# Distributional Analysis of Recursive Algorithms and Random Trees

**Problem 1:** Consider a sequence of random variables $(X_n)_{n \geq 0}$ with

$$X_n \stackrel{d}{=} X_{I_n} + n, \quad n \geq 1,$$

where $I_n$ is binomial $B(n, p)$ distributed with $0 < p < 1$ and independent of $X_0, \ldots, X_n$. Assume that $X_0 = 0$.

1. Try to find a rescaling $Y_n = (X_n - \mu(n))/\sigma(n)$ that leads to a limit equation.
   *Hint:* The function $\mu(n) = (1-p)^{-1} n$ will work, and the normal approximation of the binomial distribution is helpful.
2. Try to find a solution of the limit equation.
3. What does the scaling and the moments of that solution suggest for the expectation and variance of $X_n$?

**Problem 2:** Consider the space requirements $X_n$ of random $m$-ary search trees (cf. slides). It is known that, for a certain range of $m$, we have, as $n \to \infty$,

$$\mathbb{E}\, X_n = \frac{1}{2(\mathcal{H}_m - 1)} n + o(\sqrt{n}), \quad \mathrm{Var}(X_n) = \sigma_m n + o(n), \tag{1}$$

with a constant $\sigma_m > 0$ depending only on $m$. Use this information to rescale and to compute the limit equation.
*Hint:* The strong law of large numbers is helpful.

**Problem 3:** The number $X_n$ of recursive calls of Quickselect when selecting the smallest element from a list of $n$ elements satisfies $X_0 = X_1 = 0$ and

$$X_n \stackrel{d}{=} X_{I_n} + 1, \quad n \geq 2,$$

where $I_n$ is uniformly distributed on $\{0, \ldots, n-1\}$ and independent of $X_0, \ldots, X_{n-1}$. Prove (or believe) that, as $n \to \infty$,

$$\mathbb{E}\, X_n = \log n + O(1), \quad \mathrm{Var}(X_n) = \log n + O(1).$$

Derive the limit equation for the rescaled quantities.

**Problem 4:** Assume that $(A_1, \dots, A_K, b)$ are $L^2$-integrable random variables with $\mathbb{E}\, b = 0$ and consider

$$T : \mathcal{M}_2(0) \to \mathcal{M}_2(0), \quad \mu \mapsto \mathcal{L}\left(\sum_{r=1}^{K} A_r Z^{(r)} + b\right),$$

where $(A_1, \dots, A_K, b), Z^{(1)}, \dots, Z^{(K)}$ are independent and $\mathcal{L}(Z^{(r)}) = \mu$ for $r = 1. \dots, K$.

Prove that for all $\mu, \nu \in \mathcal{M}_2(0)$,

$$\ell_2(T(\mu), T(\nu)) \le \Big(\sum_{r=1}^{K} \mathbb{E}\, A_r^2\Big)^{1/2} \ell_2(\mu, \nu).$$

**Problem 5:** Prove that $\zeta_s$ is finite on $\mathcal{M}_s(M_1, \dots, M_m) \times \mathcal{M}_s(M_1, \dots, M_m)$ for every $s > 0$ by showing that

$$\zeta_s(\mathcal{L}(X), \mathcal{L}(Y)) \le \frac{\Gamma(1+\alpha)}{\Gamma(1+s)} (\mathbb{E}\, |X|^s + \mathbb{E}\, |Y|^s)$$

for all $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(M_1, \dots, M_m)$.
*Hint:* For $f \in \mathcal{F}_s$ use Taylor expansion as follows:

$$\begin{aligned} f(x) &= \sum_{j=0}^{m-1} \frac{f^{(j)}(0)}{j!} x^j + \int_0^1 \frac{(1-u)^{m-1} x^m}{(m-1)!} f^{(m)}(xu)\, du \\ &= \sum_{j=0}^{m} \frac{f^{(j)}(0)}{j!} x^j + \int_0^1 \frac{(1-u)^{m-1} x^m}{(m-1)!} (f^{(m)}(xu) - f^{(m)}(0))\, du. \end{aligned}$$

**Problem 6:** Assume that $X_1, \dots, X_K, Y_1, \dots, Y_K$ are independent random variables all having distributions in $\mathcal{M}_s(M_1, \dots, M_m)$. Prove that

$$\zeta_s\left(\sum_{r=1}^{K} X_r, \sum_{r=1}^{K} Y_r\right) \le \sum_{r=1}^{K} \zeta_s(X_r, Y_r).$$

**Problem 7:** Consider the sequence $(X_n)$ from Problem 1. Apply the general convergence theorem in $\mathcal{M}_p$ to verify all conjectures made in Problem 1.

**Problem 8:** Consider the space requirements $X_n$ of random $m$-ary search trees as in Problem 2. The expansions in (1) are true for $3 \le m \le 26$. Apply "A useful extension" to derive a limit law.