Recursive overlap Bernoulli distributions and an entropy concavity conjecture

Jörg Neunhäuserer
Technical University of Braunschweig
joerg.neunhaeuserer@web.de

Abstract

We introduce a family of recursively generated finite probability distributions obtained from left and right embeddings with overlaps. The construction interpolates between the classical binomial distribution and the non-overlapping Bernoulli product distribution. We derive explicit formulas for the expectation, variance, and the generating function of higher moments, and formulate a conjecture asserting that the Shannon entropy is concave. The conjecture is proved in the two extremal cases and supported by symbolic computations for numerous overlap sequences.

MSC 2020: Primary 60C05; Secondary 94A17, 60G50.

Keywords: Bernoulli distributions, recursive probability distributions, Shannon entropy, entropy concavity, overlap constructions.

1 Introduction

In this note we introduce a family of recursive overlap Bernoulli distributions generated in the following way. Let (a_n) be a strictly increasing sequence of integers with $a_1 = 2$ and let $p \in [0,1]$. Define probability vectors $P_n(p) \in \mathbb{R}^{a_n}$ by $P_1(p) = (p, 1-p)$ and

$$P_{n+1}(p) = p\, L(P_n(p)) + (1-p)\, R(P_n(p)),$$

where the "left" and "right" embedding operators $L : \mathbb{R}^{a_n} \to \mathbb{R}^{a_{n+1}}, R : \mathbb{R}^{a_n} \to \mathbb{R}^{a_{n+1}}$ are given by

$$L(P_n(p)) = (P_n(p), 0, \dots, 0), \quad R(P_n(p)) = (0, \dots, 0, P_n(p)).$$

Let $X_n(p)$ be the corresponding random variable taking values in $\{1, \dots, a_n\}$. If $(a_n) = (n+1)$ the random variables $X_n(p)$ are binomial distributed with

$$\mathbb{P}(X_n(p) = k) = \binom{n}{k-1} p^{n-k+1}(1-p)^{k-1}$$

for $k = 1, \dots, n+1$. If $(a_n) = (2^n)$ the random variables form the non-overlapping Bernoulli with

$$\mathbb{P}(X_n(p) = k) = p^{n-\sharp_n(k)}(1-p)^{\sharp_n(k)}$$

for $k = 1, \dots, 2^n$, where $\sharp_n(k)$ is the number of digits 1 in the dyadic representation of $k-1$. If $a_{n+1} \geq 2a_n$ no overlap occurs and $X_n(p)$ has the same distribution with values

in a subset of $\{1, \ldots, a_n\}$ of cardinality 2^n. If $n+1 < a_{n+1} < 2a_n$ for some n, then the two copies overlap. The overlap creates new probability distributions which interpolate between the binomial distribution and the separated Bernoulli convolutions.
In the next section we will calculate expectation, variance and moment generating function of the random variables $X_n(p)$. In section 3 we state the entropy conjecture that the Shannon entropy of $X_n(p)$ is concave in p for all $n \geq 1$ and hence an unimodal function with maximum at $p = 1/2$. We prove this conjecture in the binomial case $a_n = n+1$ and for all sequences with $a_{n+1} \geq 2a_n$. In the last section we present computational evidence that the conjecture holds in general. The motivation for this construction and the entropy conjecture is twofold. First, the construction generalizes random walks on infinite rooted self-similar graphs studied in [4]. For example, for the recursive sequence

$$a_{n+1} = a_n + a_{n-1} + 1, \qquad a_0 = 1,\ a_1 = 2,$$

we recover the directed random walk on the Fibonacci graph. Second, the construction is closely related to infinite convolved Bernoulli measures, which play an important role in fractal geometry; see Chapter 8 of [1] and the references therein. If Conjecture 3.1 is true, then the dimension function of the corresponding singular Bernoulli measures would be concave.

2 Expectation, variance and higher moments

We first give here the expectation and variance of the random variables defined in the introduction.

Proposition 2.1 *For a strictly increasing sequence of integers (a_n) with $a_1 = 2$ and $p \in [0,1]$ the expectation of $X_n(p)$ is*

$$\mathbb{E}(X_n(p)) = p + a_n(1-p)$$

the variance is

$$\mathbb{V}(X_n(p)) = p(1-p)(1 + \sum_{k=1}^{n-1}(a_{k+1} - a_k)^2)$$

Proof. From the definition we have

$$X_{n+1}(p) = \begin{cases} X_n & \text{with probability p} \\ X_n + (a_{n+1} - a_n) & \text{with probability (1-p)} \end{cases}$$

and $X_1 \in \{1, 2\}$ with probabilities p and $(1-p)$. Hence

$$\mathbb{E}(X_{n+1}(p)) = \mathbb{E}(X_n(p)) + (1-p)(a_{n+1} - a_n)$$

with $\mathbb{E}(X_1) = p + 2(1-p)$. The formula for the expectation follows since the sum is telescoping. Moreover we get

$$\mathbb{V}(X_{n+1}(p)) = \mathbb{V}(X_n(p)) + p(1-p)(a_{n+1} - a_n)^2$$

with $\mathbb{V}(X_1) = p(1-p)$. The formula for variance follows by induction. □

Now we consider the generating functions for higher moments.

Proposition 2.2 *For a strictly increasing sequence of integers (a_n) with $a_1 = 2$ and $p \in [0,1]$ the moment generating function of $X_n(p)$ is given by*

$$M_n(t) = \mathbb{E}\big[e^{tX_n}\big] = e^t\big(p + (1-p)e^t\big) \prod_{k=1}^{n-1} \Big(p + (1-p)\, e^{\,t\,(a_{k+1} - a_k)}\Big).$$

Proof. We have

$$M_{n+1}(t) = \mathbb{E}\big[e^{tX_{n+1}}\big] = p\,\mathbb{E}\big[e^{tX_n}\big] + (1-p)\,\mathbb{E}\big[e^{t(X_n + a_{n+1} - a_n)}\big].$$

$$= p\,M_n(t) + (1-p)e^{t(a_{n+1}-a_n)} M_n(t) = \big(p + (1-p)e^{t(a_{n+1}-a_n)}\big)\, M_n(t).$$

Since

$$M_1(t) = \mathbb{E}\big[e^{tX_1}\big] = p\,e^t + (1-p)\,e^{2t} = e^t\big(p + (1-p)e^t\big),$$

iterating the recursion yields the product formula. □

3 The entropy concavity conjecture

The Shannon entropy of the random variables $X_n(p)$ is given by

$$h_n(p) = H(X_n(p)) = -\sum_{k=1}^{a_n} \mathbb{P}(X_n(p) = k) \log(\mathbb{P}(X_n(p) = k)),$$

where we use the convention that $0 \log(0) = 0$, see [2]. We obviously have $h_n(0) = h_n(1)$. The symmetry

$$h_n(p) = h_n(1-p)$$

follows from

$$\mathbb{P}(X_n(p) = k) = \mathbb{P}(X_n(1-p) = a_n + 1 - k).$$

Now we state our entropy conjecture.

Conjecture 3.1 *For every strictly increasing sequence of integers (a_n) with $a_1 = 2$ and all $n \geq 1$ the entropy function h_n is concave on $[0,1]$ and hence unimodal with maximum at $1/2$.*

We prove this conjecture in two extremal cases.

Proposition 3.1 *For $a_n = n+1$ and for all sequences satisfying $a_{n+1} \geq 2a_n$ and $a_1 = 2$, Conjecture 3.1 holds.*

Proof. If $a_n = n+1$, as we already observed in the introduction $X_n(p)$ is just a shift of the Binomial distribution; $X_n(p) \sim \text{Binomial}(p, n) + 1$. Such a Shift obviously does not change the entropy. Hence the result follows from the Shepp and Olkin theorem, that states that the Shannon entropy of the binomial distribution is concave, see [3].
If $a_{n+1} \geq 2a_n$ we obtain using the probability formula from the introduction

$$h_n(p) = -\sum_{k=1}^{2^n} p^{n-\sharp_n(k)}(1-p)^{\sharp_n(k)} \log(p^{n-\sharp_n(k)}(1-p)^{\sharp_n(k)})$$

$$= -\sum_{i=0}^{n} \binom{n}{i} p^{n-i}(1-p)^i \log(p^{n-i}(1-p)^i) = -n(p\log(p) + (1-p)\log(1-p)),$$

where $\sharp_n(k)$ is the number of digits 1 in the dyadic representation of $k-1$. Now it is straightforward that the function is in fact concave. □

We have no proof of the concavity conjecture for other increasing infinite sequences. Nevertheless in the next section we obtain computational evidence for many classes of sequences.

4 Computational evidence

Given an increasing finite sequence (a_n) with $a_1 = 2$ we use a Mathematica program to test concavity of the entropy function $h_n(p)$ of the random variables $X_n(p)$. The Mathematica program used for all computations is available from the author. We found no counterexample among any of the tested sequences. We were able to test linear sequence

$$(a_n) = (2n), (3n-1), (5n-3), (10n-8)$$

up to $n = 200$. In figure 1 we display h_n and h_n'' for $n = 10, 50, 100$ in the case $(a_n) = (2n)$. We tested the polynomial sequences

$$(a_n) = (n^2+1), (n^3+1)$$

up to $n = 100$ in the first case and $n = 50$ in the second case. For the fist sequences We display h_n and h_n'' for $n = 10, 50, 100$ in figure 2. We tested the exponential increasing sequences $a_n = \lceil (3/2)^n \rceil$ and the recursively definiere sequence $a_{n+1} = a_n + a_{n-1} + 1$ with $a_0 = 1$ and $a_0 = 2$ up to $n = 30$. In the second case we include a figure of h_n and h_n''

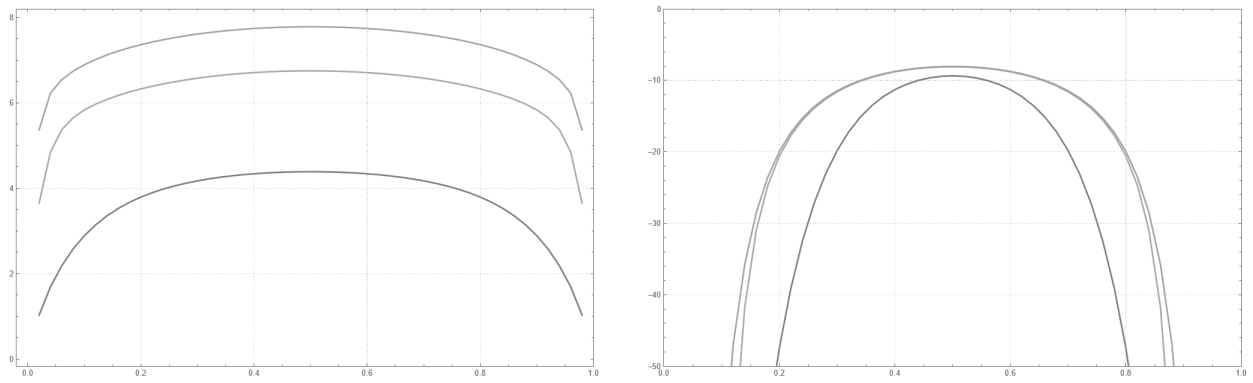

Figure 1: h_n and h_n'' for $n = 10, 50, 100$ in the case $(a_n) = (2n)$.

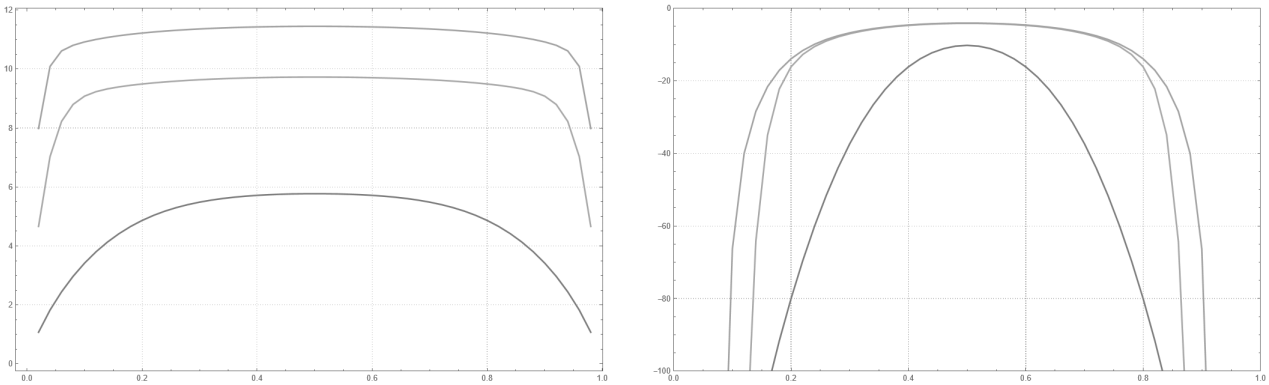

Figure 2: h_n and h_n'' for $n = 10, 50, 100$ in the case $(a_n) = (n^2 + 1)$

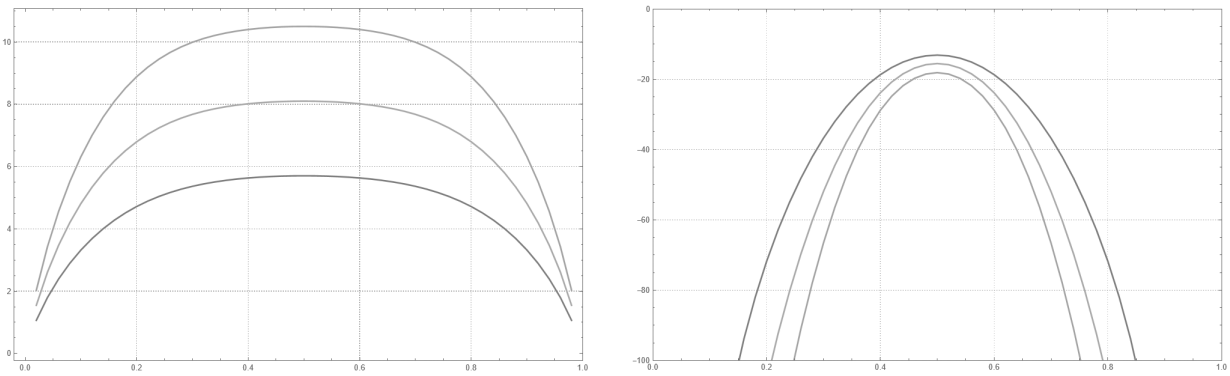

Figure 3: h_n and h_n'' for $n = 10, 15, 20$ in the case $a_{n+1} = a_n + a_{n-1} + 1$

with $n = 10, 15, 20$ for the second sequence. Moreover we testes stochastically increasing sequences up to $n = 100$. The increment of the sequence was chosen uniformly from $1, \ldots, 1000$. The random experiments suggest that, if a counterexample exists, it must occur only for considerably larger values of n or for much more irregular overlap sequences than those tested here.

References

[1] B. Bárány, K. Simon and B. Solomyak, Self-similar and Self-affine Sets and Measures, Mathematical Surveys and Monograph of the AMS, Vol. 276, (2023).

[2] C. E. Shannon, A Mathematical Theory of Communication, Bell System Technical Journal 27 (1948), 379–423, 623–656.

[3] L. Shepp and I. Olkin, Entropy of the sum of independent Bernoulli random variables and of the multinomial distribution, Proceedings of the Sixth Berkeley Symposium (1972), pp. 299–309.

[4] J. Neunhäuserer, Random walks on infinite self-similar graphs, Electronic Journal of Probability 12 (2007), 1258–1277.