THE BETA(p,1) EXTENSIONS
OF THE RANDOM (UNIFORM) CANTOR SETS

Dinis D. Pestana

FCUL, DEIO and CEAUL, Universidade de Lisboa
Campo Grande, Edifício C4, 1749–016 Lisboa, Portugal

e-mail: dinis.pestana@fc.ul.pt

and

Sandra M. Aleixo, J. Leonel Rocha

Mathematics Unit, DEC, DEQ,
Instituto Superior de Engenharia de Lisboa and CEAUL
Rua Conselheiro Emídio Navarro, 1, 1949–014 Lisboa, Portugal

e-mails: sandra.aleixo@dec.isel.ipl.pt; jrocha@deq.isel.ipl.pt

Dedicated to Professor J. Tiago Mexia on his 70th birthday

Abstract

Starting from the random extension of the Cantor middle set in [0,1], by iteratively removing the central uniform spacing from the intervals remaining in the previous step, we define random Beta(p,1)-Cantor sets, and compute their Hausdorff dimension. Next we define a deterministic counterpart, by iteratively removing the expected value of the spacing defined by the appropriate Beta(p,1) order statistics. We investigate the reasons why the Hausdorff dimension of this deterministic fractal is greater than the Hausdorff dimension of the corresponding random fractals.

Keywords: order statistics, uniform spacings, random middle third Cantor set, Beta spacings, Hausdorff dimension.

2000 Mathematics Subject Classification: 60E05, 26A30.

1. Introduction

The beta family

$$f_{p,q}(x) = K\,x^{p-1}(1-x)^{q-1}\mathbb{I}_{(0,1)}(x), \text{ with } p > 0 \text{ and } q > 0$$

plays an important role in Probability and Statistics (the appropriate norming constant being in that context $K = \frac{1}{B(p,q)}$, where

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}\mathrm{d}x$$

is the beta function or Euler's integral of the first kind), namely because of the broad range of shapes for different values of the parameters.

For special values of the parameters, $f_{p,q}$ also plays an important role in other areas of Mathematics. Namely,

- for $p = q = 2$, the *logistic parabola* $f_{2,2} = K\,x\,(1-x)$ is at the basis of the successful Verhulst model in population dynamics, that has been at the core of important developments in the area of dynamical systems. In fact, the numerical solution of the equation $x = K\,x\,(1-x)$ using the fixed point method has been at the core of fulcral developments on the theory of fractals, namely of the theory of Feigenbaum bifurcations and limiting cahotic behaviour;

- for $p = q = 1$, corresponding in Probability to the uniform law, there has been the development of stochastic extensions of the deterministic Cantor set, using self-similarity, but at each step "erasing" a uniformly distributed middle portion from each interval remaining in the previous step.

Our research has been aimed at

- using general models $f_{p,q}$ in population dynamics to show that for any pair (p, q) there exists $K_i(p, q)$ and $K_\infty(p, q)$ such that Feigenbaum bifurcations and more and more complex cyclic behaviour can be observed as $K > K_i(p, q)$ increases, until chaotic behaviour is observed for $K > K_\infty(p, q)$, [1, 2, 3, 4] and [5];

- defining and characterizing the structure of random Cantor sets when the middle sets removed at each step have a general $f_{p,q}(x) = K\,x^{p-1}$ $(1-x)^{q-1}\mathbb{I}_{(0,1)}(x)$ law, [1] and [8].

Geometric constructions of random type have been studied by several authors. The Hausdorff dimension is an important structural characteristic of fractals. Aside from the raw use of the definition, structural properties such as self-similarity can be used to compute the Hausdorff dimension of a deterministic fractals, [6, 7, 9] and [10], and those can be extended to compute the Hausdorff dimension of a random fractal, [6, 7] and references therein.

In this work we define the random middle third Cantor set, a fractal which is constructed by recursive elimination of the central spacing generated by the minimum and maximum of two observations "at random" — in the usual sense of uniformly distributed — of each interval of the previous iteration. This name is broadly justified, in the sense that the expected values of the interval extremes of each iteration coincide with the extremes of the intervals of the correspondent iteration in the construction of the deterministic middle third Cantor set. Cf. also [11], about new trends in Biology using fractal models.

The purpose of the present work is to investigate an intriguing question: although the *expected value* of what is taken out at each step in the recursive construction of the random fractal is exactly of the same size of what is taken out in the corresponding recursive step of the construction of its associated deterministic set — which, in this sense, can be regarded as the "expected fractal" — the Hausdorff dimension of a random fractal is almost surely smaller than the Hausdorff dimension of its deterministic counterpart. So, intuitively it seems that the random fractal is a lesser portion of [0,1] than the corresponding expected deterministic fractal.

In Section 2, we present the concepts and framework needed to develop our research.

In Section 3, we compute the Hausdorff dimension — in intuitive terms, a parameter that evaluates how dense a set is in $\mathbb{R}^n$, for the appropriate dimension n of the Euclidian space where it lies — of the random middle third Cantor set. This section's purpose is to show that at the first step we almost surely take a middle interval lesser than the middle interval taken out in its deterministic expected counterpart; but, on the other hand, there is a trade-off in subsequent iterations of the procedure — in fact, one more example of the effects of skewness of the parent distribution, implying that even slight differences between mean and median have far reaching consequences —, so that at the end the odds that the remaining points in the random fractal are less dense in [0,1] than the points remaining in the deterministic fractal are greater than 1.

2. RANDOM MIDDLE THIRD CANTOR SET:
PRELIMINARIES, DEFINITIONS AND RESULTS

One important evolution of twentieth century Mathematics has been the eruption of fractal geometry. Indeed, fractal sets may give a much better representation of several natural phenomena than classical geometric figures do, [6] and [10].

The middle third Cantor set, a famous example of self-similarity of Georg Cantor, is one of the most well-known and easy to construct fractals; moreover, it exhibits the most typical characteristics of fractals. This set is constructed starting from a closed interval — without loss of generality, the interval $E_0 = [0, 1]$ —, by iterative elimination of the middle subintervals of the intervals left in the previous step. Hence in the next step we obtain $E_1 = [0, \frac{1}{}$

- $E_n = \bigcup\limits_{k=1}^{2^n} \mathbb{I}_k^{(n)} = \bigcup\limits_{k=1}^{2^{n-1}} \left(\left[a_{2k-1}^{(n)}, b_{2k-1}^{(n)} \right] \cup \left[a_{2k}^{(n)}, b_{2k}^{(n)} \right] \right)$, where for each

$$k = 1, 2, \ldots, 2^{n-1},$$

$$a_{2k-1}^{(n)} = a_k^{(n-1)}; \quad b_{2k-1}^{(n)} = a_k^{(n-1)} + \frac{b_k^{(n-1)} - a_k^{(n-1)}}{3};$$

$$a_{2k}^{(n)} = b_k^{(n-1)} - \frac{b_k^{(n-1)} - a_k^{(n-1)}}{3}; \quad b_{2k}^{(n)} = b_k^{(n-1)}.$$

The middle third Cantor set is $\mathcal{C} = \bigcap\limits_{k=1}^{\infty} E_n$.

At first sight, it seems that we remove so much of the interval $[0, 1]$ during the construction of $\mathcal{C}$ that "almost nothing" remains, in the long run[*]. Indeed, $\mathcal{C}$ is a set with a non denumerable infinite number of points, containing infinite points in every neighbourhood of each one of its points. In fact, it is obvious that the middle third Cantor set $\mathcal{C}$ consists of the set of points that belong to $[0, 1]$ which, when expressed in the basis 3, do not contain the digit 1 in the corresponding series expansion, i.e., $\sum_{i=1}^{\infty} \alpha_i \, 3^{-i}$ with either $\alpha_i = 0$ or $\alpha_i = 2$, for each i. Note that to obtain E_1 from E_0 we remove all the points points with $\alpha_1 = 1$; to obtain E_2 from E_1 we remove the points with $\alpha_2 = 1$; and so on.

The Hausdorff dimension, which we formally define below for subsets from a linear set, is an important metrical invariant which carries information about the fractal, namely by providing an intuitive insight on the density of the fractal [10]:

Definition 2. Let $\mathcal{E} = \bigcap_{k=0}^{\infty} F_k$ be a fractal set constructed recursively from the set F_0, in which after the k-th iteration, the set F_k is the union of n_k intervals, each of them having length $r_k \xrightarrow[k\to\infty]{} 0$. The Hausdorff dimension of the set $\mathcal{E}$ is

---

[*] In the k-th iteration of the procedure of elimination we are taking 2^{k-1} intervals of length $\frac{1}{3^k}$; so, since the intervals that we remove are pairwise disjoint, we are taking of a total "measure", $\sum_{k=1}^{\infty} \frac{2^{k-1}}{3^k} = 1$, which suggests the rough (and indeed wrong) statement that the length of $\mathcal{C}$ is 0.

$$Dim_H \, \mathcal{E} = \lim_{k \to \infty} \frac{\ln(n_k)}{\ln\left(\frac{1}{r_k}\right)}.$$

For instance, in what regards the Cantor fractal $\mathcal{C}$, we have

$$Dim_H \, \mathcal{C} = \lim_{k \to \infty} \frac{\ln\left(2^k\right)}{\ln\left(3^k\right)} = \frac{\ln(2)}{\ln(3)} \approx 0.63093,$$

a value between 0 and 1 as expected, because the middle third Cantor set is much more than a denumerable set of points, but much less than regular (continuous) curve.

The Hausdorff dimension can be computed using the self-similarity typical of fractals, [6]. The procedure, presented by Falconer, has the double advantage of being easier to apply, because it relies on the "self-similarity ratio" observed in the recursive construction of the fractal, and of having a straightforward generalization for random fractals.

Although the procedure is in general straightforward to apply, for the sake of completeness we quote the formal result from Falconer [6]; assume that $S_1, S_2, \ldots, S_m : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ are similarities, with

$$|S_i(x) - S_i(y)| = c_i |x - y|, \; x, y \in \mathbb{R}^n$$

where $0 < c_i < 1$ (c_i is called the similarity ratio S_i). So, each S_i transforms subsets of $\mathbb{R}^n$ in geometrically similar sets.

Further, assume that for pairwise disjoint subsets there exists a non empty set V, such that

$$\bigcup_{i=1}^{m} S_i(V) \subset V$$

with V an open and limited set (this is generally referred to as *the open set condition*). Falconer's [6] result may be stated as follows:

Theorem 1. *Suppose that the open set condition is verified for the similarities S_i defined on $\mathbb{R}^n$ with ratios c_i, $(1 \leq i \leq m)$. If $\mathcal{E}$ is the invariant set satisfying $\mathcal{E} = \bigcup_{i=1}^{m} S_i(\mathcal{E})$, then $dim_H \mathcal{E} = s$ where s is the solution of the equation*

$$(1) \qquad \sum_{i=1}^{m} c_i^s = 1.$$

For the middle third Cantor set $\mathcal{C}$, in the n-th step each of the intervals $\mathbb{I}_k^{(n-1)}$, whose union is E_{n-1}, gives rise to two disjoint subintervals $\mathbb{I}_{2k-1}^{(n)} \cup \mathbb{I}_{2k}^{(n)}$, and so E_n the union of the 2^n intervals obtained by this way. The $\mathbb{I}_{2k-1}^{(n)}$ is obtained from $\mathbb{I}_k^{(n-1)}$ applying the similarity $S_1(x) = \frac{1}{}$

$$I_{2k-1}^{(n)} \cup I_{2k}^{(n)} = I_k^{(n-1)} - \left(U_{1:2}^{(n-1,k)}, U_{2:2}^{(n-1,k)} \right),$$

with $U_{1:2}^{(n-1,k)}$ and $U_{2:2}^{(n-1,k)}$ the minimum and the maximum of a random sample of dimension two of $U_k^{(n-1)} \frown Uniform(I_k^{(n-1)})$, respectively.

The random fractal connected to the random variable U, i.e., the random middle third Cantor set, is $\mathcal{F}_U = \bigcap\limits_{k=1}^{\infty} G_n$.
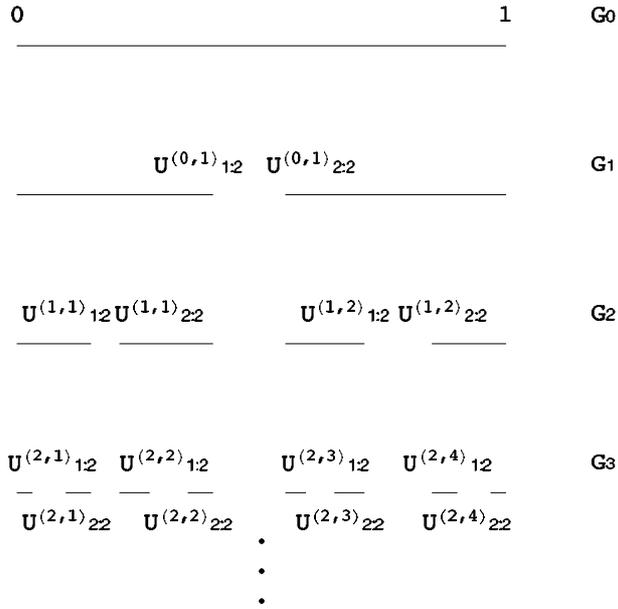

Figure 1. Construction of the random middle third Cantor set.

It is easy to see the reason why we denominate this random fractal *random middle third Cantor set*. In fact, with the help of the initial iterations in its

construction, with the obvious notations $I_1^{(1)} = [0, U_{1:2}^{(0,1)}]$, $I_2^{(1)} = [U_{2:2}^{(0,1)}, 1]$ and the removed interval or spacing $S_2 = (U_{1:2}^{(0,1)}, U_{2:2}^{(0,1)})$. It is readily established that

$$\mathbb{E}\left[U_{1:2}^{(0,1)}\right] = \int_0^1 2\, x_1 \left(1 - x_1\right) \mathrm{d}x_1 = \frac{1}{-}$$

where X and Y are the extremes of the random interval $I_k^{(n-1)}$; with $\mathbb{E}[X] = a_k^{(n-1)}$ and $\mathbb{E}[Y] = b_k^{(n-1)}$, $k = 1, \ldots, 2^{n-1}$, it follows that

$$
\begin{aligned}
\mathbb{E}\left[U_{1:2}^{(n-1,k)}\right] &= \mathbb{E}_{(X,Y)}\left[\mathbb{E}_{U_{1:2}^{(n-1,k)}|(X,Y)}\left[U_{1:2}^{(n-1,k)}\right]\right] \\[2mm]
&= \mathbb{E}_{(X,Y)}\left[X + \frac{Y - X}{3}\right] \\[2mm]
&= a_k^{(n-1)} + \frac{b_k^{(n-1)} - a_k^{(n-1)}}{3} = b_{2k-1}^{(n)}
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}\left[U_{2:2}^{(n-1,k)}\right] &= \mathbb{E}_{(X,Y)}\left[\mathbb{E}_{U_{2:2}^{(n-1,k)}|(X,Y)}\left[U_{2:2}^{(n-1,k)}\right]\right] \\[2mm]
&= \mathbb{E}_{(X,Y)}\left[Y - \frac{Y - X}{3}\right] \\[2mm]
&= b_k^{(n-1)} - \frac{b_k^{(n-1)} - a_k^{(n-1)}}{3} = a_{2k}^{(n)}
\end{aligned}
$$

as by definition $a_k^{(n-1)} = a_{2k-1}^{(n)}$ and $b_k^{(n-1)} = b_{2k}^{(n)}$, the result follows.

■

Let $\widetilde{G}_n$ be the length of the random interval G_n and $\widetilde{E}_n$ be the length of the interval E_n. As an immediately consequence of the above definitions and result, we can state the following:

Theorem 2. *The expected values of the extremes of the subsets $I_k^{(n)}$ of G_n, with $k = 1, 2, \ldots, 2^n$, in the construction of the random middle third Cantor set $\mathfrak{F}_U$, are coincident with the corresponding extremes of the subsets of E_n, in the construction of the middle third Cantor set $\mathcal{C}$, i.e., $\mathbb{E}[\widetilde{G}_n] = \widetilde{E}_n$.*

Proof. Having in mind that $I_{2k-1}^{(n)} \cup I_{2k}^{(n)} = [X, U_{1:2}^{(n-1,k)}] \cup [U_{2:2}^{(n-1,k)}, Y]$, for each $k = 1, 2, \ldots, 2^{n-1}$, we have

$$\mathbb{E}\left[\widetilde{I}_{2k-1}^{(n)} + \widetilde{I}_{2k}^{(n)}\right] = \mathbb{E}\left[\left(U_{1:2}^{(n-1,k)} - X\right) + \left(Y - U_{2:2}^{(n-1,k)}\right)\right]$$

$$= \mathbb{E}\left[U_{1:2}^{(n-1,k)}\right] - \mathbb{E}\left[X\right] + \mathbb{E}\left[Y\right] - \mathbb{E}\left[U_{2:2}^{(n-1,k)}\right]$$

$$= b_{2k-1}^{(n)} - a_k^{(n-1)} + b_k^{(n-1)} - a_{2k}^{(n)}$$

$$= \widetilde{\mathbb{I}}_{2k-1}^{(n)} + \widetilde{\mathbb{I}}_{2k}^{(n)}.$$

It follows from Definition 3 that $\mathbb{E}[\widetilde{G}_n] = \widetilde{E}_n$, as stated.

∎

3. HAUSDORFF DIMENSIONS OF $Beta(p, 1)$-CANTOR SETS

The random extension of fractals we adopted preserves one of the main features of fractality, namely self-similarity. In fact, the random Cantor set $\mathcal{F}$ that can be adequate to the following description

$$\mathcal{F} = \bigcap_{n=0}^{\infty} F_n,$$

where $[0, 1] = F_0 \supset F_1 \supset \ldots \supset F_n \supset \ldots$ is a decreasing sequence of closed intervals, where F_n is the union of 2^n closed and pairwise disjoint intervals $I^{(n)}$.

We assume the following conditions:

- Each interval $I^{(n)}$ of F_n contains two intervals of F_{n+1} (from the three intervals with random length in which F_n is divided, the middle interval is always eliminated in the following step). We designate these intervals by $I_L^{(n+1)}$ and $I_R^{(n+1)}$. The lower bound of $I_L^{(n+1)}$ is coincident with the lower bound of $I^{(n)}$ and the upper bound of $I_R^{(n+1)}$ is coincident with the upper bound of $I^{(n)}$.

- The lengths of the intervals $I_L^{(n+1)}$ and $I_R^{(n+1)}$ are random, and we inforce statistical self-similarity requiring the ratios $C_L^{(n)} = \frac{\widetilde{I}_L^{(n+1)}}{\widetilde{I}^{(n)}}$ to have the same probability distribution throughout, for any of the steps n and $n+1$, and for any interval $I^{(n)}$ of F_n, and also because of the necessity that the ratios $C_R^{(n)} = \frac{\widetilde{I}_R^{(n+1)}}{\widetilde{I}^{(n)}}$ have the same probability distribution, for any of the steps n and $n+1$, and for any interval $I^{(n)}$ of F_n. Note that, the ratios $C_L^{(n)}$ and $C_R^{(n)}$ do not necessarily have the same probability distribution and they are not independent.

As we assume that, for all steps n, with $n = 0, 1, 2, \ldots$, all the ratios $C_L^{(n)}$ have the same probability distribution, we can use in particular the ratio

$$C_1 = C_L^{(0)} = \frac{\widetilde{I}_L^{(1)}}{\widetilde{I}^{(0)}} = \frac{\widetilde{I}_L^{(1)}}{1} = \widetilde{I}_L^{(1)};$$

and similarly, as we assume that in each step the ratios $C_R^{(n)}$ do have the same probability distribution, we can use in particular the ratio

$$C_2 = C_R^{(0)} = \frac{\widetilde{I}_R^{(1)}}{\widetilde{I}^{(0)}} = \frac{\widetilde{I}_R^{(1)}}{1} = \widetilde{I}_R^{(1)}.$$

Falconer [6], proves the following result:

Theorem 3. *With probability* 1*, the random Cantor set* $\mathcal{F}$ *has Hausdorff dimension* $Dim_H \mathcal{F}$ *equal to* s*, where* s *is the solution of the equation*

$$(2) \qquad \mathbb{E}\left[C_1^s + C_2^s\right] = 1.$$

Note that, as in the procedure described in Theorem 1 to calculate the Hausdorff dimension of a deterministic fractal, based on the self-similarities caused by its recursive method of construction, s determines the expansion to which we would subject each element of F_n in order to "reconstruct" F_{n-1}. In a sense, the iteration of those "expansions" is, at each step, re-covering $F_0 = [0, 1]$.

In what concerns the random middle Cantor set, Definition 3, we have

$$C_1 = \widetilde{I}_L^{(1)} = |U_{1:2}^{(0,1)} - 0| = U_{1:2}^{(0,1)}$$

and

$$C_2 = \widetilde{I}_R^{(1)} = |1 - U_{2:2}^{(0,1)}| = 1 - U_{2:2}^{(0,1)}.$$

From

$$\mathbb{E}\left[C_1^s + C_2^s\right] = \int_0^1 x^s f_{U_{1:2}^{(0,1)}}(x)dx + \int_0^1 (1-x)^s f_{U_{2:2}^{(0,1)}}(x)dx = 4B(s+1,2) = 1$$

we conclude that

$$Dim_H \, \mathcal{F}_U = s = \frac{\sqrt{17}-3}{2} \approx 0.56155$$

almost surely.

This result is interesting, since it indicates that the random middle third Cantor set would tend to be less "dense" in [0,1] than the middle third Cantor set ($Dim_H \, \mathcal{C} \approx 0.63093$), although we have shown above that the deterministic fractal is the expectation of the corresponding random fractal.

As $C_1 = U_{1:2}^{(0,1)}$ has probability density function $f_{C_1}(x) = 2(1-x)I_{[0,1]}(x)$,

$$\mathbb{P}\left[C_1 \leq \mathbb{E}[C_1]\right] = \frac{5}{} \quad -$$

Now, suppose that we remove $S_2 = [X_{1:2}, X_{2:2}]$ from the interval $[0, 1]$, where $X_{1:2}$ and $X_{2:2}$ are the minimum and maximum, respectively, of two independent observations of the population $X \frown Beta(p, 1)$. After that, in each step, we remove the central spacing S_2 in each of the intervals remaining from the previous step. This process corresponds to a new construction of random type using the distribution Beta(p,1). In a similar way to the one used to define random middle third Cantor set in Definition 3, we define the $Beta(p, 1)$-Cantor sets as follows.

Definition 4. Let X be a Beta(p,1) random variable defined in the interval $(0, 1)$, i.e., $X \frown Beta(p, 1)$, where $X_{1:2}$ and $X_{2:2}$ are the minimum and the maximum of a random sample of dimension two of X, respectively. Let,

- $F_0 = [0, 1] = J_1^{(0)}$;

- $F_1 = F_0 - \left(X_{1:2}^{(0,1)}, X_{2:2}^{(0,1)} \right) = \left[0, X_{1:2}^{(0,1)} \right] \cup \left[X_{2:2}^{(0,1)}, 1 \right] = J_1^{(1)} \cup J_2^{(1)}$;

- $F_{n-1} = \bigcup_{k=1}^{2^{n-1}} J_k^{(n-1)}$ and $F_n = \bigcup_{k=1}^{2^n} J_k^{(n)}$, where for each

$$k = 1, 2, \ldots, 2^{n-1},$$

$$J_{2k-1}^{(n)} \cup J_{2k}^{(n)} = J_k^{(n-1)} - \left(X_{1:2}^{(n-1,k)}, X_{2:2}^{(n-1,k)} \right),$$

with $X_{1:2}^{(n-1,k)}$ and $X_{2:2}^{(n-1,k)}$ the minimum and the maximum of a random sample of dimension two of $X_k^{(n-1)} \frown Beta(p, 1, J_k^{(n-1)})$, respectively.

The random fractal connected to the random variable X, i.e., the random $Beta(p, 1)$-Cantor set, is $\mathcal{F}_{p,1} = \bigcap_{k=1}^{\infty} F_n$.

In particular, the random middle Cantor set discussed so far, in this new perspective, is $\mathcal{F}_U = \mathcal{F}_{1,1}$, since $U \frown Uniform(0,1)$ is $U \frown Beta(1,1)$.

In a correspondent deterministic approach, we consider a "mean fractal" $\mathcal{C}_{p,1}$ corresponding to $\mathcal{F}_{p,1}$, which is the intersection of the sets obtained as follows: starting from the interval $[0,1]$, we remove $[\mathbb{E}[X_{1:2}], \mathbb{E}[X_{2:2}]]$ which is the expected spacing, and this procedure is iterated. Formally, we have:

Definition 5. Let X be a Beta(p,1) random variable defined in the interval $(0,1)$, i.e., $X \frown Beta(p,1)$, where $X_{1:2}$ and $X_{2:2}$ are the minimum and the maximum of a random sample of dimension two of X, respectively. Let,

- $H_0 = [0,1] = \mathbb{J}_1^{(0)}$;

- $H_1 = H_0 - \left(\mathbb{E}\left[X_{1:2}^{(0,1)} \right], \mathbb{E}\left[X_{2:2}^{(0,1)} \right] \right) = \mathbb{J}_1^{(1)} \cup \mathbb{J}_2^{(1)}$;

- $H_{n-1} = \displaystyle\bigcup_{k=1}^{2^{n-1}} \mathbb{J}_k^{(n-1)}$ and $H_n = \displaystyle\bigcup_{k=1}^{2^n} \mathbb{J}_k^{(n)}$, where for each

$$k = 1, 2, \ldots, 2^{n-1},$$

$$\mathbb{J}_{2k-1}^{(n)} \cup \mathbb{J}_{2k}^{(n)} = \mathbb{J}_k^{(n-1)} - \left(\mathbb{E}\left[X_{1:2}^{(n-1,k)} \right], \mathbb{E}\left[X_{2:2}^{(n-1,k)} \right] \right),$$

with $X_{1:2}^{(n-1,k)}$ and $X_{2:2}^{(n-1,k)}$ the minimum and the maximum of a random sample of dimension two of $X_k^{(n-1)} \frown Beta(p,1,\mathbb{J}_k^{(n-1)})$, respectively.

The "mean fractal" or the deterministic $Beta(p,1)$-Cantor set is

$$\mathcal{C}_{p,1} = \bigcap_{k=1}^{\infty} H_n.$$

It is well known that if X is a positive random variable with distribution function $F_X(x)$, and the expectation $\mathbb{E}[X]$ exists, it can be computed using the Riemann integral of the right tail

$$\mathbb{E}[X] = \int_0^{+\infty} [1 - F_X(x)]dx.$$

The distribution functions of the minimum and of the maximum of a random sample of size two, are

$$F_{X_{1:2}}(x) = 1 - (1 - F_X(x))^2 \qquad \text{and} \qquad F_{X_{2:2}}(x) = (F_X(x))^2$$

consequently, in case the expectations of the minimum and maximum of a random sample of size two of a positive random variable with distribution function F_X do exist they may be computed as

$$\mathbb{E}[X_{1:2}] = \int_0^1 (1 - F_X(x))^2 \, dx$$

and

$$\mathbb{E}[X_{2:2}] = \int_0^1 \left(1 - (F_X(x))^2\right) dx.$$

Therefore, if $X \frown Beta(p,1)$ then

$$F_X(x) = x^p I_{(0,1)}(x) + I_{(1,\infty)}(x).$$

Consequently,

$$(3) \qquad \mathbb{E}[X_{1:2}] = \frac{2p^2}{(p+1)(2p+1)}$$

and

$$(4) \qquad \mathbb{E}[X_{2:2}] = \frac{2p}{2p+1}.$$

Let $S_2 = X_{2:2} - X_{1:2}$ be the random middle spacing, which is removed in each step of the construction of the random $Beta(p,1)$-Cantor set, the deterministic middle spacing is given by

$$\mathbb{E}[S_2] = \mathbb{E}[X_{2:2} - X_{1:2}] = \frac{2p}{(p+1)(2p+1)}.$$

In order to compute the Hausdorff dimension of the random $Beta(p, 1)$-Cantor set using the Theorem 3, we only have to bear in mind that

$$f_{X_{1:2}}(x) = \left(2px^{p-1} - 2px^{2p-1}\right)I_{(0,1)}(x)$$

and

$$f_{X_{2:2}}(x) = 2px^{2p-1}I_{(0,1)}(x).$$

Solving the equation (2) in order to s, we have

$$\mathbb{E}\left[C_1^s + C_2^s\right] = 1 \Longleftrightarrow \int_0^1 x^s f_{X_{1:2}}(x)dx + \int_0^1 x^s f_{X_{2:2}}(1-x)dx = 1$$

$$\Longleftrightarrow \int_0^1 x^s 2px^{p-1} - 2px^{2p-1}dx + \int_0^1 x^s\left(2p(1-x)^{2p-1}\right)dx = 1$$

$$\Longleftrightarrow \frac{1}{p+s} - \frac{1}{2p+s} + \frac{\Gamma(s+1)\Gamma(2p)}{\Gamma(2p+s+1)} = \frac{1}{2p}.$$

On the other hand, to determine the Hausdorff dimension of the deterministic $Beta(p, 1)$-Cantor set, we base ourselves in the Theorem 1. Note that, in each step of the deterministic Cantor set construction, the similarity ratios are $c_1 = \mathbb{E}[X_{1:2}]$ and $c_2 = 1 - \mathbb{E}[X_{2:2}]$ to the left and right intervals, respectively. The expressions of $\mathbb{E}[X_{1:2}]$ and $\mathbb{E}[X_{2:2}]$ were calculated in (3) and (4). So, the equation (1) becomes

$$(\mathbb{E}[X_{1:2}])^s + (1 - \mathbb{E}[X_{2:2}])^s = 1.$$

The probability density function of the random middle spacing S_2 is given by

$$f_{S_2}(z) = 2 \left(\frac{1}{B(p,1)} \right)^2 \int_0^{1-z} x^{p-1}(z+x)^{p-1}dx$$

$$= 2p^2 \int_0^{1-z} x^{p-1}(z+x)^{p-1}dx.$$

The probability that the random middle spacing S_2 is greater than the corresponding deterministic $\mathbb{E}[S_2]$, can be computed by the following way

$$\mathbb{P}[S_2 > \mathbb{E}[S_2]] = \int_{\mathbb{E}[S_2]}^1 f_{S_2}(z)dz.$$

In the Table 1 we can observe, for some values of p, the probability of the random middle spacing S_2 being greater than the deterministic middle spacing $\mathbb{E}[S_2]$, as well as the Hausdorff dimensions of the respective random and deterministic $Beta(p,1)$-Cantor sets. We observe that

Table 1.   Probabilities of the random middle spacing S_2 be greater than the deterministic middle spacing $\mathbb{E}[S_2]$ and Hausdorff dimensions of the random and deterministic $Beta(p,1)$-Cantor sets.

p	$\mathbb{E}[S_2]$	$Med(S_2)$	$\mathbb{P}[S_2 > \mathbb{E}[S_2]]$	$Dim_H \mathcal{F}_{p,1}$	$Dim_H \mathcal{C}_{p,1}$
0.1	0.151515	0.026300	0.293280	0.352648	0.557659
0.25	0.266667	0.172200	0.392116	0.436276	0.578127
0.5	0.333333	0.278630	0.434425	0.500000	0.600967
0.75	0.342857	0.299360	0.443826	0.536400	0.617679
1	0.333333	0.292893	0.444444	0.561553	0.630930
1.5	0.300000	0.259473	0.437901	0.595741	0.651179
2	0.266667	0.225220	0.429426	0.618907	0.666305
3	0.214286	0.173688	0.415534	0.649741	0.688046
5	0.151515	0.116921	0.399752	0.685187	0.715013
20	0.046458	0.033219	0.376770	0.761324	0.778206

- Although the Hausdorff dimensions of both the corresponding random and deterministic $Beta(p,1)$-Cantor sets increase with the parameter p, we always have

$$Dim_H \mathcal{F}_{p,1} < Dim_H \mathcal{C}_{p,1}.$$

- In what concerns the expected value of random middle spacing S_2 and the probability of the random middle spacing S_2 be greater than the deterministic middle spacing $\mathbb{E}[S_2]$, we can state that both increase with the parameter, for $p < 1$; both quantities are decreasing functions for $p \geq 1$.

So, at first sight this would seem to reinforce the intuitive (but misguided) idea that the random fractal $\mathcal{F}_{p,1}$ should have a bigger Hausdorff dimension than the correspondent "mean fractal" $\mathcal{C}_{p,1}$. But this an uneducated guess, not taking into full account the dependence issues and the trade-off in sequential steps: the less you take out at one step, the more you will probably take off in following steps.

To gain a deeper insight, we are going to evaluate the probability that the sum of the lengths of the intervals removed until the step n in the construction of the random fractal, which we shall denoted by $S_{2,R}^{(n)}$ in what follows, is greater than the sum of the lengths of the intervals removed until the step n in the construction of the correspondent "mean fractal", denoted by $S_{2,D}^{(n)}$ in what follows. This evaluation cannot be done analytically, but the evaluation is readily performed using Monte Carlo methods.

To make the Monte Carlo simulation for determining these probabilities and the correspondent 95% confidence intervals, we used in each case 5000 runs.

On the other hand, in order to compute $S_{2,D}^{(n)}$ of the "mean fractal" $\mathcal{C}_{p,1}$, observe that in the first step we obtain $[0,a] \cup [b,1]$, where $a = \mathbb{E}[X_{1:2}]$ and $b = \mathbb{E}[X_{2:2}]$. A straightforward extension is stated in the theorem that follows.

Theorem 4. *The length of the sum of the intervals removed in the construction of a "mean fractal" $\mathcal{C}_{p,1}$, until the step n, is given by*

$$S_{2,D}^{(n)} = 1 - (a + (1-b))^n, \quad with \quad n = 1, 2, \ldots$$

where $a = \mathbb{E}[X_{1:2}]$ and $b = \mathbb{E}[X_{2:2}]$.

Proof. We are going to proof the result by induction. In the step 1, the set H_1 is formed by 2^1 intervals, with total length given by $a + (1 - b) = (a + (1 - b))^1$.

Consider that, in the step n the set H_n is formed by 2^n intervals, with total length given by $(a + (1 - b))^n$, which can be rewritten as

$$\sum_{j=0}^{n} \binom{n}{j} a^j (1 - b)^{n-j}.$$

In the step $n + 1$, each one of the 2^n intervals, on the step n, lost the middle interval, given rise to two intervals of lengths

$$a \cdot a^k (1 - b)^{n-k} = a^{k+1} (1 - b)^{n-k}$$

and

$$(1 - b) \cdot a^k (1 - b)^{n-k} = a^k (1 - b)^{n+1-k},$$

respectively. Consequently, the set H_{n+1} has $2 \cdot 2^n = 2^{n+1}$ intervals. The sum of lengths of these intervals is given by

$$\sum_{j=0}^{n} \binom{n}{j} \left[a^{j+1} (1 - b)^{n-j} + a^j (1 - b)^{n+1-j} \right]$$

$$= a \sum_{j=0}^{n} \binom{n}{j} a^j (1 - b)^{n-j} + (1 - b) \sum_{j=0}^{n} \binom{n}{j} a^j (1 - b)^{n-j}$$

$$= \sum_{j=0}^{n} \binom{n}{j} a^j (1 - b)^{n-j} (a + 1 - b)$$

$$= (a + 1 - b)^n (a + 1 - b)$$

$$= (a + 1 - b)^{n+1}.$$

So, $S_{2,D}^{(n)} = 1 - (a + (1 - b))^n$, with $n = 1, 2, \ldots$, as stated. ∎

In Table 2 we compute the probability that the accumulated length of the random middle sets removed in the recursive construction of the random

Table 2. Estimated probability the probability that the sum of the lengths of the random middle intervals removed until the step n in the construction of the random fractal exceeds the sum of the lengths of the intervals removed until the step n in the construction of the correspondent "mean fractal".

Step	$Beta(0.5,1)$ Est. Prob.	95% Conf. Int.	$Beta(0.75,1)$ Est. Prob.	95% Conf. Int.	$Beta(1,1)$ Est. Prob.	95% Conf. Int.
1	0.4524	(0.4386; 0.4662)	0.4508	(0.4370; 0.4646)	0.4570	(0.4432; 0.4708)
2	0.5222	(0.5084; 0.5361)	0.5216	(0.5078; 0.5355)	0.5148	(0.5010; 0.5287)
3	0.5704	(0.5567; 0.5841)	0.5500	(0.5362; 0.5638)	0.5516	(0.5378; 0.5654)
4	0.5870	(0.5734; 0.6007)	0.5670	(0.5533; 0.5807)	0.5744	(0.5607; 0.5881)
5	0.6090	(0.5955; 0.6225)	0.5834	(0.5697; 0.5971)	0.5776	(0.5639; 0.5913)
6	0.6266	(0.6132; 0.6400)	0.5906	(0.5770; 0.6042)	0.5936	(0.5800; 0.6072)
7	0.6342	(0.6209; 0.6476)	0.5902	(0.5766; 0.6038)	0.6002	(0.5866; 0.6138)
8	0.6370	(0.6237; 0.6503)	0.5926	(0.5790; 0.6062)	0.6048	(0.5913; 0.6184)
9	0.6442	(0.6309; 0.6575)	0.5964	(0.5828; 0.6100)	0.6084	(0.5949; 0.6219)
10	0.6506	(0.6474; 0.6638)	0.5988	(0.5852; 0.6124)	0.6098	(0.5963; 0.6233)
Step	$Beta(2,1)$ Est. Prob.	95% Conf. Int.	$Beta(3,1)$ Est. Prob.	95% Conf. Int.	$Beta(20,1)$ Est. Prob.	95% Conf. Int.
1	0.4252	(0.4115; 0.4389)	0.4222	(0.4085; 0.4358)	0.3684	(0.3550; 0.3817)
2	0.4834	(0.4696; 0.4973)	0.4890	(0.4751; 0.5028)	0.4214	(0.4077; 0.4350)
3	0.5176	(0.5038; 0.5315)	0.5182	(0.5043; 0.5320)	0.4376	(0.4238; 0.4513)
4	0.5354	(0.5216; 0.5492)	0.5246	(0.5107; 0.5384)	0.4530	(0.4392; 0.4667)
5	0.5430	(0.5292; 0.5492)	0.5402	(0.5263; 0.5540)	0.4622	(0.4483; 0.4760)
6	0.5462	(0.5324; 0.5600)	0.5476	(0.5338; 0.5613)	0.4724	(0.4585; 0.4862)
7	0.5512	(0.5374; 0.5649)	0.5524	(0.5386; 0.5661)	0.4816	(0.4677; 0.4954)
8	0.5496	(0.5358; 0.5633)	0.5556	(0.5418; 0.5693)	0.4822	(0.4683; 0.4960)
9	0.5528	(0.5390; 0.5665)	0.5616	(0.5478; 0.5753)	0.4888	(0.4749; 0.5026)
10	0.5500	(0.5362; 0.5637)	0.5630	(0.5492; 0.5767)	0.4942	(0.4803; 0.5080)

$Beta(p, 1)$-Cantor set $\mathcal{F}_{p,1}$ is greater than the accumulated length of removed subintervals in the construction bof the corresponding deterministic $Beta(p, 1)$-Cantor set $\mathcal{C}_{p,1}$. While in the first step this probability is less than 0.5, in the second step, for small values of p, the odds are in favour that the length of the removed random set exceeds the length of what has been removed in its deterministic counterpart.

In fact, at each step of the recursive construction of the random fractal and of its deterministic counterpart, this pattern will apply: the probability that the accumulated length of the removed intervals in the random case exceeds the accumulated length of the removed intervals in the corresponding deterministic fractal increases steadily.

The dependence structure of order statistics, skewness and the consequent unequal mean and median contribute to this surprising reversal, and this deeper analysis of the situation shows that we should indeed expect that the random fractal be less dense in [0,1]. Thus, smaller Hausdorff dimension is a coherent result.

Aknowledgements

References

[1] S.M. Aleixo, *Analytical Methods in Probability and Probabilistic Methods in Analysis: Fractality Arising in Connection to Beta(p,q) Models, Population Dynamics and Hausdorff Dimension*, Ph. D. thesis, Univ. Lisbon, (2008).

[2] S.M. Aleixo, J.L. Rocha and D.D. Pestana, *Populational Growth Models in the Light of Symbolic Dynamics*, Proceedings of the ITI 2008, 30th International Conference on Information Technology Interfaces, Cavtat/Dubrovnik, Croatia (2008), IEEE:10.1109/ITI.2008.4588428, 311–316.

[3] S.M. Aleixo, J.L. Rocha and D.D. Pestana, *Populational Growth Models Proportional to Beta Densities with Allee Effect*, Proceedings of the BVP 2008, International Conference on Boundary Value Problems: Mathematical Models in Engineering, Biology and Medicine, Spain, American Institute of Physics **1124** (2009), 3–12.

[4] S.M. Aleixo, J.L. Rocha and D.D. Pestana, *Probabilistic Methods in Dynamical Analysis: Populations Growths Associated to Models Beta(p, q) with Allee Effect*, Proceedings of the DYNA08, International Conference on Dynamics and Applications 2008, Braga, Portugal (2008), in press.

[5] S.M. Aleixo, J.L. Rocha and D.D. Pestana, *Dynamical Behaviour on the Parameter Space: New Populational Growth Models Proportional to Beta Densities*, Proceedings of the ITI 2009, 31th International Conference on Information Technology Interfaces, Cavtat/Dubrovnik, Croatia (2009), IEEE:10.1109/ITI.2009.5196082, 213–218.

[6] K. Falconer, *Fractal Geometry, Mathematical Foundations and Applications*, John Wiley and Sons, New York 1990.

[7] Y. Pesin and H. Weiss, *On the Dimension of the Deterministic and Random Cantor-Like Sets*, Math. Res. Lett. **1** (1994), 519–529.

[8] D.D. Pestana, S.M. Aleixo and J.L. Rocha, *Hausdorff Dimension of the Random Middle Third Cantor Set*, Proceedings of the ITI 2009, 31th International Conference o n Information Technology Interfaces, Cavtat/Dubrovnik, Croatia (2009), IEEE:10.1109/ITI.2009.5196094, 279–284.

[9] J.L. Rocha and J.S. Ramos, *Weighted Kneading Theory of One-Dimensional Maps with a Hole*, Int. J. Math. Math. Sci. **38** (2004), 2019–2038.

[10] M. Schroeder, *Fractals, Chaos, Power Laws: Minutes from a Infinite Paradise*, W.H. Freeman, New York 1991.

[11] P. Waliszewski and J. Konarski, *A Mystery of the Gompertz Function*, Fractals in Biology and Medicine **IV** (2005), 277–286.