Monatsh Math (2014) 175:333–346
DOI 10.1007/s00605-014-0676-5

Order statistics of the generalised multinomial measure

Ligia L. Cristea · Helmut Prodinger

Received: 28 April 2011 / Accepted: 27 July 2014 / Published online: 29 August 2014

Abstract We study certain order statistics with respect to (probability) mass distributions of multinomial type on the unit interval. The asymptotic behaviour of the average minimum and, respectively, maximum value among n words chosen independently at random with respect to the corresponding probability measure is analysed. This is done by a combination of a method based on the Mellin transform and the depoissonisation technique.

Keywords Multinomial measure · Order Statistics · Depoissonisation · Mellin transform

Mathematics Subject Classification 30E05 · 60C05

Communicated by A. Constantin.

This paper is dedicated to the memory of Philippe Flajolet who passed away on March 22, 2011.

L. L. Cristea was supported by the Austrian Science and Research Fund (FWF), Project P20412-N18.
H. Prodinger was supported by an incentive grant from the NRF.

L. L. Cristea (✉)
Institute for Mathematics, Technische Universität Graz,
Steyrergasse 30, 8010 Graz, Austria
e-mail: strublistea@gmail.com

H. Prodinger
Mathematics Department, Stellenbosch University,
Stellenbosch 7602, South Africa
e-mail: hproding@sun.ac.za

1 Introduction

In [5] the authors introduce the multinomial measure on the unit interval in the following way. Let $q \geq 2$ be a positive integer. Denote $I = I_{0,0} = [0, 1]$ and

$$I_{n,j} = \left[\frac{j}{q^n}, \frac{j+1}{q^n}\right), \text{ for } j = 0, 1, \ldots, q^n - 2, \qquad I_{n,q^n-1} = \left[\frac{q^n-1}{q^n}, 1\right],$$

for $n = 1, 2, 3, \ldots$. Let $\mathbf{r} = (r_0, r_1, \ldots, r_{q-1})$ with $0 \leq r_i \leq 1$ and $\sum_{k=0}^{q-1} r_k = 1$.

The *multinomial measure* $\mu_{q,\mathbf{r}}$ is the probability measure on I defined by

$$\mu_{q,\mathbf{r}}(I_{n+1,qj+k}) = r_k \cdot \mu_{q,\mathbf{r}}(I_{n,j})$$

for $n = 0, 1, 2, \ldots, j = 0, 1, \ldots, q^n - 1, k = 0, 1, \ldots, q - 1$. For further details about properties of the multinomial measure we refer to [5].

In Sect. 2 we introduce the *generalised multinomial measure*. Here a generalisation consists, roughly speaking, in the fact that instead of dividing the unit interval into a finite number of subintervals of equal length, we divide it into infinitely (and denumerably) many intervals, such that the j-th interval has length pq^{j-1}, where $p = 1 - q$. One way to define the generalised multinomial measure is the following. We consider the set $\mathcal{W}$ of all (finite and infinite) words over the infinite alphabet $\mathbb{N}_0 = \{0, 1, \ldots\}$ and a probability measure $\mathbb{P}_\mathbf{r}$ defined on the set of all words. A function value associates to every word ω in $\mathcal{W}$ a real number $\mathsf{value}(\omega) \in [0, 1)$, such that the closure of the set of all such values, $\mathsf{value}(\mathcal{W})$, is the interval $[0, 1]$. Then the measure of an interval $\mu_{q,\mathbf{r}}([0, a))$, $0 \leq a \leq 1$ can be defined in a natural way as being the probability $\mathbb{P}_\mathbf{r}$ that a word of $\mathcal{W}$ has the value less than or equal to a.

Section 3 is dedicated to the study of the behaviour of the average minimum value a_n among n words of $\mathcal{W}$ chosen independently at random with respect to the multinomial measure $\mu_{\mathbf{r},q}$, for $r_j = \lambda v^j$, $j = 0, 1, \ldots.$, where $0 < v < 1$, and $v = 1 - \lambda$, which we denote by $\mu_{v,q}$. First, we establish a recursion for a_n. In the sequel, we use the exponential generating function and combine a method based on the Mellin transform (see, e.g., Flajolet et. al [3]) and the depoissonisation technique (see, e.g., Jacquet and Szpankowski [4] and Szpankowski [6]) for the study of the asymptotics of the average minimum value a_n.

In the last section the issues of the previous section are studied for the average maximum value among n words of $\mathcal{W}$ chosen independently at random with respect to the measure $\mu_{v,q}$. We note that the final formulae obtained for the asymptotics show a certain duality with respect to those of the previous section.

We mention that similar questions were also addressed by Bassino and Prodinger who studied order statistics [1], where the interest was in general q-ary expansions with missing digits, and by authors of the present work in a paper on the Cantor-Fibonacci measure [2].

2 The generalised multinomial measure

Let $\mathcal{A}$ be a denumerable set $\{a_1, a_2, \ldots\}$ which we call *alphabet*. For simplicity we will assume, without loss of generality, $\mathcal{A} = \{0, 1, \ldots\}$ along this paper, i.e., $\mathcal{A} = \mathbb{N}_0$.

We introduce some notations: Let $\mathcal{W}$ denote the set of all (finite and infinite) words over the alphabet $\mathcal{A}$ and $\mathcal{W}_m$ the set of all words of length m ($m \geq 1$) over the alphabet $\mathcal{A}$. For the integers $l, m \geq 1$, $l \geq m$ and a word $\omega \in \mathcal{W}$, $\omega = \omega_1\omega_2\ldots$ of length l or ∞ let $\omega^{(m)}$ denote the word $\omega_1\ldots\omega_m$. Obviously we have $\mathcal{W}_1 = \mathcal{A}$. We denote by $\mathcal{W}_\infty$ the set of all words of infinite length over $\mathcal{A}$.

A measure on $\mathcal{W}$ may be constructed in the following way: Let $\mathbf{r} = \{r_0, r_1, \ldots\}$ be an arbitrarily fixed sequence of real numbers such that $r_j > 0$ for all $j \geq 0$ and $\sum_{j=0}^{\infty} r_j = 1$.

We introduce a probability measure on $\mathcal{W}$ in an inductive manner.

Definition 1 For any $\omega, \omega' \in \mathcal{W}$, $\omega = \omega_1\omega_2\ldots$, and for any $k \in \mathbb{N}_0$,

$$\mathbb{P}_{\mathbf{r}}(\omega_1 = k) := r_k \quad \text{and} \quad \mathbb{P}_{\mathbf{r}}(\omega = k\omega') := r_k \cdot \mathbb{P}_{\mathbf{r}}(\omega_2\omega_3\cdots = \omega'), \tag{2.1}$$

where $k\omega$ denotes the (usual) concatenation of the letter k with the word ω'.

Now we construct a function that assigns a real value to every word of $\mathcal{W}$. Again we proceed inductively. Let $q \in (0, 1)$ be an arbitrarily fixed real number and let $p = 1 - q$. We define, for any $m \geq 1$, the function $\mathsf{value}_m : \mathcal{W}_m \to [0, 1)$, by

$$\mathsf{value}_1(k) = 1 - q^k \quad \text{and} \quad \mathsf{value}_m(k\omega) = \mathsf{value}_1(k) + pq^k \cdot \mathsf{value}_{m-1}(\omega), \tag{2.2}$$

for $\omega \in \mathcal{W}_{m-1}$.

Definition 2 The function $\mathsf{value} : \mathcal{W} \to [0, 1)$ is the (uniqe) real function with the property that for any $m \geq 1$ its restriction to $\mathcal{W}_m$ coincides with value_m.

We remark that the closure (with respect to the canonic topology on $\mathbb{R}$) of the set $\mathsf{value}(\mathcal{W})$ is the interval $[0, 1]$.

Remark An order relation on $\mathcal{W}$ denoted by $\leq^*$ can be introduced as follows:

(1) On $\mathcal{W}_1 = \mathcal{A} = \mathbb{N}_0$ $\leq^*$ coincides with the canonical order relation on $\mathbb{N}_0$.
(2) For $m \geq 2$ and $\omega, \omega' \in \mathcal{W}_m$, $\omega = \omega_1\ldots\omega_m$, $\omega' = \omega'_1\ldots\omega'_m$ we have if $\omega \leq^* \omega'$ either if $\omega_1 \leq^* \omega'_1$ or if there exists a $j \in \{1, \ldots, m-1\}$ such that $\omega_i = \omega'_i$, for all $1 \leq i \leq j$ and $\omega_{j+1} \leq^* \omega'_{j+1}$.
(3) For $\omega, \omega' \in \mathcal{W}$ we have $\omega \leq^* \omega'$ if there exists an integer $m \geq 1$ such that $\omega^{(m)} \leq^* \omega'^{(m)}$.

One can easily verify that the function value is strictly increasing with respect to $\leq^*$ and to the canonical order relation of real numbers.

The probability measure $\mathbb{P}_{\mathbf{r}}$ on $\mathcal{W}$ induces a probability measure $\mu_{\mathbf{r},q}$ on $[0, 1]$, given as follows.

Definition 3 The *generalised multinomial measure* (of parameters $\mathbf{r}$ and q) is the measure $\mu_{\mathbf{r},q}$ defined by

$$\mu_{\mathbf{r},q}\big([0,a)\big) := \mathbb{P}_{\mathbf{r}}\Big(\{\omega \in \mathcal{W} \mid \mathsf{value}(\omega) \le a\}\Big), \tag{2.3}$$

for any $a \in [0, 1]$.

Remark In the special case $r_l = q^l \cdot p$, for all $l \in \mathbb{N}_0$, one can show that $\mu_{\mathbf{r},q}$ coincides with the uniform distribution on the unit interval. Throughout this paper we consider the case $r_k = \lambda\nu^k$, where $0 < \nu < 1$ and $\lambda = 1 - \nu$.

Remark The multinomial measure can also be defined in the following equivalent manner. Given a real number $0 \le x < 1$, choose the smallest i such that $1-q^{i+1} \ge x$, and say that the first digit is i. The weight of digit i is $\lambda\nu^i$. We continue with $(x-pq^i)/q$. Moreover, if that process led to digits $d_1 d_2 \dots$, we define the value of x to be

$$(1-q^{d_1})+pq^{d_1}[(1-q^{d_2})+pq^{d_2}[(1-q^{d_3})+\cdots = \sum_{i\ge 1} p^{i-1}q^{d_1+\cdots+d_{i-1}}(1-q^{d_i}).$$

3 Order statistics of the generalised multinomial measure: the minimum

In the following we study order statistics of the function value with respect to the measure $\mu_{\mathbf{r},q}$, for $r_j = \lambda\nu^j$, $j = 0, 1, \dots$, where $0 < \nu < 1$, and $\nu = 1 - \lambda$, which we denote $\mu_{\nu,q}$.

3.1 The problem setting

We pick at random (with respect to the probability measure on $\mathcal{W}$ defined above), independently, n words from $\mathcal{W}_m$, for $n \ge 1$. We apply the function value defined above to each of the chosen words and look for the minimum among these n values. The same can be done with all random choices of n words of $\mathcal{W}_\infty$. Let us denote by $a_n^{(m)}$ the average minimal value among all possible choices of n words of length m. By taking the limit $a_n := \lim_{m\to\infty} a_n^{(m)}$ we obtain the average minimal value among all choices of n words of $\mathcal{W}_\infty$. We are interested in the study of the asymptotic behaviour of a_n, for $n \to \infty$.

The first step is to establish the recursion

$$a_n^{(m)} = \sum_{k=1}^{n} \binom{n}{k} \sum_{j=0}^{\infty} (\lambda\nu^j)^k (\nu^{j+1})^{n-k} \big(1 - q^j + pq^j \cdot a_k^{(m-1)}\big).$$

This is obtained from the relations in (2.2) based on the following idea. Let j be the minimum among the first letters of the n words, i.e., there is an integer k, $1 \le k \le n$ such that k words start with j, and the other $n-k$ words start with a letter greater than j.

By taking the limit for $m \to \infty$ in the above recursion we obtain

$$a_n = \sum_{k=1}^{n} \binom{n}{k} \lambda^k \nu^{n-k} \sum_{j=0}^{\infty} \nu^{jn} \big(1 - q^j + pq^j \cdot a_k\big).$$

This yields

$$a_n = \sum_{k=1}^{n} \binom{n}{k} \lambda^k \nu^{n-k} \left(\frac{1}{1-\nu^n} - \frac{1}{1-q\nu^n} + \frac{p}{1-q\nu^n} a_k \right),$$

and thus

$$a_n = 1 - \frac{1-\nu^n}{1-q\nu^n} + \frac{p}{1-q\nu^n} \sum_{k=1}^{n} \binom{n}{k} \lambda^k \nu^{n-k} a_k.$$

We obtain

$$a_n = \frac{p\nu^n}{1-q\nu^n} + \frac{p}{1-q\nu^n} \sum_{k=1}^{n} \binom{n}{k} \lambda^k \nu^{n-k} a_k.$$

Thus we have proven the following result.

Proposition 1 *The average minimum value among n words of infinite length over $\mathbb{N}_0$ with respect to the generalised multinomial measure $\mu_{\nu,q}$ satisfies the recursion*

$$a_n = \frac{p\nu^n}{1-q\nu^n} + \frac{p}{1-q\nu^n} \sum_{k=1}^{n} \binom{n}{k} \lambda^k \nu^{n-k} a_k, \quad \textit{for all integers } n \geq 1. \tag{3.1}$$

We set $a_0 = 0$, which is convenient for computational reasons. One can rewrite Eq. (3.1) as

$$a_n = \frac{p\nu^n}{1-p\lambda^n-q\nu^n} + \frac{p}{1-p\lambda^n-q\nu^n} \sum_{k=0}^{n-1} \binom{n}{k} \lambda^k \nu^{n-k} a_k \tag{3.2}$$

in order to compute the elements a_n inductively, for $n = 1, 2, \ldots$.

3.2 The asymptotics of the average minimum a_n

In order to study the asymptotic behaviour of the average minimum we introduce the *exponential generating function*

$$A(z) = \sum_{n\geq 0} a_n \frac{z^n}{n!}.$$

Therefore, we first rewrite Eq. (3.1) as

$$a_n(1-q\nu^n) = p\nu^n + p\sum_{k=1}^{n}\binom{n}{k}\lambda^k\nu^{n-k}a_k.$$

Then multiplication by $\frac{z^n}{n!}$ and summing up over all integers $n \geq 1$ yields

$$A(z) - qA(\nu z) = \sum_{n=1}^{\infty} p\,\frac{\nu^n z^n}{n!} + p\sum_{n=1}^{\infty}\sum_{k=1}^{n}\binom{n}{k}\lambda^k\nu^{n-k}a_k\frac{z^n}{n!},$$

and thus

$$A(z) - qA(\nu z) = p(e^{\nu z}-1) + pe^{\nu z}A(\lambda z). \tag{3.3}$$

We multiply the last equation by e^{-z} and obtain that the *Poisson transformed function* $\widehat{A}(z) = e^{-z}A(z)$ satisfies the equation

$$\widehat{A}(z) - p\widehat{A}(\lambda z) = qe^{-z}A(\nu z) + p(e^{-(1-\nu)z} - e^{-z}), \tag{3.4}$$

or, equivalently,

$$\widehat{A}(z) - p\widehat{A}(\lambda z) = R_1(z), \tag{3.5}$$

where $R_1(z) = qe^{-z}A(\nu z) + p(e^{-\lambda z} - e^{-z}) = qe^{-\lambda z}\widehat{A}(\nu z) + p(e^{-\lambda z} - e^{-z})$. As we are looking for the asymptotics of the average minimum a_n, we are going to study the behaviour of $\widehat{A}(z)$ as $z \to \infty$. This is based on the fact that $a_n \sim \mathcal{A}(n)$, which can be justified by using the technique of *depoissonisation* (for details about depoissonisation we refer to Jacquet and Szpankowski [4] and Szpankowski [6]). The idea is to extract the coefficients a_n from $A(z)$ using Cauchy's integral formula and the saddle point method. Let A^* denote the *Mellin transformed* function $\widehat{A}$, i.e.,

$$A^*(s) = \mathcal{M}[\widehat{A}(z); s] = \int_0^{\infty}\widehat{A}(z)\cdot z^{s-1}dz.$$

Then by applying the Mellin transform in Eq. (3.4) we obtain

$$A^*(s) - p\lambda^{-s}A^*(s) = R_1^*(s),$$

where $R_1^*(s)$ is the Mellin transformed function R_1 (for details regarding the Mellin transform we refer to Flajolet et al. [3]). We obtain

$$A^*(s) = \frac{R_1^*(s)}{1-p\lambda^{-s}}.$$

Now the function $\widehat{A}(z)$ can be obtained by applying the Mellin inversion formula, namely

$$\widehat{A}(z) = \frac{1}{2\pi \mathrm{i}} \int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} A^*(s) \cdot z^{-s} ds = \frac{1}{2\pi \mathrm{i}} \int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} \frac{R_1^*(s)}{1 - p\lambda^{-s}} \cdot z^{-s} ds, \tag{3.6}$$

where $0 < c < \frac{\log p}{\log \lambda}$. We shift the integral to the right and take the residues (with a negative sign) into account in order to estimate $\widehat{A}(z)$ in (3.6). The function under the integral has simple poles at $s_k = \frac{\log p}{\log \lambda} + \frac{2k\pi \mathrm{i}}{\log \lambda}$, $k \in \mathbb{Z}$. For these the residues with negative sign are

$$\frac{1}{\log \frac{1}{\lambda}} R_1^* \Big(\frac{\log p}{\log \lambda} + \frac{2k\pi \mathrm{i}}{\log \lambda} \Big) z^{-\frac{\log p}{\log \lambda} - \frac{2k\pi \mathrm{i}}{\log \lambda}},$$

with $R_1^*(s) = \int_0^\infty \big(qe^{-\lambda z} \widehat{A}(\nu z) + p(e^{-\lambda z} - e^{-z}) \big) z^{s-1} dz$.
For $k = 0$ the residue with negative sign is,

$$\frac{z^{-\frac{\log p}{\log \lambda}}}{\log \frac{1}{\lambda}} \int_0^\infty \big(qe^{-\lambda z} \widehat{A}(\nu z) + p(e^{-\lambda z} - e^{-z}) \big) \, z^{\frac{\log p}{\log \lambda} - 1} dz.$$

This term plays an important role in the asymptotic behaviour of the average minimum a_n, as the contributions from the other poles only constitute small fluctuations. By collecting all these residues into a periodic function, one gets the series

$$\frac{1}{\log \frac{1}{\lambda}} \sum_{k \in \mathbb{Z}} z^{-\log_\lambda p - \frac{2k\pi \mathrm{i}}{\log \lambda}} \int_0^\infty \big(qe^{-\lambda z} \widehat{A}(\nu z) + p(e^{-\lambda z} - e^{-z}) \big) \, z^{\log_\lambda p + \frac{2k\pi \mathrm{i}}{\log \lambda} - 1} dz.$$

Putting everything together, we have obtained the following result.

Theorem 1 *The average a_n of the minimum value among n random words of infinite length over $\mathbb{N}_0$ with respect to the generalised multinomial measure $\mu_{\nu,q}$ admits the asymptotic estimate*

$$a_n = \Phi(-\log_\lambda n) n^{-\log_\lambda p} \Big(1 + \mathcal{O}\big(\frac{1}{n}\big) \Big), \tag{3.7}$$

for $n \to \infty$, where $\Phi(x)$ is a periodic function having period 1 and known Fourier coefficients. The mean (zeroth Fourier coefficient) of Φ is given by the expression

$$\frac{1}{\log \frac{1}{\lambda}} \int_0^\infty \big(qe^{-\lambda z} \widehat{A}(\nu z) + p(e^{-\lambda z} - e^{-z}) \big) \, z^{\frac{\log p}{\log \lambda} - 1} dz. \tag{3.8}$$

Remark One can compute the integral in the zeroth Fourier coefficient numerically by taking for $\hat{\mathcal{A}}(z)$ the first few terms of its Taylor expansion, which can be found from the recurrence (3.2) for the numbers a_n. In order to do this we rewrite (3.8) as

$$\frac{q}{\log\frac{1}{\lambda}}\left(\Gamma\left(\frac{\log p}{\log\lambda}\right)+\sum_{k\geq 0}a_k\frac{\nu^k}{k!}\Gamma\left(k+\frac{\log p}{\log\lambda}\right)\right).$$

Remark For the special case when $\lambda = p$ (and thus $\mu_{\nu,q}$ is the uniform distribution on the unit interval) we obtain $a_n = \frac{1}{n+1}$, for $n \geq 1$. This can be shown by induction. From (3.2) one immediately gets $a_1 = \frac{1}{2}$. Assuming that $a_k = \frac{1}{k+1}$, for $k = 1, 2, \ldots, n-1$, the induction step is then, by the recursion in (3.2) equivalent to showing that

$$1 - p^{n+1} - q^{n+1} = (n+1)pq^n + (n+1)p\sum_{n=0}^{n-1}\binom{n}{k}p^k q^{n-k}a_k,$$

i.e.,

$$1 - p^{n+1} - q^{n+1} = (n+1)pq^n + (n+1)p\sum_{n=1}^{n-1}\binom{n}{k}p^k q^{n-k}\frac{1}{k+1},$$

which is immediately checked using the binomial formula for $(p+q)^{n+1} = 1$

and $\frac{n+1}{k+1}\binom{n}{k} = \binom{n+1}{k+1}$. Moreover, in this particular case the constant in (3.8) is

$$\frac{q}{\log\frac{1}{p}}\left(1+\sum_{n\geq 0}\frac{1}{n+1}q^n\right) = \frac{q}{\log\frac{1}{p}}\left(1-\frac{\log p}{q}-1\right) = 1.$$

4 Order statistics of the generalised multinomial measure: the maximum

4.1 The problem setting

As in the previous case, we pick at random (with respect to the probability measure on $\mathcal{W}$ defined above), independently, n words from $\mathcal{W}_m$, for $n \geq 1$. We apply the function value defined above to each of the chosen words and look for the maximum among these n values. The same can be done with all random choices of n words of $\mathcal{W}_\infty$. Let us denote in this section by $b_n^{(m)}$ the average minimal value among all possible choices of n words of length m. By taking the limit $b_n := \lim_{m\to\infty} b_n^{(m)}$ we obtain the average maximal value among all choices of n words of $\mathcal{W}_\infty$.

First, we establish the recursion

$$b_n^{(m)} = \sum_{k=1}^{n}\binom{n}{k}\sum_{j=0}^{\infty}(\lambda\nu^j)^k(1-\nu^j)^{n-k}\left(1-q^j+pq^j\cdot b_k^{(m-1)}\right), \text{ for } n \geq 1.$$

This is obtained from the relations in (2.2) based on the following idea. Let j be the maximum among the first letters of the n words, i.e., there is an integer k, $1 \le k \le n$, such that k words start with j, and the other $n-k$ words start with a letter less than j.

For $m \to \infty$ in the above recursion we obtain

$$b_n = \sum_{k=1}^{n} \binom{n}{k} \sum_{j \ge 0} (\lambda \nu^j)^k (1-\nu^j)^{n-k} (1 - q^j + p q^j b_k), \text{ for } n \ge 1. \tag{4.1}$$

4.2 The asymptotics of the average maximum b_n

As in the case of the average minimum, we are interested in the study of the asymptotic behaviour of b_n, for $n \to \infty$.

Since b_n is expected to be close to 1, we set $c_n = 1 - b_n$ for $n \ge 1$ and look for a recursion for c_n. Then, we study the asymptotic behavior of c_n. The recursion (4.1) can be rewritten as

$$1 - c_n = \sum_{k=1}^{n} \binom{n}{k} \sum_{j \ge 0} (\lambda \nu^j)^k (1-\nu^j)^{n-k} \big(1 - q^j + p q^j (1 - c_k)\big), \text{ for } n \ge 1. \tag{4.2}$$

We have

$$\begin{aligned}
&\sum_{k=1}^{n} \binom{n}{k} \sum_{j \ge 0} (\lambda \nu^j)^k (1-\nu^j)^{n-k} (1 - q^j + p q^j) \\
&\quad = \sum_{j \ge 0} \sum_{k=1}^{n} \binom{n}{k} (\lambda \nu^j)^k (1-\nu^j)^{n-k} (1 - q^{j+1}) \\
&\quad = \sum_{j \ge 0} \Big((1 + \lambda \nu^j - \nu^j)^n - (1-\nu^j)^n \Big) (1 - q^{j+1}) \\
&\quad = \sum_{j \ge 0} \Big((1 - \nu^{j+1})^n - (1-\nu^j)^n \Big) (1 - q^{j+1}) \\
&\quad = \sum_{j \ge 0} \Big((1 - \nu^{j+1})^n - (1-\nu^j)^n \Big) - \sum_{j \ge 0} \Big((1 - \nu^{j+1})^n - (1-\nu^j)^n \Big) q^{j+1} \\
&\quad = 1 - \sum_{j \ge 0} \Big((1 - \nu^{j+1})^n - (1-\nu^j)^n \Big) q^{j+1},
\end{aligned}$$

and thus we have proven the following result.

Proposition 2 *If b_n is the average maximul value among n words of infinite length over $\mathbb{N}_0$ with respect to the generalised multinomial measure $\mu_{\nu,q}$ and $c_n = 1 - b_n$,* *for $n \geq 1$, then c_n satisfies the recursion*

$$c_n = \sum_{j\geq 0}\Big((1-\nu^{j+1})^n - (1-\nu^j)^n\Big)q^{j+1} + \sum_{k=1}^{n}\binom{n}{k}\sum_{j\geq 0}(\lambda\nu^j)^k(1-\nu^j)^{n-k}pq^j c_k, \text{ for } n \geq 1. \tag{4.3}$$

In order to find an explicit formula that allows us to compute the values c_n we set $c_0 := 0$ and rewrite the above equation

$$c_n\Big(1 - \sum_{j\geq 0}(\lambda\nu^j)^n pq^j\Big) = \sum_{j\geq 0}\Big((1-\nu^{j+1})^n - (1-\nu^j)^n\Big)q^{j+1} + \sum_{k=0}^{n-1}\binom{n}{k}\sum_{j\geq 0}(\lambda\nu^j)^k(1-\nu^j)^{n-k}pq^j c_k \tag{4.4}$$

and thus obtain

$$c_n = \frac{1-q\nu^n}{1-p\lambda^n - q\nu^n}\Big(\sum_{j\geq 0}\Big((1-\nu^{j+1})^n - (1-\nu^j)^n\Big)q^{j+1} + \sum_{k=0}^{n-1}\binom{n}{k}\sum_{j\geq 0}(\lambda\nu^j)^k(1-\nu^j)^{n-k}pq^j c_k\Big), \tag{4.5}$$

which enables us to compute the value of c_n, for $n = 1, 2, \ldots$, inductively. For the exponential generating function $C(z) := \sum_{n\geq 0} c_n \frac{z^n}{n!}$ we obtain from (4.3)

$$\begin{aligned}
C(z) &= \sum_{n\geq 1}\frac{z^n}{n!}\sum_{j\geq 0}\Big((1-\nu^{j+1})^n - (1-\nu^j)^n\Big)q^{j+1} \\
&\quad + \sum_{n\geq 1}\frac{z^n}{n!}\sum_{k=0}^{n}\binom{n}{k}\sum_{j\geq 0}(\lambda\nu^j)^k(1-\nu^j)^{n-k}pq^j c_k \\
&= \sum_{j\geq 0}\sum_{n\geq 1}\Big(\frac{z^n}{n!}(1-\nu^{j+1})^n - \frac{z^n}{n!}(1-\nu^j)^n\Big)q^{j+1} \\
&\quad + \sum_{j\geq 0}\sum_{n\geq 1}\frac{z^n}{n!}\sum_{k=0}^{n}\binom{n}{k}(\lambda\nu^j)^k(1-\nu^j)^{n-k}pq^j c_k \\
&= \sum_{j\geq 0}\Big(e^{(1-\nu^{j+1})z} - e^{(1-\nu^j)z}\Big)q^{j+1} \\
&\quad + \sum_{j\geq 0}\sum_{n\geq 1}\sum_{k=0}^{n} c_k\frac{(\lambda\nu^j)^k z^k}{k!}\cdot\frac{(1-\nu^j)^{n-k}z^{n-k}}{(n-k)!}pq^j,
\end{aligned}$$

i.e.,

$$C(z) = \sum_{j\geq 0}\left(e^{(1-\nu^{j+1})z} - e^{(1-\nu^{j})z}\right)q^{j+1} + \sum_{j\geq 0} pq^{j}e^{(1-\nu^{j})z}C(\lambda\nu^{j}z). \tag{4.6}$$

Then for the Poisson transformed $\widehat{C}(z) = e^{-z}C(z)$ we have

$$\widehat{C}(z) = \sum_{j\geq 0}(e^{-\nu^{j+1}z} - e^{-\nu^{j}z})q^{j+1} + \sum_{j\geq 0} pq^{j}e^{-\nu^{j}z}C(\lambda\nu^{j}z). \tag{4.7}$$

Now we proceed as in Sect. 3 to apply the *depoissonisation*. Let $C^{*}(s) = \mathcal{M}[\widehat{C}(z); s]$ be the *Mellin transformed* function of $\widehat{C}$, with the notations in [3]. We apply the Mellin transform to Eq. (4.7) to obtain

$$C^{*}(s) = \Gamma(s)\sum_{j\geq 0} q^{j+1}\left(\nu^{-s(j+1)} - \nu^{-sj}\right) + \sum_{j\geq 0} pq^{j}\mathcal{M}[e^{-\nu^{j}z}C(\lambda\nu^{j}z); s]. \tag{4.8}$$

Since $\int_0^\infty z^{s-1}z^{n}e^{-\nu^{j}z}dz = \nu^{-j(n+s)}\Gamma(n+s)$, we obtain

$$\begin{aligned}\sum_{j\geq 0} pq^{j}\mathcal{M}[e^{-\nu^{j}z}C(\lambda\nu^{j}z); s] &= \sum_{j\geq 0}\sum_{n\geq 0} c_n\frac{\lambda^{n}\nu^{jn}}{n!}pq^{j}\nu^{-j(n+s)}\Gamma(n+s)\\ &= \frac{p}{1-q\nu^{-s}}\sum_{n\geq 0} c_n\frac{\lambda^{n}}{n!}\Gamma(n+s),\end{aligned}$$

and thus,

$$C^{*}(s) = q\Gamma(s)\left(\frac{\nu^{-s}}{1-q\nu^{-s}} - \frac{1}{1-q\nu^{-s}}\right) + \frac{p}{1-q\nu^{-s}}\sum_{n\geq 0} c_n\frac{\lambda^{n}}{n!}\Gamma(n+s) = \frac{R_2(s)}{1-q\nu^{-s}}, \tag{4.9}$$

where $R_2(s) = q\Gamma(s)(\nu^{-s} - 1) + p\sum_{n\geq 0} c_n\frac{\lambda^{n}}{n!}\Gamma(n+s)$.

Now the function $\widehat{C}(z)$ can be obtained by applying the Mellin inversion formula, namely

$$\widehat{C}(z) = \frac{1}{2\pi \mathrm{i}}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty} C^{*}(s)\cdot z^{-s}ds = \frac{1}{2\pi \mathrm{i}}\int_{c-\mathrm{i}\infty}^{c+\mathrm{i}\infty}\frac{R_2(s)}{1-q\nu^{-s}}\cdot z^{-s}ds, \tag{4.10}$$

where $0 < c < \frac{\log q}{\log \nu}$. We shift the integral to the right and take the residues with negative sign into account in order to estimate $\widehat{C}(z)$ in (4.10). The function under the integral has simple poles at $s_k = \frac{\log q}{\log \nu} + \frac{2k\pi \mathrm{i}}{\log \nu}$, $k \in \mathbb{Z}$. For these the residues with negative sign are

$$\frac{1}{\log\frac{1}{\nu}}R_2\left(\frac{\log q}{\log \nu} + \frac{2k\pi \mathrm{i}}{\log \nu}\right)z^{-\frac{\log q}{\log \nu} - \frac{2k\pi \mathrm{i}}{\log \nu}}.$$

For $k=0$ the residue with negative sign is,

$$\frac{z^{-\frac{\log q}{\log \nu}}}{\log \frac{1}{\nu}}\Big(q\Gamma\Big(\frac{\log q}{\log \nu}\Big)(\nu^{-\frac{\log q}{\log \nu}}-1)+p\sum_{n\geq 0}c_n\frac{\lambda^n}{n!}\Gamma\Big(n+\frac{\log q}{\log \nu}\Big)\Big)$$
$$=\frac{pz^{-\frac{\log q}{\log \nu}}}{\log \frac{1}{\nu}}\Big(\Gamma\Big(\frac{\log q}{\log \nu}\Big)+\sum_{n\geq 0}c_n\frac{\lambda^n}{n!}\Gamma\Big(n+\frac{\log q}{\log \nu}\Big)\Big).$$

This term plays an important role in the asymptotic behaviour of c_n, as the contributions from the other poles only constitute small fluctuations. We collect all these residues into a periodic function and obtain

$$\frac{1}{\log \frac{1}{\nu}}\sum_{k\in\mathbb{Z}}R_2\Big(\frac{\log q}{\log \nu}+\frac{2k\pi \mathrm{i}}{\log \nu}\Big)z^{-\frac{\log q}{\log \nu}-\frac{2k\pi \mathrm{i}}{\log \nu}}.$$

Putting everything together, we thus get the following result.

Theorem 2 *The average b_n of the maximum value among n random words of infinite length over $\mathbb{N}_0$ with respect to the generalised multinomial measure $\mu_{\upsilon,q}$ admits the asymptotic estimate*

$$b_n=1-\Phi(-\log_\nu n)n^{-\log_\nu q}\Big(1+\mathcal{O}(\frac{1}{n})\Big), \tag{4.11}$$

for $n\to\infty$, where $\Phi(x)$ is a periodic function having period 1 and known Fourier coefficients. The mean (zeroth Fourier coefficient) of Φ is given by the expression

$$\frac{p}{\log \frac{1}{\nu}}\Big(\Gamma\Big(\frac{\log q}{\log \nu}\Big)+\sum_{n\geq 0}c_n\frac{\lambda^n}{n!}\Gamma\Big(n+\frac{\log q}{\log \nu}\Big)\Big). \tag{4.12}$$

Remark For $\lambda=p$ we expect to get $c_n=\frac{1}{n+1}$, which indeed can be proven by induction. Then the constant in (4.12) is

$$\frac{p}{\log \frac{1}{q}}\Big(1+\sum_{n\geq 0}\frac{1}{n+1}p^n\Big)=\frac{p}{\log \frac{1}{q}}\Big(1-\frac{\log q}{p}-1\Big)=1,$$

as it should. The proof by induction can be done as follows. One easily obtains $c_1=\frac{1}{2}$ from the recursion (4.4). Assuming now that $c_k=\frac{1}{k+1}$ for all $k\geq 1$ we have to show that

$$\frac{1}{n+1}\Big(1-\sum_{j\geq 0}(pq^j)^n pq^j\Big)=\sum_{j\geq 0}\Big((1-q^{j+1})^n-(1-q^j)^n\Big)q^{j+1}$$
$$+\sum_{k=1}^{n-1}\binom{n}{k}\sum_{j\geq 0}(pq^j)^k(1-q^j)^{n-k}pq^j\frac{1}{k+1},$$

or, equivalently,

$$1-\sum_{j\geq 0}(pq^j)^{n+1}=(n+1)\sum_{j\geq 0}\Big((1-q^{j+1})^n-(1-q^j)^n\Big)q^{j+1}$$
$$+\sum_{k=1}^{n-1}\binom{n+1}{k+1}\sum_{j\geq 0}(pq^j)^{k+1}(1-q^j)^{n-k}.$$

With the binomial formula, this yields

$$1-\sum_{j\geq 0}(pq^j)^{n+1}=(n+1)\sum_{j\geq 0}\Big((1-q^{j+1})^n-(1-q^j)^n\Big)q^{j+1}$$
$$+\sum_{j\geq 0}\Big((pq^j+1-q^j)^{n+1}-(1-q^j)^{n+1}\Big)-\sum_{j\geq 0}(pq^j)^{n+1}$$
$$-(n+1)\sum_{j\geq 0}pq^j(1-q^j)^n,$$

which can be rewritten as

$$1=(n+1)\sum_{j\geq 0}\Big((1-q^{j+1})^n-(1-q^j)^n\Big)q^{j+1}$$
$$+\lim_{J\to\infty}\sum_{j=0}^{J}\Big((1-q^{j+1})^{n+1}-(1-q^j)^{n+1}\Big)-(n+1)\sum_{j\geq 0}pq^j(1-q^j)^n,$$

and, since the first term in the last sum is zero, it becomes

$$1=(n+1)\sum_{j\geq 0}\Big((1-q^{j+1})^n-(1-q^j)^n\Big)q^{j+1}$$
$$+\lim_{J\to\infty}\sum_{j=0}^{J}\Big((1-q^{j+1})^{n+1}-(1-q^j)^{n+1}\Big)-(n+1)\sum_{j\geq 0}pq^{j+1}(1-q^{j+1})^n.$$

We thus obtain

$$1=(n+1)\sum_{j\geq 0}\Big((1-q^{j+1})^n-(1-q^j)^n-p(1-q^{j+1})^n\Big)q^{j+1}$$
$$+\lim_{J\to\infty}(1-q^{J+1})^{n+1},$$

which is equivalent to

$$0=\sum_{j\geq 0}\Big(q(1-q^{j+1})^n-(1-q^j)^n\Big)q^{j+1}$$

and

$$0 = \sum_{j\geq 0}(1-q^{j+1})^n q^{j+2} - \sum_{j\geq 1}(1-q^j)^n q^{j+1}$$

which obviously holds.

References

1. Bassino, F., Prodinger, H.: (q, δ)-numeration systems with missing digits. Monatsh. für Math. **141**, 89–99 (2004)
2. Cristea, L.L., Prodinger, H.: Order statistics for the Cantor–Fibonacci distribution. Aequ. Math. **73**, 78–91 (2007)
3. Flajolet, P., Gourdon, X., Dumas, P.: Mellin transforms and asymptotics: Harmonic sums. Theoret. Comput. Sci. **144**, 3–58 (1995)
4. Jacquet, P., Szpankowski, W.: Analytical de-Poissonization and its applications. Theoret. Comput. Sci. **201**(1–2), 1–62 (1998)
5. Okada, T., Sekiguchi, T., Shiota, Y.: A Generalization of Hata–Yamaguti's results on the Takagi function II: multinomial case. Jpn. J. Indust. Appl. Math. **13**, 435–463 (1996)
6. Szpankowski, W.: Average case analysis of algorithms on sequences. Wiley-interscience series in discrete mathematics and optimization. Wiley, New York (with a foreword by Philippe Flajolet) (2001)