

Beta(p,q)-Cantor Sets: Determinism and Randomness

J. Leonel Rocha¹, Sandra M. Aleixo² and Dinis D. Pestana³

¹ Mathematics Unit, DEQ, Instituto Superior de Engenharia de Lisboa, and
CEAUL, Lisboa, Portugal

Email: jrocha@deq.isel.ipl.pt

² Mathematics Unit, DEC, Instituto Superior de Engenharia de Lisboa, and
CEAUL, Lisboa, Portugal

Email: sandra.aleixo@dec.isel.ipl.pt

³ FCUL, Universidade de Lisboa, and CEAUL, Lisboa, Portugal

Email: dinis.pestana@fc.ul.pt

Abstract: Usually randomness appears as a sophisticated extension of deterministic models, that are then presented as expectation of some class of random models (this approach is exceedingly well managed in the classical Barucha-Reid's treatise on random functions and stochastic processes). The works [1], [2], [3] and [5] summarize previous studies by the authors, using stochastic definitions of extensions of Cantor's fractal to put forward appropriate deterministic models, that in a precise sense are the expectation of a structured class of models, and investigated bifurcations, Allee' effect, and the Hausdorff dimension. $Beta(p,q)$ models, with either $p=1$ or $q=1$, or the classical Verhulst's model ($p=q=2$), proportionate interesting computable models for which computations both of Hausdorff dimension and probabilities can be explicitly evaluated, either analytically or using the Monte Carlo method.

The present extension, axed on arbitrary symbolic dynamical systems, further develops new fundamental classes of geometric constructions, and exploits the interplay of determinism and randomness on the richness of the limit fractal set, in a recursive construction. This sheds new light on the concept of Hausdorff dimensionality. We show that the dependence of the random order statistics is at the core of the apparent anomaly of consistently smaller Hausdorff dimensions of the random sets, when compared with the corresponding "expected" deterministic counterparts. We also recover Falconner's, Pesin's and Weiss' (among others) ideas on recursive geometric constructions as a straightforward approach to important issues in fractality and chaos.

Keywords: $Beta(p,q)$ densities, Random and deterministic $Beta(p,q)$ -Cantor sets, Hausdorff dimension.

1. Introduction

Our research has been aimed at defining and characterizing a new structure of

Rocha et al.

random Cantor sets when the middle sets removed at each step have a $Beta(p,q)$ law, with $q=1$, [5], and with $q=2$, [3]. In this work, we generalize these ideas to any other values of the parameter $q > 0$. In section 2, we formally define random and deterministic $Beta(p,q)$ -Cantor sets. In section 3, we developed methods to determine the Hausdorff dimension of these new sets, and in last section we present some numerical results which show that the Hausdorff dimension of a certain random $Beta(p,q)$ -Cantor set is smaller than the Hausdorff dimension of the correspondent deterministic $Beta(p,q)$ -Cantor set, as a consequence of Jensen's inequality.

2. Random and deterministic Beta(p,q)-Cantor sets

Consider that we remove the middle subinterval $S_1^{(0)} = [X_{1:2}^{(0,1)}, X_{2:2}^{(0,1)}]$, from the interval $[0, 1]$, where $X_{1:2}^{(0,1)}$ and $X_{2:2}^{(0,1)}$ are the minimum and maximum, respectively, of two independent observations of a random variable $X_1^{(0)} \cap Beta(p,q)$. After that, in each step n , we remove the middle subinterval $S_k^{(n-1)}$ in each one of the $k=1, 2, \dots, 2^{n-1}$ intervals remaining from the $n-1$ previous step. This procedure corresponds to a new geometric construction of random type using the distribution beta with shape parameters p and q , with adequate support in each interval remaining. In this work, as a similar way to the one used to define the random middle third Cantor set, random $Beta(p,1)$ -Cantor sets, [5], and random $Beta(p,2)$ -Cantor sets, [3], we present the general definition of the random $Beta(p,q)$ -Cantor sets.

Definition 1: Let $X_1^{(0)}$ be a $Beta(p,q)$ random variable defined in the interval $]0, 1[$, $X_{1:2}^{(0,1)}$ and $X_{2:2}^{(0,1)}$ be the minimum and the maximum, respectively, of a random sample of dimension two from $X_1^{(0)} \cap Beta(p,q)$. Let,

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

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

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

$$J_{2k-1}^{(n)} \cup J_{2k}^{(n)} = J_k^{(n-1)} \setminus]X_{1:2}^{(n-1,k)}, X_{2:2}^{(n-1,k)}[$$

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 $X_k^{(n-1)} \cap Beta(p, q, J_k^{(n-1)})$, respectively. The

Beta(p,q)-Cantor Sets: Determinism and Randomness

random fractal connected to a general random variable X , i.e., the random $Beta(p, q)$ -Cantor set, is

$$F_{p,q} = \bigcap_{n=0}^{\infty} F_n .$$

In a correspondent deterministic approach, we consider the intersection of the sets obtained starting from the interval $[0, 1]$, and removing iteratively the middle expected subinterval $[E[X_{1:2}^{(n-1,k)}], E[X_{2:2}^{(n-1,k)}]]$.

Therefore, formally, we can defined the deterministic $Beta(p, q)$ -Cantor set as follows:

Definition 2: Let $X_1^{(0)}$ be a $Beta(p, q)$ random variable defined in the interval $]0, 1[$, i.e., $X_1^{(0)} \cap Beta(p, q)$, where $E[X_{1:2}^{(0,1)}]$ and $E[X_{2:2}^{(0,1)}]$ are the expected values of the minimum and the maximum of a random sample of dimension two of $X_1^{(0)}$, respectively. Let,

$$H_0 = [0, 1] = J_1^{(0)} ;$$

$$H_1 = H_0 \setminus]E[X_{1:2}^{(0,1)}], E[X_{2:2}^{(0,1)}]][= [0, E[X_{1:2}^{(0,1)}]] \cup]E[X_{2:2}^{(0,1)}], 1] = J_1^{(1)} \cup J_2^{(1)} ;$$

$$H_{n-1} = \bigcup_{k=1}^{2^{n-1}} J_k^{(n-1)} \text{ and } H_n = \bigcup_{k=1}^{2^n} J_k^{(n)} , \text{ where for each } k = 1, 2, \dots, 2^{n-1} ,$$

$$J_{2k-1}^{(n)} \cup J_{2k}^{(n)} = J_k^{(n-1)} \setminus]E[X_{1:2}^{(n-1,k)}], E[X_{2:2}^{(n-1,k)}]][$$

with $E[X_{1:2}^{(n-1,k)}]$ and $E[X_{2:2}^{(n-1,k)}]$ the expected values of the minimum and the maximum of a random sample of dimension two of $X_k^{(n-1)} \cap Beta(p, q, J_k^{(n-1)})$, respectively. The ‘‘mean’’ fractal or the deterministic $Beta(p, q)$ -Cantor set, is

$$C_{p,q} = \bigcap_{n=0}^{\infty} H_n .$$

Note that, the deterministic $Beta(p, q)$ -Cantor sets $C_{p,q}$ satisfy the open set condition, [4].

Remark 3: Considering $X_k^{(n-1)} \cap Beta(p, q, J_k^{(n-1)})$, its distribution function is given by:

Rocha et al.

$$F_{X_k^{(n-1)}}(x) = \int_{a_k^{(n-1)}}^x \frac{t^{p-1}(1-t)^{q-1}}{B(p, q, \mathbf{J}_k^{(n-1)})} dt I_{\mathbf{J}_k^{(n-1)}}(x) + I_{(b_k^{(n-1)}, +\infty)}(x)$$

with $\mathbf{J}_k^{(n-1)} = [a_k^{(n-1)}, b_k^{(n-1)}]$. The mean values of the minimum and maximum of a two dimensional random sample, can be calculated as follows:

$$E[X_{1:2}^{(n-1, k)}] = \int_{a_k^{(n-1)}}^{b_k^{(n-1)}} (1 - F_X(x))^2 dx = \int_{a_k^{(n-1)}}^{b_k^{(n-1)}} \left(1 - \int_{a_k^{(n-1)}}^x \frac{t^{p-1}(1-t)^{q-1}}{B(p, q, \mathbf{J}_k^{(n-1)})} dt \right)^2 dx \quad (1)$$

and

$$E[X_{2:2}^{(n-1, k)}] = \int_{a_k^{(n-1)}}^{b_k^{(n-1)}} (F_X(x))^2 dx = \int_{a_k^{(n-1)}}^{b_k^{(n-1)}} \left(\int_{a_k^{(n-1)}}^x \frac{t^{p-1}(1-t)^{q-1}}{B(p, q, \mathbf{J}_k^{(n-1)})} dt \right)^2 dx. \quad (2)$$

3. The Hausdorff dimension of the random and deterministic Beta(p,q)-Cantor sets

The definition of random fractals presented in previous section, preserves one of the main features of fractality, namely self-similarity, but more sophisticated, in essence that the self-similarity of deterministic fractals:

- In the random $Beta(p, q)$ -Cantor set $F_{p, q}$, we have $[0, 1] = F_0 \supset F_1 \supset \dots \supset F_n \supset \dots$, a decreasing sequence of closed intervals, where F_n is the union of 2^n closed and pairwise disjoint intervals $J_k^{(n)}$.
- Each interval $J_k^{(n-1)}$ of F_{n-1} contains two intervals of F_n , since the middle interval from the three intervals with (dependent) random lengths in which F_{n-1} is divided, is always eliminated in the following step n . These intervals are denoted by $J_{2k-1}^{(n)}$ and $J_{2k}^{(n)}$, respectively. The left endpoint of $J_{2k-1}^{(n)}$ coincides with the left endpoint of $J_k^{(n-1)}$ and the right endpoint of $J_{2k}^{(n)}$ coincides with the right endpoint of $J_k^{(n-1)}$.
- The lengths of the intervals $J_{2k-1}^{(n)}$ and $J_{2k}^{(n)}$, denoted by $\tilde{J}_{2k-1}^{(n)}$ and $\tilde{J}_{2k}^{(n)}$, respectively, are random variables, and we enforce random self-similarity requiring the ratios $C_{2k-1}^{(n)} = \frac{\tilde{J}_{2k-1}^{(n)}}{\tilde{J}_k^{(n-1)}}$ to have the same probability

distribution throughout, for any of the steps $n-1$ and n , and for any interval $J_k^{(n-1)}$ of F_{n-1} , and also that the ratios $C_{2k}^{(n)} = \frac{\tilde{J}_{2k}^{(n)}}{\tilde{J}_k^{(n-1)}}$ have the same probability distribution, for any of the steps $n-1$ and n , and for any interval $J_k^{(n-1)}$ of F_{n-1} . Note that, the ratios $C_{2k-1}^{(n)}$ and $C_{2k}^{(n)}$ do not necessarily have the same probability distribution and they are not independent.

As we assume that, for all steps n , with $n = 1, 2, \dots$, all the ratios $C_{2k-1}^{(n)}$ have the same probability distribution, we can use in particular the ratio:

$$C_1 = C_1^{(1)} = \frac{\tilde{J}_1^{(1)}}{\tilde{J}_1^{(0)}} = \frac{\tilde{J}_1^{(1)}}{1} = \tilde{J}_1^{(1)} = X_{1;2}^{(0,1)} - 0 = X_{1;2}^{(0,1)} \quad (3)$$

and similarly, as we assume that in each step the ratios $C_{2k}^{(n)}$ do have the same probability distribution, the ratio is:

$$C_2 = C_2^{(1)} = \frac{\tilde{J}_2^{(1)}}{\tilde{J}_1^{(0)}} = \frac{\tilde{J}_2^{(1)}}{1} = \tilde{J}_2^{(1)} = 1 - X_{2;2}^{(0,1)}. \quad (4)$$

In view of the geometric construction of the above random sets, and of Falconer's Theorem 15.1, [4], we compute the Hausdorff dimension of the random $Beta(p, q)$ -Cantor set using the equation $\mathbf{E}[C_1^s + C_2^s] = 1$.

For $X_1^{(0)} \cap Beta(p, q)$, the probability density functions of the minimum and of the maximum of a two dimensional random sample are given by:

$$f_{X_{1;2}^{(0,1)}}(x) = 2 (1 - F_{X_1^{(0)}}(x)) f_{X_1^{(0)}}(x) \quad \text{and} \quad f_{X_{2;2}^{(0,1)}}(x) = 2 F_{X_1^{(0)}}(x) f_{X_1^{(0)}}(x),$$

where $F_{X_1^{(0)}}(x)$ and $f_{X_1^{(0)}}(x)$ are:

$$F_{X_1^{(0)}}(x) = \int_0^x \frac{1}{B(p, q)} t^{p-1} (1-t)^{q-1} dt \mathbf{I}_{(0,1)}(x) + \mathbf{I}_{(1,+\infty)}(x)$$

and

$$f_{X_1^{(0)}}(x) = \frac{1}{B(p, q)} x^{p-1} (1-x)^{q-1} \mathbf{I}_{(0,1)}(x).$$

The expected values of the transformations of the random variables C_1^s and C_2^s are given by:

$$\mathbf{E}\left[\left(X_{1;2}^{(0,1)}\right)^s\right] = \int_0^1 x^s f_{X_{1;2}^{(0,1)}}(x) dx \quad \text{and} \quad \mathbf{E}\left[\left(1 - X_{2;2}^{(0,1)}\right)^s\right] = \int_0^1 x^s f_{X_{2;2}^{(0,1)}}(1-x) dx.$$

Now, we can state the following result:

Theorem 4: Whith probability 1, the random $Beta(p, q)$ -Cantor set $F_{p, q}$ has

Rocha et al.

Hausdorff dimension $Dim_H F_{p,q}$ equal to s , where s is the solution of the equation

$$\mathbf{E}\left[\left(X_{1:2}^{(0,1)}\right)^s\right] + \mathbf{E}\left[\left(1 - X_{2:2}^{(0,1)}\right)^s\right] = 1.$$

On the other hand, to determine the Hausdorff dimension of the deterministic $Beta(p,q)$ -Cantor set, we use an analogous version of the Theorem 9.3 of Falconer, [4]. Nevertheless, it is necessary to determine the similarity ratios used in each step of the deterministic $Beta(p,q)$ -Cantor set construction. Considering that the randomness of the $Beta(p,q)$ -Cantor sets comes from the minimum and the maximum of a random variable with beta distribution, we have only two similarities.

Having in mind that, in each step n the ratios $C_{2^{k-1}}^{(n)}$ have the same probability distribution, and the ratios $C_{2^k}^{(n)}$ also have the same probability distribution, as explained before, and using (3) and (4), then:

$$\mathbf{E}[C_1] = \mathbf{E}\left[X_{1:2}^{(0,1)}\right] \text{ and } \mathbf{E}[C_2] = 1 - \mathbf{E}\left[X_{2:2}^{(0,1)}\right].$$

Therefore, the similarity ratios are $c_1 = \mathbf{E}\left[X_{1:2}^{(0,1)}\right]$ and $c_2 = 1 - \mathbf{E}\left[X_{2:2}^{(0,1)}\right]$ to the left and right intervals, respectively. The expressions of $\mathbf{E}\left[X_{1:2}^{(0,1)}\right]$ and $\mathbf{E}\left[X_{2:2}^{(0,1)}\right]$ were calculated from Eqs. (1) and (2). So, we can state the following result:

Theorem 5: Let S_i denote the similarities defined on \mathbb{R} , with ratios c_i , with $i = 1, 2$. If $C_{p,q}$ is the invariant set satisfying $C_{p,q} = \bigcup_{i=1}^2 S_i(C_{p,q})$ then

$$Dim_H C_{p,q} = s \text{ where } s \text{ is the solution of the equation}$$

$$\left(\mathbf{E}\left[X_{1:2}^{(0,1)}\right]\right)^s + \left(1 - \mathbf{E}\left[X_{2:2}^{(0,1)}\right]\right)^s = 1.$$

4. Numerical Simulations

At first sight the intuitive (but misguided) idea is that the random fractal $F_{p,q}$ should have a bigger Hausdorff dimension than the correspondent ‘‘mean fractal’’ $C_{p,q}$; In fact, the distribution of the retained spacings is skewed, and thus, the probability of retaining larger portions than in the deterministic fractal is higher than 0.5. But as a consequence, the removed portion in later steps is big, and the overall effect is that the random fractal is less dense, as explained below.

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 denote 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” $C_{p,q}$, observe that in the first step we obtain $J_1^{(1)} \cup J_2^{(1)} = [0, a] \cup [b, 1]$, where $a = a_1^{(0)} = E[X_{1:2}^{(0,1)}]$ and $b = b_1^{(0)} = E[X_{2:2}^{(0,1)}]$. A straightforward extension is stated in the theorem that follows:

Theorem 6: The length of the sum of the intervals removed in the construction of a “mean fractal” $C_{p,q}$, until the step n , is given by

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

with $n = 1, 2, \dots$, where $a = E[X_{1:2}^{(0,1)}]$ and $b = E[X_{2:2}^{(0,1)}]$.

The proof of this theorem can be seen in [2].

In Table 1, we compute the probability that the accumulated length of the random middle sets removed in the recursive construction of the random *Beta*(p,q)-Cantor set $F_{p,q}$ be greater than the accumulated length of removed subintervals in the construction of the corresponding deterministic *Beta*(p,q)-Cantor set $C_{p,q}$, for some values of p and q . While in the first steps this probability is less than 0.5, in the next steps, 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 correspondence, 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 indeed a coherent result. This is a consequence of Jensen's inequality. Although the Hausdorff dimensions of both the random and the corresponding deterministic *Beta*(p,q)-Cantor sets increase with p , we have

Rocha et al.

$$\text{Dim}_H F_{p,q} < \text{Dim}_H C_{p,q}.$$

Table 1: Estimated probability of $P[S_{2,R}^{(n)} > S_{2,D}^{(n)}]$ and respective confidence interval

Step	Beta(1,1)		Beta(2,1)		Beta(3,1)	
	Est.Prob	95% Conf. Int.	Est.Prob	95% Conf. Int.	Est.Prob	95% Conf. Int.
1	0.4570	(0.4432;0.4708)	0.4252	(0.4115;0.4389)	0.4222	(0.4085;0.4358)
2	0.5148	(0.5010;0.5287)	0.4834	(0.4696;0.4973)	0.4890	(0.4751;0.5028)
3	0.5516	(0.5378;0.5654)	0.5176	(0.5038;0.5315)	0.5182	(0.5043;0.5320)
4	0.5744	(0.5607;0.5881)	0.5354	(0.5216;0.5492)	0.5246	(0.5107;0.5384)
5	0.5776	(0.5639;0.5913)	0.5430	(0.5292;0.5492)	0.5402	(0.5263;0.5540)
6	0.5936	(0.5800;0.6072)	0.5462	(0.5324;0.5600)	0.5476	(0.5338;0.5613)
7	0.6002	(0.5866;0.6138)	0.5512	(0.5374;0.5649)	0.5524	(0.5386;0.5661)
8	0.6048	(0.5913;0.6184)	0.5496	(0.5358;0.5633)	0.5556	(0.5418;0.5693)
9	0.6084	(0.5949;0.6219)	0.5528	(0.5390;0.5665)	0.5616	(0.5478;0.5753)
10	0.6098	(0.5963;0.6233)	0.5500	(0.5362;0.5637)	0.5630	(0.5492;0.5767)
Step	Beta(1,2)		Beta(2,2)		Beta(3,2)	
	Est.Prob	95% Conf. Int.	Est.Prob	95% Conf. Int.	Est.Prob	95% Conf. Int.
1	0.4398	(0.4260;0.4536)	0.4300	(0.4163;0.4437)	0.4230	(0.4093;0.4367)
2	0.5012	(0.4873;0.5151)	0.4898	(0.4759;0.5037)	0.4836	(0.4697;0.4975)
3	0.5252	(0.5114;0.5390)	0.5128	(0.4989;0.5277)	0.5034	(0.4895;0.5173)
4	0.5374	(0.5236;0.5512)	0.5222	(0.5084;0.5360)	0.5190	(0.5052;0.5328)
5	0.5454	(0.5316;0.5592)	0.5232	(0.5094;0.5370)	0.5282	0.5144;0.5420)
6	0.5574	(0.5436;0.5712)	0.5266	(0.5128;0.5404)	0.5346	(0.5207;0.5484)
7	0.5638	(0.5501;0.5775)	0.5254	(0.5116;0.5392)	0.5388	(0.5250;0.5526)
8	0.5636	(0.5499;0.5773)	0.5266	(0.5128;0.5404)	0.5470	(0.5332;0.5608)
9	0.5666	(0.5529;0.5803)	0.5258	(0.5120;0.5396)	0.5492	(0.5354;0.5630)
10	0.5684	(0.5547;0.5821)	0.5274	(0.5136;0.5412)	0.5502	(0.5364;0.5640)

Acknowledgements

This research has been partially supported by FCT/OE, CEAUL and ISEL.

References

- [1] S. M. Aleixo, J. L. Rocha and D. D. Pestana. Populational growth models proportional to beta densities with Allee effect. *Am. Inst. of Phys.*, 1124: 3-12, 2008.
- [2] S. M. Aleixo, J. L. Rocha and D. D. Pestana. Dynamical behaviour in the parameter space: new populational growth models proportional to beta densities. In Luzar-Stiffler, V., Jarec, I. and Bekic, Z. (eds.), *Proc. of the ITI 2009, 31th Int. Conf. on Information Technology Interfaces (ITI 2009)*, 213-218, 2009.
- [3] S. M. Aleixo, J. L. Rocha and D. D. Pestana. Dynamical analysis in probabilistic models: random Beta(p,2)-Cantor sets, *Notas e Comunicações do CEAUL*, 2010.
- [4] K. Falconer. *Fractal Geometry, Mathematical Foundations and Applications*, John Wiley & Sons, New York, 1990.
- [5] D. D. Pestana, S. M. Aleixo and J. L. Rocha. The Beta(p,1) extensions of the random (uniform) Cantor sets. *Disc. Math. Prob. Stat.*, 29, 2009, in press.