Recursive Distributional Equations and Recursive Tree Processes : Lecture - II

Antar Bandyopadhyay

(Joint work with Professor David J. Aldous)

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

Mini-Workshop on Recursive Distributional Equations
J.W. Goethe Universitt Frankfurt a.M., Germany

Department of Mathematics
Chalmers University of Technology
Gothenburg, Sweden

March 8, 2005

Recursive Distributional Equation (RDE)

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

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

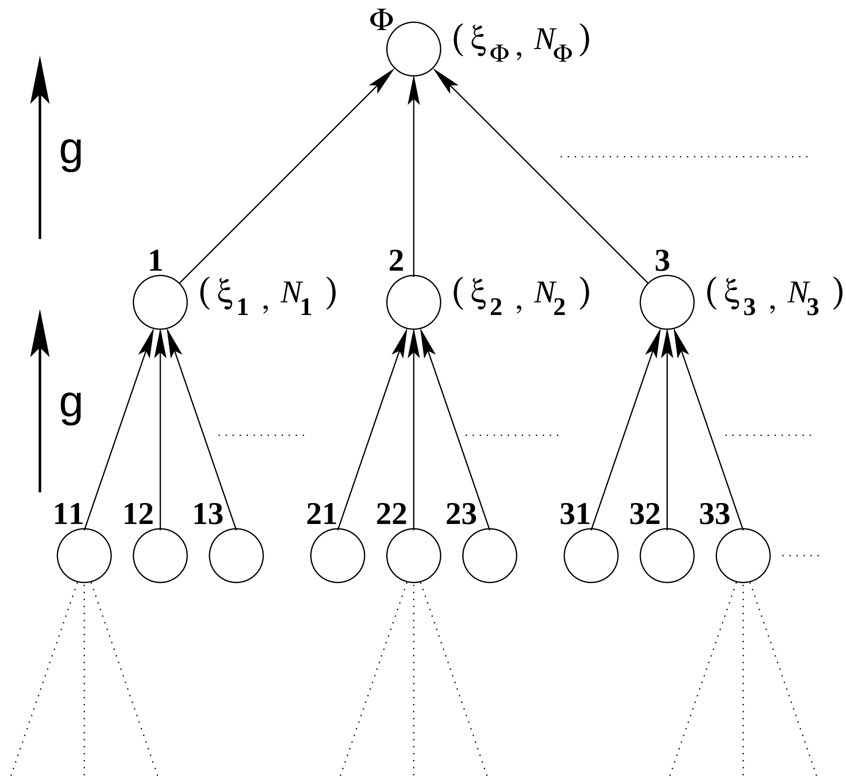
where $(X_j)_{j \geq 1}$ *are independent copies of* X *and are independent of* (ξ, N).

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

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

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.

Recursive Tree Framework (RTF)

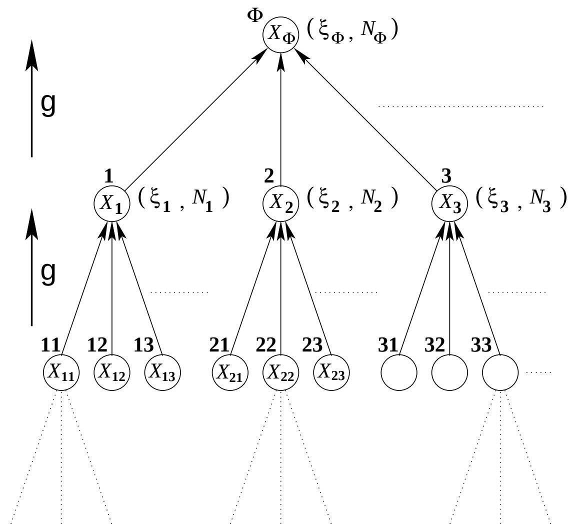


- **Skeleton :** $\mathbb{T}_\infty := (\mathcal{V}, \mathcal{E})$ is the canonical infinite tree with vertex set $\mathcal{V} := \{\mathbf{i} \,|\, \mathbf{i} \in \mathbb{N}^d,\, d \geq 1\} \cup \{\emptyset\}$, and edge set $\mathcal{E} := \{e = (\mathbf{i}, \mathbf{i}j) \,|\, \mathbf{i} \in \mathcal{V},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

- $X_\mathbf{i} \sim \mu \ \forall\ \mathbf{i} \in \mathcal{V}$.
- Fix $d \geq 0$ then $(X_\mathbf{i})_{|\mathbf{i}|=d}$ are independent.
- $X_\mathbf{i} = g\left(\xi_\mathbf{i}; X_{\mathbf{i}j}, 1 \leq j \leq^* N_\mathbf{i}\right)$ a.s. $\forall\ \mathbf{i} \in \mathcal{V}$.
- $X_\mathbf{i}$ is independent of $\{(\xi_{\mathbf{i}'}, N_{\mathbf{i}'}) \mid |\mathbf{i}'| < |\mathbf{i}|\}$ $\forall\ \mathbf{i} \in \mathcal{V}$.

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

Questions

Given a RDE

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

one can ask several questions, such as ...

(i) Does it have a solution ? (*existence*)

(ii) If yes, is it unique ? (*uniqueness*)

(iii) If μ is a solution then for what other measure say ν, do we have $T^n(\nu) \stackrel{d}{\longrightarrow} \mu$? (*domain of attraction*)

(iv) If μ is a solution then is the invariant RTP with marginal μ endogenous ? (*endogeny*)

(v) If μ is a solution the does the invariant RTP with marginal μ has trivial tail ? (*tail triviality of RTP*) [will not discuss this issue in this series]

$\cdots$ *perhaps many more !*

Endogeny

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

Endogeny vs Uniqueness

Example (Uniqueness $\not\Rightarrow$ Endogeny) :

$$X \stackrel{d}{=} \text{sign}\,(X_1) \times \xi \quad \text{on} \;\; \mathbb{R},$$

where $\xi \sim \text{Normal}\,(0,1)$ and is independent of X_1.

- Trivially the unique solution is Normal (0, 1).
- Once again invariant RTP can be indexed by non-negative integers. Let $(X_i)_{i \geq 0}$ be invariant RTP with marginal Normal (0, 1), from definition
$$X_i = \text{sign}\,(X_{i+1}) \times \xi_i \;\; \text{a.s.}$$
Let $Y_i = -X_i$, $i \geq 0$ then
$$Y_i = \text{sign}\,(Y_{i+1}) \times \xi_i \;\; \text{a.s.}$$
Thus $(Y_i)_{i \geq 0}$ is also an invariant RTP with marginal Normal (0, 1). So RTP $(X_i)_{i \geq 0}$ can not be endogenous.

Remarks :

- We will see that *uniqueness* of a "modified" RDE will imply endogeny.
- Later we will see that a RDE may have many solutions some endogenous while others are not.

Endogeny vs The Operator T

RDE - I	RDE- II
$X \stackrel{d}{=} \xi$ on $\mathbb{R}$, where $\xi \sim$ Normal $(0,1)$	$X \stackrel{d}{=}$ sign $(X_1) \times \xi$ on $\mathbb{R}$, where $\xi \sim$ Normal $(0,1)$

Observations :

- Both RDEs define the same operator T, namely mapping every probability measure to Normal $(0,1)$.
- Trivially the invariant RTP for **RDE - I** is endogenous, while that for **RDE - II** is not.
- Note T is a "very nice" function ! (It is continuous, monotone, a contraction ...)

Remark : In order to answer the question of endogeny, we need the whole structure of the RTF. In particular the function g. We will later see [Lecture-III] that if g has some "nice" properties then endogeny will follow for certain solution(s).

Bivariate Uniqueness

Consider the following **bivariate RDE**,

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

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

Note : We will denote the operator associated with this bivariate RDE by $T^{(2)}$.

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.

Uniqueness vs Bivariate Uniqueness

- It is possible to have an RDE which has unique solution but the solution fail to have bivariate uniqueness.

$$X \stackrel{d}{=} \xi + X_1 \pmod 2 \quad \text{on} \quad \{0,1\},$$

where $\xi \sim$ Bernoulli(p) for some $0 < p \leq 1$ and is independent of X_1.

  - ▶ Easy to see that the unique solution is $X \sim$ Bernoulli $\left(\frac{1}{2}\right)$.

  - ▶ It is also not hard to see that, any joint distribution (X, Y) with marginal Bernoulli $\left(\frac{1}{2}\right)$, is a solution of the bivariate RDE.

Remark : Later we will have an example (*Quicksort RDE*) where the RDE has many solutions among which only one has bivariate uniqueness.

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, suppose the bivariate uniqueness property holds. If also $T^{(2)}$ is continuous with respect to the weak convergence on the set of bivariate distributions with marginal μ, then the endogenous property holds.

(c) Further, the 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.

Proof of the Equivalence Theorem

Part (a) (Endogeny $\Rightarrow$ Bivariate Uniqueness) :

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

- ν be a solution of the bivariate RDE with marginal μ.
- Using i.i.d. innovations $(\xi_{\mathbf{i}}, N_{\mathbf{i}})_{\mathbf{i} \in \mathcal{V}}$ construct bivariate RTP $((X_{\mathbf{i}}, Y_{\mathbf{i}}))_{\mathbf{i} \in \mathcal{V}}$ such that $(X_{\mathbf{i}}, Y_{\mathbf{i}}) \sim \nu$ for all $\mathbf{i} \in \mathcal{V}$.

- Observe for every $n \geq 0$,

$$\left(X_{\emptyset}; \left(\left(\xi_{\mathbf{i}}, N_{\mathbf{i}}\right), |\mathbf{i}| \leq n\right)\right) \stackrel{d}{=} \left(Y_{\emptyset}; \left(\left(\xi_{\mathbf{i}}, N_{\mathbf{i}}\right), |\mathbf{i}| \leq n\right)\right).$$

- So if $\Lambda : S \to \mathbb{R}$ is a bounded measurable function then

$$\mathbf{E}\left[\Lambda\left(X_{\emptyset}\right) \mid \mathcal{G}_n\right] = \mathbf{E}\left[\Lambda\left(Y_{\emptyset}\right) \mid \mathcal{G}_n\right] \quad \text{a.s.},$$

where $\mathcal{G}_n$ is the σ-field generated by $\{(\xi_{\mathbf{i}}, N_{\mathbf{i}}) \mid |\mathbf{i}| \leq n\}$.

- By martingale convergence theorem and using the endogeny property we get

$$\Lambda\left(X_{\emptyset}\right) = \Lambda\left(Y_{\emptyset}\right) \quad \text{a.s.}$$

This is true for any bounded measurable function Λ, so

$$X_{\emptyset} = Y_{\emptyset} \quad \text{a.s.}$$

Part (b) : Bivariate Uniqueness $\Rightarrow$ Endogeny :

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

- Fix $\Lambda : S \to \mathbb{R}$, a bounded continuous function.
- Let $(X_{\mathbf{i}})_{\mathbf{i} \in \mathcal{V}}$ be an invariant RTP with marginal μ.
- By martingale convergence theorem

$$\mathbf{E}\left[\Lambda\left(X_\emptyset\right) \mid \mathcal{G}_n\right] \xrightarrow[\mathcal{L}_2]{\text{a.s.}} \mathbf{E}\left[\Lambda\left(X_\emptyset\right) \mid \mathcal{G}\right],$$

where $\mathcal{G}_n$ is the σ-field generated by $\left\{ (\xi_{\mathbf{i}}, N_{\mathbf{i}}) \mid |\mathbf{i}| \leq n \right\}$.

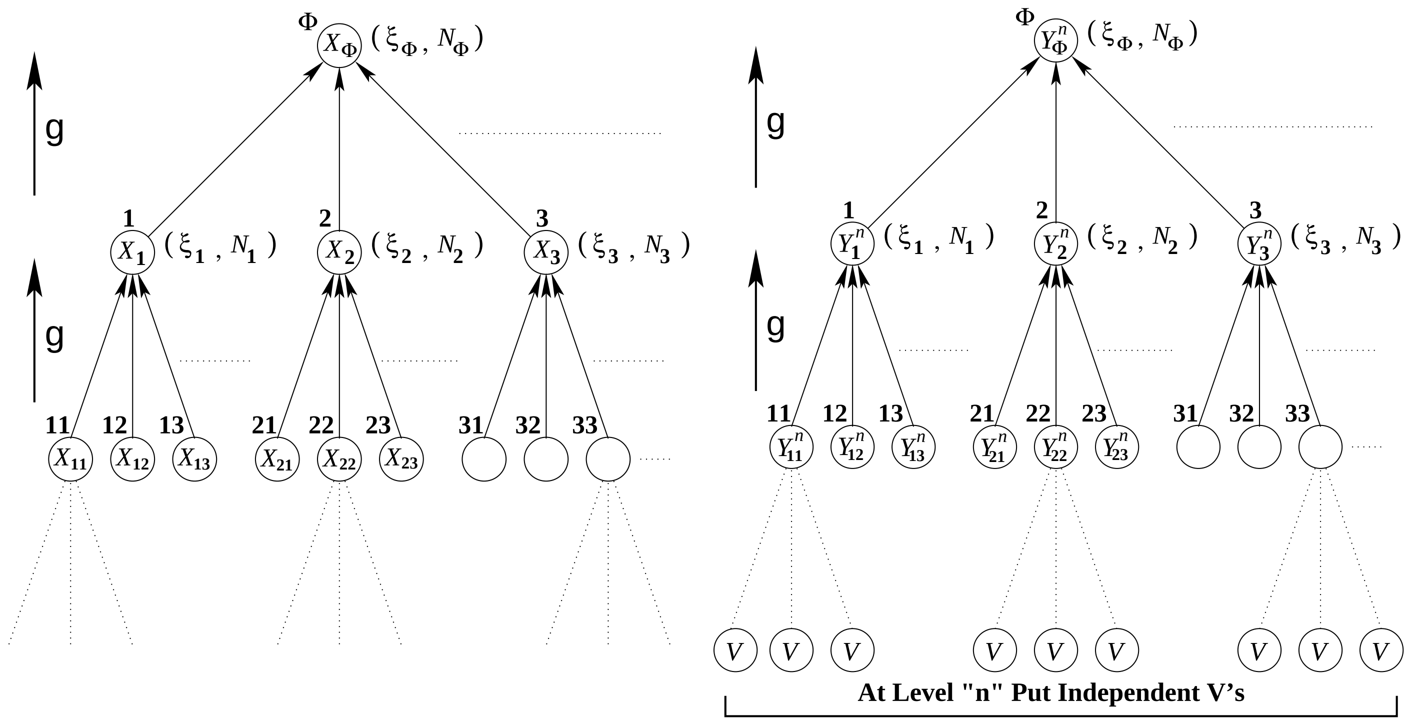
- For each $n \geq 0$ we will construct $Y_\emptyset^n$ as follows :
  - Take $(V_{\mathbf{i}})_{\mathbf{i}\in\mathcal{V}}$ i.i.d. sample from μ which are independent of the innovations $(\xi_{\mathbf{i}}, N_{\mathbf{i}})_{\mathbf{i}\in\mathcal{V}}$ as well as $(X_{\mathbf{i}})_{\mathbf{i}\in\mathcal{V}}$.
  - Fix $n \geq 0$ and define $Y_{\mathbf{i}}^n = V_{\mathbf{i}}$ for $|\mathbf{i}| = n$.
  - For $|\mathbf{i}| < n$ define $Y_{\mathbf{i}}^n$ recursively.

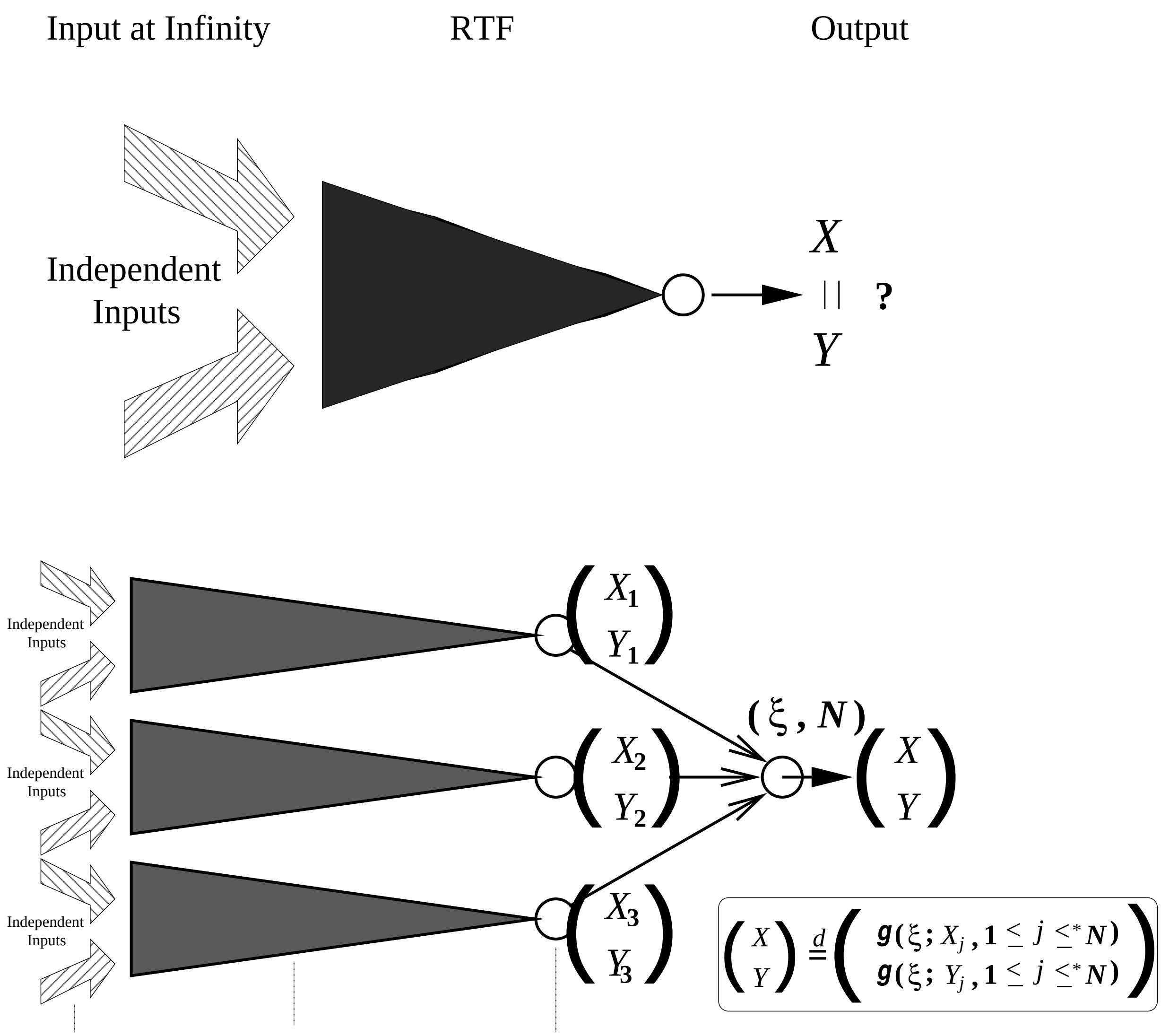
- Some immediate consequence of the construction are as follows :
  - $X_\emptyset \stackrel{d}{=} Y_\emptyset^n \stackrel{d}{=} \mu$ for every $n \geq 0$.
  - For each $n \geq 0$, the random variables $X_\emptyset$ and $Y_\emptyset^n$ when conditioned on the σ-algebra $\mathcal{G}_n$, are independent and identically distributed.
  - Moreover for each $n \geq 0$,

$$\left[\begin{array}{c} X_\emptyset \\ Y_\emptyset^{n+1} \end{array} \right] \stackrel{d}{=} T^{(2)}\left(\text{dist}\left(\left[\begin{array}{c} X_\emptyset \\ Y_\emptyset^{n} \end{array} \right]\right)\right).$$

  - Finally we also note that

$$\left[\begin{array}{c} X_\emptyset \\ Y_\emptyset^{n} \end{array} \right] \stackrel{d}{=} T^{(2)^n}(\mu \otimes \mu).$$


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}
Φ
Y^n_Φ
(ξ_Φ, N_Φ)
Y^n_1
(ξ_1, N_1)
Y^n_2
(ξ_2, N_2)
Y^n_3
(ξ_3, N_3)
Y^n_{11}
Y^n_{12}
Y^n_{13}
Y^n_{21}
Y^n_{22}
Y^n_{23}
V
At Level "n" Put Independent V's


Input at Infinity
RTF
Output
Independent Inputs
X
||
?
Y
Independent Inputs
Independent Inputs
Independent Inputs
X1
Y1
X2
Y2
X3
Y3
(ξ, N)
X
Y
(X, Y) =d (g(ξ; Xj, 1 ≤ j ≤* N), g(ξ; Yj, 1 ≤ j ≤* N))

- Consider the sequence $\left(X_\emptyset, Y_\emptyset^n\right)_{n \geq 0}$.

- It is tight because all the marginals are same which is μ.

- If $f, h : S \to \mathbb{R}$ be two bounded continuous functions then

$$\begin{aligned}
\mathbf{E}\left[f\left(X_\emptyset\right) h\left(Y_\emptyset^n\right)\right] &= \mathbf{E}\left[\mathbf{E}\left[f\left(X_\emptyset\right) h\left(Y_\emptyset^n\right) \,|\, \mathcal{G}_n\right]\right] \\
&= \mathbf{E}\left[\mathbf{E}\left[f\left(X_\emptyset\right) \,|\, \mathcal{G}_n\right] \mathbf{E}\left[h\left(X_\emptyset\right) \,|\, \mathcal{G}_n\right]\right] \\
&\to \mathbf{E}\left[\mathbf{E}\left[f\left(X_\emptyset\right) \,|\, \mathcal{G}\right] \mathbf{E}\left[h\left(X_\emptyset\right) \,|\, \mathcal{G}\right]\right]
\end{aligned}$$

- So we conclude that

$$\left(X_\emptyset, Y_\emptyset^n\right) \xrightarrow{d} \left(X^*, Y^*\right),$$

for some (X^*, Y^*) which has marginal μ.

- From the (technical) continuity assumption of $T^{(2)}$ we get that (X^*, Y^*) is a solution of the bivariate equation with marginal μ, so using bivariate uniqueness we get

$$X^* = Y^* \text{ a.s.}$$

- Let $\sigma_n^2(\Lambda) := \|\mathbf{E}[\Lambda(X_\emptyset) \mid \mathcal{G}_n] - \Lambda(X_\emptyset)\|_2^2$.

- Easy calculation shows

$$\begin{aligned}\sigma_n^2(\Lambda) &= \mathbf{E}[\mathrm{Var}(\Lambda(X_\emptyset) \mid \mathcal{G}_n)] \\ &= \tfrac{1}{2}\mathbf{E}\left[\left(\Lambda(X_\emptyset) - \Lambda\left(Y_\emptyset^n\right)\right)^2\right]\end{aligned}$$

  The last inequality follows from the simple fact that for any random variable U with finite second moment,

$$\mathrm{Var}(U) = \tfrac{1}{2}\mathbf{E}\left[(U_1 - U_2)^2\right]$$

  where (U_1, U_2) are two independent copies of U.

- Taking $n \to \infty$ limit we can then conclude that $\sigma_n^2(\Lambda) \to 0$, because $\left(X_\emptyset, Y_\emptyset^n\right) \xrightarrow{d} (X^*, X^*)$.

- So $\Lambda(X_\emptyset) = \mathbf{E}[\Lambda(X_\emptyset) \mid \mathcal{G}]$ a.s.

- This is true for every bounded continuous function Λ, so we conclude that $X_\emptyset$ is a.s. $\mathcal{G}$ measurable, proving the endogeny property.

Part (c) [“if”-part] :

- We know that the construction of $\left(Y_{\mathbf{i}}^{n}\right)_{|\mathbf{i}|\le n}$ yields $\left(X_{\emptyset}, Y_{\emptyset}^{n}\right)$ has distribution ${T^{(2)}}^{n}(\mu \otimes \mu)$.

- So we get

$$\left(X_{\emptyset}, Y_{\emptyset}^{n}\right) \xrightarrow{d} \mu^{\nearrow}.$$

- The rest follows from the previous argument.

Part (c) [“only if”-part] :

- Again work with the same construction of $\left(Y_{\emptyset}^{n}\right)_{n\geq 0}$.

- Let $\Lambda_1, \Lambda_2 : S \to \mathbb{R}$ be two bounded continuous functions.

$$\begin{aligned} & \mathbf{E}\left[\Lambda_1\left(X_{\emptyset}\right)\Lambda_2\left(Y_{\emptyset}^{n}\right)\right] \\ = \ & \mathbf{E}\left[\mathbf{E}\left[\Lambda_1\left(X_{\emptyset}\right)\Lambda_2\left(Y_{\emptyset}^{n}\right) \mid \mathcal{G}_n\right]\right] \\ = \ & \mathbf{E}\left[\mathbf{E}\left[\Lambda_1\left(X_{\emptyset}\right) \mid \mathcal{G}_n\right] \mathbf{E}\left[\Lambda_2\left(X_{\emptyset}\right) \mid \mathcal{G}_n\right]\right] \\ \to \ & \mathbf{E}\left[\mathbf{E}\left[\Lambda_1\left(X_{\emptyset}\right) \mid \mathcal{G}\right] \mathbf{E}\left[\Lambda_2\left(X_{\emptyset}\right) \mid \mathcal{G}\right]\right] \\ = \ & \mathbf{E}\left[\Lambda_1\left(X_{\emptyset}\right)\Lambda_2\left(X_{\emptyset}\right)\right] \end{aligned}$$

  The last equality follows from endogeny assumption.

- This of course then implies

$$\left(X_{\emptyset}, Y_{\emptyset}^{n}\right) \ \stackrel{d}{\longrightarrow} \ \left(X_{\emptyset}, X_{\emptyset}\right),$$

  which is same as saying

$$T^{(2)^n}\left(\mu \otimes \mu\right) \ \stackrel{d}{\longrightarrow} \ \mu^{\nearrow}.$$

Application to Solutions 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 $U \sim \text{Uniform}[0,1]$, 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

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

  [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 $X \stackrel{d}{=} Y \stackrel{d}{=} \nu * \delta_m$, 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

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

- 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.