ASYMPTOTIC ANALYSIS OF THE MOMENTS OF THE CANTOR DISTRIBUTION

Peter J. Grabner and Helmut Prodinger

Department of Mathematics
Technical University of Graz, Austria
and
Department of Algebra and Discrete Mathematics
Technical University of Vienna, Austria

Abstract. We use known recursive expressions for the moments of the Cantor distribution to derive asymptotic expansions for these moments. This is done by a combination of a method based on Mellin transform and the saddle point method.

The Cantor distribution with parameter ϑ, $0 < \vartheta < 1$ can be described by a random series

$$\frac{\bar\vartheta}{\vartheta}\sum_{i\ge 1} X_i\vartheta^i,$$

where the X_i are independent with the distribution

$$\mathbb{P}\{X_i = 0\} = \mathbb{P}\{X_i = 1\} = \frac{1}{2},$$

and $\bar\vartheta = 1-\vartheta$. The essential result of the paper [7] is the following recursion for the moments $\mathbb{E}(X^N)$:

$$\mathbb{E}(X^N) = \frac{1}{2(1-\vartheta^N)}\sum_{i=0}^{N-1}\binom{N}{i}\vartheta^i\bar\vartheta^{N-i}\mathbb{E}(X^i), \qquad N\ge 1,\ \mathbb{E}(X^0)=1. \tag{1}$$

The aim of the present note is to solve the recursion (1). We abbreviate $a_N = \mathbb{E}(X^N)$ and rewrite the recursion as

$$2a_N - \vartheta^N a_N = \sum_{i=0}^{N}\binom{N}{i}\vartheta^i\bar\vartheta^{N-i}a_i \qquad N\ge 1,\ a_0 = 1. \tag{2}$$

Key words and phrases. Cantor distribution, Mellin transform, saddle point method.

This research was supported by the Austrian-Hungarian cooperation 10U3 and by the Austrian National Bank project Nr. 4995.

The first author is supported by the Schrödinger scholarship J00936-PHY

Now we introduce the exponential generating function

$$A(z)=\sum_{N\ge0}a_N\frac{z^N}{N!}.$$

Recursion (2) translates into

$$2A(z)-2-A(\vartheta z)+1=e^{\bar\vartheta z}A(\vartheta z)-1,$$

or

$$A(z)=\frac{1+e^{\bar\vartheta z}}{2}A(\vartheta z). \tag{3}$$

We can solve this equation by iteration, and we obtain (note that $A(0)=1$)

$$A(z)=\prod_{k\ge0}\frac{1+e^{\bar\vartheta\vartheta^k z}}{2}. \tag{4}$$

For $\vartheta=\frac12$, the product collapses to

$$A(z)=\frac{e^z-1}{z},$$

so that $a_n=\frac{1}{n+1}$, which of course can be seen directly, since the Cantor distribution is then just the uniform distribution over the interval $[0,1]$. Since a_N is the coefficient of $z^N/N!$, we have solved the recursion (1). More useful, however, is the asymptotic equivalent for a_N, which we are going to derive in the sequel. For that, we need also the Poisson generating function

$$B(z)=e^{-z}A(z)=\sum_{N\ge0}b_N\frac{z^N}{N!}.$$

Equation (3) translates then into

$$B(z)=\frac{1+e^{-\bar\vartheta z}}{2}B(\vartheta z), \tag{5}$$

yielding

$$B(z)=\prod_{k\ge0}\frac{1+e^{-\vartheta^k\bar\vartheta z}}{2}.$$

We note here that the result obtained above could be deduced easily by interpreting the Cantor distribution as the distribution function of a 2-additive function given by

$$f\left(\sum_{\ell=0}^{L}\varepsilon_\ell 2^\ell\right)=\bar\vartheta\sum_{\ell=0}^{L}\varepsilon_\ell\vartheta^\ell.$$

The notion of q-additive functions was introduced in [1], where a necessary and sufficient condition for the existence of a distribution function is given. The formula (4) is an immediate consequence of Delange's formula for the Fourier transform of this distribution function. Furthermore notice that for $\vartheta < \frac{1}{2}$ the moment generating function $A(z)$ can be written as the integral

$$A(z) = \int_{\mathcal{C}} e^{zt} d\mathcal{H}(t),$$

where $\mathcal{C}$ denotes the Cantor set given by the closure of the image $f(\mathbb{N})$ and $\mathcal{H}$ denotes the normalized Hausdorff-measure of dimension $\log_{\frac{1}{\vartheta}} 2$. The function $h(x) = \int_{[0,x]\cap\mathcal{C}} d\mathcal{H}(t)$ is a singular function, i.e. it is monotonically increasing, thus differentiable almost everywhere and the derivative vanishes almost everywhere despite of the fact that $h(0) = 0$ and $h(1) = 1$.

Let us now start the asymptotic analysis of a_N. For this purpose we compute the Mellin transform (cf. [2,3]) of the logarithm of $B(z)$

$$(\log B)^*(s) = \int_0^\infty \log B(x) x^{s-1} dx = \frac{\bar{\vartheta}^{-s}}{1-\vartheta^{-s}} \int_0^\infty \log\left(\frac{1+e^{-x}}{2}\right) x^{s-1} dx \tag{6}$$

for $-1 < \Re s < 0$. The remaining integral is easily computed as

$$\int_0^\infty \log\left(\frac{1+e^{-x}}{2}\right) x^{s-1} dx = \Gamma(s)\zeta(s+1)\left(1-2^{-s}\right),$$

again for $-1 < \Re s < 0$. Thus by Mellin's inversion formula we have

$$\log B(z) = \frac{1}{2\pi i} \int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} \Gamma(s)\zeta(s+1)\left(1-2^{-s}\right) \frac{\bar{\vartheta}^{-s}}{1-\vartheta^{-s}} z^{-s} ds.$$

This formula holds for every z with $|\arg z| < \frac{\pi}{2}$. By shifting the line of integration to the right and taking the residues at $s = 0$ (double pole!) and $s = \chi_k := \frac{2k\pi i}{\log\frac{1}{\vartheta}}$ for $k \in \mathbb{Z}\setminus\{0\}$ (simple poles) into account we obtain

$$\begin{aligned}\log B(z) = &-\log_\Theta 2 \cdot \log \bar{\vartheta} z - \frac{\log 2}{2} - \frac{\log 2 \cdot \log_\Theta 2}{2}\\ &+ \frac{1}{\log\Theta} \sum_{k\in\mathbb{Z}\setminus\{0\}} \Gamma(\chi_k)\zeta(1+\chi_k)\left(1-2^{-\chi_k}\right) e^{-\chi_k \log\bar{\vartheta}} e^{2k\pi i \log_\Theta z}\\ &+ \frac{1}{2\pi i} \int_{M-i\infty}^{M+i\infty} \Gamma(s)\zeta(s+1)\left(1-2^{-s}\right) \frac{\bar{\vartheta}^{-s}}{1-\vartheta^{-s}} z^{-s} ds,\end{aligned} \tag{7}$$

where $\Theta = \frac{1}{\vartheta}$ and M is any positive real number. Since the remaining integral is convergent by the well-known asymptotic behaviour of the Γ-function, the remainder term is a $O(z^{-M})$ for any $M > 0$. From this we derive

$$B(z) = F(\log_\Theta z) z^{-\log_\Theta 2}\left(1 + O(z^{-M})\right)$$

with an infinitely differentiable 1-periodic function $F(x)$.

The "de-Poissonization" technique (cf. [4,8]) suggests the approximation $a_N \sim B(N)$. Applying it to every term of the Fourier-expansion of $F(x)$ separately (the Fourier series is uniformly convergent), we obtain

$$a_N = F(\log_\Theta N) N^{-\log_\Theta 2} \cdot \left(1 + O\left(\frac{1}{N}\right)\right). \tag{8}$$

In order to derive an explicit expression for the Fourier coefficients of $F(x)$ we take the Mellin transform of (5)

$$B^*(s) = \frac{1}{2}\Theta^s B^*(s) + \frac{1}{2}\int_0^\infty B(\vartheta x) e^{-\bar{\vartheta} x} x^{s-1} dx.$$

By the above computations we know that this transform exists for $0 < \Re s < \log_\Theta 2$. The last equation yields

$$B^*(s) = \frac{1}{2 - \Theta^s}\int_0^\infty B(\vartheta x) e^{-\bar{\vartheta} x} x^{s-1} dx,$$

where the integral converges for every s with $\Re s > 0$. By Mellin's inversion formula the Fourier coefficients of $F(x)$ are equal to the negative residues of $B^*(s)$ at $s = \log_\Theta 2 + \chi_k$, $k \in \mathbb{Z}$. These are given by

$$\hat{F}(k) = -\operatorname{Res} B^*(s)\Big|_{s=\log_\Theta 2 + \chi_k} = \frac{1}{2\log\Theta}\int_0^\infty B(\vartheta x) e^{-\bar{\vartheta} x} x^{\log_\Theta 2 - 1 + \chi_k} dx.$$

These integrals can be easily computed numerically, because they are rapidly convergent for $x \to \infty$ and the behaviour for $x \to 0$ is regular.

The value

$$\hat{F}(0) = \frac{1}{2\log\Theta}\int_0^\infty B(\vartheta x) e^{-\bar{\vartheta} x} x^{\log_\Theta 2 - 1} dx$$

is of special interest, as it is the *mean value* around which the periodic function $F(x)$ fluctuates. Since the amplitudes of these fluctuations are usually quite small, we can write

$$a_N \simeq \hat{F}(0) \cdot N^{-\log_\Theta 2},$$

which is suggestive, but not quite correct, as it ignores the fluctuations.

Remark. *Notice that $\hat{F}(0)$ can be viewed as the following limit*

$$\hat{F}(0) = \lim_{N\to\infty} \frac{1}{\log N}\sum_{n=1}^{N} a_n n^{\log_\Theta 2 - 1}.$$

This is due to the fact that $\sum_{n<N} n^{-1+it} = O(1)$ for $t \neq 0$.

To illustrate the results, we give a table of $\hat{F}(0)$, for several values of ϑ, and compare them with the values of the recursion, for $N = 50$, $N = 100$ and $N = 200$.

Furthermore we present a plot of $a_N N^{\log_{10} 2}$ versus $\log_{10} N$ for $\vartheta = 0.1$ to illustrate the fluctuating behaviour of this function.

For the reader's convenience we summarize our findings.

Theorem. *The moments* $\mathbb{E}(X^N)$ *of the Cantor distribution with parameter* ϑ *are given as the coefficient of* $\frac{z^N}{N!}$ *in*

$$\prod_{k\geq 0}\frac{1+e^{(1-\vartheta)\vartheta^k z}}{2}.$$

Asymptotically, we have

$$\mathbb{E}(X^N)=F(\log_{1/\vartheta}N)N^{-\log_{1/\vartheta}2}\cdot\left(1+O(\frac{1}{N})\right),$$

where $F(x)$ *is a periodic function of period 1 and known Fourier coefficients. The mean of* $F(x)$ *is given by*

$$-\frac{1}{2\log\vartheta}\int_0^\infty\prod_{k\geq 1}\frac{1+e^{-(1-\vartheta)\vartheta^k x}}{2}e^{-(1-\vartheta)x}x^{\log_{1/\vartheta}2-1}dx.$$

We note here that in the case $\vartheta=2^{-1/m}$ for a positive integer m the function F is constant, because all Fourier-coeffients in (7) vanish (except the one for $k=0$). This corresponds to the fact that $A(z)$ can be given explicitly by

$$A(z)=\bar{\vartheta}^{-m}2^{-\frac{m+1}{2}}z^{-m}\prod_{r=0}^{m-1}\left(e^{2\bar{\vartheta}\vartheta^r z}-1\right)\sim\bar{\vartheta}^{-m}2^{-\frac{m+1}{2}}z^{-m}e^{z},$$

which yields

$$a_N\sim\bar{\vartheta}^{-m}2^{-\frac{m+1}{2}}N^{-m}$$

for this case.

Furthermore, we note that in the context of the order statistics of the Cantor distribution a recursion occurred in [5] which was solved in [6]. In this case, a somewhat more direct approach could be used, since explicit formulæ for the quantities of interest were available, whereas here we have only the generating function $A(z)$, from which we have to extract the necessary information.

References

1. H.Delange, *Sur les fonctions q-additives ou q-multiplicatives*, Acta Arith. **21** (1972), 285–296.
2. G.Doetsch, *Handbuch der Laplace-Transformation*, Birkhäuser, Basel, 1958.
3. P.Flajolet, X.Gourdon, and P.Dumas, *Mellin transforms and asymptotics: Harmonic sums*, Theoret. Comput. Sci. (1995) (to appear).
4. P.Grabner, *Searching for losers*, Random Structures and Algorithms **4** (1993), 99–110.
5. J.R.M.Hosking, *Moments of order statistics of the Cantor distribution*, Statistics and Probability Letters **19** (1994), 161–165.
6. A.Knopfmacher and H.Prodinger, *Explicit and asymptotic formulæ for the expected values of the order statistics ot the Cantor distribution*, submitted (1994).
7. F.R.Lad and W.F.C.Taylor, *The moments of the Cantor distribution*, Statistics and Probability Letters **13** (1992), 307–310.
8. B.Rais, P.Jacquet, and W.Szpankowski, *Limiting distributions for the depth in Patricia tries*, SIAM J. Discr. Mathematics **6** (1993), 197–213.

P. Grabner
Institut für Mathematik A
TU Graz
Steyrergasse 30
A-8010 Vienna
Austria

H. Prodinger
Institut für Algebra und Diskrete Mathematik
TU Vienna
Wiedner Hauptstrasse 8–10
A-1040 Vienna
Austria

E-mail address: grabner@ftug.dnet.tu--graz.ac.at
proding@rsmb.tuwien.ac.at