TRAVELLING WAVES AND THE DISTRIBUTION OF THE HEIGHT OF BINARY SEARCH TREES*

Michael Drmota

Inst. of Discrete Mathematics and Geometry

Vienna University of Technology, A 1040 Wien, Austria

michael.drmota@tuwien.ac.at

www.dmg.tuwien.ac.at/drmota/

* joint work with **Brigitte Chauvin** (Université de Versailles)

ÖMG-DMV Kongress Mathematik 2005, Klagenfurt, September 18–23, 2005

Outline of the Talk

- Random Bisection Problem
- Binary Search Trees
- Results
- Travelling Wave
- Intersection Property

Random Bisection Problem

Random Bisection Problem

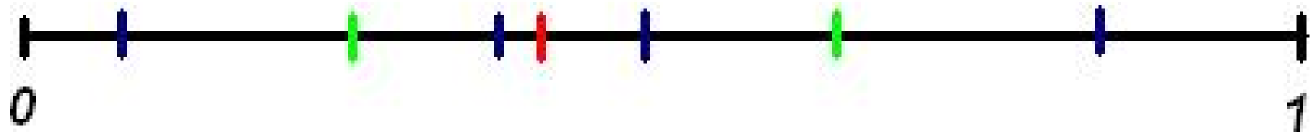
Random Bisection Problem

Random Bisection Problem

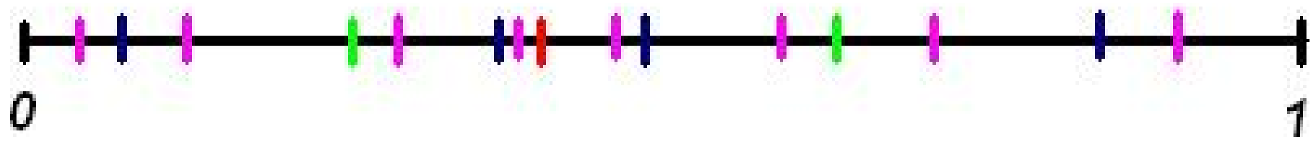
V is a random variable on $[0, 1]$

Random Bisection Problem

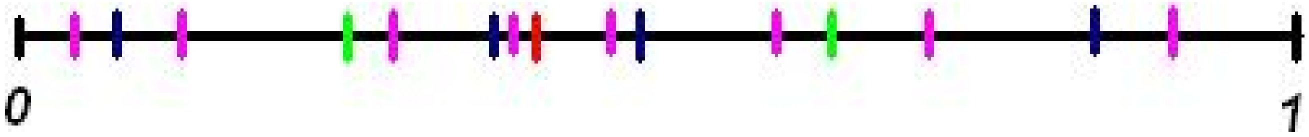
Random Bisection Problem



Random Bisection Problem



Random Bisection Problem



etc.

Random Bisection Problem

Analytic Description:

- x initial length (it was 1 in our example)

- $P_k(x,\ell)$ the probability that after k steps each of the 2^k fragments after is shorter (or equal) than ℓ

- $P_k(x,\ell) = P_k(x/\ell,1) = \overline{P}_k(x/\ell)$

- $\overline{P}_0(x) = 1$ for $0 \le x < 1$, $\overline{P}_0(x) = 0$ for $x \ge 1$

$$\boxed{\overline{P}_{k+1}(x) = \mathbf{E}\left(\overline{P}_k(xV)\overline{P}_k(x(1-V))\right)}$$

Random Bisection Problem

t-Beta-Distribution for V:

$$f(x) = \frac{(2t+1)!}{(t!)^2} \boxed{x^t(1-x)^t}$$

density of distribution of V ($t \geq 0$ integer parameter)

Random Bisection Problem

Transfer Operator T:

$$\boxed{(\mathbf{T}F)(x) := \mathbf{E}\left(F(xV)F(x(1-V))\right)}$$

With this definition we have

$$\boxed{\overline{P}_{k+1} = \mathbf{T}\overline{P}_k}$$

with $\overline{P}_0(x) = 1$ for $0 \leq x < 1$ and $\overline{P}_0(x) = 0$ for $x \geq 1$

Random Bisection Problem

Laplace Transform:

Set

$$L_F(u) := \int_0^\infty e^{-xu} F(x)\, dx \qquad (u > 0)$$

Then

$$\boxed{L_{\mathbf{T}F}(u)^{(2t+1)} = -\frac{(2t+1)!}{(t!)^2} \left(L_F(u)^{(t)}\right)^2}$$

Random Bisection Problem

Laplace Transform:

Set

$$L_F(u) := \int_0^\infty e^{-xu} F(x)\, dx \qquad (u > 0)$$

Then

$$\boxed{L_{\mathbf{T}F}(u)^{(2t+1)} = -\frac{(2t+1)!}{(t!)^2} \left(L_F(u)^{(t)} \right)^2}$$

and consequently

$$\boxed{\boxed{L_{\overline{P}_{k+1}}(u)^{(2t+1)} = -\frac{(2t+1)!}{(t!)^2} \left(L_{\overline{P}_k}(u)^{(t)} \right)^2}}$$

Binary Search Trees

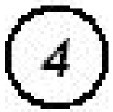
Storing Data:

4, 6, 3, 5, 1, 8, 2, 7

Binary Search Trees

Storing Data:

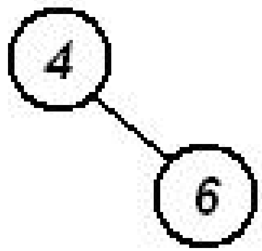
6, 3, 5, 1, 8, 2, 7



Binary Search Trees

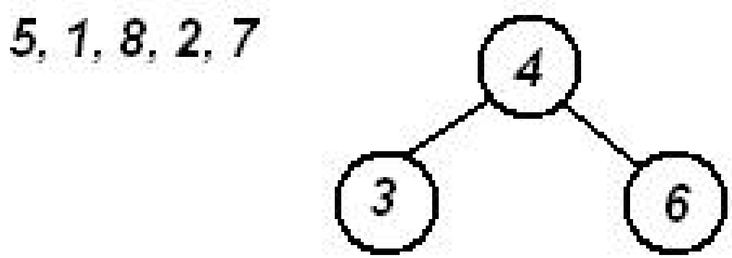
Storing Data:

3, 5, 1, 8, 2, 7



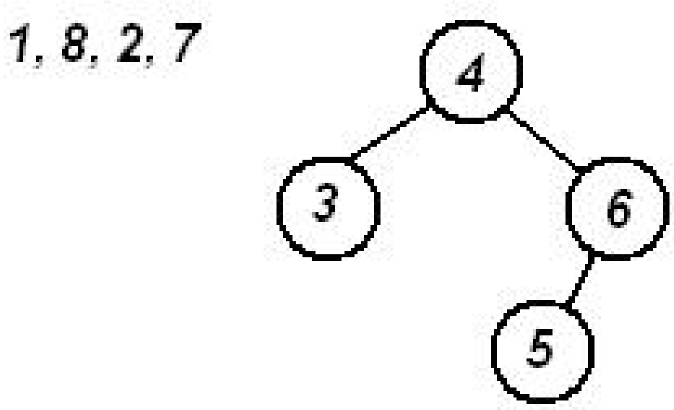
Binary Search Trees

Storing Data:



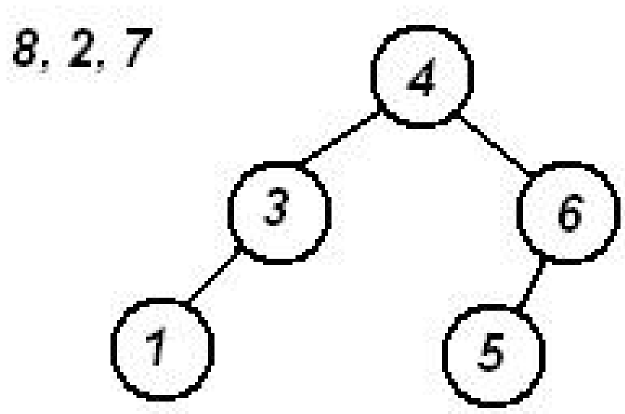
Binary Search Trees

Storing Data:



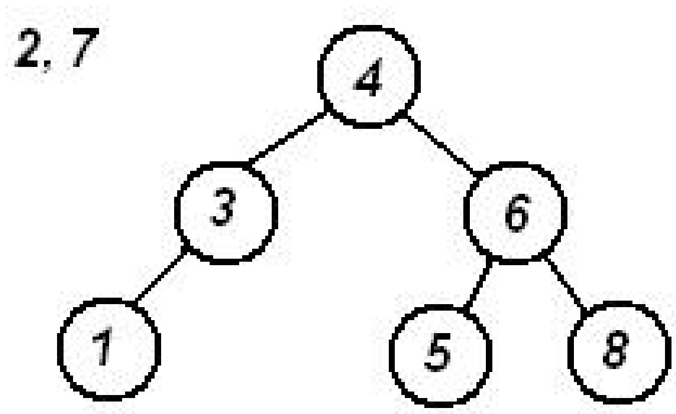
Binary Search Trees

Storing Data:



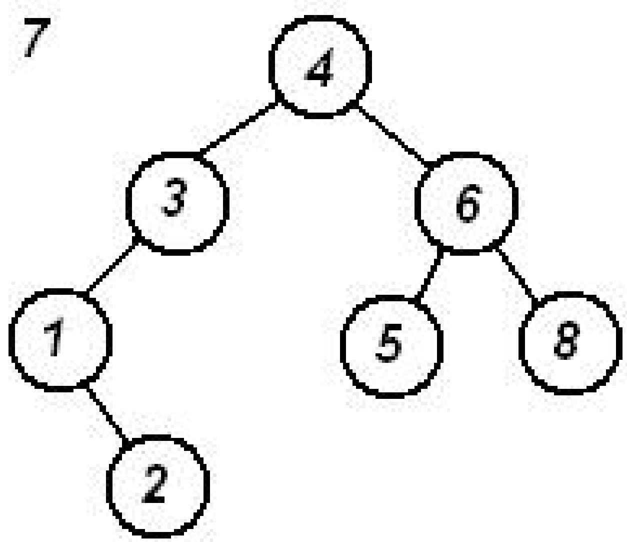
Binary Search Trees

Storing Data:



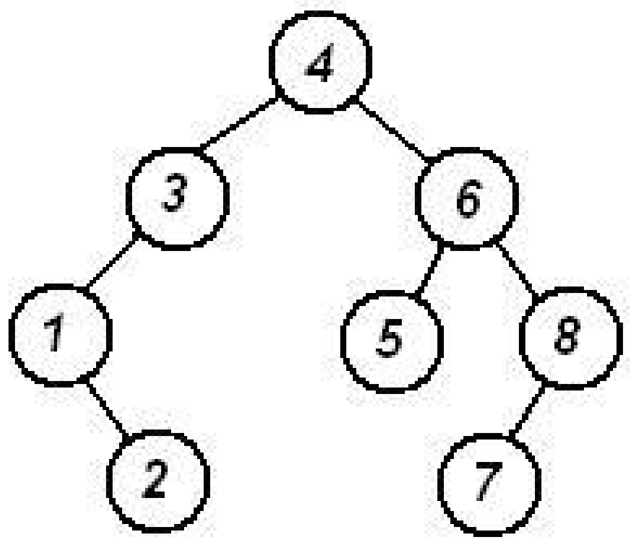
Binary Search Trees

Storing Data:



Binary Search Trees

Storing Data:



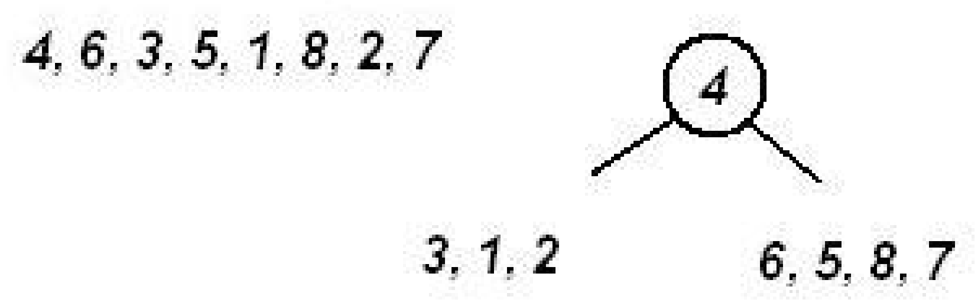
Binary Search Trees

Quicksort:

4, 6, 3, 5, 1, 8, 2, 7

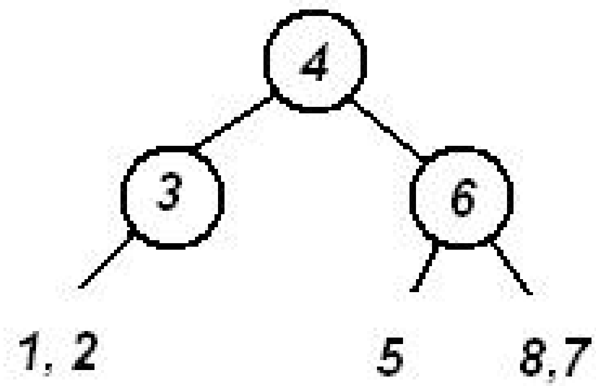
Binary Search Trees

Quicksort:



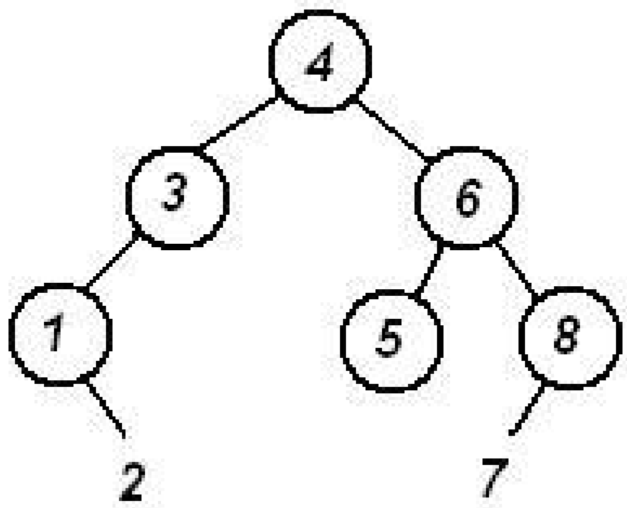
Binary Search Trees

Quicksort:



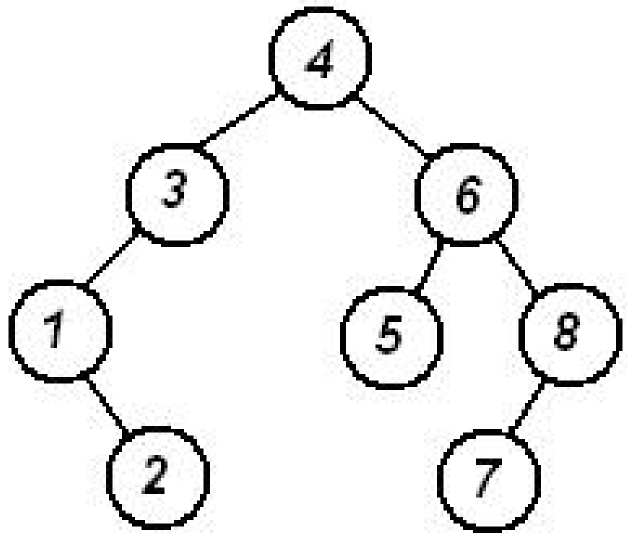
Binary Search Trees

Quicksort:



Binary Search Trees

Quicksort:



Binary Search Trees

Median of 3 – Quicksort:

4, 6, 3, 5, 1, 8, 2, 7

Binary Search Trees

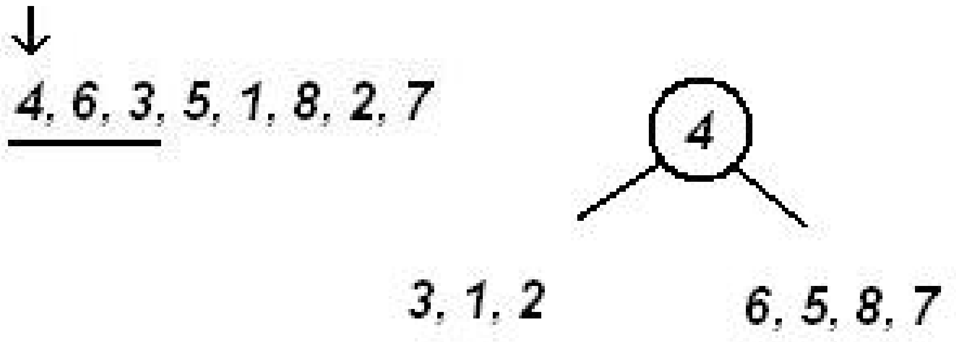
Median of 3 – Quicksort:

↓

4, 6, 3, 5, 1, 8, 2, 7

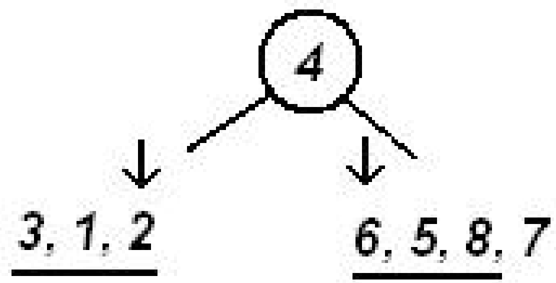
Binary Search Trees

Median of 3 – Quicksort:



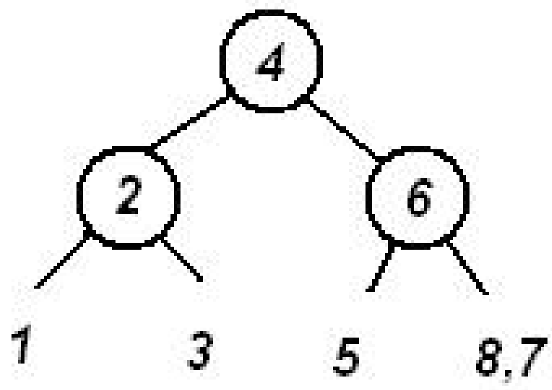
Binary Search Trees

Median of 3 – Quicksort:



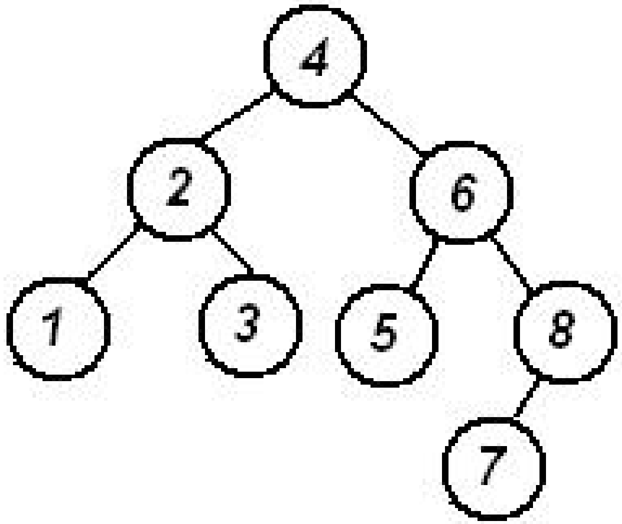
Binary Search Trees

Median of 3 – Quicksort:



Binary Search Trees

Median of 3 – Quicksort:



Binary Search Trees

Probabilistic Model:

Every permutation of $\{1, 2, \ldots, n\}$ is equally likely.

$\longrightarrow$ probability distribution on binary trees of size n

$\longrightarrow$ every parameter on trees is a **random variable**

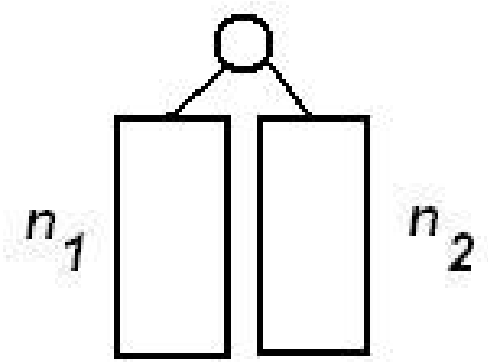
Notation

H_n ... **height** of trees (of size n)

Binary Search Trees

Height of Medien of $(2t+1)$-Quicksort:
(fringe balanced binary search trees)

$$\mathbf{Pr}\{H_n \le k+1\} = \sum_{n_1+n_2=n-1} \frac{\binom{n_1}{t}\binom{n_2}{t}}{\binom{n}{2t+1}} \mathbf{Pr}\{H_{n_1} \le k\} \cdot \mathbf{Pr}\{H_{n_2} \le k\}$$



Binary Search Trees

Generating Functions:

$$y_k(x) = \sum_{n \geq 0} \mathbf{Pr}\{H_n \leq k\} \cdot x^n$$

$$y_{k+1}(x)^{(2t+1)} = \frac{(2t+1)!}{(t!)^2} \left(y_k(x)^{(t)} \right)^2$$

with initial conditions $y_0(x) = 1, \; y_k(0) = 1$.

Random Bisection Problem versus Binary Search Trees

Comparision:

Random bisection problem

$$\boxed{\boxed{L_{\overline{P}_{k+1}}(u)^{(2t+1)} = -\frac{(2t+1)!}{(t!)^2}\left(L_{\overline{P}_k}(u)^{(t)}\right)^2}}$$

Height of (fringe balanced) binary search trees

$$\boxed{\boxed{y_{k+1}(x)^{(2t+1)} = \frac{(2t+1)!}{(t!)^2}\left(y_k(x)^{(t)}\right)^2}}$$

Results

Theorem 1

Let x_k be defined by $\overline{P}_k(x_k) = \frac{1}{2}$. Then there exists a continuous function $F(x)$ such that (uniformly for $x \geq 0$ as $k \to \infty$)

$$\boxed{\overline{P}_k(x) = F(x/x_k) + o(1)}.$$

More precisely, we have

$$x_k = e^{\rho k + \Theta(\log k)}$$

for some $\rho > 0$ (defined on the next slide) and $F(x)$ is uniquely defined by $F(1) = \frac{1}{2}$ and by the relation

$$\boxed{F(x/\rho) = (\mathbf{T}F)(x)}.$$

Results

Definition of ρ:

Let $\beta > 0$ be the solution of

$$\sum_{j=0}^{t} \log(\beta + t + 1 + j) - \log(2t!) = \sum_{j=0}^{t} \frac{\beta}{\beta + t + 1 + j}.$$

Then

$$\rho = \sum_{j=0}^{t} \frac{1}{\beta + t + 1 + j}.$$

Results

Theorem 2

We have (uniformly for $k \geq 0$ as $n \to \infty$)

$$\boxed{\mathbf{Pr}\{H_n \leq k\} = F(n/c_k) + o(1)},$$

where c_k satisfies $c_k \sim c y_k(1)$ (for some $c > 0$) and

$$c_k = e^{\rho k + o(k)}$$

Furthermore,

$$\mathbf{E}\, H_n = \max\{k \geq 0 : c_k \leq n\} + O(1) \sim \frac{1}{\log \rho} \cdot \log n$$

and

$$\boxed{\mathbf{Pr}\{|H_n - \mathbf{E}\, H_n| > y\} = O(e^{-\eta y})}.$$

In particular we have, as $n \to \infty$, $\mathbf{Var}\, H_n = O(1)$.

Travelling Wave

First Observation:

If $F(x)$ satisfies

$$F(x/\rho) = (\mathbf{T}F)(x).$$

(for some $\rho > 0$) then

$$F_k(x) := F(x/\rho^k)$$

satisfies the recurrence

$$\boxed{F_{k+1}(x) = (\mathbf{T}F_k)(x)}.$$

However, $F_0(x) = F(x) \neq \overline{P}_0(x) = \mathbf{1}_{[0,1]}(x)$

Travelling Wave

Second Observation:

Set

$$\Phi(u) = \int_0^\infty F(x)e^{-xu}\,dx$$

and

$$\tilde{y}_k(x) = e^{\rho k}\Phi(e^{\rho k}(1-x)).$$

Then

$$\boxed{\boxed{\tilde{y}_{k+1}(x)^{(2t+1)} = \frac{(2t+1)!}{(t!)^2}\left(\tilde{y}_k(x)^{(t)}\right)^2}}.$$

Travelling Wave

Solution of $F(x/\rho) = (\mathbf{T}F)(x)$:

Set $A_1 = (e^\rho V)^\beta$ and $A_2 = (e^\rho(1 - V))^\beta$ and suppose that $X \geq 0$ satisfies the **stochastic fixed point equation**:

$$\boxed{Y \stackrel{d}{=} A_1 Y_1 + A_2 Y_2}$$

(where Y_1 and Y_2 have the same distribution as Y and $Y_1, Y_2, (A_1, A_2)$ are independent).

Then

$$F(x) = \mathbf{E}\, e^{-x^\beta Y}$$

satisfies

$$F(x/\rho) = (\mathbf{T}F)(x).$$

Travelling Wave

Proposition (Biggins, Kyprianou, Durrett, Liggett, ...)

Set

$$v(\alpha) = \log\left(\mathbf{E}\left(\sum_{i\geq 1} A_i^\alpha\right)\right)$$

and suppose that $v(0) > 0$, that $\alpha = 1$ is contained in the interior of $\{\alpha : v(\alpha) < \infty\}$, and that

$$v(1) = v'(1) = 0.$$

Then the stochastic fixed point equation

$$Y \stackrel{d}{=} \sum_{i\geq 1} A_i Y_i$$

has (up to scaling) a unique non-negative solution and the Laplace transform $\Phi(x) = \mathbf{E}\, e^{xY}$ satisfies

$$\lim_{x\to 0+} \frac{1-\Phi(x)}{-x\log x} = c_1$$

for some constant $c_1 > 0$.

Travelling Wave

Remark 1.

The condition $v(1) = v'(1) = 0$ constitutes a **critical case**.

If $v(1) = 0$ and $v'(1) < 0$ then there is also a solution (that can be also obtained quite easily by a contraction argument).

If $v(1) = 0$ and $v'(1) > 0$ then there is **no solution**.

Remark 2.

The condition for β is just a reformulation that we are in the critial case.

Intersection Property

Point process:

$$Z = \sum_{j=1}^{N} \delta_{X_j},$$

Example: $N = 2$, $X_1 = \log(1/V)$, $X_2 = \log(1/(1-V))$.

Transform $\mathbf{T}$ (for distributions functions):

$$(\mathbf{T}G)(x) = \mathbf{E}\left(\prod_{j=1}^{N} G(x - X_j)\right).$$

Example: $G(x) = F(e^{-x})$: $F(x) = \mathbf{E}(F(xV)F(x(1-V)))$.

Intersection Property

Intersection property:

Suppose that $F(x)$ and $G(x)$ are continuous distribution functions such that the difference $F(x) - G(x)$ has exactly one zero. Then the difference $(\mathbf{T}\,F)(x) - (\mathbf{T}\,G)(x)$ has at most one zero.

Intersection Property

Lemma.

Suppose that V is t-beta distributed and $\mathbf{T}$ is defined by $(\mathbf{T}F)(x) = \mathbf{E}(F(xV)F(x(1-V)))$.

Then the Laplace transforms $\Phi(u) = \int_0^\infty F(x)e^{-xu}\,dx$ satisfy an *intersection property*.

This property is the **key property** for the proof of Theorems 1 and 2.

It is not clear whether this is also true on the level of distributions functions?

Intersection Property

Theorem 3

Let $G_0(x) = 0$ for $x < 0$ and $G_0(x) = 1$ for $x \geq 0$ and set $G_{k+1} = \mathbf{T}\, G_k$, that is,

$$G_{k+1}(x) = \mathbf{E}\left(\prod_{j=1}^{N} G_k(x - X_j) \right).$$

If $\mathbf{T}$ satisfies the *intersection property* then there exists $w(x)$ such that (uniformly for real x as $k \to \infty$)

$$\boxed{G_k(x) = w(x - m(k)) + o(1)},$$

where $m(k)$ is defined by $G_k(m(k)) = \frac{1}{2}$.

More precisely, we have

$$\boxed{m(k) = kc + o(k)}.$$

for some constant $c > 0$ and $w(x)$ satisfies

$$\boxed{w(x) = \mathbf{E}\left(\prod_{j=1}^{N} w(x + c - X_j)\right)}.$$

Thank You!