Branching Distributional Equations and their Applications

Mariana Olvera-Cravioto

UNC Chapel Hill

molvera@unc.edu

August 20th, 2018

A first example of a distributional equation

- Consider a small ice-cream cart serviced by a single person.
- Customers arrive according to a renewal process having i.i.d. interarrival times $\{\tau_i : i \geq 1\}$, i.e., τ_i is the time between the $(i-1)$th and ith customers.
- Customers are served in the order in which they arrive.
- Let χ_i denote the service time of the ith customer, i.e., the time it takes her/him to place their order and pay.
- Assume the $\{\chi_i : i \geq 1\}$ are i.i.d. and independent of the $\{\tau_i\}$.
- Let W_n denote the waiting time, prior to being served, of the nth customer.

A first example of a distributional equation

- Consider a small ice-cream cart serviced by a single person.
- Customers arrive according to a renewal process having i.i.d. interarrival times $\{\tau_i : i \geq 1\}$, i.e., τ_i is the time between the $(i-1)$th and ith customers.
- Customers are served in the order in which they arrive.
- Let χ_i denote the service time of the ith customer, i.e., the time it takes her/him to place their order and pay.
- Assume the $\{\chi_i : i \geq 1\}$ are i.i.d. and independent of the $\{\tau_i\}$.
- Let W_n denote the waiting time, prior to being served, of the nth customer.
- Then,

$$W_{n+1} = \max\{0, W_n + \chi_n - \tau_{n+1}\}$$

The Lindley equation

- Let $X_n = \chi_n - \tau_{n+1}$ and assume that the ice-cream cart starts with no customers at time zero, then:

$$W_1 = 0, \qquad W_2 = \max\{0, X_1\}, \qquad W_3 = \max\{0, X_2, X_1 + X_2\},$$

and in general,

The Lindley equation

- Let $X_n = \chi_n - \tau_{n+1}$ and assume that the ice-cream cart starts with no customers at time zero, then:

$$W_1 = 0, \qquad W_2 = \max\{0, X_1\}, \qquad W_3 = \max\{0, X_2, X_1 + X_2\},$$

and in general,

$$\begin{aligned} W_{n+1} &= \max\{0, X_n, X_n + X_{n-1}, \dots, X_n + \cdots + X_1\} \\ &\stackrel{D}{=} \max\{0, X_1, X_1 + X_2, \dots, X_1 + \cdots + X_n\} \end{aligned}$$

- It follows that if $E[X_1] < 0$ there exists $0 \leq W < \infty$ such that

$$W_n \Rightarrow W, \qquad n \to \infty.$$

- We now obtain the well-known Lindley equation:

$$W \stackrel{D}{=} \max\{0, W + X\},$$

where X has the same distribution as X_1.

The Galton-Watson process

- Consider an individual who during his lifetime has a random number of offspring distributed according to F.
- Each of his offspring has a number of of offspring also distributed according to F, independently of everything else.
- Let Z_n denote the number of individuals in the nth generation of this process, with $Z_0 = 1$ representing the original individual.
- $\{Z_n : n \geq 1\}$ is called a **Galton-Watson process**.
- If we let $\{N_i : i \geq 0\}$ be i.i.d. with common distribution F, then

$$Z_{n+1} \stackrel{D}{=} \sum_{i=1}^{Z_n} N_i \stackrel{D}{=} \sum_{j=1}^{N_0} Z_{n,j},$$

where the $\{Z_{n,j} : j \geq 1\}$ are i.i.d. copies of Z_n independent of N_0.

A branching distributional equation

- Provided $\mu = E[N_1] < \infty$, $M_n = Z_n/\mu^n$ is a nonnegative martingale, and therefore there exists a limit $M < \infty$ such that

$$M_n \to M, \qquad \text{a.s. as } n \to \infty.$$

- And from the second distributional identity we obtain that M must satisfy

$$M \stackrel{D}{=} \frac{1}{\mu} \sum_{j=1}^{N} M_j,$$

where N has distribution F and the $\{M_j : j \geq 1\}$ are i.i.d. copies of M independent of N.

- This equation is a special case of the **homogeneous smoothing transform**.

Towards more general equations

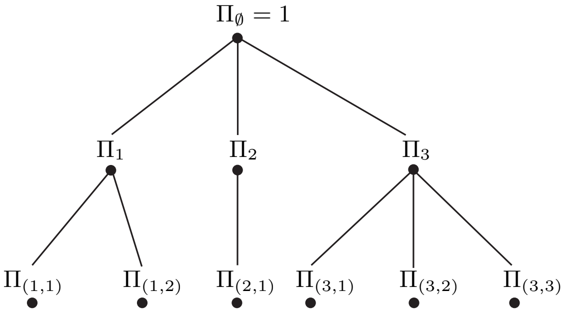
- In order to describe more general branching distributional equations we introduce first the **weighted branching process**.
- Consider a Galton-Watson process as before.
- To track individuals in the family, label the original individual as $\emptyset$, and give each individual in the kth generation, $k \geq 1$, a label of the form $(i_1, i_2, \ldots, i_k)$.
- Let $N_{(i_1,\ldots,i_k)}$ denote the number of offspring that individual $(i_1, \ldots, i_k)$ has, define $A_0 = 1$, and for $k \geq 1$:

$$A_k = \left\{(i_1, \ldots, i_k) : (i_1, \ldots, i_{k-1}) \in A_{k-1}, 1 \leq i_k \leq N_{(i_1,\ldots,i_{k-1})}\right\}.$$

- Note that $Z_n = |A_n|$.
- Each individual $\mathbf{i} = (i_1, \ldots, i_k)$ is also assigned a mark $(N_{\mathbf{i}}, Q_{\mathbf{i}}, C_{(\mathbf{i},1)}, \ldots C_{(\mathbf{i},N_{\mathbf{i}})})$, where $(\mathbf{i}, j) = (i_1, \ldots, i_k, j)$.
- Assume that the marks $\left\{(N_{\mathbf{i}}, Q_{\mathbf{i}}, C_{(\mathbf{i},1)}, \ldots C_{(\mathbf{i},N_{\mathbf{i}})})\right\}$ are i.i.d.

Weighted branching trees

- The weighted branching tree $\mathcal{T}$:



- Each node in the tree has a weight $\Pi_{(i_1,\ldots,i_n)}$ defined via the recursion

$$\Pi_{i_1} = C_{i_1}, \qquad \Pi_{(i_1,\ldots,i_n)} = C_{(i_1,\ldots,i_n)}\Pi_{(i_1,\ldots,i_{n-1})}, \qquad n \geq 2,$$

and $\Pi = 1$ is the weight of the root node.

The linear equation or smoothing transform

- Define the total sum of the weights up to the kth generation:

$$R^{(k)} = \sum_{j=0}^{k} \sum_{\mathbf{i} \in A_j} \Pi_{\mathbf{i}} Q_{\mathbf{i}}$$

- By looking at the whole tree up to the kth generation in terms of the subtrees of depth $k-1$ that are rooted in A_1 we obtain that

$$R^{(k)} = Q_{\emptyset} + \sum_{j=1}^{N_{\emptyset}} C_j R_j^{(k-1)},$$

where the $R_j^{(k-1)}$ are the sums of the weights of the subtrees whose roots are the nodes in A_1.

The linear equation or smoothing transform

- Assuming that $R^{(k)} \Rightarrow R$ as $k \to \infty$ for some finite r.v. R, we obtain the **non-homogeneous smoothing transform**:

$$R \stackrel{D}{=} Q + \sum_{j=1}^{N} C_j R_j,$$

  where the $\{R_j\}$ are i.i.d. copies of R, independent of the vector $(N, Q, \{C_j\})$.

- Similarly, it can be shown that if everything is non-negative and $E\left[\sum_{j=1}^{N} C_j\right] = 1$ then $W^{(k)} = \sum_{\mathbf{i} \in A_k} \Pi_{\mathbf{i}} Q_{\mathbf{i}} \to W$ as $k \to \infty$ and W solves the **homogeneous smoothing transform**:

$$W \stackrel{D}{=} \sum_{j=1}^{N} C_j W_j$$

The maximum equation or high-order Lindley equation

- Assume that the marks $\{(N_{\mathbf{i}}, Q_{\mathbf{i}}, C_{(\mathbf{i},1)}, \ldots C_{(\mathbf{i},N_{\mathbf{i}})})\}$ are non-negative.
- Now define the maximum weight up to the kth generation:

$$R^{(k)} = \bigvee_{j=0}^{k} \bigvee_{\mathbf{i} \in A_j} \Pi_{\mathbf{i}} Q_{\mathbf{i}}$$

- The same analysis as before gives that

$$R^{(k)} = Q_{\emptyset} \vee \bigvee_{j=1}^{N_{\emptyset}} C_j R_j^{(k-1)},$$

where the $\{R_j^{(k-1)} : j \geq 1\}$ are i.i.d. with the same distribution as $R^{(k-1)}$.

The maximum equation or high-order Lindley equation

- Assuming again that $R^{(k)} \Rightarrow R$ as $k \to \infty$ for some finite r.v. R, we obtain the **maximum equation**:

$$R \stackrel{D}{=} Q \vee \bigvee_{j=1}^{N} C_j R_j,$$

where the $\{R_j\}$ are i.i.d. copies of R, independent of the vector $(N, Q, \{C_j\})$.

- Taking logarithm on both sides and setting $Q \equiv 1$, $W = \log R$, and $X_j = \log C_j$, we obtain the **high-order Lindley equation**:

$$W \stackrel{D}{=} \max \left\{ 0, \max_{1 \le j \le N} (X_j + W_j) \right\}$$

The general case

- We can consider branching distributional equations of the form

$$R \stackrel{\mathcal{D}}{=} \Phi\left(Q, N, \{C_i\}, \{R_i\}\right),$$

where the $\{R_i\}$ are i.i.d. copies of R, independent of $(Q, N, \{C_i\})$, and Φ is a deterministic map.

- In general, the solutions to these type of equations are NOT unique!
- For all our examples, it is easy to show that a sufficient condition for a solution to exist is that there exists a $0 < \beta \leq 1$ such that $E\left[\sum_{j=1}^{N} |C_i|^{\beta}\right] < 1$, in which case

$$R = \sum_{j=0}^{\infty} \sum_{\mathbf{i} \in A_j} \Pi_{\mathbf{i}} Q_{\mathbf{i}} \qquad \text{and} \qquad R = \bigvee_{j=0}^{\infty} \bigvee_{\mathbf{i} \in A_j} \Pi_{\mathbf{i}} Q_{\mathbf{i}}$$

solve the linear and maximum equations, resp.

- These solutions are called known as the **attracting endogenous solutions**, however, infinitely many more solutions may exist.

Other examples

- Discounted tree sums:

$$R \stackrel{D}{=} Q + \bigvee_{i=1}^{N} C_i R_i$$

- Ising model:

$$R \stackrel{D}{=} Q + \sum_{i=1}^{N} \mathsf{atanh}(\mathsf{tanh}(\beta)\mathsf{tanh}(R_i))$$

A first open problem

- Can you find explicit choices for the distribution of $(Q, N, \{C_i\})$ for which the distribution of either

$$R = \sum_{j=0}^{\infty} \sum_{\mathbf{i} \in A_j} \Pi_{\mathbf{i}} Q_{\mathbf{i}} \qquad \text{and} \qquad R = \bigvee_{j=0}^{\infty} \bigvee_{\mathbf{i} \in A_j} \Pi_{\mathbf{i}} Q_{\mathbf{i}}$$

  can be explicitly computed?
- *Hint:* The non-branching case of the maximum equation is Lindley's equation, which does have a very well-known explicit special case. Can you think of any for the branching case?

Computing the distribution of the solutions

- Computing analytically the distribution of the attracting endogenous solution to branching distributional equations is usually impossible.
- Hence, we need to find numerical ways of doing so.
- Note that when the existence of a solution to the equation

$$R \stackrel{D}{=} \Phi\left(Q, N, \{C_i\}, \{R_i\}\right)$$

can be established, we can expect to find a solution by iterating the equation, i.e., by constructing

$$R^{(k)} \stackrel{D}{=} \Phi\left(Q, N, \{C_i\}, \{R_i^{(k-1)}\}\right)$$

starting with some initial distribution for $\{R_i^{(0)}\}$.

In terms of measures and trees

- Let T be the map that takes the probability measure ν to the probability measure $T(\nu)$ describing the distribution of

$$\Phi\left(Q, N, \{C_i\}, \{X_i\}\right),$$

  where the $\{X_i\}$ are i.i.d. and distributed according to ν, independent of $(Q, N, \{C_i\})$.
- Then, the iteration process described earlier corresponds to taking ν_0 to be the distribution of the $\{R_i^{(0)}\}$ and $\nu_k = T(\nu_{k-1})$ be the distribution of $R^{(k)}$.
- In terms of weighted branching trees, the process is also equivalent to constructing a weighted tree of depth k and iteratively compute $R^{(k)}$ starting at the leaves and moving upwards to the root.
- The latter leads to a **naïve Monte Carlo approach**. (Why naïve?)

A naïve Monte Carlo approach

- In many cases we can show that there exists $0 < c < 1$ and $p \geq 1$ such that
$$E\left[\left|R - R_{\emptyset}^{(k)}\right|^p\right] = O\left(c^k\right).$$
- Hence, it suffices to compute the distribution of $R^{(k)} = R_{\emptyset}^{(k)}$ for large k.
- In principle, we could simulate a tree for k generations (k large) and use
$$R \approx R^{(k)}.$$
- For the linear and maximum recursion this would imply computing
$$R^{(k)} = \sum_{j=0}^{k} \sum_{\mathbf{i} \in A_j} \Pi_{\mathbf{i}} Q_{\mathbf{i}} \qquad \text{and} \qquad R^{(k)} = \bigvee_{j=0}^{k} \bigvee_{\mathbf{i} \in A_j} \Pi_{\mathbf{i}} Q_{\mathbf{i}}.$$

A naïve Monte Carlo approach

- In many cases we can show that there exists $0 < c < 1$ and $p \geq 1$ such that
$$E\left[\left|R - R_{\emptyset}^{(k)}\right|^{p}\right] = O\left(c^{k}\right).$$
- Hence, it suffices to compute the distribution of $R^{(k)} = R_{\emptyset}^{(k)}$ for large k.
- In principle, we could simulate a tree for k generations (k large) and use
$$R \approx R^{(k)}.$$
- For the linear and maximum recursion this would imply computing
$$R^{(k)} = \sum_{j=0}^{k} \sum_{\mathbf{i} \in A_j} \Pi_{\mathbf{i}} Q_{\mathbf{i}} \qquad \text{and} \qquad R^{(k)} = \bigvee_{j=0}^{k} \bigvee_{\mathbf{i} \in A_j} \Pi_{\mathbf{i}} Q_{\mathbf{i}}.$$
- **Problem:** This approach requires an average of $(E[N])^{k}$ i.i.d. copies of $(Q, N, \{C_i\})$!