Max-type Recursive Distributional Equations and Associated Recursive Tree Processes

Antar Bandyopadhyay

(Joint work with Professor David J. Aldous)

(Work done at UC, Berkeley, Dept. of Statistics)

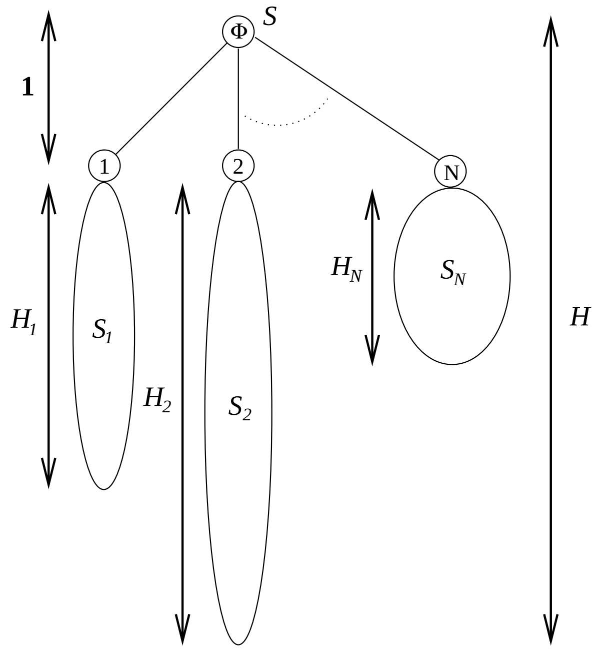
Institute for Mathematics and Its Applications
University of Minnesota

(http://www.ima.umn.edu/~antar)

November 7, 2003

An Easy Examples

Consider a *(sub)-critical* Galton-Watson branching process with the progeny distribution N, so $\mathbf{E}[N] \leq 1$; we assume $\mathbf{P}(N = 1) < 1$.



Height of the Tree : Let $H :=$ 1+ height of the G-W tree, then $H < \infty$ a.s., and

$$H \stackrel{d}{=} 1 + \max(H_1, H_2, \ldots, H_N) \quad \text{on} \quad \mathbb{N},$$

where $(H_j)_{j \geq 1}$ are i.i.d. with same law as of H and are independent of N.

We will call such an equation a *Recursive Distributional Equation* (RDE).

Typical features of RDE

$$X \stackrel{d}{=} 1 + \max(X_1, X_2, \ldots, X_N) \quad \text{on } \mathbb{N}.$$

- **Unknown :** Distribution of X.
- **Known :** The distribution of N (which may or may not be random) and some manipulation (which may be random but then with known distribution) of known and unknown quantities !
- **What RDE is doing :** To find a distribution μ such that when we take i.i.d. samples from it and only use N many of them (where N is independent of the samples) and do the manipulation then we end up with another sample from μ.

Remark : In the case $N = 1$ a.s. it reduces to the question of finding a stationary distribution of a Markov chain.

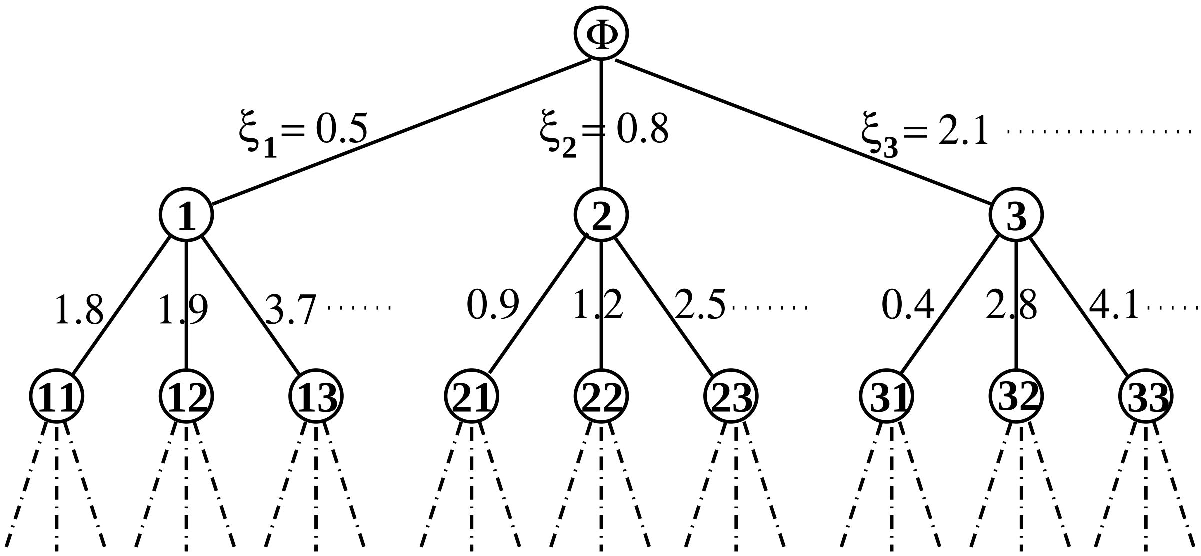
Two main uses of RDEs

- **Direct use :** The RDE is used directly to define a distribution. Examples include
  - ▶ The height (and also the size) of a (sub)-critical Galton-Watson tree (our first example).
  - ▶ The Quicksort distribution.
  - ▶ Discounted tree sums / inhomogeneous percolation on trees.

- **Indirect use:** The RDE is used to define some auxiliary variables which help in defining/characterizing some other quantity of interest. Various different short of applications, to name a few
  - ▶ 540° *argument* ! (will give an example).
  - ▶ Determining critical points and scaling laws (will not give an example).

Example of a 540° argument Optimal Matching Problem on PWIT

PWIT (**P**oisson **W**eighted **I**nfinite **T**ree) **:**



For the edges $(\mathbf{i}, \mathbf{i}j)$ where $j \geq 1$ the weights are points of Poisson point process of rate 1 on $(0, \infty)$ written as $\left(\xi_{\mathbf{i}j}\right)_{j \geq 1}$. The processes are independent as $\mathbf{i}$ varies.

Problem : Find a matching (*invariant*) which is optimal in the sense that it minimizes the average edge weight.

Remark : This problem in some sense (can be made rigorous) is the limit of the random assignment problem on the complete graph K_n with i.i.d. Exponential(n) edge-weights. [Aldous, 1992, 2001].

540° argument

- **Step 1 :** For each vertex $\mathbf{i}$ of the **PWIT**, let $\mathbf{T}^{\mathbf{i}}$ be the infinite tree rooted at $\mathbf{i}$ containing only its descendents. We define the quantities
  - $W_{\mathbf{i}} :=$ *Total* weight of optimal matching on $\mathbf{T}^{\mathbf{i}}$.
  - $\widetilde{W}_{\mathbf{i}} :=$ *Total* weight of optimal matching on $\mathbf{T}^{\mathbf{i}} \setminus \{\mathbf{i}\}$.
  - $X_{\mathbf{i}} := W_{\mathbf{i}} - \widetilde{W}_{\mathbf{i}}$. **Note :** $X_{\mathbf{i}} = \infty - \infty$!
- **Step 2 :** Assuming these quantities make sense one can write the following *recursion*

$$X_{\emptyset} = \min_{j \geq 1} (\xi_j - X_j)$$

  where $(X_j)_{j \geq 1}$ are i.i.d. with same law as of $X_{\emptyset}$, and are independent of $(\xi_j)_{j \geq 1}$ which are points of a Poisson point process of rate 1 on $(0, \infty)$.

- **Step 3 :** One can show [Aldous, 2001]
  - The RDE is well defined and has unique solution as the *Logistic distribution*. We will call this RDE the *Logistic* RDE.
  - Now we can reconstruct (rigorously) the optimal matching using the variables $X_{\mathbf{i}}$. For example, match root $\emptyset$ with $\arg\min_{j \geq 1} (\xi_j - X_j)$.

General Setup

- Let $(S, \mathfrak{S})$ be a measurable space, and $\mathcal{P}$ be the collection of all probabilities on $(S, \mathfrak{S})$.

- Let (ξ, N) be a pair of random variables such that N takes values in $\{0, 1, 2, \ldots; \infty\}$.

- Let $(X_j)_{j \geq 1}$ be **i.i.d** S-valued random variables, which are independent of (ξ, N).

- $g(\cdot)$ is a S-valued measurable function with appropriate domain.

Recursive Distributional Equation (RDE)

Definition 1 *The following* fixed-point equation *on* $\mathcal{P}$ *is called a* Recursive Distributional Equation (RDE)

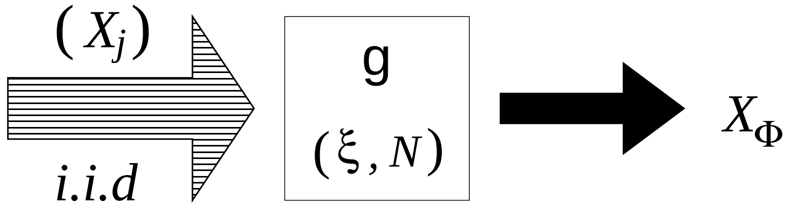
$$X \stackrel{d}{=} g\left(\xi; X_j, 1 \leq j \leq^* N\right), \quad \textit{on} \;\; S$$

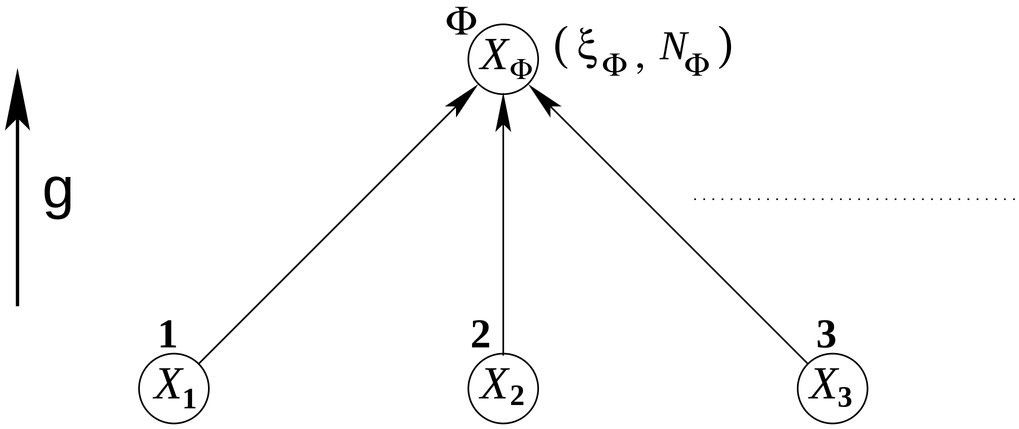
where $(X_j)_{j \geq 1}$ *are i.i.d. with same law as of* X *and are independent of* (ξ, N).

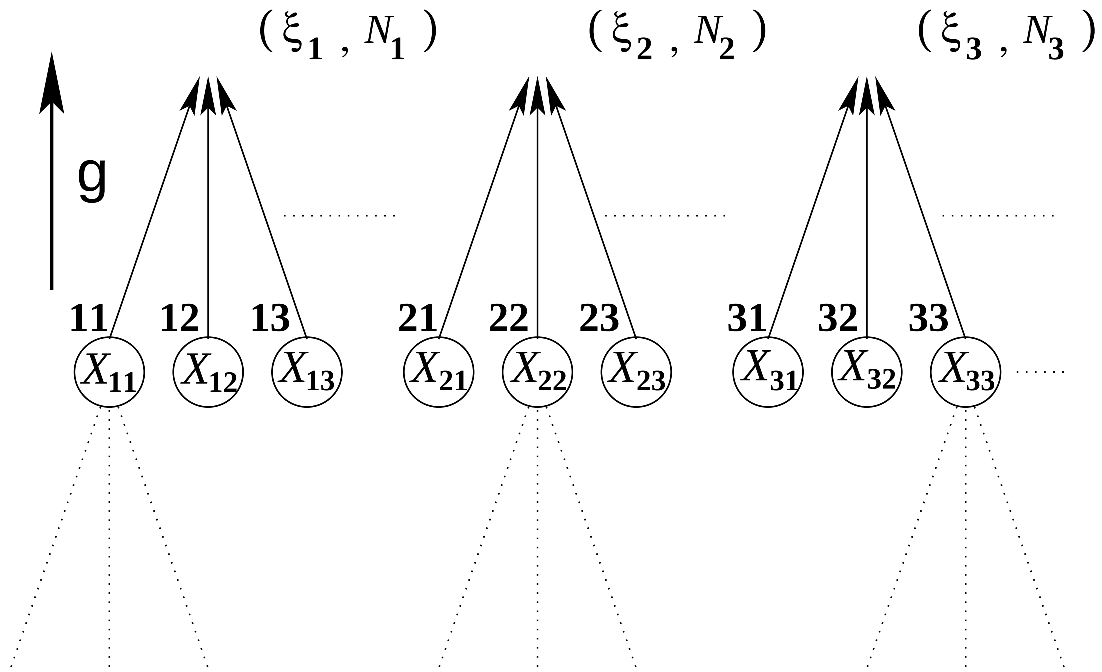
Remark : A more conventional (analysis) way of writing the equation would be

$$\mu = T(\mu)$$

where T is the operator associated with the above equation, which depends on the function g and the joint distribution of the pair (ξ, N), and μ is the (unknown) law of X.

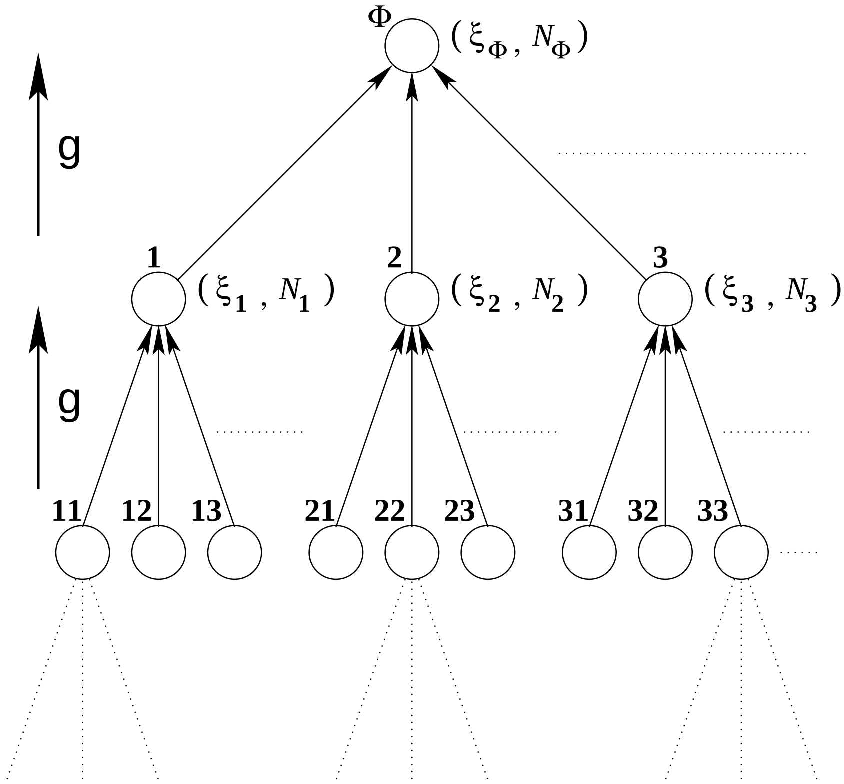



Φ
X_Φ
(ξ_Φ, N_Φ)
g
1
2
3
X_1
X_2
X_3



(ξ 1 , N 1)
(ξ 2 , N 2)
(ξ 3 , N 3)
g
11
12
13
21
22
23
31
32
33
X11
X12
X13
X21
X22
X23
X31
X32
X33

Recursive Tree Framework (RTF)

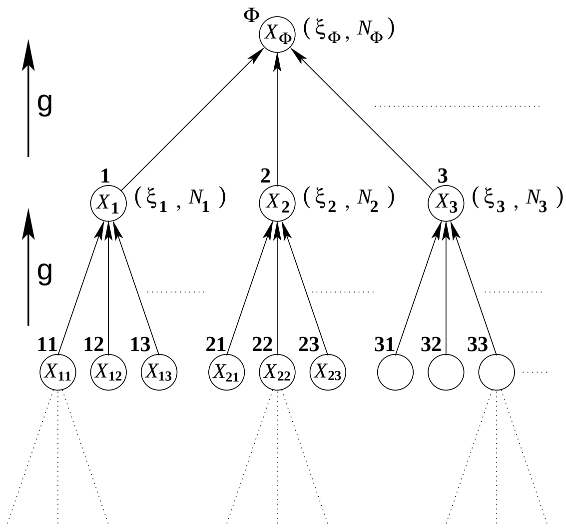


- **Skeleton :** $\mathbb{T}_\infty := (\mathcal{V}, \mathcal{E})$ is the canonical infinite tree with vertex set $\mathcal{V} := \left\{ \mathbf{i} \,|\, \mathbf{i} \in \mathbb{N}^d,\ d \geq 1 \right\} \cup \{\emptyset\}$, and edge set $\mathcal{E} := \{ e = (\mathbf{i}, \mathbf{i}j) \,|\, \mathbf{i} \in \mathcal{V}, j \in \mathbb{N} \}$, and root $\emptyset$.

- **Innovations :** Collection of **i.i.d** pairs $\{(\xi_{\mathbf{i}}, N_{\mathbf{i}}) \,|\, \mathbf{i} \in \mathcal{V}\}$.

- **Function :** The function $g(\cdot)$.

Recursive Tree Process (RTP)



Consider a **RTF** and let μ be a solution of the associated **RDE** . A collection of S-valued random variables $(X_{\mathbf{i}})_{\mathbf{i}\in\mathcal{V}}$ is called an invariant *Recursive Tree Process (RTP)* with marginal μ if

- $X_{\mathbf{i}} \sim \mu \;\; \forall\, \mathbf{i} \in \mathcal{V}$.
- $X_{\mathbf{i}} = g\left(\xi_{\mathbf{i}}; X_{\mathbf{i}j}, 1 \leq j \leq^* N_{\mathbf{i}}\right) \;\; \forall\, \mathbf{i} \in \mathcal{V}$.
- $X_{\mathbf{i}}$ is independent of $\{(\xi_{\mathbf{i}'}, N_{\mathbf{i}'}) \mid |\mathbf{i}'| < |\mathbf{i}|\,\}$, for all $\mathbf{i} \in \mathcal{V}\backslash\{\emptyset\}$, where $|\mathbf{i}| = d$ if $\mathbf{i} \in \mathbb{N}^d$.

Remark : Using *Kolmogorov's consistency*, an invariant RTP with marginal μ exists if and only if μ is a solution of the associated RDE.

Endogeny

Natural Question : Does $X_\emptyset$ only depend on the innovation process (the *data*) $(\xi_{\mathbf{i}}, N_{\mathbf{i}})_{\mathbf{i}\in\mathcal{V}}$?

Definition 2 *Let $\mathcal{G}$ be the σ-field generated by the* innovation process $\{\,(\xi_{\mathbf{i}}, N_{\mathbf{i}})\,|\,\mathbf{i}\in\mathcal{V}\,\}$*. We will say an invariant RTP is* endogenous *if $X_\emptyset$ is $\mathcal{G}$-measurable.*

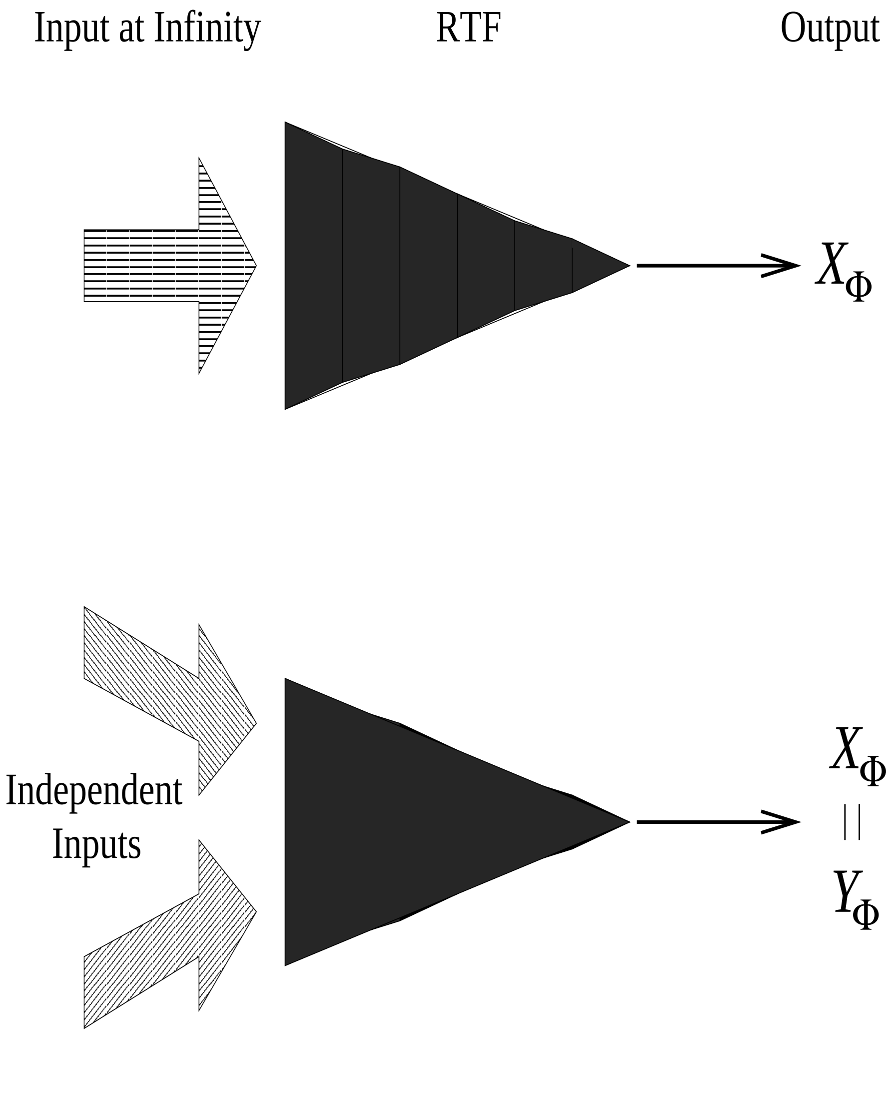
Motivations

- Presence / absence of *external* randomness.
- Influence of the boundary at infinity !
- Relation with *long-range independence* (Gamarnik et al, 2003).

One easy fact to built our confidence

Remark : Associated with a RTF there is a Galton-Watson branching process tree rooted at $\emptyset$ defined only through $\{N_{\mathbf{i}} \,|\, \mathbf{i} \in \mathcal{V}\}$, call it $\mathcal{T}$. Essentially any associated invariant RTP lives on $\mathcal{T}$.

Proposition 1 *If $\mathcal{T}$ is almost surely finite (equivalently $\mathbf{E}[N] \leq 1$) then the associated RDE has unique solution and the RTP is endogenous.*


Input at Infinity
RTF
Output
X_Φ
Independent
Inputs
X_Φ
||
Y_Φ

Bivariate Uniqueness

Consider the following **bivariate RDE**,

$$\begin{pmatrix} X \\ \\ Y \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} g\left(\xi; X_j, 1 \leq j \leq^* N\right) \\ \\ g\left(\xi; Y_j, 1 \leq j \leq^* N\right) \end{pmatrix}$$

where $(X_j, Y_j)_{j\geq 1}$ are i.i.d and has the same law as of (X, Y), and are independent of the innovation (ξ, N).

Definition 3 *An invariant RTP with marginal μ has* **bivariate uniqueness** *property if the above* bivariate RDE *has unique solution as $X = Y$ a.s on the space of joint probabilities with both marginals μ.*

An Equivalence Theorem

Theorem 1 *Suppose the* S *is a* Polish *space. Consider an invariant RTP with marginal distribution* μ.

(a) If the endogenous property holds then the bivariate uniqueness property holds.

(b) Conversely, (under some technical conditions) if the bivariate uniqueness property holds and then the endogenous property holds.

(c) If $T^{(2)}$ *be the operator associated with the* bivariate RDE *then endogenous property holds if and only if*

$$T^{(2)^n}(\mu \otimes \mu) \xrightarrow{d} \mu^{\nearrow},$$

where $\mu \otimes \mu$ *is the product measure, and* $\mu^{\nearrow}$ *is the measure concentrated on the diagonal with both marginal* μ.

Remark : Similar kind of result appears in the study of Gibbs measures and Markov random fields.

Back to Logistic RDE

- Recall the *Logistic RDE* associated with the optimal matching problem on **PWIT**,

$$X \stackrel{d}{=} \min_{j \geq 1} \left(\xi_j - X_j \right)$$

where $(\xi_j)_{j \geq 1}$ are points of a rate 1 Poisson point process on $(0, \infty)$.

- Also recall that it has *unique* solution as the *Logistic distribution*, given by

$$\mathbf{P}\left(X \leq x\right) = \frac{1}{1 + e^{-x}}, \;\; x \in \mathbb{R}.$$

Theorem 2 *The invariant* Logistic RTP *has bivariate uniqueness property, that is, the following* bivariate RDE *has* unique *solution as $X = Y$ a.s with* Logistic *marginal*

$$\begin{pmatrix} X \\ \\ Y \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} \min\limits_{j \geq 1}(\xi_j - X_j) \\ \\ \min\limits_{j \geq 1}(\xi_j - Y_j) \end{pmatrix}$$

where $(\xi_j)_{j \geq 1}$ are points of a rate 1 *Poisson point process on* $(0, \infty)$*, and $(X_j, Y_j)_{j \geq 1}$ are* i.i.d *with same joint distribution as of (X, Y), and are independent of $(\xi_j)_{j \geq 1}$.*

Corollary 2.1 *The invariant* Logistic RTP *is endogenous.*

Outline of the proof of Theorem 2

- The marginals of X and Y satisfy the Logistic RDE, so $X \stackrel{d}{=} Y$, and $X \sim$ Logistic distribution.

- Note that $(\xi_j; (X_j, Y_j))_{j \geq 1}$ is a Poisson point process on $(0, \infty) \times \mathbb{R}^2$, with mean intensity $dt\, \nu(d(x, y))$, where $\nu = \text{Law}(X, Y)$. Let $G(x, y) := \mathbf{P}(X > x, Y > y)$, then

$$G(x, y) = \overline{H}(x)\, \overline{H}(y) \exp\left(\int_0^\infty G(t - x, t - y)\, dt \right),$$

where $\overline{H}(x) = e^{-x}/(1 + e^{-x})$ is the tail for the Logistic distribution function.

- To prove $X = Y$ a.s, it is enough to show that $X \wedge Y \stackrel{d}{=} X$.

- Let $g(x) := G(x, x) = \mathbf{P}(X \wedge Y > x)$, then

$$g(x) = \overline{H}^2(x) \exp\left(\int_{-x}^\infty g(s)\, ds \right).$$

It is enough to prove that $g = \overline{H}$ is the unique solution.

$$g(x) = \overline{H}^2(x)\ \exp\left(\int_{-x}^{\infty} g(s)\, ds\right)$$

- Define $\mathfrak{F} := \left\{ f : \mathbb{R} \to [0,1] \mid \overline{H}^2(x) \le f(x) \le \overline{H}(x) \right\}$, and $T : \mathfrak{F} \to \mathfrak{F}$ as
$$T(f)(x) := \overline{H}^2(x)\ \exp\left(\int_{-x}^{\infty} f(s)\, ds\right).$$

- Notice that, $g \in \mathfrak{F}$ and $g = T(g)$.

- It is enough to show that T has unique fixed point as $\overline{H}$ on $\mathfrak{F}$.

- Define a natural partial order, say, " $\preccurlyeq$ " on $\mathfrak{F}$ as $f \preccurlyeq h$ iff $f(x) \le h(x),\ \forall\, x \in \mathbb{R}$.

- T is a monotone operator on $(\mathfrak{F}, \preccurlyeq)$, that is, $T(f) \preccurlyeq T(h) \iff f \preccurlyeq h$.

- Let $f_0 := \overline{H}^2$, define recursively $f_{n+1} := T(f_n)$.

- Observe that $f_n \preccurlyeq g \preccurlyeq \overline{H}$, so it is enough to prove that $f_n(x) \to \overline{H}(x)$ pointwise.

- Using induction one can show

$$f_n(x) = \overline{H}(x) \exp\left(-\beta_{n-1}\left(\overline{H}(x)\right)\right), \; n \geq 1;$$

where

$$\beta_n(s) = \int_s^1 \frac{1}{w}\left(1 - e^{-\beta_{n-1}(1-w)}\right) dw, \; n \geq 1,$$

with $\beta_0(s) = 1 - s$.

- Note that $\beta_n(1) = 0, \; \forall \, n$.

- It is enough to show that $\beta_n(x) \to 0$ pointwise.

$$\beta_n(s) = \int_s^1 \frac{1}{w}\left(1 - e^{-\beta_{n-1}(1-w)}\right) dw, \ n \geq 1$$

- Easy calculation shows that $\beta_n(s) \downarrow$ pointwise. Let $L(s) := \lim_{n\to\infty} \beta_n(s)$.

- $L(1) = 0$ and L satisfy the integral equation

$$L(s) = \int_s^1 \frac{1}{w}\left(1 - e^{-L(1-w)}\right) dw.$$

- Enough to show that $L \equiv 0$.

- Consider $\eta(w) := (1-w)e^{L(1-w)} + we^{-L(w)} - 1$, then it is easy to see that $\eta(0) = \eta(1) = 0$, and further

$$\eta'(w) = e^{-L(w)}\left[2 - \left(e^{L(1-w)} + e^{-L(1-w)}\right)\right] \leq 0.$$

Thus $\eta \equiv 0 \iff L \equiv 0$.