random walk (i.e. a random walk which makes steps of $+1$ with probability $1/2$ and steps of $-1$ with probability $1/2$) conditioned to return to the origin for the first time at time $2n$.

Hint: first consider the unconditioned Galton-Watson tree with this offspring distribution.

(b) It’s straightforward to interpolate linearly to get a continuous function $\tilde{C}^n : [0, 2n] \to \mathbb{R}_+$. Let $\tilde{T}^n$ be the $\mathbb{R}$-tree encoded by this linear interpolation. Show that $d_{GH}(T^n, \tilde{T}^n) \leq \frac{1}{2}$.

Hint: notice that $T^n$ considered as a metric space has only $n$ points, whereas $\tilde{T}^n$ is an $\mathbb{R}$-tree and consists of uncountably many points. Draw a picture and find a correspondence.

(c) Suppose that we have continuous excursions $f : [0, 1] \to \mathbb{R}_+ \text{ and } g : [0, 1] \to \mathbb{R}_+$ which encode $\mathbb{R}$-trees $T_f$ and $T_g$. For $t \in [0, 1]$, let $p_f(t)$ be the image of $t$ in the tree $T_f$ and similarly for $p_g(t)$. Now define a correspondence

$$R = \{(x, y) \in T_f \times T_g : x = p_f(t), y = p_g(t) \text{ for some } t \in [0, 1]\}.$$

Show that $\text{dis}(R) \leq 4\|f - g\|_\infty$.

Hint: recall how the metric in an $\mathbb{R}$-tree is related to the function encoding it.

(d) Observe that the variance of the step-size in a simple symmetric random walk is 1. Hence, by Kaigh’s theorem, we have

$$\frac{1}{\sqrt{2(n - 1)}}(C^n(2(n - 1)t), 0 \leq t \leq 1) \xrightarrow{d} (e(t), 0 \leq t \leq 1)$$

as $n \to \infty$. Use this, (b) and (c) to prove directly that $\frac{1}{\sqrt{n}}T^n$ converges to a constant multiple of the Brownian CRT in the Gromov-Hausdorff sense.

Hint: you may want to use Skorokhod’s representation theorem in order to work on a probability space where the convergence (*) occurs almost surely.

5. **(Generating a component of the Erdős–Rényi random graph conditionally on its size)**

Let $\tilde{G}_m^p$ be a connected graph on vertices labelled by $[m]$ generated as follows:

- Pick a random tree $\tilde{T}_m^p$ such that
  \[
  \mathbb{P}(\tilde{T}_m^p = T) \propto (1 - p)^{-a(T)}, \quad T \in \mathcal{T}.
  \]

- Add each of the $a(\tilde{T}_m^p)$ permitted edges independently with probability $p$.

Show that $\tilde{G}_m^p$ has the same distribution as a component of $G(n, p)$ conditioned to have vertex set $[m]$.

6. **(Configuration model)**

Fix a degree sequence $d_1, \ldots, d_n$. Show that the probability of generating a particular multigraph $G$ with these degrees is

\[
\frac{1}{(\ell_n - 1)!! \prod_{i=1}^n d_i!} \prod_{e \in E(G)} \text{mult}(e)!
\]

where $\ell_n = \sum_{i=1}^n d_i$, $\text{sl}(G)$ is the number of self-loops in $G$ and $\text{mult}(e)$ is the multiplicity of the edge $e \in E(G)$.

7. **(Size-biased random variables)**

Suppose that $X$ is a non-negative integer-valued random variable with finite mean, and let $X^*$ have the *size-biased distribution*

\[
\mathbb{P}(X^* = k) = \frac{k \mathbb{P}(X = k)}{\mathbb{E}[X]}, \quad k \geq 1.
\]

Show that $X^* \overset{d}{=} 1 + X$ if and only if $X$ has a Poisson distribution.

8. **(Size-biased order and absolute continuity)**

Suppose that $X_1, X_2, \ldots, X_n$ are i.i.d. random variables with finite mean, and let $(\tilde{X}_1^n, \tilde{X}_2^n, \ldots, \tilde{X}_m^n)$ be the same random variables in size-biased random order. Let $X_1^*, X_2^*, \ldots$ be i.i.d. with the (true) size-biased distribution. Show that for $m < n$,

\[
\mathbb{P}(\tilde{X}_1^n = k_1, \tilde{X}_2^n = k_2, \ldots, \tilde{X}_m^n = k_m)
= \frac{n! \mu^m}{(n-m)!} \left[ \prod_{i=1}^m \frac{1}{\sum_{j=i}^m k_j + \Xi_{n-m}} \right] \mathbb{P}(X_1^* = k_1, X_2^* = k_2, \ldots, X_m^* = k_m),
\]

where $\Xi_{n-m} \overset{d}{=} X_1 + \ldots + X_{n-m}$. 