Endogeny for Recursive Tree Processes : Application to Quicksort RDE

Antar Bandyopadhyay

(Joint work with Prof. David J. Aldous)

[Work done at UC, Berkeley and IMA, Minneapolis]

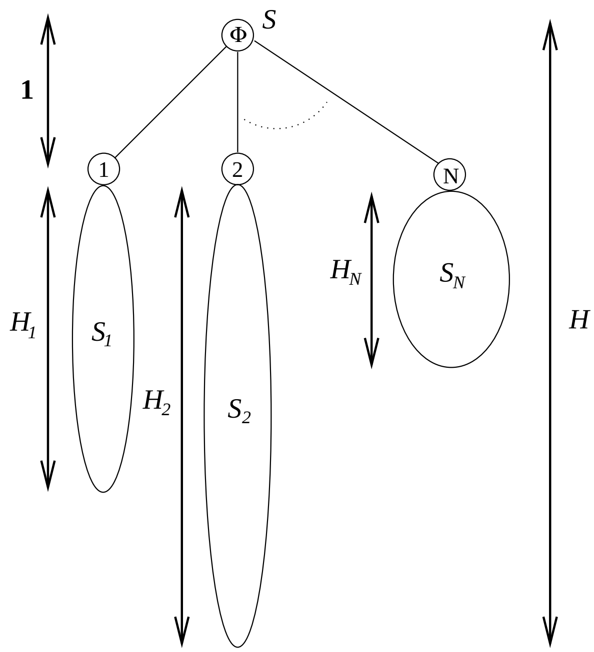
Kiel-Göteborg Workshop on Probability and Combinatorics

Department of Mathematics
Chalmers University of Technology
Götoborg, Sweden

November 13, 2004

Three Examples

Examples 1 (Height of a GW-Branching Tree) : 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 equation a *Recursive Distributional Equations* (RDE).

Example 2 (Quicksort Algorithm/Distribution) :

- Select the first number from a pile of n numbers and divide the other $(n-1)$ numbers into two piles, according to *less* than or *bigger* than the first number.

- Recursively sort the two piles (which are now smaller in size).

- $X(n) :=$ # comparisons needed to sort n numbers starting from a uniform random permutation of $[n]$. Then

$$X(n) \stackrel{d}{=} X_1(U_n) + X_2(n-1-U_n) + (n-1),$$

where $X_1(\cdot)$ and $X_2(\cdot)$ are i.i.d. with same law as of $X(\cdot)$ and are independent of U_n which is uniform on $\{0, 1, 2, \ldots, n-1\}$.

- Rösler (1990) showed $\mathbf{E}[X(n)] \sim 2n \log n$ and moreover

$$\frac{X(n) - 2n \log n}{n} \xrightarrow{d} Y,$$

where distribution of Y satisfies the RDE

$$Y \stackrel{d}{=} UY_1 + (1-U)Y_2 + C(U) \quad \text{on } \mathbb{R},$$

where Y_1 and Y_2 are i.i.d. with same law as of Y and are independent of $U \sim$ Uniform$[0,1]$, and $C(u) := 1 + 2u \log u + 2(1-u) \log(1-u)$.

Examples 3 (Worst-Case Time of `FIND`) :

$$T \stackrel{d}{=} 1 + \max\left(UT_1, (1-U)T_2\right) \quad \text{on} \ \ \mathbb{R}_+$$

where (T_1, T_2) are i.i.d. copies of T and are independent of $U \sim \text{Uniform}[0,1]$.

- Studied by Grübel and Rösler (1996) and Devroye (2001).
- It gives the asymptotic distribution of the number of comparisons needed for the worst case of the `FIND` algorithm of Hoare (1961) after scaling.
- It has unique solution, which has all moments finite, and supported on $[2, \infty)$.

Typical features of RDEs

$$\text{Ex. 1}: \quad X \overset{d}{=} 1 + \max(X_1, X_2, \ldots, X_N) \quad \text{on } \mathbb{N}$$
$$\text{Ex. 2}: \quad X \overset{d}{=} UX_1 + (1-U)X_2 + C(U) \quad \text{on } \mathbb{R}$$
$$\text{Ex. 3}: \quad X \overset{d}{=} 1 + \max(UX_1, (1-U)X_2) \quad \text{on } \mathbb{R}_+$$

- **Unknown Quantity :** Distribution of X.
- **Known Quantities :**
  - $N \leq \infty$ which may or may not be random (e.g. $N \equiv 2$ in Ex. 2 & 3).
  - Possibly some more randomness whose distribution is known (e.g. U in the Ex. 2 & 3).
  - How we combine the known and unknown randomness (e.g. "$1 + \max$" operation in Ex. 1).
- **What is the RDE doing ?** To find a distribution μ such that when we take i.i.d. samples $(X_j)_{j\geq 1}$ 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 $X \sim \mu$.

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

Two main uses of RDEs

- **Direct use :** The RDE is used directly to define a distribution. Examples include,
  - ▶ The height of a (sub)-critical Galton-Watson tree (Ex. 1).
  - ▶ The Quicksort distribution (Ex. 2).
  - ▶ Discounted tree sums / inhomogeneous percolation on trees (Ex. 3 is a special case).
  - ▶ ... *and many others*.

- **Indirect use:** The RDE is used to define some auxiliary variables which help in defining/characterizing some other quantity of interest. Among others the following two type of applications are of special interest (*but we will not discuss these in this talk*),
  - ▶ 540° *argument* !
  - ▶ Determining critical points and scaling laws.

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)

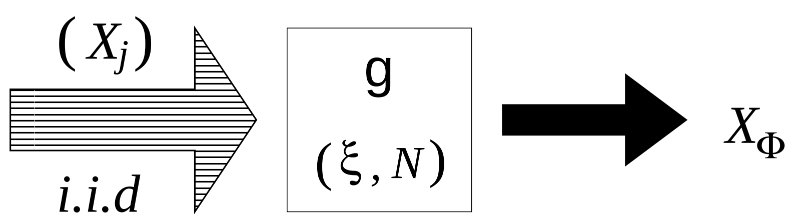
$$X \stackrel{d}{=} g\left(\xi; X_j, 1 \leq j \leq^* N\right) \quad \textit{on} \quad S,$$

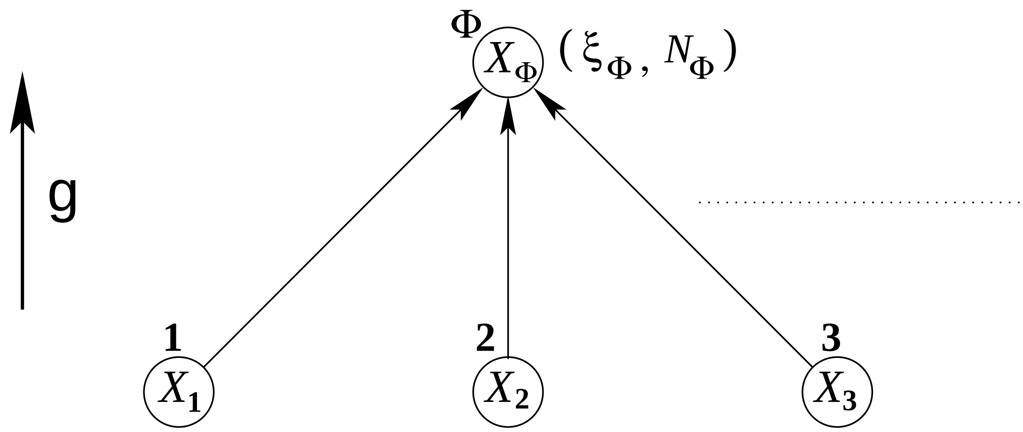
where $(X_j)_{j\geq 1}$ *are independent copies of* X *and are independent of* (ξ, N)*.*

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

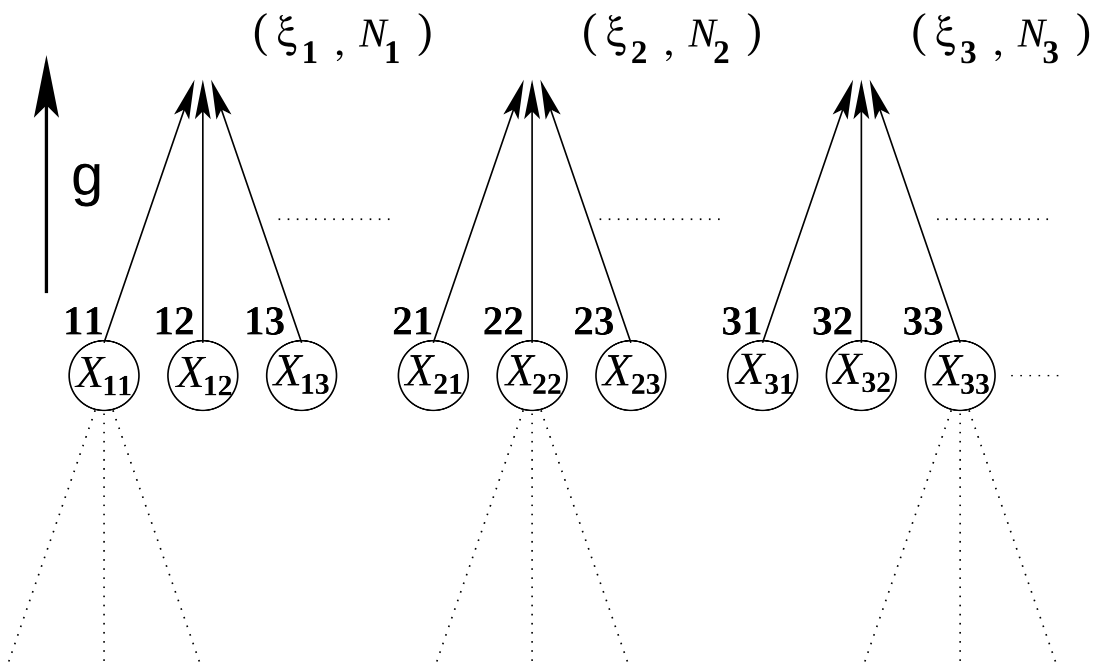
$$\mu = \mathbf{T}(\mu)$$

where $\mathbf{T}$ 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.



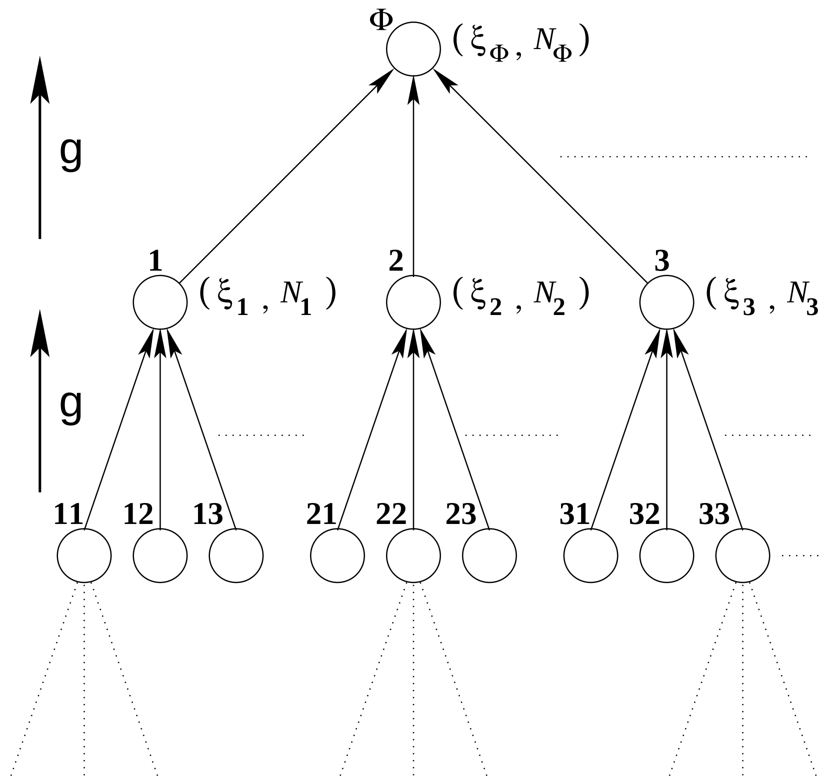


g
Φ
X_Φ
(ξ_Φ, N_Φ)
1
2
3
X_1
X_2
X_3



(ξ1 , N1)
(ξ2 , N2)
(ξ3 , N3)
g
11
12
13
21
22
23
31
32
33
X11
X12
X13
X21
X22
X23
X31
X32
X33

Recursive Tree Framework (RTF)

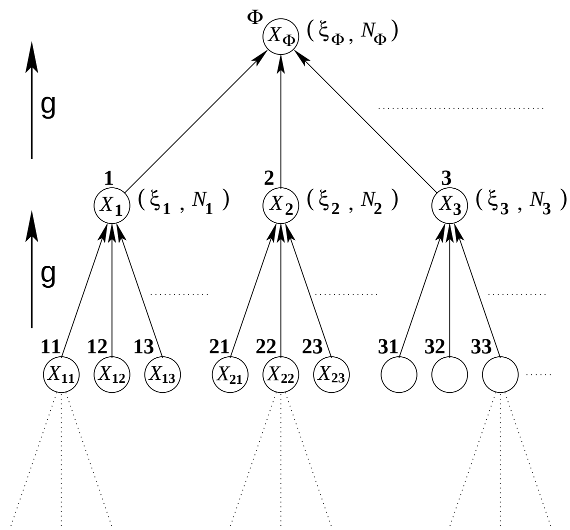


- **Skeleton :** $\mathbb{T}_\infty := (\mathcal{V}, \mathcal{E})$ is the canonical infinite tree with vertex set $\mathcal{V} := \left\{\mathbf{i} \,|\, \mathbf{i} \in \mathbb{N}^d,\, d \geq 1 \right\} \cup \{\emptyset\}$, and edge set $\mathcal{E} := \{e = (\mathbf{i}, \mathbf{i}j) \,|\, \mathbf{i}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

- $X_{\mathbf{i}} \sim \mu \;\; \forall \; \mathbf{i} \in \mathcal{V}$.
- $X_{\mathbf{i}} = g\left(\xi_{\mathbf{i}}; X_{\mathbf{i}j}, 1 \leq j \leq^* N_{\mathbf{i}}\right)$ a.s. $\forall \; \mathbf{i} \in \mathcal{V}$.
- $X_{\mathbf{i}}$ 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 !
- Sometime can be used for characterization of certain solutions (*we will see how this works for Quick-sort distribution*).

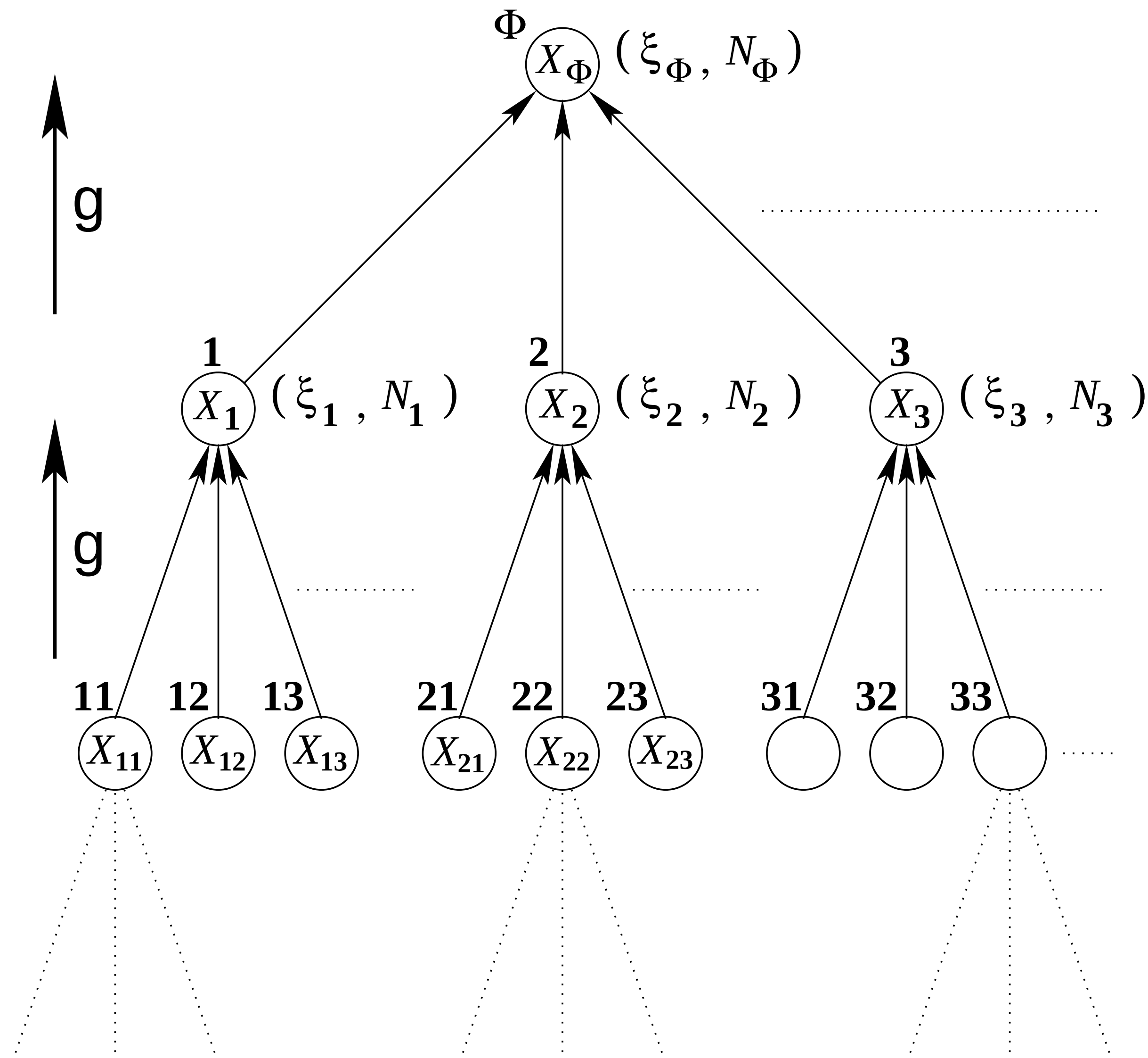
One easy fact to built our confidence

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 $\mathcal{T}$.

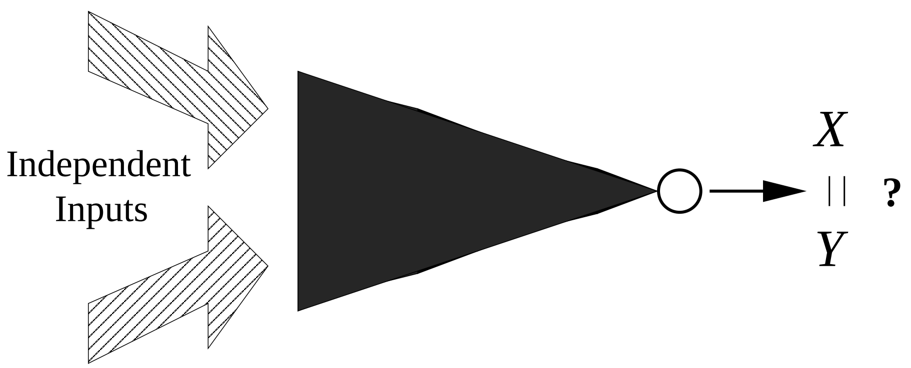
Proposition 1 *If $\mathcal{T}$ is almost surely finite (equivalently $\mathbf{E}[N] \leq 1$) then the associated RDE has unique solution and the RTP is endogenous.*

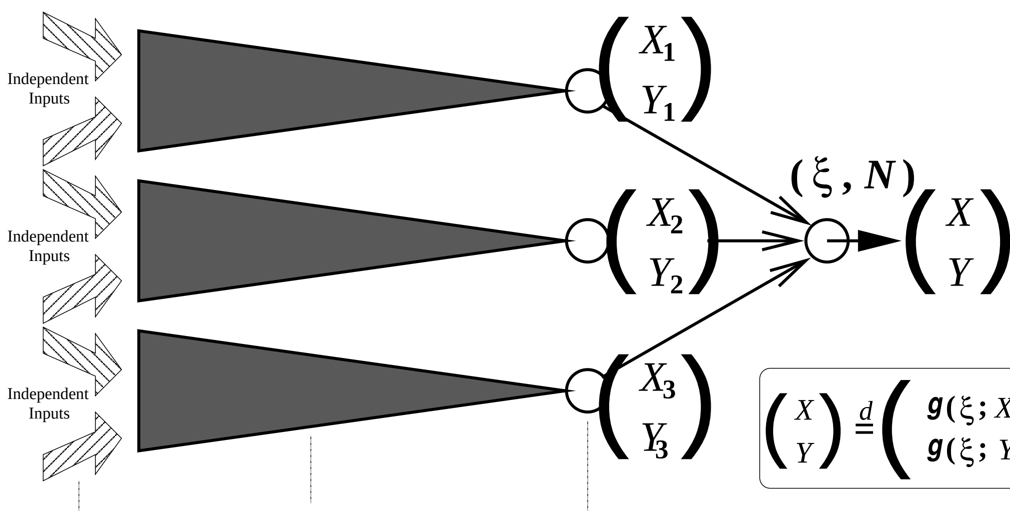
Remarks :

- The RDEs in Ex. 1 have unique solutions and it is endogenous.
- Perhaps the simplest example of a RDE with no non-trivial endogenous solution is the following
$$X \stackrel{d}{=} \frac{X_1 + X_2}{\sqrt{2}}.$$
The solution set is the Normal$(0, \sigma^2)$ family. But the associated RTF has *no randomness* involved and hence none of the non-trivial RTP is endogenous.
- The Quicksort RDE also has binary branching and hence a priory we can not say any thing about uniqueness/endogeny.


Φ
X_Φ
(ξ_Φ, N_Φ)
g
g
1
2
3
X_1
(ξ_1, N_1)
X_2
(ξ_2, N_2)
X_3
(ξ_3, N_3)
11
12
13
21
22
23
31
32
33
X_{11}
X_{12}
X_{13}
X_{21}
X_{22}
X_{23}

Input at Infinity RTF Output





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

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

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

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

Remark : Results of similar type can also be found in the study of Gibbs measures and Markov random fields.

Successful Use and/or Application of Endogeny

- **Characterization :** Sometime one can show that only the "*fundamental*" solution(s) of a RDE is(are) endogenous.
  - ▶ We will show that for the *Quicksort RDE* the limiting *Quicksort* distribution and its translates are the only endogenous solutions.

- 540° **argument :** (*will not discuss these*)
  - ▶ Application to *random assignment problem*.
  - ▶ Application to *frozen percolation* process on infinite regular trees.

Solution Set of the Quicksort RDE

Recall that the *Quicksort RDE* is given by

$$X \stackrel{d}{=} UX_1 + (1-U)X_2 + C(U) \quad \text{on} \quad \mathbb{R},$$

where (X_1, X_2) are i.i.d. copies of X and are independent of $U \sim$ Uniform$[0,1]$, and $C(u) := 1 + 2u \log u + 2(1-u)\log(1-u)$.

Known :

- If X is a solution then so is $(m + X)$ for any $m \in \mathbb{R}$.
- There is a unique solution with $\mathbf{E}[X] = 0$ and $\mathbf{E}\left[X^2\right] < \infty$ [Rösler, 1992].
- Let ν be the solution with mean zero and finite variance then the set of all solutions is given by

  $$\left\{\nu * \text{Cauchy}\left(m, \sigma^2\right) \,\middle|\, m \in \mathbb{R}, \, \sigma^2 \in \mathbb{R}_+ \right\}$$

  [Fill and Janson, 2000]
- Note that the only mean zero solution is ν.

Theorem 2 *A solution of the Quicksort RDE is* endogenous *if and only if* $\sigma^2 = 0$.

Remark : In other words, the solution ν and its translates are the only endogenous solutions.

Proof of Theorem 2

- We will use the bivariate uniqueness technique.

- Let $\mu = \nu * \text{Cauchy}\left(m, \sigma^2\right)$ be a solution of the Quicksort RDE. Consider the bivariate RDE

$$\begin{pmatrix} X \\ Y \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} UX_1 + (1-U)X_2 + C(U) \\ UY_1 + (1-U)Y_2 + C(U) \end{pmatrix},$$

  where $(X_j, Y_j)_{j=1,2}$ are i.i.d. copies of (X, Y) and are independent of $U \sim \text{Uniform}[0,1]$. Further assume $X \stackrel{d}{=} Y \stackrel{d}{=} \mu$.

Proof of the "if"-part

$$\begin{pmatrix} X \\ Y \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} UX_1 + (1-U)X_2 + C(U) \\ UY_1 + (1-U)Y_2 + C(U) \end{pmatrix}$$

- We assume $\sigma^2 = 0$.
- Let $D = X - Y$ and similarly define D_1 and D_2.
- Then $D \stackrel{d}{=} UD_1 + (1-U)D_2$ on $\mathbb{R}$.
- Since $\sigma^2 = 0$, so $X \stackrel{d}{=} Y \stackrel{d}{=} \nu$, thus D has finite second moment.
- Simple calculation then shows $\mathbf{E}[D] = 0 = \mathbf{E}\left[D^2\right]$.
- Thus $X = Y$ a.s., that is, bivariate uniqueness holds.

Proof of the "only if"-part

$$\begin{pmatrix} X \\ Y \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} UX_1 + (1-U)X_2 + C(U) \\ UY_1 + (1-U)Y_2 + C(U) \end{pmatrix}$$

- Suppose $\sigma^2 > 0$.

- We will show that $(Q+Z, Q+W)$ is a solution of the bivariate equation, where Z and W are i.i.d. Cauchy $\left(m, \sigma^2\right)$ and are independent of $Q \sim \nu$.

- Observe that if Z_1 and Z_2 are i.i.d. Cauchy $\left(m, \sigma^2\right)$ and are independent of $U \sim$ Uniform$[0,1]$ then

$$Z = UZ_1 + (1-U)Z_2$$

  is also Cauchy $\left(m, \sigma^2\right)$ and it is independent of U (follows by computing the characteristic function).

- Take $(Z_1, Z_2; W_1, W_2)$ i.i.d. Cauchy $\left(m, \sigma^2\right)$; (Q_1, Q_2) i.i.d. copies of $Q \sim \nu$; and $U \sim$ Uniform$[0, 1]$. All are independent.

- Define $X_j := Q_j + Z_j$ and $Y_j := Q_j + W_j$, $j \in \{1, 2\}$.

- Let $Q := UQ_1 + (1 - U)Q_2 + C(U)$ then $Q \sim \nu$.

- If $Z := UZ_1 + (1 - U)Z_2$ and $W := UW_1 + (1 - U)W_2$ then

$$\begin{array}{ccccc} Q & + & Z & = & UX_1 + (1 - U)X_2 + C(U) \\ Q & + & W & = & UY_1 + (1 - U)Y_2 + C(U) \end{array}$$

- But Z and W are i.i.d. Cauchy $\left(m, \sigma^2\right)$ and are independent of Q.

- Thus $(Q + Z, Q + W)$ is a non-trivial solution of the bivariate RDE and hence bivariate uniqueness fails.