Random trees and their scaling limits: complements and exercises

[Version of September 12, 2016]

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

1. **(Weak convergence)** Suppose that $(X_n)_{n\geq 0}$ and X are real-valued random variables. Show that
$$\mathbb{P}(X_n \leq x) \to \mathbb{P}(X \leq x) \text{ as } n \to \infty$$
for all points x of continuity of the right-hand side if and only if
$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)] \text{ as } n \to \infty$$
for all bounded continuous functions $f : \mathbb{R} \to \mathbb{R}$.
Hint: in one direction, you may find it useful to consider a continuous approximation to the indicator function
$$\mathbb{1}_{\{y\leq x\}} = \begin{cases} 1 & \text{if } y \leq x \\ 0 & \text{otherwise.} \end{cases}$$
In the other, you may want to use Skorokhod's representation theorem. See Durrett, **Probability: theory and examples** *if you get stuck!*

2. **(Leaves in a uniform random tree)** Let N_n be the number of leaves in a uniform random tree T_n on n labelled vertices. In this exercise, we will find a limit in probability for N_n/n as $n \to \infty$. It will be useful to write $N_n = \sum_{i=1}^n I_i$, where
$$I_i = \begin{cases} 1 & \text{if the vertex labelled } i \text{ is a leaf} \\ 0 & \text{otherwise.} \end{cases}$$

   (a) Using Cayley's formula, show that for any $i \in \{1, 2, \ldots, n\}$,
$$\mathbb{P}(i \text{ is a leaf}) = \frac{(n-1)^{n-2}}{n^{n-2}}.$$

   (b) If the indicator random variables $I_1, I_2, \ldots, I_n$ were independent, we would now be able to use the weak law of large numbers to find a limit in probability for N_n/n. But it is not quite the case that the random variables are independent. Show that for $i \neq j$,
$$\mathbb{P}(i \text{ is a leaf and } j \text{ is a leaf}) = \frac{(n-2)^{n-2}}{n^{n-2}}.$$

   (c) Find $\operatorname{var}(I_i)$ and $\operatorname{cov}(I_i, I_j)$ for $i \neq j$ and, hence, find $\operatorname{var}(N_n/n)$.

(d) Using Chebyshev's inequality, prove that N_n/n converges in probability as $n \to \infty$, and specify the limit.

3. **(Inhomogeneous Poisson processes)**

(a) Suppose that $(P(t), t \geq 0)$ is an inhomogeneous Poisson process of rate $\lambda = (\lambda(t), t \geq 0)$, and let $(Q(t), t \geq 0)$ be an ordinary Poisson process of rate 1. Show that

$$(P(t), t \geq 0) \stackrel{d}{=} (Q(\Lambda(t)), t \geq 0),$$

where $\Lambda(t) := \int_0^t \lambda(x)dx$.

(b) Suppose that $C_1, C_2, \ldots$ are the points of an inhomogeneous Poisson process of rate $\lambda(t) = t$, $t \geq 0$. Use (a) to show that if $E_1, E_2, \ldots$ are i.i.d. Exponential(1) then

$$(C_k,\ k \geq 1) \stackrel{d}{=} \left(\sqrt{2 \sum_{i=1}^{k} E_i},\ k \geq 1 \right).$$

For much more about general Poisson processes, see the classic book by J. F. C. Kingman, **Poisson processes**, Oxford University Press (1993).

4. **(Conditioned Galton-Watson trees)** Let T be a Galton-Watson tree with offspring distribution p and total progeny N. Prove the following statements.

(a) If $p(k) = 2^{-k-1}$, $k \geq 0$ (i.e. Geometric(1/2) offspring distribution) then conditional on $N = n$, the tree is uniform on the set of ordered trees with n vertices.

(b) If $p(0) = 1/2$ and $p(2) = 1/2$ then, conditional on $N = n$ (for n odd), the tree is uniform on the set of (complete) binary trees.

5. **(The total population of a Poisson-Galton-Watson process)** Let N be the total population size in a Galton-Watson branching process with Poisson(λ) offspring distribution, for some $\lambda \in (0, 1]$. Recall that $N = \inf\{k \geq 0 : X(k) = -1\}$, where the depth-first walk $(X(k), k \geq 0)$ is a random walk with step distribution $\nu(k) = e^{-\lambda}\lambda^{k+1}/(k+1)!$, $k \geq -1$.

(a) Consider a possible path for X which first hits -1 at time n. There are n different cyclic rearrangements of the n steps

$$X(1) - X(0), X(2) - X(1), \ldots, X(n) - X(n-1)$$

of X i.e.

$$X(i+1) - X(i), X(i+2) - X(i+1), \ldots, X(i+n) - X(i+n-1),$$

for $0 \leq i \leq n-1$, where the indices are taken mod n. We always have $\sum_{k=0}^{n-1}(X(i+k+1) - X(i+k)) = -1$. Show that only one of these cyclic rearrangements results in the walk hitting -1 for the first time at n.

(b) Use this to argue that we must have

$$\mathbb{P}(N = n) = \frac{1}{n}\mathbb{P}(X(n) = -1).$$

(c) Deduce from this that

$$\mathbb{P}(N = n) = \frac{(\lambda n)^{n-1}e^{-\lambda n}}{n!}, \quad n \geq 1.$$

This is known as the *Borel distribution*.

6. **(Weak ladder heights for skip-free random walks)** The random walks which occur as depth-first walks of Galton-Watson trees have the special property that they are *skip-free to the left*, which means that they have step distribution concentrated on $\{-1, 0, 1, 2, \ldots\}$. It turns out that these random walks have some particularly nice properties.

Let $(X(k), k \geq 0)$ be a random walk with step distribution $\nu(k)$, $k \geq -1$. Assume that $\sum_{k\geq -1} k\nu(k) = 0$ and that $\sum_{k\geq -1} k^2\nu(k) = \sigma^2 < \infty$. Suppose $X(0) = 0$ and let $T = \inf\{k \geq 1 : X(k) \geq 0\}$. This is called the *first weak (ascending) ladder time*. The random walk is recurrent, and so $T < \infty$ a.s. and it follows that $X(T) < \infty$ a.s. Write $\mu(k) = \mathbb{P}(X(T) = k)$ for $k \geq 0$, the law of the *first weak ladder height*.

If the first step of the random walk is to 0 or above, then $T = 1$ and $X(T)$ is simply the new location of the random walk.

If the first step is to -1, on the other hand, things can be more involved. In general, the random walk may now make several excursions which go below -1 and stay below it before returning to -1. Finally the walk leaves -1, perhaps initially going downwards, but eventually reaching $\{0, 1, 2, \ldots\}$ *without hitting* -1 *again*. Indeed, using the strong Markov property, we can see that the random walk makes a geometrically distributed number of excursions which return to -1 before it hits $\{0, 1, \ldots\}$, where the parameter of the geometric is (by translation-invariance) $\mathbb{P}(X(T) > 0)$.

(a) By conditioning on the first step of the random walk, and using the above considerations, show that for $k \geq 0$,

$$\mathbb{P}(X(T) = k) = \nu(k) + \nu(-1)\mathbb{P}(X(T) = k + 1 | X(T) > 0).$$

(b) Show that for $k \geq 0$,

$$\mathbb{P}(X(T) = k) = \sum_{j=0}^{\infty} \left(\frac{\nu(-1)}{\mathbb{P}(X(T) > 0)} \right)^j \nu(k + j).$$

(c) Show directly that

$$\sum_{k=0}^{\infty} \bar{\nu}(k) = 1$$

where $\bar{\nu(k)} = \sum_{j=k}^{\infty} \nu(j)$.

(d) Using the fact that $\sum_{k=0}^{\infty} \mathbb{P}(X(T) = k) = 1$, deduce carefully that we must have

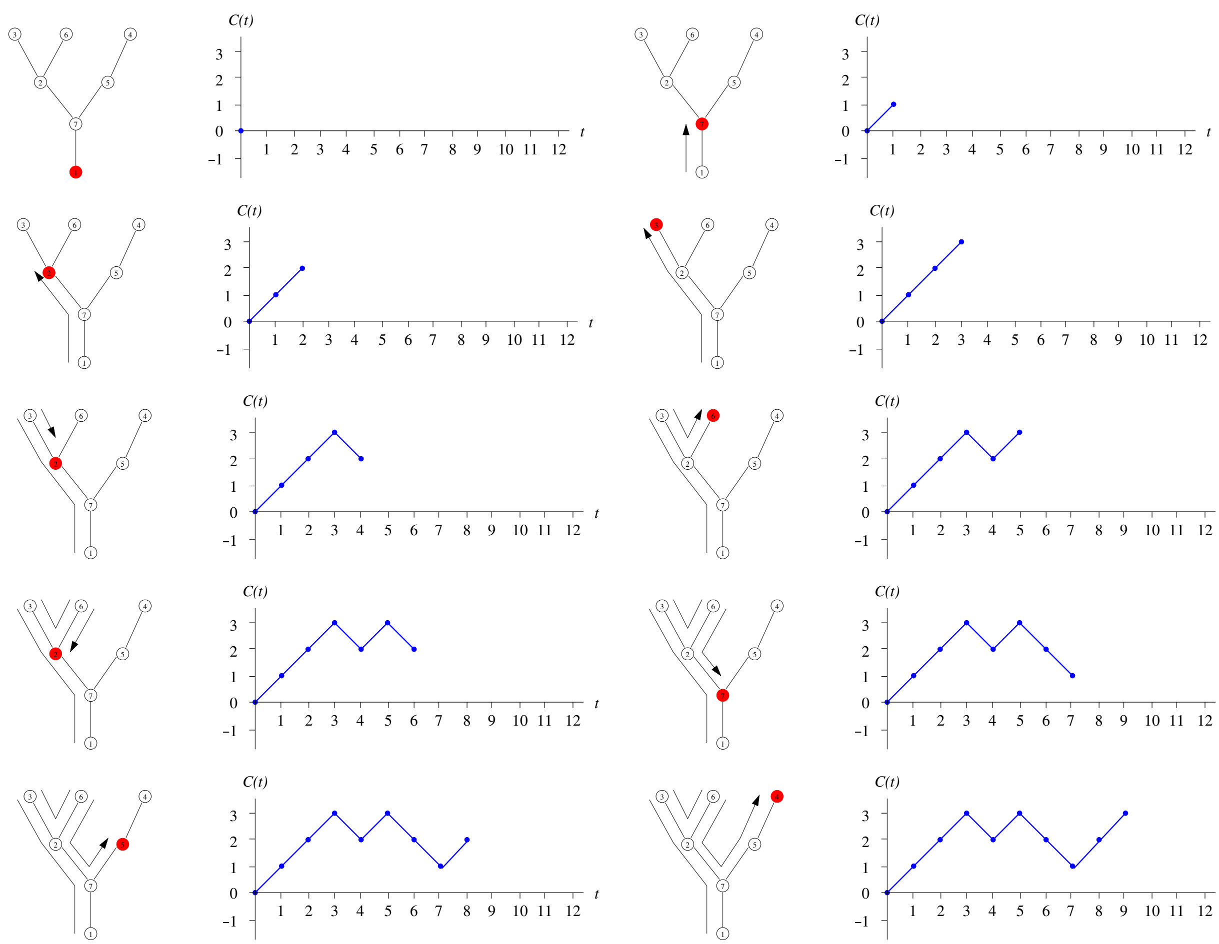
$$\mathbb{P}(X(T) > 0) = \nu(-1)$$

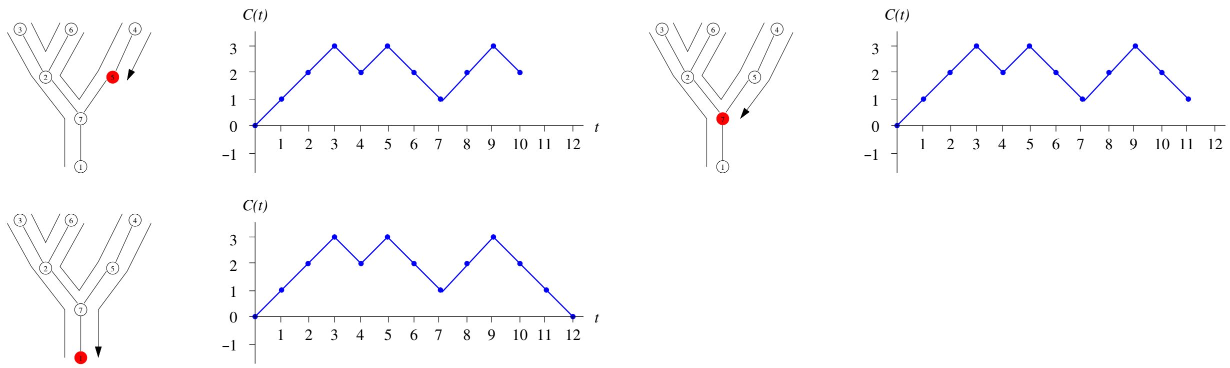
and hence that $\mathbb{P}(X(T) = k) = \bar{\nu}(k)$ for $k \geq 0$.
Hint: you may want to use a probability generating function.

(e) Finally, show that $\mathbb{E}[X(T)] = \sigma^2/2$.

This calculation is inspired by one by Jean-François Marckert & Abdelkader Mokkadem, **The depth first processes of Galton-Watson trees converge to the same Brownian excursion**, *Annals of Probability* **31** (2003), pp.1655-1678. *They credit the argument to Feller.*

7. **(Contour function and convergence to the BCRT)** In lectures, 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 $\mathbf{t} \in \mathbf{T}$, we imagine a particle tracing the outline of the tree from left to right at speed 1. (In the example below, we have an unordered labelled tree, but use the canonical ordering given by the vertex labels to think of it as an ordered tree.)





Let T_n be a conditioned Galton-Watson tree with offspring distribution $p(k) = \left(\frac{1}{2}\right)^{k+1}, k \geq 0$, as in Question 4(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

$$d_{\mathrm{GH}}(T_n, \tilde{T}_n) \leq \frac{1}{2}.$$

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 $f : [0,1] \to \mathbb{R}_+$ and $g : [0,1] \to \mathbb{R}_+$ which encode $\mathbb{R}$-trees $\mathcal{T}_f$ and $\mathcal{T}_g$. For $t \in [0,1]$, let $p_f(t)$ be the image of t in the tree $\mathcal{T}_f$ and similarly for $p_g(t)$. Now define a correspondence

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

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

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

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 will want to use Skorokhod's representation theorem in order to work on a probability space where the convergence (*) occurs almost surely.

A beautiful exposition of this approach is given by Jean-François Le Gall & Grégory Miermont, **Scaling limits of random trees and planar maps**, Lecture notes for the Clay Mathematical Institute Summer School in Buzios, July 11 - August 7, 2010, available at `http://perso.ens-lyon.fr/gregory.miermont/Cours_Buzios.pdf`.

8. **(Dirichlet distribution and gamma random variables)**

   (a) Show that if
   $$(D_1, D_2, \ldots, D_n) \sim \text{Dir}(a_1, a_2, \ldots, a_n)$$
   and $G \sim \text{Gamma}(\sum_{i=1}^n a_i, 1)$ are independent then
   $$G \times (D_1, D_2, \ldots, D_n) \stackrel{d}{=} (G_1, G_2, \ldots, G_n),$$
   where $G_1 \sim \text{Gamma}(a_1, 1), G_2 \sim \text{Gamma}(a_2, 1), \ldots, G_n \sim \text{Gamma}(a_n, 1)$ are independent.

   (b) Show, moreover, that
   $$\left(\frac{G_1}{\sum_{i=1}^n G_i}, \frac{G_2}{\sum_{i=1}^n G_i}, \ldots, \frac{G_n}{\sum_{i=1}^n G_i} \right) \stackrel{d}{=} (D_1, D_2, \ldots, D_n)$$
   and is independent of $\sum_{i=1}^n G_i \sim \text{Gamma}(\sum_{i=1}^n a_i, 1)$.

9. **(A Mexican flag variant of Pólya's urn)** Suppose that you have an urn initially containing a green ball, a white ball and a red ball. At each time-step, pick a ball uniformly at random from the urn, look at its colour, and put it back into the urn with two extra balls of the same colour. Let $N_1(k), N_2(k), N_3(k)$ denote the number of green, white and red balls respectively at step $k \geq 0$, so that $(N_1(0), N_2(0), N_3(0)) = (1, 1, 1)$. Observe that at step k, there are $2k + 3$ balls in the urn. Write $\mathbf{N}(k) = (N_1(k), N_2(k), N_3(k))$.

   (a) Show that $\left(\frac{1}{2k+3}\mathbf{N}(k), k \geq 0\right)$ is a (vector-valued) martingale (i.e. $\left(\frac{1}{2k+3}N_i(k), k \geq 0\right)$ is a martingale for $i = 1, 2, 3$).

   (b) Deduce that
   $$\frac{1}{2k+3}\mathbf{N}(k) \to \mathbf{P} \text{ a.s. as } k \to \infty,$$
   for some non-negative random vector $\mathbf{P} = (P_1, P_2, P_3)$ such that $P_1 + P_2 + P_3 = 1$.

   (c) For $k \geq 1$, let B_k denote the index of the colour picked at step k. Show that for $n_1, n_2, n_3 \geq 0$ such that $n_1 + n_2 + n_3 = n$,
   $$\mathbb{P}\left(B_1 = \cdots = B_{n_1} = 1, B_{n_1+1} = \cdots = B_{n_1+n_2} = 2, B_{n_1+n_2+1} = \cdots = B_n = 3\right)$$
   $$= \frac{(2n_1)!(2n_2)!(2n_3)!(n+1)!}{n_1!n_2!n_3!(2n+2)!}.$$

(d) Show that for a fixed sequence $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in \{1,2,3\}^n$ and any permutation $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$,

$$\mathbb{P}\left(B_1 = b_1, B_2 = b_2, \ldots, B_n = b_n\right) = \mathbb{P}\left(B_{\sigma(1)} = b_1, B_{\sigma(2)} = b_2, \ldots, B_{\sigma(n)} = b_n\right).$$

This says that the random variables $(B_1, \ldots, B_n)$ *are* exchangeable.

(e) Deduce that

$$\mathbb{P}\left(N_1(n) = 2n_1 + 1, N_2(n) = 2n_2 + 1, N_3(n) = 2n_3 + 1\right) = \frac{\binom{2n_1}{n_1}\binom{2n_2}{n_2}\binom{2n_3}{n_3}}{(n+1)\binom{2n+2}{n+1}}.$$

(f) Now fix $x_1, x_2, x_3 \geq 0$ such that $x_1 + x_2 + x_3 = 1$. Using Stirling's formula, show that

$$\begin{aligned} &\mathbb{P}\left(N_1(n) = 2\lfloor nx_1 \rfloor + 1, N_2(n) = 2\lfloor nx_2 \rfloor + 1, N_3(n) = 2\lfloor nx_3 \rfloor + 1\right) \\ &\quad \sim C n^{-2} x_1^{-1/2} x_2^{-1/2} x_3^{-1/2}, \end{aligned}$$

where C is a constant.

(g) Deduce that $\mathbf{P} \sim \mathrm{Dir}(1/2, 1/2, 1/2)$.

For a fascinating account of exchangeability and its consequences, see the classic paper by J. F. C. Kingman, **Uses of exchangeability**, *Annals of Probability*, **6**(2), pp.183–197 (1978).

10. (**Recursive distributional equations**)

(a) Although you perhaps didn't think of it like this, you are probably already familiar with one recursive distributional equation: fix $\lambda \in (0,1)$ and let X be a real-valued random variable such that

$$X \stackrel{d}{=} \sqrt{\lambda} X_1 + \sqrt{1-\lambda} X_2,$$

where X_1 and X_2 are independent copies of X. Show that $X \sim \mathrm{N}(0, \sigma^2)$ is a solution for any $\sigma^2 > 0$.

(b) Suppose that the integer-valued random variable X satisfies the recursive distributional equation

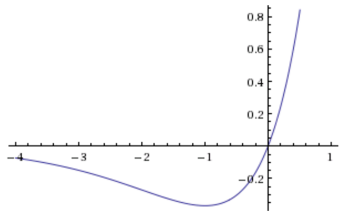
$$X \stackrel{d}{=} 1 + \sum_{i=1}^{P} X_i, \qquad (1)$$

where $X_1, X_2, \ldots$ are i.i.d. copies of X, independent of P which has Poisson(λ) distribution for some $\lambda \in (0, 1]$.

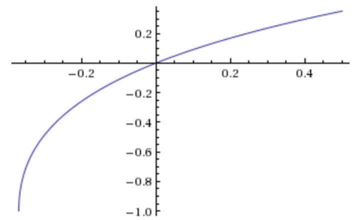
(i) Let $G(s) = \mathbb{E}\left[s^X\right]$, $s \in [-1, 1]$, be the probability generating function of X. Show that

$$G(s) = s \exp(\lambda G(s) - \lambda). \qquad (2)$$

Let $f(x) = xe^x$, $x \in \mathbb{R}$. The inverse relation of f (considered as a complex function) is the so-called *Lambert W function*. Since f is not injective, there are many solutions, and so W is multivalued (except at 0). If we are only interested in $x \geq -1$, however, we see that there is a unique inverse, which is usually called W_0. From the graph, you can see that $W_0(y) \geq -1$ and $W_0(y)$ is defined for $y \geq -1/e$.



$f(x) = xe^x$, $x \in \mathbb{R}$.



$W_0(y)$, $y \geq -1/e$.

Returning now to the equation (2), we see that we can rewrite it as

$$-\lambda G(s) \exp(-\lambda G(s)) = -\lambda s e^{-\lambda}$$

and so

$$G(s) = -\frac{1}{\lambda} W_0(-\lambda s e^{-\lambda}), \quad s \in [-1, 1].$$

So there is a unique solution to (1), and this is its probability generating function.

(ii) W_0 has series expansion $W_0(t) = \sum_{n=1}^{\infty} \frac{(-n)^{n-1} t^n}{n!}$, with radius of convergence $1/e$. Use this to find the probability mass function of X.

(iii) What is the relationship to your answer to Question 5? Why?

For much more about recursive distributional equations, see the survey paper by David Aldous & Antar Bandyopadhyay, **A survey of max-type recursive distributional equations**, *Annals of Probability* **15**(2) (2005) pp.1047–1110.