# COMBINATORIAL PROBLEMS RELATED TO SEQUENCES WITH REPEATED ENTRIES

**Margaret Lyn Archibald**

School of Mathematics
University of the Witwatersrand
Johannesburg
South Africa

Under the supervision of
Professor A. Knopfmacher and Professor H. Prodinger.

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy.

Johannesburg, 2005.

# Declaration

I declare that this thesis is my own, unaided work. It is being submitted for the degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other university.

________________________

Margaret Lyn Archibald

______ day of ____________________ 2005

# Abstract


Sequences of numbers have important applications in the field of Computer Science. As a result they have become increasingly regarded in Mathematics, since analysis can be instrumental in investigating algorithms.

Three concepts are discussed in this thesis, all of which are concerned with 'words' or 'sequences' of natural numbers where repeated letters are allowed:

- **The number of distinct values in a sequence with geometric distribution**

  In Part I, a sample which is geometrically distributed is considered, with the objective of counting how many different letters occur at least once in the sample. It is concluded that the number of distinct letters grows like $\log n$ as $n \to \infty$. This is then generalised to the question of how many letters occur at least $b$ times in a word.

- **The position of the maximum (and/or minimum) in a sequence with geometric distribution**

  Part II involves many variations on the central theme which addresses the question: "What is the probability that the maximum in a geometrically distributed sample occurs in the first $d$ letters of a word of length $n$?" (assuming $d \leq n$). Initially, $d$ is considered fixed, but in later chapters $d$ is allowed to grow with $n$. It is found that for $1 \leq d = o(n)$, the results are the same as when $d$ is fixed.

- **The average depth of a key in a binary search tree formed from a sequence with repeated entries**

  Lastly, in Part III, random sequences are examined where repeated letters are allowed. First, the average left-going depth of the first one is found, and later the right-going path to the first $r$ if the alphabet is $\{1, \ldots, r\}$ is examined. The final chapter uses a merge (or 'shuffle') operator to obtain the average depth of an arbitrary node, which can be expressed in terms of the left-going and right-going depths.

# Acknowledgements

I hereby acknowledge the financial support of the NRF for both my staff development grant and the NRF/CNRS Doctoral Program which provided me with an opportunity to conduct my research partially in France.

I would like to thank Julien Clement and Frédérique Bassino for their help and advice during my stay in France, and the Université de Marne-la-Vallée which was my host for part of 2004.

Lastly, thank you to my supervisors Arnold Knopfmacher and Helmut Prodinger for their continued support, help, insight, dedication and availability throughout this project.

# Contents

# List of Figures

# List of Tables

# Symbols and Formulæ

$$p + q = 1, \qquad p, q \geq 0$$

$$Q = q^{-1}$$

$$L = \log Q$$

$$\log x = \ln x$$

$$n^* = n(Q - 1)$$

$$\chi_k = \frac{2k\pi\boldsymbol{i}}{L}, \qquad \text{for all} \quad k \in \mathbb{Z}\backslash\{0\}$$

Note: We denote $\sqrt{-1}$ as $\boldsymbol{i}$ to distinguish it from the $i$ we use as an index.

Multinomial:

$$\binom{n}{n_1, \ldots, n_r} = \frac{n!}{n_1! \cdots n_r!}, \qquad \text{where } n_1 + \cdots + n_r = n$$

Euler's constant:

$$\gamma = 0.57721$$

Gamma function:

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt, \qquad \Re(z) > 0$$

$$\Gamma(n) = (n-1)!, \qquad n \in \mathbb{Z}$$

Psi/Digamma function:

$$\psi(z) = \frac{d}{dz} \log \Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}$$

Harmonic number:

$$H_n = \sum_{k=1}^{n} \frac{1}{k} = \psi(n+1) + \gamma$$

Harmonic number approximation:

$$H_n = \log n + \gamma + o(1), \qquad \text{as } n \to \infty$$

Second order harmonic number:

$$H_n^{(2)} = \sum_{k=1}^{n} \frac{1}{k^2}$$

Harmonic number identities:

$$\sum_{k=1}^{n} H_k = (n+1)H_n - n,$$

$$\sum_{k=1}^{n} H_k^{(2)} = (n+1)H_n^{(2)} - H_n$$

$$\sum_{k=1}^{n} \frac{1}{n-k+1} H_k = H_{n+1}^2 - H_{n+1}^{(2)}$$

$$\sum_{k=1}^{n} \frac{1}{k} H_k = \sum_{1\le j\le k\le n} \frac{1}{jk} = \frac{1}{2}\bigl(H_n^2 + H_n^{(2)}\bigr)$$

$$\sum_{1\le j<k\le n} \frac{1}{jk} = \frac{1}{2}\bigl(H_n^2 - H_n^{(2)}\bigr)$$

$$\begin{aligned}\sum_{k=1}^{j} \frac{H_{n+k-j}}{k} + \sum_{k=1}^{n+1-j} \frac{H_{j+k-1}}{k} &= \frac{1}{2}\bigl(H_j^2 + H_j^{(2)}\bigr) + \frac{1}{2}\bigl(H_{n+1-j}^2 + H_{n+1-j}^{(2)}\bigr) \\ &\quad + H_j H_{n+1-j} + \frac{1}{j(n+1-j)} \\ &\quad + \frac{n+1}{j(n+1-j)}\bigl(H_n - H_j - H_{n+1-j}\bigr)\end{aligned}$$

Mellin transform:

$$f^*(s) = \int_0^\infty f(x) x^{s-1} dx, \qquad s \in \langle -u, -v \rangle$$

Inverse Mellin transform:

$$f(x) = \frac{1}{2\boldsymbol{i}\pi} \int_{c-\boldsymbol{i}\infty}^{c+\boldsymbol{i}\infty} f^*(s) x^{-s} ds, \qquad -u < c < -v$$

Harmonic sum rule for Mellin transforms

$$\sum_i \lambda_i f(\mu_i x) \qquad \text{transforms to} \qquad \Bigl(\sum_i \lambda_i \mu_i^{-s}\Bigr) \cdot f^*(s) \qquad (\mu_i > 0)$$

Power rule for Mellin transforms

$$x^\rho f(x^\theta) \qquad \text{transforms to} \qquad \frac{1}{\theta} f^*\Bigl(\frac{s+\rho}{\theta}\Bigr)$$

Ordinary generating function (OGF)

$$A(z) = \sum_{n\ge 0} a_n z^n$$

Exponential generating function (EGF)

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

# Chapter 1

# Thesis synopsis

Words (sequences of natural numbers) are discussed, and various parameters are analysed with the help of symbolic equations, generating functions, probabilities, Rice's method, Mellin transforms and the Combinatorial Laplace transform. Expressing ideas symbolically is always a useful way of getting a more concrete, intuitive understanding of abstract concepts, and translating these ideas into generating functions allows many calculations to take place. "Generating functions are more than a technical tool used to solve recurrences and compute moments – they are a necessary and natural link between the algorithms that are our objects of study and analytic methods that are necessary to discover their properties. Generating functions serve both as a combinatorial tool to facilitate counting and as an analytic tool to develop precise estimates for quantities of interest." ([36, page 82])

The parameters of interest in this thesis for which we make use of this powerful tool are discussed in detail below. After the generating function has been determined, often further manipulations are required to obtain the (sometimes asymptotic) results. Unless otherwise stated, all asymptotic results are taken as $n \to \infty$. Rice's method (introduced in Chapter 2) is used frequently in the theorems from which this thesis is composed. This method allows us to approximate an alternating sum expression of the form

$$\sum_{k=1}^{n} \binom{n}{k} (-1)^k f(k), \qquad (n \in \mathbb{N})$$

whose magnitude would otherwise be difficult to estimate. However this method is very specific and it is sometimes necessary to make use of other techniques. In Part II we see such a case, and the Mellin transform is used. This transform (along with its inverse) allows the user to approximate more diverse expressions than Rice's method. Finally, the 'shuffle' operator is introduced in Part III, which necessitates the use of the Laplace transform which converts between ordinary and exponential generating functions. The original generating functions in this case are ordinary, but in order to produce a 'shuffle', a product of exponential generating functions is required. The use of exponential generating functions ensures that the ordering of the letters does not pose a problem within the generating function, a property used elsewhere in the thesis too – see, for example, Chapter 3 in Part I.

## 1.1 Distinct values

How many distinct values could one expect in a geometrically distributed sample (defined below) of length $n$? This question is addressed in Chapter 3, and generalised in Chapter 4 to the number of values in such a sample that appear at least twice – or even, say, five times or more.

Words (sequences) of length $n$ are considered. The entries in the sequence or *letters* are natural numbers which occur independently of each other (i.e., the letter appearing in the first position has no effect on the letter in the second, third or eleventh positions). Each letter occurs with geometric probability (i.e., letter $j$ will occur with probability $pq^j$ for $p+q=1$ – see Chapter 2). Thus any natural number could occur, but smaller numbers (letters) are more common. The expected value and variance are found for the number of letters occurring at least once in a word. That is, we predict that we will have $\mathbb{E}(d_n)$ different letters in a random (geometrically distributed) word of length $n$, and then find how closely the numbers of distinct values are clustered around this mean (this quantity is the variance, expressed by $\mathbb{V}(d_n)$). Then both the expected value and variance are found for the number of letters occurring at least $b$ times in a word.

Consider the specific example 182122211211212161142643171131, which is a word of length $n=30$ created by Mathematica from geometric random variables where $p=q=\frac{1}{2}$. There are seven distinct values, namely $\{1,2,3,4,6,7,8\}$. Five of the letters occur at least twice ($\{1,2,3,4,6\}$), and two of these letters ({1,2}) occur eight times or more. If we let $Q=\frac{1}{q}$, $L=\log Q$ and $\gamma \approx 0.57721$, then according to the main term in the results, we would expect approximately

$$\begin{aligned}\log_Q n + \frac{\gamma}{L} + \log_Q(Q-1) - \frac{1}{2} &= \log_2 30 + \frac{\gamma}{\log 2} + \log_2 1 - \frac{1}{2}\\ &= 5.23964\end{aligned}$$

distinct values and approximately

$$\begin{aligned}\log_Q n + \frac{\gamma}{L} + \log_Q(Q-1) - \frac{1}{2} - \frac{1}{L}H_{b-1} &= \log_2 30 + \frac{\gamma}{\log 2} + \log_2 1 - \frac{1}{2} - \frac{1}{\log 2}H_1 \\ &= 3.79694\end{aligned}$$

letters occurring at least twice. For the number of letters appearing in the word eight times or more, we would expect about

$$\begin{aligned}\log_Q n + \frac{\gamma}{L} + \log_Q(Q-1) - \frac{1}{2} - \frac{1}{L}H_{b-1} &= \log_2 30 + \frac{\gamma}{\log 2} + \log_2 1 - \frac{1}{2} - \frac{1}{\log 2}H_7 \\ &= 1.49893.\end{aligned}$$

Hence these random experimental values, though enabling one to understand the problem better, give results greater than those we would expect (the average). This can be explained by the fact that these results are asymptotic as $n \to \infty$, so for $n = 30$ we are unlikely to get accurate results.

The distinct value problem can also be viewed as a problem of balls in urns or boxes. After throwing $n$ balls into an infinite row of boxes which are initially empty, with the relevant probability that a ball falls into a certain box (for example, if $p = q = \frac{1}{2}$, it is twice as likely that the ball will land in the box closest to you as in the next box along), the analogous question would then read: "How many boxes are non-empty after $n$ balls have been thrown?". In the more general case, we would ask: "How many boxes contain at least $b$ balls after $n$ have been thrown?". It is necessary to assume that the thrower has perfect aim and that the probability of the ball landing outside every box is zero.

Chapters 2 and 5 introduce and conclude this portion of the thesis.

## 1.2 Maxima and minima

The position of the maximum and/or minimum of a geometrically distributed sample of length $n$ is now considered. After introducing the problem in Chapter 6, Chapter 7 is used to address the question: "What is the probability that the maximum value in a geometrically distributed sample lies in the first position?". Then, by extending the possible positions of the maximum to the first $d$ places in the word, this is generalised in Chapter 8. At first, $d$ is considered fixed, but later chapters allow $d$ to grow with $n$.

It is necessary to consider different cases because repeated maxima can either be allowed or not. In the first scenario, two cases must be considered. We refer to these as 'weak' and 'strict', which correspond respectively to allowing the maximum to appear again in the word (any number of times) and not allowing another recurrence of the maximum. For example, the words 422131221113 and 311213212131 both satisfy the 'weak' condition, but only the former satisfies the 'strict' condition. For the more general scenario of having the maximum in the first $d$ letters of the word, four cases need to be considered. This is because the weak/strict classification applies to both the first $d$ letters and also (independently) to the remaining $n-d$ letters in the word. Table 1.1 gives examples of each case. The different cases are denoted by an ordered tuple where the first entry corresponds to the first $d$ letters of the word and the second refers to the rest of the word. In Table 1.1, we take $n$ to be 20 and let $d$ be 6 and assume that each letter $(j)$ occurs in the word with probability $\left(\frac{1}{2}\right)^j$. (For geometric distribution, we assume each letter $j$ occurs with probability $pq^{j-1}$ where $p+q=1$. Here, we choose the values $p=q=\frac{1}{2}$.) The maximum in each case is the letter 3. Note that these cases are not disjoint. For example, any of the first three cases would also fall under the (weak, weak) classification, but not necessarily vice versa.

| Classification | Example |
|---|---|
| (strict, strict) | 11231211122111212211 |
| (weak, strict) | 13231311122121211112 |
| (strict, weak) | 12311221312122113311 |
| (weak, weak) | 32111322121122111231 |

Table 1.1: Examples of words in each of the four cases where $n=20$ and $d=6$.

The same cases are considered for the minimum value (see Chapters 9 and 10) with different results due to the geometric probabilities attached to each letter (or natural number). For these cases, the restrictions have more influence because of the geometric distribution: the probabilities decrease as the value of the letter increases. Hence smaller letters occur more frequently than larger letters. Because of this, a large letter is much more likely to occur only once than a small letter and consequently we expect there to be a greater difference between classifications here than when the maximum was considered. The results for these cases are exact and, unlike the restricted maximum cases, do not require asymptotic analysis.

In Chapter 11, the probability that the minimum of the first $d$ letters is greater than (and possibly equal to) all letters in the rest of the word is found. Again, four cases are considered and as in Chapters 7 and 8, asymptotics are involved in obtaining the results. In all of the above categories, it can be seen that the second component in each tuple plays more of a role than the first, for $n$ large. This is due to the fact that $d$ is fixed relative to $n$, so as $n$ gets larger, the second part of the word (all letters from position $d+1$ onwards) dominates the first $d$ letters. But what if $d$ grows with $n$?

The remaining chapters in Part II do not assume $d$ is fixed. In Chapter 12, $d$ is allowed to grow linearly with $n$ – i.e., $d = \alpha n$ for $0 < \alpha \leq 1$. Finally in Chapter 13, $d$ grows with $n$ according to the relationship $d = \alpha n^{\gamma}$ for $0 < \alpha \leq 1$ and $0 < \gamma < 1$. It is found that the results for the $d$ fixed cases hold in this chapter too, and further manipulations show that these results hold for $1 \leq d = o(n)$.

Part II is concluded in Chapter 14 with a brief analysis of the results.

## 1.3 Binary search trees

For this section of the thesis, the sequences are no longer geometrically distributed. Sequences of length $n$ are created from letters $\{1, \ldots, r\}$ (which can appear more than once in the word) according to two models.

The first is the 'multiset' model. For this we assume that we know how many appearances each letter makes in the word, i.e., we assume that we know $n_i$ for $i \in \{1, \ldots, r\}$ in the multiset $\{n_1 \cdot 1 \ ; \ n_2 \cdot 2 \ ; \ \ldots \ ; \ n_r \cdot r\}$.

The second model is the 'probability' model. A probability is attached to each letter in the alphabet $\{1, \ldots, r\}$, and we assume each letter occurs independently of all the rest. A more symbolic generating function is used, and in general the calculations are easier because of the restriction that $P_{[1,r]} = p_1 + \cdots + p_r = 1$ (i.e., the total sum of all probabilities in the finite alphabet will be one).

The results sought are as follows:

- The average left-going depth of the first 1, and its variance;

- The average right-going depth of the first $r$, and its variance;
- The average depth of an arbitrary key $\alpha$, and its variance.

For the 'left-going' and 'right-going' cases, the binary search tree (which contains equal/repeated keys) corresponding to each sequence is built as follows: the first letter of the word is the root. Subsequent nodes are inserted as children, to the left if they are strictly less than the parent node and to the right if they are larger than or equal to the parent node. In this way we create a left-going branch only when we have a strict left-to-right minimum, and we create a right-going branch when we have a weak left-to-right maximum. See Figure 1.1.

For the average depth of the key $\alpha$, we assume that from a sequence with repeated letters, we create a binary search tree with distinct nodes. Thus keys smaller than the parent node will be inserted to its left and those larger will be inserted to the right. Those equal to the parent node will be passed over. As an example, consider Figure 1.1 which shows the two different methods of creating a binary search tree from the sequence 23131.


Figure 1.1: Two binary search trees corresponding to the sequence 23131.

Generating functions are used to express the situation in each case, and then the moments are calculated by partial differentiation. The variables in the multiset model case are $z$ (which counts all letters); $u$ (which counts all relevant left/right-going branches), and $x_i$, $i \in \{1, \ldots, r\}$ (where $x_i$ counts how many times the letter $i$ appears in the word). In the probability model, variables $z$ and $u$ have the same role, but we replace each $x_i$ with a $p_i$ which is the probability the letter $i$ occurs in the word. Thus only the coefficients of $z$ and $u$ are needed, as each $p_i$ has a set value which can be substituted directly into the expression.

Binomials and multinomials emerge in the course of the calculations, and identities are used to simplify these. Other identities used are harmonic number identities (see [12, 19, 23, 37]).

For both models, the expectation and variance are found in each of the three cases. The multiset model produces an exact form in terms of $n_i$, $i \in \{1, \ldots, r\}$, where $n_i$

is the number of times the letter $i$ occurs in the sequence, and $n_1+\cdots+n_r = n$, the length of the sequence. The probability model gives an asymptotic approximation in terms of $p_i$, $i \in \{1, \ldots, r\}$ where $p_i$ represents the probability of letter $i$ occurring in the input sequence, and $p_1 + \cdots + p_r = 1$. In all cases, the results from the two models are asymptotically equal, which can be intuitively understood by thinking of the probability of the letter $i$ occurring in the multiset model as $\frac{n_i}{n}$.

# Part I

# Distinct Values

# Chapter 2

# Introduction

We consider words $x_1x_2\cdots x_n$ with letters $x_j \in \{1,2,\dots\}$. The letter $i$ occurs with (geometric) probability $pq^{i-1}$ where $p+q=1$, and the letters are considered to be independent, so that the word $x_1x_2\cdots x_n$ appears with probability

$$(pq^{x_1-1})(pq^{x_2-1})\cdots(pq^{x_n-1}) = (p/q)^n q^{x_1+\cdots+x_n}.$$

In this way larger letters occur less frequently than smaller letters, and if we consider the case where $p=q=\frac{1}{2}$, then about half of the letters must be 1, and letter 1 occurs twice as often as letter 2 which occurs twice as often as letter 3 and so forth.

The combinatorics of geometric random variables has gained importance because of applications in computer science. We mention just two areas: **skiplists** [5, 27, 33] and **probabilistic counting** [9, 16].

Some of the previous studies relating to combinatorics of geometric random variables are as follows. In [29] the number of left-to-right maxima was investigated in the model of *words* (strings) $a_1\cdots a_n$, where the letters $a_j \in \mathbb{N}$ are independently generated according to the geometric distribution described above. H.-K. Hwang and his collaborators obtained further results about this limiting behaviour in [3]. The two parameters 'value' and 'position' of the $r$th left-to-right maximum for geometric random variables were considered in a subsequent paper [21]. Other combinatorial questions have been considered in [25, 28, 30, 31].

The following question is addressed: "How many different letters appear in words of length $n$, generated by geometric random variables?" For this parameter $(d_n)$, we derive expectation and variance. We use the following notation: $Q := \frac{1}{q}$, $L := \log Q$, $n^* := n(Q-1)$, $\gamma \approx 0.57721$ (Euler's constant), and $\chi_k := \frac{2k\pi \boldsymbol{i}}{L}$ for $k \in \mathbb{Z}$, $k \neq 0$. We use $\delta(x)$ to represent a periodic function with mean zero.

These results (expectation and variance of this quantity) have been found previously for the special case of $p = q = \frac{1}{2}$, see [13]. The results in this thesis extend this idea to any values of $p$ and $q$ where $p + q = 1$ and $p, q \geq 0$.

We then generalise this question as follows: "How many letters appear at least $b$ times, where $b \geq 1$ is a design parameter?". Using this notation, $b = 1$ is the previous case.

In the asymptotic formulæ that we derive, there appear periodic oscillations, due to poles of certain functions at $z = \chi_k := \frac{2k\pi \boldsymbol{i}}{L}$, $k \in \mathbb{Z}$, $k \neq 0$. They are usually tiny, but play an essential role especially in the variance. In some cases, there are no fluctuations in the variance, see [32].

A technique from complex analysis which we make use of frequently hereafter is called 'Rice's method'. This method can be used to approximate alternating sums as follows. The lemma states (see [10, 29, 38])

**Lemma 1** *Let $\mathcal{C}$ be a curve surrounding the points $1, 2, \ldots, n$ in the complex plane, and let $f(z)$ be analytic inside $\mathcal{C}$. Then*

$$\sum_{k=1}^{n} \binom{n}{k} (-1)^k f(k) = -\frac{1}{2\pi \boldsymbol{i}} \int_{\mathcal{C}} [n; z] f(z) dz, \tag{2.1}$$

*where*

$$[n; z] = \frac{(-1)^{n-1} n!}{z(z-1)\cdots(z-n)} = \frac{\Gamma(n+1)\Gamma(-z)}{\Gamma(n+1-z)}.$$

(The expression $[n; z]$ is sometimes called the kernel.) By extending the contour of integration, it turns out that under suitable growth conditions (see [10]) the asymptotic expansion of our alternating sum is given by

$$\sum \operatorname{Res}([n; z] f(z)) + \text{smaller order terms},$$

where the sum is taken over all poles different from $1, \ldots, n$. Poles that lie more to the left lead to smaller terms in the asymptotic expansion.

# Chapter 3

# Classical case: The distinct value problem

Consider the following bivariate generating function:

$$F(z,u) := \prod_{i\geq 1}\big(1+u(e^{zpq^{i-1}}-1)\big) = \prod_{i\geq 0}\big(1+u(e^{zpq^{i}}-1)\big). \tag{3.1}$$

Suppose the total number of letters in a word is represented by $n$, and $k$ represents the number of distinct values appearing in that word, then the coefficient of $\frac{z^n}{n!}u^k$ is the probability that a word of length $n$ has $k$ distinct values. The function in (3.1) is an exponential generating function in terms of $z$, hence the factor $n!$. It is a probability generating function in terms of the variable $u$, and thus differentiating partially with respect to this variable will lead to the expected number of distinct letters in any word of length $n$. If letter $i$ occurs at least once, then this will be accommodated by the presence of the $u$ in front of the expression $(e^{zpq^{i-1}}-1)$ which represents all non-empty 'sets' of letter $i$ which occur in the word. The initial 1 inside the product denotes the empty set – used if the letter $i$ does not appear in the word. The problem of letters appearing at different places in the word is overcome by the use of the exponential generating function.

Note that substituting $u=1$ into this function gives $e^z$ (since all probabilities sum to 1, see (3.4) below), which is to be expected because it reduces to a generating function whose coefficients represent the probability that a word of length $n$ has no restrictions.

The theorems which follow are proved in this chapter.

**Theorem 3.1** *The number of distinct letters in a word of length $n$ whose letters* *occur independently and with geometric probability, is on average*

$$\mathbb{E}(d_n) = \log_Q n + \frac{\gamma}{L} + \log_Q(Q-1) - \frac{1}{2} + \delta_E(\log_Q n^*) + O\left(\frac{1}{n}\right), \tag{3.2}$$

*as* $n \to \infty$, *where* $n^* = n(Q-1)$ *and*

$$\delta_E(x) = -\frac{1}{L} \sum_{k \neq 0} \Gamma(-\chi_k) e^{2k\pi \boldsymbol{i} x},$$

*is defined in equation (3.8).*

**Theorem 3.2** *The variance of the number of distinct letters in a word of length* $n$ *of geometric random variables is*

$$\mathbb{V}(d_n) = \log_Q 2 + \delta_V(\log_Q n^*) + o(1), \tag{3.3}$$

*as* $n \to \infty$, *where*

$$\delta_V(x) = \delta_E(x + \log_Q 2) - \delta_E(x),$$

*with* $\delta_E(x)$ *from (3.8).*

## 3.1 The expected value (classical case)

Let $d_n$ be the number of distinct values in a word of length $n$, and let $\mathbb{E}(d_n)$ represent the expected value of this quantity. Then (see [12])

$$\mathbb{E}(d_n) = n![z^n] \frac{\partial}{\partial u} F(z,u) \Big|_{u=1},$$

where $F(z,u)$ is defined as in (3.1). We have

$$\begin{aligned} \prod_{i \geq 0} \left(1 + u(e^{zpq^i} - 1)\right)\Big|_{u=1} &= \prod_{i \geq 0} e^{zpq^i} \\ &= e^{zp} e^{zpq} e^{zpq^2} \cdots \\ &= e^{z(p+pq+pq^2+\cdots)} \\ &= e^z. \end{aligned} \tag{3.4}$$

Also note that the derivative of a product can be written as a sum, whose summand (in this case) includes a product we know. So for $f_i(z,u) := 1 + u(e^{zpq^i} - 1)$,

$$\frac{d}{du} \prod_{i \geq 0} f_i(z,u) = \sum_{j \geq 0} \frac{\prod_{i \geq 0} f_i(z,u)}{f_j(z,u)} \frac{d}{du} f_j(z,u),$$

since by the product rule for derivatives, to differentiate a product we must keep all terms the same, except one which we differentiate. Then all of these terms are summed. Thus the derivative of this product is the sum of all the terms where a factor has been removed from the product, and has been replaced by its derivative. Using this and (3.4), we get

$$\begin{aligned}
\mathbb{E}(d_n) &= n![z^n]\frac{\partial}{\partial u}\prod_{i\geq 0}\left(1+u(e^{zpq^i}-1)\right)\Big|_{u=1}\\
&= n![z^n]e^z\sum_{i\geq 0}\frac{e^{zpq^i}-1}{e^{zpq^i}}\\
&= n![z^n]e^z\sum_{i\geq 0}(1-e^{-zpq^i})\\
&= n![z^n]\sum_{i\geq 0}\left(e^z-e^{z(1-pq^i)}\right)\\
&= \sum_{i\geq 0}\left(1-(1-pq^i)^n\right)\\
&= \sum_{i\geq 0}\left(1-\sum_{k=0}^{n}\binom{n}{k}(-pq^i)^k\right)\\
&= \sum_{k=1}^{n}\sum_{i\geq 0}\binom{n}{k}(-1)^{k-1}p^kq^{ik}\\
&= \sum_{k=1}^{n}\binom{n}{k}(-1)^{k-1}p^k\sum_{i\geq 0}q^{ik}\\
&= \sum_{k=1}^{n}\binom{n}{k}(-1)^{k-1}\frac{p^k}{1-q^k}\\
&= \sum_{k=1}^{n}\binom{n}{k}(-1)^{k}\frac{-(1-Q^{-1})^k}{1-Q^{-k}}\\
&= \sum_{k=1}^{n}\binom{n}{k}(-1)^{k}\frac{-(Q-1)^k}{Q^k-1}.
\end{aligned}$$

This gives an expression for the expectation. However, we cannot easily see what the number of distinct values would be from this form. To get a better idea we approximate this alternating sum using Rice's method, which is described in the introduction and requires the function to be written in this form (i.e., a finite alternating sum of a binomial and a function of the index of summation). We use Lemma 1 (see (2.1)) to approximate this alternating sum (i.e., the expected value). The poles we look at (from definition (3.5) below) occur at $z = 0$ and $z = \chi_k := \frac{2k\pi \boldsymbol{i}}{L}$, $k \in \mathbb{Z}\backslash\{0\}$. All other poles lead to smaller terms.

The first pole we will deal with is at $z = 0$. For

$$f(z) := -\frac{(Q-1)^z}{Q^z - 1}, \tag{3.5}$$

and (see Lemma 1)

$$[n; z] = \frac{(-1)^{n-1} n!}{z(z-1)\cdots(z-n)},$$

we can see that $[n; z]f(z)$ has a double pole at $z = 0$. We thus expand everything to two terms. Firstly, we have ([29]):

$$[n; z] = \frac{(-1)^{n-1} n!}{z(z-1)\cdots(z-n)} \sim -\frac{1}{z}(1 + zH_n),$$

where $H_n = \sum_{i=1}^{n} \frac{1}{i}$ is the $n$th harmonic number. We expand $f(z)$ to get:

$$\begin{aligned}
f(z) &= -\frac{(Q-1)^z}{Q^z - 1} \\
&= -\frac{e^{z \log(Q-1)}}{e^{z \log Q} - 1} \\
&\sim -\frac{1 + z\log(Q-1)}{1 + z\log Q + \frac{z^2 \log^2 Q}{2} - 1} \\
&= -\frac{1 + z\log(Q-1)}{z \log Q(1 + \frac{z \log Q}{2})} \\
&\sim -\frac{1}{z \log Q}(1 + z\log(Q-1))\left(1 - \frac{z \log Q}{2}\right) \\
&= -\frac{1}{zL}(1 + z\log(Q-1))\left(1 - \frac{zL}{2}\right).
\end{aligned}$$

To calculate the residue at $z = 0$ we consider the coefficient of $z^{-1}$ in $[n; z]f(z)$,

$$\begin{aligned}
&[z^{-1}]\frac{1}{z}(1 + zH_n)\frac{1}{zL}(1 + z\log(Q-1))\left(1 - \frac{zL}{2}\right) \\
&\quad = [z]\frac{1}{L}(1 + zH_n)(1 + z\log(Q-1))\left(1 - \frac{zL}{2}\right) \\
&\quad = \frac{1}{L}\left(H_n + \log(Q-1) - \frac{L}{2}\right) \\
&\quad = \frac{1}{L}\left(\log n + \gamma + \log(Q-1) - \frac{L}{2}\right) + O\left(\frac{1}{n}\right), \qquad \text{as } n \to \infty \\
&\quad = \log_Q n + \frac{\gamma}{L} + \log_Q(Q-1) - \frac{1}{2} + O\left(\frac{1}{n}\right),
\end{aligned}$$

where $\gamma$ is Euler's constant and the harmonic numbers are given by $\log n + \gamma + O\left(\frac{1}{n}\right)$ as $n \to \infty$. But $f(z) = -\frac{(Q-1)^z}{Q^z-1}$ also has simple poles at $z = \chi_k = \frac{2k\pi \boldsymbol{i}}{L}$, $k \in \mathbb{Z}$, $k \neq 0$. For $\varepsilon := z - \chi_k$, we have

$$f(z) = -\frac{(Q-1)^z}{Q^z - 1} = -\frac{(Q-1)^{\varepsilon + \chi_k}}{Q^{\varepsilon + \chi_k} - 1} = -\frac{(Q-1)^{\varepsilon}(Q-1)^{\chi_k}}{Q^{\varepsilon} Q^{\chi_k} - 1}.$$

Now,

$$Q^{\chi_k} = (e^{\log Q})^{\frac{2k\pi \boldsymbol{i}}{L}} = e^{2k\pi \boldsymbol{i}} = 1, \tag{3.6}$$

so

$$f(z) = (Q-1)^{\chi_k}\Big(-\frac{(Q-1)^{\varepsilon}}{Q^{\varepsilon}-1}\Big).$$

Since

$$-\frac{(Q-1)^{\varepsilon}}{Q^{\varepsilon}-1} = -\frac{e^{\varepsilon\log(Q-1)}}{e^{\varepsilon\log Q}-1} \sim -\frac{1}{1+\varepsilon L-1} = -\frac{1}{\varepsilon L}, \tag{3.7}$$

we have that the residue of $f(z)$ is

$$[\varepsilon^{-1}](Q-1)^{\chi_k}\Big(-\frac{1}{\varepsilon L}\Big) = -\frac{1}{L}(Q-1)^{\chi_k}.$$

From [1], we can see that

$$[n;\chi_k] = \frac{\Gamma(-\chi_k)\Gamma(n+1)}{\Gamma(n+1-\chi_k)} = \Gamma(-\chi_k)n^{\chi_k}\Big(1+O\Big(\frac{1}{n}\Big)\Big),$$

and

$$(Q-1)^{\chi_k}n^{\chi_k} = e^{(\log n^*)\chi_k} = e^{2k\pi\boldsymbol{i}\log_Q n^*},$$

which means that we can write the main term of the fluctuations as $\delta_E(\log_Q n^*)$, with

$$\delta_E(x) := -\frac{1}{L}\sum_{k\neq 0}\Gamma(-\chi_k)e^{2k\pi\boldsymbol{i}x}. \tag{3.8}$$

We thus have a formula for the expected value so we can approximate the number of distinct letters in a word of length $n$ as:

$$\mathbb{E}(d_n) = \log_Q n + \frac{\gamma}{L} + \log_Q(Q-1) - \frac{1}{2} + \delta_E(\log_Q n^*) + O\Big(\frac{1}{n}\Big),$$

for $\delta_E(x)$ as in (3.8). This concludes the proof of Theorem 3.1. ■

## Remark:

It is of interest to compare this result with the mean of the *largest value* in a geometrically distributed sample of $n$ letters, denoted by $\mathbb{E}(M_n)$, due to Szpankowski and Rego [39]:

$$\mathbb{E}(M_n) = \log_Q n + \frac{\gamma}{L} + \frac{1}{2} + \delta_E(\log_Q n) + O\Big(\frac{1}{n}\Big), \qquad \text{as } n \to \infty. \tag{3.9}$$

Ignoring the small fluctuating terms we see that the expected *number of missing values* in the range 1 up to $\mathbb{E}(M_n)$ is asymptotically given by

$$\mathbb{E}(M_n) - \mathbb{E}(d_n) \approx 1 - \log_Q(Q-1). \tag{3.10}$$

Observe that as $Q$ goes from 1 to $\infty$ (or as $q$ goes from 1 to 0), expression (3.10) goes monotonically from infinity to 0. Thus $\mathbb{E}(d_n) \to \mathbb{E}(M_n)$ as $q = Q^{-1} \to 0$, which is intuitively clear, since the limiting word is just the sequence $111 \cdots 1$ with only one distinct value.

Our expected value is sandwiched between (3.9) and the number of consecutive non-empty boxes (equivalently the first value which does not occur in our sample). The case $q = \frac{1}{2}$ is dealt with in [9], where this value was given as

$$\mathbb{E}(c_n) = \log_2 n + \log_2 \varphi + P(\log_2 n) + o(1)$$

for $\varphi = 0.77351\ldots$ and periodic function $P(x)$ with period 1 and amplitude bounded by $10^{-5}$. We can now see that all three grow like $\log_2 n$ for $Q = 2$. Thus it is the constants that determine the (intuitive) ordering $\mathbb{E}(c_n) \leq \mathbb{E}(d_n) \leq \mathbb{E}(M_n)$. We calculate the constants numerically (correct to three decimal places) for the case $Q = 2$ to see by how much each expected value differs from the next.

| | $\mathbb{E}(c_n)$ | $\mathbb{E}(d_n)$ | $\mathbb{E}(M_n)$ |
|---|---|---|---|
| Constant | $\log_2 \varphi$ | $\frac{\gamma}{L} + \log_2 1 - \frac{1}{2}$ | $\frac{\gamma}{L} + \frac{1}{2}$ |
| Numerical value | $-0,371$ | 0,333 | 1,333 |

Table 3.1: Numerical values of the constant terms of $\mathbb{E}(c_n)$, $\mathbb{E}(d_n)$ and $\mathbb{E}(M_n)$.

## 3.2 The variance (classical case)

The formula for variance from a probability generating function is given in [12], with a similar application in [29].

$$\mathbb{V}(d_n) = n![z^n]\frac{\partial^2}{\partial u^2}F(z,u)\Big|_{u=1} + \mathbb{E}(d_n) - \mathbb{E}^2(d_n).$$

Again, the generating function we deal with is (see (3.1))

$$F(z,u) = \prod_{i\geq 0}\big(1 + u(e^{zpq^i} - 1)\big),$$

and so the first term in the variance expression can be calculated as follows: Let

$$f_i(z,u) := 1 + u(e^{zpq^i} - 1),$$

then since (from (3.4))

$$\prod_{i\geq 0} \left(1 + u(e^{zpq^i} - 1)\right)\Big|_{u=1} = e^z,$$

the second moment is

$$\begin{aligned} n![z^n]\frac{\partial^2}{\partial u^2}F(z,u)\Big|_{u=1} &= n![z^n]\frac{\partial^2}{\partial u^2}\prod_{i\geq 1} f_i(z,u)\Big|_{u=1} \\ &= n![z^n]\Bigg[\prod_{i\geq 1} f_i(z,u) 2\sum_{j<k} \frac{\frac{\partial}{\partial u}f_j(z,u)}{f_j(z,u)} \cdot \frac{\frac{\partial}{\partial u}f_k(z,u)}{f_k(z,u)} \\ &\qquad + \prod_{i\geq 1} f_i(z,u) \sum_j \frac{\frac{\partial^2}{\partial u^2}f_j(z,u)}{f_j(z,u)}\Bigg|_{u=1}\Bigg]. \end{aligned} \tag{3.11}$$

The second term in (3.11) is zero: any second partial derivative with respect to $u$ will be zero, as each $f_j(z,u)$ is linear with respect to $u$. Thus

$$\begin{aligned} n![z^n]\frac{\partial^2}{\partial u^2}F(z,u)\Big|_{u=1} &= n![z^n]2e^z\sum_{j<k}\frac{e^{zpq^j}-1}{e^{zpq^j}}\cdot\frac{e^{zpq^k}-1}{e^{zpq^k}} \\ &= n![z^n]2e^z\sum_{j<k}(1-e^{-zpq^j})(1-e^{-zpq^k}) \\ &= n![z^n]2\sum_{j<k}(e^z - e^{z(1-pq^j)} - e^{z(1-pq^k)} + e^{z(1-pq^j-pq^k)}) \\ &= 2\sum_{j<k}(1-(1-pq^j)^n - (1-pq^k)^n + (1-pq^j-pq^k)^n). \end{aligned}$$

This quantity can be split up (preserving convergence) as follows in order to be dealt with in two parts:

$$\underbrace{2\sum_{j<k}\left[1-(1-pq^k)^n\right]}_{\text{(a)}} + \underbrace{2\sum_{j<k}\left[(1-pq^j-pq^k)^n - (1-pq^j)^n\right]}_{\text{(b)}}. \tag{3.12}$$

The reason for this is that now the summand of (a) is independent of $j$ and can be dealt with separately from (b) which requires a slightly different approach. The factor of two is temporarily ignored.

**Part (a)**

Since $1-(1-pq^k)^n$ is independent of $j$,

$$\sum_{k\geq 0}\sum_{j=0}^{k-1}\left[1-(1-pq^k)^n\right] = \sum_{k\geq 0} k\left[1-(1-pq^k)^n\right].$$

This can be rewritten as an alternating sum so that Rice's method can be used. Using the binomial expansion we get

$$\begin{aligned}
\sum_{k\geq 0} k\big[1-(1-pq^k)^n\big] &= \sum_{k\geq 0} k\Big[1-\sum_{i=0}^{n}\binom{n}{i}(-pq^k)^i\Big] \\
&= \sum_{k\geq 0} k\Big[-\sum_{i=1}^{n}\binom{n}{i}(-pq^k)^i\Big] \\
&= \sum_{k\geq 0} k\sum_{i=1}^{n}\binom{n}{i}(-1)^{i+1}p^i q^{ki} \\
&= \sum_{i=1}^{n}\binom{n}{i}(-1)^{i+1}p^i\sum_{k\geq 0} k(q^i)^k \\
&= \sum_{i=1}^{n}\binom{n}{i}(-1)^{i+1}p^i\frac{q^i}{(1-q^i)^2} \\
&= \sum_{i=1}^{n}\binom{n}{i}(-1)^{i}\frac{-(pq)^i}{(1-q^i)^2} \\
&= \sum_{i=1}^{n}\binom{n}{i}(-1)^{i}\frac{-((1-q)q)^i}{(1-q^i)^2} \\
&= \sum_{i=1}^{n}\binom{n}{i}(-1)^{i}\frac{-(Q-1)^i}{(Q^i-1)^2}.
\end{aligned}$$

So $f(z)=\frac{-(Q-1)^z}{(Q^z-1)^2}$ and we have a triple pole at $z=0$ as $[n;z]$ has a simple pole and $\frac{-(Q-1)^z}{(Q^z-1)^2}$ has a double pole. If we expand to three terms we get

$$\frac{(-1)^{n-1}n!}{z(z-1)\cdots(z-n)} \sim -\frac{1}{z}\left(1+zH_n+z^2\frac{H_n^2+H_n^{(2)}}{2}\right),$$

(from [29]), and

$$\begin{aligned}
\frac{-(Q-1)^z}{(Q^z-1)^2} &= \frac{-e^{\log(Q-1)^z}}{(e^{\log Q^z}-1)^2} \\
&= \frac{-e^{z\log(Q-1)}}{(e^{z\log Q}-1)^2} \\
&\sim -\frac{1+z\log(Q-1)+\frac{z^2\log^2(Q-1)}{2}}{(1+z\log Q+\frac{z^2\log^2 Q}{2}+\frac{z^3\log^3 Q}{6}-1)^2} \\
&= -\frac{1+z\log(Q-1)+\frac{z^2\log^2(Q-1)}{2}}{z^2L^2(1+\frac{zL}{2}+\frac{z^2L^2}{6})^2} \\
&\sim -\frac{1}{z^2L^2}\left(1+z\log(Q-1)+\frac{z^2\log^2(Q-1)}{2}\right) \\
&\qquad\qquad \cdot\left(1-\left(\frac{zL}{2}+\frac{z^2L^2}{6}\right)+\frac{z^2L^2}{4}\right)^2
\end{aligned}$$

$$\sim -\frac{1}{z^2L^2}\left(1 + z\log(Q-1) + \frac{z^2\log^2(Q-1)}{2}\right)\left(1 - zL + \frac{5z^2L^2}{12}\right).$$

We now briefly note that as $n \to \infty$ (see [36])

$$H_n^2 = \left(\log n + \gamma + O\left(\frac{1}{n}\right)\right)^2 = \log^2 n + 2\gamma\log n + \gamma^2 + o(1), \tag{3.13}$$

and

$$H_n^{(2)} = \frac{\pi^2}{6} + O\left(\frac{1}{n}\right). \tag{3.14}$$

The residue for the triple pole at $z = 0$ as $n \to \infty$ is

$$\begin{aligned}
&[z^{-1}]\frac{1}{z^3L^2}\left(1 + z\log(Q-1) + \frac{z^2\log^2(Q-1)}{2}\right)\left(1 - zL + \frac{5z^2L^2}{12}\right)\\
&\qquad\qquad\qquad\qquad\cdot\left(1 + zH_n + z^2\frac{H_n^2 + H_n^{(2)}}{2}\right)\\
&= [z^2]\frac{1}{L^2}\left(1 + z\log(Q-1) + \frac{z^2\log^2(Q-1)}{2}\right)\left(1 - zL + \frac{5z^2L^2}{12}\right)\\
&\qquad\qquad\qquad\qquad\cdot\left(1 + zH_n + z^2\frac{H_n^2 + H_n^{(2)}}{2}\right)\\
&= \frac{1}{L^2}\left(\frac{\log^2(Q-1)}{2} + \frac{5L^2}{12} + \frac{H_n^2 + H_n^{(2)}}{2} - L\log(Q-1) + H_n\log(Q-1) - LH_n\right)\\
&= \frac{\log_Q^2(Q-1)}{2} + \frac{5}{12} + \frac{H_n^2 + H_n^{(2)}}{2L^2} - \log_Q(Q-1) + \frac{\log_Q(Q-1)H_n}{L} - \frac{H_n}{L}\\
&= \frac{\log_Q^2(Q-1)}{2} + \frac{5}{12} + \frac{1}{2L^2}(\log^2 n + 2\gamma\log n + \gamma^2) + \frac{1}{2L^2}\left(\frac{\pi^2}{6}\right) - \log_Q(Q-1)\\
&\qquad + \frac{\log_Q(Q-1)(\log n + \gamma)}{L} - \frac{\log n + \gamma}{L} + o(1) \quad \text{(by (3.13) and (3.14))}\\
&= \frac{1}{2}\log_Q^2 n + \frac{\gamma}{L}\log_Q n + \log_Q(Q-1)\log_Q n - \log_Q n + \frac{1}{2}\log_Q^2(Q-1)\\
&\qquad - \log_Q(Q-1) + \frac{\gamma}{L}\log_Q(Q-1) + \frac{5}{12} + \frac{\pi^2}{12L^2} + \frac{\gamma^2}{2L^2} - \frac{\gamma}{L} + o(1).
\end{aligned}$$

Now $f(z)$ also has double poles at $z = \chi_k = \frac{2k\pi \boldsymbol{i}}{L}$, $k \in \mathbb{Z}$, $k \neq 0$. By letting $\varepsilon := z - \chi_k$, we can do the following

$$f(z) = \frac{-(Q-1)^z}{(Q^z-1)^2} = \frac{-(Q-1)^{\varepsilon+\chi_k}}{(Q^{\varepsilon+\chi_k}-1)^2} = \frac{-(Q-1)^{\chi_k}(Q-1)^{\varepsilon}}{(Q^{\varepsilon}Q^{\chi_k}-1)^2}.$$

By (3.6), $Q^{\chi_k} = 1$, so

$$f(z) = (Q-1)^{\chi_k}\left(-\frac{(Q-1)^{\varepsilon}}{(Q^{\varepsilon}-1)^2}\right).$$

We have already expanded the fraction to three terms for the pole at $z = 0$, so we merely note the expansion to two terms as

$$f(z) \sim (Q-1)^{\chi_k}\frac{-1}{\varepsilon^2L^2}(1 + \varepsilon\log(Q-1))(1 - \varepsilon L). \tag{3.15}$$

Lastly,

$$[n;\chi_k] = \frac{\Gamma(-z)\Gamma(n+1)}{\Gamma(n+1-z)}$$

needs to be expanded to 2 terms around $z = \chi_k$. The term $\Gamma(n+1)$ is just a constant in this case. Using a Taylor expansion we can write

$$\Gamma(-z) \sim \Gamma(-\chi_k) - \Gamma'(-\chi_k)(z-\chi_k) = \Gamma(-\chi_k)\big[1 - \psi(-\chi_k)(z-\chi_k)\big],$$

where $\psi(x)$ is the Digamma function, and similarly

$$\Gamma(n+1-z) \sim \Gamma(n+1-\chi_k)\big[1 - \psi(n+1-\chi_k)(z-\chi_k)\big].$$

This means that with the same substitution as before ($\varepsilon := z - \chi_k$), we have

$$[n;\chi_k] \sim \Gamma(n+1)\frac{\Gamma(-\chi_k)}{\Gamma(n+1-\chi_k)}\big[1 - \psi(-\chi_k)\varepsilon + \psi(n+1-\chi_k)\varepsilon\big],$$

around $\varepsilon = 0$. We approximate the $\psi$ function by [1, page 259]

$$\psi(n+1-\chi_k) \sim \log(n+1-\chi_k) = \log\Big(n\Big(1 + \frac{1-\chi_k}{n}\Big)\Big) \sim \log n,$$

as $n \to \infty$, so that

$$\begin{aligned}
[n;\chi_k] &\sim \Gamma(n+1)\frac{\Gamma(-\chi_k)}{\Gamma(n+1-\chi_k)}\big[1 - \psi(-\chi_k)\varepsilon + \varepsilon\log n\big] \\
&= \Gamma(-\chi_k)\frac{\Gamma(n+1)}{\Gamma(n+1-\chi_k)}\big[1 - \psi(-\chi_k)\varepsilon + \varepsilon\log n\big] \\
&\sim \Gamma(-\chi_k)n^{\chi_k}\big[1 - \psi(-\chi_k)\varepsilon + \varepsilon\log n\big],
\end{aligned} \tag{3.16}$$

as $n \to \infty$. If we put (3.15) and (3.16) together we get

$$[n;\chi_k]f(z) \sim (Q-1)^{\chi_k}\frac{-1}{\varepsilon^2L^2}\Gamma(-\chi_k)n^{\chi_k}\big[1-\psi(-\chi_k)\varepsilon+\varepsilon\log n\big](1+\varepsilon\log(Q-1))(1-\varepsilon L),$$

as $n \to \infty$, and by rewriting $(Q-1)^{\chi_k}n^{\chi_k}$ as

$$(Q-1)^{\chi_k}n^{\chi_k} = e^{\chi_k \log n^*} = e^{2k\pi\boldsymbol{i}\log_Q n^*},$$

we get the residue (coefficient of $\varepsilon^{-1}$) from the poles at $z = \chi_k$, $k \neq 0$ to be (asymptotically)

$$\begin{aligned}
&\sum_{k\neq 0} e^{2k\pi\boldsymbol{i}\log_Q n^*}\Gamma(-\chi_k)\frac{-1}{L^2}\big[-\psi(-\chi_k) + \log n + \log(Q-1) - L\big] \\
&\quad= \frac{-1}{L^2}\sum_{k\neq 0} e^{2k\pi\boldsymbol{i}\log_Q n^*}\Gamma(-\chi_k)\big[-\psi(-\chi_k) + \log n + \log(Q-1) - L\big] \\
&\quad= \frac{-1}{L}\sum_{k\neq 0} e^{2k\pi\boldsymbol{i}\log_Q n^*}\Gamma(-\chi_k)\Big[\log_Q n - \frac{\psi(-\chi_k)}{L} + \log_Q(Q-1) - 1\Big].
\end{aligned}$$

So the total result for Part (a) is:

$$
\begin{aligned}
&\sum_{k\geq 0}\sum_{j=0}^{k-1}\left[1-(1-pq^k)^n\right] \\
&\quad= \frac{1}{2}\log_Q^2 n + \frac{\gamma}{L}\log_Q n + \log_Q(Q-1)\log_Q n - \log_Q n + \frac{1}{2}\log_Q^2(Q-1) \\
&\quad\quad - \log_Q(Q-1) + \frac{\gamma}{L}\log_Q(Q-1) + \frac{5}{12} + \frac{\pi^2}{12L^2} + \frac{\gamma^2}{2L^2} - \frac{\gamma}{L} \\
&\quad\quad - \frac{1}{L}\sum_{k\neq 0} e^{2k\pi \boldsymbol{i}\log_Q n^*}\Gamma(-\chi_k)\left[\log_Q n - \frac{\psi(-\chi_k)}{L} + \log_Q(Q-1) - 1\right] + o(1).
\end{aligned}
\tag{3.17}
$$

**Part (b)**

Applying the Binomial Theorem to (b) gives

$$
\begin{aligned}
&\sum_{j<k}\left[(1-pq^j-pq^k)^n - (1-pq^j)^n\right] \\
&\qquad\qquad = \sum_{j<k}\left[\sum_{i\geq 0}\binom{n}{i}(-pq^j-pq^k)^i - \sum_{i\geq 0}\binom{n}{i}(-pq^j)^i\right] \\
&\qquad\qquad = \sum_{i=1}^{n}\binom{n}{i}(-1)^{i-1}\sum_{j<k}\left[(pq^j)^i - (pq^j+pq^k)^i\right]
\end{aligned}
\tag{3.18}
$$

(since term $i=0$ is zero). This is now written in the correct form for Rice's method to be used, where

$$
\begin{aligned}
f(z) &= -\sum_{j<k}\left[(pq^j)^z - (pq^j+pq^k)^z\right] \\
&= -\sum_{j<k}(pq^j)^z\left[1-(1+q^{k-j})^z\right] \\
&= -\sum_{j\geq 0}(pq^j)^z\sum_{m\geq 1}\left[1-(1+q^m)^z\right], \qquad \text{for } m := k-j \\
&= -p^z\sum_{j\geq 0}(q^z)^j\sum_{m\geq 1}\left[1-(1+q^m)^z\right] \\
&= -\frac{p^z}{1-q^z}\, g(z),
\end{aligned}
$$

if we define

$$g(z) := \sum_{m\geq 1}\left[1-(1+q^m)^z\right].$$

In order to find out more about the function $g(z)$, we expand $g(z)$ around $z=0$

$$g(z) = \sum_{m\geq 1}\left[1-(1+q^m)^z\right]$$

$$
\begin{aligned}
&= \sum_{m\ge1}\left[1-e^{\log(1+q^m)^z}\right] \\
&= \sum_{m\ge1}\left[1-e^{z\log(1+q^m)}\right] \\
&= \sum_{m\ge1}\left[1-\left(1+z\log(1+q^m)+\frac{z^2\log^2(1+q^m)}{2}+\cdots\right)\right] \\
&= \sum_{m\ge1}\left[-z\log(1+q^m)-\frac{z^2\log^2(1+q^m)}{2}+\cdots\right] \\
&= z\sum_{m\ge1}-\log(1+q^m)-\frac{z^2}{2}\sum_{m\ge1}\log^2(1+q^m)+\cdots \\
&= z\sum_{m\ge1}\sum_{k\ge1}\frac{(-1)^k(q^m)^k}{k}-\frac{z^2}{2}\sum_{m\ge1}\left(\sum_{k\ge1}\frac{(-1)^{k+1}(q^m)^k}{k}\right)^2+\cdots \\
&= z\sum_{k\ge1}\frac{(-1)^k}{k}\sum_{m\ge1}(q^m)^k-\frac{z^2}{2}\sum_{k\ge1}\sum_{j\ge1}\frac{(-1)^{k+j}}{kj}\sum_{m\ge1}q^{mk+mj}+\cdots \\
&= \underbrace{z\sum_{k\ge1}\frac{(-1)^k}{k}\frac{q^k}{1-q^k}}_{\alpha}+z^2\underbrace{\left(-\frac{1}{2}\sum_{k\ge1}\sum_{j\ge1}\frac{(-1)^{k+j}}{kj}\frac{q^{k+j}}{1-q^{k+j}}\right)}_{\beta}+\cdots.
\end{aligned}
$$

With the aid of Mathematica, $\alpha$ and $\beta$ can be evaluated to give constants for fixed values of $q$. For example, see Table 3.2 below for these values for $q=\frac{1}{2}$ and $q=\frac{1}{3}$.

| Constant | $\alpha$ | $\beta$ |
|---|---|---|
| Definition | $\sum_{k\ge1}\frac{(-1)^k}{k}\frac{q^k}{1-q^k}$ | $-\frac{1}{2}\sum_{k\ge1}\sum_{j\ge1}\frac{(-1)^{k+j}}{kj}\frac{q^{k+j}}{1-q^{k+j}}$ |
| At $q=\frac{1}{2}$ | $-0.868877$ | $-0.116506$ |
| At $q=\frac{1}{3}$ | $-0.447844$ | $-0.047677$ |

Table 3.2: The values of the constants $\alpha$ and $\beta$ when $q=\frac{1}{2}$ and $q=\frac{1}{3}$.

So $g(z)$ can be written as $g(z)=\alpha z+\beta z^2+\cdots$ where $\alpha,\beta,\ldots$ are constants. A polynomial does not have a pole at $z=0$, and thus when we apply Rice's method to (3.18) we have a simple pole at $z = 0$ since there is one pole in $[n; z]$ and none in $f(z)$ (there would be one pole in the expression $-\frac{p^z}{1-q^z} = -\frac{(Q-1)^z}{Q^z-1}$, but it is cancelled by the zero of $g(0)$). Consequently we expand everything to one term, giving

$$\frac{(-1)^{n-1}n!}{z(z-1)\cdots(z-n)} \sim -\frac{1}{z},$$
$$-\frac{(Q-1)^z}{Q^z-1} \sim -\frac{1}{zL},$$

and

$$g(z) \sim \alpha z.$$

The residue for $z = 0$ is thus

$$[z^{-1}]\Big(-\frac{1}{z}\Big)\Big(-\frac{1}{zL}\Big)\alpha z = [z]\frac{\alpha z}{L} = \frac{\alpha}{L}.$$

But $-\frac{(Q-1)^z}{Q^z-1}$ also has simple poles at $z = \chi_k = \frac{2k\pi \boldsymbol{i}}{L}$, $k \in \mathbb{Z}$, $k \neq 0$. We rearrange $g(z)$ as follows to get the contribution of $g(z)$ at $\chi_k$.

$$\begin{aligned}
g(\chi_k) &= \sum_{m\geq 1}\big[1-(1+q^m)^{\chi_k}\big] \\
&= \sum_{m\geq 1}\Big[1-\sum_{l\geq 0}\binom{\chi_k}{l}(q^m)^l\Big] \\
&= -\sum_{m\geq 1}\sum_{l\geq 1}\binom{\chi_k}{l}(q^m)^l \\
&= -\sum_{l\geq 1}\binom{\chi_k}{l}\sum_{m\geq 1}(q^l)^m \\
&= -\sum_{l\geq 1}\binom{\chi_k}{l}\frac{q^l}{1-q^l} \\
&= -\sum_{l\geq 1}\binom{\chi_k}{l}\frac{l}{Q^l-l}.
\end{aligned}$$

To work out the residue of $-\frac{(Q-1)^z}{Q^z-1}$, we again let $\varepsilon := z - \chi_k$, and use the same method as in Part (a):

$$-\frac{(Q-1)^z}{Q^z-1} = -\frac{(Q-1)^{\varepsilon+\chi_k}}{Q^{\varepsilon+\chi_k}-1} = -(Q-1)^{\chi_k}\frac{(Q-1)^{\varepsilon}}{Q^{\varepsilon}-1},$$

since $Q^{\chi_k} = 1$ as in (3.6). Now

$$\frac{(Q-1)^{\varepsilon}}{Q^{\varepsilon}-1} \sim \frac{1}{\varepsilon L},$$

and so the residue is $\frac{1}{L}$. As in Part (a)

$$(Q-1)^{\chi_k}[n;\chi_k] \sim \Gamma(-\chi_k)e^{2k\pi \boldsymbol{i}\log_Q n^*},$$

and so the main term of the contribution from the simple poles at $z = \chi_k$ is

$$-\frac{1}{L}\sum_{k\neq 0} g(\chi_k)\Gamma(-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*},$$

which means that the total result for Part (b) is

$$\sum_{j<k}\left[(1-pq^j-pq^k)^n-(1-pq^j)^n\right] = \frac{\alpha}{L} - \frac{1}{L}\sum_{k\neq 0} g(\chi_k)\Gamma(-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*} + o(1).$$

**The variance resolved**

Now that (a) and (b) have been found, we can return to the variance calculations, which in addition to the second moment (equation (3.12)), must include the following two terms (the approximate expected value is given in (3.2)).

$$\mathbb{E}(d_n) = \log_Q n + \frac{\gamma}{L} + \log_Q(Q-1) - \frac{1}{2} + \delta_E(\log_Q n^*) + O\left(\frac{1}{n}\right).$$

and

$$\begin{aligned}\mathbb{E}^2(d_n) = {} & \log_Q^2 n + 2\log_Q n\,\delta_E(\log_Q n^*) + 2\log_Q n\log_Q(Q-1) + \frac{2\,\gamma\log_Q n}{L} \\ & - \log_Q n - \log_Q(Q-1) + \frac{2\,\gamma\log_Q(Q-1)}{L} + \log_Q^2(Q-1) \\ & + 2\log_Q(Q-1)\,\delta_E(\log_Q n^*) + \frac{1}{4} - \frac{\gamma}{L} + \frac{\gamma^2}{L^2} - \delta_E(\log_Q n^*) \\ & + \frac{2\,\gamma\,\delta_E(\log_Q n^*)}{L} + \delta_E^2(\log_Q n^*) + o(1).\end{aligned}$$

We can now put all of these together (remembering that Part (a) and Part (b) must include a factor of two) to get

$$\begin{aligned}\mathbb{V}(d_n) &= n![z^n]\frac{\partial^2}{\partial u^2}F(z,u)\Big|_{u=1} + \mathbb{E}(d_n) - \mathbb{E}^2(d_n) \\ &= \log_Q^2 n + \frac{2\gamma}{L}\log_Q n + 2\log_Q(Q-1)\log_Q n - 2\log_Q n + \log_Q^2(Q-1) \\ &\quad - 2\log_Q(Q-1) + \frac{2\gamma}{L}\log_Q(Q-1) + \frac{5}{6} + \frac{\pi^2}{6L^2} + \frac{\gamma^2}{L^2} - \frac{2\gamma}{L} \\ &\quad - \frac{2}{L}\sum_{k\neq 0} e^{2k\pi\boldsymbol{i}\log_Q n^*}\Gamma(-\chi_k)\Big[\log_Q n - \frac{\psi(-\chi_k)}{L} + \log_Q(Q-1) - 1\Big] \\ &\quad + \frac{2\alpha}{L} - \frac{2}{L}\sum_{k\neq 0} g(\chi_k)\Gamma(-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*} \\ &\quad + \log_Q n + \frac{\gamma}{L} + \log_Q(Q-1) - \frac{1}{2} + \delta_E(\log_Q n^*) \\ &\quad - \Big[\log_Q^2 n + 2\log_Q n\delta_E(\log_Q n^*) + 2\log_Q n\log_Q(Q-1) + \frac{2\gamma\log_Q n}{L}\end{aligned}$$

$$
\begin{aligned}
&- \log_Q n - \log_Q(Q-1) + \frac{2\gamma \log_Q(Q-1)}{L} + \log_Q^2(Q-1) \\
&+ 2\log_Q(Q-1)\delta_E(\log_Q n^*) + \frac{1}{4} - \frac{\gamma}{L} + \frac{\gamma^2}{L^2} - \delta_E(\log_Q n^*) \\
&+ \frac{2\gamma\delta_E(\log_Q n^*)}{L} + \delta_E^2(\log_Q n^*)\Big] + o(1) \\
= &\frac{1}{12} + \frac{\pi^2}{6L^2} + \frac{2\alpha}{L} \\
&- \frac{2}{L}\sum_{k\neq 0} e^{2k\pi \boldsymbol{i} \log_Q n^*}\Gamma(-\chi_k)\Big[\log_Q n - \frac{\psi(-\chi_k)}{L} + \log_Q(Q-1) - 1\Big] \\
&- \frac{2}{L}\sum_{k\neq 0} g(\chi_k)\Gamma(-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*} + \delta_E(\log_Q n^*) - 2\log_Q n\delta_E(\log_Q n^*) \\
&- 2\log_Q(Q-1)\delta_E(\log_Q n^*) + \delta_E(\log_Q n^*) - \frac{2\gamma\delta_E(\log_Q n^*)}{L} \\
&- \delta_E^2(\log_Q n^*) + o(1),
\end{aligned}
$$

where

$$\delta_E(x) = -\frac{1}{L}\sum_{k\neq 0}\Gamma(-\chi_k)e^{2k\pi\boldsymbol{i}x}.$$

We can split up the $\delta_E^2(\log_Q n^*)$ term into a constant term (the mean of the fluctuating function) and a fluctuating function of period 1 and mean zero (see [17] and (4.9)). Let

$$
\begin{aligned}
\delta_E^2(\log_Q n^*) &:= [\delta_E^2]_0 + \hat{\delta}_E(\log_Q n^*) \\
&= \frac{\pi^2}{6L^2} + \frac{1}{12} - \log_Q 2 - \frac{2}{L}\sum_{h\geq 1}\frac{(-1)^{h-1}}{h(Q^h-1)} + \hat{\delta}_E(\log_Q n^*).
\end{aligned}
$$

Then we have $\left(\text{for } \alpha = \sum_{k\geq 1}\frac{(-1)^k}{k}\frac{q^k}{1-q^k}\right)$

$$
\begin{aligned}
\mathbb{V}(d_n) &= \log_Q 2 + \frac{2\alpha}{L} + \frac{2}{L}\sum_{h\geq 1}\frac{(-1)^{h-1}}{h(Q^h-1)} + \delta_V(\log_Q n^*) + o(1) \\
&= \log_Q 2 + \delta_V(\log_Q n^*) + o(1),
\end{aligned}
$$

where

$$
\begin{aligned}
\delta_V(x) &:= \frac{2}{L}\sum_{k\neq 0}\Gamma(-\chi_k)e^{2k\pi\boldsymbol{i}x}\Big[\frac{\psi(-\chi_k)}{L} - g(\chi_k) + \frac{\gamma}{L}\Big] - \hat{\delta}_E(x) \\
&= \delta_E(x + \log_Q 2) - \delta_E(x), \qquad (3.19)
\end{aligned}
$$

for $\delta_E(x)$ as in (3.8), and with

$$g(x) = -\sum_{l\geq 1}\binom{x}{l}\frac{q^l}{1-q^l}.$$

Appendix A shows the simplification of (3.19), which concludes the proof of Theorem 3.2. ■

### 3.2.1 Extreme cases of $\alpha$

For interest we look at the extreme cases of $\alpha$ in $g(z) = \alpha z + \beta z^2 + \cdots$. As $q \to 0$, $\alpha \to 0$. As $q \to 1$, then if $q = e^{-t}$, we can instead consider $t \to 0$. I.e.,

$$\alpha = \sum_{k\geq 1} \frac{(-1)^k}{k} \frac{q^k}{1-q^k} = \sum_{k\geq 1} \frac{(-1)^k}{k} \frac{e^{-tk}}{1-e^{-tk}} = -\sum_{k\geq 1} \frac{(-1)^{k-1}}{k} \frac{1}{e^{tk}-1}.$$

This can be found in the appendix of [20], and by defining it as a function of $t$, say $h(t)$ we get the following result from that paper, which makes use of Mellin transforms to get:

$$\alpha = h(t) = -\frac{\pi^2}{12t} + \frac{\log 2}{2} - \frac{t}{24} + h\left(\frac{2\pi^2}{t}\right).$$

This identity holds for $0 < t < 2\pi^2$. We are interested in what happens as $t \to 0$. Since

$$h(t) = -\sum_{k\geq 1} \frac{(-1)^{k-1}}{k} \frac{1}{e^{tk}-1},$$

it can be seen that $h\left(\frac{2\pi^2}{t}\right) \to 0$ as $t \to 0$, and thus the last term in the expression for $\alpha$ is small enough to be insignificant. The remaining three terms provide an approximation for $\alpha$ near $q = 1$ where $t = \log \frac{1}{q}$.

# Chapter 4

# General case: The number of letters occurring at least $b$ times

We now generalise this idea and consider the number of values which appear at least $b$ times in a word. Our probability generating function needs to be extended. To do this we need to ensure that $u$ only takes into account those values that occur at least $b$ times. So it is necessary to subtract the letters that occur fewer times, and add them elsewhere. For example, for $b = 3$, we would have

$$F_3(z,u) := \prod_{i\geq 0}\left(1 + zpq^i + \frac{(zpq^i)^2}{2} + u\left(e^{zpq^i} - 1 - zpq^i - \frac{(zpq^i)^2}{2}\right)\right),$$

where the term $zpq^i$ corresponds to the letter $i$ occurring exactly once in the word and $\frac{(zpq^i)^2}{2}$ corresponds to the letter $i$ occurring exactly twice in the word. For general $b$ we have

$$F_b(z,u) := \prod_{i\geq 0}\left(\sum_{k=0}^{b-1}\frac{(zpq^i)^k}{k!} + u\left(e^{zpq^i} - \sum_{k=0}^{b-1}\frac{(zpq^i)^k}{k!}\right)\right). \tag{4.1}$$

The results proved in this chapter are presented below.

**Theorem 4.1** *The expected number of digits occurring at least b times in a word of length n whose letters are independently generated with geometric probability is*

$$\mathbb{E}_b(d_n) = \log_Q n + \frac{\gamma}{L} + \log_Q(Q-1) - \frac{1}{2} - \frac{1}{L}H_{b-1} + \delta_{E_b}(\log_Q n^*) + O\left(\frac{1}{n}\right), \tag{4.2}$$

*as $n \to \infty$, where*

$$\delta_{E_b}(x) = \frac{1}{L}\sum_{j\neq 0}\frac{e^{2j\pi \boldsymbol{i} x}}{\chi_j}\frac{\Gamma(b-\chi_j)}{\Gamma(b)},$$

*as in (4.9).*

**Theorem 4.2** *The variance of this quantity is*

$$
\begin{aligned}
\mathbb{V}_b(d_n) = \log_Q 2 &+ \frac{2}{L}\sum_{i\ge 1}\frac{(-1)^{i+b-1}}{i(Q^i-1)}\binom{i+b-1}{i}\binom{i-1}{b-1} \\
&- \frac{2}{L}\sum_{j=1}^{b-1}\frac{1}{2j}\binom{2j}{j}\sum_{h\ge 0}\binom{-2j}{h}\frac{1}{Q^{h+j}-1} + \frac{2}{L}\sum_{h\ge 1}\frac{(-1)^{h-1}}{h(Q^h-1)} \\
&- \frac{1}{L}\sum_{j=1}^{b-1}\frac{1}{2j}\binom{2j}{j}2^{-2j} + \delta_{V_b}(\log_Q n^*) + o(1), \qquad (4.3)
\end{aligned}
$$

*as $n \to \infty$. The fluctuating function $\delta_{V_b}(x)$ is defined in (4.30).*

## 4.1 The expected value (general case)

To find the average number of letters occurring $b$ times or more in a sequence of length $n$ whose letters occur independently and with geometric distribution, we differentiate (4.1) partially with respect to $u$, then replace $u$ by 1 and find the coefficient of $z^n$ (not forgetting the $n!$ since this is an exponential generating function).

$$
\begin{aligned}
\mathbb{E}_b(d_n) &= n![z^n]\frac{\partial}{\partial u}\prod_{i\ge 0}\left(\sum_{k=0}^{b-1}\frac{(zpq^i)^k}{k!} + u\left(e^{zpq^i} - \sum_{k=0}^{b-1}\frac{(zpq^i)^k}{k!}\right)\right)\Bigg|_{u=1} \\
&= n![z^n]\sum_{i\ge 0}\frac{e^z\left(e^{zpq^i} - \sum_{k=0}^{b-1}\frac{(zpq^i)^k}{k!}\right)}{e^{zpq^i}} \\
&= n![z^n]\sum_{i\ge 0}\frac{e^{z(1+pq^i)} - e^z\sum_{k=0}^{b-1}\frac{(zpq^i)^k}{k!}}{e^{zpq^i}} \\
&= n![z^n]\sum_{i\ge 0}\left(e^z - e^{z(1-pq^i)}\sum_{k=0}^{b-1}\frac{(zpq^i)^k}{k!}\right) \\
&= \sum_{i\ge 0}\left(1 - n![z^n]\left(e^{z(1-pq^i)}\sum_{k=0}^{b-1}\frac{(zpq^i)^k}{k!}\right)\right) \\
&= \sum_{i\ge 0}\left(1 - \sum_{k=0}^{b-1}\binom{n}{k}(1-pq^i)^{n-k}(pq^i)^k\right) \\
&= \sum_{i\ge 0}\left(1 - \sum_{k=0}^{b-1}\binom{n}{k}\sum_{j\ge 0}\binom{n-k}{j}(-pq^i)^j(pq^i)^k\right) \\
&= \sum_{i\ge 0}\left(1 - \sum_{j\ge 0}\binom{n}{j}(-pq^i)^j\right) - \sum_{i\ge 0}\sum_{k=1}^{b-1}\binom{n}{k}\sum_{j\ge 0}\binom{n-k}{j}(-pq^i)^j(pq^i)^k
\end{aligned}
$$

$$= -\sum_{i\ge 0}\sum_{j\ge 1}\binom{n}{j}(-pq^i)^j - \sum_{i\ge 0}\sum_{k=1}^{b-1}\binom{n}{k}\sum_{j\ge 0}\binom{n-k}{j}(-pq^i)^j(pq^i)^k$$

$$= -\sum_{i\ge 0}\sum_{j\ge 1}\binom{n}{j}(-1)^j p^j q^{ij} - \sum_{i\ge 0}\sum_{k=1}^{b-1}\binom{n}{k}\sum_{j\ge 0}\binom{n-k}{j}(-1)^j p^{j+k} q^{i(j+k)}$$

$$= -\sum_{j\ge 1}\binom{n}{j}(-1)^j p^j \sum_{i\ge 0} q^{ij} - \sum_{k=1}^{b-1}\binom{n}{k}\sum_{j\ge 0}\binom{n-k}{j}(-1)^j p^{j+k}\sum_{i\ge 0} q^{i(j+k)}$$

$$= \underbrace{\sum_{j=1}^{n}\binom{n}{j}(-1)^{j-1}\frac{p^j}{1-q^j}}_{\triangle} - \sum_{k=1}^{b-1}\binom{n}{k}\underbrace{\sum_{j=0}^{n-k}\binom{n-k}{j}(-1)^j\frac{p^{j+k}}{1-q^{j+k}}}_{\triangledown}. \qquad (4.4)$$

In the above expression, $\triangle$ is our original expected value (i.e., for the number of distinct values) and Rice's method (with the contour of integration surrounding $0,\ldots,N$, for $N = n-k$) can be used for $\triangledown$:

$$\triangledown = \sum_{j=0}^{n-k}\binom{n-k}{j}(-1)^j\frac{p^{j+k}}{1-q^{j+k}} = \sum_{j=0}^{N}\binom{N}{j}(-1)^j\frac{(1-Q^{-1})^{j+k}}{1-Q^{-(j+k)}}.$$

We have the function

$$f_k(z) := \frac{(1-Q^{-1})^{z+k}}{1-Q^{-(z+k)}} = \frac{(Q-1)^{z+k}}{Q^{z+k}-1}, \qquad (4.5)$$

which has a simple pole at $z=-k$, since $k\ge 1$. The contribution of $[N;z]$ around $z=-k$ is

$$\begin{aligned}[n-k;-k] &= \frac{(-1)^{n-k-1}(n-k)!}{(-k)(-k-1)\cdots(-k-(n-k))}\\ &= \frac{(-1)^{n-k-1}(n-k)!}{(-1)^{n-k+1}(k)(k+1)\cdots(n)}\\ &= \frac{(n-k)!(k-1)!}{n!}. \qquad (4.6)\end{aligned}$$

To expand $f_k(z)$ to one term around the simple pole at $z=-k$, let $\varepsilon := z+k$ and then use (3.7) to obtain

$$f_k(z) \sim \frac{1}{\varepsilon L},$$

and so the residue is

$$[\varepsilon^{-1}]\frac{1}{\varepsilon L}\frac{(n-k)!(k-1)!}{n!} = \frac{(n-k)!(k-1)!}{Ln!}.$$

This can now be substituted into (4.4) as the inner sum ($\triangledown$), giving

$$\sum_{k=1}^{b-1}\binom{n}{k}\frac{(n-k)!(k-1)!}{Ln!} = \sum_{k=1}^{b-1}\frac{1}{Lk} = \frac{1}{L}H_{b-1}. \qquad (4.7)$$

Lastly, we need to calculate the fluctuations contributed by the simple poles at $z+k=\chi_j$, $j\in\mathbb{Z}$, $j\neq 0$. We have

$$f_k(z)=\frac{(Q-1)^{z+k}}{Q^{z+k}-1},$$

so for $\varepsilon := z+k-\chi_j$, and from (3.6) and (3.7),

$$f_k(z)=\frac{(Q-1)^{\varepsilon+\chi_j}}{Q^{\varepsilon+\chi_j}-1}=(Q-1)^{\chi_j}\frac{(Q-1)^{\varepsilon}}{Q^{\varepsilon}-1}\sim(Q-1)^{\chi_j}\frac{1}{\varepsilon L},$$

so the residue is

$$(Q-1)^{\chi_j}\frac{1}{L}.$$

The contribution of $[N;z]$ around $z=-k+\chi_j$ is (refer to [1])

$$\begin{aligned}[n-k;-k+\chi_j]&=\frac{\Gamma(k-\chi_j)\Gamma(n-k+1)}{\Gamma(n-k+1+k-\chi_j)}\\&=\frac{\Gamma(k-\chi_j)\Gamma(n-k+1)}{\Gamma(n+1-\chi_j)}\\&=\Gamma(k-\chi_j)n^{\chi_j-k}\left(1+O\left(\frac{1}{n}\right)\right).\end{aligned}\tag{4.8}$$

Again we can write

$$(Q-1)^{\chi_j}n^{\chi_j}=e^{(\log n^*)\chi_j}=e^{2\pi\boldsymbol{i}j\log_Q n^*}.$$

This means that for each value of $k$ we have a main term contribution of

$$\frac{1}{L}\sum_{j\neq 0}\Gamma(k-\chi_j)n^{-k}e^{2\pi\boldsymbol{i}j\log_Q n^*}.$$

We sum this to get

$$\sum_{k=1}^{b-1}\binom{n}{k}\frac{1}{L}\sum_{j\neq 0}\Gamma(k-\chi_j)n^{-k}e^{2\pi\boldsymbol{i}j\log_Q n^*},$$

which we can subtract from the $\delta_E$ function in the case $b=1$ (see (3.8)) to get the fluctuations to be

$$\begin{aligned}\frac{-1}{L}\sum_{j\neq 0}e^{2j\pi\boldsymbol{i}\log_Q n^*}&\left(\Gamma(-\chi_j)+\sum_{k=1}^{b-1}\binom{n}{k}\Gamma(k-\chi_j)n^{-k}\right)\\&=\frac{-1}{L}\sum_{k=0}^{b-1}\binom{n}{k}n^{-k}\sum_{j\neq 0}e^{2j\pi\boldsymbol{i}\log_Q n^*}\Gamma(k-\chi_j)\\&=\frac{-1}{L}\sum_{k=0}^{b-1}\frac{n^{\underline{k}}}{k!n^k}\sum_{j\neq 0}e^{2j\pi\boldsymbol{i}\log_Q n^*}\Gamma(k-\chi_j)\end{aligned}$$

$$= \frac{-1}{L} \sum_{k=0}^{b-1} \sum_{j \neq 0} e^{2j\pi \boldsymbol{i} \log_Q n^*} \frac{\Gamma(k-\chi_j)}{k!} \left(1 + O\left(\frac{1}{n}\right)\right)$$

$$= \frac{-1}{L} \sum_{j \neq 0} e^{2j\pi \boldsymbol{i} \log_Q n^*} \sum_{k=0}^{b-1} \frac{\Gamma(k-\chi_j)}{k!} \left(1 + O\left(\frac{1}{n}\right)\right)$$

$$= \frac{-1}{L} \sum_{j \neq 0} e^{2j\pi \boldsymbol{i} \log_Q n^*} \left(-\frac{1}{\chi_j} \frac{\Gamma(b-\chi_j)}{\Gamma(b)}\right) \left(1 + O\left(\frac{1}{n}\right)\right)$$

$$= \frac{1}{L} \sum_{j \neq 0} \frac{e^{2j\pi \boldsymbol{i} \log_Q n^*}}{\chi_j} \frac{\Gamma(b-\chi_j)}{\Gamma(b)} \left(1 + O\left(\frac{1}{n}\right)\right).$$

Thus the expected number of digits occurring at least $b$ times in a word is (see (3.2) and (4.7))

$$\mathbb{E}_b(d_n) = \log_Q n + \frac{\gamma}{L} + \log_Q(Q-1) - \frac{1}{2} - \frac{1}{L} H_{b-1} + \delta_{E_b}(\log_Q n^*) + O\left(\frac{1}{n}\right),$$

as $n \to \infty$, where

$$\delta_{E_b}(x) := \frac{1}{L} \sum_{j \neq 0} \frac{e^{2j\pi \boldsymbol{i} x}}{\chi_j} \frac{\Gamma(b-\chi_j)}{\Gamma(b)}. \tag{4.9}$$

This concludes the proof of Theorem 4.1. ∎

## 4.2 The variance (general case)

We use the same formula as before, namely:

$$\mathbb{V}_b(d_n) = n![z^n] \frac{\partial^2}{\partial u^2} F_b(z,u)\Big|_{u=1} + \mathbb{E}_b(d_n) - \mathbb{E}_b^2(d_n),$$

for $F_b(z,u)$ as defined in equation (4.1). We define

$$f_i(z,u) := \sum_{k=0}^{b-1} \frac{(zpq^i)^k}{k!} + u\left(e^{zpq^i} - \sum_{k=0}^{b-1} \frac{(zpq^i)^k}{k!}\right).$$

The second factorial moment can be calculated in a similar fashion to (3.11), namely:

$$\frac{\partial^2}{\partial u^2} F_b(z,u)\Big|_{u=1}$$

$$= \frac{\partial^2}{\partial u^2} \prod_{i \geq 0} f_i(z,u)\Big|_{u=1}$$

$$= \prod_{i \geq 0} f_i(z,u) \sum_{0 \leq l < j} \frac{2\frac{\partial}{\partial u} f_l(z,u) \frac{\partial}{\partial u} f_j(z,u)}{f_l(z,u) f_j(z,u)} + \prod_{i \geq 0} f_i(z,u) \sum_j \frac{\frac{\partial^2}{\partial u^2} f_j(z,u)}{f_j(z,u)}$$

$$
\begin{aligned}
&= 2e^z \sum_{0\le l<j} \frac{\left(e^{zpq^l} - \sum_{k=0}^{b-1} \frac{(zpq^l)^k}{k!}\right)\left(e^{zpq^j} - \sum_{k=0}^{b-1} \frac{(zpq^j)^k}{k!}\right)}{e^{zpq^l}e^{zpq^j}} \\
&= 2e^z \sum_{0\le l<j} \left(1 - e^{-zpq^l} \sum_{k=0}^{b-1} \frac{(zpq^l)^k}{k!}\right)\left(1 - e^{-zpq^j} \sum_{k=0}^{b-1} \frac{(zpq^j)^k}{k!}\right) \\
&= 2 \sum_{0\le l<j} \left[e^z - e^{z(1-pq^l)} \sum_{k=0}^{b-1} \frac{(zpq^l)^k}{k!} - e^{z(1-pq^j)} \sum_{k=0}^{b-1} \frac{(zpq^j)^k}{k!}\right. \\
&\qquad\qquad \left. + e^{z(1-pq^l-zpq^j)} \sum_{k=0}^{b-1} \frac{(zpq^l)^k}{k!} \sum_{k=0}^{b-1} \frac{(zpq^j)^k}{k!}\right] \\
&= 2 \sum_{0\le l<j} \left[e^z - e^{z(1-pq^j)} \sum_{k=0}^{b-1} \frac{(zpq^j)^k}{k!}\right] \qquad (4.10) \\
&\quad + 2 \sum_{0\le l<j} \left[e^{z(1-pq^l-zpq^j)} \sum_{k=0}^{b-1} \frac{(zpq^l)^k}{k!} \sum_{k=0}^{b-1} \frac{(zpq^j)^k}{k!} - e^{z(1-pq^l)} \sum_{k=0}^{b-1} \frac{(zpq^l)^k}{k!}\right]. \qquad (4.11)
\end{aligned}
$$

**Expression (4.10)**

The expression given by (4.10) is in fact two terms which are treated together because of their mutual independence of the index $l$. Hence

$$
2 \sum_{0\le l<j} \left[e^z - e^{z(1-pq^j)} \sum_{k=0}^{b-1} \frac{(zpq^j)^k}{k!}\right] = 2 \sum_{0\le j} j \left[e^z - e^{z(1-pq^j)} \sum_{k=0}^{b-1} \frac{(zpq^j)^k}{k!}\right].
$$

We want the coefficient of $\frac{z^n}{n!}$ of these exponential generating functions, which is 1 for $e^z$, and for the other functions we look at a term for any $k$ between 0 and $b-1$ to get a coefficient of:

$$
\begin{aligned}
n![z^n]e^{z(1-pq^j)} \frac{(zpq^j)^k}{k!} &= \frac{n!(pq^j)^k}{k!}[z^n]z^k e^{z(1-pq^j)} \\
&= \frac{n!(pq^j)^k}{k!}[z^{n-k}]e^{z(1-pq^j)} \\
&= \frac{n!(pq^j)^k}{k!} \frac{(1-pq^j)^{n-k}}{(n-k)!} \\
&= \binom{n}{k} (pq^j)^k (1-pq^j)^{n-k}.
\end{aligned}
$$

The coefficient of the expanded sum is thus

$$
2 \sum_{j\ge 0} j \left[1 - (1-pq^j)^n - npq^j(1-pq^j)^{n-1} - \cdots - \binom{n}{b-1} (pq^j)^{b-1} (1-pq^j)^{n-(b-1)}\right].
$$

Now this can be split up into

$$2\sum_{j\geq 0} j\big[1-(1-pq^j)^n\big], \tag{4.12}$$

which is known from (3.17), and

$$-2\sum_{j\geq 0} j\left[npq^j(1-pq^j)^{n-1}+\cdots+\binom{n}{b-1}(pq^j)^{b-1}(1-pq^j)^{n-(b-1)}\right]. \tag{4.13}$$

A typical term in (4.13) is:

$$\begin{aligned}
-2\binom{n}{s}\sum_{j\geq 0} j(pq^j)^s(1-pq^j)^{n-s} &= -2\binom{n}{s}\sum_{j\geq 0} j(pq^j)^s\sum_{h=0}^{n-s}\binom{n-s}{h}(-1)^h(pq^j)^h\\
&= -2\binom{n}{s}\sum_{h=0}^{n-s}\binom{n-s}{h}(-1)^h\sum_{j\geq 0} j(pq^j)^{h+s}\\
&= -2\binom{n}{s}\sum_{h=0}^{n-s}\binom{n-s}{h}(-1)^h\frac{(pq)^{h+s}}{(1-q^{h+s})^2}\\
&= -2\binom{n}{s}\sum_{h=0}^{N}\binom{N}{h}(-1)^h\frac{(pq)^{h+s}}{(1-q^{h+s})^2},
\end{aligned}$$

(where $N := n-s$) for which there is a double pole at $z=-s$. Again Rice's method can be used. Let

$$f(z) := \frac{(pq)^{z+s}}{(1-q^{z+s})^2},$$

and let $\varepsilon := z+s$. Then we expand around $\varepsilon = 0$ to two terms:

$$\begin{aligned}
\frac{(pq)^\varepsilon}{(1-q^\varepsilon)^2} &= \frac{((1-Q^{-1})Q^{-1})^\varepsilon}{(1-Q^{-\varepsilon})^2}\\
&= \frac{(Q-1)^\varepsilon}{(Q^\varepsilon-1)^2}\\
&= \frac{e^{\varepsilon\log(Q-1)}}{(e^{\varepsilon\log Q}-1)^2}\\
&\sim \frac{1+\varepsilon\log(Q-1)}{(1+\varepsilon\log Q+\frac{\varepsilon^2\log^2 Q}{2}-1)^2}\\
&= \frac{1+\varepsilon\log(Q-1)}{\varepsilon^2L^2(1+\frac{\varepsilon L}{2})^2}\\
&\sim \frac{1}{\varepsilon^2L^2}\big(1+\varepsilon\log(Q-1)\big)\left(1-\frac{\varepsilon L}{2}\right)^2\\
&\sim \frac{1}{\varepsilon^2L^2}\big(1+\varepsilon\log(Q-1)-\varepsilon L\big).
\end{aligned} \tag{4.14}$$

The Taylor expansion of $[N;z]$ around $z=-s$ (i.e., around $\varepsilon = 0$) to two places (this can be done by Mathematica) is

$$[n-s;-s] \sim \frac{\Gamma(n-s+1)\Gamma(s)}{\Gamma(n+1)}\big[1+(z+s)\psi(n+1)-(z+s)\psi(s)\big]$$

$$= \frac{(n-s)!(s-1)!}{n!}\big[1+\varepsilon\,\psi(n+1)-\varepsilon\,\psi(s)\big]. \tag{4.15}$$

We can now calculate the residue by multiplying (4.14) and (4.15) and looking at the coefficient of $\varepsilon^{-1}$.

$$\begin{aligned}
&[\varepsilon^{-1}]\frac{1}{\varepsilon^2L^2}\big(1+\varepsilon\log(Q-1)-\varepsilon L\big)\frac{(n-s)!(s-1)!}{n!}\big[1+\varepsilon\,\psi(n+1)-\varepsilon\,\psi(s)\big]\\
&\quad= \frac{(n-s)!(s-1)!}{n!}[\varepsilon]\frac{1}{L^2}\big(1+\varepsilon\log(Q-1)-\varepsilon L\big)\big[1+\varepsilon\,\psi(n+1)-\varepsilon\,\psi(s)\big]\\
&\quad= \frac{(n-s)!(s-1)!}{n!}\frac{1}{L^2}\big(\log(Q-1)-L+\psi(n+1)-\psi(s)\big).
\end{aligned} \tag{4.16}$$

We also have double poles at $z+s=\chi_k$, $k\in\mathbb{Z}$, $k\neq 0$. Let $\varepsilon := z+s-\chi_k$, then using (3.6), the function $f(z)$ can be written as

$$\begin{aligned}
f(z) &= \frac{(pq)^{z+s}}{(1-q^{z+s})^2}\\
&= \frac{((1-Q^{-1})Q^{-1})^{\varepsilon+\chi_k}}{(1-Q^{-(\varepsilon+\chi_k)})^2}\\
&= \frac{Q^{2(\varepsilon+\chi_k)}Q^{-\varepsilon-\chi_k}(1-Q^{-1})^{\varepsilon+\chi_k}}{Q^{2(\varepsilon+\chi_k)}(1-Q^{-\varepsilon-\chi_k})^2}\\
&= \frac{(Q-1)^{\varepsilon+\chi_k}}{(Q^{\varepsilon+\chi_k}-1)^2}\\
&= (Q-1)^{\chi_k}\frac{(Q-1)^{\varepsilon}}{(Q^{\varepsilon}-1)^2},
\end{aligned} \tag{4.17}$$

since $Q^{\chi_k}=1$ from (3.6). Expanding the fraction to two terms, we have

$$\begin{aligned}
\frac{(Q-1)^{\varepsilon}}{(Q^{\varepsilon}-1)^2} &= \frac{e^{\varepsilon\log(Q-1)}}{(e^{\varepsilon\log Q}-1)^2}\\
&\sim \frac{1+\varepsilon\log(Q-1)}{(1+\varepsilon\log Q+\frac{\varepsilon^2\log^2 Q}{2}-1)^2}\\
&= \frac{1+\varepsilon\log(Q-1)}{(\varepsilon\log Q+\frac{\varepsilon^2\log^2 Q}{2})^2}\\
&= \frac{1+\varepsilon\log(Q-1)}{\varepsilon^2L^2(1+\frac{\varepsilon L}{2})^2}\\
&\sim \frac{1}{\varepsilon^2L^2}(1+\varepsilon\log(Q-1))\left(1-\frac{\varepsilon L}{2}\right)\left(1-\frac{\varepsilon L}{2}\right).
\end{aligned}$$

The $[N;z]$ factor ($N=n-s$) expanded to two terms around $z=\chi_k-s$ (i.e., around $\varepsilon=0$) is

$$\begin{aligned}
&[n-s;\chi_k-s]\\
&\sim \frac{\Gamma(n-s+1)\Gamma(s-\chi_k)}{\Gamma(n+1-\chi_k)}\big[1+(z+s-\chi_k)\psi(n-\chi_k+1)-(z+s-\chi_k)\psi(s-\chi_k)\big].
\end{aligned}$$

We put these together (including the factor $(Q-1)^{\chi_k}$ from (4.17)) to get the main term of the residue,

$$
\begin{aligned}
&[\varepsilon^{-1}](Q-1)^{\chi_k}\frac{1}{\varepsilon^2L^2}(1+\varepsilon\log(Q-1))\Big(1-\frac{\varepsilon L}{2}\Big)\Big(1-\frac{\varepsilon L}{2}\Big)\\
&\qquad\cdot\frac{\Gamma(n-s+1)\Gamma(s-\chi_k)}{\Gamma(n+1-\chi_k)}\big[1+\varepsilon\,\psi(n-\chi_k+1)-\varepsilon\,\psi(s-\chi_k)\big]\\
&=(Q-1)^{\chi_k}\frac{1}{L^2}\frac{\Gamma(n-s+1)\Gamma(s-\chi_k)}{\Gamma(n+1-\chi_k)}[\varepsilon](1+\varepsilon\log(Q-1))\Big(1-\frac{\varepsilon L}{2}\Big)\Big(1-\frac{\varepsilon L}{2}\Big)\\
&\qquad\qquad\qquad\qquad\cdot\big[1+\varepsilon\,\psi(n-\chi_k+1)-\varepsilon\,\psi(s-\chi_k)\big]\\
&=(Q-1)^{\chi_k}\frac{1}{L^2}\frac{\Gamma(n-s+1)\Gamma(s-\chi_k)}{\Gamma(n+1-\chi_k)}\\
&\qquad\quad\cdot\big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\big)\\
&\sim\frac{1}{L^2}\Gamma(s-\chi_k)(Q-1)^{\chi_k}n^{\chi_k-s}\big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\big)\\
&\sim\frac{1}{L^2}\Gamma(s-\chi_k)n^{-s}e^{2k\pi\boldsymbol{i}\log_Q(n(Q-1))}\big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\big),
\end{aligned}
$$

which holds for all $k\neq0$, and so the residue at each $\chi_k$ is asymptotic to

$$
\frac{1}{L^2}\Gamma(s-\chi_k)n^{-s}e^{2k\pi\boldsymbol{i}\log_Q n^*}\big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\big),
$$

(recall $n^*=n(Q-1)$), which can be summed over all $k\neq0$ to get

$$
\begin{aligned}
&\sum_{k\neq0}\frac{1}{L^2}\Gamma(s-\chi_k)n^{-s}e^{2k\pi\boldsymbol{i}\log_Q n^*}\big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\big)\\
&=n^{-s}\frac{1}{L^2}\sum_{k\neq0}\Gamma(s-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\big).
\end{aligned}
$$

This result can be combined with equation (4.16) to give the total residues for a typical term as asymptotic to

$$
\begin{aligned}
-2\binom{n}{s}\Bigg[&\frac{(n-s)!(s-1)!}{n!}\frac{1}{L^2}\big(\log(Q-1)-L+\psi(n+1)-\psi(s)\big)\\
&\quad+n^{-s}\frac{1}{L^2}\sum_{k\neq0}\Gamma(s-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\\
&\qquad\qquad\cdot\big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\big)\Bigg]\\
=&-\frac{2}{s}\frac{1}{L^2}\big(\log(Q-1)-L+\psi(n+1)-\psi(s)\big)\\
&-2\binom{n}{s}n^{-s}\frac{1}{L^2}\sum_{k\neq0}\Gamma(s-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\\
&\qquad\qquad\cdot\big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\big)\\
\sim&-\frac{2}{s}\frac{1}{L^2}\big(\log(Q-1)-L+\psi(n+1)-\psi(s)\big)
\end{aligned}
$$

$$
\begin{aligned}
&- 2\binom{n}{s}\frac{\Gamma(n-s+1)}{\Gamma(n+1)}\frac{1}{L^2}\sum_{k\neq 0}\Gamma(s-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\\
&\qquad\qquad \cdot \big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\big)\\
&= -\frac{2}{s}\frac{1}{L^2}\big(\log(Q-1)-L+\psi(n+1)-\psi(s)\big)\\
&\quad - 2\binom{n}{s}\frac{(n-s)!}{n!}\frac{1}{L^2}\sum_{k\neq 0}\Gamma(s-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\\
&\qquad\qquad \cdot \big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\\
&= -\frac{2}{sL^2}\big(\log(Q-1)-L+\psi(n+1)-\psi(s)\big)\\
&\quad - \frac{2}{s!L^2}\sum_{k\neq 0}\Gamma(s-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\\
&\qquad\qquad \cdot \big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\big).
\end{aligned}
$$

Since we have $b-1$ of these terms added together, we can now sum them to get (from (4.13))

$$
\begin{aligned}
&\sum_{s=1}^{b-1}\Bigg[-\frac{2}{L^2 s}\big(\log(Q-1)-L+\psi(n+1)-\psi(s)\big)\\
&\quad - \frac{2}{s!L^2}\sum_{k\neq 0}\Gamma(s-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\\
&\qquad\qquad \cdot \big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\big)\Bigg]\\
&= -\frac{2}{L^2}\sum_{s=1}^{b-1}\frac{1}{s}\big(\log(Q-1)-L+\psi(n+1)-\psi(s)\big)\\
&\quad - \frac{2}{L^2}\sum_{s=1}^{b-1}\frac{1}{s!}\sum_{k\neq 0}\Gamma(s-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\\
&\qquad\qquad \cdot \big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\big)\\
&= -\frac{2}{L^2}\sum_{s=1}^{b-1}\frac{1}{s}\big(\log(Q-1)-L+\psi(n+1)\big)+\frac{2}{L^2}\sum_{s=1}^{b-1}\frac{1}{s}\psi(s)\\
&\quad - \frac{2}{L^2}\sum_{s=1}^{b-1}\frac{1}{s!}\sum_{k\neq 0}\Gamma(s-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\\
&\qquad\qquad \cdot \big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\big)\\
&= -\frac{2}{L^2}\big(\log(Q-1)-L+\psi(n+1)\big)H_{b-1}+\frac{2}{L^2}\sum_{s=1}^{b-1}\frac{\psi(s)}{s}\\
&\quad - \frac{2}{L^2}\sum_{s=1}^{b-1}\frac{1}{s!}\sum_{k\neq 0}\Gamma(s-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\\
&\qquad\qquad \cdot \big(\log(Q-1)-L+\psi(n-\chi_k+1)-\psi(s-\chi_k)\big).
\end{aligned}
$$

Expression (4.12) was dealt with in the classical variance discussion (Part (a)), where it was shown that

$$2\sum_{j\geq 0} j\big[1-(1-pq^j)^n\big] = 2\sum_{i=1}^{n}\binom{n}{i}(-1)^i\frac{-(Q-1)^i}{(Q^i-1)^2}$$

whose residue is

$$\begin{aligned}
&\log_Q^2(Q-1)+\frac{5}{6}+\log_Q^2 n+\frac{2\gamma}{L}\log_Q n+\frac{\gamma^2}{L^2}+\frac{\pi^2}{6L^2}-2\log_Q(Q-1)\\
&+2\log_Q(Q-1)\log_Q n+\frac{2\gamma}{L}\log_Q(Q-1)-2\log_Q n-\frac{2\gamma}{L}\\
&+\frac{2}{L}\sum_{k\neq 0}\Gamma(-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\Big[-\log_Q n+\frac{\psi(-\chi_k)}{L}-\log_Q(Q-1)+1\Big]+o(1).
\end{aligned}$$

Since $\psi(n-\chi_k+1)\sim\log n$, we have that the total residue for (4.10) is

$$\begin{aligned}
&-\frac{2}{L}\big(\log_Q(Q-1)-1+\log_Q n\big)H_{b-1}+\frac{2}{L^2}\sum_{s=1}^{b-1}\frac{\psi(s)}{s}\\
&\quad+\log_Q^2(Q-1)+\frac{5}{6}+\log_Q^2 n+\frac{2\gamma}{L}\log_Q n+\frac{\gamma^2}{L^2}+\frac{\pi^2}{6L^2}-2\log_Q(Q-1)\\
&\quad+2\log_Q(Q-1)\log_Q n+\frac{2\gamma}{L}\log_Q(Q-1)-2\log_Q n-\frac{2\gamma}{L}\\
&\quad\underbrace{-\frac{2}{L}\sum_{k\neq 0}\Gamma(-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\log_Q n}_{\flat}+\frac{2}{L}\sum_{k\neq 0}\Gamma(-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\frac{\psi(-\chi_k)}{L}\\
&\quad-\frac{2}{L}\sum_{k\neq 0}\Gamma(-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\log_Q(Q-1)+\frac{2}{L}\sum_{k\neq 0}\Gamma(-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\\
&\quad-\frac{2}{L}\sum_{s=1}^{b-1}\frac{1}{s!}\sum_{k\neq 0}\Gamma(s-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\log_Q(Q-1)\\
&\quad+\frac{2}{L}\sum_{s=1}^{b-1}\frac{1}{s!}\sum_{k\neq 0}\Gamma(s-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\underbrace{-\frac{2}{L}\sum_{s=1}^{b-1}\frac{1}{s!}\sum_{k\neq 0}\Gamma(s-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\log_Q n}_{\sharp}\\
&\quad+\frac{2}{L}\sum_{s=1}^{b-1}\frac{1}{s!}\sum_{k\neq 0}\Gamma(s-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n^*}\frac{\psi(s-\chi_k)}{L}+o(1).
\end{aligned}$$

Although it is not clear now why the following changes are made, it will become apparent when cancelling terms in the variance. We take terms $\flat$ and $\sharp$ from the above formula, and use [12] get

$$\begin{aligned}
\flat+\sharp&=\frac{-2}{L}\log_Q n\sum_{k\neq 0}e^{2k\pi\boldsymbol{i}\log_Q n^*}\Big(\Gamma(-\chi_k)+\sum_{s=1}^{b-1}\frac{1}{s!}\Gamma(s-\chi_k)\Big)\\
&=\frac{-2}{L}\log_Q n\sum_{k\neq 0}e^{2k\pi\boldsymbol{i}\log_Q n^*}\sum_{s=0}^{b-1}\frac{\Gamma(s-\chi_k)}{s!}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{-2}{L}\log_Q n \sum_{k\neq 0} e^{2k\pi \boldsymbol{i}\log_Q n^*} \sum_{s=0}^{b-1} \frac{\Gamma(s+a)}{s!} \qquad (\text{let } a := -\chi_k)\\
&= \frac{-2}{L}\log_Q n \sum_{k\neq 0} e^{2k\pi \boldsymbol{i}\log_Q n^*} \sum_{s=0}^{b-1} \frac{(s+a-1)!}{s!}\\
&= \frac{-2\,(a-1)!}{L}\log_Q n \sum_{k\neq 0} e^{2k\pi \boldsymbol{i}\log_Q n^*} \sum_{s=0}^{b-1} \binom{s+a-1}{a-1}\\
&= \frac{-2\,(a-1)!}{L}\log_Q n \sum_{k\neq 0} e^{2k\pi \boldsymbol{i}\log_Q n^*} \binom{b+a-1}{a}\\
&= \frac{-2}{L}\log_Q n \sum_{k\neq 0} e^{2k\pi \boldsymbol{i}\log_Q n^*} \frac{(b+a-1)!}{a\,(b-1)!}\\
&= \frac{-2}{L}\log_Q n \sum_{k\neq 0} e^{2k\pi \boldsymbol{i}\log_Q n^*} \frac{\Gamma(b+a)}{a\,\Gamma(b)}\\
&= \frac{2}{L}\log_Q n \sum_{k\neq 0} e^{2k\pi \boldsymbol{i}\log_Q n^*} \frac{\Gamma(b-\chi_k)}{\chi_k\,\Gamma(b)}\\
&= 2\log_Q n\delta_{E_b}(x).
\end{aligned}
$$

Now the final residue for expression (4.10) is

$$
\begin{aligned}
&-\frac{2}{L}\big(\log_Q(Q-1) - 1 + \log_Q n\big) H_{b-1} + \frac{2}{L^2}\sum_{s=1}^{b-1}\frac{\psi(s)}{s}\\
&+\log_Q^2(Q-1) + \frac{5}{6} + \log_Q^2 n + \frac{2\gamma}{L}\log_Q n + \frac{\gamma^2}{L^2} + \frac{\pi^2}{6L^2} - 2\log_Q(Q-1)\\
&+2\log_Q(Q-1)\log_Q n + \frac{2\gamma}{L}\log_Q(Q-1) - 2\log_Q n - \frac{2\gamma}{L} + 2\log_Q n\delta_{E_b}(x)\\
&+\frac{2}{L}\sum_{k\neq 0}\Gamma(-\chi_k)e^{2k\pi \boldsymbol{i}\log_Q n^*}\Big[\frac{\psi(-\chi_k)}{L} - \log_Q(Q-1) + 1\Big] \qquad (4.18)\\
&-\frac{2}{L}\sum_{s=1}^{b-1}\frac{1}{s!}\sum_{k\neq 0}\Gamma(s-\chi_k)e^{2k\pi \boldsymbol{i}\log_Q n^*}\bigg(\log_Q(Q-1) - 1 - \frac{\psi(s-\chi_k)}{L}\bigg) + o(1).
\end{aligned}
$$

**Expression (4.11)**

We now turn our attention to the other portion of the second factorial moment, namely the expression in (4.11). We have

$$
\begin{aligned}
\Upsilon &:= 2\sum_{0\le l<j}\Bigg[e^{z(1-pq^l-zpq^j)}\sum_{k=0}^{b-1}\frac{(zpq^l)^k}{k!}\sum_{k=0}^{b-1}\frac{(zpq^j)^k}{k!} - e^{z(1-pq^l)}\sum_{k=0}^{b-1}\frac{(zpq^l)^k}{k!}\Bigg]\\
&= 2\sum_{0\le l<j}\Bigg[e^{z(1-pq^l-pq^j)}\bigg(1+\cdots+\frac{(zpq^l)^{b-1}}{(b-1)!}\bigg)\underbrace{\bigg(1+\cdots+\frac{(zpq^j)^{b-1}}{(b-1)!}\bigg)}
\end{aligned}
$$

$$- e^{z(1-pq^l)}\left(1+\cdots+\frac{(zpq^l)^{b-1}}{(b-1)!}\right)\Bigg].$$

The bracketed factor is split up into two parts. We write

$$\begin{aligned}
\Upsilon &= 2\sum_{0\le l<j}\Bigg[e^{z(1-pq^l-pq^j)}\left(1+\cdots+\frac{(zpq^l)^{b-1}}{(b-1)!}\right) - e^{z(1-pq^l)}\left(1+\cdots+\frac{(zpq^l)^{b-1}}{(b-1)!}\right)\\
&\qquad + e^{z(1-pq^l-pq^j)}\left(1+\cdots+\frac{(zpq^l)^{b-1}}{(b-1)!}\right)\left(zpq^j+\cdots+\frac{(zpq^j)^{b-1}}{(b-1)!}\right)\Bigg]\\
&= 2\sum_{0\le l<j}\Bigg[\left(1+\cdots+\frac{(zpq^l)^{b-1}}{(b-1)!}\right)\left(e^{z(1-pq^l-pq^j)}-e^{z(1-pq^l)}\right)\\
&\qquad + e^{z(1-pq^l-pq^j)}\left(1+\cdots+\frac{(zpq^l)^{b-1}}{(b-1)!}\right)\left(zpq^j+\cdots+\frac{(zpq^j)^{b-1}}{(b-1)!}\right)\Bigg].
\end{aligned}$$

We now call the first sum $P$ and the second sum $R$. Thus

$$P := 2\sum_{0\le l<j}\left[\left(1+\cdots+\frac{(zpq^l)^{b-1}}{(b-1)!}\right)\left(e^{z(1-pq^l-pq^j)}-e^{z(1-pq^l)}\right)\right], \tag{4.19}$$

and

$$R := 2\sum_{0\le l<j}\left[e^{z(1-pq^l-pq^j)}\left(1+\cdots+\frac{(zpq^l)^{b-1}}{(b-1)!}\right)\left(zpq^j+\cdots+\frac{(zpq^j)^{b-1}}{(b-1)!}\right)\right]. \tag{4.20}$$

**Dealing with $P$**

Let $P_s$ be a typical term of (4.19), where $s\in\{0,1,\dots,b-1\}$.

$$P_s := 2\sum_{0\le l<j}\left[\frac{(zpq^l)^s}{s!}\left(e^{z(1-pq^l-pq^j)}-e^{z(1-pq^l)}\right)\right].$$

In this way we simplify the expression and can extract coefficients more easily. We now look at the coefficients of this expression (which is an exponential generating function, hence the $n!$).

$$\begin{aligned}
n![z^n]P_s &= n![z^n]2\sum_{0\le l<j}\frac{z^s(pq^l)^s}{s!}\left(e^{z(1-pq^l-pq^j)}-e^{z(1-pq^l)}\right)\\
&= 2\sum_{0\le l<j} n![z^n]\frac{z^s(pq^l)^s}{s!}\sum_{k\ge 0}\left(\frac{z^k(1-pq^l-pq^j)^k}{k!}-\frac{z^k(1-pq^l)^k}{k!}\right)\\
&= 2\sum_{0\le l<j} n![z^n]\sum_{k\ge 0}\frac{z^{k+s}}{k!s!}(pq^l)^s\left((1-pq^l-pq^j)^k-(1-pq^l)^k\right)\\
&= 2\sum_{0\le l<j} n![z^n]\sum_{n\ge s}\frac{z^n}{(n-s)!s!}(pq^l)^s\left((1-pq^l-pq^j)^{n-s}-(1-pq^l)^{n-s}\right)\\
&= 2\sum_{0\le l<j}\frac{n!}{s!(n-s)!}(pq^l)^s\left((1-pq^l-pq^j)^{n-s}-(1-pq^l)^{n-s}\right), \quad \text{for } n\ge s
\end{aligned}$$

$$
\begin{aligned}
&= 2\sum_{0\le l<j}\binom{n}{s}(pq^l)^s\big((1-pq^l-pq^j)^{n-s}-(1-pq^l)^{n-s}\big)\\
&= 2\sum_{0\le l<j}\binom{n}{s}(pq^l)^s\left[\sum_{k=0}^{n-s}\binom{n-s}{k}(-pq^l-pq^j)^k-\sum_{k=0}^{n-s}\binom{n-s}{k}(-pq^l)^k\right]\\
&= 2\sum_{0\le l<j}\binom{n}{s}(pq^l)^s\sum_{k=0}^{n-s}\binom{n-s}{k}(-1)^k\big((pq^l+pq^j)^k-(pq^l)^k\big)\\
&= 2\binom{n}{s}\sum_{k=0}^{n-s}\binom{n-s}{k}(-1)^k\sum_{0\le l<j}(pq^l)^s(pq^l)^k\big((1+q^{j-l})^k-1\big)\\
&= 2\binom{n}{s}\sum_{k=0}^{n-s}\binom{n-s}{k}(-1)^k\sum_{l\ge0}(pq^l)^s(pq^l)^k\sum_{h\ge1}\big((1+q^h)^k-1\big)\quad(h:=j-l)\\
&= 2\binom{n}{s}\sum_{k=0}^{n-s}\binom{n-s}{k}(-1)^k\sum_{l\ge0}p^{s+k}q^{(s+k)l}\sum_{h\ge1}\big((1+q^h)^k-1\big)\\
&= 2\binom{n}{s}\sum_{k=0}^{n-s}\binom{n-s}{k}(-1)^k\frac{p^{s+k}}{1-q^{s+k}}\sum_{h\ge1}\big((1+q^h)^k-1\big)\\
&= 2\binom{n}{s}\sum_{k=0}^{N}\binom{N}{k}(-1)^k\frac{p^{s+k}}{1-q^{s+k}}\sum_{h\ge1}\big((1+q^h)^k-1\big),
\end{aligned}
\tag{4.21}
$$

where $N := n-s$. We can now use Rice's method. If

$$H(k) := \sum_{h\ge1}\big((1+q^h)^k-1\big),$$

then $H(z)$ has no poles, so we only need to look at the fraction,

$$\frac{p^{s+k}}{1-q^{s+k}} = \frac{(1-q)^{s+k}}{1-q^{s+k}} = \frac{(1-Q^{-1})^{s+k}}{1-Q^{-(s+k)}} = \frac{(Q-1)^{s+k}}{Q^{s+k}-1},$$

which is the same fraction as that in the expected value for the number of values occurring at least $b$ times (see definition (4.5)). We thus use the expansions from that calculation for the poles at $z=-s$ and $z=-s+\chi_k$ (here we use variable $s$ instead of $k$). In the first case we had (in (3.7))

$$f_s(z) := \frac{(Q-1)^{s+z}}{Q^{s+z}-1} \sim \frac{1}{L(s+z)},$$

making the residue $\frac{1}{L}$. The (exact) contribution of quantity $[N;z]$ around $z=-s$ was also calculated above as being (see equation (4.6))

$$[n-s;-s] = \frac{(n-s)!(s-1)!}{n!}.$$

We must also calculate the contribution of the new quantity $H(z)$.

$$H(z) = \sum_{h\ge1}\big((1+q^h)^z-1\big),$$

so

$$
\begin{aligned}
H(-s) &= \sum_{h\geq 1}\left((1+q^h)^{-s}-1\right)\\
&= \sum_{h\geq 1}\left(\frac{1}{(1+q^h)^s}-1\right)\\
&= \sum_{h\geq 1}\left(\sum_{i\geq 0}\binom{i+s-1}{i}(-q^h)^i-1\right)\\
&= \sum_{h\geq 1}\sum_{i\geq 1}\binom{i+s-1}{i}(-q^h)^i\\
&= \sum_{i\geq 1}\binom{i+s-1}{i}(-1)^i\sum_{h\geq 1}(q^i)^h\\
&= \sum_{i\geq 1}\binom{i+s-1}{i}(-1)^i\frac{q^i}{1-q^i}\\
&= \sum_{i\geq 1}\binom{i+s-1}{i}(-1)^i\frac{1}{Q^i-1}.
\end{aligned}
$$

The total residue from the pole at $z=-s$ is thus

$$
\frac{1}{L}\frac{(n-s)!(s-1)!}{n!}H(-s),
$$

and by substituting this back into the expression for the coefficients of $P_s$ (see (4.21)), and summing on $s$, we get

$$
\begin{aligned}
&\sum_{s=0}^{b-1}2\binom{n}{s}\frac{1}{L}\frac{(n-s)!(s-1)!}{n!}H(-s)\\
&\quad= \sum_{s=0}^{b-1}2\binom{n}{s}\frac{1}{L}\frac{(n-s)!(s-1)!}{n!}\sum_{i\geq 1}\binom{i+s-1}{i}(-1)^i\frac{1}{Q^i-1}\\
&\quad= \frac{2}{L}\sum_{i\geq 1}(-1)^i\frac{1}{Q^i-1}\sum_{s=0}^{b-1}\binom{n}{s}\frac{(n-s)!(s-1)!}{n!}\binom{i+s-1}{i}\\
&\quad= \frac{2}{L}\sum_{i\geq 1}(-1)^i\frac{1}{Q^i-1}\sum_{s=0}^{b-1}\frac{1}{i}\binom{i+s-1}{i-1}\\
&\quad= \frac{2}{L}\sum_{i\geq 1}(-1)^i\frac{1}{Q^i-1}\frac{1}{i}\binom{i+b-1}{i}\\
&\quad= \frac{2}{L}\sum_{i\geq 1}\frac{(-1)^i}{i(Q^i-1)}\binom{i+b-1}{i}. \qquad (4.22)
\end{aligned}
$$

For the poles occurring at $z=\chi_k-s$ we calculated in equations (3.6) and (3.7) that for $\varepsilon := z+s-\chi_k$ we have

$$
f_s(z)\sim (Q-1)^{\chi_k}\frac{1}{\varepsilon L},
$$

and $[N; z]$ around $z = \chi_k - s$ was also dealt with above in (4.8),

$$[n-s; \chi_k - s] \sim \Gamma(s-\chi_k) n^{\chi_k - s}.$$

Again we need to calculate the contribution of the new quantity $H(z)$,

$$\begin{aligned}
H(\chi_k - s) &= \sum_{h\ge 1} \left( (1+q^h)^{\chi_k - s} - 1 \right) \\
&= \sum_{h\ge 1} \left( \frac{1}{(1+q^h)^{s-\chi_k}} - 1 \right) \\
&= \sum_{h\ge 1} \left( \sum_{i\ge 0} \binom{i+s-\chi_k-1}{i} (-q^h)^i - 1 \right) \\
&= \sum_{h\ge 1} \sum_{i\ge 1} \binom{i+s-\chi_k-1}{i} (-q^h)^i \\
&= \sum_{i\ge 1} \binom{i+s-\chi_k-1}{i} (-1)^i \sum_{h\ge 1} (q^i)^h \\
&= \sum_{i\ge 1} \binom{i+s-\chi_k-1}{i} (-1)^i \frac{1}{Q^i - 1},
\end{aligned}$$

to get the fluctuating residues

$$\begin{aligned}
&\sum_{k\ne 0} (Q-1)^{\chi_k} \frac{1}{L} \Gamma(s-\chi_k) n^{\chi_k - s} H(\chi_k - s) \\
&\qquad = n^{-s} \frac{1}{L} \sum_{k\ne 0} (Q-1)^{\chi_k} n^{\chi_k} \Gamma(s-\chi_k) H(\chi_k - s) \\
&\qquad = n^{-s} \frac{1}{L} \sum_{k\ne 0} e^{2k\pi \boldsymbol{i} \log_Q n^*} \Gamma(s-\chi_k) H(\chi_k - s).
\end{aligned}$$

These can also be substituted into the expression for the coefficients of $P_s$ (see (4.21)) and summed to get

$$\begin{aligned}
&\sum_{s=0}^{b-1} 2 \binom{n}{s} n^{-s} \frac{1}{L} \sum_{k\ne 0} e^{2k\pi \boldsymbol{i} \log_Q n^*} \Gamma(s-\chi_k) H(\chi_k - s) \\
&\quad \sim \frac{2}{L} \sum_{s=0}^{b-1} \binom{n}{s} \frac{\Gamma(n-s+1)}{\Gamma(n+1)} \sum_{k\ne 0} e^{2k\pi \boldsymbol{i} \log_Q n^*} \Gamma(s-\chi_k) H(\chi_k - s) \\
&\quad = \frac{2}{L} \sum_{s=0}^{b-1} \frac{1}{s!} \sum_{k\ne 0} e^{2k\pi \boldsymbol{i} \log_Q n^*} \Gamma(s-\chi_k) H(\chi_k - s). \qquad (4.23)
\end{aligned}$$

Altogether we have that the coefficient for the quantity $P$ is (from (4.22) and (4.23))

$$n![z^n]P = \frac{2}{L} \sum_{i\ge 1} \frac{(-1)^i}{i(Q^i - 1)} \binom{i+b-1}{i}$$

$$+\frac{2}{L}\sum_{s=0}^{b-1}\frac{1}{s!}\sum_{k\neq 0}e^{2k\pi\boldsymbol{i}\log_Q n^*}\Gamma(s-\chi_k)H(\chi_k-s)+o(1). \tag{4.24}$$

**Dealing with $R$**

Previously, see (4.20), $R$ was defined to be

$$R=2\sum_{0\le l<j}\left[e^{z(1-pq^l-pq^j)}\Big(1+\cdots+\frac{(zpq^l)^{b-1}}{(b-1)!}\Big)\Big(zpq^j+\cdots+\frac{(zpq^j)^{b-1}}{(b-1)!}\Big)\right].$$

A typical term is

$$R_{st}:=2\sum_{0\le l<j}e^{z(1-pq^l-pq^j)}\frac{(zpq^l)^s}{s!}\frac{(zpq^j)^t}{t!},$$

with $0\le s\le b-1$ and $1\le t\le b-1$. We follow the same procedure as for $P$,

$$\begin{aligned}
n![z^n]R_{st}&=n![z^n]2\sum_{0\le l<j}e^{z(1-pq^l-pq^j)}\frac{(zpq^l)^s}{s!}\frac{(zpq^j)^t}{t!}\\
&=n![z^n]2\sum_{0\le l<j}\sum_{k\ge 0}\frac{z^k(1-pq^l-pq^j)^k}{k!}\frac{(zpq^l)^s}{s!}\frac{(zpq^j)^t}{t!}\\
&=2\sum_{0\le l<j}\sum_{k\ge 0}n![z^n]\frac{z^k(1-pq^l-pq^j)^k}{k!}\frac{z^s(pq^l)^s}{s!}\frac{z^t(pq^j)^t}{t!}\\
&=2\sum_{0\le l<j}\sum_{k\ge 0}\frac{n!}{k!s!t!}[z^n]z^{k+s+t}(1-pq^l-pq^j)^k(pq^l)^s(pq^j)^t\\
&=2\sum_{0\le l<j}\frac{n!}{(n-s-t)!s!t!}(1-pq^l-pq^j)^{n-s-t}(pq^l)^s(pq^j)^t\\
&=2\sum_{0\le l<j}\frac{n!}{(n-s-t)!s!t!}(pq^l)^s(pq^j)^t\sum_{k=0}^{n-s-t}\binom{n-s-t}{k}(-pq^l-pq^j)^k\\
&=2\frac{n!}{(n-s-t)!s!t!}\sum_{k=0}^{n-s-t}\binom{n-s-t}{k}(-1)^kp^{k+s+t}\sum_{0\le l<j}q^{ls+jt+lk}(1+q^{j-l})^k\\
&=2\frac{n!}{(n-s-t)!s!t!}\sum_{k=0}^{n-s-t}\binom{n-s-t}{k}(-1)^kp^{k+s+t}\sum_{l\ge 0}q^{l(s+t+k)}\sum_{h\ge 1}q^{ht}(1+q^h)^k\\
&=2\frac{n!}{(n-s-t)!s!t!}\sum_{k=0}^{N}\binom{N}{k}(-1)^k\frac{p^{k+s+t}}{1-q^{k+s+t}}\sum_{h\ge 1}q^{ht}(1+q^h)^k,
\end{aligned}\tag{4.25}$$

for $h:=j-l$ and $N:=n-s-t$ and we define

$$H_t(k):=\sum_{h\ge 1}q^{ht}(1+q^h)^k$$

before continuing by collecting residues for Rice's method. The function $H_t(z)$ has no poles, so we only need to consider the poles contributed by the function

$$f(z) := \frac{p^{z+s+t}}{1-q^{z+s+t}} = \frac{(Q-1)^{z+s+t}}{Q^{z+s+t}-1},$$

which has simple poles at $z+s+t=0$ and at $z+s+t=\chi_k$ for all $k \neq 0$. Expanding around the poles at $z=-s-t$ gives a residue of $\frac{1}{L}$ (computed in (3.7)). The kernel $[N; z]$ can be calculated exactly at $z=-s-t$ as (refer to (4.6))

$$\begin{aligned}
[n-s-t; -s-t] &= \frac{(-1)^{n-s-t-1}(n-s-t)!}{(-s-t)(-s-t-1)\cdots(-s-t-(n-s-t))} \\
&= \frac{(n-s-t)!}{(s+t)(s+t+1)\cdots(n)} \\
&= \frac{(n-s-t)!(s+t-1)!}{n!}.
\end{aligned}$$

The function $H_t(z)$ evaluated at $z=-s-t$ is (see [36, page 83])

$$\begin{aligned}
H_t(-s-t) &= \sum_{h\geq 1} q^{ht}(1+q^h)^{-s-t} \\
&= \sum_{h\geq 1} q^{ht}\frac{1}{(1+q^h)^{s+t}} \\
&= \sum_{h\geq 1} q^{ht}\sum_{i\geq 0}\binom{i+s+t-1}{i}(-q^h)^i \\
&= \sum_{i\geq 0}\binom{i+s+t-1}{i}(-1)^i\sum_{h\geq 1} q^{(t+i)h} \\
&= \sum_{i\geq 0}\binom{i+s+t-1}{i}(-1)^i\frac{q^{t+i}}{1-q^{t+i}} \\
&= \sum_{i\geq 0}\binom{i+s+t-1}{i}(-1)^i\frac{1}{Q^{t+i}-1}.
\end{aligned}$$

This means that the residue from the pole at $z=-s-t$ is

$$\frac{(n-s-t)!(s+t-1)!}{Ln!}H_t(-s-t).$$

This can be substituted into the expression for the coefficients of $R_{st}$ (from (4.25)) and summed on $s$ and $t$ to get

$$\begin{aligned}
&\sum_{s=0}^{b-1}\sum_{t=1}^{b-1} 2\frac{n!}{(n-s-t)!s!\,t!}\frac{(n-s-t)!(s+t-1)!}{Ln!}H_t(-s-t) \\
&\quad = \frac{2}{L}\sum_{s=0}^{b-1}\sum_{t=1}^{b-1}\frac{(s+t-1)!}{s!\,t!}H_t(-s-t)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{2}{L}\sum_{s=0}^{b-1}\sum_{t=1}^{b-1}\frac{(s+t-1)!}{s!\,t!}\sum_{i\ge 0}\binom{i+s+t-1}{i}(-1)^i\frac{1}{Q^{t+i}-1}\\
&= \frac{2}{L}\sum_{i\ge 0}(-1)^i\sum_{t=1}^{b-1}\frac{1}{Q^{t+i}-1}\sum_{s=0}^{b-1}\frac{(s+t-1)!}{s!\,t!}\binom{i+s+t-1}{i}\\
&= \frac{2}{L}\sum_{i\ge 0}(-1)^i\sum_{t=1}^{b-1}\frac{1}{Q^{t+i}-1}\sum_{s=0}^{b-1}\frac{(i+s+t-1)!}{i!\,s!\,t!}\\
&= \frac{2}{L}\sum_{i\ge 0}(-1)^i\sum_{t=1}^{b-1}\frac{1}{Q^{t+i}-1}\frac{(i+t-1)!}{t!\,i!}\sum_{s=0}^{b-1}\binom{i+s+t-1}{s}\\
&= \frac{2}{L}\sum_{i\ge 0}(-1)^i\sum_{t=1}^{b-1}\frac{1}{Q^{t+i}-1}\frac{(i+t-1)!}{t!\,i!}\binom{i+b+t-1}{t+i}. \qquad (4.26)
\end{aligned}
$$

By letting $\varepsilon := z+s+t-\chi_k$, we expand $f(z) = \frac{(Q-1)^{z+s+t}}{Q^{z+s+t}-1}$ around $\varepsilon = 0$ as in (3.6) and (3.7) to get an approximation of

$$f(z) \sim \frac{1}{\varepsilon L}(Q-1)^{\chi_k}.$$

The quantity $[N; z]$ can be calculated at $z = \chi_k - s - t$ as (see (4.8))

$$[n-s-t; \chi_k - s - t] \sim \Gamma(-\chi_k + s + t) n^{\chi_k - s - t}.$$

The function $H_t(z)$ evaluated at $z = \chi_k - s - t$ is

$$
\begin{aligned}
H_t(\chi_k - s - t) &= \sum_{h\ge 1} q^{ht}\frac{1}{(1+q^h)^{s+t-\chi_k}}\\
&= \sum_{h\ge 1} q^{ht}\sum_{i\ge 0}\binom{i+s+t-\chi_k-1}{i}(-q^h)^i\\
&= \sum_{i\ge 0}\binom{i+s+t-\chi_k-1}{i}(-1)^i\sum_{h\ge 1}q^{(t+i)h}\\
&= \sum_{i\ge 0}\binom{i+s+t-\chi_k-1}{i}(-1)^i\frac{1}{Q^{t+i}-1}.
\end{aligned}
$$

Putting these quantities together, we get the main term of the fluctuating residues from the poles at each $z = \chi_k - s - t$ to be

$$
\begin{aligned}
&[\varepsilon^{-1}](Q-1)^{\chi_k}\frac{1}{\varepsilon L}\Gamma(-\chi_k+s+t)n^{\chi_k-s-t}H_t(\chi_k-s-t)\\
&\quad = \frac{1}{L}(Q-1)^{\chi_k}n^{\chi_k}\Gamma(-\chi_k+s+t)n^{-s-t}H_t(\chi_k-s-t)\\
&\quad = \frac{1}{L}e^{2k\pi \boldsymbol{i}\log_Q n^*}\Gamma(-\chi_k+s+t)n^{-s-t}H_t(\chi_k-s-t),
\end{aligned}
$$

which means that the residue fluctuations are given asymptotically by

$$\frac{1}{L}\sum_{k\ne 0} e^{2k\pi \boldsymbol{i}\log_Q n^*}\Gamma(-\chi_k+s+t)n^{-s-t}H_t(\chi_k-s-t).$$

We now substitute this quantity into our expression for $n![z^n]R_{st}$ (from (4.25)) and sum on $s$ and $t$ to get

$$
\begin{aligned}
&\sum_{s=0}^{b-1}\sum_{t=1}^{b-1} 2\frac{n!}{(n-s-t)!s!t!}\frac{1}{L}\sum_{k\neq 0} e^{2k\pi \boldsymbol{i}\log_Q n^*}\Gamma(-\chi_k+s+t)n^{-s-t}H_t(\chi_k-s-t)\\
&\quad=\frac{2}{L}\sum_{s=0}^{b-1}\sum_{t=1}^{b-1}\frac{n!}{(n-s-t)!s!t!}n^{-s-t}\sum_{k\neq 0} e^{2k\pi \boldsymbol{i}\log_Q n^*}\Gamma(-\chi_k+s+t)H_t(\chi_k-s-t)\\
&\quad\sim\frac{2}{L}\sum_{s=0}^{b-1}\sum_{t=1}^{b-1}\frac{n!}{(n-s-t)!s!t!}\frac{\Gamma(n-s-t+1)}{\Gamma(n+1)}\\
&\qquad\qquad\qquad\cdot\sum_{k\neq 0} e^{2k\pi \boldsymbol{i}\log_Q n^*}\Gamma(-\chi_k+s+t)H_t(\chi_k-s-t)\\
&\quad=\frac{2}{L}\sum_{s=0}^{b-1}\sum_{t=1}^{b-1}\frac{n!}{(n-s-t)!s!t!}\frac{(n-s-t)!}{n!}\\
&\qquad\qquad\qquad\cdot\sum_{k\neq 0} e^{2k\pi \boldsymbol{i}\log_Q n^*}\Gamma(-\chi_k+s+t)H_t(\chi_k-s-t)\\
&\quad=\frac{2}{L}\sum_{s=0}^{b-1}\sum_{t=1}^{b-1}\frac{1}{s!\,t!}\sum_{k\neq 0} e^{2k\pi \boldsymbol{i}\log_Q n^*}\Gamma(-\chi_k+s+t)H_t(\chi_k-s-t). \qquad (4.27)
\end{aligned}
$$

Then (4.26) and (4.27) give the coefficient of $R$ asymptotically as $n\to\infty$:

$$
\begin{aligned}
n![z^n]R &= \frac{2}{L}\sum_{i\geq 0}(-1)^i\sum_{t=1}^{b-1}\frac{1}{Q^{t+i}-1}\frac{(i+t-1)!}{t!\,i!}\binom{i+b+t-1}{t+i} \qquad (4.28)\\
&\quad+\frac{2}{L}\sum_{s=0}^{b-1}\sum_{t=1}^{b-1}\frac{1}{s!\,t!}\sum_{k\neq 0} e^{2k\pi \boldsymbol{i}\log_Q n^*}\Gamma(-\chi_k+s+t)H_t(\chi_k-s-t)+o(1).
\end{aligned}
$$

The variance in the general case can thus be obtained by combining these results in the following way: (4.10) + (4.24) + (4.28) + (4.2) − $(4.2)^2$. We would expect the main term of the variance to be a constant with some small fluctuations which can be written as a delta function. The definition of $\delta_{E_b}(x)$ can be found in (4.9).

$$
\begin{aligned}
\mathbb{V}_b(d_n) &= n![z^n]\frac{\partial^2}{\partial u^2}F_b(z,u)\Big|_{u=1}+\mathbb{E}_b(d_n)-\mathbb{E}_b^2(d_n)\\
&= -\frac{2}{L}\big(\log_Q(Q-1)-1+\log_Q n\big)H_{b-1}+\frac{2}{L^2}\sum_{s=1}^{b-1}\frac{\psi(s)}{s}\\
&\quad+\log_Q^2(Q-1)+\frac{5}{6}+\log_Q^2 n+\frac{2\gamma}{L}\log_Q n+\frac{\gamma^2}{L^2}+\frac{\pi^2}{6L^2}-2\log_Q(Q-1)\\
&\quad+2\log_Q(Q-1)\log_Q n+\frac{2\gamma}{L}\log_Q(Q-1)-2\log_Q n-\frac{2\gamma}{L}\\
&\quad+2\log_Q n\,\delta_{E_b}(\log_Q n^*)+\frac{2}{L}\sum_{k\neq 0}\Gamma(-\chi_k)e^{2k\pi \boldsymbol{i}\log_Q n^*}\Big[\frac{\psi(-\chi_k)}{L}-\log_Q(Q-1)+1\Big]
\end{aligned}
$$

$$
\begin{aligned}
&- \frac{2}{L} \sum_{s=1}^{b-1} \frac{1}{s!} \sum_{k \neq 0} \Gamma(s-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q n^*} \left( \log_Q(Q-1) - 1 - \frac{\psi(s-\chi_k)}{L} \right) \\
&+ \frac{2}{L} \sum_{i \geq 1} \frac{(-1)^i}{i(Q^i-1)} \binom{i+b-1}{i} + \frac{2}{L} \sum_{s=0}^{b-1} \frac{1}{s!} \sum_{k \neq 0} e^{2k\pi \boldsymbol{i} \log_Q n^*} \Gamma(s-\chi_k) H(\chi_k - s) \\
&+ \frac{2}{L} \sum_{i \geq 0} (-1)^i \sum_{t=1}^{b-1} \frac{1}{Q^{t+i}-1} \frac{(i+t-1)!}{t!\, i!} \binom{i+b+t-1}{t+i} \\
&+ \frac{2}{L} \sum_{s=0}^{b-1} \sum_{t=1}^{b-1} \frac{1}{s!\, t!} \sum_{k \neq 0} e^{2k\pi \boldsymbol{i} \log_Q n^*} \Gamma(-\chi_k + s + t) H_t(\chi_k - s - t) \\
&+ \log_Q n + \frac{\gamma}{L} + \log_Q(Q-1) - \frac{1}{2} - \frac{1}{L} H_{b-1} + \delta_{E_b}(\log_Q n^*) - \frac{1}{4} + \frac{\gamma}{L} - \frac{\gamma^2}{L^2} \\
&- \frac{H_{b-1}}{L} + \frac{2\,\gamma\, H_{b-1}}{L^2} - \frac{H_{b-1}^2}{L^2} - \log_Q^2 n - 2\, \log_Q n\, \log_Q(Q-1) - \log_Q^2(Q-1) \\
&+ \log_Q n - \frac{2\,\gamma}{L} \log_Q n + \frac{2}{L} H_{b-1} \log_Q n + \log_Q(Q-1) - \frac{2}{L} \gamma \log_Q(Q-1) \\
&+ \frac{2}{L} H_{b-1} \log_Q(Q-1) + \delta_{E_b}(\log_Q n^*) - \frac{2}{L} \gamma\, \delta_{E_b}(\log_Q n^*) + \frac{2}{L} H_{b-1}\, \delta_{E_b}(\log_Q n^*) \\
&- 2\, \log_Q n\, \delta_{E_b}(\log_Q n^*) - 2\, \log_Q(Q-1)\, \delta_{E_b}(\log_Q n^*) - \delta_{E_b}^2(\log_Q n^*) + o(1) \\
= &\ \frac{1}{12} + \frac{\pi^2}{6L^2} + \underbrace{\frac{2\,\gamma\, H_{b-1}}{L^2}}_{\boxed{4}} - \underbrace{\frac{H_{b-1}^2}{L^2}}_{\boxed{5}} + \underbrace{\frac{2}{L^2} \sum_{s=1}^{b-1} \frac{\psi(s)}{s}}_{\boxed{3}} + \underbrace{\frac{2}{L} \sum_{i \geq 1} \frac{(-1)^i}{i(Q^i-1)} \binom{i+b-1}{i}}_{\boxed{1}} \\
&+ \underbrace{\frac{2}{L} \sum_{i \geq 0} (-1)^i \sum_{t=1}^{b-1} \frac{1}{Q^{t+i}-1} \frac{(i+t-1)!}{t!\, i!} \binom{i+b+t-1}{t+i}}_{\boxed{2}} \\
&+ \frac{2}{L} \sum_{s=0}^{b-1} \sum_{t=1}^{b-1} \frac{1}{s!\, t!} \sum_{k \neq 0} e^{2k\pi \boldsymbol{i} \log_Q n^*} \Gamma(-\chi_k + s + t) H_t(\chi_k - s - t) \\
&+ \frac{2}{L} \sum_{k \neq 0} \Gamma(-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q n^*} \left[ \frac{\psi(-\chi_k)}{L} - \log_Q(Q-1) + 1 \right] \\
&- \frac{2}{L} \sum_{s=1}^{b-1} \frac{1}{s!} \sum_{k \neq 0} \Gamma(s-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q n^*} \left( \log_Q(Q-1) - 1 - \frac{\psi(s-\chi_k)}{L} \right) \\
&+ \frac{2}{L} \sum_{s=0}^{b-1} \frac{1}{s!} \sum_{k \neq 0} e^{2k\pi \boldsymbol{i} \log_Q n^*} \Gamma(s-\chi_k) H(\chi_k - s) + 2\delta_{E_b}(\log_Q n^*) - \frac{2}{L} \gamma\, \delta_{E_b}(\log_Q n^*) \\
&+ \frac{2}{L} H_{b-1}\, \delta_{E_b}(\log_Q n^*) - 2\, \log_Q(Q-1)\, \delta_{E_b}(\log_Q n^*) - \underbrace{\delta_{E_b}^2(\log_Q n^*)}_{\boxed{6}} + o(1).
\end{aligned}
$$

There are still a few simplifications that can be made to the variance. We can express $\boxed{1}$ and $\boxed{2}$ more simply by removing the $i = 0$ term from $\boxed{2}$ and including the $t=0$ term $\big($i.e., $\boxed{1}$ $\big)$ in its place. Then

$$\begin{aligned}
\boxed{1}+\boxed{2} &= \frac{2}{L}\sum_{i\geq 1}\frac{(-1)^i}{i(Q^i-1)}\binom{i+b-1}{i} \\
&\quad + \frac{2}{L}\sum_{i\geq 0}(-1)^i\sum_{t=1}^{b-1}\frac{1}{Q^{t+i}-1}\frac{(i+t-1)!}{t!\,i!}\binom{i+b+t-1}{t+i} \\
&= \frac{2}{L}\sum_{t=1}^{b-1}\frac{1}{(Q^t-1)}\frac{(t-1)!}{t!}\binom{b+t-1}{t} \\
&\quad + \frac{2}{L}\sum_{i\geq 1}(-1)^i\sum_{t=0}^{b-1}\frac{1}{Q^{t+i}-1}\frac{(i+t-1)!}{t!\,i!}\binom{i+b+t-1}{t+i} \\
&= \frac{2}{L}\sum_{t=1}^{b-1}\frac{1}{t(Q^t-1)}\binom{b+t-1}{t} \\
&\quad + \frac{2}{L}\sum_{i\geq 1}(-1)^i\sum_{t=0}^{b-1}\frac{1}{t(Q^{t+i}-1)}\binom{i+t-1}{i}\binom{i+b+t-1}{t+i} \\
&= \frac{2}{L}\sum_{t=1}^{b-1}\frac{1}{t(Q^t-1)}\binom{b+t-1}{t} \\
&\quad + \frac{2}{L}\sum_{t=0}^{b-1}\sum_{i\geq 1}\frac{(-1)^i}{Q^{t+i}-1}\frac{1}{i+t}\binom{i+t}{i}\binom{i+b+t-1}{t+i} \\
&= \frac{2}{L}\sum_{t=1}^{b-1}\frac{1}{t(Q^t-1)}\binom{b+t-1}{t} + \frac{2}{L}\sum_{t=0}^{b-1}\sum_{i\geq t+1}\frac{(-1)^{i-t}}{Q^i-1}\frac{1}{i}\binom{i}{i-t}\binom{i+b-1}{i} \\
&= \frac{2}{L}\sum_{t=1}^{b-1}\frac{1}{t(Q^t-1)}\binom{b+t-1}{t} \\
&\quad + \frac{2}{L}\sum_{t=0}^{b-1}\left[\sum_{i\geq 1}\frac{(-1)^{i-t}}{Q^i-1}\frac{1}{i}\binom{i}{i-t}\binom{i+b-1}{i} - [t\geq 1]\frac{1}{Q^t-1}\frac{1}{t}\binom{t+b-1}{t}\right].
\end{aligned}$$

Note here that for $t=0$, this term does not need to be removed. Hence

$$\begin{aligned}
\boxed{1}+\boxed{2} &= \frac{2}{L}\sum_{t=1}^{b-1}\frac{1}{t(Q^t-1)}\binom{b+t-1}{t} + \frac{2}{L}\sum_{t=0}^{b-1}\sum_{i\geq 1}\frac{(-1)^{i-t}}{Q^i-1}\frac{1}{i}\binom{i}{i-t}\binom{i+b-1}{i} \\
&\qquad - \frac{2}{L}\sum_{t=1}^{b-1}\frac{1}{t(Q^t-1)}\binom{t+b-1}{t} \\
&= \frac{2}{L}\sum_{t=0}^{b-1}\sum_{i\geq 1}\frac{(-1)^{i-t}}{i(Q^i-1)}\binom{i}{i-t}\binom{i+b-1}{i} \\
&= \frac{2}{L}\sum_{i\geq 1}\frac{(-1)^i}{i(Q^i-1)}\binom{i+b-1}{i}\sum_{t=0}^{b-1}(-1)^t\binom{i}{t} \\
&= \frac{2}{L}\sum_{i\geq 1}\frac{(-1)^i}{i(Q^i-1)}\binom{i+b-1}{i}(-1)^{b-1}\binom{i-1}{b-1} \qquad \text{(identity from [12])}
\end{aligned}$$

$$= \frac{2}{L}\sum_{i\geq 1}\frac{(-1)^{i+b-1}}{i(Q^i-1)}\binom{i+b-1}{i}\binom{i-1}{b-1}.$$

The harmonic numbers can be expressed as $H_n = \psi(n+1)+\gamma$. Thus

$$\begin{aligned}
\boxed{3} &= \frac{2}{L^2}\sum_{s=1}^{b-1}\frac{\psi(s)}{s}\\
&= \frac{2}{L^2}\sum_{s=1}^{b-1}\frac{1}{s}(H_{s-1}-\gamma)\\
&= \frac{2}{L^2}\sum_{s=1}^{b-1}\frac{1}{s}\left(H_s-\frac{1}{s}\right)-\frac{2\gamma}{L^2}\sum_{s=1}^{b-1}\frac{1}{s} \qquad \left(H_{s-1}=H_s-\frac{1}{s}\right)\\
&= \frac{2}{L^2}\sum_{s=1}^{b-1}\frac{1}{s}H_s-\frac{2}{L^2}\sum_{s=1}^{b-1}\frac{1}{s^2}-\frac{2\gamma}{L^2}H_{b-1}\\
&= \frac{2}{L^2}\frac{1}{2}\big(H_{b-1}^2+H_{b-1}^{(2)}\big)-\frac{2}{L^2}H_{b-1}^{(2)}-\frac{2\gamma}{L^2}H_{b-1}\\
&= \frac{1}{L^2}H_{b-1}^2-\frac{1}{L^2}H_{b-1}^{(2)}-\frac{2\gamma}{L^2}H_{b-1},
\end{aligned}$$

which means we cancel the terms $\boxed{4}-\boxed{5}$. It is also necessary to look at the term $\boxed{6}$, whose mean is non-zero. In order to use results from [17] we rewrite

$$\begin{aligned}
\delta_{E_b}(x) &= \frac{1}{L}\sum_{k\neq 0}\frac{e^{2k\pi\boldsymbol{i}x}}{\chi_k}\frac{\Gamma(b-\chi_k)}{\Gamma(b)}\\
&= \frac{1}{L}\sum_{k\neq 0}\frac{e^{2k\pi\boldsymbol{i}x}}{\chi_k}\frac{(b-\chi_k-1)!}{(b-1)!}\\
&= \frac{1}{L}\sum_{k\neq 0}\frac{e^{2k\pi\boldsymbol{i}x}}{\chi_k}(-\chi_k)!\binom{b-\chi_k-1}{b-1}\\
&= \frac{1}{L}\sum_{k\neq 0}e^{2k\pi\boldsymbol{i}x}\big(-(-\chi_k-1)!\big)\binom{b-\chi_k-1}{b-1}\\
&= -\frac{1}{L}\sum_{k\neq 0}e^{2k\pi\boldsymbol{i}x}\Gamma(-\chi_k)\binom{b-\chi_k-1}{b-1}, \qquad (4.29)
\end{aligned}$$

which is exactly the function dealt with in [17]. In [17] the square of this function is split into two parts – a constant (the mean of the square of the function) and the remaining periodic function of period 1 and mean zero. (In the classical case the special case of this function was when $b=1$.) We write

$$\delta_{E_b}^2(x) = [\delta_{E_b}^2]_0+\tilde{\delta}_{E_b}(x),$$

where

$$\tilde{\delta}_{E_b}(x) = \frac{1}{L^2}\sum_{k\neq 0}\sum_{j\neq 0,\neq k}\frac{\Gamma(b-\chi_k)}{\chi_k\Gamma(b)}\frac{\Gamma(b-\chi_{k-j})}{\chi_{k-j}\Gamma(b)}e^{2\pi\boldsymbol{i}kx}.$$

Also,

$$[\delta^2_{E_b}]_0 = \frac{\pi^2}{6L^2} + \frac{1}{12} - \log_Q 2 + \frac{2}{L}\sum_{j=1}^{b-1}\frac{1}{2j}\binom{2j}{j}\sum_{h\ge 0}\binom{-2j}{h}\frac{1}{Q^{h+j}-1}$$
$$- \frac{2}{L}\sum_{h\ge 1}\frac{(-1)^{h-1}}{h(Q^h-1)} + \frac{1}{L}\sum_{j=1}^{b-1}\frac{1}{2j}\binom{2j}{j}2^{-2j} - \frac{H^{(2)}_{b-1}}{L^2}.$$

Therefore, the general variance is

$$\mathbb{V}_b(d_n) = \log_Q 2 + \frac{2}{L}\sum_{i\ge 1}\frac{(-1)^{i+b-1}}{i(Q^i-1)}\binom{i+b-1}{i}\binom{i-1}{b-1}$$
$$- \frac{2}{L}\sum_{j=1}^{b-1}\frac{1}{2j}\binom{2j}{j}\sum_{h\ge 0}\binom{-2j}{h}\frac{1}{Q^{h+j}-1} + \frac{2}{L}\sum_{h\ge 1}\frac{(-1)^{h-1}}{h(Q^h-1)}$$
$$- \frac{1}{L}\sum_{j=1}^{b-1}\frac{1}{2j}\binom{2j}{j}2^{-2j} + \delta_{V_b}(\log_Q n^*) + o(1),$$

where

$$\delta_{V_b}(x) = \frac{2}{L}\sum_{k\ne 0}\Gamma(-\chi_k)e^{2k\pi \boldsymbol{i}x}\left[\frac{\psi(-\chi_k)}{L} - \log_Q(Q-1) + 1\right] \tag{4.30}$$
$$- \frac{2}{L}\sum_{s=1}^{b-1}\frac{1}{s!}\sum_{k\ne 0}\Gamma(s-\chi_k)e^{2k\pi \boldsymbol{i}x}\big(\log_Q(Q-1) - 1 - \frac{\psi(s-\chi_k)}{L}\big)$$
$$+ \frac{2}{L}\sum_{s=0}^{b-1}\frac{1}{s!}\sum_{k\ne 0}e^{2k\pi \boldsymbol{i}x}\Gamma(s-\chi_k)H(\chi_k - s)$$
$$+ \frac{2}{L}\sum_{s=0}^{b-1}\sum_{t=1}^{b-1}\frac{1}{s!\,t!}\sum_{k\ne 0}e^{2k\pi \boldsymbol{i}x}\Gamma(-\chi_k+s+t)H_t(\chi_k-s-t)$$
$$+ 2\delta_{E_b}(x) - \frac{2}{L}\gamma\,\delta_{E_b}(x) + \frac{2}{L}H_{b-1}\,\delta_{E_b}(x) - 2\,\log_Q(Q-1)\,\delta_{E_b}(x) - \tilde{\delta}_{E_b}(x).$$

This concludes the proof of Theorem 4.2. ■

## 4.3 The mean and variance for large $b$

To examine this variance result as $b \to \infty$, we can use results from [17] which state that

$$-\frac{1}{L}\sum_{j=1}^{b-1}\frac{1}{2j}\binom{2j}{j}2^{-2j} = -\frac{\log 2}{L} + \frac{1}{\sqrt{\pi}}b^{-\frac{1}{2}} + O\big(b^{-\frac{3}{2}}\big)$$

and (for any $\varepsilon > 0$)

$$-\frac{2}{L}\sum_{j=1}^{b-1}\frac{1}{2j}\binom{2j}{j}\sum_{h\ge 0}\binom{-2j}{h}\frac{1}{Q^{h+j}-1}$$

$$
\begin{aligned}
&= -\frac{2}{L}\sum_{m\geq 1}\log(1+Q^{-m}) + O\Bigg(\Big(\frac{4}{Q(1+Q^{-1})^2}-\varepsilon\Big)^b\Bigg)\\
&= -\frac{2}{L}\sum_{h\geq 1}\frac{(-1)^{h-1}}{h(Q^h-1)} + O\Bigg(\Big(\frac{4}{Q(1+Q^{-1})^2}-\varepsilon\Big)^b\Bigg),
\end{aligned}
$$

whose big-$O$ term is exponentially small as $b \to \infty$. From [15],

$$
\frac{2}{L}\sum_{i\geq 1}\frac{(-1)^{i+b-1}}{i(Q^i-1)}\binom{i+b-1}{i}\binom{i-1}{b-1} = O\Bigg(\Big(\frac{4}{Q(1+Q^{-1})^2}-\varepsilon\Big)^b\Bigg)
$$

for any $\varepsilon > 0$ and is likewise exponentially small. Thus as $b \to \infty$, the constant in the asymptotic expansion of the variance is

$$
\frac{1}{\sqrt{\pi}}b^{-\frac{1}{2}} + O\big(b^{-\frac{3}{2}}\big) + O\Bigg(\Big(\frac{4}{Q(1+Q^{-1})^2}-\varepsilon\Big)^b\Bigg) + \delta_{V_b}(x) = O\big(b^{-\frac{1}{2}}\big).
$$

# Chapter 5

# Conclusion

Sequences of natural numbers (or 'words') were considered, where each letter occurred with geometric probability. It was shown that the expected value for the number of distinct values in a geometrically distributed sample of length $n$ is asymptotic to $\log_Q n$ as $n \to \infty$. If letters are required to occur more often (say, $b$ times) then the expected value is dependent on $b$ and as $b$ gets large the expected value decreases logarithmically with respect to $b$ (it will decrease asymptotically according to the term $-\log_Q b$). For $b = 1$, the classical case, this extra term (the only term dependent on $b$) disappears.

The variance is small, with the calculations becoming more intricate when $b$ is larger than 1. In addition to the main term in the classical case – which is equal to $\log_Q 2$ – the general variance also has various sums which were shown to be $O\big(b^{-\frac{1}{2}}\big)$ for large $b$. Again, substituting 1 for $b$ resorts back to the classical case. An extension of this work can be found in [26].

# Part II

# Maxima and Minima

# Chapter 6

# Introduction

Given a word where the letters are natural numbers, we consider these letters to occur independently and with geometric probability. So for $p + q = 1$, each letter $i$ appears in the word with probability $pq^{i-1}$.

"What is the probability that the maximum in a word occurs in the first position?" This question is addressed in Chapter 7 and later is generalised in Chapter 8 to finding the probability that the maximum occurs in the first $d$ positions of a word. We take words of length $n$ and require $d \leq n$. To begin with $d$ is considered fixed. The same scenarios are then considered for the minimum in Chapters 9 and 10. Another generalisation dealt with in Chapter 11 is that the minimum value of the first $d$ letters is greater than ('strict') and possibly equal to ('weak') all other values in the word. All of these ideas can be interpreted in the strict and weak sense.

A similar concept has been looked at for compositions, see [22]. Also, the probability that there is a single winner in a geometrically distributed sample is looked at in [6] (see Chapter 7). The method in this thesis, however, is to use generating functions and Rice's method [2, 16, 21, 25, 28, 29, 30, 31] to obtain both the main term and the periodic fluctuations which appear. The first case is dealt with in more detail than the rest, as the process is similar in each case.

It must be noted that in Chapters 9 and 10, there are elementary probabilistic explanations for the results given. It is thus unnecessary to use these techniques, but they are included for consistency of method, and labelled 'Proposition' rather than 'Theorem'. An example of the probabilistic argument, as suggested by one of the external examiners, would be as follows: Suppose $X_1, X_2, \ldots, X_n$ are independently and identically distributed geometric random variables, with $p$ and $q$ as defined above. Then $P(X = k) = pq^{k-1}$ and $P(X > k) = pq^k + pq^{k+1} + pq^{k+2} + \cdots = pq^k(1 + q + q^2 + \cdots) = pq^k \frac{1}{1-q} = q^k$. Hence the probability that a strict minimum occurs in the first position is given by

$$\begin{aligned}\sum_{j\geq 1} P(X_1 = j, X_m > j, 2 \leq m \leq n) &= \sum_{j\geq 1} P(X = j)(P(X > j))^{n-1} \\ &= \sum_{j\geq 1} pq^{j-1}(q^j)^{n-1} \\ &= \frac{p}{q}\frac{q^n}{1-q^n} \\ &= \frac{Q-1}{Q^n-1},\end{aligned}$$

as in Proposition 9.1. For Proposition 9.2 the use of the non-strict inequality ($X_m \geq j$) gives the result, and the propositions in Chapter 10 follow by summing over the first $d$ positions.

Having established these results, a follow-up question might be "what is the probability that the maximum in a sequence of length $n$ occurs in the first $d$ positions, if $d$ grows with $n$?". We use Mellin transforms to deal with the cases when $d = \alpha n$ for $0 < \alpha \leq 1$ (in Chapter 12), and when $d = \alpha n^\gamma$ for $0 < \gamma < 1$ (in Chapter 13).

Rice's method was discussed previously. It makes use of Lemma 1 which is recalled below.

*Let $\mathcal{C}$ be a curve surrounding the points $1, 2, \ldots, n$ in the complex plane, and let $f(z)$ be analytic inside $\mathcal{C}$. Then*

$$\sum_{k=1}^{n} \binom{n}{k} (-1)^k f(k) = -\frac{1}{2\pi \boldsymbol{i}} \int_{\mathcal{C}} [n; z] f(z) dz,$$

*where*

$$[n; z] = \frac{(-1)^{n-1} n!}{z(z-1)\cdots(z-n)} = \frac{\Gamma(n+1)\Gamma(-z)}{\Gamma(n+1-z)}. \qquad \square$$

By extending the contour of integration, it turns out that under suitable growth conditions (see [10]) the asymptotic expansion of our alternating sum is given by

$$\sum \operatorname{Res}([n; z] f(z)) + \text{smaller order terms},$$

where the sum is taken over all poles different from $1, \ldots, n$. Poles that lie more to the left lead to smaller terms in the asymptotic expansion. The symbol $\boldsymbol{i}$ is used to represent $\sqrt{-1}$ (as opposed to $i$ used elsewhere as an index).

The other technique used is the Mellin transform. The Mellin transform of a function $f(x)$ is a function of the complex variable $s$ and is denoted by an asterisk. It is defined by

$$f^*(s) := \int_0^\infty f(x) x^{s-1} dx. \tag{6.1}$$

The Mellin transform exists in a strip in the complex plane, denoted $\langle -u, -v\rangle$ where $-u < \Re(s) < -v$ if $\Re(s)$ denotes the real part of the complex number $s$. To find the boundary values of the strip, two limits are taken on $f(x)$. For the left boundary of the strip, if $f(x) = O(x^u)$ as $x \to 0$, then $-u$ bounds how small the real part of $s$ can be. The largest value of the real part of $s$ is given by $-v$ if $f(x) = O(x^v)$ as $x \to \infty$.

In order to re-establish the original function $f(x)$, the inversion formula must be used. The inverse of the Mellin transform is

$$f(x) = \frac{1}{2\pi \boldsymbol{i}} \int_{c-\boldsymbol{i}\infty}^{c+\boldsymbol{i}\infty} f^*(s) x^{-s} ds \tag{6.2}$$

where $c \in \mathbb{R}$ is in the fundamental strip. This is sometimes written as

$$f(x) = \frac{1}{2\pi \boldsymbol{i}} \int_{(c)} f^*(s) x^{-s} ds. \tag{6.3}$$

It is in considering this inverse transform that we approximate functions using residue calculus. We look at the poles of the integrand. If we are interested in where the parameter is large, we move the contour right and collect negative residues. If the parameter of interest needs to be small, positive residues are collected by moving the contour line left. In the examples which follow, the parameter of interest is the length of the word $n$, and since we wish to approximate for large $n$, we move the contour right until we pass the first poles of the integrand and then evaluate the negative residues.

References involving Mellin transforms include [7, 8, 10, 11, 20, 24, 38].

In Table 6.1 below are examples of words created randomly according to a geometric distribution on certain values of $q$. The words were created randomly on Mathematica (Using 'GeometricDistribution' from the statistics package 'DiscreteDistributions'):

| $q$ | Random word |
|---|---|
| $\frac{1}{100}$ | 11111111111111111111 |
| $\frac{1}{2}$ | 13211312151511122112 |
| $\frac{99}{100}$ | 168;45;3;200;83;85;49;79;43;127;17;1;2;140;123;7;116;77;58;81 |

Table 6.1: Examples of random words with given geometric probabilities.

# Chapter 7

# The maximum in the first position

We break this problem into two parts, one case ('strict') where we do not allow the maximum ($k$) to appear anywhere else in the word, and the other ('weak') where $k$ is allowed to appear elsewhere, as long as it first appears in position 1. As an example, consider the words 4111322131 and 4141324112. The maximum in both is 4, which appears only in the first position in the former but in positions 1,3 and 7 in the latter.

The results which follow are proved in this chapter. A different (probabilistic) approach with similar results can be found in [6] following a problem stated in [34]. Eisenberg, Stengle and Strang ([6]) consider the probability that there is only one maximum in a geometrically distributed sample, and by multiplying this result by $\frac{1}{n}$, we have a similar idea to that in Theorem 7.1. Similarly in Chapter 8 one can multiply by $\frac{d}{n}$ to obtain a more general result.

**Theorem 7.1** *The probability that the only maximum value in a word of length $n$ is in the first position of the word is*

$$P_s(M) \sim \frac{1 - Q^{-1}}{Ln}(1 + \delta(n)), \tag{7.1}$$

*as $n \to \infty$, where $Q = q^{-1}$, $L = \log Q$, and*

$$\delta(x) = \sum_{k \neq 0} \Gamma(1 - \chi_k) e^{2k\pi \boldsymbol{i} \log_Q x}$$

*as defined in (7.5).*

**Theorem 7.2** *The probability that the maximum value in a word of length $n$ is in the first position and possibly other positions is*

$$P_w(M) \sim \frac{Q - 1}{Ln}(1 + \delta(n)), \tag{7.2}$$

*as* $n \to \infty$, *where*

$$\delta(x) = \sum_{k\neq 0} \Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q x}.$$

## 7.1 Maximum in the first position – (strict)

We wish to construct words whose largest element appears in the first position and nowhere else in the word (i.e., the strict case). We interpret this situation symbolically as

$$\bigcup_{k\geq 1} k\{1,\dots,k-1\}^*.$$

The union is taken over all natural numbers $k$, where in each case $k$ represents the maximum value in the word. We start with the maximum $k$ and this is followed by a sequence of any length from the set of letters $\{1,\dots,k-1\}$ in any order, with repeats allowed.

We translate this symbolic equation into a generating function whose coefficients represent the probability that a word of length $n$ has its maximum in the first position. It is

$$F_M^{(s)}(z) := \sum_{k\geq 1} zpq^{k-1} \frac{1}{1-\sum\limits_{j=1}^{k-1} zpq^{j-1}},$$

where the sum on $j$ is a telescoping series (since $p = 1-q$), and thus

$$\begin{aligned}
F_M^{(s)}(z) &= \sum_{k\geq 1} zpq^{k-1} \frac{1}{1-z\sum\limits_{j=1}^{k-1}(q^{j-1}-q^j)} \\
&= \sum_{k\geq 1} zpq^{k-1} \frac{1}{1-z(1-q^{k-1})} \\
&= \sum_{k\geq 1} zpq^{k-1} \sum_{j\geq 0} z^j (1-q^{k-1})^j \\
&= \sum_{k\geq 1}\sum_{j\geq 0} z^{j+1} pq^{k-1}(1-q^{k-1})^j.
\end{aligned}$$

We want to consider the coefficient of $z^n$ (i.e., $n := j+1$ so $n-1 = j$) in this ordinary generating function, as this will give us the probability that a word of length $n$ has a strict maximum in the first position.

$$\begin{aligned}
[z^n]F_M^{(s)}(z) &= \sum_{k\geq 1} pq^{k-1}(1-q^{k-1})^{n-1} \\
&= \sum_{k\geq 1} pq^{k-1} \sum_{j=0}^{n-1} \binom{n-1}{j} (-q^{k-1})^j
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k\geq 1}\sum_{j=0}^{n-1}\binom{n-1}{j}(-1)^j p q^{(k-1)(j+1)} \\
&= \sum_{j=0}^{n-1}\binom{n-1}{j}(-1)^j p \sum_{k\geq 1} q^{(k-1)(j+1)} \\
&= \sum_{j=0}^{n-1}\binom{n-1}{j}(-1)^j p \sum_{K\geq 0} (q^{j+1})^K \quad \text{(where } K := k-1) \\
&= \sum_{j=0}^{n-1}\binom{n-1}{j}(-1)^j p \frac{1}{1-q^{j+1}} \\
&= \sum_{j=0}^{n-1}\binom{n-1}{j}(-1)^j \frac{1-Q^{-1}}{1-Q^{-(j+1)}}.
\end{aligned}
$$

We can now use complex analysis to evaluate the alternating sum asymptotically. The method is called 'Rice's method' and was discussed in Part I. We use the same lemma (Lemma 1, see (2.1)), which allows us to express a sum such as this as an integral.

In our case, the sum begins at zero, so we start with a contour surrounding points $0, 1, \ldots, n-1$. The function whose poles we examine is $f(z) := \frac{1-Q^{-1}}{1-Q^{-(z+1)}}$, and thus the poles we need to consider are at $z+1 = 0$ and $z+1 = \chi_k$, where $\chi_k = \frac{2k\pi \boldsymbol{i}}{L}$ for all $k \in \mathbb{Z}\setminus\{0\}$, $(L = \log Q)$. For the first of these poles we have (expanding around $z = -1$)

$$
\frac{1-Q^{-1}}{1-Q^{-z-1}} = \frac{1-Q^{-1}}{1-e^{(-z-1)L}} \sim \frac{1-Q^{-1}}{1-(1+(-z-1)L)} = \frac{1-Q^{-1}}{(z+1)L},
$$

so the residue is $\frac{1-Q^{-1}}{L}$. The contribution of the kernel is

$$
\begin{aligned}
[n-1;-1] &= \frac{(-1)^{n-1-1}(n-1)!}{(-1)(-1-1)\cdots(-1-(n-1))} \\
&= \frac{(-1)^{n-2}(n-1)!}{(-1)(-2)\cdots(-n)} \\
&= \frac{(-1)^{n-2}(n-1)!}{(-1)^n n!} \\
&= \frac{1}{n}. \qquad (7.3)
\end{aligned}
$$

Thus for the pole at $z = -1$ we have $\frac{1-Q^{-1}}{nL}$. For the other poles we let $\varepsilon := z+1-\chi_k$. Then (since $Q^{\chi_k} = 1$ by (3.6))

$$
\frac{1-Q^{-1}}{1-Q^{-z-1}} = \frac{1-Q^{-1}}{1-Q^{-\varepsilon-\chi_k}} = \frac{1-Q^{-1}}{1-Q^{-\varepsilon}} = \frac{1-Q^{-1}}{1-e^{-\varepsilon L}} \sim \frac{1-Q^{-1}}{1-(1-\varepsilon L)} = \frac{1-Q^{-1}}{\varepsilon L},
$$

so again we have a residue of $\frac{1-Q^{-1}}{L}$. In this case, the contribution of the kernel is asymptotic as $n \to \infty$.

$$\begin{aligned}
[n-1; \chi_k - 1] &= \frac{\Gamma(n-1+1)\Gamma(-\chi_k+1)}{\Gamma(n-1+1-\chi_k+1)} \\
&= \Gamma(1-\chi_k)\frac{\Gamma(n)}{\Gamma(n-\chi_k+1)} \\
&\sim \Gamma(1-\chi_k)n^{\chi_k-1} \qquad \big(\text{see [1, page 257]}\big), \\
&= \frac{1}{n}\Gamma(1-\chi_k)e^{\chi_k \log n}. \qquad (7.4)
\end{aligned}$$

So altogether for the remaining poles we have

$$\frac{1-Q^{-1}}{Ln}\sum_{k\neq 0}\Gamma(-\chi_k+1)e^{2k\pi\boldsymbol{i}\log_Q n}.$$

Putting these results together, it can be seen that the expected value is asymptotic to

$$\frac{1-Q^{-1}}{Ln} + \frac{1-Q^{-1}}{Ln}\sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n}, \qquad \text{as } n \to \infty.$$

Since the fluctuating function will occur frequently, we define it as

$$\delta(x) := \sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q x}. \qquad (7.5)$$

We can thus approximate the probability that the strict maximum is in the first position as

$$P_s(M) \sim \frac{1-Q^{-1}}{Ln}(1+\delta(n)).$$

This concludes the proof of Theorem 7.1. ■

## 7.2 Maximum in the first position – (weak)

If we allow the maximum to occur elsewhere in the word too, we change the symbolic equation slightly, to

$$\bigcup_{k\geq 1} k\{1,\ldots,k\}^*,$$

since the sequence which follows the initial $k$ may now include the letter $k$. This can be given as the generating function

$$F_M^{(w)}(z) := \sum_{k\geq 1} zpq^{k-1}\frac{1}{1-\sum\limits_{j=1}^{k} zpq^{j-1}}.$$

As before, we can rewrite this as

$$
\begin{aligned}
F_M^{(w)}(z) &= \sum_{k\geq 1} zpq^{k-1}\frac{1}{1-z(1-q^k)} \\
&= \sum_{k\geq 1} zpq^{k-1}\sum_{j\geq 0} z^j(1-q^k)^j \\
&= \sum_{k\geq 1}\sum_{j\geq 0} z^{j+1}pq^{k-1}(1-q^k)^j
\end{aligned}
$$

which makes it easier to find the coefficient of $z^n$:

$$
\begin{aligned}
[z^n]F_M^{(w)}(z) &= [z^n]\sum_{k\geq 1}\sum_{j\geq 0} z^{j+1}pq^{k-1}(1-q^k)^j \\
&= \sum_{k\geq 1} pq^{k-1}(1-q^k)^{n-1} \\
&= \sum_{k\geq 1} pq^{k-1}\sum_{j=0}^{n-1}\binom{n-1}{j}(-q^k)^j \\
&= \sum_{j=0}^{n-1}\binom{n-1}{j}(-1)^j\frac{p}{q}\sum_{k\geq 1}(q^{j+1})^k \\
&= \sum_{j=0}^{n-1}\binom{n-1}{j}(-1)^j\frac{p}{q}\frac{q^{j+1}}{1-q^{j+1}} \\
&= \sum_{j=0}^{n-1}\binom{n-1}{j}(-1)^j\frac{p}{q}\frac{1}{Q^{j+1}-1} \\
&= \sum_{j=0}^{n-1}\binom{n-1}{j}(-1)^j\frac{Q-1}{Q^{j+1}-1}.
\end{aligned}
$$

Again we use Rice's method and look at the poles of the function $f(z) := \frac{Q-1}{Q^{z+1}-1}$. We must consider the poles at $z=-1$ and $z+1=\chi_k$. The first of these gives

$$
f(z) = \frac{Q-1}{e^{(z+1)L}-1} \sim \frac{Q-1}{(z+1)L}, \qquad \text{near } z=-1,
$$

and the kernel is the same as in the previous example (see (7.3)),

$$
[n-1,-1] = \frac{1}{n}.
$$

Thus the residue is $\frac{Q-1}{Ln}$. For the remaining poles, if $\varepsilon := z+1-\chi_k$, then

$$
f(z) = \frac{Q-1}{Q^{z+1}-1} = \frac{Q-1}{Q^{\varepsilon+\chi_k}-1} = \frac{Q-1}{Q^{\varepsilon}-1} = \frac{Q-1}{e^{\varepsilon L}-1} \sim \frac{Q-1}{\varepsilon L}, \tag{7.6}
$$

and the coefficient of $\varepsilon^{-1}$ in (7.6) is $\frac{Q-1}{L}$. The kernel is again the same, so (from (7.4))

$$
[n-1;\chi_k-1] \sim \frac{1}{n}\Gamma(1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n},
$$

and we sum this over all non-zero $k$ to get

$$\frac{Q-1}{Ln}\sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi \boldsymbol{i}\log_Q n}.$$

In all, we get an expected probability asymptotic to

$$P_w(M) \sim \frac{Q-1}{Ln}(1+\delta(n)),$$

as $n \to \infty$. Recall $\delta(x) = \sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi \boldsymbol{i}\log_Q x}$ from equation (7.5).

Hence Theorem 7.2 is proved. ■

The only difference between this and the strict case is the $Q-1$ instead of the $1-Q^{-1}$. Since $Q > 1 \Rightarrow Q-1 > 1-Q^{-1}$, the weak case will be slightly larger. This is to be expected as it is more likely that we have a $k$ in position 1 if there is more than one $k$ in the word. That is, if there is one $k$ in a word of length $n$, there is a small chance of it being exactly in position 1. However, if there are, say, 3 $k$'s in a word of length $n$, then there is a greater chance that one of them is in position 1.

# Chapter 8

# The maximum in the first $d$ positions

This idea can be viewed in four different ways, all of which end up with similar results asymptotically. These asymptotic results are proportional to $d$ (in fact, are the previous results with a factor $d$) as there are $d$ chances that $k$ will be where we want it to be, so the probability increases. For the problem to make sense, we need a word of length $d$ or more (where, for the time being, $d$ is fixed and independent of $n$) in all of these scenarios. We can have the (strict, strict) case where $k$ can only appear once in the whole word, and this appearance must be in the first $d$ letters of the word. Next, we can allow $k$ to appear more than once in the first $d$ positions, but never in the rest of the word. We will call this the (weak, strict) case. Alternatively, we can allow the letter $k$ to appear any number of times in the rest of the word, but only once in the first $d$ places – i.e., the (strict, weak) case. Lastly we can let $k$ be anywhere in the word, any number of times, as long as it occurs at least once in the first $d$ places. We call this case (weak, weak).

The following results are presented:

**Theorem 8.1** *The probability that the single maximum value in a geometrically distributed sample of length $n$ appears in the first $d$ positions of a word is asymptotic (as $n \to \infty$) to*

$$P_{ss}(M) \sim \frac{(1-Q^{-1})d}{Ln}(1+\delta(n)), \tag{8.1}$$

*where $\delta(x) = \sum_{k\neq 0} \Gamma(1-\chi_k)e^{2k\pi \boldsymbol{i} \log_Q x}$ as defined in equation (7.5).*

**Theorem 8.2** *The probability that the maximum value in a geometrically distributed sample of length $n$ appears in the first $d$ positions of a word any number of* *times (but at least once), and nowhere else in the word is approximately*

$$P_{ws}(M) \sim \frac{(1-Q^{-1})d}{Ln}(1+\delta(n)), \tag{8.2}$$

*as $n \to \infty$, where $\delta(x)$ is defined in equation (7.5).*

**Theorem 8.3** *The probability that the maximum value in a geometrically distributed sample of length $n$ appears only once in the first $d$ positions of a word and any number of times in the remaining $n-d$ letters is asymptotic to*

$$P_{sw}(M) \sim \frac{(Q-1)d}{Ln}(1+\delta(n)), \tag{8.3}$$

*as $n \to \infty$, where $\delta(x)$ is defined in (7.5).*

**Theorem 8.4** *The probability that the maximum value in a geometrically distributed sample of length $n$ appears at least once in the first $d$ positions of a word and any number of times anywhere else in the word is asymptotic to*

$$P_{ww}(M) \sim \frac{(Q-1)d}{Ln}(1+\delta(n)), \tag{8.4}$$

*as $n \to \infty$, for $\delta(x)$ from (7.5).*

## 8.1 Maximum in the first $d$ positions – (strict, strict)

In this case we have strictly only one maximum ($k$) in the word, which has to appear somewhere in the first $d$ places. All other letters in the word must be natural numbers which are less than or equal to $k-1$. Symbolically we represent the first $d$ letters as

$$A_k^{(s)} := \bigcup_{i=0}^{d-1} \{1,\dots,k-1\}^i k \{1,\dots,k-1\}^{d-1-i}, \tag{8.5}$$

and thus all possible words with this restriction can be symbolised by

$$\bigcup_{k\geq 1} A_k^{(s)}\{1,\dots,k-1\}^*.$$

The generating function then becomes

$$F_M^{(s,s)}(z) := \sum_{k\geq 1}\sum_{i=0}^{d-1}\left(\sum_{j=1}^{k-1} zpq^{j-1}\right)^i zpq^{k-1}\left(\sum_{j=1}^{k-1} zpq^{j-1}\right)^{d-1-i} \frac{1}{1-\sum\limits_{j=1}^{k-1} zpq^{j-1}}$$

$$= \sum_{k\ge1}\sum_{i=0}^{d-1} z^i(1-q^{k-1})^i zpq^{k-1}z^{d-1-i}(1-q^{k-1})^{d-1-i}\frac{1}{1-z(1-q^{k-1})}$$

$$= \sum_{k\ge1}\sum_{i=0}^{d-1} z^d(1-q^{k-1})^i pq^{k-1}(1-q^{k-1})^{d-1-i}\sum_{j\ge0} z^j(1-q^{k-1})^j,$$

and thus

$$\begin{aligned}
[z^n]F_M^{(s,s)}(z) &= [z^n]\sum_{k\ge1}\sum_{i=0}^{d-1} z^d(1-q^{k-1})^i pq^{k-1}(1-q^{k-1})^{d-1-i}\sum_{j\ge0} z^j(1-q^{k-1})^j \\
&= \sum_{k\ge1}\sum_{i=0}^{d-1}(1-q^{k-1})^i pq^{k-1}(1-q^{k-1})^{d-1-i}(1-q^{k-1})^{n-d} \\
&= \sum_{k\ge1}\sum_{i=0}^{d-1}(1-q^{k-1})^{n-1}pq^{k-1} \\
&= \sum_{k\ge1} d(1-q^{k-1})^{n-1}pq^{k-1} \\
&= \sum_{k\ge1} d\sum_{h=0}^{n-1}\binom{n-1}{h}(-q^{k-1})^h pq^{k-1} \\
&= \sum_{h=0}^{n-1} d\binom{n-1}{h}(-1)^h\frac{p}{q^{h+1}}\sum_{k\ge1} q^{(h+1)k} \\
&= \sum_{h=0}^{n-1} d\binom{n-1}{h}(-1)^h\frac{p}{q^{h+1}}\frac{q^{h+1}}{1-q^{h+1}} \\
&= \sum_{h=0}^{n-1} d\binom{n-1}{h}(-1)^h\frac{Q^{h+1}(1-Q^{-1})}{Q^{h+1}-1}.
\end{aligned}$$

We use Rice on $f(z) := \frac{Q^{z+1}(1-Q^{-1})}{Q^{z+1}-1}$, and consider poles at $z+1=0$ and $z+1=\chi_k$. To expand $f(z)$ about $z=-1$, we let $\varepsilon := z+1$. Then

$$f(z) = \frac{Q^{\varepsilon}(1-Q^{-1})}{Q^{\varepsilon}-1} = \frac{e^{\varepsilon L}(1-Q^{-1})}{e^{\varepsilon L}-1} \sim \frac{1-Q^{-1}}{\varepsilon L} \tag{8.6}$$

and the residue is $\frac{1-Q^{-1}}{L}$. This is combined with the kernel contribution (see (7.3))

$$[n-1,-1] = \frac{1}{n},$$

so that the first pole gives us

$$\frac{1-Q^{-1}}{Ln}.$$

The remainder of the poles use $\varepsilon := z+1-\chi_k$, so

$$f(z) = \frac{Q^{\varepsilon+\chi_k}(1-Q^{-1})}{Q^{\varepsilon+\chi_k}-1} = \frac{Q^{\varepsilon}(1-Q^{-1})}{Q^{\varepsilon}-1} \sim \frac{1-Q^{-1}}{\varepsilon L},$$

as in (8.6) and the residue is $\frac{1-Q^{-1}}{L}$. The kernel is

$$[n-1, \chi_k - 1] \sim \frac{1}{n}\Gamma(1-\chi_k)e^{2k\pi \boldsymbol{i} \log_Q n},$$

from (7.4) which means that the remaining poles give

$$\frac{1-Q^{-1}}{Ln}\sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi \boldsymbol{i} \log_Q n}.$$

In total, we have the probability asymptotic to

$$\frac{(1-Q^{-1})d}{Ln} + \frac{(1-Q^{-1})d}{Ln}\sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi \boldsymbol{i} \log_Q n}$$

and thus

$$P_{ss}(M) \sim \frac{(1-Q^{-1})d}{Ln}(1+\delta(n)),$$

where $\delta(x)$ is defined as in equation (7.5). This brings us to the end of the proof of Theorem 8.1. ∎

## 8.2 Maximum in the first $d$ positions – (weak, strict)

For this scenario we require that there is at least one $k$ in the first $d$ places, and possibly more. However, the rest of the word may only have letters from the set $\{1, \ldots, k-1\}$. We express the first $d$ letters symbolically as

$$A_k^{(w)} = \bigcup_{i=0}^{d-1}\{1, \ldots, k-1\}^i k\{1, \ldots, k\}^{d-1-i}. \tag{8.7}$$

Note that we must fix the first $k$ to avoid counting the same sequence more than once. Altogether we get

$$\bigcup_{k\geq 1} A_k^{(w)}\{1, \ldots, k-1\}^*.$$

Thus the generating function is

$$\begin{aligned}
F_M^{(w,s)}(z) &:= \sum_{k\geq 1}\sum_{i=0}^{d-1}\left(\sum_{j=1}^{k-1} zpq^{j-1}\right)^i zpq^{k-1}\left(\sum_{j=1}^{k} zpq^{j-1}\right)^{d-1-i}\frac{1}{1-\sum\limits_{j=1}^{k-1} zpq^{j-1}} \\
&= \sum_{k\geq 1}\sum_{i=0}^{d-1} z^i(1-q^{k-1})^i zpq^{k-1}z^{d-1-i}(1-q^k)^{d-1-i}\frac{1}{1-z(1-q^{k-1})}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k\geq 1} z^d \sum_{i=0}^{d-1} (1-q^{k-1})^i p q^{k-1} (1-q^k)^{d-1-i} \sum_{j\geq 0} z^j (1-q^{k-1})^j \\
&= \sum_{k\geq 1} z^{d+j} \sum_{i=0}^{d-1} \sum_{j\geq 0} (1-q^{k-1})^{i+j} p q^{k-1} (1-q^k)^{d-1-i}.
\end{aligned}
$$

Before considering coefficients of this generating function, note that although we make use of Rice's method here, we could just as well have used a Mellin transform. We keep to Rice for consistency, noting that Mellin transforms are discussed later in detail (when we make $d$ dependent on $n$), and the equivalent calculation for this generating function, and that of the (weak, weak) case, can be found in Appendix B.

$$
\begin{aligned}
&[z^n] F_M^{(w,s)}(z) \\
&= \sum_{k\geq 1} \sum_{i=0}^{d-1} (1-q^{k-1})^{i+n-d} p q^{k-1} (1-q^k)^{d-1-i} \qquad (8.8) \\
&= \sum_{k\geq 1} \sum_{i=0}^{d-1} \sum_{h=0}^{i+n-d} \binom{i+n-d}{h} (-q^{k-1})^h p q^{k-1} \sum_{l=0}^{d-1-i} \binom{d-1-i}{l} (-q^k)^l \\
&= \sum_{i=0}^{d-1} \sum_{h=0}^{i+n-d} \binom{i+n-d}{h} \sum_{l=0}^{d-1-i} \binom{d-1-i}{l} (-1)^{l+h} \frac{p}{q^{h+1}} \sum_{k\geq 1} q^{k(l+1+h)} \\
&= \sum_{i=0}^{d-1} \sum_{h=0}^{i+n-d} \binom{i+n-d}{h} \sum_{l=0}^{d-1-i} \binom{d-1-i}{l} (-1)^{l+h} \frac{p}{q^{h+1}} \frac{q^{l+1+h}}{1-q^{l+1+h}} \\
&= \sum_{i=0}^{d-1} \sum_{l=0}^{d-1-i} \binom{d-1-i}{l} (-1)^l \underbrace{\sum_{h=0}^{i+n-d} \binom{i+n-d}{h} (-1)^h \frac{Q^{h+1}(1-Q^{-1})}{Q^{l+1+h}-1}}. \qquad (8.9)
\end{aligned}
$$

Again, Rice's method is useful to evaluate the sum in the brace, and we consider poles of the function $f(z) := \frac{Q^{z+1}(1-Q^{-1})}{Q^{z+1+l}-1}$. These occur at $z+1+l=0$ and $z+1+l=\chi_k$. We start with the former, and define $\varepsilon := z+1+l$. Then

$$
\begin{aligned}
f(z) &= \frac{Q^{\varepsilon-l}(1-Q^{-1})}{Q^{\varepsilon}-1} \\
&= \frac{e^{\varepsilon L} Q^{-l}(1-Q^{-1})}{e^{\varepsilon L}-1} \\
&\sim \frac{Q^{-l}(1-Q^{-1})}{\varepsilon L} \qquad (8.10)
\end{aligned}
$$

and so the residue is $\frac{Q^{-l}(1-Q^{-1})}{L}$. For $N := n+i-d$, the contribution of the kernel is

$$
[N; -1-l] = \frac{(-1)^{N-1} N!}{(-1-l)(-1-l-1)\cdots(-1-l-N)}
$$

$$
\begin{aligned}
&= \frac{(-1)^{N-1}N!}{(-1)^{N+1}(l+1)(l+2)\cdots(l+N+1)} \\
&= \frac{N!}{(l+1)(l+2)\cdots(l+N+1)} \\
&= \frac{N!l!}{(l+N+1)!} \\
&= \frac{1}{(l+N+1)\binom{N+l}{l}},
\end{aligned}
\tag{8.11}
$$

and so from this pole we get

$$
\frac{Q^{-l}(1-Q^{-1})}{L}\frac{1}{(l+1+N)\binom{N+l}{l}}.
$$

The other poles give us (for $\varepsilon := z+1+l-\chi_k$)

$$
\begin{aligned}
f(z) &= \frac{Q^{\varepsilon-l+\chi_k}(1-Q^{-1})}{Q^{\varepsilon+\chi_k}-1} \\
&= \frac{Q^{\varepsilon-l}(1-Q^{-1})}{Q^{\varepsilon}-1} \\
&\sim \frac{Q^{-l}(1-Q^{-1})}{\varepsilon L},
\end{aligned}
$$

by (3.6) and (8.10), so the residue is $\frac{Q^{-l}(1-Q^{-1})}{L}$. The kernel can be expressed in terms of Gamma functions. If we assume that $d$ is fixed (i.e., independent of $n$), we get

$$
\begin{aligned}
[N;\chi_k-1-l] &= \frac{\Gamma(n+i-d+1)\Gamma(-\chi_k+1+l)}{\Gamma(n+i-d+1-\chi_k+1+l)} \\
&= \Gamma(-\chi_k+1+l)\frac{\Gamma(n+i-d+1)}{\Gamma(n+i-d+1-\chi_k+1+l)} \\
&\sim \Gamma(-\chi_k+1+l)n^{\chi_k-1-l} \qquad \text{(see [1, page 257])} \\
&= \frac{1}{n^{l+1}}\Gamma(-\chi_k+1+l)e^{\chi_k\log n} \\
&= \frac{1}{n^{l+1}}\Gamma(l+1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n}.
\end{aligned}
\tag{8.12}
$$

Thus we have

$$
\frac{Q^{-l}(1-Q^{-1})}{Ln^{1+l}}\sum_{k\neq 0}\Gamma(1+l-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n}
$$

for the remaining poles, and altogether the probability from (8.9) is

$$
\begin{aligned}
\sum_{i=0}^{d-1}\sum_{l=0}^{d-1-i}\binom{d-1-i}{l}(-1)^l\Bigg(&\frac{Q^{-l}(1-Q^{-1})}{L}\frac{1}{(l+1+N)\binom{N+l}{l}} \\
&+\frac{Q^{-l}(1-Q^{-1})}{Ln^{1+l}}\sum_{k\neq 0}\Gamma(1+l-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n}\Bigg).
\end{aligned}
\tag{8.13}
$$

It can now be seen that the $l = 0$ term dominates. This is because the $l = 0$ term is of order $\frac{1}{N}$, the $l = 1$ term is $O(\frac{1}{N^2})$, the $l = 2$ term is $O(\frac{1}{N^3})$ etc. Thus we need only consider this term when discussing the asymptotics. By substituting $l = 0$ in the above expression, we get (recall that $d$ is fixed)

$$\begin{aligned}
&\frac{1-Q^{-1}}{L}\sum_{i=0}^{d-1}\left(\frac{1}{N+1}+\frac{1}{n}\sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n}\right)\\
&\qquad=\frac{1-Q^{-1}}{L}\sum_{i=0}^{d-1}\frac{1}{n+i-d+1}+\frac{1-Q^{-1}}{L}\sum_{i=0}^{d-1}\frac{1}{n}\sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n}\\
&\qquad\sim\frac{(1-Q^{-1})d}{Ln}+\frac{(1-Q^{-1})d}{Ln}\sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n}. \qquad (8.14)
\end{aligned}$$

Hence

$$P_{ws}(M)\sim\frac{(1-Q^{-1})d}{Ln}(1+\delta(n)).$$

This is the asymptotic expansion of the probability in the (weak, strict) case, where $\delta(x)$ is defined in (7.5). This concludes the proof of Theorem 8.2. ■

## 8.3 Maximum in the first $d$ positions – (strict, weak)

Here we assume that $k$ only appears once in the first $d$ positions, but can occur any number of times in the rest of the word. Therefore we again use (see (8.5))

$$A_k^{(s)}=\bigcup_{i=0}^{d-1}\{1,\dots,k-1\}^i k\{1,\dots,k-1\}^{d-1-i},$$

and in total we have

$$\bigcup_{k\geq 1}A_k^{(s)}\{1,\dots,k\}^*.$$

The generating function is thus

$$\begin{aligned}
F_M^{(s,w)}(z) &:= \sum_{k\geq 1}\sum_{i=0}^{d-1}\left(\sum_{j=1}^{k-1}zpq^{j-1}\right)^i zpq^{k-1}\left(\sum_{j=1}^{k-1}zpq^{j-1}\right)^{d-1-i}\frac{1}{1-\sum\limits_{j=1}^{k}zpq^{j-1}}\\
&=\sum_{k\geq 1}\sum_{i=0}^{d-1}z^i(1-q^{k-1})^i zpq^{k-1}z^{d-1-i}(1-q^{k-1})^{d-1-i}\frac{1}{1-z(1-q^k)}\\
&=\sum_{k\geq 1}\sum_{i=0}^{d-1}\sum_{j\geq 0}z^{d+j}(1-q^{k-1})^{d-1}pq^{k-1}(1-q^k)^j
\end{aligned}$$

$$= \sum_{k\geq 1}\sum_{j\geq 0} dz^{d+j}(1-q^{k-1})^{d-1}pq^{k-1}(1-q^k)^j.$$

We are interested in the coefficient of $z^n$, hence

$$\begin{aligned}
[z^n]F_M^{(s,w)}(z) &= \sum_{k\geq 1} d(1-q^{k-1})^{d-1}pq^{k-1}(1-q^k)^{n-d} && (8.15)\\
&= \sum_{k\geq 1} d\sum_{l=0}^{d-1}\binom{d-1}{l}(-q^{k-1})^l pq^{k-1}\sum_{h=0}^{n-d}\binom{n-d}{h}(-q^k)^h\\
&= d\sum_{l=0}^{d-1}\binom{d-1}{l}(-1)^l\sum_{h=0}^{n-d}\binom{n-d}{h}(-1)^h\frac{p}{q^{l+1}}\sum_{k\geq 1}q^{k(h+l+1)}\\
&= d\sum_{l=0}^{d-1}\binom{d-1}{l}(-1)^l\sum_{h=0}^{n-d}\binom{n-d}{h}(-1)^h\frac{p}{q^{l+1}}\frac{q^{h+l+1}}{1-q^{h+l+1}}\\
&= d\sum_{l=0}^{d-1}\binom{d-1}{l}(-1)^l\sum_{h=0}^{n-d}\binom{n-d}{h}(-1)^h\frac{Q^{l+1}(1-Q^{-1})}{Q^{h+l+1}-1}. && (8.16)
\end{aligned}$$

Again, to find the alternating sum on $h$ asymptotically, we use Rice's method on the function $f(z) := \frac{Q^{l+1}(1-Q^{-1})}{Q^{z+l+1}-1}$. The simple poles we need to consider are at $z+l+1=0$ and $z+l+1=\chi_k$, and for both we define $N := n-d$. For the first, let $\varepsilon := z+l+1$, then

$$f(z) = \frac{Q^{l+1}(1-Q^{-1})}{Q^{\varepsilon}-1} = \frac{Q^{l+1}(1-Q^{-1})}{e^{\varepsilon L}-1} \sim \frac{Q^{l+1}(1-Q^{-1})}{\varepsilon L}, \tag{8.17}$$

with residue $\frac{Q^{l+1}(1-Q^{-1})}{L}$. We join this with

$$[N;-l-1] = \frac{1}{(l+N+1)\binom{N+l}{l}},$$

(see (8.11)) to get the result for the first pole as

$$\frac{Q^{l+1}(1-Q^{-1})}{L(l+N+1)\binom{N+l}{l}}.$$

For the remaining poles (where $z+l+1=\chi_k$ for all $k\neq 0$), let $\varepsilon := z+l+1-\chi_k$. Then since $Q^{\chi_k}=1$ by (3.6), the asymptotics are the same as the previous pole – see (8.17) – so

$$f(z) = \frac{Q^{l+1}(1-Q^{-1})}{Q^{\varepsilon+\chi_k}-1} = \frac{Q^{l+1}(1-Q^{-1})}{Q^{\varepsilon}-1} \sim \frac{Q^{l+1}(1-Q^{-1})}{\varepsilon L},$$

and again we have a residue of $\frac{Q^{l+1}(1-Q^{-1})}{L}$. The kernel is similar to (8.12):

$$[N;\chi_k-1-l] = \Gamma(-\chi_k+1+l)\frac{\Gamma(n-d+1)}{\Gamma(n-d+1-\chi_k+1+l)}$$

$$\begin{aligned} &\sim \Gamma(-\chi_k + 1 + l) n^{\chi_k - 1 - l} \qquad \text{(for } d \text{ fixed)} \\ &= \frac{1}{n^{l+1}} \Gamma(l + 1 - \chi_k) e^{2k\pi \boldsymbol{i} \log_Q n}. \end{aligned}$$

These poles thus contribute

$$\frac{Q^{l+1}(1 - Q^{-1})}{Ln^{l+1}} \sum_{k \neq 0} \Gamma(1 + l - \chi_k) e^{2k\pi \boldsymbol{i} \log_Q n}.$$

So to find the probability asymptotically, we add these together and put them inside the previous sums from (8.16), i.e.,

$$\begin{aligned} d \sum_{l=0}^{d-1} \binom{d-1}{l} (-1)^l \frac{Q^{l+1}(1 - Q^{-1})}{L} \\ \cdot \left( \frac{1}{(l + N + 1)\binom{N+l}{l}} + \frac{1}{n^{l+1}} \sum_{k \neq 0} \Gamma(1 + l - \chi_k) e^{2k\pi \boldsymbol{i} \log_Q n} \right). \end{aligned}$$

Again, due to the fraction $\frac{1}{(l+N+1)\binom{N+l}{l}}$, we can see that the $l = 0$ term is the largest for large $n$, and thus we can simplify the asymptotic approximation of the probability (as $n \to \infty$) to

$$\begin{aligned} &\frac{Q(1 - Q^{-1})}{L} d \left( \frac{1}{N + 1} + \frac{1}{n} \sum_{k \neq 0} \Gamma(1 - \chi_k) e^{2k\pi \boldsymbol{i} \log_Q n} \right) \\ &= \frac{Q - 1}{L} \frac{d}{n - d + 1} + \frac{(Q - 1)d}{Ln} \sum_{k \neq 0} \Gamma(1 - \chi_k) e^{2k\pi \boldsymbol{i} \log_Q n} \\ &\sim \frac{(Q - 1)d}{Ln} + \frac{(Q - 1)d}{Ln} \sum_{k \neq 0} \Gamma(1 - \chi_k) e^{2k\pi \boldsymbol{i} \log_Q n}. \end{aligned}$$

Thus

$$P_{sw}(M) \sim \frac{(Q - 1)d}{Ln} (1 + \delta(n)),$$

as $n \to \infty$, for $\delta(x)$ as defined in (7.5). The proof of Theorem 8.3 has thus been completed. ■

## 8.4 Maximum in the first $d$ positions – (weak, weak)

The requirement for this case is that there is at least one $k$ in any position between 1 and $d$. All other letters in the word can be anything in the alphabet $\{1, 2, \ldots, k\}$. We express the first $d$ letters symbolically as (refer to (8.7))

$$A_k^{(w)} = \bigcup_{i=0}^{d-1} \{1, \ldots, k - 1\}^i k \{1, \ldots, k\}^{d-1-i}.$$

Since we do not want to count $k$ more than once, we secure the position of the first $k$ in the first $d$ letters. Now the union is taken over all $k$ – i.e., over all words where $k$ is the maximum, and we include the rest of the word ($\{1,\ldots,k\}^*$), giving:

$$\bigcup_{k\geq 1} A_k^{(w)}\{1,\ldots,k\}^*.$$

In terms of generating functions, this reads as

$$\begin{aligned}
F_M^{(w,w)}(z) &:= \sum_{k\geq 1}\sum_{i=0}^{d-1}\left(\sum_{j=1}^{k-1} zpq^{j-1}\right)^i zpq^{k-1}\left(\sum_{j=1}^{k} zpq^{j-1}\right)^{d-1-i}\frac{1}{1-\sum\limits_{j=1}^{k} zpq^{j-1}}\\
&= \sum_{k\geq 1}\sum_{i=0}^{d-1} z^i(1-q^{k-1})^i zpq^{k-1}z^{d-1-i}(1-q^k)^{d-1-i}\frac{1}{1-z(1-q^k)}\\
&= \sum_{k\geq 1} z^d\sum_{i=0}^{d-1}(1-q^{k-1})^i pq^{k-1}(1-q^k)^{d-1-i}\frac{1}{1-z(1-q^k)}.
\end{aligned}$$

We are interested in a word of length $n$, and thus consider the coefficient of $z^n$ in this series.

$$\begin{aligned}
&[z^n]F_M^{(w,w)}(z)\\
&= [z^{n-d}]\sum_{k\geq 1}\sum_{i=0}^{d-1}(1-q^{k-1})^i pq^{k-1}(1-q^k)^{d-1-i}\sum_{j\geq 0}(z(1-q^k))^j\\
&= \sum_{k\geq 1}\sum_{i=0}^{d-1}(1-q^{k-1})^i pq^{k-1}(1-q^k)^{n-1-i} \qquad (8.18)\\
&= \sum_{k\geq 1}\sum_{i=0}^{d-1}\sum_{l=0}^{i}\binom{i}{l}(-q^{k-1})^l pq^{k-1}\sum_{h=0}^{n-1-i}\binom{n-1-i}{h}(-q^k)^h\\
&= \sum_{i=0}^{d-1}\sum_{l=0}^{i}\binom{i}{l}(-1)^l\frac{p}{q^{l+1}}\sum_{h=0}^{n-1-i}\binom{n-1-i}{h}(-1)^h\sum_{k\geq 1}q^{(l+1+h)k}\\
&= \sum_{i=0}^{d-1}\sum_{l=0}^{i}\binom{i}{l}(-1)^l Q^{l+1}(1-Q^{-1})\sum_{h=0}^{n-1-i}\binom{n-1-i}{h}(-1)^h\frac{1}{Q^{l+1+h}-1}. \qquad (8.19)
\end{aligned}$$

Not surprisingly, Rice's method is used to evaluate the alternating sum on $h$ asymptotically. Our contour surrounds points $0,\ldots,n-1-i$ in the complex plane. For $N := n-1-i$, we can rewrite the relevant sum as

$$\sum_{h=0}^{N}\binom{N}{h}(-1)^h\frac{1}{Q^{l+1+h}-1}.$$

So the function we consider for residue collection is $f(z) := \frac{1}{Q^{l+1+z}-1}$, and therefore the poles we deal with are at $l+1+z=0$ and $l+1+z=\chi_k$ ($\chi_k$ as before is defined for all non-zero integers $k$). If $\varepsilon := z+1+l$ then expanding around $\varepsilon = 0$ produces

$$f(z) = \frac{1}{Q^{\varepsilon}-1} = \frac{1}{e^{\varepsilon L}-1} \sim \frac{1}{1+\varepsilon L - 1} = \frac{1}{\varepsilon L},$$

and so the residue is $\frac{1}{L}$. The contribution of $[N; z]$ at $z = -1-l$ is (see (8.11))

$$[N; -l-1] = \frac{1}{(l+1+N)\binom{N+l}{l}}.$$

For the poles at $z = \chi_k - 1 - l$ we put $\varepsilon := z + l + 1 - \chi_k$ and again we have

$$f(z) \sim \frac{1}{\varepsilon L}$$

and so the residue is $\frac{1}{L}$. For the kernel (see (8.12) and recall that $d$ is independent of $n$),

$$\begin{aligned} [N; \chi_k - 1 - l] &= \Gamma(1+l-\chi_k)\frac{\Gamma(n-i)}{\Gamma(n-i-\chi_k+1+l)} \\ &\sim \Gamma(1+l-\chi_k) n^{\chi_k - 1 - l} \\ &= \frac{1}{n^{l+1}}\Gamma(1+l-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q n}. \end{aligned} \tag{8.20}$$

So for the poles at $z = \chi_k - 1 - l$, the contribution is

$$\frac{1}{Ln^{l+1}} \sum_{k \neq 0} \Gamma(1+l-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q n}.$$

Altogether the expected value is

$$\begin{aligned} &\sum_{i=0}^{d-1} \sum_{l=0}^{i} \binom{i}{l} (-1)^l Q^{l+1}(1-Q^{-1}) \\ &\qquad \cdot \left[ \frac{1}{L}\frac{1}{(l+N+1)\binom{N+l}{l}} + \frac{1}{Ln^{l+1}} \sum_{k \neq 0} \Gamma(1+l-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q n} \right] \\ &= \sum_{i=0}^{d-1} \sum_{l=0}^{i} \binom{i}{l} (-1)^l Q^{l+1}(1-Q^{-1}) \frac{1}{L}\frac{1}{(l+N+1)\binom{N+l}{l}} \\ &\quad + \sum_{i=0}^{d-1} \sum_{l=0}^{i} \binom{i}{l} (-1)^l Q^{l+1}(1-Q^{-1}) \frac{1}{Ln^{l+1}} \sum_{k \neq 0} \Gamma(1+l-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q n}. \end{aligned}$$

Again, the $l = 0$ term dominates, leaving

$$\begin{aligned} &= \sum_{i=0}^{d-1} \frac{Q-1}{L(n-i)} + \sum_{i=0}^{d-1} (Q-1) \frac{1}{Ln} \sum_{k \neq 0} \Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q n} \\ &\sim \frac{(Q-1)d}{Ln} + \frac{(Q-1)d}{Ln} \sum_{k \neq 0} \Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q n}, \end{aligned}$$

as $n \to \infty$. Therefore

$$P_{ww}(M) \sim \frac{(Q-1)d}{Ln}(1 + \delta(n)),$$

for $\delta(x)$ as in (7.5). Theorem 8.4 is now proved. ∎

It can now be seen that the cases with $k$ allowed to repeat in the 'rest' of the word have probabilities which are larger than those which are strict for the second part. As in Chapter 7, this is due to the factor $Q - 1$ as opposed to $1 - Q^{-1}$. As $n \to \infty$ the length of the 'rest' of the word dominates the first $d$ letters, and thus our results are dependent on the second restriction rather than the first. Some numerical examples (of the main term only, for $Q = 2$) are given below for a comparison:

| | (strict, strict)/(weak, strict) | (strict, weak)/(weak, weak) |
|---|---|---|
| $n = 10$, $d = 4$ | 0.28854 | 0.57708 |
| $n = 100$, $d = 4$ | 0.02885 | 0.05771 |

Table 8.1: Numerical values (main term) for the maximum in the first $d$ positions.

Also, the results in cases (strict, strict) and (weak, strict) are exact replicas of the strict case (7.1) in Chapter 7, except for the extra factor of $d$ which indicates that we have the same probability in each of the first $d$ positions. Similarly, multiplying (7.2) by $d$ gives the results in the (strict, weak) and (weak, weak) cases in this chapter, as the probability in (7.2) applies to each position of the first $d$ letters in a word, and the total probability is the sum of $d$ of these.

# Chapter 9

# The minimum in the first position

Suppose we now restrict the set of words we consider to those where the minimum value ($j$) lies in the first position. We can either allow this minimum to repeat in the rest of the word or not. If we allow repeats, we call this the 'weak' case, whereas the 'strict' case refers to the scenario when $j$ does not appear again.

**Proposition 9.1** *The probability that the minimum value in a geometrically distributed sample of length $n$ appears in the first position and nowhere else in the word is*

$$P_s(m) = \frac{Q-1}{Q^n - 1}. \tag{9.1}$$

**Proposition 9.2** *The probability that the minimum value in a geometrically distributed sample of length $n$ appears in the first position and any number of times elsewhere else in the word is*

$$P_w(m) = \frac{Q^{n-1}(Q-1)}{Q^n - 1}. \tag{9.2}$$

## 9.1 Minimum in the first position – (strict)

Here we assume that the minimum ($j$) appears only once in the word – and that it is in the first position of that word. We can express this idea symbolically as

$$\bigcup_{j\geq 1} j\{j+1, j+2, \ldots\}^*$$

where the union on $j$ signifies all possible values of the minimum (all $j \in \mathbb{N}$); the minimum in position 1 is followed by a sequence of any length but consisting only of letters whose values exceed $j$. As a generating function this can be expressed as

$$
\begin{aligned}
F_m^{(s)}(z) &:= \sum_{j\ge 1} zpq^{j-1}\frac{1}{1-\sum\limits_{k=j+1}^{\infty} zpq^{k-1}}\\
&= \sum_{j\ge 1} zpq^{j-1}\frac{1}{1-zq^j} \qquad \text{(telescoping series, } p=1-q\text{)}\\
&= \sum_{j\ge 1} zpq^{j-1}\sum_{k\ge 0} z^k (q^j)^k\\
&= \sum_{k\ge 0} z^{k+1}\frac{p}{q}\frac{q^{k+1}}{1-q^{k+1}},
\end{aligned}
$$

whose coefficient of $z^n$ is thus

$$
P_s(m) = [z^n]F_m^{(s)}(z) = \frac{p}{q}\frac{q^n}{1-q^n} = \frac{Q-1}{Q^n-1}.
$$

Here we have the result exactly, with no fluctuations (as we did in the case with maxima). This completes the proof of Proposition 9.1. ■

## 9.2 Minimum in the first position – (weak)

Now we allow the minimum to repeat – i.e., the same situation as before, only now $j$ can appear more than once in the word. Symbolically

$$
\bigcup_{j\ge 1} j\{j, j+1, \dots\}^*
$$

which translates to

$$
\begin{aligned}
F_m^{(w)}(z) &:= \sum_{j\ge 1} zpq^{j-1}\frac{1}{1-\sum\limits_{k=j}^{\infty} zpq^{k-1}}\\
&= \sum_{j\ge 1} zpq^{j-1}\frac{1}{1-zq^{j-1}}\\
&= \sum_{j\ge 1} zpq^{j-1}\sum_{k\ge 0} z^k (q^{j-1})^k\\
&= \sum_{k\ge 0} z^{k+1}\frac{p}{1-q^{k+1}}.
\end{aligned}
$$

The coefficient of $z^n$ is

$$
P_w(m) = [z^n]F_m^{(w)}(z) = \frac{p}{1-q^n} = \frac{Q^{n-1}(Q-1)}{Q^n-1}.
$$

This proves Proposition 9.2. ■

By comparing these two results, we can see that the weak case is larger (since $Q > 1$), which is to be expected since we are including more words as possibilities. Also, because the sample is geometrically distributed, it is highly unlikely that such a sample will have a strict maximum at all, let alone have it occurring in the first position.

# Chapter 10

# The minimum in the first $d$ positions

As in Chapter 8, we have four different scenarios here, since we can apply our weak/strict classification to the first $d$ letters as well as to the rest of the word. We require a word of length $d$ or more. The first case we look at is where we only allow the minimum ($j$) to appear once in the word – and its position is restricted to the first $d$ places in the word. We call this (strict, strict). Secondly we let $j$ occur more than once in the word, but only in the first $d$ places – (weak, strict). Allowing $j$ to occur in the rest of the word, but only once in the first $d$ positions, gives the case (strict, weak). Finally, we can let $j$ occur any number of times, anywhere in the word, as long as it occurs at least once in the first $d$ letters – namely (weak, weak).

The results which follow are proved in this chapter. These results are exact (which is also the case in Chapter 9).

**Proposition 10.1** *The probability that the one and only time the minimum value of a geometrically distributed sample of length $n$ occurs in the sample is in the first $d$ positions, is*

$$P_{ss}(m) = \frac{(Q-1)d}{Q^n - 1}. \tag{10.1}$$

**Proposition 10.2** *The probability that the minimum value of a geometrically distributed sample of length $n$ occurs at least once in the first $d$ positions of the sample, but never in the last $n-d$ positions is*

$$P_{ws}(m) = \frac{Q^d - 1}{Q^n - 1}. \tag{10.2}$$

**Proposition 10.3** *The probability that the minimum value of a geometrically distributed sample of length $n$ occurs once only in the first $d$ positions of the sample and any number of times elsewhere in the sample, is*

$$P_{sw}(m) = \frac{Q^{n-d}(Q-1)d}{Q^n - 1}. \tag{10.3}$$

**Proposition 10.4** *The probability that the minimum value of a geometrically distributed sample of length $n$ occurs anywhere in the sample, but at least once in the first $d$ positions, is*

$$P_{ww}(m) = \frac{1 - Q^{-d}}{1 - Q^{-n}}. \tag{10.4}$$

## 10.1 Minimum in the first $d$ positions – (strict, strict)

Suppose we have a geometrically distributed word of length $n$. What is the probability that its minimum value ($j$) appears once in the first part of the word but never again? Symbolically we express the first part (the first $d$ letters) of any such word as

$$A_j^{(s)} := \bigcup_{i=0}^{d-1} \{j+1, \ldots\}^i j \{j+1, \ldots\}^{d-1-i} \tag{10.5}$$

and so all possible words of this type could be expressed as

$$\bigcup_{j \geq 1} A_j^{(s)} \{j+1, \ldots\}^*.$$

The generating function is

$$\begin{aligned}
F_m^{(s,s)}(z) &:= \sum_{j\geq 1} \sum_{i=0}^{d-1} \left( \sum_{k=j+1}^{\infty} zpq^{k-1} \right)^i zpq^{j-1} \left( \sum_{k=j+1}^{\infty} zpq^{k-1} \right)^{d-1-i} \frac{1}{1 - \sum\limits_{k=j+1}^{\infty} zpq^{k-1}} \\
&= \sum_{j\geq 1} \sum_{i=0}^{d-1} z^i (q^j)^i zpq^{j-1} z^{d-1-i} (q^j)^{d-1-i} \frac{1}{1 - zq^j} \\
&= \sum_{j\geq 1} \sum_{i=0}^{d-1} z^d pq^{jd-1} \sum_{k\geq 0} z^k q^{jk} \\
&= \sum_{j\geq 1} \sum_{k\geq 0} dz^{k+d} pq^{j(k+d)-1},
\end{aligned}$$

and thus the probability that the strict minimum occurs only once in the first $d$ positions is

$$[z^n] F_m^{(s,s)}(z) = \sum_{j\geq 1} dpq^{jn-1}$$

$$= d\frac{p}{q}\sum_{j\geq 1} q^{jn}$$

$$= d\frac{p}{q}\frac{q^n}{1-q^n}.$$

Hence

$$P_{ss}(m) = \frac{(Q-1)d}{Q^n - 1},$$

and Proposition 10.1 is proved. ■

This result is the same as the minimum in the first position (strict, see Chapter 9), but multiplied by $d$, as there are $d$ opportunities for $j$ to occur somewhere in the first $d$ positions.

## 10.2 Minimum in the first $d$ positions – (weak, strict)

Suppose we allow $j$ to recur within the first $d$ but not thereafter. We express the first $d$ letters as

$$A_j^{(w)} := \bigcup_{i=0}^{d-1} \{j+1,\dots\}^i j \{j, j+1, \dots\}^{d-1-i}, \tag{10.6}$$

where the position of the first $j$ is fixed. This can be substituted into the symbolic equation for all such words, namely

$$\bigcup_{j\geq 1} A_j^{(w)} \{j+1,\dots\}^*.$$

Then we can define the generating function as

$$F_m^{(w,s)}(z) := \sum_{j\geq 1}\sum_{i=0}^{d-1}\left(\sum_{k=j+1}^{\infty} zpq^{k-1}\right)^i zpq^{j-1}\left(\sum_{k=j}^{\infty} zpq^{k-1}\right)^{d-1-i} \frac{1}{1-\sum\limits_{k=j+1}^{\infty} zpq^{k-1}}$$

$$= \sum_{j\geq 1}\sum_{i=0}^{d-1} z^i (q^j)^i zpq^{j-1} z^{d-1-i}(q^{j-1})^{d-1-i}\frac{1}{1-zq^j}$$

$$= \sum_{j\geq 1}\sum_{i=0}^{d-1} z^d pq^{jd-d+i}\sum_{k\geq 0} z^k q^{jk}$$

$$= \sum_{k\geq 0} z^{k+d}pq^{-d}\sum_{j\geq 1} q^{j(k+d)}\sum_{i=0}^{d-1} q^i$$

$$= \sum_{k\geq 0} z^{k+d}pq^{-d}\frac{q^{k+d}}{1-q^{k+d}}\frac{1-q^d}{1-q},$$

and thus

$$[z^n]F_m^{(w,s)}(z) = pq^{-d}\frac{q^n}{1-q^n}\frac{1-q^d}{1-q}$$
$$= p\frac{1}{q^{-n}-1}\frac{q^{-d}-1}{1-q}.$$

Consequently,

$$P_{ws}(m) = \frac{Q^d - 1}{Q^n - 1},$$

which completes the proof of Proposition 10.2. $\blacksquare$

## 10.3 Minimum in the first $d$ positions – (strict, weak)

If we now allow $j$ to appear only once in the first $d$ letters of a geometrically distributed sample of length $n$, but any number of times in the rest, then (from (10.5))

$$A_j^{(s)} = \bigcup_{i=0}^{d-1}\{j+1,\ldots\}^i j\{j+1,\ldots\}^{d-1-i},$$

as in the (strict, strict) case, but altogether we include a $j$ in the starred sequence to get

$$\bigcup_{j\geq 1} A_j^{(s)}\{j, j+1,\ldots\}^*.$$

The generating function is

$$F_m^{(s,w)}(z) := \sum_{j\geq 1}\sum_{i=0}^{d-1}\left(\sum_{k=j+1}^{\infty} zpq^{k-1}\right)^i zpq^{j-1}\left(\sum_{k=j+1}^{\infty} zpq^{k-1}\right)^{d-1-i}\frac{1}{1-\sum\limits_{k=j}^{\infty} zpq^{k-1}}$$
$$= \sum_{j\geq 1}\sum_{i=0}^{d-1} z^i(q^j)^i zpq^{j-1}z^{d-1-i}(q^j)^{d-1-i}\frac{1}{1-zq^{j-1}}$$
$$= \sum_{j\geq 1}\sum_{i=0}^{d-1} z^d pq^{jd-1}\sum_{k\geq 0} z^k q^{k(j-1)}$$
$$= \sum_{k\geq 0} dz^{k+d}pq^{-k-1}\sum_{j\geq 1} q^{j(k+d)}$$
$$= \sum_{k\geq 0} dz^{k+d}pq^{-k-1}\frac{q^{k+d}}{1-q^{k+d}},$$

so the probability is

$$[z^n]F_m^{(s,w)}(z) = dpq^{-(n-d)-1}\frac{q^n}{1-q^n}$$

$$
\begin{aligned}
&= dpq^{d-1}\frac{1}{1-q^n} \\
&= d(1-Q^{-1})Q^{1-d}\frac{1}{1-Q^{-n}}.
\end{aligned}
$$

Hence

$$P_{sw}(m) = \frac{Q^{n-d}(Q-1)d}{Q^n-1},$$

which concludes the proof of Proposition 10.3. ■

## 10.4 Minimum in the first $d$ positions – (weak, weak)

Now we consider the final option. We let $j$ occur anywhere, as long as it appears at least once in the first $d$ places.

$$A_j^{(w)} = \bigcup_{i=0}^{d-1} \{j+1,\ldots\}^i j\{j,j+1,\ldots\}^{d-1-i},$$

as in (10.6) where the first $j$'s position is fixed, but others may occur to the right. All such words are represented by

$$\bigcup_{j\geq 1} A_j^{(w)}\{j,j+1,\ldots\}^*.$$

The generating function is

$$
\begin{aligned}
F_m^{(w,w)}(z) &:= \sum_{j\geq 1}\sum_{i=0}^{d-1}\left(\sum_{k=j+1}^{\infty} zpq^{k-1}\right)^i zpq^{j-1}\left(\sum_{k=j}^{\infty} zpq^{k-1}\right)^{d-1-i}\frac{1}{1-\sum\limits_{k=j}^{\infty} zpq^{k-1}} \\
&= \sum_{j\geq 1}\sum_{i=0}^{d-1} z^i(q^j)^i zpq^{j-1}z^{d-1-i}(q^{j-1})^{d-1-i}\frac{1}{1-zq^{j-1}} \\
&= \sum_{j\geq 1}\sum_{i=0}^{d-1} z^d pq^{jd-d+i}\sum_{k\geq 0} z^k q^{k(j-1)} \\
&= \sum_{k\geq 0} z^{k+d}pq^{-d-k}\sum_{j\geq 1} q^{j(k+d)}\sum_{i=0}^{d-1} q^i \\
&= \sum_{k\geq 0} z^{k+d}pq^{-d-k}\frac{q^{k+d}}{1-q^{k+d}}\frac{1-q^d}{1-q} \\
&= \sum_{k\geq 0} z^{k+d}\frac{1-q^d}{1-q^{k+d}},
\end{aligned}
$$

and so in this case the coefficient of $z^n$ (i.e., the probability) is

$$P_{ww}(m) = [z^n]F_m^{(w,w)}(z) = \frac{1-q^d}{1-q^n} = \frac{1-Q^{-d}}{1-Q^{-n}}.$$

This proves Proposition 10.4. ■

It can be seen by substituting even small values for $n$ and $d$, that these different situations are what we would expect. Again, for a fixed $d$ and large $n$, we can see that it is the classification of the 'rest' of the word that takes precedence, i.e., the (strict, strict) and (weak, strict) cases are in a different order of magnitude to the (strict, weak) and (weak, weak) cases. For example, if we take $Q = 2$, we get:

| | (strict, strict) | (weak, strict) | (strict, weak) | (weak, weak) |
|---|---|---|---|---|
| $n = 10$, $d = 4$ | 0.00391 | 0.01466 | 0.25024 | 0.93842 |
| $n = 100$, $d = 4$ | $3.1554 * 10^{-30}$ | $1.183 * 10^{-29}$ | 0.25 | 0.9375 |

Table 10.1: Numerical values for the minimum in the first $d$ positions.

# Chapter 11

# The minimum of the first *d* is the maximum of the rest

We require the minimum value ($j$, possibly repeated) of the first $d$ letters to be either strictly greater than or greater than or equal to the maximum of the rest of the word. Again there are four cases, which are all combinations of the pair (strict, weak).

The theorems below are proved in this chapter.

**Theorem 11.1** *The probability that the strict (i.e., occurs only once) minimum value of the first d positions is the strict maximum value of the remaining letters in a geometrically distributed sample of length n is asymptotic to*

$$P_{ss}(mM) \sim \frac{(Q-1)d!}{Ln^dQ^d} + \frac{(Q-1)d}{Ln^dQ^d}\delta_d(n), \tag{11.1}$$

*as $n \to \infty$, where*

$$\delta_d(x) = \sum_{k\neq 0} \Gamma(d-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q x}, \tag{11.2}$$

*which is defined in (11.10).*

**Theorem 11.2** *The probability that the weak (i.e., possibly repeated) minimum value of the first d positions is the strict maximum value of the remaining letters in an n-letter geometrically distributed sample is approximately*

$$P_{ws}(mM) \sim \frac{(Q^d-1)(d-1)!}{Ln^dQ^d} + \frac{(Q^d-1)}{Ln^dQ^d}\delta_d(n), \tag{11.3}$$

*as $n \to \infty$, where $\delta_d(x) = \sum_{k\neq 0} \Gamma(d-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q x}$ as in (11.10).*

**Theorem 11.3** *In a geometrically distributed sample of length $n$, the probability that the strict minimum value of the first $d$ letters is the weak maximum value of the remaining letters is asymptotic (as $n \to \infty$) to*

$$P_{sw}(mM) \sim \frac{(Q-1)d!}{Ln^d} + \frac{(Q-1)d}{Ln^d}\delta_d(n), \tag{11.4}$$

*where $\delta_d(x)$ is defined in (11.10).*

**Theorem 11.4** *The asymptotic probability that the weak minimum value of the first $d$ positions is the same as the weak maximum value of the rest in a geometrically distributed sample of length $n$ is*

$$P_{ww}(mM) \sim \frac{(Q^d-1)(d-1)!}{Ln^d} + \frac{Q^d-1}{Ln^d}\delta_d(n), \tag{11.5}$$

*as $n \to \infty$ for $\delta_d(x)$ as in (11.10).*

## 11.1 Minimum of first $d$ is greater than the rest – (strict, strict)

Here we do not allow $j$ to occur anywhere else in the word. That is, $j$ occurs only once in the first $d$ letters, and is strictly the minimum there. It does not occur in the rest of the word either, and nor does any letter larger than $j$.

We can express the first $d$ letters symbolically as we did in Chapter 10 (see (10.5))

$$A_j^{(s)} = \bigcup_{i=0}^{d-1} \{j+1, \ldots\}^i j \{j+1, \ldots\}^{d-1-i},$$

so that our total symbolic equation is

$$\bigcup_{j \geq 1} A_j^{(s)} \{1, \ldots, j-1\}^*.$$

This can be translated into the generating function

$$\begin{aligned}
F_d^{(s,s)}(z) &:= \sum_{j\geq 1}\sum_{i=0}^{d-1}\left(\sum_{h=j+1}^{\infty} zpq^{h-1}\right)^i zpq^{j-1}\left(\sum_{h=j+1}^{\infty} zpq^{h-1}\right)^{d-1-i} \frac{1}{1-\sum\limits_{h=1}^{j-1} zpq^{h-1}} \\
&= \sum_{j\geq 1}\sum_{i=0}^{d-1} z^i (q^j)^i zpq^{j-1} z^{d-1-i} (q^j)^{d-1-i} \frac{1}{1-z(1-q^{j-1})} \\
&= \sum_{j\geq 1}\sum_{i=0}^{d-1} z^d pq^{ij+j-1+jd-j-ij} \sum_{h\geq 0} z^h (1-q^{j-1})^h
\end{aligned}$$

$$= \sum_{h\geq 0}\sum_{j\geq 1} dz^{d+h}pq^{jd-1}(1-q^{j-1})^h.$$

We wish to examine the coefficient of $z^n$:

$$\begin{aligned}
[z^n]F_d^{(s,s)}(z) &= \sum_{j\geq 1} dpq^{jd-1}(1-q^{j-1})^{n-d} \\
&= \sum_{j\geq 1} dpq^{jd-1}\sum_{l=0}^{n-d}\binom{n-d}{l}(-q^{j-1})^l \\
&= \sum_{l=0}^{n-d}\binom{n-d}{l}(-1)^l dpq^{-l-1}\sum_{j\geq 1} q^{j(d+l)} \\
&= \sum_{l=0}^{n-d}\binom{n-d}{l}(-1)^l dpq^{-l-1}\frac{q^{d+l}}{1-q^{d+l}} \\
&= d\sum_{l=0}^{n-d}\binom{n-d}{l}(-1)^l \frac{Q^l(Q-1)}{Q^{d+l}-1} \\
&= \frac{(Q-1)d}{Q^d}\sum_{l=0}^{n-d}\binom{n-d}{l}(-1)^l \frac{Q^{d+l}}{Q^{d+l}-1}. \qquad (11.6)
\end{aligned}$$

We can now make use of Rice's method (the contour surrounds points $0, 1, \ldots, n-d$) to evaluate this asymptotically. We consider the function $f(z) := \frac{Q^{d+z}}{Q^{d+z}-1}$. The poles about which we must expand are the simple poles at $z+d = 0$ and $z+d = \chi_k$ where $k \neq 0$. If $\varepsilon := z + d$, the former gives us

$$f(z) = \frac{Q^\varepsilon}{Q^\varepsilon - 1} = \frac{e^{\varepsilon L}}{e^{\varepsilon L}-1} \sim \frac{1}{\varepsilon L}, \qquad (11.7)$$

so the residue is $\frac{1}{L}$. The kernel is

$$\begin{aligned}
[n-d; -d] &= \frac{(-1)^{n-d-1}(n-d)!}{(-d)(-d-1)\cdots(-d-(n-d))} \\
&= \frac{(-1)^{n-d-1}(n-d)!}{(-d)(-(d+1))\cdots(-n)} \\
&= \frac{(-1)^{n-d-1}(n-d)!}{(-1)^{n-d-1}(d)(d+1)\cdots(n)} \\
&= \frac{(-1)^{n-d-1}(n-d)!}{(-1)^{n-d-1}(d)(d+1)\cdots(n)} \\
&= \frac{(n-d)!(d-1)!}{n!} \\
&= \frac{1}{d\binom{n}{d}}.
\end{aligned}$$

Since we are interested in $n$ large, we can approximate this (as $n \to \infty$, see [1, page 257]) to:

$$[n-d; -d] = \frac{1}{d\binom{n}{d}} = \frac{d!(n-d)!}{dn!} = \frac{(d-1)!\Gamma(n-d+1)}{\Gamma(n+1)} \sim \frac{(d-1)!}{n^d}. \qquad (11.8)$$

So for the simple pole at $z+d=0$, we need to multiply by the $\frac{(Q-1)d}{Q^d}$ in (11.6) to get

$$\frac{(Q-1)d!}{Ln^dQ^d}.$$

For the other simple poles (at $z+d=\chi_k$) let $\varepsilon := z+d-\chi_k$, then since $Q^{\chi_k}=1$ by (3.6) we have

$$f(z) = \frac{Q^{\varepsilon+\chi_k}}{Q^{\varepsilon+\chi_k}-1} = \frac{Q^{\varepsilon}}{Q^{\varepsilon}-1} \sim \frac{1}{\varepsilon L}$$

giving a residue of $\frac{1}{L}$ as in (11.7). This time the kernel is

$$\begin{aligned} [n-d;\chi_k-d] &= \frac{\Gamma(d-\chi_k)\Gamma(n-d+1)}{\Gamma(n+1-\chi_k)} \\ &\sim \Gamma(d-\chi_k)n^{\chi_k-d} \\ &= \frac{1}{n^d}\Gamma(d-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n}, \end{aligned} \tag{11.9}$$

and thus we get the contribution of the remaining poles as (see (11.6))

$$\frac{(Q-1)d}{Ln^dQ^d}\sum_{k\neq 0}\Gamma(d-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n}.$$

Adding these together will give the probability as asymptotic to

$$P_{ss}(mM) \sim \frac{(Q-1)d!}{Ln^dQ^d} + \frac{(Q-1)d}{Ln^dQ^d}\delta_d(n),$$

as $n\to\infty$, where

$$\delta_d(x) := \sum_{k\neq 0}\Gamma(d-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q x}. \tag{11.10}$$

This concludes the proof of Theorem 11.1. ■

## 11.2 Minimum of first $d$ is greater than the rest – (weak, strict)

Here we allow $j$ to appear any number of times (and at least once) in the first $d$ letters, but it is not allowed to appear in the rest of the word. From (10.6) we have

$$A_j^{(w)} = \bigcup_{i=0}^{d-1}\{j+1,\ldots\}^i j\{j,\ldots\}^{d-1-i},$$

then all possible words would be

$$\bigcup_{j\geq 1} A_j^{(w)}\{1,\ldots,j-1\}^*.$$

Translating this into a generating function gives

$$
\begin{aligned}
F_d^{(w,s)}(z) &:= \sum_{j\ge 1}\sum_{i=0}^{d-1}\left(\sum_{h=j+1}^{\infty} zpq^{h-1}\right)^{i} zpq^{j-1}\left(\sum_{h=j}^{\infty} zpq^{h-1}\right)^{d-1-i}\frac{1}{1-\sum\limits_{h=1}^{j-1} zpq^{h-1}} \\
&= \sum_{j\ge 1}\sum_{i=0}^{d-1} z^i (q^j)^i zpq^{j-1} z^{d-1-i}(q^{j-1})^{d-1-i}\frac{1}{1-z(1-q^{j-1})} \\
&= \sum_{j\ge 1}\sum_{i=0}^{d-1} z^d pq^{jd-d+i}\sum_{h\ge 0} z^h (1-q^{j-1})^h \\
&= \sum_{j\ge 1}\sum_{h\ge 0} z^{h+d} pq^{jd-d}(1-q^{j-1})^h \sum_{i=0}^{d-1} q^i \\
&= \sum_{j\ge 1}\sum_{h\ge 0} z^{h+d} pq^{jd-d}(1-q^{j-1})^h \frac{1-q^d}{1-q} \\
&= (q^{-d}-1)\sum_{j\ge 1}\sum_{h\ge 0} z^{h+d} q^{jd}(1-q^{j-1})^h,
\end{aligned}
$$

so

$$
\begin{aligned}
[z^n]F_d^{(w,s)}(z) &= (q^{-d}-1)\sum_{j\ge 1} q^{jd}(1-q^{j-1})^{n-d} \\
&= (q^{-d}-1)\sum_{j\ge 1} q^{jd}\sum_{l=0}^{n-d}\binom{n-d}{l}(-q^{j-1})^l \\
&= (q^{-d}-1)\sum_{l=0}^{n-d}\binom{n-d}{l}(-1)^l q^{-l}\sum_{j\ge 1} q^{j(d+l)} \\
&= (q^{-d}-1)\sum_{l=0}^{n-d}\binom{n-d}{l}(-1)^l q^{-l}\frac{q^{d+l}}{1-q^{d+l}} \\
&= (1-q^d)\sum_{l=0}^{n-d}\binom{n-d}{l}(-1)^l \frac{1}{1-q^{d+l}} \\
&= \sum_{l=0}^{n-d}\binom{n-d}{l}(-1)^l \frac{1-Q^{-d}}{1-Q^{-d-l}} \\
&= \sum_{l=0}^{n-d}\binom{n-d}{l}(-1)^l \frac{Q^{d+l}(1-Q^{-d})}{Q^{d+l}-1} \\
&= \frac{Q^d-1}{Q^d}\sum_{l=0}^{n-d}\binom{n-d}{l}(-1)^l \frac{Q^{d+l}}{Q^{d+l}-1}.
\end{aligned}
$$

Rice's method tells us to examine $f(z) := \frac{Q^{d+z}}{Q^{d+z}-1}$ at $z+d=0$ and $z+d=\chi_k$, so we expand about $\varepsilon := z+d$ to get (see (11.7))

$$f(z) = \frac{Q^{\varepsilon}}{Q^{\varepsilon}-1} \sim \frac{1}{\varepsilon L}$$

so we end up with $\frac{1}{L}$. From equation (11.8)

$$[n-d;-d] \sim \frac{(d-1)!}{n^d}.$$

For $\varepsilon := z + d - \chi_k$, from (3.6) and (11.7)

$$f(z) = \frac{Q^{\varepsilon+\chi_k}}{Q^{\varepsilon+\chi_k}-1} \sim \frac{1}{\varepsilon L}$$

with residue $\frac{1}{L}$, and (from (11.9))

$$[n-d;\chi_k-d] \sim \frac{1}{n^d}\Gamma(d-\chi_k)e^{2k\pi \boldsymbol{i} \log_Q n}.$$

As a result we get

$$P_{ws}(mM) = [z^n]F_d^{(w,s)}(z) \sim \frac{(Q^d-1)(d-1)!}{Ln^dQ^d} + \frac{(Q^d-1)}{Ln^dQ^d}\delta_d(n),$$

(as $n \to \infty$) for $\delta_d(x) = \sum_{k\neq 0}\Gamma(d-\chi_k)e^{2k\pi \boldsymbol{i} \log_Q x}$ as in (11.10). This completes the proof of Theorem 11.2. ∎

## 11.3 Minimum of first $d$ is greater than or equal to the rest – (strict, weak)

If we now consider a word in which $j$ is the strict minimum of the first $d$ letters, but that we allow any of $1, 2, \ldots, j$ in the rest of the word then we have the same $A_j^{(s)}$ as in (10.5), namely

$$A_j^{(s)} = \bigcup_{i=0}^{d-1}\{j+1,\ldots\}^i j\{j+1,\ldots\}^{d-1-i},$$

which is part of the overall symbolic equation

$$\bigcup_{j\geq 1} A_j^{(s)}\{1,\ldots,j\}^*.$$

This translates into the generating function

$$F_d^{(s,w)}(z) := \sum_{j\geq 1}\sum_{i=0}^{d-1}\left(\sum_{h=j+1}^{\infty} zpq^{h-1}\right)^i zpq^{j-1}\left(\sum_{h=j+1}^{\infty} zpq^{h-1}\right)^{d-1-i}\frac{1}{1-\sum\limits_{h=1}^{j} zpq^{h-1}}$$

$$= \sum_{j\geq 1}\sum_{i=0}^{d-1} z^i(q^j)^i zpq^{j-1}z^{d-1-i}(q^j)^{d-1-i}\frac{1}{1-z(1-q^j)}$$

$$= \sum_{j\geq 1}\sum_{i=0}^{d-1} z^d pq^{jd-1} \sum_{h\geq 0} z^h (1-q^j)^h$$
$$= \sum_{h\geq 0}\sum_{j\geq 1} dz^{d+h} pq^{jd-1}(1-q^j)^h,$$

of which the coefficient is

$$\begin{aligned}
[z^n]F_d^{(s,w)}(z) &= \sum_{j\geq 1} dpq^{jd-1}(1-q^j)^{n-d} \\
&= \sum_{j\geq 1} dpq^{jd-1} \sum_{l=0}^{n-d} \binom{n-d}{l} (-q^j)^l \\
&= \sum_{l=0}^{n-d} \binom{n-d}{l} (-1)^l dpq^{-1} \sum_{j\geq 1} q^{j(d+l)} \\
&= \sum_{l=0}^{n-d} \binom{n-d}{l} (-1)^l dpq^{-1} \frac{q^{d+l}}{1-q^{d+l}} \\
&= \sum_{l=0}^{n-d} \binom{n-d}{l} (-1)^l d \frac{Q-1}{Q^{d+l}-1} \\
&= (Q-1)d \sum_{l=0}^{n-d} \binom{n-d}{l} (-1)^l \frac{1}{Q^{d+l}-1}.
\end{aligned}$$

We consider the poles of $f(z) := \frac{1}{Q^{d+z}-1}$ at $z+d=0$ and at $z+d=\chi_k$. Expanding around $\varepsilon = 0$ (where $\varepsilon := z+d$) gives

$$f(z) = \frac{1}{Q^\varepsilon - 1} = \frac{1}{e^{\varepsilon L} - 1} \sim \frac{1}{\varepsilon L}, \tag{11.11}$$

with residue $\frac{1}{L}$. As in the (strict, strict) case, see equation (11.8),

$$[n-d, -d] \sim \frac{(d-1)!}{n^d}.$$

For $\varepsilon := z + d - \chi_k$, assuming (3.6) and (11.11), we have

$$f(z) = \frac{1}{Q^{\varepsilon+\chi_k} - 1} \sim \frac{1}{\varepsilon L}$$

so the residue is also $\frac{1}{L}$, and (as in (11.9))

$$[n-d; \chi_k - d] \sim \frac{1}{n^d}\Gamma(d-\chi_k)e^{2k\pi \boldsymbol{i} \log_Q n}.$$

Altogether as $n \to \infty$, the probability is given by

$$P_{sw}(mM) \sim \frac{(Q-1)d!}{Ln^d} + \frac{(Q-1)d}{Ln^d}\delta_d(n),$$

where $\delta_d(x) = \sum\limits_{k\neq 0} \Gamma(d-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q x}$. This proves Theorem 11.3. ■

## 11.4 Minimum of first $d$ is greater than or equal to the rest – (weak, weak)

We now consider the case where $j$ can appear more than once in the first $d$ letters, and can also appear any number of times in the rest of the word. So we have a word whose first $d$ letters contain anything from $\{j, j+1, \ldots\}$, with at least one $j$, and the remainder of the word consists of anything from $\{1, 2, \ldots, j\}$. We start with (see (10.6))

$$A_j^{(w)} = \bigcup_{i=0}^{d-1} \{j+1, \ldots\}^i j \{j, \ldots\}^{d-1-i},$$

to get all such words symbolised by:

$$\bigcup_{j\geq 1} A_j^{(w)} \{1, \ldots, j\}^*.$$

The generating function is thus

$$\begin{aligned}
F_d^{(w,w)}(z) &:= \sum_{j\geq 1}\sum_{i=0}^{d-1}\left(\sum_{h=j+1}^{\infty} zpq^{h-1}\right)^i zpq^{j-1}\left(\sum_{h=j}^{\infty} zpq^{h-1}\right)^{d-1-i} \frac{1}{1-\sum\limits_{h=1}^{j} zpq^{h-1}} \\
&= \sum_{j\geq 1}\sum_{i=0}^{d-1} z^i (q^j)^i zpq^{j-1} z^{d-1-i} (q^{j-1})^{d-1-i} \frac{1}{1-z(1-q^j)} \\
&= \sum_{j\geq 1}\sum_{i=0}^{d-1} z^d pq^{jd-d+i} \sum_{h\geq 0} z^h (1-q^j)^h \\
&= \sum_{h\geq 0} z^{d+h} \sum_{j\geq 1} pq^{jd-d}(1-q^j)^h \sum_{i=0}^{d-1} q^i \\
&= \sum_{h\geq 0} z^{d+h} \sum_{j\geq 1} pq^{jd-d}(1-q^j)^h \frac{1-q^d}{1-q} \\
&= \sum_{h\geq 0} z^{d+h} \sum_{j\geq 1} q^{jd-d}(1-q^j)^h (1-q^d).
\end{aligned}$$

The probability is then

$$\begin{aligned}
[z^n]F_d^{(w,w)}(z) &= \sum_{j\geq 1} q^{jd-d}(1-q^j)^{n-d}(1-q^d) \\
&= (1-q^d)\sum_{j\geq 1} q^{jd-d} \sum_{l=0}^{n-d} \binom{n-d}{l} (-q^j)^l \\
&= (1-q^d)q^{-d} \sum_{l=0}^{n-d} \binom{n-d}{l} (-1)^l \sum_{j\geq 1} q^{jl+jd}
\end{aligned}$$

$$
\begin{aligned}
&= (1-q^d)q^{-d}\sum_{l=0}^{n-d}\binom{n-d}{l}(-1)^l\frac{q^{l+d}}{1-q^{l+d}}\\
&= \sum_{l=0}^{n-d}\binom{n-d}{l}(-1)^l\frac{Q^d-1}{Q^{l+d}-1}\\
&= (Q^d-1)\sum_{l=0}^{n-d}\binom{n-d}{l}(-1)^l\frac{1}{Q^{l+d}-1}.
\end{aligned}
$$

We now find the alternating sum asymptotically by looking at $f(z) := \frac{1}{Q^{z+d}-1}$ at $z+d=0$ and $z+d=\chi_k$. Using (11.11), we define $\varepsilon := z+d$ to get

$$f(z) = \frac{1}{Q^\varepsilon - 1} = \frac{1}{e^{\varepsilon L}-1} \sim \frac{1}{\varepsilon L},$$

so the contribution is $\frac{1}{L}$, and for $\varepsilon := z+d-\chi_k$ (by (3.6)),

$$f(z) = \frac{1}{Q^{\varepsilon+\chi_k}-1} = \frac{1}{Q^\varepsilon - 1} = \frac{1}{e^{\varepsilon L}-1} \sim \frac{1}{\varepsilon L},$$

with the same contribution. Again the kernels are

$$[n-d; -d] \sim \frac{(d-1)!}{n^d}$$

from (11.8) and

$$[n-d; \chi_k - d] \sim \frac{1}{n^d}\Gamma(d-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n},$$

from (11.9). So the probability in the final (weak, weak) case tends to

$$P_{ww}(mM) \sim \frac{(Q^d-1)(d-1)!}{Ln^d} + \frac{Q^d-1}{Ln^d}\delta_d(n)$$

as $n \to \infty$ if $\delta_d(x) = \sum_{k\neq 0}\Gamma(d-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q x}$. Theorem 11.4 is proved. ■

# Chapter 12

# The maximum in the first $d$ positions for $d = \alpha n$

Here, as in Chapter 8, we consider the probability of having the maximum in the first $d$ positions. This time, however, we suppose that $d$ grows linearly with $n$: define $d = \alpha n$, where $0 < \alpha \leq 1$. Do the previous results still hold? In the (strict, strict) case for the maximum, we had a dominant term of $\frac{pd}{Ln}$ when $d$ was fixed, and so by replacing the $d$ with $\alpha n$ gives a probability of $\frac{p\alpha}{L}$. If $\alpha = 1$, then $d = n$ so we have split the word into a first part of $n$ letters and an empty second part. This is equivalent to considering the probability of having a single maximum anywhere in a geometrically distributed sample of size $n$. Our result is then $\frac{p}{L}$, which corresponds to an equivalent result in [18, page 3], where the probability is found of having a single winner (highest value) in a sample of players who each toss a coin until a head is obtained.

If $d$ were independent of $n$ then the (strict, weak) case for the maximum should give the same result. This result was $\frac{pd}{qLn}$ which becomes $\frac{p}{qL}$ for $d = n$. We have an extra factor of $q$ in the denominator, implying that $d$ does indeed depend on $n$.

In both the (weak, weak) and (weak, strict) cases we would expect a probability of 1 for $d = n$, as we are considering the probability that the maximum in a sample can occur any number of times (i.e., there are no restrictions on the word). The two results are again $\frac{p}{L}$ and $\frac{p}{qL}$, which are not equal to 1 in general.

The discussion above suggests that the previous results do not hold for at least three of the cases if $d$ is proportional to $n$. This is true, and it is due to the limit as $n \to \infty$. If we consider the (weak, strict) case of the maximum (see Chapter 8), we see that the expected value expression involves two sums on $i$ and $l$ respectively. For $d$ fixed the $l = 0$ term was dominant. This can be seen by substituting a few values of $l$ into the dominant term of equation (8.13) to see that for $d$ independent of $n$, each of these is of order $O(n^{-(l+1)})$. If however $d = \alpha n$, then this is not the case (due to the presence of the $d$ and the fact that both the upper limits of the sums depend on $d$, all of which introduce more $n$'s into the expression) and each term is of order $O(\frac{1}{n})$. So the $l = 0$ term does not necessarily dominate and the rest of the argument will not be valid. Also, towards the end of the calculations, several approximations are made which assume that $d$ is independent of $n$. (For example, see the asymptotics (8.12) and (8.14) in the (weak, strict) case.) Only in the (strict, strict) case is this not a problem, as our dominant term is not troublesome and only involves a single sum on $i$ of a summand independent of $i$. Also, in this case, the asymptotics for $n$ large do not depend on $d$. For the other three cases we make use of a different technique from complex analysis called the 'Mellin' transform. Why? "Harmonic sums surface recurrently in the context of analytic combinatorics and Mellin transforms are a method of choice for coping with them" ([11, page 575]).

Mellin transforms were discussed in more detail in Chapter 6 and allow us to transform a function (say $f(x)$) into an integral which exists in a strip in the complex plane, determined by the behaviour of $f(x)$ as $x \to 0^+$ and $x \to \infty$. We write

$$f^*(s) := \int_0^\infty f(x)x^{s-1}dx$$

which exists for $s$ in $\langle -u, -v \rangle$, see (6.1). Then the inversion formula can be invoked, to give a contour integral on the complex variable $s$ which can be approximated by residue calculus. As in (6.2)

$$f(x) = \frac{1}{2\pi \boldsymbol{i}} \int_{c-\boldsymbol{i}\infty}^{c+\boldsymbol{i}\infty} f^*(s)x^{-s}ds$$

for $-u < c < -v$. We denote this inversion with a shorthand of (see (6.3))

$$f(x) = \frac{1}{2\pi \boldsymbol{i}} \int_{(c)} f^*(s)x^{-s}ds.$$

The results proved in this chapter are stated below and use the assumption that $d = \alpha n$ for $0 < \alpha \leq 1$.

**Theorem 12.1** *The probability that the maximum of a geometrically distributed sample occurs at least once in the first $d$ letters, and never again is*

$$P_{ws}^{\alpha n}(M) \sim \frac{1}{L}\log\left(\frac{1}{1-\alpha(1-Q^{-1})}\right) + \frac{1}{L}\big(\delta_0(n(1-\alpha p)) - \delta_0(n)\big), \tag{12.1}$$

*as $n \to \infty$, where*

$$\delta_0(x) = \sum_{k\neq 0} \Gamma(-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q x},$$

*as defined in (12.12).*

**Theorem 12.2** *The probability that the maximum of a geometrically distributed sample occurs only once in the first $d$ letters, but any number of times in the rest of the word is*

$$P_{sw}^{\alpha n}(M) \sim \frac{\alpha(Q-1)}{L(1+\alpha(Q-1))}\big(1+\delta(n(q+p\alpha))\big), \tag{12.2}$$

*as $n \to \infty$, where*

$$\delta(x) = \sum_{k\neq 0} \Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q x},$$

*from (7.5).*

**Theorem 12.3** *The probability that the maximum of a geometrically distributed sample occurs at least once in the first $d$ letters, and any number of times in the rest of the word is*

$$P_{ww}^{\alpha n}(M) \sim \frac{\log(1+\alpha(Q-1))}{L} + \frac{1}{L}\left(\delta_0(n) - \delta_0\left(n\Big(\frac{q+\alpha p}{q}\Big)\right)\right), \tag{12.3}$$

*as $n \to \infty$, where*

$$\delta_0(x) = \sum_{k\neq 0} \Gamma(-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q x},$$

*as in (12.12).*

## 12.1 Case (weak, strict), for $d = \alpha n$

Suppose $d = \alpha n$ for $0 < \alpha \leq 1$ in the (weak, strict) case of the maximum. We start with equation (8.8), with $d$ replaced by $\alpha n$.

$$\begin{aligned}
[z^n]F_M^{(w,s)}(z) &= \sum_{k\geq 1}\sum_{i=0}^{\alpha n-1} (1-q^{k-1})^{i+n-\alpha n} pq^{k-1}(1-q^k)^{\alpha n-1-i} \\
&= \sum_{k\geq 1}(1-q^{k-1})^{n(1-\alpha)} pq^{k-1}(1-q^k)^{\alpha n-1}\sum_{i=0}^{\alpha n-1}\left(\frac{1-q^{k-1}}{1-q^k}\right)^i.
\end{aligned}$$

Note that

$$\sum_{i=0}^{\alpha n-1}\left(\frac{1-q^{k-1}}{1-q^k}\right)^i = \frac{1-\left(\frac{1-q^{k-1}}{1-q^k}\right)^{\alpha n}}{1-\frac{1-q^{k-1}}{1-q^k}}$$

$$
\begin{aligned}
&= \frac{(1-q^k)^{\alpha n} - (1-q^{k-1})^{\alpha n}}{(1-q^k)^{\alpha n}} \cdot \frac{1-q^k}{1-q^k-(1-q^{k-1})} \\
&= \frac{(1-q^k)^{\alpha n} - (1-q^{k-1})^{\alpha n}}{(1-q^k)^{\alpha n-1} q^{k-1}(1-q)}, \qquad (12.4)
\end{aligned}
$$

and thus, (since $p = 1-q$, and $(1-a)^n \sim e^{-an}$ for small $a$)

$$
\begin{aligned}
[z^n]F_M^{(w,s)}(z) &= \sum_{k\geq 1} (1-q^{k-1})^{n(1-\alpha)} p q^{k-1} (1-q^k)^{\alpha n-1} \frac{(1-q^k)^{\alpha n} - (1-q^{k-1})^{\alpha n}}{(1-q^k)^{\alpha n-1} q^{k-1} p} \\
&= \sum_{k\geq 1} (1-q^{k-1})^{n(1-\alpha)} \big[(1-q^k)^{\alpha n} - (1-q^{k-1})^{\alpha n}\big] \\
&\sim \sum_{k\geq 1} e^{-q^{k-1} n(1-\alpha)} \big[e^{-q^k \alpha n} - e^{-q^{k-1}\alpha n}\big] \\
&= \sum_{k\geq 1} \Big[e^{-nq^{k-1}(1-\alpha(1-q))} - e^{-nq^{k-1}}\Big] \\
&= \sum_{k\geq 1} \Big[e^{-nq^{k-1}(1-\alpha p)} - e^{-nq^{k-1}}\Big]. \qquad (12.5)
\end{aligned}
$$

We are now in a position to use Mellin transforms to find an approximation. Unfortunately, taking the Mellin transform of the exponential function gives a fundamental strip of $\langle 0, \infty\rangle$. This fundamental strip can only be valid where it overlaps the interval of convergence of the geometric series in (12.8), which is $(-\infty, 0)$. Note that the fundamental strip does not include endpoints, and thus the intersection is empty. We therefore subtract one from each of the exponentials in (12.5), keeping the total value the same. The Mellin transform is now valid in the strip $\langle -1, 0\rangle$. We define a function of $x$ (which replaces $n$) as follows:

$$
f(x) := \sum_{k\geq 1} \big[(e^{-xq^{k-1}(1-\alpha p)} - 1) - (e^{-xq^{k-1}} - 1)\big].
$$

The Mellin transform of this function can be found using two rules (see [11, page 576]). By 'linearity',

$$
\sum_i \lambda_i f_i(x) \qquad \text{transforms to} \qquad \sum_i \lambda_i f_i^*(s), \qquad (12.6)
$$

so in our example the sum on $k$ is taken care of. (In this case $\lambda_i$ is simply one.) The 'scaling' rule for Mellin transforms states that for $\mu > 0$,

$$
f(\mu x) \qquad \text{transforms to} \qquad \mu^{-s} f^*(s). \qquad (12.7)
$$

The scaling rule is applicable in this case, where $f(x) = e^{-x}$ and $\mu$ is respectively $q^{k-1}(1-\alpha p)$ and $q^{k-1}$. The Mellin transform of this function is thus:

$$
f^*(s) = \sum_{k\geq 1} \big[\big(q^{k-1}(1-\alpha p)\big)^{-s}\Gamma(s) - (q^{k-1})^{-s}\Gamma(s)\big]
$$

$$
\begin{aligned}
&= \sum_{k\geq 1} q^{s(1-k)}\Gamma(s)\big[(1-\alpha p)^{-s}-1\big] \\
&= q^s\Gamma(s)\big[(1-\alpha p)^{-s}-1\big]\sum_{k\geq 1}(q^{-s})^k \\
&= q^s\Gamma(s)\big[(1-\alpha p)^{-s}-1\big]\frac{q^{-s}}{1-q^{-s}}, \qquad \text{for } \Re(s)<0 \qquad (12.8) \\
&= \Gamma(s)\big[(1-\alpha p)^{-s}-1\big]\frac{1}{1-q^{-s}}.
\end{aligned}
$$

The final fundamental strip is the overlap between the strip found previously and the real $s$ values for which the geometric sum converges. In this case our fundamental strip is $\langle -1, 0\rangle$. We choose a value inside this, say $-\frac{1}{2}$, with which to perform our inverse Mellin transform:

$$
f(x) = \frac{1}{2\pi\boldsymbol{i}}\int_{-\frac{1}{2}-\boldsymbol{i}\infty}^{-\frac{1}{2}+\boldsymbol{i}\infty}\Gamma(s)\big[(1-\alpha p)^{-s}-1\big]\frac{1}{1-q^{-s}}x^{-s}ds.
$$

We approximate this by moving the contour right and collecting negative residues. The first poles we encounter are at $s=0$ (which would be a double pole except that one of them cancels with the factor $(1-\alpha p)^{-s}-1$, so it is simple) and the simple poles at $s=\chi_k=\frac{2k\pi\boldsymbol{i}}{L}$, $k\neq 0$. As usual the former contributes the main term and the rest contribute the fluctuations which are comparatively extremely small. As $s\to 0$,

$$
\Gamma(s)\sim\frac{1}{s},
$$

$$
(1-\alpha p)^{-s}-1\sim 1-s\log(1-\alpha p)-1=-s\log(1-\alpha p),
$$

$$
\frac{1}{1-q^{-s}}=\frac{1}{1-e^{-s\log q}}\sim\frac{1}{1-(1-s\log q)}=\frac{1}{s\log q} \qquad (12.9)
$$

and

$$
x^{-s}\sim 1.
$$

Thus the negative residue is

$$
-[s^{-1}]\frac{1}{s}\big(-s\log(1-\alpha p)\big)\frac{1}{s\log q}=\frac{\log(1-\alpha p)}{\log q}. \qquad (12.10)
$$

(If $\alpha=1$ and thus $d=n$, we have the expected result of probability one.) We also have simple poles at $s=\chi_k$, for $k\neq 0$. Let $\varepsilon := s-\chi_k$ then at $\varepsilon=0$,

$$
\Gamma(s)=\Gamma(\varepsilon+\chi_k)=\Gamma(\chi_k),
$$

$$
(1-\alpha p)^{-s}-1=(1-\alpha p)^{-\chi_k-\varepsilon}-1=(1-\alpha p)^{-\chi_k}-1,
$$

$$x^{-s} = x^{-\chi_k - \varepsilon} = x^{-\chi_k}.$$

Since expanding $\frac{1}{1-q^{-s}}$ around $s = 0$ is the same as expanding $\frac{1}{1-q^{-\varepsilon}}$ around $\varepsilon = 0$ (see (12.9)) and $Q^{\chi_k} = 1$ (see (3.6)),

$$\frac{1}{1-q^{-s}} = \frac{1}{1-q^{-\chi_k-\varepsilon}} = \frac{1}{1-q^{-\varepsilon}} \sim \frac{1}{\varepsilon \log q}.$$

Hence the total negative residues for the simple poles at $s = \chi_k$, $k \neq 0$ are (recall $L = \log Q$)

$$\begin{aligned}
&\sum_{k\neq 0}(-1)[\varepsilon^{-1}]\Gamma(\chi_k)\big[(1-\alpha p)^{-\chi_k} - 1\big]\frac{1}{\varepsilon \log q}x^{-\chi_k} \\
&\quad = \frac{1}{L}\sum_{k\neq 0}\Gamma(\chi_k)\big[(1-\alpha p)^{-\chi_k} - 1\big]x^{-\chi_k} \\
&\quad = \frac{1}{L}\sum_{k\neq 0}\Gamma(\chi_k)\big[e^{-\chi_k \log(1-\alpha p)} - 1\big]e^{-\chi_k \log x} \\
&\quad = \frac{1}{L}\sum_{k\neq 0}\Gamma(\chi_k)\big[e^{-2k\pi \boldsymbol{i} \log_Q(1-\alpha p)} - 1\big]e^{-2k\pi \boldsymbol{i} \log_Q x} \\
&\quad = \frac{1}{L}\sum_{k\neq 0}\Gamma(\chi_k)\big[e^{-2k\pi \boldsymbol{i} \log_Q(x(1-\alpha p))} - e^{-2k\pi \boldsymbol{i} \log_Q x}\big] \\
&\quad = \frac{1}{L}\sum_{k\neq 0}\Gamma(-\chi_k)\big[e^{2k\pi \boldsymbol{i} \log_Q(x(1-\alpha p))} - e^{2k\pi \boldsymbol{i} \log_Q x}\big]. \qquad (12.11)
\end{aligned}$$

(The sign of each $k$ can be changed since we are summing over all non-zero $k$.) Now, substitute back $x = n$ and put (12.10) and (12.11) together to get the probability of having a weak maximum in the first $d$ positions which does not repeat in the rest of the word (where $d = \alpha n$ grows with $n$):

$$\frac{\log(1-\alpha p)}{\log q} + \frac{1}{L}\sum_{k\neq 0}\Gamma(-\chi_k)\big[e^{2k\pi \boldsymbol{i} \log_Q(n(1-\alpha p))} - e^{2k\pi \boldsymbol{i} \log_Q n}\big].$$

Thus

$$P_{ws}^{\alpha n}(M) \sim \frac{1}{L}\log\left(\frac{1}{1-\alpha(1-Q^{-1})}\right) + \frac{1}{L}\big(\delta_0(n(1-\alpha p)) - \delta_0(n)\big),$$

where

$$\delta_0(x) := \sum_{k\neq 0}\Gamma(-\chi_k)e^{2k\pi \boldsymbol{i} \log_Q x}. \qquad (12.12)$$

This concludes the proof of Theorem 12.1. ■

Note that if $\alpha = 1$ then $d = n$ which represents a word of length $n$ with a (possibly repeated) maximum. This is the same as a word of length $n$ with no restrictions, which will occur with a probability of 1. Replacing $\alpha$ by 1 in the main term yields a probability of 1.

For interest we plot the dominant term of this result for $\alpha$ from 0 to 1. It is interesting to note that different values of $q$ give differently-shaped graphs.

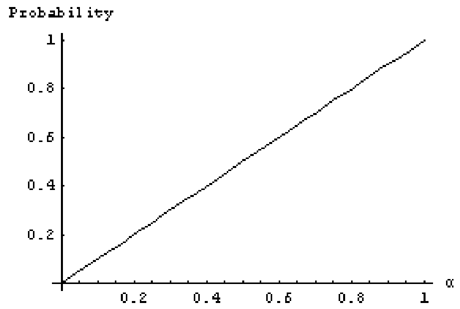


Figure 12.1: Probability of a (weak, strict) maximum for $d = \alpha n$, where $q = \frac{999}{1000}$.

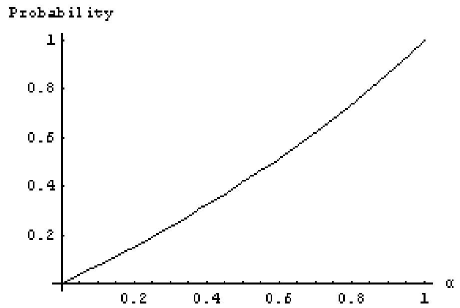


Figure 12.2: Probability of a (weak, strict) maximum for $d = \alpha n$, where $q = \frac{1}{2}$.

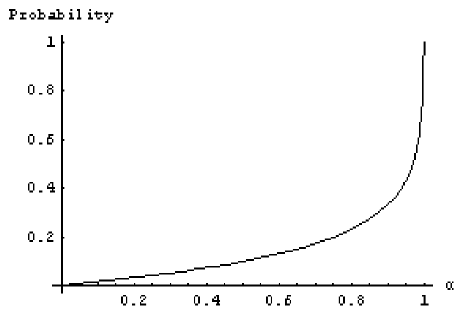


Figure 12.3: Probability of a (weak, strict) maximum for $d = \alpha n$, where $q = \frac{1}{1000}$.

The graphs above demonstrate that for large $q$ (close to 1), the graph is practically linear, tending to be more exponential as $q \to 0$. Why would this be the case? Because of the geometric distribution, $q$ small means that in practice we have a word made up almost entirely of ones (for example, for $q = \frac{1}{1000}$, if we considered a word of length $n = 1000$, we would expect only one of the thousand letters not to be a one). In this case, having the maximum occurring in the first $d$ positions is unlikely for small $\alpha$ (i.e., $d$ relatively small compared with $n$), simply because the maximum hardly ever occurs, whereas for $\alpha$ near 1 ($d$ near $n$) the probability improves. On the other hand, for $q$ large (close to one) we have a more even spread (larger letters occurring with greater variety), tending towards a permutation of the natural numbers as $n \to \infty$. We would expect a linear graph when plotting these probabilities for permutations, because the maximum in a permutation is just as likely to occur anywhere, and so the probability of it being in the first part of the word will grow linearly with $\alpha$. Examples of words of length 20 are given in Table 6.1.

## 12.2 Case (strict, weak), for $d = \alpha n$

From (8.15), replacing $d$ with $\alpha n$ gives

$$\begin{aligned}
[z^n]F_M^{(s,w)}(z) &= \sum_{k\geq 1} \alpha n(1-q^{k-1})^{\alpha n-1} pq^{k-1}(1-q^k)^{n-\alpha n} \\
&= \sum_{k\geq 1} \alpha npq^{k-1}(1-q^{k-1})^{\alpha n-1}(1-q^k)^{n(1-\alpha)} \\
&\sim \sum_{k\geq 1} \alpha npq^{k-1}(1-q^{k-1})^{\alpha n}(1-q^k)^{n(1-\alpha)} \\
&\sim \sum_{k\geq 1} \alpha npq^{k-1}e^{-(\alpha n)q^{k-1}}e^{-n(1-\alpha)q^k} \\
&= \sum_{k\geq 1} \alpha npq^{k-1}e^{-(\alpha n)q^{k-1}-n(1-\alpha)q^k} \\
&= \sum_{k\geq 1} \alpha npq^{k-1}e^{-nq^{k-1}(\alpha+q-q\alpha)} \\
&= \sum_{k\geq 1} \alpha npq^{k-1}e^{-nq^{k-1}(q+p\alpha)}.
\end{aligned}$$

We are now ready to take the Mellin transform. Define

$$f(x) := \sum_{k\geq 1} \alpha xpq^{k-1}e^{-xq^{k-1}(q+p\alpha)}.$$

Again, the linearity and scaling rules (see (12.6) and (12.7)) can be used. In this case, because of the factor of $x$, the 'power' rule is also used. This rule is stated in (13.8), and is responsible for the $s+1$ replacing the expected $s$ in the transform to follow.

$$\begin{aligned}
f^*(s) &= \sum_{k\geq 1} \alpha pq^{k-1}(q^{k-1})^{-(s+1)}(q+p\alpha)^{-(s+1)}\Gamma(s+1) \\
&= \alpha p(q+p\alpha)^{-(s+1)}\Gamma(s+1)\sum_{k\geq 1}(q^{k-1})^{-s}
\end{aligned}$$

$$
\begin{aligned}
&= \alpha p q^s (q+p\alpha)^{-(s+1)}\Gamma(s+1)\sum_{k\geq 1} q^{-sk} \\
&= \alpha p q^s (q+p\alpha)^{-(s+1)}\Gamma(s+1)\frac{q^{-s}}{1-q^{-s}}, \qquad \text{for } \Re(s)<0 \\
&= \alpha p (q+p\alpha)^{-(s+1)}\Gamma(s+1)\frac{1}{1-q^{-s}}.
\end{aligned}
$$

The fundamental strip is the overlap of the interval $(-\infty, 0)$ and the fundamental strip of $xe^{-x}$ which has a left boundary at $-1$, since

$$
\lim_{x\to 0} xe^{-x} = 0, \quad \text{and} \quad xe^{-x} \sim x - x^2, \quad \text{which grows like } x^1 \text{ as } x \to 0,
$$

and a right boundary of $\infty$, since

$$
\lim_{x\to\infty} xe^{-x} = 0 = O(x^{-m}) \text{ for any positive } m.
$$

Thus the fundamental strip for $f(x)$ is $\langle -1, 0\rangle$. We choose the contour integral from $-\frac{1}{2} - \boldsymbol{i}\infty$ to $-\frac{1}{2} + \boldsymbol{i}\infty$, and perform the inverse Mellin transform to get:

$$
f(x) = \frac{1}{2\pi \boldsymbol{i}} \int_{(-\frac{1}{2})} \alpha p (q+p\alpha)^{-(s+1)}\Gamma(s+1)\frac{1}{1-q^{-s}} x^{-s} ds.
$$

By moving the contour right (since we are interested in $x$ large), the first poles we reach are the simple pole at $s=0$, as well as the simple poles at $s=\chi_k$, $k\neq 0$. From (12.9),

$$
\frac{1}{1-q^{-s}} \sim \frac{1}{s\log q},
$$

and so the negative residue is

$$
\begin{aligned}
-[s^{-1}]\alpha p (q+p\alpha)^{-1}\Gamma(1)\frac{1}{s\log q} &= -\alpha p (q+p\alpha)^{-1}\Gamma(1)\frac{1}{\log q} \\
&= \alpha p (q+p\alpha)^{-1}\frac{1}{\log Q} \\
&= \frac{\alpha p}{L(q+p\alpha)}. \qquad (12.13)
\end{aligned}
$$

That is the main term, but there are also fluctuations which come from the negative residues of the poles at $s=\chi_k$. Let $\varepsilon := s - \chi_k$, then around $\varepsilon = 0$ we get (see (12.9) and (3.6))

$$
\frac{1}{1-q^{-s}} \sim \frac{1}{\varepsilon \log q},
$$

and so the residues are:

$$
\begin{aligned}
&-\sum_{k\neq 0}[\varepsilon^{-1}]\alpha p (q+p\alpha)^{-(\chi_k+1)}\Gamma(\chi_k+1)\frac{1}{\varepsilon\log q}x^{-\chi_k} \\
&\qquad = -\sum_{k\neq 0}\alpha p (q+p\alpha)^{-(\chi_k+1)}\Gamma(\chi_k+1)\frac{1}{\log q}x^{-\chi_k}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\alpha p}{L} \sum_{k \neq 0} (q + p\alpha)^{-(\chi_k + 1)} \Gamma(\chi_k + 1) x^{-\chi_k} \\
&= \frac{\alpha p}{L(q + p\alpha)} \sum_{k \neq 0} \Gamma(\chi_k + 1)(x(q + p\alpha))^{-\chi_k} \\
&= \frac{\alpha p}{L(q + p\alpha)} \sum_{k \neq 0} \Gamma(\chi_k + 1) e^{-\chi_k \log(x(q+p\alpha))} \\
&= \frac{\alpha p}{L(q + p\alpha)} \sum_{k \neq 0} \Gamma(\chi_k + 1) e^{-2k\pi \boldsymbol{i} \log_Q(x(q+p\alpha))} \\
&= \frac{\alpha p}{L(q + p\alpha)} \sum_{k \neq 0} \Gamma(1 - \chi_k) e^{2k\pi \boldsymbol{i} \log_Q(x(q+p\alpha))}. \qquad (12.14)
\end{aligned}
$$

The expressions (12.13) and (12.14) give a total probability in the (strict, weak) case of

$$
\frac{\alpha p}{L(q + p\alpha)} + \frac{\alpha p}{L(q + p\alpha)} \sum_{k \neq 0} \Gamma(1 - \chi_k) e^{2k\pi \boldsymbol{i} \log_Q(n(q+p\alpha))},
$$

and so

$$
P_{sw}^{\alpha n}(M) \sim \frac{\alpha(Q - 1)}{L(1 + \alpha(Q - 1))} \big(1 + \delta(n(q + p\alpha))\big),
$$

for $\delta(x) = \sum_{k \neq 0} \Gamma(1 - \chi_k) e^{2k\pi \boldsymbol{i} \log_Q x}$ as in (7.5). Hence the proof of Theorem 12.2 is completed. ■

To check the above result we again put $\alpha = 1$, but in this case we do not expect a probability of 1, since we want the probability of getting a strict maximum in a word of length $n$ (for $d = n$, the first strict/weak classification is valid for the entire word). This probability has been found in [18] to be $\frac{p}{L}$. For $\alpha = 1$:

$$
\frac{\alpha p}{L(q + p\alpha)} = \frac{p}{L(q + p)} = \frac{p}{L},
$$

as required. Again we look at three different graphs, corresponding to various values of $q$.

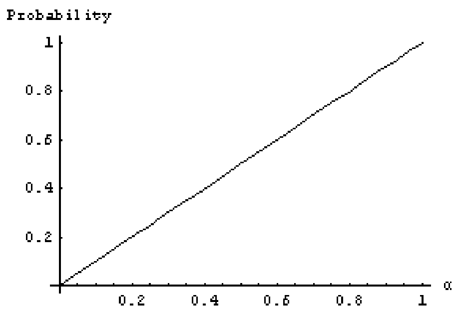


Figure 12.4: Probability of a (strict, weak) maximum for $d = \alpha n$, where $q = \frac{999}{1000}$.

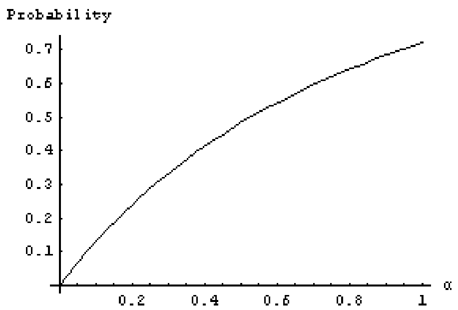


Figure 12.5: Probability of a (strict, weak) maximum for $d = \alpha n$, where $q = \frac{1}{2}$.

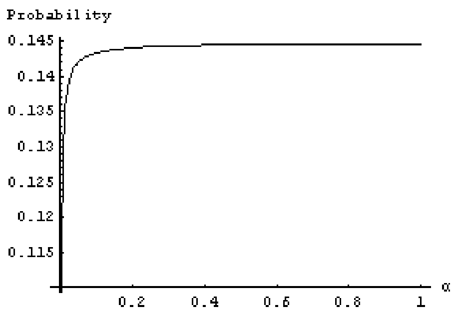


Figure 12.6: Probability of a (strict, weak) maximum for $d = \alpha n$, where $q = \frac{1}{1000}$.

Here the graphs follow the same pattern as in the (weak, strict) case in that the larger the value of $q$, the straighter the line. However, the graphs here are convex rather than concave. The probability reaches its maximum (not necessarily one) quickly, as a strict maximum is very unlikely to occur anyway, but if it does it will almost certainly be a weak maximum of the rest of the word. Also, the maximum probability decreases as $q$ decreases. This is because for small $q$ we expect a majority of ones and thus we are unlikely to find a strict maximum, whatever the value of $d$.

## 12.3 Case (weak, weak), for $d = \alpha n$

Using the approximation $(1-a)^n \sim e^{-an}$ for small $a$, we start with (8.18) to get

$$
\begin{aligned}
[z^n]F_M^{(w,w)}(z) &= \sum_{k\geq 1}\sum_{i=0}^{\alpha n-1}(1-q^{k-1})^i pq^{k-1}(1-q^k)^{n-1-i} \\
&= \sum_{k\geq 1} pq^{k-1}(1-q^k)^{n-1}\sum_{i=0}^{\alpha n-1}\left(\frac{1-q^{k-1}}{1-q^k}\right)^i \\
&= \sum_{k\geq 1}(1-q^k)^{n(1-\alpha)}\big[(1-q^k)^{\alpha n}-(1-q^{k-1})^{\alpha n}\big], \quad \text{see (12.4)} \\
&\sim \sum_{k\geq 1} e^{-nq^k(1-\alpha)}\big[e^{-\alpha n q^k}-e^{-\alpha n q^{k-1}}\big] \\
&= \sum_{k\geq 1}\big[e^{-nq^k}-e^{-nq^{k-1}(q+\alpha p)}\big] \\
&= \sum_{k\geq 1}\big[(e^{-nq^k}-1)-(e^{-nq^{k-1}(q+\alpha p)}-1)\big].
\end{aligned}
$$

As in the (weak, strict) case, we subtract one from each exponential so that the fundamental strip is non-empty i.e., there is an overlap of the strip where the Mellin transform of $e^{-cx}$ exists (namely $\langle -1, 0\rangle$) and the geometric sum (to follow) converges (for $\Re(s) < 0$). We define

$$
f(x) := \sum_{k\geq 1}\big[(e^{-xq^k}-1)-(e^{-xq^{k-1}(q+\alpha p)}-1)\big],
$$

so that (by (12.6) and (12.7))

$$
\begin{aligned}
f^*(s) &= \sum_{k\geq 1}\big[q^{-sk}\Gamma(s)-(q^{k-1})^{-s}(q+\alpha p)^{-s}\Gamma(s)\big] \\
&= \Gamma(s)\sum_{k\geq 1}\big[q^{-sk}-q^{-sk}q^s(q+\alpha p)^{-s}\big] \\
&= \Gamma(s)\big[1-q^s(q+\alpha p)^{-s}\big]\sum_{k\geq 1}q^{-sk} \\
&= \Gamma(s)\big[1-q^s(q+\alpha p)^{-s}\big]\frac{q^{-s}}{1-q^{-s}}, \qquad \text{for } \Re(s)<0,
\end{aligned}
$$

exists in the strip $\langle -1, 0\rangle$. We can thus rewrite $f(x)$ as the contour integral

$$
f(x) = \frac{1}{2\pi \boldsymbol{i}}\int_{(-\frac{1}{2})}\Gamma(s)\big[1-q^s(q+\alpha p)^{-s}\big]\frac{q^{-s}}{1-q^{-s}}x^{-s}ds.
$$

Moving the contour right to collect negative residues, the first poles we encounter are simple poles which occur at $s=0$ and $s=\chi_k$, $k\neq 0$. The former contributes the main term and the latter the fluctuations. At $s=0$:

$$
\Gamma(s) \sim \frac{1}{s},
$$

$$\begin{aligned}1 - q^s(q+\alpha p)^{-s} &= 1 - \Big(\frac{q+\alpha p}{q}\Big)^{-s} = 1 - e^{-s\log\left(\frac{q+\alpha p}{q}\right)} \\ &\sim 1 - \Big(1 - s\log\Big(\frac{q+\alpha p}{q}\Big)\Big) = s\log\Big(\frac{q+\alpha p}{q}\Big),\end{aligned}$$

$$\frac{q^{-s}}{1-q^{-s}} \sim \frac{1}{s\log q}, \tag{12.15}$$

and

$$x^{-s} = 1.$$

Thus the negative residue at $s=0$ is

$$-[s^{-1}]\frac{1}{s} s\log\Big(\frac{q+\alpha p}{q}\Big)\frac{1}{s\log q} = \frac{1}{L}\log\Big(\frac{q+\alpha p}{q}\Big). \tag{12.16}$$

At $s = \chi_k$, let $\varepsilon := s - \chi_k$. Then expanding around $\varepsilon = 0$ gives (see (12.15))

$$\frac{q^{-s}}{1-q^{-s}} \sim \frac{1}{\varepsilon\log q},$$

and so the sum of the negative residues is

$$\begin{aligned}&-\sum_{k\neq 0}[\varepsilon^{-1}]\Gamma(\chi_k)\big[1 - q^{\chi_k}(q+\alpha p)^{-\chi_k}\big]\frac{1}{\varepsilon\log q}x^{-\chi_k} \\ &= -\sum_{k\neq 0}\Gamma(\chi_k)\big[1 - q^{\chi_k}(q+\alpha p)^{-\chi_k}\big]\frac{1}{\log q}x^{-\chi_k} \\ &= \frac{1}{L}\sum_{k\neq 0}\Gamma(\chi_k)\Big[1 - \Big(\frac{q+\alpha p}{q}\Big)^{-\chi_k}\Big]x^{-\chi_k} \\ &= \frac{1}{L}\sum_{k\neq 0}\Gamma(\chi_k)\Big[1 - e^{-\chi_k\log(\frac{q+\alpha p}{q})}\Big]e^{-\chi_k\log x} \\ &= \frac{1}{L}\sum_{k\neq 0}\Gamma(\chi_k)\Big[e^{-2k\pi\boldsymbol{i}\log_Q x} - e^{-2k\pi\boldsymbol{i}\log_Q\left(x(\frac{q+\alpha p}{q})\right)}\Big] \\ &= \frac{1}{L}\sum_{k\neq 0}\Gamma(-\chi_k)\Big[e^{2k\pi\boldsymbol{i}\log_Q x} - e^{2k\pi\boldsymbol{i}\log_Q\left(x(\frac{q+\alpha p}{q})\right)}\Big]. \end{aligned} \tag{12.17}$$

Therefore from (12.16) and (12.17), the total probability in the (weak, weak) case is

$$\frac{1}{L}\log\Big(\frac{q+\alpha p}{q}\Big) + \frac{1}{L}\sum_{k\neq 0}\Gamma(-\chi_k)\Big[e^{2k\pi\boldsymbol{i}\log_Q n} - e^{2k\pi\boldsymbol{i}\log_Q\left(n(\frac{q+\alpha p}{q})\right)}\Big].$$

Consequently

$$P_{ww}^{\alpha n}(M) \sim \frac{\log(1+\alpha(Q-1))}{L} + \frac{1}{L}\Big(\delta_0(n) - \delta_0\Big(n\Big(\frac{q+\alpha p}{q}\Big)\Big)\Big),$$

where $\delta_0(x) = \sum_{k\neq 0}\Gamma(-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q x}$ as in (12.12). Thus the proof of Theorem 12.3 is complete. ■

This case, (weak, weak), is similar to the (weak, strict) case above, and when $\alpha = 1$, $d = n$ and we are calculating the probability of having a weak maximum in a word of length $n$ – that is – a word with no restrictions. If we put $\alpha = 1$ into the main term above, we get

$$\frac{\log(1+\alpha(Q-1))}{L} = \frac{\log(1+(Q-1))}{L} = \frac{\log Q}{L} = 1,$$

as expected. The three graphs in this case will be

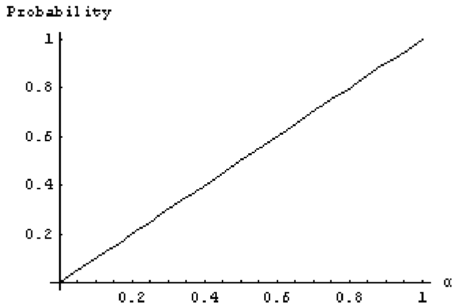


Figure 12.7: Probability of a (weak, weak) maximum for $d = \alpha n$, where $q = \frac{999}{1000}$.

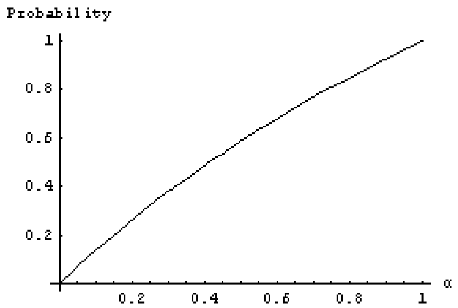


Figure 12.8: Probability of a (weak, weak) maximum for $d = \alpha n$, where $q = \frac{1}{2}$.

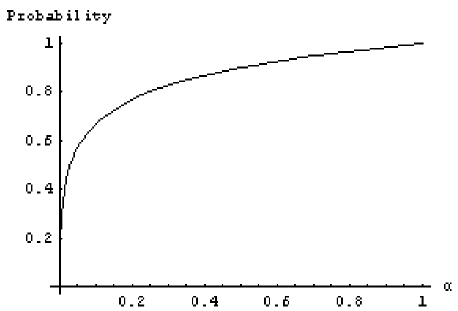


Figure 12.9: Probability of a (weak, weak) maximum for $d = \alpha n$, where $q = \frac{1}{1000}$.

For $q$ large, the same explanation of linearity for permutations holds. For $q$ small the graph is convex since there are so few distinct letters that a weak maximum of the first $d$ will still most likely continue to be a weak maximum of the rest, hence the maximum is reached quite soon. Note that the curve is more gentle than in the previous case due to the weak classification.

# Chapter 13

# The maximum in the first $d$ positions for $1 \le d = o(n)$

What if $d$ is dependent on $n$, but does not grow linearly with $n$? For example, take $d = \alpha n^\gamma$ for $0 < \gamma < 1$. It can be shown that the results are the same as when $d$ is fixed. We can thus refer back to the step in the calculations for $d$ fixed where the $d = \alpha n$ calculations failed. The important stage is when the main term of the probability is given by the expression

$$\underbrace{d\sum_{l=0}^{d-1}\binom{d-1}{l}}_{\mathrm{i}}(-1)^l\frac{Q^{l+1}(1-Q^{-1})}{L}\underbrace{\frac{1}{(n-d+1+l)\binom{n-d+l}{l}}}_{\mathrm{ii}}. \tag{13.1}$$

This is the (strict, weak) case but the others are similar. For $d$ fixed, it can be seen that the $l = 0$ term dominates, since each term in the sum on $l$ is $O\big(\frac{1}{n^{l+1}}\big)$. For $d$ proportional to $n$, each term is $O(\frac{1}{n})$, so none clearly dominates, and Mellin transforms are required to find the result (see Chapter 12). But what if $d = \alpha n^\gamma$ for $0 < \gamma < 1$, or even $d = \frac{n}{\log n}$? It turns out that for $1 \le d = o(n)$, we get the same results as the $d$ fixed case. The explanation is given below.

Suppose we let $f(n) = o(n)$ for some $f(n)$ such that $f(n) \to \infty$ as $n \to \infty$. Then we can write $d = \frac{n}{f(n)}\big(= o(n)\big)$. In general, a typical term in the sum on $l$ is $O\big(\frac{1}{f^{l+1}(n)}\big)$. This is because the final fraction (ii) in (13.1) will be $O\big(\frac{1}{n^{l+1}}\big)$, and the first part (i) will be $O\big(\frac{n^{l+1}}{f^{l+1}(n)}\big)$. Thus since $f(n) \to \infty$ as $n \to \infty$, the $l = 0$ term dominates and the same calculations as in the $d$ fixed case hold. The results listed below and proved in this chapter support this (using Mellin transforms) for $d = \alpha n^\gamma$ where $0 < \gamma < 1$. We consider the same three cases as in Chapter 12.

**Theorem 13.1** *The probability that the maximum in a geometrically distributed sample appears at least once in the first $d = \alpha n^\gamma$ letters, but not in the rest of the* *sample is*

$$P_{ws}^{\alpha n^\gamma}(M) \sim \frac{(1-Q^{-1})\alpha n^\gamma}{Ln}\big(1+\delta(n)\big), \tag{13.2}$$

*as* $n \to \infty$*, where*

$$\delta(x) = \sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q x},$$

*as defined in (7.5).*

**Theorem 13.2** *The probability that the maximum in a geometrically distributed sample appears exactly once in the first* $d = \alpha n^\gamma$ *letters, and any number of times in the rest of the sample is*

$$P_{sw}^{\alpha n^\gamma}(M) \sim \frac{(Q-1)\alpha n^\gamma}{Ln}(1+\delta(n)), \tag{13.3}$$

*as* $n \to \infty$*, for* $\delta(x) = \sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q x}$ *as in (7.5).*

**Theorem 13.3** *The probability that the maximum in a geometrically distributed sample appears at least once in the first* $d = \alpha n^\gamma$ *letters, and any number of times in the rest of the sample is*

$$P_{ww}^{\alpha n^\gamma}(M) \sim \frac{(Q-1)\alpha n^\gamma}{Ln}(1+\delta(n)) \tag{13.4}$$

*as* $n \to \infty$*, where* $\delta(x) = \sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q x}$ *is defined in (7.5).*

## 13.1 Case (weak, strict), for $d = \alpha n^\gamma$

Again we start with equation (8.8) for the (weak, strict) case of the maximum, and replace $d$ by $\alpha n^\gamma$ where $0 < \alpha \leq 1$ and $0 < \gamma < 1$. Then

$$\begin{aligned}
[z^n]F_M^{(w,s)}(z) &= \sum_{k\geq 1}\sum_{i=0}^{d-1}(1-q^{k-1})^{i+n-d}pq^{k-1}(1-q^k)^{d-1-i} \\
&= \sum_{k\geq 1}(1-q^{k-1})^{n-\alpha n^\gamma}pq^{k-1}(1-q^k)^{\alpha n^\gamma-1}\sum_{i=0}^{\alpha n^\gamma-1}\left(\frac{1-q^{k-1}}{1-q^k}\right)^i.
\end{aligned}$$

The sum on $i$ can be simplified as follows

$$\begin{aligned}
\sum_{i=0}^{\alpha n^\gamma-1}\left(\frac{1-q^{k-1}}{1-q^k}\right)^i &= \frac{1-\left(\frac{1-q^{k-1}}{1-q^k}\right)^{\alpha n^\gamma}}{1-\frac{1-q^{k-1}}{1-q^k}} \\
&= \frac{(1-q^k)^{\alpha n^\gamma}-(1-q^{k-1})^{\alpha n^\gamma}}{(1-q^k)^{\alpha n^\gamma}}\cdot\frac{1-q^k}{1-q^k-(1-q^{k-1})}
\end{aligned}$$

$$= \frac{(1-q^k)^{\alpha n^\gamma} - (1-q^{k-1})^{\alpha n^\gamma}}{(1-q^k)^{\alpha n^\gamma - 1} q^{k-1} p}. \tag{13.5}$$

Substituting this back into the initial equation allows us to cancel the denominator:

$$\begin{aligned}
[z^n]F_M^{(w,s)}(z) &= \sum_{k\geq 1}(1-q^{k-1})^{n-\alpha n^\gamma} p q^{k-1}(1-q^k)^{\alpha n^\gamma - 1}\frac{(1-q^k)^{\alpha n^\gamma} - (1-q^{k-1})^{\alpha n^\gamma}}{(1-q^k)^{\alpha n^\gamma - 1} q^{k-1} p} \\
&= \sum_{k\geq 1}(1-q^{k-1})^{n-\alpha n^\gamma}\left[(1-q^k)^{\alpha n^\gamma} - (1-q^{k-1})^{\alpha n^\gamma}\right] \\
&= \sum_{k\geq 1}\left[(1-q^{k-1})^{n-\alpha n^\gamma}(1-q^k)^{\alpha n^\gamma} - (1-q^{k-1})^{n}\right] \\
&\sim \sum_{k\geq 1}\left[e^{-q^{k-1}(n-\alpha n^\gamma)} e^{-q^k \alpha n^\gamma} - e^{-q^{k-1}n}\right]
\end{aligned}$$

using the approximation $(1-a)^n \sim e^{-an}$ for small $a$. We are now ready to use the Mellin transform. We define the function we want to transform as

$$f(x) := \sum_{k\geq 1}\left[e^{-q^{k-1}x + \alpha p q^{k-1} x^\gamma} - e^{-q^{k-1}x}\right]. \tag{13.6}$$

In this case the first exponential is a function of two different powers of $x$. To better understand how to deal with this, we first consider the simplified expression $e^{-x+\sqrt{x}}$, and use a series expansion to write

$$e^{-x+\sqrt{x}} = e^{-x}\sum_{i\geq 0}\frac{1}{i!}(\sqrt{x})^i.$$

So instead of transforming $e^{-x+\sqrt{x}}$, we transform $\sum\limits_{i\geq 0}\frac{1}{i!}x^{\frac{i}{2}}e^{-x}$. This can be done using the 'harmonic sum rule' (see [11, page 576]), namely for $\mu_i > 0$,

$$\sum_i \lambda_i f(\mu_i x) \qquad \text{transforms to} \qquad \Big(\sum_i \lambda_i \mu_i^{-s}\Big)\cdot f^*(s), \tag{13.7}$$

and the 'power rule', which states that

$$x^\rho f(x^\theta) \qquad \text{transforms to} \qquad \frac{1}{\theta} f^*\Big(\frac{s+\rho}{\theta}\Big). \tag{13.8}$$

Then ('$\mathcal{M}$' denotes the Mellin transform)

$$\mathcal{M}\big(x^{\frac{i}{2}}e^{-x}\big) = \Gamma\Big(s+\frac{i}{2}\Big),$$

and so

$$\mathcal{M}\big(e^{-x+\sqrt{x}}\big) = \sum_i \frac{1}{i!}\Gamma\Big(s+\frac{i}{2}\Big).$$

Now all that remains is to generalise the power $\frac{1}{2}$ to $\gamma$ and introduce some constant coefficients. Thus

$$\mathcal{M}\big(e^{-ax+bx^\gamma}\big) = \mathcal{M}\Bigg(e^{-ax}\sum_{i\geq 0}\frac{1}{i!}(bx^\gamma)^i\Bigg)$$

$$
\begin{aligned}
&= \sum_{i\geq 0} \frac{1}{i!} b^i \mathcal{M}\big(e^{-ax} x^{\gamma i}\big) \\
&= \sum_{i\geq 0} \frac{1}{i!} b^i a^{-s-\gamma i} \Gamma(s+\gamma i).
\end{aligned}
$$

In our case, $a = q^{k-1}$ and $b = \alpha p q^{k-1}$. We first rewrite the function in (13.6) as

$$
\begin{aligned}
f(x) &= \sum_{k\geq 1} e^{-q^{k-1}x} \big(e^{\alpha p q^{k-1} x^{\gamma}} - 1\big) \\
&= \sum_{k\geq 1} e^{-q^{k-1}x} \left( \sum_{j\geq 0} \frac{(\alpha p q^{k-1} x^{\gamma})^j}{j!} - 1 \right) \\
&= \sum_{k\geq 1} e^{-q^{k-1}x} \sum_{j\geq 1} \frac{(\alpha p q^{k-1} x^{\gamma})^j}{j!} \\
&= \sum_{j\geq 1} \frac{(\alpha p q^{-1})^j}{j!} \sum_{k\geq 1} q^{jk} x^{\gamma j} e^{-q^{k-1}x}.
\end{aligned}
$$

Then using the above we get the Mellin transform of $f(x)$ to be

$$
\begin{aligned}
f^*(s) &= \sum_{j\geq 1} \frac{(\alpha p q^{-1})^j}{j!} \sum_{k\geq 1} q^{jk} (q^{k-1})^{-s-\gamma j} \Gamma(s+\gamma j) \\
&= \sum_{j\geq 1} \frac{(\alpha p q^{-1})^j}{j!} \Gamma(s+\gamma j) q^{s+\gamma j} \sum_{k\geq 1} q^{k(j-s-\gamma j)} \\
&= \sum_{j\geq 1} \frac{(\alpha p q^{-1})^j}{j!} \Gamma(s+\gamma j) q^{s+\gamma j} \frac{q^{j-s-\gamma j}}{1-q^{j-s-\gamma j}}, \qquad \text{for } \Re(s) < j(1-\gamma) \\
&= \sum_{j\geq 1} \frac{(\alpha p)^j}{j!} \Gamma(s+\gamma j) \frac{1}{1-q^{j-s-\gamma j}}.
\end{aligned}
$$

Now we must find the fundamental strip. To do that we examine what happens to $f(x)$ as $x \to 0$ and when $x \to \infty$. Around $x = 0$, we need only consider

$$
e^{-x}(e^{x^{\gamma}} - 1) \sim e^{-x}(1 + x^{\gamma} - 1) = x^{\gamma} e^{-x} \sim x^{\gamma},
$$

which tells us that the left boundary of the fundamental strip is $-\gamma$. We also have

$$
\lim_{x\to\infty} e^{-x}(e^{x^{\gamma}} - 1) = 0 = O(x^{-m}) \text{ for any positive } m,
$$

so the right boundary will be $\infty$. However, we must also consider the convergence of the sum on $k$. For this to converge we need the real part of $s$ to lie in the interval $(-\infty, j(1-\gamma))$. The intersection of these gives the strip in which the Mellin transform exists, i.e., $\langle -\gamma, \infty\rangle \cap \langle -\infty, j(1-\gamma)\rangle = \langle -\gamma, j(1-\gamma)\rangle$. To continue with the inversion formula, we must pick a value in this range. Since $j \geq 1$ and $0 < \gamma < 1$, we can choose this value to be 0. The inverse Mellin transform (which gives us back our original function as a contour integral) will then be

$$f(x) = \frac{1}{2\pi \boldsymbol{i}} \sum_{j\geq 1} \frac{(\alpha p)^j}{j!} \int_{(0)} \Gamma(s+\gamma j) \frac{1}{1-q^{j-s-\gamma j}} x^{-s} ds$$

which can be approximated using residue calculus. Since we are interested in $x$ ($= n$, the size of a word) large, we collect negative residues by moving the contour to the right. Possible poles lie at $s+\gamma j = 0, -1, -2, \ldots$ or at $j - s - \gamma j = \chi_k, \ \forall\ k$. Moving right from 0 means that the first poles we encounter will be at $j = 1$ (so $s = 1-\gamma$) when $k = 0$, and $1 - s - \gamma = \chi_k$ when $k \neq 0$. The first of these gives the dominant pole, and for $\varepsilon := s + \gamma - 1$, expanding around $\varepsilon = 0$ gives (set $j = 1$)

$$\Gamma(s+\gamma) = \Gamma(\varepsilon + 1) = \Gamma(1) = 1,$$
$$\frac{1}{1-q^{1-s-\gamma}} = \frac{1}{1-q^{-\varepsilon}} \sim \frac{1}{\varepsilon \log q}, \qquad \text{(by 12.9)}$$
$$x^{-s} = x^{-\varepsilon-1+\gamma} = x^{\gamma-1}.$$

The contribution from this pole is thus

$$-[\varepsilon^{-1}](\alpha p) \frac{1}{\varepsilon \log q} x^{\gamma-1} = \frac{\alpha p x^{\gamma-1}}{L}. \tag{13.9}$$

For the remaining simple poles, define $\varepsilon := \chi_k - 1 + s + \gamma$. Then, around $\varepsilon = 0$:

$$\Gamma(s+\gamma) = \Gamma(\varepsilon + 1 - \chi_k) = \Gamma(1-\chi_k),$$
$$\frac{1}{1-q^{1-s-\gamma}} \sim \frac{1}{\varepsilon \log q}, \qquad \text{from (12.9)}$$
$$x^{-s} = x^{\chi_k-\varepsilon-1+\gamma} = x^{\chi_k+\gamma-1}.$$

Therefore in total, for all of the poles at $1 - s - \gamma = \chi_k$, the negative residue is:

$$\begin{aligned} \sum_{k\neq 0} (-1)[\varepsilon^{-1}](\alpha p)\Gamma(1-\chi_k) \frac{1}{\varepsilon \log q} x^{\chi_k+\gamma-1} &= \frac{\alpha p x^{\gamma}}{Lx} \sum_{k\neq 0} \Gamma(1-\chi_k) x^{\chi_k} \\ &= \frac{\alpha p x^{\gamma}}{Lx} \sum_{k\neq 0} \Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q x}. \end{aligned} \tag{13.10}$$

Replacing $x$ with $n$ in (13.9) and (13.10) gives the final probability in the (weak, strict) case as asymptotic to

$$\frac{\alpha p n^{\gamma}}{Ln} + \frac{\alpha p n^{\gamma}}{Ln} \sum_{k\neq 0} \Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q n},$$

as $n \to \infty$. Hence

$$P_{ws}^{\alpha n^{\gamma}}(M) \sim \frac{(1-Q^{-1})\alpha n^{\gamma}}{Ln} \big(1 + \delta(n)\big),$$

where $\delta(x) = \sum_{k\neq 0} \Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q x}$ as in (7.5). This concludes the proof of Theorem 13.1. ■

## 13.2 Case (strict, weak), for $d = \alpha n^\gamma$

We begin in the same manner as always. From (8.15), as $n \to \infty$

$$\begin{aligned}
[z^n]F_M^{(s,w)}(z) &= \sum_{k\geq 1} \alpha n^\gamma (1-q^{k-1})^{\alpha n^\gamma - 1} pq^{k-1}(1-q^k)^{n-\alpha n^\gamma} \\
&\sim \sum_{k\geq 1} \alpha n^\gamma (1-q^{k-1})^{\alpha n^\gamma} pq^{k-1}(1-q^k)^{n-\alpha n^\gamma} \\
&\sim \sum_{k\geq 1} \alpha n^\gamma e^{-q^{k-1}\alpha n^\gamma} pq^{k-1} e^{-q^k(n-\alpha n^\gamma)} \\
&= \sum_{k\geq 1} \alpha pq^{k-1} n^\gamma e^{-q^k n - pq^{k-1}\alpha n^\gamma}.
\end{aligned}$$

In order to perform the transform, we define the function:

$$f(x) := \sum_{k\geq 1} \alpha pq^{k-1} x^\gamma e^{-q^k x - pq^{k-1}\alpha x^\gamma}.$$

Again we have two different powers of $x$ in the exponential function, and we transform this function (using the 'harmonic sum rule', and the power rule with $\theta = 1$ and $\rho = \gamma(i+1)$) to

$$\begin{aligned}
f^*(s) := \mathcal{M}&\left( \sum_{k\geq 1} \alpha pq^{k-1} x^\gamma e^{-q^k x} e^{-pq^{k-1}\alpha x^\gamma} \right) \\
&= \sum_{k\geq 1} \alpha pq^{k-1} \mathcal{M}\left( x^\gamma e^{-q^k x} \sum_{i\geq 0} \frac{1}{i!}(-pq^{k-1}\alpha x^\gamma)^i \right) \\
&= \sum_{k\geq 1} \alpha pq^{k-1} \sum_{i\geq 0} \frac{1}{i!}(-pq^{k-1}\alpha)^i \mathcal{M}\big(x^{\gamma(i+1)} e^{-q^k x}\big) \\
&= \sum_{k\geq 1} \alpha pq^{k-1} \sum_{i\geq 0} \frac{1}{i!}(-pq^{k-1}\alpha)^i (q^k)^{-s-\gamma(i+1)} \Gamma(s+\gamma(i+1)) \\
&= \sum_{i\geq 0} \frac{1}{i!}(-pq^{-1}\alpha)^i \alpha pq^{-1} \Gamma(s+\gamma(i+1)) \sum_{k\geq 1} q^{k(i+1-s-\gamma(i+1))} \\
&= \sum_{i\geq 0} \frac{1}{i!}(-pq^{-1}\alpha)^i \alpha pq^{-1} \Gamma(s+\gamma(i+1)) \frac{q^{i+1-s-\gamma(i+1)}}{1-q^{i+1-s-\gamma(i+1)}} \\
&= \sum_{i\geq 0} \frac{1}{i!}(-p\alpha)^i \alpha p \Gamma(s+\gamma(i+1)) \frac{q^{-s-\gamma(i+1)}}{1-q^{i+1-s-\gamma(i+1)}},
\end{aligned}$$

for $\Re(s) < (1-\gamma)(i+1)$. We determine the fundamental strip as follows. Ignoring constants, our function behaves like $x^\gamma e^{-x-x^\gamma}$, so we consider the behaviour around $x$ at 0 and $x$ at infinity. First the series expansion around $x = 0$ is

$$x^\gamma e^{-x-x^\gamma} \sim x^\gamma(1 + (-x - x^\gamma)) = O(x^\gamma),$$

so the left border of the fundamental strip is $-\gamma$. Second,

$$\lim_{x\to\infty}\frac{x^\gamma}{e^{x+x^\gamma}}=\lim_{x\to\infty}\frac{\gamma x^{1-\gamma}}{e^{x+x^\gamma}(1+\gamma x^{1-\gamma})}=0=O(x^{-m})$$

for any positive $m$, thus the right border of the fundamental strip is $\infty$. The transform $f^*(s)$ exists in the intersection of the domain of convergence of the generalised Dirichlet series and the fundamental strip of $f^*(s)$. This is the overlap between $\langle-\gamma,\infty\rangle$ and $\langle-\infty,(i+1)(1-\gamma)\rangle$, so the transform exists in the strip $\langle-\gamma,(i+1)(1-\gamma)\rangle$. We can choose an $x$ value of 0 again. Then the inverse transform will be

$$f(x)=\frac{1}{2\pi\boldsymbol{i}}\sum_{i\geq 0}\frac{1}{i!}(-p\alpha)^i\alpha p\int_{(0)}\Gamma(s+\gamma(i+1))q^{-s-\gamma(i+1)}\frac{1}{1-q^{i+1-s-\gamma(i+1)}}x^{-s}ds.$$

To approximate this integral, we move the contour right and collect negative residues. In doing this, the first pole we pass is when $i=0$. For the dominant pole, let $\varepsilon:=s+\gamma-1$, and expand around $\varepsilon=0$.

$$\Gamma(s+\gamma)=\Gamma(\varepsilon+1)=\Gamma(1)=1,$$

$$q^{-s-\gamma}=q^{-\varepsilon-1}=q^{-1},$$
$$\frac{1}{1-q^{1-s-\gamma}}\sim\frac{1}{\varepsilon\log q},\qquad\text{(see (12.9))}$$
$$x^{-s}=x^{-\varepsilon+\gamma-1}=x^{\gamma-1}.$$

Putting these together gives a negative residue of

$$-[\varepsilon^{-1}]\alpha pq^{-1}\frac{1}{\varepsilon\log q}x^{\gamma-1}=\frac{\alpha px^{\gamma-1}}{qL}.\tag{13.11}$$

For the other simple poles (which lead to the fluctuations), let $\varepsilon:=s+\gamma-1+\chi_k$. Then around $\varepsilon=0$,

$$\Gamma(s+\gamma)=\Gamma(\varepsilon+1-\chi_k)=\Gamma(1-\chi_k),$$

$$q^{-s-\gamma}=q^{\chi_k-\varepsilon-1}=q^{\chi_k-1},$$
$$\frac{1}{1-q^{1-s-\gamma}}\sim\frac{1}{\varepsilon\log q},\qquad\text{(from (12.9))}$$
$$x^{-s}=x^{\chi_k-\varepsilon+\gamma-1}=x^{\chi_k+\gamma-1}.$$

Altogether, for all values of $k$ except 0, the (negative) residues of these remaining poles will be

$$\sum_{k\neq 0}(-1)[\varepsilon^{-1}]\alpha p\Gamma(1-\chi_k)q^{\chi_k-1}\frac{1}{\varepsilon\log q}x^{\chi_k+\gamma-1}$$

$$
\begin{aligned}
&= \frac{\alpha p x^{\gamma}}{Lqx} \sum_{k\neq 0} \Gamma(1-\chi_k) q^{\chi_k} x^{\chi_k} \\
&= \frac{\alpha p x^{\gamma}}{Lqx} \sum_{k\neq 0} \Gamma(1-\chi_k) e^{\chi_k \log(qx)} \\
&= \frac{\alpha p x^{\gamma}}{Lqx} \sum_{k\neq 0} \Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i}(-\log_Q Q+\log_Q x)} \\
&= \frac{\alpha p x^{\gamma}}{Lqx} \sum_{k\neq 0} \Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i}(\log_Q x-1)} \\
&\sim \frac{\alpha p x^{\gamma}}{Lqx} \sum_{k\neq 0} \Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i}\log_Q x}. \qquad (13.12)
\end{aligned}
$$

We write the probability in terms of $n$ rather than $x$ and sum (13.11) and (13.12) to get

$$
\frac{\alpha p n^{\gamma}}{qLn} + \frac{\alpha p n^{\gamma}}{Lqn} \sum_{k\neq 0} \Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i}\log_Q n}.
$$

Therefore

$$
P_{sw}^{\alpha n^{\gamma}}(M) \sim \frac{(Q-1)\alpha n^{\gamma}}{Ln}(1+\delta(n)),
$$

as $n \to \infty$ for $\delta(x) = \sum_{k\neq 0} \Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i}\log_Q x}$ from (7.5). Hence the proof of Theorem 13.2 is complete. ■

## 13.3 Case (weak, weak), for $d = \alpha n^{\gamma}$

From (8.18), the coefficient of the generating function is

$$
\begin{aligned}
[z^n]F_M^{(w,w)}(z) &= \sum_{k\geq 1} \sum_{i=0}^{\alpha n^{\gamma}-1} (1-q^{k-1})^i p q^{k-1} (1-q^k)^{n-1-i} \\
&= \sum_{k\geq 1} p q^{k-1}(1-q^k)^{n-1} \sum_{i=0}^{\alpha n^{\gamma}-1} \left(\frac{1-q^{k-1}}{1-q^k}\right)^i \\
&= \sum_{k\geq 1} p q^{k-1}(1-q^k)^{n-1} \frac{(1-q^k)^{\alpha n^{\gamma}} - (1-q^{k-1})^{\alpha n^{\gamma}}}{(1-q^k)^{\alpha n^{\gamma}-1} p q^{k-1}} \qquad \text{(by (13.5))} \\
&= \sum_{k\geq 1} (1-q^k)^{n-\alpha n^{\gamma}} \left[(1-q^k)^{\alpha n^{\gamma}} - (1-q^{k-1})^{\alpha n^{\gamma}}\right] \\
&= \sum_{k\geq 1} \left[(1-q^k)^{n} - (1-q^k)^{n-\alpha n^{\gamma}}(1-q^{k-1})^{\alpha n^{\gamma}}\right] \\
&\sim \sum_{k\geq 1} \left[e^{-q^k n} - e^{-q^k(n-\alpha n^{\gamma})} e^{-q^{k-1}\alpha n^{\gamma}}\right] \\
&= \sum_{k\geq 1} \left[e^{-q^k n} - e^{-q^k n - \alpha n^{\gamma} p q^{k-1}}\right].
\end{aligned}
$$

In order to perform the Mellin transform, we define

$$\begin{aligned} f(x) &:= \sum_{k\geq 1}\left[e^{-q^k x} - e^{-q^k x - pq^{k-1}\alpha x^\gamma}\right] \\ &= \sum_{k\geq 1} e^{-q^k x}\left(1 - \sum_{j\geq 0}\frac{(-pq^{k-1}\alpha x^\gamma)^j}{j!}\right) \\ &= -\sum_{k\geq 1} e^{-q^k x}\sum_{j\geq 1}\frac{(-pq^{k-1}\alpha x^\gamma)^j}{j!} \\ &= -\sum_{j\geq 1}\frac{(-pq^{-1}\alpha)^j}{j!}\sum_{k\geq 1} q^{kj}x^{\gamma j}e^{-q^k x}. \end{aligned}$$

This can be transformed to

$$\begin{aligned} f^*(s) &= -\sum_{j\geq 1}\frac{(-pq^{-1}\alpha)^j}{j!}\sum_{k\geq 1} q^{kj}(q^k)^{-s-\gamma j}\Gamma(s+\gamma j) \\ &= -\sum_{j\geq 1}\frac{(-pq^{-1}\alpha)^j}{j!}\Gamma(s+\gamma j)\sum_{k\geq 1} q^{k(j-s-\gamma j)} \\ &= -\sum_{j\geq 1}\frac{(-pq^{-1}\alpha)^j}{j!}\Gamma(s+\gamma j)\frac{q^{j-s-\gamma j}}{1-q^{j-s-\gamma j}}, \qquad \text{for } \Re(s) < j(1-\gamma) \\ &= -\sum_{j\geq 1}\frac{(-p\alpha)^j}{j!}\Gamma(s+\gamma j)q^{-s-\gamma j}\frac{1}{1-q^{j-s-\gamma j}}, \end{aligned}$$

which again exists in the strip $\langle -\gamma, \infty\rangle \cap \langle -\infty, j(1-\gamma)\rangle = \langle -\gamma, j(1-\gamma)\rangle$. We can position our line of integration at 0 on the real axis ($s \in \mathbb{C}$), then inverting this transform gives

$$f(x) = \frac{-1}{2\pi\boldsymbol{i}}\sum_{j\geq 1}\frac{(-p\alpha)^j}{j!}\int_{(0)}\Gamma(s+\gamma j)q^{-s-\gamma j}\frac{1}{1-q^{j-s-\gamma j}}x^{-s}ds.$$

Moving the contour right means that the first poles we reach are at $j = 1$. When $1 - s - \gamma = \chi_k$, we have the fluctuations, but we begin with the main term. For $\varepsilon := s + \gamma - 1$, around $\varepsilon = 0$:

$$\Gamma(s+\gamma) = \Gamma(\varepsilon + 1) = \Gamma(1) = 1,$$

$$\begin{gathered} q^{-s-\gamma} = q^{-1-\varepsilon} = q^{-1}, \\ \frac{1}{1-q^{1-s-\gamma}} \sim \frac{1}{\varepsilon\log q}, \qquad \text{(see (12.9))} \\ x^{-s} = x^{\gamma-1-\varepsilon} = x^{\gamma-1}. \end{gathered}$$

The negative residue of the dominant term is thus:

$$-[\varepsilon^{-1}](-1)(-p\alpha)q^{-1}\frac{1}{\varepsilon\log q}x^{\gamma-1} = \frac{p\alpha x^{\gamma-1}}{qL}. \tag{13.13}$$

For the fluctuations, we look at the simple poles at $1 - s - \gamma = \chi_k$. Define $\varepsilon := -1 + s + \gamma + \chi_k$, to get expansions around $\varepsilon = 0$.

$$\Gamma(s+\gamma) = \Gamma(\varepsilon + 1 - \chi_k) = \Gamma(1 - \chi_k),$$

$$\begin{aligned} q^{-s-\gamma} &= q^{-1-\varepsilon+\chi_k} = q^{\chi_k - 1}, \\ \frac{1}{1 - q^{1-s-\gamma}} &\sim \frac{1}{\varepsilon \log q}, \qquad \text{(from (12.9))} \\ x^{-s} &= x^{\gamma - 1 - \varepsilon + \chi_k} = x^{\gamma + \chi_k - 1}. \end{aligned}$$

Putting these results together means that we get the following contributions from all non-zero values of $k$ ($k = 0$ is the dominant pole).

$$\begin{aligned} -[\varepsilon^{-1}] &\sum_{k\neq 0} (-1)(-p\alpha)\Gamma(1-\chi_k) q^{\chi_k - 1} \frac{1}{\varepsilon \log q} x^{\gamma + \chi_k - 1} \\ &= \frac{p\alpha x^\gamma}{qLx} \sum_{k \neq 0} \Gamma(1-\chi_k) q^{\chi_k} x^{\chi_k} \\ &= \frac{p\alpha x^\gamma}{qLx} \sum_{k \neq 0} \Gamma(1-\chi_k) e^{\chi_k \log(qx)} \\ &= \frac{p\alpha x^\gamma}{qLx} \sum_{k \neq 0} \Gamma(1-\chi_k) e^{\chi_k (\log q + \log x)} \\ &= \frac{p\alpha x^\gamma}{qLx} \sum_{k \neq 0} \Gamma(1-\chi_k) e^{\chi_k (\log x - 1)} \\ &\sim \frac{p\alpha x^\gamma}{qLx} \sum_{k \neq 0} \Gamma(1-\chi_k) e^{\chi_k \log x} \\ &= \frac{p\alpha x^\gamma}{qLx} \sum_{k \neq 0} \Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q x}. \end{aligned} \tag{13.14}$$

In terms of the length of the word ($n$) this probability is asymptotic to (see (13.13) and (13.14))

$$\frac{p\alpha n^\gamma}{qLn} + \frac{p\alpha n^\gamma}{qLn} \sum_{k \neq 0} \Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q n},$$

which gives a result of

$$P_{ww}^{\alpha n^\gamma}(M) \sim \frac{(Q-1)\alpha n^\gamma}{Ln}(1 + \delta(n)),$$

as $n \to \infty$, with $\delta(x)$ as defined in (7.5). This concludes both the proof of Theorem 13.3 and the chapter. ■

# Chapter 14

# Conclusion

Before concluding Part II, we make some observations. The first is an investigation into how some results are affected as $q$ changes. For this we use the probability of the minimum being strict in the first $d$ positions and weak in the rest of the word. Secondly, a comparison is made of the results for the maximum cases. We look at two possible categories in which $d$ can lie.

## 14.1 What happens as $q$ moves from 0 to 1?

It is interesting to look at what happens as the value of $q$ changes, and particularly what happens at the extreme values of $q$, namely 0 and 1. We have already discussed what it means intuitively to have $q = \frac{1}{2}$: we expect about half of our values to be 1, about a quarter to be 2, an eighth to be 3, and so forth. Since $p+q = 1$, if $q = 0$ then $p = 1$ and this means that the probability of a 1 occurring is 1. All other letters occur with probability 0. Thus we have a word consisting only of 1's. As $q$ gets larger more and more larger letters are allowed. So as $q$ tends to 1 the probability of every letter becomes smaller and smaller and we expect each letter in the alphabet to occur only once. At this extremum the probabilities tend to $0 = \frac{1}{\infty}$, so as $n \to \infty$ the word would tend to a permutation of all the natural numbers.

An interesting example of what happens between these two extrema is the probability in the (strict, weak) case for the minimum value occurring in the first $d$ positions of the word. The graphs below demonstrate this for different (fixed) values of $d$. On the horizontal axis, $q$ ranges from 0 to 1 and thus the left and right endpoints represent the scenarios discussed above. The vertical axis represents the probability for the (strict, weak) case for the minimum. This is the probability of having a strict minimum in the first $d$ letters of the word, which is a weak minimum in the rest of the word. We choose $n = 1000000$ and plot the graphs for $d = 1, 2, 3, 4, 5$.

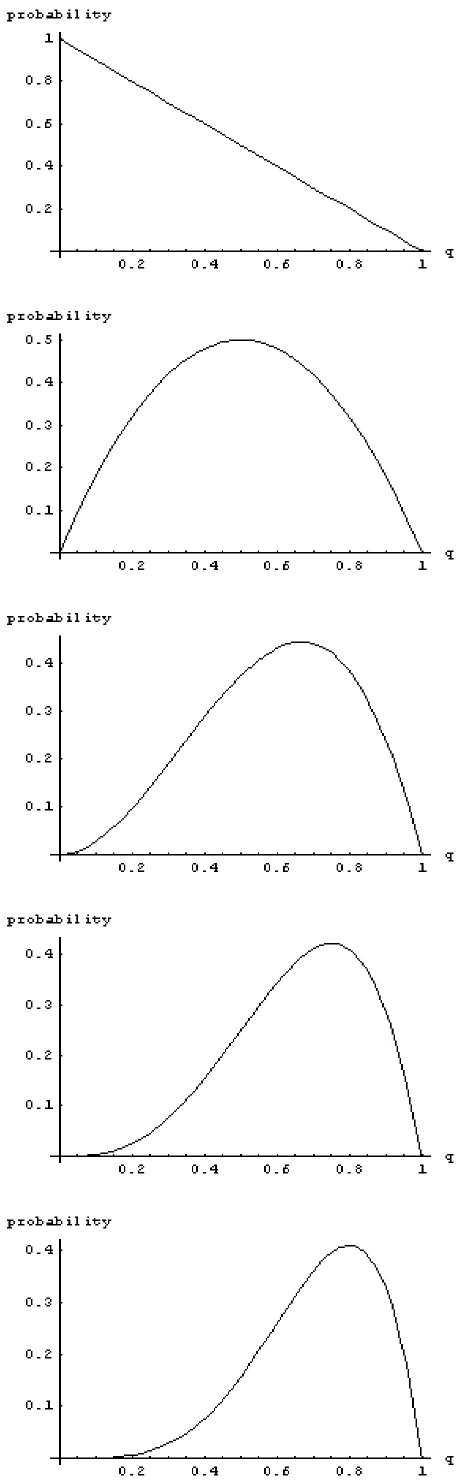


Figure 14.1: Graphs of (strict, weak) minimum probability for $d = 1, \ldots, 5$.

From the above graphs, it appears that the peak occurs at around $\frac{d-1}{d}$ if $d \geq 2$.

The result from Chapter 10 (see 10.3) states that the probability is

$$\frac{Q^{n-d}(Q-1)d}{Q^n-1} = \frac{Q^{-d}(Q-1)d}{1-Q^{-n}} \sim Q^{-d}(Q-1)d$$

as $n \to \infty$. We can express this function in terms of lowercase $q$ as

$$f(q) := (q^{-1}-1)dq^d,$$

which can then be differentiated to confirm this hypothesis:

$$f'(q) = d(d-1)q^{d-2} - d^2q^{d-1}.$$

To find the turning point (where the tangent is horizontal) we put $f'(q) = 0$, and divide by $dq^{d-2}$ to get

$$0 = d - 1 - dq.$$

Solving for $q$ gives the result.

This result indicates that we are mostly likely to get a geometrically distributed word with a strict minimum in the first $d$ positions where the minimum is allowed to repeat in the rest of the word if $q = \frac{d-1}{d}$ (for $d \geq 2$). This comes from the fact that we attach a geometric probability to each letter in the first $d$ positions of a word. If $q = \frac{1}{2}$ then we would expect about half the letters in the word to be 1's. Similarly in the first $d$ letters. So if $d = 2$, we expect a strict minimum if we have a single 1 and one other letter. This situation is maximised by putting $q = \frac{d-1}{d} = \frac{1}{2}$.

The other cases are similar. If $d = 3$, and $q = \frac{d-1}{d} = \frac{2}{3}$, then the probability of getting a 1 in the first $d$ positions is $pq^{j-1} = pq^0 = \frac{1}{3}(\frac{2}{3})^0 = \frac{1}{3}$. I.e., 1 out of every 3 letters will be a 1 (a strict minimum), and more specifically, 1 out of the first $d = 3$ letters will be a strict minimum.

This carries to the cases for general $d \geq 2$. We are most likely to get a word of these specifications (i.e., with a strict minimum in the first $d$ letters which can recur in the remaining letters) if $q = \frac{d-1}{d}$. This is because the probability of getting a 1 in the first $d$ positions is $\frac{1}{d}$, since $p + q = 1$ and the geometric probability attached to 1 is $pq^0 = p$ and $p = 1 - q = 1 - \frac{d-1}{d}$. This is exactly what we require: that 1 out of $d$ letters is a 1 and the rest are any other letters.

If $q < \frac{d-1}{d}$, then the probability of getting a 1 increases, so we are more likely to get a second (or third etc.) occurrence of the minimum (the most likely is 1) in the first $d$ letters of the word which means this word does not fit the specifications. If $q > \frac{d-1}{d}$, then the probability of getting a 1 has decreased and so we are more likely to get a word does not have a 1 in the first $d$ letters of the word and this does not suit the specifications either.

## 14.2 Analysis of results for the maximum

We are now ready to compare the results we have obtained in finding the probability that the maximum in a word (either 'strict' or 'weak') occurs in the first $d$ positions of a word of length $n$. We considered two categories. The first was the original case, where $d$ was fixed relative to $n$, but also where $d = o(n)$ (sometimes denoted $d << n$). The other category was for $d$ proportional to $n$, that is, if we let $d := \alpha n$ where $0 < \alpha \leq 1$.

For the first category, four cases were considered, whereas for $d$ proportional to $n$ we only considered three. This was because the same method worked for both categories in the (strict, strict) case. The (weak, strict), (strict, weak) and (weak, weak) cases required a different method when $d$ grows linearly with $n$. This is the case because of the asymptotics, as explained in the previous chapters.

Table 14.1 shows the dominant term for the results of the two categories for the four cases, expressed in terms of $Q(=\frac{1}{q})$.

| Case | (strict, strict) | (weak, strict) | (strict, weak) | (weak, weak) |
|---|---|---|---|---|
| $1 \leq d = o(n)$ | $\frac{(1-Q^{-1})d}{Ln}$ | $\frac{(1-Q^{-1})d}{Ln}$ | $\frac{(Q-1)d}{Ln}$ | $\frac{(Q-1)d}{Ln}$ |
| $d = \alpha n$ | $\frac{(1-Q^{-1})\alpha}{L}$ | $\frac{1}{L}\log\frac{1}{1-\alpha(1-Q^{-1})}$ | $\frac{\alpha(Q-1)}{L(1+\alpha(Q-1))}$ | $\frac{\log(1+\alpha(Q-1))}{L}$ |

Table 14.1: Summary of maximum results – main term only.

If we consider $\alpha$ small (i.e., close to 0) in the second category, we should get similar solutions to category one (in which $d$ is always small relative to $n$ for $n$ large). We thus determine what these dominant terms look like asymptotically as $\alpha \to 0$. We use approximations $\log(1+x) \sim x$ and $\frac{1}{1-x} \sim 1$ as $x \to 0$ (see [36]). Suppose $d = \alpha n$, then for the (weak, strict) case, we have

$$\frac{\log(1-\alpha(1-Q^{-1}))}{\log Q^{-1}} \sim \frac{-\alpha(1-Q^{-1})}{\log Q^{-1}} = \frac{\alpha(1-Q^{-1})}{L}.$$

For the (strict, weak) case, we find that

$$\frac{\alpha(Q-1)}{L(1+\alpha(Q-1))} \sim \frac{\alpha(Q-1)}{L},$$

and for the (weak, weak) case,

$$\frac{\log(1+\alpha(Q-1))}{L} \sim \frac{\alpha(Q-1)}{L}.$$

By replacing $\alpha$ by $\frac{d}{n}$, it can be seen that each of these corresponds to the results when $1 \leq d = o(n)$ (see Table 14.1 above).

## 14.3 Concluding remarks

In Part II we found the probabilities of having the maximum and/or minimum occurring in specific positions in a word of length $n$ whose letters are natural numbers which occur independently and with geometrically probability. It was found that the weak/strict classification had more effect on the results for the minima than the maxima, and also that on the whole the classification had more sway in the latter part of the word rather than the first $d$ positions. This is because in general $n$ is considerably larger than $d$, and usually that $n \to \infty$. The minima probabilities (for $d$ fixed) were $O(q^n)$ if the second restriction (referring to the last $n-d$ letters of the word) was strict, and a constant (relative to $n$) if it was weak. The probability that the minimum value of the first $d$ letters was the maximum of the rest was $O(\frac{1}{n^d})$ in all cases for $d$ fixed. For the probability of finding the maximum occurring in the first $d$ positions of a word, we considered more options and let $d$ grow with $n$. The two categories discussed were $1 \leq d = o(n)$ and $d = \alpha n$ where $0 < \alpha \leq 1$. For the former the probabilities were either $\frac{(1-q)d}{Ln}$ or $\frac{(1-q)d}{Lqn}$, depending on whether the second restriction was strict or weak. For small $\alpha$, the second category's results were the same.

# Part III

# Binary Search Trees

# Chapter 15

# Introduction

We examine binary search trees formed from sequences with equal entries. A binary search tree is a planar tree where each node has a maximum of 2 children. These trees are created from input sequences (where repeats are allowed) as follows: the first element in the sequence is the root of the tree and thereafter elements which are strictly less than the parent node are placed to the left (as the left child) and those greater than *or equal to* the parent node, are inserted as the right child. For example, the binary search tree of the sequence 323123411343 would be drawn as follows:

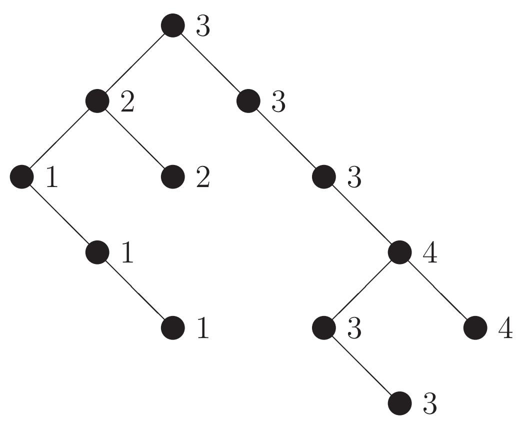


Figure 15.1: The binary search tree of sequence 323123411343.

The sequences (words) are now no longer generated with geometric distribution, but rather are created from finite alphabets according to two models, as discussed in Section 15.1.

Previous research on this topic includes:

Burge (1976) [4] uses a recursive argument to find the expected value of the left-going depth (number of left-going branches from root to key) of the first 1 inserted into a binary tree produced from the multiset $\{m_1{\cdot}1\ ;\ m_2{\cdot}2\ ;\ m_3{\cdot}3\ ;\ \ldots\ ;\ m_n{\cdot}n\}$. Burge uses similar methods to find the right-going depths of keys and ultimately the expected depth.

Sedgewick (1977) [35] wrote a paper on this subject from the point of view of a computer scientist. This paper deals with the quicksort applications of this analysis, and the author notes that "we can be fairly certain that conclusions that we draw based on the number of comparisons will carry through to the total running time".

Kemp (1996) [14] also looks at the left/right depth of a given key using two models (the same models used in this thesis) – one where the input sequence is composed of elements of a given multiset with all possible sequences equally likely, and one where the input sequence is $n$ elements chosen (independently) from a finite set of elements, each with some specified probability. Kemp's approach makes use of recurrences, whereas the approach outlined below is symbolic.

## 15.1 Method

We describe the situation in a similar way to Kemp, but use generating functions rather than probabilistic recurrence arguments (see [14]) to find the expected value and variance.

### 15.1.1 The 'multiset' model

The first model we use assumes that we have input sequences of length $n$, formed from the multiset $\{n_1 \cdot 1\ ;\ n_2 \cdot 2\ ;\ \ldots\ ;\ n_r \cdot r\}$. That is, we assume we know how many times the letter occurs in the sequence, and we let $n_i$ denote the number of times the letter $i$ occurs in the sequence. The multinomial $\binom{n}{n_1,\ldots,n_r}$ expresses how many sequences there are, and all are equally likely to occur. We have $n_1 + n_2 + \cdots + n_r = n$. It suffices to consider the alphabet $\{1, 2, \ldots, r\}$, as we are only interested in the letters relative to each other. Any other alphabet with such an ordering would be dealt with in the same way, ('we use the notation "the $i$th smallest key" and "the key $i$" synonymously' [14, page 40]). Hence the assumption that $n_i > 0$, for $i \in \{1, \ldots, r\}$. Manipulating the bivariate generating function of all such sequences makes it possible for us to find the parameters we want (the three different cases we consider are discussed below), and dividing by the number of possible sequences gives the expected value. Then further manipulation allows us to find the variance.

### 15.1.2 The 'probability' model

The second model is a probabilistic model, sometimes called the 'memoryless' model or the 'Bernoulli' model. A probability is attached to every letter in the alphabet, so the letter $i$ would appear in the sequence with probability $p_i$. The sequence of length $n$ consists of letters chosen independently from this alphabet. We assume that the probabilities of the letters in the alphabet add up to 1, and that each probability is non-zero. The probability distribution function is thus well-defined.

## 15.2 Cases

The three parameters or 'cases' discussed are:

### 15.2.1 Left-going depth of the first 1

The left-going depth of the first 1 is the number of left-going branches from the root to the node corresponding to the first 1. It is numerically equal to one fewer than the number of strict left-to-right minima. This is because we create a left-going branch only when we have a strict left-to-right minima, but we will end up with one extra because the first letter of the word (first element of the sequence) will also count as a strict left-to-right minimum, but does not create a left-going branch.

In Figure 15.1, the left-going depth of the first 1 is 2, and the number of (strict) left-to-right minima in the corresponding sequence (**32**3**1**123411343) is 3.

The result obtained in Theorem 16.1 below gives the expected value of the left-going depth of the first 1 in all binary search trees formed from a particular multiset. Theorem 16.2 gives the same result for a particular alphabet whose letters have specific probabilities. As an example, consider the sequences 323123411343, 212411433333, 432343321131, and 123123443331. All of these sequences are built from the multiset $\{3 \cdot 1\ ;\ 2 \cdot 2\ ;\ 5 \cdot 3\ ;\ 2 \cdot 4\}$, or could be drawn independently from the alphabet $\{1, 2, 3, 4\}$ where 1 occurs with probability $\frac{1}{4}$; 2 and 4 occur with probability $\frac{1}{6}$; and 3 occurs with probability $\frac{5}{12}$. The left-going depths of the first 1 in each sequence are respectively 2,1,3, and 0. To find the average one needs to consider all possible sequences, add all the left-going depths of the first 1, and divide by the total number of sequences (in this case $\binom{12}{3,2,5,2} = 166320$). The result is not necessarily a whole number.

### 15.2.2 Right-going depth of the first $r$

If our alphabet is $\{1, 2, \ldots, r\}$, finding the right-going depth of the first $r$ is equivalent to finding the number of weak left-to-right maxima up to the first occurrence of $r$, subtract one. In this case the fact that we allow equal keys is more relevant than in the previous case (finding the left-going depth of the first 1) because of the way we form the tree by putting keys strictly less than on the left and greater than *or equal to* on the right of the parent node.

For example, in Figure 15.1, (where the alphabet is $\{1, \ldots, 4\}$) the right-going depth of the first 4 is 3, and the number of (weak) left-to-right maxima up to the first occurrence of 4 in the sequence (**3**2**3**12**34**11343) is 4.

Again, this is a specific example, and the expected value (see Theorems 17.1 and 17.2) would be the average of all such results.

### 15.2.3 Expected depth of an arbitrary key $\alpha$

This is created by summing the right-going and left-going depths. We use the idea of a 'shuffle' operator. Here, we only allow one appearance of each letter in the binary search tree. First, we calculate the left/right-going depths (i.e., the (strict) left-to-right minima of elements $\{\alpha+1, \ldots, r\}$ and the (strict) left-to-right maxima of elements $\{1, \ldots, \alpha-1\}$) and then 'shuffle' these sequences and concatenate them with the first $\alpha$ followed by the rest of the sequence. This idea is described in more detail in Chapter 18.

## 15.3 Notation

The expected value and variance are found for each of the three cases. Also, there are two versions of each result which correspond to the multiset model (an exact result) and the probability model (an asymptotic result). The following notation is used: $\mathbb{E}$ and $\mathbb{V}$ represent the expectation and variance. Their subscripts 'lg', 'rg' and '$\alpha$' denote the three cases, namely left-going depth of the first 1, right-going depth of the first $r$, where the alphabet is $\{1, \ldots, r\}$ and average depth of any key $\alpha$. The model used is represented by the superscript 'm' or 'p'. For the multiset model, $n_i$ stands for how many times the letter $i$ appears in the sequence. In the probability model, $p_i$ represents the probability that the letter $i$ will occur in the sequence. As a shorthand, we denote $N_{[2,5]} = n_2 + n_3 + n_4 + n_5$ or $P_{[4,4]} = p_4$. We assume $N_{[i,j]} = P_{[i,j]} = 0$ for $i > j$, and we let $n$ be the length of the sequence.

# Chapter 16

# Left-going depth of first 1

## 16.1 Introduction

Sequences are created according to the multiset model (number of appearances of each letter fixed) and the probability model (each letter appears independently with a certain probability). Generating functions then express the situation and the moments are found by differentiating partially.

The left-going depth of the first 1 is the number of left-going branches which must be followed from the root of the tree to the first node labelled 1. This is also the longest path on the left-most side of the tree. By counting the number of nodes which must be passed while travelling from the root to this node, one obtains a count of exactly one more than the number of branches. The result is found by counting these nodes, each of which corresponds to a strict left-to-right minimum in the input sequence.

The results of this chapter are:

**Theorem 16.1** *The expected value of the left-going depth of the first one is (multiset model)*

$$\mathbb{E}_{\mathrm{lg}}^{\mathrm{m}} = \sum_{k=2}^{r} \frac{n_k}{N_{[1,k]}}.$$

**Theorem 16.2** *According to the probability model, the expectation of the left-going depth of the first one is*

$$\mathbb{E}_{\mathrm{lg}}^{\mathrm{p}} = \sum_{i=2}^{r} \frac{p_i}{P_{[1,i]}}\left(1 - P_{[i+1,r]}^{n}\right) \sim \sum_{i=2}^{r} \frac{p_i}{P_{[1,i]}},$$

*as $n \to \infty$.*

**Theorem 16.3** *The left-going depth of the first one has a variance of*

$$\mathbb{V}_{\mathrm{lg}}^{\mathrm{m}} = \sum_{k=2}^{r} \frac{n_k}{N_{[k,r]}} - \sum_{k=2}^{r} \frac{n_k^2}{N_{[k,r]}^2},$$

*by the multiset model.*

**Theorem 16.4** *There exists a real $\lambda$, $0 \le \lambda < 1$, such that the variance of the left-going depth of the first one is*

$$\begin{aligned}\mathbb{V}_{\mathrm{lg}}^{\mathrm{p}} &= \left( \sum_{i=2}^{r} \frac{p_i}{P_{[1,i]}} - \sum_{i=2}^{r} \frac{p_i^2}{P_{[1,i]}^2} \right) \bigl(1 + O(\lambda^n)\bigr) \\ &\sim \sum_{i=2}^{r} \frac{p_i}{P_{[1,i]}} - \sum_{i=2}^{r} \frac{p_i^2}{P_{[1,i]}^2},\end{aligned}$$

*as $n \to \infty$, by the probability model.*

## 16.2 Left-going expectation – multiset model

Suppose we have an alphabet of $\{1, 2, \ldots, r\}$ from which we construct binary search trees of $n$ nodes with repeats, where we have $n_1$ 1's, $n_2$ 2's etc. so that $n_1 + n_2 + \cdots + n_r = n$ and sequences are formed from the multiset $\{n_1 {\cdot} 1 \ ; \ n_2 {\cdot} 2 \ ; \ \ldots \ ; \ n_r {\cdot} r\}$. Finding the left-going depth of the first 1 in the binary search tree is the same as counting the number of strict left-to-right minima and subtracting one. We can express all possible words of this form (i.e., of length $n$ with letters chosen from our alphabet) symbolically as

$$\begin{aligned}\{1, \ldots, r\}^* = \bigl(\varepsilon + r\{r\}^*\bigr)\bigl(\varepsilon + (r-1)\{r-1, r\}^*\bigr)\bigl(\varepsilon + (r-2)\{r-2, r-1, r\}^*\bigr) \\ \cdots \bigl(\varepsilon + 1\{1, \ldots, r\}^*\bigr).\end{aligned}$$

Note that the presence of $\varepsilon$ indicates that we may not have an $r$ (respectively, an $r-1, \ldots, 1$) in the sequence. This does not affect the result but simplifies the calculations. All it means is that we create more words than we need initially but afterwards we look at the coefficient only for the cases that make sense in our problem (i.e., $n_i > 0$).

This symbolic equation can be expressed as a generating function where $z$ counts every letter, $u$ counts all (except the last) of the left-to-right minima (which will correspond to the relevant left-going branches of the corresponding tree), and $x_1, \ldots, x_r$ respectively mark the number of 1's, ..., $r$'s. We use the shorthand notation of $X_{[i,j]} = x_i + x_{i+1} + \cdots + x_{j-1} + x_j$ and write

$$f(z,u,x_1,\ldots,x_r) := \prod_{i=2}^{r}\left(1+\frac{zux_i}{1-zX_{[i,r]}}\right)\left(1+\frac{zx_1}{1-zX_{[1,r]}}\right). \tag{16.1}$$

Here $u$ can be seen to count only the values $r, r-1, \ldots, 2$ which are strict left-to-right minima.

For the first moment we want the partial derivative of $f$ with respect to $u$. Since derivatives of products in general are quite tedious, we use the logarithmic derivative instead. This means differentiating a sum rather than a product. We use $(\log f)' = \frac{f'}{f}$ to give $f' = f(\log f)'$, which we then evaluate at $u = 1$ to get the moment. Now if $u = 1$, then $f$ is a telescoping series, so

$$\begin{aligned}
f\big|_{u=1} &= \prod_{i=1}^{r}\left(1+\frac{zx_i}{1-zX_{[i,r]}}\right)\\
&= \prod_{i=1}^{r}\left(\frac{1-zX_{[i,r]}+zx_i}{1-zX_{[i,r]}}\right)\\
&= \prod_{i=1}^{r}\left(\frac{1-zX_{[i+1,r]}}{1-zX_{[i,r]}}\right)\\
&= \left(\frac{1-zX_{[2,r]}}{1-zX_{[1,r]}}\right)\left(\frac{1-zX_{[3,r]}}{1-zX_{[2,r]}}\right)\left(\frac{1-zX_{[4,r]}}{1-zX_{[3,r]}}\right)\cdots\left(\frac{1-zX_{[r+1,r]}}{1-zX_{[r,r]}}\right)\\
&= \frac{1}{1-zX_{[1,r]}},
\end{aligned} \tag{16.2}$$

since $X_{[r,r]} = x_r$ and $X_{[r+1,r]} = 0$. This is to be expected as we consider all possibilities if we do not put a restriction on the number of left-to-right minima (i.e., if $u = 1$). The other factor is

$$\begin{aligned}
\frac{\partial}{\partial u}\log f\Big|_{u=1} &= \sum_{i=2}^{r}\frac{\partial}{\partial u}\log\left(1+\frac{zux_i}{1-zX_{[i,r]}}\right)\Bigg|_{u=1}\\
&= \sum_{i=2}^{r}\frac{\partial}{\partial u}\log\left(\frac{1-zX_{[i,r]}+zux_i}{1-zX_{[i,r]}}\right)\Bigg|_{u=1}\\
&= \sum_{i=2}^{r}\frac{1}{\frac{1-zX_{[i,r]}+zux_i}{1-zX_{[i,r]}}}\cdot\frac{zx_i}{1-zX_{[i,r]}}\Bigg|_{u=1}\\
&= \sum_{i=2}^{r}\frac{zx_i}{1-zX_{[i,r]}+zux_i}\Bigg|_{u=1}\\
&= \sum_{i=2}^{r}\frac{zx_i}{1-zX_{[i,r]}+zx_i}\\
&= \sum_{i=2}^{r}\frac{zx_i}{1-zX_{[i+1,r]}}.
\end{aligned} \tag{16.3}$$

(Both of the above calculations can also be done by Maple or similar.) Thus

$$\frac{\partial}{\partial u} f\Big|_{u=1} = \frac{1}{1 - zX_{[1,r]}} \sum_{i=2}^{r} \frac{zx_i}{1 - zX_{[i+1,r]}}.$$

In order to get the coefficient of $z^n$, we consider a typical term in this sum to which we can apply partial fraction decomposition to get

$$\begin{aligned}
&[z^n]\frac{zx_i}{(1 - zX_{[1,r]})(1 - zX_{[i+1,r]})} \\
&= [z^n]\frac{x_i}{X_{[1,i]}}\left(\frac{1}{1 - zX_{[1,r]}} - \frac{1}{1 - zX_{[i+1,r]}}\right) \\
&= \frac{x_i}{X_{[1,i]}}\left(X_{[1,r]}^n - X_{[i+1,r]}^n\right) \\
&= \frac{x_i}{X_{[1,i]}}\left(\left(X_{[1,i]} + X_{[i+1,r]}\right)^n - X_{[i+1,r]}^n\right) \\
&= \frac{x_i}{X_{[1,i]}}\left(\sum_{k=0}^{n}\binom{n}{k} X_{[1,i]}^k X_{[i+1,r]}^{n-k} - X_{[i+1,r]}^n\right) \\
&= \frac{x_i}{X_{[1,i]}}\sum_{k=1}^{n}\binom{n}{k} X_{[1,i]}^k X_{[i+1,r]}^{n-k} \\
&= x_i \sum_{k=1}^{n}\binom{n}{k} X_{[1,i]}^{k-1} X_{[i+1,r]}^{n-k} \\
&= x_i \sum_{k=1}^{n}\binom{n}{k} \sum_{j_1+\cdots+j_i=k-1}\binom{k-1}{j_1,\ldots,j_i} x_1^{j_1}\cdots x_i^{j_i} \\
&\qquad\cdot \sum_{j_{i+1}+\cdots+j_r=n-k}\binom{n-k}{j_{i+1},\ldots,j_r} x_{i+1}^{j_{i+1}}\cdots x_r^{j_r} \\
&= \sum_{k=1}^{n}\sum_{j_1+\cdots+j_i=k-1}\sum_{j_{i+1}+\cdots+j_r=n-k}\binom{n}{k}\binom{k-1}{j_1,\ldots,j_i}\binom{n-k}{j_{i+1},\ldots,j_r} \\
&\qquad\qquad \cdot x_1^{j_1}\cdots x_i^{j_i+1} x_{i+1}^{j_{i+1}}\cdots x_r^{j_r}.
\end{aligned}$$

To find the expected value or first moment, we want to the find the coefficient of $x_p^{n_p}$, $\forall\, p = 1, \ldots, r$ in the above expression. We do this by equating all $n_p$'s with the $j_p$'s, except for $n_i$ which is $j_i + 1$. Thus (recall $n = n_1 + \cdots + n_r$, and let $N_{[i,r]} := n_i + \cdots + n_r$)

$$\begin{aligned}
&[z^n x_1^{n_1}\cdots x_r^{n_r}]\frac{\partial}{\partial u} f\Big|_{u=1} \\
&= [z^n x_1^{n_1}\cdots x_r^{n_r}]\sum_{i=2}^{r}\frac{zx_i}{(1 - zX_{[1,r]})(1 - zX_{[i+1,r]})} \\
&= \sum_{i=2}^{r}\binom{n}{j_1+\cdots+j_i+1}\binom{j_1+\cdots+j_i}{j_1,\ldots,j_i}\binom{j_{i+1}+\cdots+j_r}{j_{i+1},\ldots,j_r}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=2}^{r} \binom{n}{N_{[1,i]}} \binom{N_{[1,i]}-1}{n_1,\ldots,n_i-1} \binom{N_{[i+1,r]}}{n_{i+1},\ldots,n_r} \\
&= \sum_{i=2}^{r} \frac{n!(N_{[1,i]}-1)!(N_{[i+1,r]})!}{(N_{[1,i]})!(N_{[i+1,r]})!n_1!\cdots(n_i-1)!n_{i+1}!\cdots n_r!} \\
&= \sum_{i=2}^{r} \frac{n_i}{N_{[1,i]}} \binom{n}{n_1,\ldots,n_r}.
\end{aligned}
$$

To get the expected value, we divide through by the total number of possibilities for words of length $n$ from the alphabet $\{1,\ldots,r\}$, i.e., $\binom{n}{n_1,\ldots,n_r}$. Thus the expected value for the left-going depth of the first 1 – i.e., the length of the path consisting only of left-going branches is

$$
\mathbb{E}_{\mathrm{lg}}^{\mathrm{m}} = \sum_{i=2}^{r} \frac{n_i}{N_{[1,i]}}, \tag{16.4}
$$

as in [4, page 453]. This concludes the proof of Theorem 16.1. ■

## 16.3 Left-going expectation – probability model

The same solution can be found using the probability model which also uses generating functions but the calculations are simpler as probabilities are substituted for the 'place-holder' variables $x_1,\ldots,x_r$. It works as follows: Suppose we wish to generate all possible words of length $n$, where letter $i$ occurs with probability $p_i$. We express the generating function in a different way. The first line is the previous definition and the second is the new step.

$$
\begin{aligned}
f(z,u,x_1,\ldots,x_r) &:= \prod_{i=2}^{r} \left(1+\frac{zux_i}{1-zX_{[i,r]}}\right)\left(1+\frac{zx_1}{1-zX_{[1,r]}}\right) \\
&= \sum_{w\in A^*} u^{l(w)} z^{|w|} x_1^{|w|_1}\cdots x_r^{|w|_r},
\end{aligned} \tag{16.5}
$$

where $w$ is any word from alphabet $A=\{1,\ldots,r\}$, $l(w)$ is the left-going depth of the first one (we need not know what this function is defined to be), $|w|$ is the length of the word (previously called $n$) and $|w|_i$ is the number of times the letter $i$ occurs in the word, which was denoted $n_i$ for the multiset model. With this model, the expected value for a word of length $|w|=n$ can be found using

$$
\mathbb{E}_{\mathrm{lg}}^{\mathrm{p}} = [z^n]\frac{\partial}{\partial u} f(z,u,p_1,\ldots,p_r)\Big|_{u=1}
$$

where $p_i$ is the probability of the letter $i$ occurring. So we have an easy way to find the expected value, especially since we have just performed this derivative in the previous model. So in this case the expected value is

$$\mathbb{E}_{\text{lg}}^{\text{p}} = [z^n]\frac{\partial}{\partial u}\prod_{i=2}^{r}\left(1+\frac{zup_i}{1-zP_{[i,r]}}\right)\left(1+\frac{zp_1}{1-zP_{[1,r]}}\right),$$

whose partial derivative was found via logarithmic differentiation to be a product of (refer to (16.2) and (16.3) from Section 16.2)

$$f\Big|_{u=1} = \frac{1}{1-zP_{[1,r]}}$$

and

$$\frac{\partial}{\partial u}\log f\Big|_{u=1} = \sum_{i=2}^{r}\frac{zp_i}{1-zP_{[i+1,r]}}.$$

Again, using partial fractions this can be rewritten as (note that $0 < p_i < 1$ here is a probability and not a counter (i.e., $\in \mathbb{N}$) for the number of $i$'s as in [4])

$$\begin{aligned}
\mathbb{E}_{\text{lg}}^{\text{p}} &= [z^n]\sum_{i=2}^{r}\frac{1}{1-zP_{[1,r]}}\frac{zp_i}{1-zP_{[i+1,r]}} \\
&= [z^n]\sum_{i=2}^{r}\frac{p_i}{P_{[1,i]}}\left(\frac{1}{1-zP_{[1,r]}}-\frac{1}{1-zP_{[i+1,r]}}\right) \\
&= \sum_{i=2}^{r}\frac{p_i}{P_{[1,i]}}\big(P_{[1,r]}^n - P_{[i+1,r]}^n\big).
\end{aligned}$$

Now, since the $p_i$'s are probabilities, we have $P_{[1,r]} = p_1 + \cdots + p_r = 1$. Also, $P_{[i+1,r]} < 1$, hence as $n \to \infty$, $P_{[i+1,r]}^n \to 0$, leaving the result:

$$\mathbb{E}_{\text{lg}}^{\text{p}} \sim \sum_{i=2}^{r}\frac{p_i}{P_{[1,i]}}, \tag{16.6}$$

as $n \to \infty$. The proof of Theorem 16.2 is thus complete. ■

It can now be seen that we get a similar result for both models. However, the former is in terms of $n_i$'s, which are integer values representing how many $i$'s there are in the word, whereas the latter is in terms of $p_i$'s which are fractional values representing probabilities. Intuitively it seems plausible that by associating $n_i$ with $p_i n$ for large $n$ (and $i = 1, \ldots, r$), the results of Theorems 16.1 and 16.2 should not be too far from each other. We thus compare expression (16.4) with (16.6) as follows:

$$\sum_{i=2}^{r}\frac{n_i}{N_{[1,i]}} \sim \sum_{i=2}^{r}\frac{p_i n}{p_1 n + p_2 n + p_i n} = \sum_{i=2}^{r}\frac{p_i}{P_{[1,i]}},$$

as $n \to \infty$.

## 16.4 Left-going variance – multiset model

We have already found the expected value of the left-going depth of the first 1 in a binary tree with repeats allowed using generating functions. We now find the variance by first finding the second moment using this generating function. The generating function is still (see equation (16.1), $X_{[i,j]} = x_i + \cdots + x_j$ etc.)

$$f(z, u, x_1, \ldots, x_r) = \prod_{i=2}^{r} \left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\left(1 + \frac{zx_1}{1 - zX_{[1,r]}}\right),$$

and the variance is given by

$$\begin{aligned}\mathbb{V}_{\mathrm{lg}}^{\mathrm{m}} &:= [z^n x_1^{n_1} \cdots x_r^{n_r}]\frac{\partial^2}{\partial u^2} f(z, 1, x_1, \ldots, x_r) + [z^n x_1^{n_1} \cdots x_r^{n_r}]\frac{\partial}{\partial u} f(z, 1, x_1, \ldots, x_r) \\ &\quad - \left([z^n x_1^{n_1} \cdots x_r^{n_r}]\frac{\partial}{\partial u} f(z, 1, x_1, \ldots, x_r)\right)^2. \end{aligned} \tag{16.7}$$

To find the second moment (i.e., the first term in (16.7)), we can use Maple (or use the same trick as before, twice over i.e., since $(\log f)'' = f''f^{-1} - (f')^2 f^{-2}$, we have $f'' = f(\log f)'' + f'(\log f)'$ which is easier to calculate). Either way, we end up with (for $X = x_1 + \cdots + x_r$)

$$f''|_{u=1} = \frac{1}{1 - zX}\left(\left(\sum_{i=2}^{r} \frac{zx_i}{1 - zX_{[i+1,r]}}\right)^2 - \sum_{i=2}^{r} \frac{z^2 x_i^2}{(1 - zX_{[i+1,r]})^2}\right). \tag{16.8}$$

A typical term of the first sum would be $\frac{zx_i}{1-zX_{[i+1,r]}} \cdot \frac{zx_j}{1-zX_{[j+1,r]}}$, where $i$ and $j$ both run from 2 to $r$. All the terms where $i = j$ will be of the form $\frac{z^2x_i^2}{(1-zX_{[i+1,r]})^2}$, where $i$ runs from 2 to $r$. These are all cancelled by the second sum, leaving only those where $i \neq j$. Because this is symmetric (i.e., $i$ and $j$ can be swapped to give the same term again), we can include a factor of 2 and write this as

$$\begin{aligned} f''|_{u=1} &= 2\frac{1}{1 - zX}\sum_{i=2}^{r}\sum_{j=i+1}^{r} \frac{zx_i}{1 - zX_{[i+1,r]}} \cdot \frac{zx_j}{1 - zX_{[j+1,r]}} \\ &= 2\sum_{i=2}^{r}\sum_{j=i+1}^{r} \frac{z^2 x_i x_j}{(1 - zX)(1 - zX_{[i+1,r]})(1 - zX_{[j+1,r]})} \\ &= 2\sum_{i=2}^{r}\sum_{j=i+1}^{r} z^2 x_i x_j \sum_{k\geq 0} z^k X^k \sum_{l\geq 0} z^l X_{[i+1,r]}^l \sum_{m\geq 0} z^m X_{[j+1,r]}^m. \end{aligned}$$

We are interested in coefficients of this quantity, and we start by looking at that of $z^n$, and then consider the $x_i$'s. Put $n := k + l + m + 2$ to get

$$[z^n]f''|_{u=1} = 2\sum_{i=2}^{r}\sum_{j=i+1}^{r} x_i x_j \sum_{k\geq n-2} X^k \sum_{n-k-l-2\geq 0} X_{[i+1,r]}^l X_{[j+1,r]}^{n-k-l-2}$$

$$= 2\sum_{i=2}^{r}\sum_{j=i+1}^{r}\sum_{k=0}^{n-2}\sum_{l=0}^{n-k-2} \underbrace{x_i x_j X^k X_{[i+1,r]}^l X_{[j+1,r]}^{n-k-l-2}}_{\spadesuit}.$$

For ease of notation we simplify the bracketed portion on its own, so

$$\begin{aligned}
\spadesuit &= x_i x_j (X_{[1,i]} + X_{[i+1,r]})^k X_{[i+1,r]}^l X_{[j+1,r]}^{n-k-l-2} \\
&= x_i x_j \sum_{s=0}^{k} \binom{k}{s} X_{[1,i]}^s X_{[i+1,r]}^{k-s} X_{[i+1,r]}^l X_{[j+1,r]}^{n-k-l-2} \\
&= x_i x_j \sum_{s=0}^{k} \binom{k}{s} X_{[1,i]}^s X_{[i+1,r]}^{k-s+l} X_{[j+1,r]}^{n-k-l-2} \\
&= x_i x_j \sum_{s=0}^{k} \binom{k}{s} X_{[1,i]}^s (X_{[i+1,j]} + X_{[j+1,r]})^{k-s+l} X_{[j+1,r]}^{n-k-l-2} \\
&= x_i x_j \sum_{s=0}^{k} \binom{k}{s} X_{[1,i]}^s \sum_{t=0}^{k-s+l} \binom{k-s+l}{t} X_{[i+1,j]}^t X_{[j+1,r]}^{k-s+l-t} X_{[j+1,r]}^{n-k-l-2} \\
&= x_i x_j \sum_{s=0}^{k} \binom{k}{s} X_{[1,i]}^s \sum_{t=0}^{k-s+l} \binom{k-s+l}{t} X_{[i+1,j]}^t X_{[j+1,r]}^{k-s+l-t+n-k-l-2} \\
&= x_i x_j \sum_{s=0}^{k} \sum_{t=0}^{k-s+l} \binom{k}{s} \binom{k-s+l}{t} X_{[1,i]}^s X_{[i+1,j]}^t X_{[j+1,r]}^{n-s-t-2} \\
&= x_i x_j \sum_{s=0}^{k} \sum_{t=0}^{k-s+l} \binom{k}{s} \binom{k-s+l}{t} \sum_{b_1+\cdots+b_i=s} \binom{b_1+\cdots+b_i}{b_1,\ldots,b_i} x_1^{b_1}\cdots x_i^{b_i} \qquad (16.9) \\
&\quad \cdot \sum_{b_{i+1}+\cdots+b_j=t} \binom{b_{i+1}+\cdots+b_j}{b_{i+1},\ldots,b_j} x_{i+1}^{b_{i+1}}\cdots x_j^{b_j} \\
&\quad \cdot \sum_{b_{j+1}+\cdots+b_r=n-s-t-2} \binom{b_{j+1}+\cdots+b_r}{b_{j+1},\ldots,b_r} x_{j+1}^{b_{j+1}}\cdots x_r^{b_r}.
\end{aligned}$$

We are now ready to take the coefficient of $x_m^{n_m}$ for $m = 1, \ldots, r$. We can equate $b_m = n_m$ for all $m$ values except $i$ and $j$, for which we have, respectively: $n_i = b_i + 1$ and $n_j = b_j + 1$ owing to the presence of the factor $x_i x_j$ in line (16.9) above. Consider the complete expression to get

$$\begin{aligned}
&[z^n x_1^{n_1} \cdots x_r^{n_r}] f''|_{u=1} \\
&= 2\sum_{i=2}^{r}\sum_{j=i+1}^{r}\sum_{k=0}^{n-2}\sum_{l=0}^{n-k-2} \binom{k}{b_1+\cdots+b_i} \binom{k+l-(b_1+\cdots+b_i)}{b_{i+1}+\cdots+b_j} \\
&\qquad \cdot \binom{b_1+\cdots+b_i}{b_1,\ldots,b_i} \binom{b_{i+1}+\cdots+b_j}{b_{i+1},\ldots,b_j} \binom{b_{j+1}+\cdots+b_r}{b_{j+1},\ldots,b_r}
\end{aligned}$$

$$= 2\sum_{i=2}^{r}\sum_{j=i+1}^{r}\underbrace{\sum_{k=0}^{n-2}\binom{k}{N_{[1,i]}-1}\overbrace{\sum_{l=0}^{n-k-2}\binom{k+l-(N_{[1,i]}-1)}{N_{[i+1,j]}-1}}^{\clubsuit}}_{\diamondsuit}$$

$$\cdot\binom{N_{[1,i]}-1}{n_1,\ldots,n_i-1}\binom{N_{[i+1,j]}-1}{n_{i+1},\ldots,n_j-1}\binom{N_{[j+1,r]}}{n_{j+1},\ldots,n_r}. \tag{16.10}$$

By letting $w := k+l-(N_{[1,i]}-1)$, we can write

$$\begin{aligned}
\clubsuit &= \sum_{l=0}^{n-k-2}\binom{k+l-(N_{[1,i]}-1)}{N_{[i+1,j]}-1}\\
&= \sum_{w=k-(N_{[1,i]}-1)}^{k-(N_{[1,i]}-1)+n-k-2}\binom{w}{N_{[i+1,j]}-1}\\
&= \sum_{w=k-(N_{[1,i]}-1)}^{n-(N_{[1,i]}-1)-2}\binom{w}{N_{[i+1,j]}-1}\\
&= \sum_{w=k-(N_{[1,i]}-1)}^{N_{[i+1,r]}-1}\binom{w}{N_{[i+1,j]}-1}\\
&= \sum_{w=0}^{N_{[i+1,r]}-1}\binom{w}{N_{[i+1,j]}-1} - \sum_{w=0}^{k-N_{[1,i]}}\binom{w}{N_{[i+1,j]}-1}\\
&= \binom{N_{[i+1,r]}}{N_{[i+1,j]}} - \binom{k-N_{[1,i]}+1}{N_{[i+1,j]}},
\end{aligned}$$

by 'upper summation' in [12, page 174]. Making use of 'upper summation' again and the identity $\binom{r}{m}\binom{m}{k} = \binom{r}{k}\binom{r-k}{m-k}$ ('trinomial revision' [12, page 174]) we can write

$$\begin{aligned}
\diamondsuit &= \sum_{k=0}^{n-2}\binom{k}{N_{[1,i]}-1}\left[\binom{N_{[i+1,r]}}{N_{[i+1,j]}} - \binom{k-N_{[1,i]}+1}{N_{[i+1,j]}}\right]\\
&= \sum_{k=0}^{n-2}\binom{k}{N_{[1,i]}-1}\binom{N_{[i+1,r]}}{N_{[i+1,j]}} - \sum_{k=0}^{n-2}\binom{k}{N_{[1,i]}-1}\binom{k-N_{[1,i]}+1}{N_{[i+1,j]}}\\
&= \binom{n-1}{N_{[1,i]}}\binom{N_{[i+1,r]}}{N_{[i+1,j]}} - \sum_{k=0}^{n-2}\binom{k}{N_{[1,j]}-1}\binom{N_{[1,j]}-1}{N_{[1,i]}-1}\\
&= \binom{n-1}{N_{[1,i]}}\binom{N_{[i+1,r]}}{N_{[i+1,j]}} - \binom{n-1}{N_{[1,j]}}\binom{N_{[1,j]}-1}{N_{[1,i]}-1}.
\end{aligned}$$

The quantity $\diamondsuit$ can now be substituted back into equation (16.10) to get

$$[z^n x_1^{n_1}\cdots x_r^{n_r}]f''|_{u=1}$$

$$= 2\sum_{i=2}^{r}\sum_{j=i+1}^{r}\binom{n-1}{N_{[1,i]}}\binom{N_{[i+1,r]}}{N_{[i+1,j]}}$$

$$
\begin{aligned}
&\qquad \cdot \binom{N_{[1,i]}-1}{n_1,\ldots,n_i-1}\binom{N_{[i+1,j]}-1}{n_{i+1},\ldots,n_j-1}\binom{N_{[j+1,r]}}{n_{j+1},\ldots,n_r} \\
&\quad - 2\sum_{i=2}^{r}\sum_{j=i+1}^{r}\binom{n-1}{N_{[1,j]}}\binom{N_{[1,j]}-1}{N_{[1,i]}-1} \\
&\qquad \cdot \binom{N_{[1,i]}-1}{n_1,\ldots,n_i-1}\binom{N_{[i+1,j]}-1}{n_{i+1},\ldots,n_j-1}\binom{N_{[j+1,r]}}{n_{j+1},\ldots,n_r} \\
&= 2\sum_{i=2}^{r}\sum_{j=i+1}^{r}\frac{(n-1)!N_{[i+1,r]}!}{N_{[1,i]}!(N_{[i+1,r]}-1)!N_{[i+1,j]}!N_{[j+1,r]}!} \\
&\qquad \cdot \frac{(N_{[1,i]}-1)!(N_{[i+1,j]}-1)!N_{[j+1,r]}!}{n_1!\cdots(n_i-1)!n_{i+1}!\cdots(n_j-1)!n_{j+1}!\cdots n_r!} \\
&\quad - 2\sum_{i=2}^{r}\sum_{j=i+1}^{r}\frac{(n-1)!(N_{[1,j]}-1)!}{N_{[1,j]}!(N_{[j+1,r]}-1)!(N_{[1,i]}-1)!N_{[i+1,j]}!} \\
&\qquad \cdot \frac{(N_{[1,i]}-1)!(N_{[i+1,j]}-1)!N_{[j+1,r]}!}{n_1!\cdots(n_i-1)!n_{i+1}!\cdots(n_j-1)!n_{j+1}!\cdots n_r!} \\
&= 2\sum_{i=2}^{r}\sum_{j=i+1}^{r}\frac{n_in_jN_{[i+1,r]}}{nN_{[1,i]}N_{[i+1,j]}}\binom{n}{n_1,\ldots,n_r} - 2\sum_{i=2}^{r}\sum_{j=i+1}^{r}\frac{n_in_jN_{[j+1,r]}}{nN_{[1,j]}N_{[i+1,j]}}\binom{n}{n_1,\ldots,n_r} \\
&= 2\sum_{i=2}^{r}\sum_{j=i+1}^{r}\frac{n_in_jN_{[i+1,r]}N_{[1,j]}-n_in_jN_{[j+1,r]}N_{[1,i]}}{nN_{[1,i]}N_{[1,j]}N_{[i+1,j]}}\binom{n}{n_1,\ldots,n_r} \\
&= 2\sum_{i=2}^{r}\sum_{j=i+1}^{r}\frac{n_in_j(N_{[i+1,j]}+N_{[j+1,r]})(N_{[1,i]}+N_{[i+1,j]})-n_in_jN_{[j+1,r]}N_{[1,i]}}{nN_{[1,i]}N_{[1,j]}N_{[i+1,j]}}\binom{n}{n_1,\ldots,n_r} \\
&= 2\sum_{i=2}^{r}\sum_{j=i+1}^{r}n_in_j\frac{N_{[i+1,j]}(N_{[1,i]}+N_{[i+1,j]})+N_{[j+1,r]}N_{[i+1,j]}}{nN_{[1,i]}N_{[1,j]}N_{[i+1,j]}}\binom{n}{n_1,\ldots,n_r} \\
&= 2\sum_{i=2}^{r}\sum_{j=i+1}^{r}n_in_j\frac{N_{[i+1,j]}N_{[1,r]}}{nN_{[1,i]}N_{[1,j]}N_{[i+1,j]}}\binom{n}{n_1,\ldots,n_r} \\
&= 2\sum_{i=2}^{r}\sum_{j=i+1}^{r}\frac{n_in_j}{N_{[1,i]}N_{[1,j]}}\binom{n}{n_1,\ldots,n_r},
\end{aligned}
$$

since $n = N_{[1,r]}$. Dividing by the total number of words of such a multiset, $\binom{n}{n_1,\ldots,n_r}$, gives us the second moment. We now recall (equation (16.7)) that to calculate the variance we also need to add the expected value and subtract the square of the expected value (see equation (16.4)). So we have a variance of:

$$
2\sum_{i=2}^{r}\sum_{j=i+1}^{r}\frac{n_in_j}{N_{[1,i]}N_{[1,j]}} + \sum_{i=2}^{r}\frac{n_i}{N_{[i,r]}} - \left(\sum_{i=2}^{r}\frac{n_i}{N_{[i,r]}}\right)^2. \tag{16.11}
$$

As in the beginning of this calculation, we use the idea of splitting up a squared sum into terms which are perfect squares and those which are not. It can be seen that the terms in the second moment (i.e., the first term in (16.11)) correspond to the terms in the squared expected value which are not squares. This simplifies the variance to

$$\mathbb{V}_{\text{lg}}^{\text{m}} = \sum_{i=2}^{r} \frac{n_i}{N_{[i,r]}} - \sum_{i=2}^{r} \frac{n_i^2}{N_{[i,r]}^2}. \tag{16.12}$$

This concludes the proof of Theorem 16.3. ■

## 16.5 Left-going variance – probability model

We now calculate the variance using the probability model, which we recall from the expected value calculations. The generating function is (see equation (16.5))

$$\begin{aligned} f(z,u,x_1,\dots,x_r) &= \prod_{i=2}^{r} \left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\left(1 + \frac{zx_1}{1 - zX_{[1,r]}}\right) \\ &= \sum_{w \in A^*} u^{l(w)} z^{|w|} x_1^{|w|_1} \cdots x_r^{|w|_r}, \end{aligned}$$

where $w$ is the word from alphabet $A = \{1,\dots,r\}$, $l(w)$ is the left-going depth of the first one, $|w|$ is the length of the word (this was called $n$ before) and $|w|_i$ is the number of times the letter $i$ occurs in the word. With this model, the variance for a word of length $|w| = n$ is

$$\begin{aligned} \mathbb{V}_{\text{lg}}^{\text{p}} = [z^n] \frac{\partial^2}{\partial u^2} f(z,u,p_1,\dots,p_r)\Big|_{u=1} + [z^n] \frac{\partial}{\partial u} f(z,u,p_1,\dots,p_r)\Big|_{u=1} \\ - \left([z^n] \frac{\partial}{\partial u} f(z,u,p_1,\dots,p_r)\Big|_{u=1}\right)^2. \end{aligned} \tag{16.13}$$

Again, $p_i$ is the probability that the letter $i$ occurs. Equation (16.6) gives the second and third terms, and we use our previously-found derivative (above, see equation (16.8)) for the first term (the second moment), which we calculate now.

$$\begin{aligned} &[z^n] \frac{\partial^2}{\partial u^2} f(z,u,p_1,\dots,p_r)\Big|_{u=1} \\ &\quad = [z^n] \frac{1}{1 - zP} \left(\left(\sum_{i=2}^{r} \frac{zp_i}{1 - zP_{[i+1,r]}}\right)^2 - \sum_{i=2}^{r} \frac{z^2 p_i^2}{(1 - zP_{[i+1,r]})^2}\right). \end{aligned}$$

As explained in the multiset model, the squared terms (where we multiply a term by itself) in the first sum cancel with the second sum. This leaves all the other terms (twice each due to symmetry). The triple product can be decomposed into partial fractions, as shown below.

$$\begin{aligned} &[z^n] \frac{\partial^2}{\partial u^2} f(z,u,p_1,\dots,p_r)\Big|_{u=1} \\ &\quad = [z^n] \frac{2}{1 - zP} \sum_{i=2}^{r} \sum_{j=i+1}^{r} \frac{zp_i}{1 - zP_{[i+1,r]}} \cdot \frac{zp_j}{1 - zP_{[j+1,r]}} \end{aligned}$$

$$
\begin{aligned}
&= [z^n] \sum_{i=2}^{r} \sum_{j=i+1}^{r} \frac{2z^2 p_i p_j}{(1-zP)(1-zP_{[i+1,r]})(1-zP_{[j+1,r]})} \\
&= [z^n] \sum_{i=2}^{r} \sum_{j=i+1}^{r} \left[ \frac{2p_i p_j}{(1-zP)P_{[1,i]}P_{[1,j]}} - \frac{2p_i p_j}{(1-zP_{[i+1,r]})P_{[i+1,j]}P_{[1,i]}} \right. \\
&\qquad\qquad \left. + \frac{2p_i p_j}{(1-zP_{[j+1,r]})P_{[1,j]}P_{[i+1,j]}} \right] \\
&= \sum_{i=2}^{r} \sum_{j=i+1}^{r} \left[ \frac{2p_i p_j}{P_{[1,i]}P_{[1,j]}} P^n - \frac{2p_i p_j}{P_{[i+1,j]}P_{[1,i]}} P^n_{[i+1,r]} + \frac{2p_i p_j}{P_{[1,j]}P_{[i+1,j]}} P^n_{[j+1,r]} \right].
\end{aligned}
$$

Finally, we note that $P = p_1 + \cdots + p_r = 1$ and also that $P_{[i+1,r]} < 1$ and $P_{[j+1,r]} < 1$. This inequality is strict since $p_i \neq 0$, $\forall\, i = 2, \ldots, r$. So if we consider what happens as $n \to \infty$,

$$
\lim_{n\to\infty} P^n = 1, \qquad \lim_{n\to\infty} P^n_{[i+1,r]} = 0, \quad \text{and} \quad \lim_{n\to\infty} P^n_{[j+1,r]} = 0,
$$

which means that we do not need to consider the second two terms if we take the limit as $n \to \infty$. Thus

$$
[z^n] \frac{\partial^2}{\partial u^2} f(z, u, p_1, \ldots, p_r) \Big|_{u=1} \sim \sum_{i=2}^{r} \sum_{j=i+1}^{r} \frac{2p_i p_j}{P_{[1,i]}P_{[1,j]}},
$$

as $n \to \infty$. To get the variance we must add the expected value and subtract its square. These quantities can be obtained from the probability model expected value section (see (16.6)).

$$
\mathbb{V}^{\mathrm{p}}_{\mathrm{lg}} \sim \sum_{i=2}^{r} \sum_{j=i+1}^{r} \frac{2p_i p_j}{P_{[1,i]}P_{[1,j]}} + \sum_{i=2}^{r} \frac{p_i}{P_{[1,i]}} - \left( \sum_{i=2}^{r} \frac{p_i}{P_{[1,i]}} \right)^2,
$$

as $n \to \infty$. Again, cancellations can take place to simplify this expression. The first sum represents all the terms of the last sum (the square of the expectation) which are not squares, and thus both of these disappear, leaving

$$
\mathbb{V}^{\mathrm{p}}_{\mathrm{lg}} \sim \sum_{i=2}^{r} \frac{p_i}{P_{[1,i]}} - \sum_{i=2}^{r} \frac{p_i^2}{P^2_{[1,i]}}, \tag{16.14}
$$

as $n \to \infty$, which completes the proof of Theorem 16.4. ■

As in the expected value results, replacing $n_i$ by $p_i \cdot n$ in equation (16.12) will yield the result in (16.14).

# Chapter 17

# Right-going depth of first $r$

## 17.1 Introduction

We find the expected value and variance of the right-going depth of the first $r$ in a binary search tree formed from a word drawn from the alphabet $\{1, \ldots, r\}$ and where repeats are allowed. In this case the fact that we allow equal keys is more relevant than in the previous case (finding the left-going depth of the first 1) because of the way we form the tree by putting keys strictly less than the parent node to its left and those greater than or equal to the parent node to the right. Multivariate generating functions are used to determine this expected value. We use $n$ to represent the total number of letters in the word and we denote the number of 1's by $n_1$, the number of 2's by $n_2$ and so on. Thus $n_1 + n_2 + \cdots + n_r = n$.

In this chapter, we prove the following:

**Theorem 17.1** *The multiset model gives the expected value of the right-going depth of the first $r$ as*

$$\mathbb{E}^{\mathrm{m}}_{\mathrm{rg}} = \sum_{i=1}^{r-1} \frac{n_i}{N_{[i+1,r]} + 1}.$$

**Theorem 17.2** *Using the probability model, the expected value of the right-going depth of the first $r$ is*

$$\mathbb{E}^{\mathrm{p}}_{\mathrm{rg}} \sim \sum_{i=1}^{r-1} \frac{p_i}{P_{[i+1,r]}},$$

*as $n \to \infty$.*

**Theorem 17.3** *The variance of the right-going depth of the first $r$ according to the multiset model is*

$$\mathbb{V}^{\mathrm{m}}_{\mathrm{rg}} = 2\sum_{i=1}^{r-1} \frac{n_i(n_i-1)}{(N_{[i+1,r]}+1)(N_{[i+1,r]}+2)} + \sum_{i=1}^{r-1} \frac{n_i}{N_{[i+1,r]}+1} - \sum_{i=1}^{r-1} \frac{n_i^2}{(N_{[i+1,r]}+1)^2}.$$

**Theorem 17.4** *By the probability model, this variance is*

$$\mathbb{V}^{\mathrm{p}}_{\mathrm{rg}} \sim \sum_{i=1}^{r-1} \frac{p_i}{P_{[i+1,r]}} + \sum_{i=1}^{r-1} \frac{p_i^2}{P^2_{[i+1,r]}},$$

*as $n \to \infty$.*

## 17.2 Right-going expectation – multiset model

One way to express symbolically all words from the alphabet $\{1, \dots, r\}$ is

$$(\varepsilon + 1\{1\}^*)(\varepsilon + 2\{1,2\}^*)(\varepsilon + 3\{1,2,3\}^*)...(\varepsilon + r\{1,..,r\}^*),$$

where $\varepsilon$ represents an empty sub-word. This symbolic expression can be translated into a function which will generate all possible words from the alphabet. In this generating function, $z$ will count every letter in the word (or every node/key in the tree) and $u$ will count only those nodes which will cause a right-going branch (this corresponds to the weak left-to-right maxima up to – but not including – the first occurrence of $r$). Thus we have a probability generating function with respect to $u$, and the expected value and variance can be calculated by differentiating with respect to $u$. The other variables are $x_1, x_2, \dots, x_r$ where $x_i$ counts the number of times the letter $i$ appears in the word. So for this situation we have the following generating function (where we write $X_{[i,j]} = x_i + \cdots + x_j$):

$$f(z, u, x_1, \dots, x_r) := \prod_{i=1}^{r-1} \left(1 + \frac{zux_i}{1 - z(X_{[1,i-1]} + ux_i)}\right)\left(1 + \frac{zx_r}{1 - zX_{[1,r]}}\right). \quad (17.1)$$

Thus $u$ will count only the right-going branches (the number of weak left-to-right maxima, not including those corresponding to $r$'s in the word). Now we are interested in finding the partial derivative with respect to $u$ and evaluating this at $u = 1$. We then want the coefficient of $z^n x_1^{n_1} x_2^{n_2} \cdots x_r^{n_r}$. Now, $(\log f)' = \frac{f'}{f}$ implies that $f' = f(\log f)'$. In this way we can change the log product into a sum of logs and find the derivative more easily. The derivative is the product of

$$f\big|_{u=1} = \prod_{i=1}^{r} \left(1 + \frac{zx_i}{1 - zX_{[1,i]}}\right)$$

$$
\begin{aligned}
&= \prod_{i=1}^{r} \left( \frac{1 - zX_{[1,i]} + zx_i}{1 - zX_{[1,i]}} \right) \\
&= \prod_{i=1}^{r} \left( \frac{1 - zX_{[1,i-1]}}{1 - zX_{[1,i]}} \right) \\
&= \left( \frac{1}{1 - zx_1} \right) \left( \frac{1 - zx_1}{1 - zX_{[1,2]}} \right) \left( \frac{1 - zX_{[1,2]}}{1 - zX_{[1,3]}} \right) \cdots \left( \frac{1 - zX_{[1,r-1]}}{1 - zX_{[1,r]}} \right) \\
&= \frac{1}{1 - zX_{[1,r]}},
\end{aligned}
$$

(which represents the generating function for all words drawn from the alphabet $\{1, \ldots, r\}$ with no restrictions) and

$$
\begin{aligned}
\frac{\partial}{\partial u} \log f \Big|_{u=1} &= \frac{\partial}{\partial u} \log \prod_{i=1}^{r-1} \left( 1 + \frac{zux_i}{1 - z\left(X_{[1,i-1]} + ux_i\right)} \right) \left( 1 + \frac{zx_r}{1 - zX_{[1,r]}} \right) \Big|_{u=1} \\
&= \sum_{i=1}^{r-1} \frac{\partial}{\partial u} \log \left( 1 + \frac{zux_i}{1 - z\left(X_{[1,i-1]} + ux_i\right)} \right) \Big|_{u=1} \\
&= \sum_{i=1}^{r-1} \frac{1}{1 + \frac{zux_i}{1 - z\left(X_{[1,i-1]} + ux_i\right)}} \frac{\partial}{\partial u} \left( 1 + \frac{zux_i}{1 - z\left(X_{[1,i-1]} + ux_i\right)} \right) \Big|_{u=1} \\
&= \sum_{i=1}^{r-1} \frac{1}{1 + \frac{zux_i}{1 - z\left(X_{[1,i-1]} + ux_i\right)}} \frac{\partial}{\partial u} \Big[ zux_i \big(1 - z(X_{[1,i-1]} + ux_i)\big)^{-1} \Big] \Big|_{u=1} \\
&= \sum_{i=1}^{r-1} \frac{1}{\frac{1 - z\left(X_{[1,i-1]} + ux_i\right) + zux_i}{1 - z\left(X_{[1,i-1]} + ux_i\right)}} \Big[ zx_i \big(1 - z(X_{[1,i-1]} + ux_i)\big)^{-1} \\
&\qquad + zux_i(-1)\big(1 - z(X_{[1,i-1]} + ux_i)\big)^{-2}(-zx_i) \Big] \Big|_{u=1} \\
&= \sum_{i=1}^{r-1} \frac{1 - z\left(X_{[1,i-1]} + ux_i\right)}{1 - z\left(X_{[1,i-1]} + ux_i\right) + zux_i} \\
&\qquad \cdot \left[ \frac{zx_i\big(1 - z(X_{[1,i-1]} + ux_i)\big) + zux_i(-1)(-zx_i)}{\big(1 - z(X_{[1,i-1]} + ux_i)\big)^2} \right] \Bigg|_{u=1} \\
&= \sum_{i=1}^{r-1} \frac{1 - z\left(X_{[1,i-1]} + ux_i\right)}{1 - z\left(X_{[1,i-1]} + ux_i\right) + zux_i} \\
&\qquad \cdot \left[ \frac{zx_i - z^2x_iX_{[1,i-1]} - z^2x_i^2u + z^2ux_i^2}{\big(1 - z(X_{[1,i-1]} + ux_i)\big)^2} \right] \Bigg|_{u=1} \\
&= \sum_{i=1}^{r-1} \frac{1 - z\left(X_{[1,i-1]} + ux_i\right)}{1 - z\left(X_{[1,i-1]} + ux_i\right) + zux_i} \frac{zx_i - z^2x_iX_{[1,i-1]}}{\big(1 - z(X_{[1,i-1]} + ux_i)\big)^2} \Bigg|_{u=1} \\
&= \sum_{i=1}^{r-1} \frac{1 - zX_{[1,i-1]} - zx_i}{1 - zX_{[1,i-1]} - zx_i + zx_i} \frac{zx_i - z^2x_iX_{[1,i-1]}}{\big(1 - z(X_{[1,i-1]} + x_i)\big)^2} \\
&= \sum_{i=1}^{r-1} \frac{1 - zX_{[1,i]}}{1 - zX_{[1,i-1]}} \frac{zx_i(1 - zX_{[1,i-1]})}{\big(1 - zX_{[1,i]}\big)^2}
\end{aligned}
$$

$$= \sum_{i=1}^{r-1} \frac{zx_i}{1 - zX_{[1,i]}}.$$

So our expression becomes

$$f'\Big|_{u=1} = \frac{1}{1 - zX_{[1,r]}} \sum_{i=1}^{r-1} \frac{zx_i}{1 - zX_{[1,i]}} \tag{17.2}$$
$$= \sum_{i=1}^{r-1} \left( \frac{x_i}{X_{[i+1,r]}} \frac{1}{1 - zX_{[1,r]}} - \frac{x_i}{X_{[i+1,r]}} \frac{1}{1 - zX_{[1,i]}} \right),$$

as a partial fraction decomposition. From this we wish to determine the coefficient of $z^n x_1^{n_1} x_2^{n_2} \cdots x_r^{n_r}$. We start by finding the coefficient of $z^n$:

$$\begin{aligned}
[z^n]\frac{\partial}{\partial u} f\Big|_{u=1} &= \sum_{i=1}^{r-1} \frac{x_i}{X_{[i+1,r]}} \left( X_{[1,r]}^n - X_{[1,i]}^n \right) \\
&= \sum_{i=1}^{r-1} \frac{x_i}{X_{[i+1,r]}} \left( \sum_{k=0}^{n} \binom{n}{k} X_{[i+1,r]}^k X_{[1,i]}^{n-k} - X_{[1,i]}^n \right) \\
&= \sum_{i=1}^{r-1} x_i \sum_{k=1}^{n} \binom{n}{k} X_{[i+1,r]}^{k-1} X_{[1,i]}^{n-k} \\
&= \sum_{i=1}^{r-1} x_i \sum_{k=1}^{n} \binom{n}{k} \sum_{j_{i+1}+\cdots+j_r = k-1} \binom{k-1}{j_{i+1}, \ldots, j_r} x_{i+1}^{j_{i+1}} \cdots x_r^{j_r} \\
&\qquad \cdot \sum_{j_1+\cdots+j_i = n-k} \binom{n-k}{j_1, \ldots, j_i} x_1^{j_1} \cdots x_i^{j_i} \\
&= \sum_{i=1}^{r-1} \sum_{k=1}^{n} \sum_{j_{i+1}+\cdots+j_r=k-1} \sum_{j_1+\cdots+j_i=n-k} \\
&\qquad \cdot \binom{n}{k} \binom{k-1}{j_{i+1}, \ldots, j_r} \binom{n-k}{j_1, \ldots, j_i} x_1^{j_1} \cdots x_i^{j_i+1} x_{i+1}^{j_{i+1}} \cdots x_r^{j_r},
\end{aligned}$$

and then equate $n_p = j_p, \forall\, p \in \{1, \ldots, r\}$, $p \neq i$, and $n_i = j_i + 1$ to get the coefficient of $z^n x_1^{n_1} x_2^{n_2} \cdots x_r^{n_r}$ to be

$$\begin{aligned}
&\sum_{i=1}^{r-1} \binom{n}{N_{[i+1,r]} + 1} \binom{N_{[i+1,r]}}{n_{i+1}, \ldots, n_r} \binom{N_{[1,i]} - 1}{n_1, \ldots, n_i - 1} \\
&= \sum_{i=1}^{r-1} \frac{n!}{(N_{[i+1,r]} + 1)!(N_{[1,i]} - 1)!} \frac{N_{[i+1,r]}!}{n_{i+1}! \ldots n_r!} \frac{(N_{[1,i]} - 1)!}{n_1! \ldots (n_i - 1)!} \\
&= \sum_{i=1}^{r-1} \frac{n_i n!}{(N_{[i+1,r]} + 1) n_1! \ldots n_i! n_{i+1}! \ldots n_r!} \\
&= \sum_{i=1}^{r-1} \frac{n_i}{N_{[i+1,r]} + 1} \binom{n}{n_1, \ldots, n_r}.
\end{aligned}$$

If we then divide by the number of all possible words $\binom{n}{n_1,\dots,n_r}$, we get the expected value of the right-going depth of the first $r$ in a word of length $n$ with $n_i$ occurrences of the letter $i$, $i \in \{1,\dots,r\}$. It is

$$\mathbb{E}^{\mathrm{m}}_{\mathrm{rg}} = \sum_{i=1}^{r-1} \frac{n_i}{N_{[i+1,r]}+1}. \tag{17.3}$$

This is the same result as in [4], but with slightly different notation. Theorem 17.1 is thus proved. ■

In the next chapter, a similar expression is found, but with an $n_i$ instead of the one in the denominator. This corresponds to the fact that for this case we insert all the letters from the input sequence into the tree, whereas to find an arbitrary key $\alpha$, we construct a binary search tree where all the keys are distinct, from an input sequence with repeated keys.

## 17.3 Right-going expectation – probability model

The generating function for the probability model in the right-going case is

$$\begin{aligned} f(z,u,x_1,\dots,x_r) &= \prod_{i=1}^{r-1}\left(1+\frac{zux_i}{1-z\big(X_{[1,i-1]}+ux_i\big)}\right)\left(1+\frac{zx_r}{1-zX_{[1,r]}}\right) \\ &= \sum_{w\in A^*} z^{|w|}u^{r(w)}x_1^{|w|_1}x_2^{|w|_2}\cdots x_r^{|w|_r}, \end{aligned} \tag{17.4}$$

where $w$ is a word with length $|w|$ and $|w|_i$ is the number of $i$'s appearing in the word $w$. We choose letters from the alphabet $A = \{1,\dots,r\}$ to form each word. In this case we want $u$ to count the number of right-going branches from the root to the first occurrence of letter $r$ in a binary search tree with equal keys. This is formed from an input sequence whose (possibly repeated) letters are chosen from the alphabet $\{1,\dots,r\}$. Note that it is not necessary to know the definition of the function $r(w)$ which measures the right-going path to the first $r$. Also note that this is not necessarily the longest path to the right as in the left-going case, since we may have repeats of the letter $r$ which would lead to subsequent right-going branches which are not counted.

Now that we have expressed the generating function as a sum, we can find the expected value of this right-going depth in a tree with these criteria by taking the partial derivative with respect to $u$, which will make $r(w)$ a coefficient. We then substitute $u=1$ into the function as $u$ is no longer needed. After that we find the coefficient of $z^n$ which means we are interested in words of length $n$, and lastly we substitute $p_i$ for $x_i$, where $p_i$ is the probability with which the letter $i$ occurs in the word. In this way, we multiply the right-going depth of the first $r$ $(r(w))$ by the probability of each tree occurring – i.e., the product of the probabilities of all the letters. We multiply the probability $p_1$ $|w|_1$ times, and find the product of this with the probability $p_2$ $|w|_2$ times etc. up to letter $r$ whose probability is $p_r$ and we must multiply this to our product $|w|_r$ times.

More concisely, this means that

$$\mathbb{E}_{\mathrm{rg}}^{\mathrm{p}} := [z^n]\frac{\partial}{\partial u} f(z, u, p_1, \ldots, p_r)\Big|_{u=1}.$$

The fact that we have already found this partial derivative helps us here, since we know that (see multiset model above, equation (17.2))

$$\begin{aligned}
\mathbb{E}_{\mathrm{rg}}^{\mathrm{p}} &= [z^n]\frac{1}{1 - zP_{[1,r]}} \sum_{i=1}^{r-1} \frac{zp_i}{(1 - zP_{[1,i]})} \\
&= [z^n]\sum_{i=1}^{r-1} \left( \frac{p_i}{P_{[i+1,r]}} \frac{1}{(1 - zP_{[1,r]})} - \frac{p_i}{P_{[i+1.r]}} \frac{1}{(1 - zP_{[1,i]})} \right) \\
&= \sum_{i=1}^{r-1} \left( \frac{p_i}{P_{[i+1,r]}} P_{[1,r]}^n - \frac{p_i}{P_{[i+1.r]}} P_{[1,i]}^n \right).
\end{aligned}$$

We now recall that $P_{[1,r]}^n = 1^n = 1$ and that $P_{[1,i]} < 1,\ \forall\ i = 1, \ldots, r-1$, and so $P_{[1,i]}^n \to 0$ as $n \to \infty$. Thus

$$\mathbb{E}_{\mathrm{rg}}^{\mathrm{p}} \sim \sum_{i=1}^{r-1} \frac{p_i}{P_{[i+1,r]}}, \tag{17.5}$$

as $n \to \infty$. This concludes the proof of Theorem 17.2. ■

Now suppose we refer back to the multiset model result in equation (17.3). By associating $n_i$ with $np_i$ for large $n$, we can write

$$\begin{aligned}
\mathbb{E}_{\mathrm{rg}}^{\mathrm{m}} &= \sum_{i=1}^{r-1} \frac{n_i}{1 + N_{[i+1,r]}} \\
&\sim \sum_{i=1}^{r-1} \frac{np_i}{1 + np_{i+1} + \cdots + np_r} \\
&= \sum_{i=1}^{r-1} \frac{p_i}{\frac{1}{n} + p_{i+1} + \cdots + p_r} \\
&\sim \sum_{i=1}^{r-1} \frac{p_i}{P_{[i+1,r]}}, \qquad \text{as } n \to \infty,
\end{aligned}$$

and this corresponds to the probability model result in equation (17.5).

## 17.4 Right-going variance – multiset model

We now investigate the variance of the right-going depth of the first $r$ in a binary search tree with equal keys whose alphabet is $\{1,\ldots,r\}$. The generating function is the same as in equation (17.1):

$$f(z,u,x_1,\ldots,x_r) := \prod_{i=1}^{r-1}\left(1+\frac{zux_i}{1-z\left(X_{[1,i-1]}+ux_i\right)}\right)\left(1+\frac{zx_r}{1-zX_{[1,r]}}\right),$$

from which we will find

$$\begin{aligned}\mathbb{V}_{\text{rg}}^{\text{m}} := {}&[z^n x_1^{n_1}\cdots x_r^{n_r}]\frac{\partial^2}{\partial u^2}f(z,1,x_1,\ldots,x_r)+[z^n x_1^{n_1}\cdots x_r^{n_r}]\frac{\partial}{\partial u}f(z,1,x_1,\ldots,x_r)\\ &-\left([z^n x_1^{n_1}\cdots x_r^{n_r}]\frac{\partial}{\partial u}f(z,1,x_1,\ldots,x_r)\right)^2. \end{aligned} \tag{17.6}$$

To do this we start with the second moment and use Maple to give us:

$$\begin{aligned}\frac{\partial^2}{\partial u^2}f(z,1,x_1,\ldots,x_r) &= f(z,1,x_1,\ldots,x_r)\cdot\frac{\partial^2}{\partial u^2}\log f(z,1,x_1,\ldots,x_r)\\ &\quad+\frac{\partial}{\partial u}f(z,1,x_1,\ldots,x_r)\cdot\frac{\partial}{\partial u}\log f(z,1,x_1,\ldots,x_r)\\ &=\frac{1}{1-zX_{[1,r]}}\sum_{i=1}^{r-1}\frac{\partial}{\partial u}\frac{zx_i}{(1-zX_{[1,i-1]}-zux_i)}\bigg|_{u=1}\\ &\quad+\frac{1}{1-zX_{[1,r]}}\left(\sum_{i=1}^{r-1}\frac{zx_i}{(1-zX_{[1,i-1]}-zux_i)}\right)^2\bigg|_{u=1}\\ &=\frac{1}{1-zX_{[1,r]}}\sum_{i=1}^{r-1}\frac{z^2x_i^2}{(1-zX_{[1,i-1]}-zux_i)^2}\bigg|_{u=1}\\ &\quad+\frac{1}{1-zX_{[1,r]}}\left(\sum_{i=1}^{r-1}\frac{zx_i}{(1-zX_{[1,i-1]}-zux_i)}\right)^2\bigg|_{u=1}\\ &=\frac{1}{1-zX_{[1,r]}}\left[\sum_{i=1}^{r-1}\frac{z^2x_i^2}{(1-zX_{[1,i]})^2}+\left(\sum_{i=1}^{r-1}\frac{zx_i}{(1-zX_{[1,i]})}\right)^2\right]\\ &=\underbrace{2\sum_{i=1}^{r-1}\frac{z^2x_i^2}{(1-zX_{[1,r]})(1-zX_{[1,i]})^2}}_{A} \qquad (17.7)\\ &\quad\underbrace{+2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{z^2x_ix_j}{(1-zX_{[1,r]})(1-zX_{[1,i]})(1-zX_{[1,j]})}}_{B}.\end{aligned}$$

We deal with $A$ and $B$ separately, and note that $X := X_{[1,r]}$.

$$A := 2\sum_{i=1}^{r-1}\frac{z^2x_i^2}{(1-zX)(1-zX_{[1,i]})^2}$$

$$
\begin{aligned}
&= 2\sum_{i=1}^{r-1} z^2 x_i^2 \sum_{j\geq 0} z^j X^j \left(\sum_{k\geq 0} z^k X_{[1,i]}^k\right)^2 \\
&= 2\sum_{i=1}^{r-1} z^2 x_i^2 \sum_{j\geq 0} z^j X^j \sum_{k\geq 0}\sum_{l=0}^{k} X_{[1,i]}^l X_{[1,i]}^{k-l} z^k \\
&= 2\sum_{i=1}^{r-1} z^2 x_i^2 \sum_{j\geq 0} z^j X^j \sum_{k\geq 0}\sum_{l=0}^{k} X_{[1,i]}^k z^k \\
&= 2\sum_{i=1}^{r-1} z^2 x_i^2 \sum_{j\geq 0} z^j X^j \sum_{k\geq 0}(k+1) X_{[1,i]}^k z^k.
\end{aligned}
$$

So

$$
\begin{aligned}
[z^n]A &= 2\sum_{i=1}^{r-1} x_i^2 \sum_{n-j-2\geq 0}(n-j-2+1)X^j X_{[1,i]}^{n-j-2} \\
&= 2\sum_{i=1}^{r-1} x_i^2 \sum_{j=0}^{n-2}(n-j-1)X^j X_{[1,i]}^{n-j-2} \\
&= 2\sum_{i=1}^{r-1} x_i^2 \sum_{j=0}^{n-2}(n-j-1)(X_{[1,i]}+X_{[i+1,r]})^j X_{[1,i]}^{n-j-2} \\
&= 2\sum_{i=1}^{r-1} x_i^2 \sum_{j=0}^{n-2}(n-j-1)\sum_{m=0}^{j}\binom{j}{m} X_{[1,i]}^m X_{[i+1,r]}^{j-m} X_{[1,i]}^{n-j-2} \\
&= 2\sum_{i=1}^{r-1} x_i^2 \sum_{j=0}^{n-2}(n-j-1)\sum_{m=0}^{j}\binom{j}{m} X_{[i+1,r]}^{j-m} X_{[1,i]}^{m+n-j-2} \\
&= 2\sum_{i=1}^{r-1} x_i^2 \sum_{j=0}^{n-2}(n-j-1)\sum_{m=0}^{j}\binom{j}{m} \sum_{k_{i+1}+\cdots+k_r=j-m}\binom{j-m}{k_{i+1},\ldots,k_r} x_{i+1}^{k_{i+1}}\cdots x_r^{k_r} \\
&\qquad \cdot \sum_{k_1+\cdots+k_i=m+n-j-2}\binom{m+n-j-2}{k_1,\ldots,k_i} x_1^{k_1}\cdots x_i^{k_i} \\
&= 2\sum_{i=1}^{r-1}\sum_{j=0}^{n-2}\sum_{m=0}^{j}\sum_{k_{i+1}+\cdots+k_r=j-m}\;\sum_{k_1+\cdots+k_i=m+n-j-2}(n-j-1) \\
&\qquad \cdot \binom{j}{m}\binom{j-m}{k_{i+1},\ldots,k_r}\binom{m+n-j-2}{k_1,\ldots,k_i} x_1^{k_1}\cdots x_{i-1}^{k_{i-1}} x_i^{k_i+2} x_{i+1}^{k_{i+1}}\cdots x_r^{k_r}.
\end{aligned}
$$

Thus for $k_s = n_s$, $\forall\, s \neq i$ and $k_i + 2 = n_i$,

$$
\begin{aligned}
&[z^n x_1^{n_1}\cdots x_r^{n_r}]A \\
&= 2\sum_{i=1}^{r-1}\sum_{j=0}^{n-2}(n-j-1)\binom{j}{k_{i+1}+\cdots+k_r}\binom{k_{i+1}+\cdots+k_r}{k_{i+1},\ldots,k_r}\binom{k_1+\cdots+k_i}{k_1,\ldots,k_i} \\
&= 2\sum_{i=1}^{r-1}\sum_{j=0}^{n-2}(n-(j+1))\binom{j}{N_{[i+1,r]}}\binom{N_{[i+1,r]}}{n_{i+1},\ldots,n_r}\binom{N_{[1,i]}-2}{n_1,\ldots,n_i-2}
\end{aligned}
$$

$$= 2\sum_{i=1}^{r-1}\binom{N_{[i+1,r]}}{n_{i+1},\ldots,n_r}\binom{N_{[1,i]}-2}{n_1,\ldots,n_i-2}\left[n\sum_{j=0}^{n-2}\binom{j}{N_{[i+1,r]}}-\sum_{j=0}^{n-2}(j+1)\binom{j}{N_{[i+1,r]}}\right].$$

From [12, page 174]:

- $\binom{r}{k} = \frac{r}{k}\binom{r-1}{k-1}$ for integer $k \neq 0$

- $\sum\limits_{k=0}^{n}\binom{k}{m} = \binom{n+1}{m+1}$ for integers $m, n \geq 0$

to simplify the following (note that $\binom{0}{d} = 0$ for $d \neq 0$):

$$\begin{aligned}
&n\sum_{j=0}^{n-2}\binom{j}{N_{[i+1,r]}}-\sum_{j=0}^{n-2}(j+1)\binom{j}{N_{[i+1,r]}}\\
&= n\binom{n-1}{N_{[i+1,r]}+1}-\sum_{j=0}^{n-2}(N_{[i+1,r]}+1)\binom{j+1}{N_{[i+1,r]}+1}\\
&= (N_{[i+1,r]}+2)\binom{n}{N_{[i+1,r]}+2}-(N_{[i+1,r]}+1)\sum_{j=0}^{n-2}\binom{j+1}{N_{[i+1,r]}+1}\\
&= (N_{[i+1,r]}+2)\binom{n}{N_{[i+1,r]}+2}-(N_{[i+1,r]}+1)\sum_{J=0}^{n-1}\binom{J}{N_{[i+1,r]}+1}\\
&= (N_{[i+1,r]}+2)\binom{n}{N_{[i+1,r]}+2}-(N_{[i+1,r]}+1)\binom{n}{N_{[i+1,r]}+2}\\
&= \binom{n}{N_{[i+1,r]}+2}\left[(N_{[i+1,r]}+2)-(N_{[i+1,r]}+1)\right]\\
&= \binom{n}{N_{[i+1,r]}+2}.
\end{aligned}$$

So we can substitute this back into the coefficient of $A$ to get

$$\begin{aligned}
&[z^n x_1^{n_1}\cdots x_r^{n_r}]A\\
&= 2\sum_{i=1}^{r-1}\binom{n}{N_{[i+1,r]}+2}\binom{N_{[1,i]}-2}{n_1,\ldots,n_i-2}\binom{N_{[i+1,r]}}{n_{i+1},\ldots,n_r}\\
&= 2\sum_{i=1}^{r-1}\frac{n!(N_{[1,i]}-2)!N_{[i+1,r]}!}{(N_{[i+1,r]}+2)!(n-(N_{[i+1,r]}+2))!n_1!\cdots n_{i-1}!(n_i-2)!n_{i+1}!\cdots n_r!}\\
&= 2\sum_{i=1}^{r-1}\frac{n!(N_{[1,i]}-2)!N_{[i+1,r]}!}{(N_{[i+1,r]}+2)!(n_1+\cdots+n_i-2)!n_1!\cdots n_{i-1}!(n_i-2)!n_{i+1}!\cdots n_r!}\\
&= 2\sum_{i=1}^{r-1}\frac{n_i(n_i-1)}{(N_{[i+1,r]}+1)(N_{[i+1,r]}+2)}\binom{n}{n_1,\ldots,n_r},
\end{aligned}$$

which gives the contribution of the A term to the second moment after dividing by $\binom{n}{n_1,\ldots,n_r}$. The contribution is thus

$$2\sum_{i=1}^{r-1}\frac{n_i(n_i-1)}{(N_{[i+1,r]}+1)(N_{[i+1,r]}+2)}. \tag{17.8}$$

On the other hand, we have (from (17.7))

$$\begin{aligned}
B &:= 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{z^2x_ix_j}{(1-zX)(1-zX_{[1,i]})(1-zX_{[1,j]})}\\
&= 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1} z^2x_ix_j\sum_{k\ge 0}z^kX^k\sum_{l\ge 0}z^lX_{[1,i]}^l\sum_{m\ge 0}z^mX_{[1,j]}^m,
\end{aligned}$$

and thus

$$\begin{aligned}
[z^n]B &= 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}x_ix_j\sum_{k=0}^{n-2}\sum_{n-l-k-2\ge 0}X^kX_{[1,i]}^lX_{[1,j]}^{n-l-k-2}\\
&= 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}x_ix_j\sum_{k=0}^{n-2}\sum_{l=0}^{n-k-2}(X_{[1,j]}+X_{[j+1,r]})^kX_{[1,i]}^lX_{[1,j]}^{n-l-k-2}\\
&= 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}x_ix_j\sum_{k=0}^{n-2}\sum_{l=0}^{n-k-2}\sum_{s=0}^{k}\binom{k}{s}X_{[1,j]}^sX_{[j+1,r]}^{k-s}X_{[1,i]}^lX_{[1,j]}^{n-l-k-2}\\
&= 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}x_ix_j\sum_{k=0}^{n-2}\sum_{l=0}^{n-k-2}\sum_{s=0}^{k}\binom{k}{s}X_{[j+1,r]}^{k-s}X_{[1,i]}^lX_{[1,j]}^{s+n-l-k-2}\\
&= 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}x_ix_j\sum_{k=0}^{n-2}\sum_{l=0}^{n-k-2}\sum_{s=0}^{k}\binom{k}{s}X_{[j+1,r]}^{k-s}X_{[1,i]}^l(X_{[1,i]}+X_{[i+1,j]})^{s+n-l-k-2}\\
&= 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}x_ix_j\sum_{k=0}^{n-2}\sum_{l=0}^{n-k-2}\sum_{s=0}^{k}\binom{k}{s}X_{[j+1,r]}^{k-s}X_{[1,i]}^l\\
&\quad\cdot\sum_{t=0}^{s+n-l-k-2}\binom{s+n-l-k-2}{t}X_{[1,i]}^tX_{[i+1,j]}^{s+n-l-k-2-t}\\
&= 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}x_ix_j\sum_{k=0}^{n-2}\sum_{l=0}^{n-k-2}\sum_{s=0}^{k}\sum_{t=0}^{s+n-l-k-2}\\
&\quad\cdot\binom{s+n-l-k-2}{t}\binom{k}{s}X_{[j+1,r]}^{k-s}X_{[1,i]}^{t+l}X_{[i+1,j]}^{s+n-l-k-2-t}\\
&= 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}x_ix_j\sum_{k=0}^{n-2}\sum_{l=0}^{n-k-2}\sum_{s=0}^{k}\sum_{t=0}^{s+n-l-k-2}\binom{s+n-l-k-2}{t}\binom{k}{s}\\
&\quad\cdot\sum_{b_{j+1}+\cdots+b_r=k-s}\binom{k-s}{b_{j+1},\ldots,b_r}x_{j+1}^{b_{j+1}}\cdots x_r^{b_r}\sum_{b_1+\cdots+b_i=t+l}\binom{t+l}{b_1,\ldots,b_i}x_1^{b_1}\cdots x_i^{b_i}\\
&\quad\cdot\sum_{b_{i+1}+\cdots+b_j=s+n-l-k-t-2}\binom{s+n-l-k-t-2}{b_{i+1},\ldots,b_j}x_{i+1}^{b_{i+1}}\cdots x_j^{b_j}
\end{aligned}$$

$$= 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\sum_{k=0}^{n-2}\sum_{l=0}^{n-k-2}\sum_{s=0}^{k}\sum_{t=0}^{s+n-l-k-2}\sum_{b_{j+1}+\cdots+b_r=k-s}\sum_{b_1+\cdots+b_i=t+l}\sum_{b_{i+1}+\cdots+b_j=s+n-l-k-t-2}$$
$$\cdot\binom{s+n-l-k-2}{t}\binom{k}{s}\binom{s+n-l-k-t-2}{b_{i+1},\ldots,b_j}\binom{t+l}{b_1,\ldots,b_i}\binom{k-s}{b_{j+1},\ldots,b_r}$$
$$\cdot\, x_{j+1}^{b_{j+1}}\cdots x_r^{b_r}x_1^{b_1}\cdots x_i^{b_i+1}x_{i+1}^{b_{i+1}}\cdots x_j^{b_j+1}.$$

Now we can write

$$[z^n x_1^{n_1}\cdots x_r^{n_r}]B = 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\overbrace{\sum_{k=0}^{n-2}\sum_{l=0}^{n-k-2}\binom{N_{[1,j]}-l-2}{N_{[i+1,j]}-1}}^{d} \tag{17.9}$$
$$\cdot\binom{k}{N_{[j+1,r]}}\binom{N_{[i+1,j]}-1}{n_{i+1},\ldots,n_j-1}\binom{N_{[1,i]}-1}{n_1,\ldots,n_i-1}\binom{N_{[j+1,r]}}{n_{j+1},\ldots,n_r},$$

from which we extract and simplify the quantity

$$\begin{aligned}
d &:= \sum_{k=0}^{n-2}\sum_{l=0}^{n-k-2}\binom{N_{[1,j]}-l-2}{N_{[i+1,j]}-1}\binom{k}{N_{[j+1,r]}}\\
&= \sum_{k=0}^{n-2}\binom{k}{N_{[j+1,r]}}\sum_{l=0}^{n-k-2}\binom{N_{[1,j]}-l-2}{N_{[i+1,j]}-1}\\
&= \sum_{k=0}^{n-2}\binom{k}{N_{[j+1,r]}}\sum_{l=0}^{n-k-2}\binom{N_{[1,j]}-2-(n-k-2)+l}{N_{[i+1,j]}-1}\\
&= \sum_{k=0}^{n-2}\binom{k}{N_{[j+1,r]}}\sum_{l=0}^{n-k-2}\binom{k+l-n_{j+1}-\cdots-n_r}{N_{[i+1,j]}-1}\\
&= \sum_{k=0}^{n-2}\binom{k}{N_{[j+1,r]}}\sum_{t=k-n_{j+1}-\cdots-n_r}^{N_{[1,j]}-2}\binom{t}{N_{[i+1,j]}-1}\\
&\qquad (\text{where } t := k+l-n_{j+1}-\cdots-n_r)\\
&= \sum_{k=0}^{n-2}\binom{k}{N_{[j+1,r]}}\left[\sum_{t=0}^{N_{[1,j]}-2}\binom{t}{N_{[i+1,j]}-1}-\sum_{t=0}^{k-n_{j+1}-\cdots-n_r-1}\binom{t}{N_{[i+1,j]}-1}\right]\\
&= \binom{n-1}{N_{[j+1,r]}+1}\binom{N_{[1,j]}-1}{N_{[i+1,j]}}-\binom{N_{[i+1,r]}}{N_{[j+1,r]}}\sum_{k=0}^{n-2}\binom{k}{N_{[i+1,r]}}\\
&= \binom{n-1}{N_{[j+1,r]}+1}\binom{N_{[1,j]}-1}{N_{[i+1,j]}}-\binom{N_{[i+1,r]}}{N_{[j+1,r]}}\binom{n-1}{N_{[i+1,r]}+1}.
\end{aligned}$$

Thus we have, from (17.9)

$$[z^n x_1^{n_1}\cdots x_r^{n_r}]B$$
$$= 2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\binom{N_{[i+1,j]}-1}{n_{i+1},\ldots,n_j-1}\binom{N_{[1,i]}-1}{n_1,\ldots,n_i-1}\binom{N_{[j+1,r]}}{n_{j+1},\ldots,n_r}$$

$$
\begin{aligned}
&\cdot\left[\binom{n-1}{N_{[j+1,r]}+1}\binom{N_{[1,j]}-1}{N_{[i+1,j]}}-\binom{N_{[i+1,r]}}{N_{[j+1,r]}}\binom{n-1}{N_{[i+1,r]}+1}\right]\\
&=2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{(N_{[i+1,j]}-1)!(N_{[1,i]}-1)!N_{[j+1,r]}!}{n_1!\cdots(n_i-1)!n_{i+1}!\cdots(n_j-1)!n_{j+1}!\cdots n_r!}\\
&\quad\cdot\frac{(n-1)!(N_{[1,j]}-1)!}{(N_{[j+1,r]}+1)!(N_{[1,j]}-2)!N_{[i+1,j]}!(N_{[1,i]}-1)!}\\
&\quad-2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{(N_{[i+1,j]}-1)!(N_{[1,i]}-1)!N_{[j+1,r]}!}{n_1!\cdots(n_i-1)!n_{i+1}!\cdots(n_j-1)!n_{j+1}!\cdots n_r!}\\
&\quad\cdot\frac{N_{[i+1,r]}!(n-1)!}{N_{[j+1,r]}!N_{[i+1,j]}!(N_{[i+1,r]}+1)!(N_{[1,i]}-2)!}\\
&=2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{n_in_j(N_{[1,j]}-1)}{n(N_{[j+1,r]}+1)N_{[i+1,j]}}\frac{n!}{n_1!\cdots n_i!n_{i+1}!\cdots n_j!n_{j+1}!\cdots n_r!}\\
&\quad-2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{n_in_j(N_{[1,i]}-1)}{nN_{[i+1,j]}(N_{[i+1,r]}+1)}\frac{n!}{n_1!\cdots n_i!n_{i+1}!\cdots n_j!n_{j+1}!\cdots n_r!}\\
&=2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{n_in_j}{(N_{[j+1,r]}+1)(N_{[i+1,r]}+1)}\binom{n}{n_1,\ldots,n_r}\\
&\quad\cdot\left[\frac{(N_{[1,j]}-1)(N_{[i+1,r]}+1)-(N_{[1,i]}-1)(N_{[j+1,r]}+1)}{nN_{[i+1,j]}}\right]\\
&=2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{n_in_j}{(N_{[j+1,r]}+1)(N_{[i+1,r]}+1)}\binom{n}{n_1,\ldots,n_r}.
\end{aligned}
$$

The B contribution to the second moment is then

$$
2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{n_in_j}{(N_{[i+1,r]}+1)(N_{[j+1,r]}+1)}. \tag{17.10}
$$

Together, A and B give the second moment as (see (17.8) and (17.10))

$$
2\sum_{i=1}^{r-1}\frac{n_i(n_i-1)}{(N_{[i+1,r]}+1)(N_{[i+1,r]}+2)}+2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{n_in_j}{(N_{[i+1,r]}+1)(N_{[j+1,r]}+1)}.
$$

For the variance, we also need to include two more terms (see (17.6)). We use equation (17.3) to write

$$
\begin{aligned}
\mathbb{V}_{\mathrm{rg}}^{\mathrm{m}}&=2\sum_{i=1}^{r-1}\frac{n_i(n_i-1)}{(N_{[i+1,r]}+1)(N_{[i+1,r]}+2)}+\underbrace{2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{n_in_j}{(N_{[i+1,r]}+1)(N_{[j+1,r]}+1)}}\\
&\quad+\sum_{i=1}^{r-1}\frac{n_i}{N_{[i+1,r]}+1}-\left(\sum_{i=1}^{r-1}\frac{n_i}{N_{[i+1,r]}+1}\right)^2.
\end{aligned}
$$

The square of the expectation includes the bracketed terms which then cancel leaving only each term squared, i.e.,

$$\mathbb{V}^{\mathrm{m}}_{\mathrm{rg}} = 2\sum_{i=1}^{r-1} \frac{n_i(n_i-1)}{(N_{[i+1,r]}+1)(N_{[i+1,r]}+2)} + \sum_{i=1}^{r-1} \frac{n_i}{N_{[i+1,r]}+1} - \sum_{i=1}^{r-1} \frac{n_i^2}{(N_{[i+1,r]}+1)^2}. \tag{17.11}$$

This is the exact formula. The asymptotic expression of this as $n \to \infty$ is simpler. We have

$$\sum_{i=1}^{r-1} \frac{n_i(n_i-1)}{(N_{[i+1,r]}+1)(N_{[i+1,r]}+2)} \sim \sum_{i=1}^{r-1} \frac{n_i^2}{(N_{[i+1,r]}+1)^2},$$

and

$$\frac{n_i}{N_{[i+1,r]}+1} \sim \frac{n_i}{N_{[i+1,r]}},$$

and thus

$$\mathbb{V}^{\mathrm{m}}_{\mathrm{rg}} \sim \sum_{i=1}^{r-1} \frac{n_i}{N_{[i+1,r]}} + \sum_{i=1}^{r-1} \frac{n_i^2}{(N_{[i+1,r]})^2},$$

as $n \to \infty$. This completes the proof of Theorem 17.3. ■

## 17.5 Right-going variance – probability model

To find the variance using the probability model, we use (see equation (17.4))

$$\begin{aligned} f(z,u,x_1,\dots,x_r) &:= \prod_{i=1}^{r-1} \left(1 + \frac{zux_i}{1 - z(X_{[1,i-1]} + ux_i)}\right)\left(1 + \frac{zx_r}{1 - zx_{[1,r]}}\right) \\ &= \sum_{w \in A^*} z^{|w|} u^{r(w)} x_1^{|w|_1} x_2^{|w|_2} \cdots x_r^{|w|_r}, \end{aligned}$$

and

$$\begin{aligned} \mathbb{V}^{\mathrm{p}}_{\mathrm{rg}} &:= [z^n]\frac{\partial^2}{\partial^2 u} f(z,u,p_1,\dots,p_r)\Big|_{u=1} + [z^n]\frac{\partial}{\partial u} f(z,u,p_1,\dots,p_r)\Big|_{u=1} \\ &\quad - ([z^n]\frac{\partial}{\partial u} f(z,u,p_1,\dots,p_r)\Big|_{u=1})^2. \end{aligned} \tag{17.12}$$

We only lack the second moment (i.e., the first term of (17.12)), and for this we use the calculations from the multiset model. We have (see (17.7))

$$\begin{aligned} &[z^n]\frac{\partial^2}{\partial^2 u} f(z,u,p_1,\dots,p_r)\Big|_{u=1} \\ &= \underbrace{[z^n]\sum_{i=1}^{r-1} \frac{2z^2 p_i^2}{(1 - zP_{[1,r]})(1 - zP_{[1,i]})^2}}_{\dagger} + \underbrace{[z^n]\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1} \frac{2z^2 p_i p_j}{(1 - zP_{[1,r]})(1 - zP_{[1,i]})(1 - zP_{[1,j]})}}_{\ddagger}, \end{aligned}$$

and we know that by expanding this expression in terms of partial fractions we can find the coefficient of $z^n$. This method is quite long so we start by making a few remarks which will shorten the calculations.

- We can write

  $$\begin{aligned}&[z^n]\frac{z^2}{(1-az)(1-bz)(1-cz)}\\&=[z^n]\left(\frac{1}{(1-az)(a-b)(a-c)}+\frac{1}{(1-bz)(b-a)(b-c)}+\frac{1}{(1-cz)(c-b)(c-a)}\right)\\&=\frac{1}{(a-b)(a-c)}a^n+\frac{1}{(b-a)(b-c)}b^n+\frac{1}{(c-b)(c-a)}c^n.\end{aligned}$$

  In our case we have this situation where $a=P_{[1,r]}=1$, $b=P_{[1,i]}<1$ and $c=P_{[1,j]}<1$ (for $i,j<r$). As $n\to\infty$, we have $a^n\to 1$, $b^n\to 0$ and $c^n\to 0$. Thus it is unnecessary to do all the calculations – we need only include the terms of the partial fraction expansion which have the factor $\frac{1}{1-zP_{[1,r]}}$ (or $P^n_{[1,r]}$). This takes care of the double sum (from ‡) in the second moment above.

- For the terms which have a repeated factor (see †) in the denominator this idea remains the same. We have

  $$\begin{aligned}&[z^n]\frac{z^2}{(1-az)(1-bz)^2}\\&=[z^n]\left(\frac{1}{(1-az)(a-b)^2}+\frac{a-2b}{(1-bz)b(a-b)^2}+\frac{1}{(1-bz)^2b(b-a)}\right)\\&=\frac{1}{(a-b)^2}a^n+\frac{a-2b}{b(a-b)^2}b^n+\frac{1}{b(b-a)}(n+1)b^n.\end{aligned}$$

  Again we have that $a=P_{[1,r]}=1$ and $b=P_{[1,i]}<1$ since $i<r$, and $p_i>0$, $\forall\, i=1,\ldots,r$. Thus since $a^n\to 1$, $b^n\to 0$ and $nb^n\to 0$ as $n\to\infty$, we need only consider the terms of the form $\frac{1}{(1-az)(a-b)^2}$.

We can now give an asymptotic approximation of the second moment (as $n\to\infty$):

$$[z^n]\frac{\partial^2}{\partial^2 u}f(z,u,p_1,\ldots,p_r)\Big|_{u=1}\sim\sum_{i=1}^{r-1}\frac{2p_i^2}{P^2_{[i+1,r]}}+\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{2p_ip_j}{P_{[i+1,r]}P_{[j+1,r]}}.$$

For the variance we must also include the other two terms in (17.12), so

$$\begin{aligned}\mathbb{V}^{\mathrm{p}}_{\mathrm{rg}}&\sim 2\sum_{i=1}^{r-1}\frac{p_i^2}{P^2_{[i+1,r]}}+2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{p_ip_j}{P_{[i+1,r]}P_{[j+1,r]}}+\sum_{i=1}^{r-1}\frac{p_i}{P_{[i+1,r]}}-\left(\sum_{i=1}^{r-1}\frac{p_i}{P_{[i+1,r]}}\right)^2\\&=2\sum_{i=1}^{r-1}\frac{p_i^2}{P^2_{[i+1,r]}}+2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{p_ip_j}{P_{[i+1,r]}P_{[j+1,r]}}\end{aligned}$$

$$+\sum_{i=1}^{r-1}\frac{p_i}{P_{[i+1,r]}}-\sum_{i=1}^{r-1}\frac{p_i^2}{P_{[i+1,r]}^2}-2\sum_{i=1}^{r-1}\sum_{j=i+1}^{r-1}\frac{p_ip_j}{P_{[i+1,r]}P_{[j+1,r]}},$$

and thus

$$\mathbb{V}_{\mathrm{rg}}^{\mathrm{p}}\sim\sum_{i=1}^{r-1}\frac{p_i}{P_{[i+1,r]}}+\sum_{i=1}^{r-1}\frac{p_i^2}{P_{[i+1,r]}^2}, \tag{17.13}$$

as $n\to\infty$. The proof of Theorem 17.4 is thus complete. ∎

Once again, a link can be shown between the two results 17.11 and 17.13 as $n\to\infty$, replacing $n_i$ by $np_i$.

# Chapter 18

# Expected depth of an arbitrary node $\alpha$

## 18.1 Introduction

In this section we investigate the cost of searching for a key $\alpha$ in the binary search tree formed from a word $(w)$ of length $n$ made up of letters (possibly repeated) from the alphabet $\{1, \ldots, r\}$. The cost can also be thought of as the number of comparisons or length of path from the root to the node $\alpha$, as in the previous cases. In this case, allowing each element to appear only once in the binary search tree means that we consider strict left-to-right maxima (and minima, though this is no different to Chapter 16). Note that previously in the right-going section we included all nodes in the tree and thus considered weak left-to-right maxima.

Why is this different to the distinct key case? If only the distinct keys are allowed into the tree, the binary search tree will always only have $r$ nodes. However, since it was formed from a multiset of $\{1, \ldots, r\}$, each tree will appear with a different probability than if it originated from a sequence with distinct keys. Consider the set $\{1, 2, 3\}$ and the multiset $\{1 \cdot 1\ ;\ 1 \cdot 2\ ;\ 2 \cdot 3\}$. For the former, there are $3! = 6$ possible sequences, giving rise to 5 different binary search trees, see Figure 18.1 below.

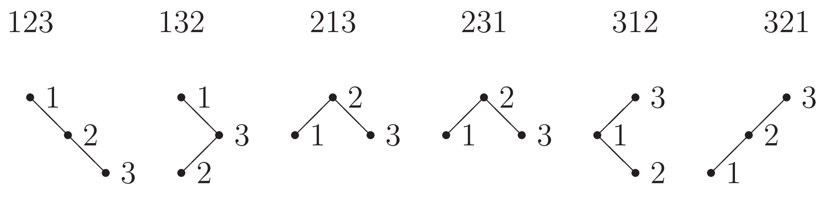


Figure 18.1: The binary search trees from the set $\{1, 2, 3\}$.

Of the five trees in Figure 18.1, all have the probability $\frac{1}{6}$, except the tree corresponding to the sequences 213 and 231, which occurs with a probability of $\frac{1}{3}$. However, if we look at the $\binom{4}{1,1,2} = 12$ sequences from the set $\{1,2,3,3\}$, we get:

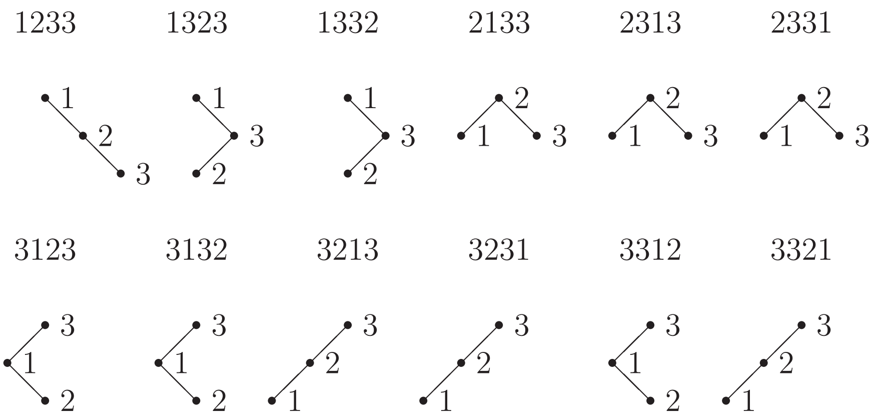


Figure 18.2: The binary search trees from the set $\{1,2,3,3\}$.

Figure 18.2 shows that even though we get the same five trees resulting, their probabilities have changed. In this case all trees occur with probability $\frac{1}{4}$, except for the first (corresponding to sequence 1233) and the second (corresponding to sequences 1323 and 1332) which occur with probabilities $\frac{1}{12}$ and $\frac{1}{6}$ respectively. Thus searching for the average depth of a certain key will be different, even if the number and shape of the trees are the same. For example, the average depths of 1,2 and 3 in that order are $\frac{5}{6}$, 1 and $\frac{5}{6}$ in Figure 18.1 and 1, $\frac{7}{6}$ and $\frac{7}{12}$ in Figure 18.2. (The average depth of 3, the repeated letter has decreased significantly.)

Again, generating functions are used to express the situation, and similar techniques to Chapters 16 and 17 are used. The alphabet is divided into two distinct sub-alphabets at the value $\alpha$ – those letters larger than $\alpha$ and those letters smaller than $\alpha$. Thus we use the fact that we already know how to find the left-going depth (left-to-right minima) of the smallest element in a sub-alphabet, and the right-going depth (left-to-right maxima) of the largest element in a sub-alphabet.

We use the notation $r = \text{card}(A)$, i.e., $A = \{1, \ldots, r\}$. So if $w \in A^*$ then the binary search tree of $w$ has every symbol of $w$ inserted once only. (This differs from the general quicksort analyses which only deal with binary search trees formed from permutations of sequences of distinct letters.) The question we address is: "what are the expected value and variance of the cost of finding node $\alpha$", using each of these models?

To answer this question we use the left-going and right-going depth, and a tool called the 'shuffle' product, which is simply the product of two exponential generating functions (EGFs), but which produces a 'shuffle' between two words whose alphabets are distinct. So by applying the shuffle product to two words we end up with all possible combinations of the original words with the letters interwoven, but with the original order within the original two words unchanged. For example, take the two words $ab$ and $cd$. If we shuffle these like cards, we get $\binom{4}{2} = 6$ solutions: $\{abcd, acbd, acdb, cabd, cadb, cdab\}$. The definition of the shuffle product is

$$au \sqcup\!\sqcup\, bv := a(u \sqcup\!\sqcup\, bv) + b(au \sqcup\!\sqcup\, v).$$

The product of exponential generating functions produces this shuffle product. However, the original generating functions for calculating the left-going and right-going depths are ordinary generating functions (OGFs), so at some point we need to change from ordinary generating functions to exponential generating functions, perform the shuffle, and then convert the result back again. To do this we use the combinatorial Laplace transform. Ultimately we want to apply the shuffle product to two languages, which we assume to be distinct. This is why an exponential generating function product is the answer, because each exponential generating function represents a language, and we multiply the two to get all possible combinations of words from these languages. The output will thus be a language of all words which were formed from a shuffle of two words, one from each of the original languages.

The (combinatorial) Laplace transform states the following (see [36, page 92]): Given an EGF $\hat{A}(x)$ for a sequence $\{a_k\}$, the ordinary generating function for the sequence is given by

$$\int_0^\infty \hat{A}(zt)e^{-t}dt$$

if the integral exists. To show this we consider the exponential generating function $\hat{A}(z) = \sum\limits_{n\geq 0} a_n \frac{z^n}{n!}$, and then show that

$$\sum_{n\geq 0} a_n z^n = \int_0^\infty \hat{A}(zt)e^{-t}dt.$$

Firstly,

$$\int_0^\infty \hat{A}(zt)e^{-t}dt = \int_0^\infty \sum_{n\geq 0} a_n \frac{z^n t^n}{n!} e^{-t}dt = \sum_{n\geq 0} a_n \frac{z^n}{n!} \int_0^\infty t^n e^{-t}dt$$

and by Euler's integral which defines the Gamma function as $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$, we have that $\int_0^\infty t^n e^{-t}dt = n!$ and we are done. To change from an ordinary generating function to an exponential generating function one uses the inverse Laplace transform. This is done using Hankel's contour integral ([1, page 255]), which says (note that we use '$\boldsymbol{i}$' for the complex number instead of '$i$' to differentiate from the index we use later):

$$\frac{1}{\Gamma(z)} = \frac{\boldsymbol{i}}{2\pi} \int_C (-t)^{-z} e^{-t} dt, \qquad (|z| < \infty),$$

i.e.,

$$\frac{1}{n!} = \frac{1}{\Gamma(n+1)} = \frac{\boldsymbol{i}}{2\pi} \int_C (-t)^{-n-1} e^{-t} dt. \tag{18.1}$$

So whereas when we transform from an exponential generating function to an ordinary generating function we multiply by $n!$, to transform from an ordinary generating function to an exponential generating function, we must divide by $n!$, or multiply by the Hankel contour integral in (18.1). So if $A(z) = \sum_{n\geq 0} a_n z^n$, then our corresponding exponential generating function (with the same coefficients) is

$$\begin{aligned}
\sum_{n\geq 0} a_n z^n \frac{\boldsymbol{i}}{2\pi} \int_C (-t)^{-n-1} e^{-t} dt &= \frac{\boldsymbol{i}}{2\pi} \int_C \sum_{n\geq 0} a_n z^n (-t)^{-n} (-t)^{-1} e^{-t} dt \\
&= \frac{-\boldsymbol{i}}{2\pi} \int_C \sum_{n\geq 0} a_n \Big(\frac{z}{-t}\Big)^n \frac{1}{t} e^{-t} dt \\
&= \frac{1}{2\pi \boldsymbol{i}} \int_C A\Big(-\frac{z}{t}\Big) \frac{1}{te^t} dt \\
&= \frac{1}{2\pi \boldsymbol{i}} \int_C \sum_{n\geq 0} a_n \Big(-\frac{z}{t}\Big)^n \frac{1}{te^t} dt \\
&= \sum_{n\geq 0} a_n z^n \frac{1}{2\pi \boldsymbol{i}} \int_C (-t)^{-n} \frac{1}{te^t} dt \\
&= \sum_{n\geq 0} a_n z^n \frac{\boldsymbol{i}}{2\pi} \int_C (-t)^{-n-1} e^{-t} dt \\
&= \sum_{n\geq 0} a_n \frac{z^n}{n!},
\end{aligned}$$

from equation (18.1).

As an example consider the ordinary generating function $A(z) := \sum_{n\geq 0} z^n = \frac{1}{1-z}$ which has coefficients of 1 everywhere. The exponential generating function with coefficients of 1 which we expect is $e^z$. We have that the exponential generating function is

$$\begin{aligned}
\frac{1}{2\pi \boldsymbol{i}} \int_C A\Big(-\frac{z}{t}\Big) \frac{1}{te^t} dt &= \frac{1}{2\pi \boldsymbol{i}} \int_C \frac{1}{1 - (-\frac{z}{t})} \frac{1}{te^t} dt \\
&= \frac{1}{2\pi \boldsymbol{i}} \int_C \frac{1}{t\big(1 - (-\frac{z}{t})\big)} \frac{1}{e^t} dt
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2\pi \boldsymbol{i}} \int\limits_C \frac{f(t)}{t-(-z)} dt, \qquad \text{where } f(t) = e^{-t} \\
&= f(-z) \\
&= e^z,
\end{aligned}
$$

as required.

Using the above, the theorems below are proved in the sections to follow. As in the previous two chapters, the results correspond asymptotically as $n \to \infty$.

**Theorem 18.1** *The expected depth of some $\alpha \in \{1, \ldots, r\}$ is given by the exact formula*

$$\mathbb{E}^{\mathrm{m}}_{\alpha} = \sum_{i=1}^{\alpha-1} \frac{n_i}{N_{[i,\alpha]}} + \sum_{i=\alpha+1}^{r} \frac{n_i}{N_{[\alpha,i]}}.$$

**Theorem 18.2** *Alternatively, the expectation of the depth of $\alpha$ can be expressed asymptotically in terms of probabilities:*

$$\mathbb{E}^{\mathrm{p}}_{\alpha} \sim \sum_{i=1}^{\alpha-1} \frac{p_i}{P_{[i,\alpha]}} + \sum_{i=\alpha+1}^{r} \frac{p_i}{P_{[\alpha,i]}}$$

*as $n \to \infty$.*

**Theorem 18.3** *The variance of the depth of some $\alpha \in \{1, \ldots, r\}$ can be expressed (by the multiset model) as*

$$
\begin{aligned}
\mathbb{V}^{\mathrm{m}}_{\alpha} = 2\sum_{i=1}^{\alpha-1}\sum_{k=1}^{i-1} &\left( \frac{n_i n_k}{N_{[i,\alpha-1]}N_{[k,\alpha-1]}} - \frac{n_\alpha n_i n_k}{N_{[k,i-1]}N_{[i,\alpha-1]}N_{[i,\alpha]}} - \frac{n_\alpha n_i n_k}{N_{[k,i-1]}N_{[k,\alpha-1]}N_{[k,\alpha]}} \right) \\
+ 2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1} &\left( \frac{n_i n_k}{N_{[\alpha+1,i]}N_{[\alpha+1,k]}} + \frac{n_\alpha n_i n_k}{N_{[k+1,i]}N_{[\alpha+1,i]}N_{[\alpha,i]}} - \frac{n_\alpha n_i n_k}{N_{[k+1,i]}N_{[\alpha+1,k]}N_{[\alpha,k]}} \right) \\
+ 2\sum_{i=1}^{\alpha-1}\sum_{j=\alpha+1}^{r} &\left( \frac{n_i n_j}{N_{[i,\alpha-1]}N_{[\alpha+1,j]}} - \frac{n_\alpha n_i n_j}{N_{[i,\alpha-1]}N_{[\alpha+1,j]}N_{[\alpha,j]}} \right. \\
&\left. \quad - \frac{n_\alpha n_i n_j}{N_{[i,\alpha-1]}N_{[\alpha+1,j]}N_{[i,\alpha]}} - \frac{n_\alpha n_i n_j}{N_{[i,\alpha-1]}N_{[\alpha+1,j]}N_{[i,j]}} \right) \\
+ \sum_{i=1}^{\alpha-1} \frac{n_i}{N_{[i,\alpha]}} &+ \sum_{i=\alpha+1}^{r} \frac{n_i}{N_{[\alpha,i]}} \\
- \left( \sum_{i=1}^{\alpha-1} \frac{n_i}{N_{[i,\alpha]}} \right)^2 &- 2\sum_{i=1}^{\alpha-1} \frac{n_i}{N_{[i,\alpha]}} \sum_{i=\alpha+1}^{r} \frac{n_i}{N_{[\alpha,i]}} - \left( \sum_{i=\alpha+1}^{r} \frac{n_i}{N_{[\alpha,i]}} \right)^2.
\end{aligned}
$$

**Theorem 18.4** *Using the probability model, the variance of the depth of $\alpha$ as $n \to \infty$ is*

$$\mathbb{V}^{\mathrm{p}}_{\alpha} \sim 2\sum_{i=1}^{\alpha-1}\sum_{k=1}^{i-1}\left(\frac{p_i p_k}{P_{[i,\alpha-1]}P_{[k,\alpha-1]}} - \frac{p_\alpha p_i p_k}{P_{[k,i-1]}P_{[i,\alpha-1]}P_{[i,\alpha]}} + \frac{p_\alpha p_i p_k}{P_{[k,i-1]}P_{[k,\alpha-1]}P_{[k,\alpha]}}\right)$$
$$+ 2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\left(\frac{p_i p_k}{P_{[\alpha+1,i]}P_{[\alpha+1,k]}} + \frac{p_\alpha p_i p_k}{P_{[k+1,i]}P_{[\alpha+1,i]}P_{[\alpha,i]}} - \frac{p_\alpha p_i p_k}{P_{[k+1,i]}P_{[\alpha+1,k]}P_{[\alpha,k]}}\right)$$
$$+ 2\sum_{i=1}^{\alpha-1}\sum_{j=\alpha+1}^{r}\left(\frac{p_i p_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}} - \frac{p_\alpha p_i p_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[\alpha,j]}}\right.$$
$$\left. - \frac{p_\alpha p_i p_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[i,\alpha]}} + \frac{p_\alpha p_i p_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[i,j]}}\right)$$
$$+ \sum_{i=1}^{\alpha-1}\frac{p_i}{P_{[i,\alpha]}} + \sum_{i=\alpha+1}^{r}\frac{p_i}{P_{[\alpha,i]}}$$
$$- \left(\sum_{i=1}^{\alpha-1}\frac{p_i}{P_{[i,\alpha]}}\right)^2 - 2\sum_{i=1}^{\alpha-1}\frac{p_i}{P_{[i,\alpha]}}\sum_{i=\alpha+1}^{r}\frac{p_i}{P_{[\alpha,i]}} - \left(\sum_{i=\alpha+1}^{r}\frac{p_i}{P_{[\alpha,i]}}\right)^2.$$

## 18.2 Expectation – multiset model

We start by considering the cost function of finding some element $\alpha$ in our binary search tree. We define the generating function as:

$$f(z,u,x_1,\ldots,x_r) := \left[N_{\max}(z,u,X_{[1,\alpha-1]}) ш N_{\min}(z,u,X_{[\alpha+1,r]})\right] z x_\alpha \cdot \frac{1}{1-zX},$$

where $X = x_1 + x_2 + \cdots + x_r$ and ш indicates the shuffle product. The shuffle product is the usual $\cdot$ for exponential generating functions if the alphabets are distinct. This shuffle product takes place between ordinary generating functions $N_{\max}$ (which counts the number of left-to-right maxima in the letters smaller than $\alpha$ to the left of the first $\alpha$) and $N_{\min}$ (which counts the number of left-to-right minima the letters larger than $\alpha$ to the left of the first $\alpha$). The factor $zx_\alpha$ represents the first occurrence of $\alpha$, and the remaining factor of $\frac{1}{1-zX}$ represents everything to the right of the first $\alpha$ which can be of any length and which consists of any letters from 1 to $r$ (with repeats). The variables are as follows: $z$ counts all letters, $u$ counts all left-to-right maxima (resp. minima), and $x_1$ counts all ones, $x_2$ counts all twos etc.. We also use the shorthand $X_{[3,5]} = x_3 + x_4 + x_5$, $X_{[4,4]} = x_4$ and so on. Notice that $u$ does not appear after the first $\alpha$ since the depth of the first $\alpha$ is only dependent on the letters which occur to the left of $\alpha$ in the word.

Now we need to define the ordinary generating functions $N_{\max}$ and $N_{\min}$. We want $N_{\max}(z,u,X_{[1,\alpha-1]})$ to count the number of left-to-right maxima in all the letters strictly smaller than $\alpha$ which appear to the left of the left-most $\alpha$ in the word. If we let $\varepsilon$ represent an empty word, we can express these letters symbolically as

$$\big(\varepsilon + 1\{1\}^*\big)\big(\varepsilon + 2\{1,2\}^*\big)\big(\varepsilon + 3\{1,2,3\}^*\big)\cdots\big(\varepsilon + (\alpha-1)\{1,\ldots,\alpha-1\}^*\big),$$

which can be translated into the generating function

$$N_{\max}(z,u,X_{[1,\alpha-1]}) := \prod_{i=1}^{\alpha-1}\left(1 + \frac{zux_i}{1-zX_{[1,i]}}\right).$$

This is a product of factors which correspond to the bracketed factors in the symbolic equation above. Each factor corresponds to a new letter (one larger than the previous one), with the possibility of a new left-to-right maximum occurring if such a letter occurs in such a position. If it is a left-to-right maximum, then both $z$ and $u$ will count the $x_i$, and this is followed by a sequence, $\left(\frac{1}{1-zX_{[1,i]}}\right)$, which can have any element less than or equal to that element. If it is not a left-to-right maximum, then the only contribution from that factor is the 1 (or symbolically the $\varepsilon$). Note that this letter can still occur in the word, either as part of the $\frac{1}{1-zX_{[1,i+1]}}$ factor in the subsequent bracket, or as part of the $\frac{1}{1-zX}$ factor after the first $\alpha$). This only means that it appears to the right of a larger value and thus is not a left-to-right maximum.

For the other case we want to count the number of left-to-right minima among those letters which appear to the left of the left-most $\alpha$ which are strictly greater than $\alpha$. We use a similar argument to translate the symbols

$$\big(\varepsilon + r\{r\}^*\big)\big(\varepsilon + (r-1)\{r,r-1\}^*\big)\cdots\big(\varepsilon + (\alpha+1)\{r,r-1,\ldots,\alpha+1\}^*\big),$$

into the generating function

$$N_{\min}(z,u,X_{[\alpha+1,r]}) := \prod_{i=\alpha+1}^{r}\left(1 + \frac{zux_i}{1-zX_{[i,r]}}\right).$$

Now that these two functions have been defined we can rewrite the cost function more explicitly:

$$f(z,u,x_1,\ldots,x_r) = \left[\prod_{i=1}^{\alpha-1}\left(1+\frac{zux_i}{1-zX_{[1,i]}}\right) ш \prod_{i=\alpha+1}^{r}\left(1+\frac{zux_i}{1-zX_{[i,r]}}\right)\right]\frac{zx_\alpha}{1-zX}. \tag{18.2}$$

Our next task is to find the expected value, which for probability generating functions such as these means we will need to find the first order partial derivative with respect to $u$. First we note that

$$\frac{\partial}{\partial u}(f ш g)\Big|_{u=1} = \left(\frac{\partial}{\partial u}f\Big|_{u=1}\right) ш g\big|_{u=1} + f\big|_{u=1} ш \left(\frac{\partial}{\partial u}g\Big|_{u=1}\right). \tag{18.3}$$

So we let $f := N_{\min}$ and $g := N_{\max}$ in (18.3), and deal with the shuffle product by making use of the inverse Laplace transform and the Laplace transform in that order. Each of these four expressions is now written in a form which can easily be transformed into a corresponding exponential generating function so that the shuffle product can be performed. For the first factor of interest we have:

$$
\begin{aligned}
&\frac{\partial}{\partial u} N_{\min}\Big|_{u=1} \\
&\quad = \frac{\partial}{\partial u} \log N_{\min}\Big|_{u=1} \cdot N_{\min}\Big|_{u=1} \\
&\quad = \frac{\partial}{\partial u} \log \prod_{i=\alpha+1}^{r} \left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\Bigg|_{u=1} \cdot \prod_{i=\alpha+1}^{r} \left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\Bigg|_{u=1} \\
&\quad = \sum_{i=\alpha+1}^{r} \frac{\partial}{\partial u} \log \left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\Bigg|_{u=1} \cdot \prod_{i=\alpha+1}^{r} \left(1 + \frac{zx_i}{1 - zX_{[i,r]}}\right) \\
&\quad = \sum_{i=\alpha+1}^{r} \frac{\partial}{\partial u} \log \left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\Bigg|_{u=1} \cdot \prod_{i=\alpha+1}^{r} \left(\frac{1 - zX_{[i,r]} + zx_i}{1 - zX_{[i,r]}}\right) \\
&\quad = \sum_{i=\alpha+1}^{r} \frac{\partial}{\partial u} \log \left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\Bigg|_{u=1} \cdot \prod_{i=\alpha+1}^{r} \left(\frac{1 - zX_{[i+1,r]}}{1 - zX_{[i,r]}}\right) \\
&\quad = \sum_{i=\alpha+1}^{r} \frac{\partial}{\partial u} \log \left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\Bigg|_{u=1} \\
&\qquad \cdot \left(\frac{1 - zX_{[\alpha+2,r]}}{1 - zX_{[\alpha+1,r]}}\right)\left(\frac{1 - zX_{[\alpha+3,r]}}{1 - zX_{[\alpha+2,r]}}\right) \cdots \left(\frac{1 - zX_{[r,r]}}{1 - zX_{[r-1,r]}}\right)\left(\frac{1}{1 - zX_{[r,r]}}\right) \\
&\quad = \sum_{i=\alpha+1}^{r} \frac{\partial}{\partial u} \log \left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\Bigg|_{u=1} \cdot \left(\frac{1}{1 - zX_{[\alpha+1,r]}}\right) \\
&\quad = \sum_{i=\alpha+1}^{r} \frac{\partial}{\partial u} \log \left(\frac{1 - zX_{[i,r]} + zux_i}{1 - zX_{[i,r]}}\right)\Bigg|_{u=1} \cdot \left(\frac{1}{1 - zX_{[\alpha+1,r]}}\right) \\
&\quad = \sum_{i=\alpha+1}^{r} \frac{1}{\frac{1 - zX_{[i,r]} + zux_i}{1 - zX_{[i,r]}}} \cdot \frac{zx_i}{1 - zX_{[i,r]}}\Bigg|_{u=1} \cdot \left(\frac{1}{1 - zX_{[\alpha+1,r]}}\right) \\
&\quad = \frac{1}{1 - zX_{[\alpha+1,r]}} \sum_{i=\alpha+1}^{r} \frac{1 - zX_{[i,r]}}{1 - zX_{[i+1,r]}} \cdot \frac{zx_i}{1 - zX_{[i,r]}} \\
&\quad = \frac{1}{1 - zX_{[\alpha+1,r]}} \sum_{i=\alpha+1}^{r} \frac{zx_i}{1 - zX_{[i+1,r]}}.
\end{aligned}
$$

We make use of partial fraction decompositions (this can be done by hand or on Maple) to get

$$
\frac{\partial}{\partial u} N_{\min}\Big|_{u=1} = \sum_{i=\alpha+1}^{r} \frac{x_i}{X_{[\alpha+1,i]}} \Bigg( \underbrace{\frac{1}{1 - zX_{[\alpha+1,r]}} - \frac{1}{1 - zX_{[i+1,r]}}}_{\heartsuit} \Bigg).
$$

Now $\heartsuit$ is the difference of two ordinary generating functions of the form $\frac{1}{1-bz}$ which can be transformed into corresponding exponential generating functions (with the same coefficients) of the form $e^{bz}$. This is a useful form to have it in because we now want to find the product of these exponential generating functions (assuming the alphabets are distinct, which they are in this case, being either all smaller than $\alpha$ for $N_{\max}$, or all larger than $\alpha$ for $N_{\min}$). It is also then easy to transform back into an ordinary generating function of the form $\frac{1}{1-cz}$. So if we transform the ordinary generating function $\frac{\partial}{\partial u} N_{\min}\big|_{u=1}$ into an exponential generating function with the same coefficients, and let 'E' denote the exponential generating function and 'd' the differentiation, we get:

$$N_{\min}^{\mathrm{Ed}} := \sum_{i=\alpha+1}^{r} \frac{x_i}{X_{[\alpha+1,i]}} \big( e^{zX_{[\alpha+1,r]}} - e^{zX_{[i+1,r]}} \big). \tag{18.4}$$

The next factor of interest is the ordinary generating function

$$\begin{aligned}
N_{\max}\big|_{u=1} &= \prod_{i=1}^{\alpha-1} \left( 1 + \frac{zux_i}{1-zX_{[1,i]}} \right)\bigg|_{u=1} \\
&= \prod_{i=1}^{\alpha-1} \left( 1 + \frac{zx_i}{1-zX_{[1,i]}} \right) \\
&= \prod_{i=1}^{\alpha-1} \left( \frac{1-zX_{[1,i-1]}}{1-zX_{[1,i]}} \right) \\
&= \left( \frac{1}{1-zX_{[1,1]}} \right) \left( \frac{1-zX_{[1,1]}}{1-zX_{[1,2]}} \right) \left( \frac{1-zX_{[1,2]}}{1-zX_{[1,3]}} \right) \\
&\qquad \cdots \left( \frac{1-zX_{[1,\alpha-3]}}{1-zX_{[1,\alpha-2]}} \right) \left( \frac{1-zX_{[1,\alpha-2]}}{1-zX_{[1,\alpha-1]}} \right) \\
&= \frac{1}{1-zX_{[1,\alpha-1]}},
\end{aligned}$$

which transforms to the exponential generating function

$$N_{\max}^{\mathrm{E}} := e^{zX_{[1,\alpha-1]}}. \tag{18.5}$$

So the product of (18.4) and (18.5) gives us the first term in (18.3), namely the exponential generating function of $\left( \frac{\partial}{\partial u} N_{\min}\Big|_{u=1} \right) ш \, N_{\max}\big|_{u=1}$, which is (note this is the usual product and is thus commutative):

$$\begin{aligned}
&e^{zX_{[1,\alpha-1]}} \cdot \sum_{i=\alpha+1}^{r} \frac{x_i}{X_{[\alpha+1,i]}} \big( e^{zX_{[\alpha+1,r]}} - e^{zX_{[i+1,r]}} \big) \\
&= \sum_{i=\alpha+1}^{r} \frac{x_i}{X_{[\alpha+1,i]}} e^{zX_{[1,\alpha-1]}} \big( e^{zX_{[\alpha+1,r]}} - e^{zX_{[i+1,r]}} \big) \\
&= \sum_{i=\alpha+1}^{r} \frac{x_i}{X_{[\alpha+1,i]}} \big( e^{zX_{[1,\alpha-1]}} e^{zX_{[\alpha+1,r]}} - e^{zX_{[1,\alpha-1]}} e^{zX_{[i+1,r]}} \big)
\end{aligned}$$

$$= \sum_{i=\alpha+1}^{r} \frac{x_i}{X_{[\alpha+1,i]}} \left( e^{z\left(X_{[1,\alpha-1]}+X_{[\alpha+1,r]}\right)} - e^{z\left(X_{[1,\alpha-1]}+X_{[i+1,r]}\right)} \right).$$

By doing this we have performed the shuffle and all that remains is to transform back into an ordinary generating function. So

$$\begin{aligned} &\left(\frac{\partial}{\partial u} N_{\min}\Big|_{u=1}\right) ш\, N_{\max}\Big|_{u=1} \\ &= \sum_{i=\alpha+1}^{r} \frac{x_i}{X_{[\alpha+1,i]}} \left( \frac{1}{1 - z\left(X_{[1,\alpha-1]} + X_{[\alpha+1,r]}\right)} - \frac{1}{1 - z\left(X_{[1,\alpha-1]} + X_{[i+1,r]}\right)} \right), \end{aligned} \tag{18.6}$$

which we will leave in this form because later we will need to find the coefficient of $z^n$. We now turn our attention to the second term in (18.3), which is $N_{\min}\big|_{u=1} ш \left(\frac{\partial}{\partial u} N_{\max}\Big|_{u=1}\right)$ and which we will deal with in a very similar way. Firstly,

$$\begin{aligned} N_{\min}\Big|_{u=1} &= \prod_{i=\alpha+1}^{r} \left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\Bigg|_{u=1} \\ &= \prod_{i=\alpha+1}^{r} \left(1 + \frac{zx_i}{1 - zX_{[i,r]}}\right) \\ &= \prod_{i=\alpha+1}^{r} \left(\frac{1 - zX_{[i,r]} + zx_i}{1 - zX_{[i,r]}}\right) \\ &= \prod_{i=\alpha+1}^{r} \left(\frac{1 - zX_{[i+1,r]}}{1 - zX_{[i,r]}}\right) \\ &= \left(\frac{1 - zX_{[\alpha+2,r]}}{1 - zX_{[\alpha+1,r]}}\right)\left(\frac{1 - zX_{[\alpha+3,r]}}{1 - zX_{[\alpha+2,r]}}\right)\left(\frac{1 - zX_{[\alpha+4,r]}}{1 - zX_{[\alpha+3,r]}}\right) \\ &\qquad \cdots \left(\frac{1 - zX_{[r,r]}}{1 - zX_{[r-1,r]}}\right)\left(\frac{1}{1 - zX_{[r,r]}}\right) \\ &= \frac{1}{1 - zX_{[\alpha+1,r]}}, \end{aligned}$$

whose corresponding exponential generating function is

$$N_{\min}^{\mathrm{E}} := e^{zX_{[\alpha+1,r]}}. \tag{18.7}$$

The other operand of the shuffle operator in this case is the ordinary generating function

$$\begin{aligned} \frac{\partial}{\partial u} N_{\max}\Big|_{u=1} &= \frac{\partial}{\partial u} \prod_{i=1}^{\alpha-1} \left(1 + \frac{zux_i}{1 - zX_{[1,i]}}\right)\Bigg|_{u=1} \\ &= \frac{\partial}{\partial u} \log \prod_{i=1}^{\alpha-1} \left(1 + \frac{zux_i}{1 - zX_{[1,i]}}\right)\Bigg|_{u=1} \cdot \prod_{i=1}^{\alpha-1} \left(1 + \frac{zux_i}{1 - zX_{[1,i]}}\right)\Bigg|_{u=1} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{\alpha-1} \frac{\partial}{\partial u} \log\left(1 + \frac{zux_i}{1 - zX_{[1,i]}}\right)\Big|_{u=1} \cdot \prod_{i=1}^{\alpha-1} \left(\frac{1 - zX_{[1,i]} + zx_i}{1 - zX_{[1,i]}}\right) \\
&= \sum_{i=1}^{\alpha-1} \frac{\partial}{\partial u} \log\left(\frac{1 - zX_{[1,i]} + zux_i}{1 - zX_{[1,i]}}\right)\Big|_{u=1} \cdot \prod_{i=1}^{\alpha-1} \left(\frac{1 - zX_{[1,i-1]}}{1 - zX_{[1,i]}}\right) \\
&= \sum_{i=1}^{\alpha-1} \frac{\partial}{\partial u} \log\left(\frac{1 - zX_{[1,i]} + zux_i}{1 - zX_{[1,i]}}\right)\Big|_{u=1} \\
&\quad \cdot \left(\frac{1}{1 - zX_{[1,1]}}\right)\left(\frac{1 - zX_{[1,1]}}{1 - zX_{[1,2]}}\right) \cdots \left(\frac{1 - zX_{[1,i-3]}}{1 - zX_{[1,\alpha-2]}}\right)\left(\frac{1 - zX_{[1,\alpha-2]}}{1 - zX_{[1,\alpha-1]}}\right) \\
&= \sum_{i=1}^{\alpha-1} \frac{\partial}{\partial u} \log\left(\frac{1 - zX_{[1,i]} + zux_i}{1 - zX_{[1,i]}}\right)\Big|_{u=1} \cdot \frac{1}{1 - zX_{[1,\alpha-1]}} \\
&= \sum_{i=1}^{\alpha-1} \frac{1}{\frac{1 - zX_{[1,i]} + zux_i}{1 - zX_{[1,i]}}} \frac{zx_i}{1 - zX_{[1,i]}}\Big|_{u=1} \cdot \frac{1}{1 - zX_{[1,\alpha-1]}} \\
&= \sum_{i=1}^{\alpha-1} \frac{1 - zX_{[1,i]}}{1 - zX_{[1,i-1]}} \frac{zx_i}{1 - zX_{[1,i]}} \frac{1}{1 - zX_{[1,\alpha-1]}} \\
&= \sum_{i=1}^{\alpha-1} \frac{zx_i}{1 - zX_{[1,i-1]}} \frac{1}{1 - zX_{[1,\alpha-1]}}.
\end{aligned}
$$

Under partial fraction decomposition, this becomes

$$
\frac{\partial}{\partial u} N_{\max}\Big|_{u=1} = \sum_{i=1}^{\alpha-1} \frac{x_i}{X_{[i,\alpha-1]}} \left(\frac{1}{1 - zX_{[1,\alpha-1]}} - \frac{1}{1 - zX_{[1,i-1]}}\right),
$$

which is easy to transform into an exponential generating function as before, giving

$$
N_{\max}^{\mathrm{Ed}} := \sum_{i=1}^{\alpha-1} \frac{x_i}{X_{[i,\alpha-1]}} \left(e^{zX_{[1,\alpha-1]}} - e^{zX_{[1,i-1]}}\right). \tag{18.8}
$$

The product of (18.7) and (18.8) gives us the exponential generating function of the second term in (18.3) (i.e., performs the shuffle product of the two ordinary generating functions), which is

$$
\begin{aligned}
&e^{zX_{[\alpha+1,r]}} \cdot \sum_{i=1}^{\alpha-1} \frac{x_i}{X_{[i,\alpha-1]}} \left(e^{zX_{[1,\alpha-1]}} - e^{zX_{[1,i-1]}}\right) \\
&= \sum_{i=1}^{\alpha-1} \frac{x_i}{X_{[i,\alpha-1]}} \left(e^{zX_{[\alpha+1,r]}} e^{zX_{[1,\alpha-1]}} - e^{zX_{[\alpha+1,r]}} e^{zX_{[1,i-1]}}\right) \\
&= \sum_{i=1}^{\alpha-1} \frac{x_i}{X_{[i,\alpha-1]}} \left(e^{z\left(X_{[1,\alpha-1]} + X_{[\alpha+1,r]}\right)} - e^{z\left(X_{[1,i-1]} + X_{[\alpha+1,r]}\right)}\right).
\end{aligned}
$$

This can be transformed back using the Laplace transform, to get the ordinary generating function:

$$
N_{\min}\Big|_{u=1} ш \left(\frac{\partial}{\partial u} N_{\max}\Big|_{u=1}\right)
$$

$$= \sum_{i=1}^{\alpha-1} \frac{x_i}{X_{[i,\alpha-1]}} \left( \frac{1}{1 - z\big(X_{[1,\alpha-1]} + X_{[\alpha+1,r]}\big)} - \frac{1}{1 - z\big(X_{[1,i-1]} + X_{[\alpha+1,r]}\big)} \right). \quad (18.9)$$

Therefore, by adding (18.6) and (18.9), we have the first-order partial derivative (with respect to $u$) of the shuffle product in the cost function (18.2). We must still include the final factor $\frac{zx_\alpha}{1-zX}$, but since this is independent of $u$ (and thus does not need to be differentiated) this is a simple product, giving

$$\frac{zx_\alpha}{1-zX} \left[ \sum_{i=1}^{\alpha-1} \frac{x_i}{X_{[i,\alpha-1]}} \left( \frac{1}{1 - z\big(X_{[1,\alpha-1]} + X_{[\alpha+1,r]}\big)} - \frac{1}{1 - z\big(X_{[1,i-1]} + X_{[\alpha+1,r]}\big)} \right) \right.$$
$$\left. + \sum_{i=\alpha+1}^{r} \frac{x_i}{X_{[\alpha+1,i]}} \left( \frac{1}{1 - z\big(X_{[1,\alpha-1]} + X_{[\alpha+1,r]}\big)} - \frac{1}{1 - z\big(X_{[1,\alpha-1]} + X_{[i+1,r]}\big)} \right) \right]. \quad (18.10)$$

Multiplying through by the factor $\frac{zx_\alpha}{1-zX}$ produces four terms of the form $\frac{az}{(1-bz)(1-cz)}$, each of which can be decomposed into partial fractions. Then the coefficient can be found, and thus ultimately the expected value. We make use of Maple for the partial fraction decompositions. For the first term,

$$\sum_{i=1}^{\alpha-1} \frac{x_i x_\alpha}{X_{[i,\alpha-1]}} \frac{z}{(1-zX)\big(1 - z\big(X_{[1,\alpha-1]} + X_{[\alpha+1,r]}\big)\big)}$$
$$= \sum_{i=1}^{\alpha-1} \frac{x_i x_\alpha}{X_{[i,\alpha-1]}} \left( \frac{1}{x_\alpha(1-zX)} - \frac{1}{x_\alpha\big(1 - z\big(X_{[1,\alpha-1]} + X_{[\alpha+1,r]}\big)\big)} \right)$$
$$= \sum_{i=1}^{\alpha-1} \frac{x_i}{X_{[i,\alpha-1]}} \left( \frac{1}{1-zX} - \frac{1}{1 - z\big(X_{[1,\alpha-1]} + X_{[\alpha+1,r]}\big)} \right). \quad (18.11)$$

The second term is

$$-\sum_{i=1}^{\alpha-1} \frac{x_i x_\alpha}{X_{[i,\alpha-1]}} \frac{z}{(1-zX)\big(1 - z\big(X_{[1,i-1]} + X_{[\alpha+1,r]}\big)\big)}$$
$$= -\sum_{i=1}^{\alpha-1} \frac{x_i x_\alpha}{X_{[i,\alpha-1]} X_{[i,\alpha]}} \left( \frac{1}{1-zX} - \frac{1}{1 - z\big(X_{[1,i-1]} + X_{[\alpha+1,r]}\big)} \right), \quad (18.12)$$

and the third term is

$$\sum_{i=\alpha+1}^{r} \frac{x_\alpha x_i}{X_{[\alpha+1,i]}} \frac{z}{(1-zX)\big(1 - z\big(X_{[1,\alpha-1]} + X_{[\alpha+1,r]}\big)\big)}$$
$$= \sum_{i=\alpha+1}^{r} \frac{x_i}{X_{[\alpha+1,i]}} \left( \frac{1}{1-zX} - \frac{1}{1 - z\big(X_{[1,\alpha-1]} + X_{[\alpha+1,r]}\big)} \right). \quad (18.13)$$

Finally the fourth term is

$$-\sum_{i=\alpha+1}^{r} \frac{x_\alpha x_i}{X_{[\alpha+1,i]}} \frac{z}{(1-zX)\big(1 - z\big(X_{[1,\alpha-1]} + X_{[i+1,r]}\big)\big)}$$

$$= -\sum_{i=\alpha+1}^{r} \frac{x_\alpha x_i}{X_{[\alpha+1,i]}X_{[\alpha,i]}}\left(\frac{1}{1-zX} - \frac{1}{1-z\big(X_{[1,\alpha-1]}+X_{[i+1,r]}\big)}\right). \qquad (18.14)$$

Once we have the coefficient of $z^n x_1^{n_1}\cdots x_r^{n_r}$ in each of the four quantities, we divide by the number of possible words to get the expected value. First we note that of the two terms in (18.11) and (18.13), the second in each case has no $x_\alpha$ (i.e., $n_\alpha = 0$) and thus these terms need not be counted. So we are left with one term each for (18.11) and (18.13) and two terms each for (18.12) and (18.14). We deal with these by first considering the coefficient of $z^n$ and then rewriting the expressions in terms of multinomial expansions so that the 'trouble' terms in the denominator can be cancelled. All that then remains is to expand further in terms of multinomial expansions to get the coefficient of $x_1^{n_1}\cdots x_r^{n_r}$. Only the methods for (18.12) and (18.13) are included, since (18.11) and (18.14) are similar. For (18.13) we have:

$$\begin{aligned}
&[z^n]\sum_{i=\alpha+1}^{r}\frac{x_i}{X_{[\alpha+1,i]}}\frac{1}{1-zX}\\
&\quad=\sum_{i=\alpha+1}^{r}\frac{x_i}{X_{[\alpha+1,i]}}X^n\\
&\quad=\sum_{i=\alpha+1}^{r}\frac{x_i}{X_{[\alpha+1,i]}}\sum_{k_1+k_2+k_3=n}\binom{n}{k_1,k_2,k_3}(X_{[1,\alpha]})^{k_1}(X_{[\alpha+1,i]})^{k_2}(X_{[i+1,r]})^{k_3}\\
&\quad=\sum_{i=\alpha+1}^{r}x_i\sum_{k_1+k_2+k_3=n}\binom{n}{k_1,k_2,k_3}(X_{[1,\alpha]})^{k_1}(X_{[\alpha+1,i]})^{k_2-1}(X_{[i+1,r]})^{k_3}\\
&\quad=\sum_{i=\alpha+1}^{r}x_i\sum_{k_1+k_2+k_3=n}\binom{n}{k_1,k_2,k_3}\sum_{b_1+b_2+\cdots+b_\alpha=k_1}\binom{k_1}{b_1,b_2,\ldots,b_\alpha}x_1^{b_1}\cdots x_\alpha^{b_\alpha}\\
&\qquad\cdot\sum_{b_{\alpha+1}+\cdots+b_i=k_2-1}\binom{k_2-1}{b_{\alpha+1},\ldots,b_i}x_{\alpha+1}^{b_{\alpha+1}}\cdots x_i^{b_i}\\
&\qquad\cdot\sum_{b_{i+1}+\cdots+b_r=k_3}\binom{k_3}{b_{i+1},\ldots,b_r}x_{i+1}^{b_{i+1}}\cdots x_r^{b_r}. \qquad (18.15)
\end{aligned}$$

To find the coefficient of $x_1^{n_1}\cdots x_r^{n_r}$, we can equate all $n_m$ to $b_m$ for all $m \in \{1,\ldots,r\}$ except for $m = i$, when $n_i = b_i + 1$ (because of the extra $x_i$ in the first line of (18.15)). The coefficient is then

$$\begin{aligned}
&\sum_{i=\alpha+1}^{r}\binom{n}{b_1+b_2+\cdots+b_\alpha,\ b_{\alpha+1}+\cdots+b_i+1,\ b_{i+1}+\cdots+b_r}\\
&\qquad\cdot\binom{b_1+b_2+\cdots+b_\alpha}{b_1,b_2,\ldots,b_\alpha}\binom{b_{\alpha+1}+\cdots+b_i}{b_{\alpha+1},\ldots,b_i}\binom{b_{i+1}+\cdots+b_r}{b_{i+1},\ldots,b_r}\\
&=\sum_{i=\alpha+1}^{r}\binom{n}{N_{[1,\alpha]},N_{[\alpha+1,i]},N_{[i+1,r]}}
\end{aligned}$$

$$
\begin{aligned}
&\quad \cdot \binom{N_{[1,\alpha]}}{n_1, n_2, \ldots, n_\alpha} \binom{N_{[\alpha+1,i]} - 1}{n_{\alpha+1}, \ldots, n_i - 1} \binom{N_{[i+1,r]}}{n_{i+1}, \ldots, n_r} \\
&= \sum_{i=\alpha+1}^{r} \frac{n!}{N_{[1,\alpha]}! N_{[\alpha+1,i]}! N_{[i+1,r]}!} \\
&\quad \cdot \frac{N_{[1,\alpha]}!(N_{[\alpha+1,i]} - 1)! N_{[i+1,r]}!}{n_1! n_2! \cdots n_\alpha! n_{\alpha+1}! \cdots (n_i - 1)! n_{i+1}! \cdots n_r!} \\
&= \sum_{i=\alpha+1}^{r} \frac{n_i}{N_{[\alpha+1,i]}} \frac{n!}{n_1! n_2! \cdots n_\alpha! n_{\alpha+1}! \cdots n_i! n_{i+1}! \cdots n_r!} \\
&= \sum_{i=\alpha+1}^{r} \frac{n_i}{N_{[\alpha+1,i]}} \binom{n}{n_1, \ldots, n_r!}.
\end{aligned}
$$

For (18.12) we have (ignoring the negative):

$$
\begin{aligned}
&[z^n] \sum_{i=1}^{\alpha-1} \frac{x_i x_\alpha}{X_{[i,\alpha-1]} X_{[i,\alpha]}} \left( \frac{1}{1 - zX} - \frac{1}{1 - z\big(X_{[1,i-1]} + X_{[\alpha+1,r]}\big)} \right) \\
&= \sum_{i=1}^{\alpha-1} \frac{x_i x_\alpha}{X_{[i,\alpha-1]} X_{[i,\alpha]}} \Big( X^n - \big(X_{[1,i-1]} + X_{[\alpha+1,r]}\big)^n \Big) \\
&= \sum_{i=1}^{\alpha-1} \frac{x_i x_\alpha}{X_{[i,\alpha-1]} X_{[i,\alpha]}} \Big( \big(X_{[1,i-1]} + X_{[i,\alpha]} + X_{[\alpha+1,r]}\big)^n - \big(X_{[1,i-1]} + X_{[\alpha+1,r]}\big)^n \Big) \\
&= \sum_{i=1}^{\alpha-1} \frac{x_i x_\alpha}{X_{[i,\alpha-1]} X_{[i,\alpha]}} \left( \sum_{k=0}^{n} \binom{n}{k} X_{[i,\alpha]}^k \big(X_{[1,i-1]} + X_{[\alpha+1,r]}\big)^{n-k} - \big(X_{[1,i-1]} + X_{[\alpha+1,r]}\big)^n \right) \\
&= \sum_{i=1}^{\alpha-1} \frac{x_i x_\alpha}{X_{[i,\alpha-1]} X_{[i,\alpha]}} \sum_{k=1}^{n} \binom{n}{k} X_{[i,\alpha]}^k \big(X_{[1,i-1]} + X_{[\alpha+1,r]}\big)^{n-k} \\
&= \sum_{i=1}^{\alpha-1} \frac{x_i x_\alpha}{X_{[i,\alpha-1]}} \sum_{k=1}^{n} \binom{n}{k} X_{[i,\alpha]}^{k-1} \big(X_{[1,i-1]} + X_{[\alpha+1,r]}\big)^{n-k} \\
&= \sum_{i=1}^{\alpha-1} \frac{x_i x_\alpha}{X_{[i,\alpha-1]}} \sum_{k=1}^{n} \binom{n}{k} \sum_{j=0}^{k-1} \binom{k-1}{j} X_{[i,\alpha-1]}^j x_\alpha^{k-1-j} \big(X_{[1,i-1]} + X_{[\alpha+1,r]}\big)^{n-k} \\
&= \sum_{i=1}^{\alpha-1} x_i x_\alpha \sum_{k=1}^{n} \binom{n}{k} \sum_{j=0}^{k-1} \binom{k-1}{j} X_{[i,\alpha-1]}^{j-1} x_\alpha^{k-1-j} \big(X_{[1,i-1]} + X_{[\alpha+1,r]}\big)^{n-k} \\
&= \sum_{i=1}^{\alpha-1} x_i x_\alpha \sum_{k=1}^{n} \binom{n}{k} \sum_{j=0}^{k-1} \binom{k-1}{j} \sum_{b_i + \cdots + b_{\alpha-1} = j-1} \binom{j-1}{b_i, \ldots, b_{\alpha-1}} x_i^{b_i} \cdots x_{\alpha-1}^{b_{\alpha-1}} x_\alpha^{k-1-j} \\
&\quad \cdot \sum_{b_1 + \cdots + b_{i-1} + b_{\alpha+1} \cdots b_r = n-k} \binom{n-k}{b_1, \ldots, b_{i-1}, b_{\alpha+1}, \ldots, b_r} x_1^{b_1} \cdots x_{i-1}^{b_{i-1}} x_{\alpha+1}^{b_{\alpha+1}} \cdots x_r^{b_r}.
\end{aligned}
$$

At this stage we can find the coefficient of $x_1^{n_1} \cdots x_r^{n_r}$ by equating $n_m$ with $b_m$ for all $m \in \{1, \ldots, r\}$ except $i$ and $\alpha$. For these we have $n_i = b_i + 1$ and $n_\alpha = k - j$.

We end up with ($n = n_1 + \cdots + n_r$)

$$\sum_{i=1}^{\alpha-1} \binom{n}{N_{[1,i-1]} + N_{[\alpha+1,r]}} \binom{n - (N_{[1,i-1]} + N_{[\alpha+1,r]}) - 1}{n_i - 1 + \cdots + n_{\alpha-1} + 1}$$
$$\cdot \binom{n_i - 1 + \cdots + n_{\alpha-1}}{n_i - 1, \ldots, n_{\alpha-1}} \binom{N_{[1,i-1]} + N_{[\alpha+1,r]}}{n_1, \ldots, n_{i-1}, n_{\alpha+1}, \ldots, n_r}$$
$$= \sum_{i=1}^{\alpha-1} \binom{n}{N_{[i,\alpha]}} \binom{N_{[i,\alpha]} - 1}{N_{[i,\alpha-1]}} \binom{N_{[i,\alpha-1]} - 1}{n_i - 1, \ldots, n_{\alpha-1}} \binom{N_{[1,i-1]} + N_{[\alpha+1,r]}}{n_1, \ldots, n_{i-1}, n_{\alpha+1}, \ldots, n_r}$$
$$= \sum_{i=1}^{\alpha-1} \frac{n!}{N_{[i,\alpha]}!(n - n_i - \cdots - n_\alpha)!} \frac{(N_{[i,\alpha]} - 1)!}{N_{[i,\alpha-1]}!(N_{[i,\alpha]} - 1 - N_{[i,\alpha-1]})!}$$
$$\cdot \frac{(N_{[i,\alpha-1]} - 1)!}{(n_i - 1)!, \ldots, n_{\alpha-1}!} \frac{(N_{[1,i-1]} + N_{[\alpha+1,r]})!}{n_1!, \ldots, n_{i-1}!, n_{\alpha+1}!, \ldots, n_r!}$$
$$= \sum_{i=1}^{\alpha-1} \frac{n_i n_\alpha}{N_{[i,\alpha]} N_{[i,\alpha-1]}} \binom{n}{n_1, \ldots, n_r}.$$

The method is the same for (18.11) and (18.14), and after dividing by the number of all possible words $\binom{n}{n_1,\ldots,n_r}$ we end up with:

$$\mathbb{E}^{\mathrm{m}}_\alpha = \sum_{i=1}^{\alpha-1} \left( \frac{n_i}{N_{[i,\alpha-1]}} - \frac{n_i n_\alpha}{N_{[i,\alpha]} N_{[i,\alpha-1]}} \right) + \sum_{i=\alpha+1}^{r} \left( \frac{n_i}{N_{[\alpha+1,i]}} - \frac{n_i n_\alpha}{N_{[\alpha,i]} N_{[\alpha+1,i]}} \right)$$
$$= \sum_{i=1}^{\alpha-1} \frac{n_i N_{[i,\alpha]} - n_i n_\alpha}{N_{[i,\alpha]} N_{[i,\alpha-1]}} + \sum_{i=\alpha+1}^{r} \frac{n_i N_{[\alpha,i]} - n_i n_\alpha}{N_{[\alpha,i]} N_{[\alpha+1,i]}}$$
$$= \sum_{i=1}^{\alpha-1} \frac{n_i N_{[i,\alpha-1]}}{N_{[i,\alpha]} N_{[i,\alpha-1]}} + \sum_{i=\alpha+1}^{r} \frac{n_i N_{[\alpha+1,i]}}{N_{[\alpha,i]} N_{[\alpha+1,i]}},$$

so that

$$\mathbb{E}^{\mathrm{m}}_\alpha = \sum_{i=1}^{\alpha-1} \frac{n_i}{N_{[i,\alpha]}} + \sum_{i=\alpha+1}^{r} \frac{n_i}{N_{[\alpha,i]}}. \tag{18.16}$$

This concludes the proof of Theorem 18.1. ■

## 18.3 Expectation – probability model

As we did in the left-going and right-going cases, we can now use a few of the calculations from the multiset model to help us find the result using the probability model. In this case we have (see equation (18.2)):

$$f(z, u, x_1, \ldots, x_r) = \left[ \prod_{i=1}^{\alpha-1} \left(1 + \frac{zux_i}{1 - zX_{[1,i]}}\right) ш \prod_{i=\alpha+1}^{r} \left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right) \right] \frac{zx_\alpha}{1 - zX}$$
$$= \sum_{w \in A^*} z^{|w|} u^{a(w)} x_1^{|w|_1} x_2^{|w|_2} \cdots x_r^{|w|_r} \tag{18.17}$$

where $w$ is a word of length $|w|$ from alphabet $A = \{1, \dots, r\}$, $|w|_i$ is the number of $i$'s in the word, and $a(w)$ is the depth of $\alpha$. The expected value is

$$\mathbb{E}^{\mathrm{p}}_{\alpha} = [z^n]\frac{\partial}{\partial u} f(z, u, p_1, \dots, p_r)\Big|_{u=1},$$

where $p_i$ is the probability of letter $i$ appearing in the word (input sequence).

We have

$$\begin{aligned}
\mathbb{E}^{\mathrm{p}}_{\alpha} = [z^n]\frac{zp_\alpha}{1-zP}\Bigg[&\sum_{i=1}^{\alpha-1}\frac{p_i}{P_{[i,\alpha-1]}}\left(\frac{1}{1-z\big(P_{[1,\alpha-1]}+P_{[\alpha+1,r]}\big)}-\frac{1}{1-z\big(P_{[1,i-1]}+P_{[\alpha+1,r]}\big)}\right)\\
&+\sum_{i=\alpha+1}^{r}\frac{p_i}{P_{[\alpha+1,i]}}\left(\frac{1}{1-z\big(P_{[1,\alpha-1]}+P_{[\alpha+1,r]}\big)}-\frac{1}{1-z\big(P_{[1,\alpha-1]}+P_{[i+1,r]}\big)}\right)\Bigg]\\
= [z^n]\Bigg[&\sum_{i=1}^{\alpha-1}\frac{p_i}{P_{[i,\alpha-1]}}\left(\frac{1}{1-zP}-\frac{1}{1-z\big(P_{[1,\alpha-1]}+P_{[\alpha+1,r]}\big)}\right)\\
&-\sum_{i=1}^{\alpha-1}\frac{p_ip_\alpha}{P_{[i,\alpha-1]}}\left(\frac{1}{P_{[i,\alpha]}(1-zP)}-\frac{1}{P_{[i,\alpha]}\big(1-z\big(P_{[1,i-1]}+P_{[\alpha+1,r]}\big)\big)}\right)\\
&+\sum_{i=\alpha+1}^{r}\frac{p_i}{P_{[\alpha+1,i]}}\left(\frac{1}{1-zP}-\frac{1}{1-z\big(P_{[1,\alpha-1]}+P_{[\alpha+1,r]}\big)}\right)\\
&-\sum_{i=\alpha+1}^{r}\frac{p_\alpha p_i}{P_{[\alpha+1,i]}}\left(\frac{1}{P_{[\alpha,i]}(1-zP)}-\frac{1}{P_{[\alpha,i]}\big(1-z\big(P_{[1,\alpha-1]}+P_{[i+1,r]}\big)\big)}\right)\Bigg].
\end{aligned}$$

Since $P = p_1 + p_2 + \cdots + p_r = 1$ and $P_{[i,j]} < 1$ in all other cases, the coefficient of $z^n$ of the second term in each sum will go to zero as $n \to \infty$. This leaves us with

$$\begin{aligned}
\mathbb{E}^{\mathrm{p}}_{\alpha} \sim [z^n]\Bigg[&\sum_{i=1}^{\alpha-1}\frac{p_i}{P_{[i,\alpha-1]}}\frac{1}{1-zP}-\sum_{i=1}^{\alpha-1}\frac{p_ip_\alpha}{P_{[i,\alpha-1]}}\frac{1}{P_{[i,\alpha]}(1-zP)}\\
&+\sum_{i=\alpha+1}^{r}\frac{p_i}{P_{[\alpha+1,i]}}\frac{1}{1-zP}-\sum_{i=\alpha+1}^{r}\frac{p_\alpha p_i}{P_{[\alpha+1,i]}}\frac{1}{P_{[\alpha,i]}(1-zP)}\Bigg]\\
&= \sum_{i=1}^{\alpha-1}\frac{p_i}{P_{[i,\alpha-1]}}P^n-\sum_{i=1}^{\alpha-1}\frac{p_ip_\alpha}{P_{[i,\alpha-1]}P_{[i,\alpha]}}P^n+\sum_{i=\alpha+1}^{r}\frac{p_i}{P_{[\alpha+1,i]}}P^n-\sum_{i=\alpha+1}^{r}\frac{p_\alpha p_i}{P_{[\alpha+1,i]}P_{[\alpha,i]}}P^n\\
&= \sum_{i=1}^{\alpha-1}\frac{p_i}{P_{[i,\alpha-1]}}-\sum_{i=1}^{\alpha-1}\frac{p_ip_\alpha}{P_{[i,\alpha-1]}P_{[i,\alpha]}}+\sum_{i=\alpha+1}^{r}\frac{p_i}{P_{[\alpha+1,i]}}-\sum_{i=\alpha+1}^{r}\frac{p_\alpha p_i}{P_{[\alpha+1,i]}P_{[\alpha,i]}}\\
&= \sum_{i=1}^{\alpha-1}\frac{p_iP_{[i,\alpha]}-p_ip_\alpha}{P_{[i,\alpha-1]}P_{[i,\alpha]}}+\sum_{i=\alpha+1}^{r}\frac{p_iP_{[\alpha,i]}-p_\alpha p_i}{P_{[\alpha+1,i]}P_{[\alpha,i]}}\\
&= \sum_{i=1}^{\alpha-1}\frac{p_iP_{[i,\alpha-1]}}{P_{[i,\alpha-1]}P_{[i,\alpha]}}+\sum_{i=\alpha+1}^{r}\frac{p_iP_{[\alpha+1,i]}}{P_{[\alpha+1,i]}P_{[\alpha,i]}}.
\end{aligned}$$

Therefore

$$\mathbb{E}^{\mathrm{p}}_{\alpha} \sim \sum_{i=1}^{\alpha-1}\frac{p_i}{P_{[i,\alpha]}}+\sum_{i=\alpha+1}^{r}\frac{p_i}{P_{[\alpha,i]}}, \tag{18.18}$$

as $n \to \infty$, which completes the proof of Theorem 18.2. ■

## 18.4 Expectation – all keys distinct

Now we make the substitutions $n_i = 1$ for all $i \in \{1, \ldots, r\}$ which will mean that we consider the case where equal keys are not allowed. These calculations have been done previously and we expect an average search cost of

$$2H_n - 3 + \frac{2H_n}{n} \tag{18.19}$$

(see [36, page 249]). The final term is given as negative in [36], but it is obtained by dividing

$$2(n+1)(H_{n+1} - 1) - 2n,$$

(the result of Theorem 5.5 [36, page 247]), by $n$ and adding 1, which produces a positive third term. The result stated above gives us the average search cost for finding key $\alpha$ in a binary search tree. Thus we must sum our result over $\alpha = 1, \ldots, r$ and divide by $r$ to get the average. We also note that if all keys are distinct, the length of the word ($n$) is the same as the length of the alphabet ($r$), as we assume all keys are used exactly once. Thus from (18.16) we have (where 'd' stands for distinct)

$$\begin{aligned}
\mathbb{E}_\alpha^{\mathrm{d}} &= \sum_{\alpha=1}^{n} \left( \sum_{i=1}^{\alpha-1} \frac{1}{\alpha - i + 1} + \sum_{i=\alpha+1}^{n} \frac{1}{i - \alpha + 1} \right) \frac{1}{n} \\
&= \sum_{\alpha=1}^{n} \left( \sum_{k=2}^{\alpha} \frac{1}{k} + \sum_{k=2}^{n-\alpha+1} \frac{1}{k} \right) \frac{1}{n} \\
&= \sum_{\alpha=1}^{n} \big( H_\alpha - 1 + H_{n-\alpha+1} - 1 \big) \frac{1}{n}.
\end{aligned}$$

Since

$$\sum_{k=1}^{n} H_k = (n+1)H_n - n, \tag{18.20}$$

from [23], we can continue to get

$$\begin{aligned}
\mathbb{E}_\alpha^{\mathrm{d}} &= \sum_{\alpha=1}^{n} \big( H_\alpha - 1 + H_{n-\alpha+1} - 1 \big) \frac{1}{n} \\
&= \left( \sum_{\alpha=1}^{n} H_\alpha + \sum_{\alpha=1}^{n} H_{n-\alpha+1} - \sum_{\alpha=1}^{n} 2 \right) \frac{1}{n} \\
&= \Big( (n+1)H_n - n + \sum_{i=1}^{n} H_i - 2n \Big) \frac{1}{n}
\end{aligned}$$

$$
\begin{aligned}
&= \Big((n+1)H_n - n + (n+1)H_n - n - 2n\Big)\frac{1}{n} \\
&= 2H_n + \frac{2H_n}{n} - 4.
\end{aligned}
$$

Finally note that looking at left-to-right maxima/minima implies that this is the average search cost less one. Graphically, this means that we were counting branches and not nodes in the binary search tree. This can easily be remedied by adding one to the result, which gives us equation (18.19).

## 18.5 Variance – multiset model

Again we start with the generating function (see equation (18.2))

$$
\begin{aligned}
f(z,u,x_1,\ldots,x_r) &= \big[N_{\max}(z,u,X_{[1,\alpha-1]}) ш N_{\min}(z,u,X_{[\alpha+1,r]})\big]\frac{zx_\alpha}{1-zX} \\
&= \left[\prod_{i=1}^{\alpha-1}\left(1+\frac{zux_i}{1-zX_{[1,i]}}\right) ш \prod_{i=\alpha+1}^{r}\left(1+\frac{zux_i}{1-zX_{[i,r]}}\right)\right]\frac{zx_\alpha}{1-zX}.
\end{aligned}
$$

To find the variance we need to find

$$
\begin{aligned}
\mathbb{V}_\alpha^{\mathrm{m}} := [z^n x_1^{n_1}\cdots x_r^{n_r}]\frac{\partial^2}{\partial u^2}f(z,1,x_1,\ldots,x_r) + [z^n x_1^{n_1}\cdots x_r^{n_r}]\frac{\partial}{\partial u}f(z,1,x_1,\ldots,x_r) \\
- \left([z^n x_1^{n_1}\cdots x_r^{n_r}]\frac{\partial}{\partial u}f(z,1,x_1,\ldots,x_r)\right)^2. \qquad (18.21)
\end{aligned}
$$

The expected value calculations will give us the second and third terms, so we concentrate on the first term (i.e., the second moment) which requires us to perform a second-order derivative. For the time being we ignore the factor $\frac{zx_\alpha}{1-zX}$ which is independent of $u$, and we let

$$
f := N_{\max} = \prod_{i=1}^{\alpha-1}\left(1+\frac{zux_i}{1-zX_{[1,i]}}\right),
$$

and

$$
g := N_{\min} = \prod_{i=\alpha+1}^{r}\left(1+\frac{zux_i}{1-zX_{[i,r]}}\right),
$$

to write

$$
\begin{aligned}
\frac{\partial^2}{\partial u^2}(f ш g)\Big|_{u=1} &= \frac{\partial^2}{\partial u^2}f\Big|_{u=1} ш g\Big|_{u=1} && (18.22)\\
&\quad + 2\frac{\partial}{\partial u}f\Big|_{u=1} ш \frac{\partial}{\partial u}g\Big|_{u=1} && (18.23)\\
&\quad + \frac{\partial^2}{\partial u^2}g\Big|_{u=1} ш f\Big|_{u=1}. && (18.24)
\end{aligned}
$$

Of the above, only the two second-order partial derivatives in (18.22) and (18.24) have not been already calculated (see (18.4), (18.5), (18.7) and (18.8)). We thus look at those now (recall that $f'' = f(\log f)'' + f'(\log f)')$:

$$
\begin{aligned}
\frac{\partial^2}{\partial u^2} f\Big|_{u=1} &= \frac{\partial^2}{\partial u^2} N_{\max}\Big|_{u=1} \\
&= \frac{\partial^2}{\partial u^2} \prod_{i=1}^{\alpha-1} \left(1 + \frac{zux_i}{1 - zX_{[1,i]}}\right)\Bigg|_{u=1} \\
&= \prod_{i=1}^{\alpha-1} \left(1 + \frac{zux_i}{1 - zX_{[1,i]}}\right)\Bigg|_{u=1} \cdot \frac{\partial^2}{\partial u^2} \sum_{i=1}^{\alpha-1} \log\left(1 + \frac{zux_i}{1 - zX_{[1,i]}}\right)\Bigg|_{u=1} \\
&\quad + \frac{\partial}{\partial u} \prod_{i=1}^{\alpha-1} \left(1 + \frac{zux_i}{1 - zX_{[1,i]}}\right)\Bigg|_{u=1} \cdot \frac{\partial}{\partial u} \sum_{i=1}^{\alpha-1} \log\left(1 + \frac{zux_i}{1 - zX_{[1,i]}}\right)\Bigg|_{u=1} \\
&= \frac{1}{1 - zX_{[1,\alpha-1]}} \left( - \sum_{i=1}^{\alpha-1} \frac{z^2 x_i^2}{(1 - zX_{[1,i-1]})^2} \right) \\
&\quad + \frac{1}{1 - zX_{[1,\alpha-1]}} \left( \sum_{i=1}^{\alpha-1} \frac{zx_i}{1 - zX_{[1,i-1]}} \right)^2 \\
&= \frac{1}{1 - zX_{[1,\alpha-1]}} \left( \left( \sum_{i=1}^{\alpha-1} \frac{zx_i}{1 - zX_{[1,i-1]}} \right)^2 - \sum_{i=1}^{\alpha-1} \frac{z^2 x_i^2}{(1 - zX_{[1,i-1]})^2} \right).
\end{aligned}
$$

The square of the sum will give us all combinations of terms such as

$$
\frac{z^2 x_i x_k}{(1 - zX_{[1,i-1]})(1 - zX_{[k-1]})}.
$$

The terms for which $i = k$ will be cancelled by the second sum, thus:

$$
\begin{aligned}
&\frac{\partial^2}{\partial u^2} f\Big|_{u=1} = \frac{\partial^2}{\partial u^2} N_{\max}\Big|_{u=1} \\
&= \frac{1}{1 - zX_{[1,\alpha-1]}} \left( \left( \sum_{i=1}^{\alpha-1} \frac{zx_i}{1 - zX_{[1,i-1]}} \right)^2 - \sum_{i=1}^{\alpha-1} \frac{z^2 x_i^2}{(1 - zX_{[1,i-1]})^2} \right) \\
&= 2 \sum_{i=1}^{\alpha-1} \sum_{k=1}^{i-1} \Bigg( \frac{x_i x_k}{(X_{[1,\alpha-1]} - X_{[1,i-1]})(X_{[1,\alpha-1]} - X_{[1,k-1]})(1 - zX_{[1,\alpha-1]})} \\
&\qquad - \frac{x_i x_k}{(X_{[1,i-1]} - X_{[1,k-1]})(X_{[1,\alpha-1]} - X_{[1,i-1]})(1 - zX_{[1,i-1]})} \\
&\qquad + \frac{x_i x_k}{(X_{[1,i-1]} - X_{[1,k-1]})(X_{[1,\alpha-1]} - X_{[1,k-1]})(1 - zX_{[1,k-1]})} \Bigg) \\
&= 2 \sum_{i=1}^{\alpha-1} \sum_{k=1}^{i-1} \Bigg( \frac{x_i x_k}{X_{[i,\alpha-1]} X_{[k,\alpha-1]} (1 - zX_{[1,\alpha-1]})} - \frac{x_i x_k}{X_{[k,i-1]} X_{[i,\alpha-1]} (1 - zX_{[1,i-1]})} \\
&\qquad + \frac{x_i x_k}{X_{[k,i-1]} X_{[k,\alpha-1]} (1 - zX_{[1,k-1]})} \Bigg),
\end{aligned}
$$

(decomposed into partial fractions in order to perform the shuffle). Similarly

$$\begin{aligned}
\frac{\partial^2}{\partial u^2} g\Big|_{u=1} &= \frac{\partial^2}{\partial u^2} N_{\min}\Big|_{u=1} \\
&= \frac{\partial^2}{\partial u^2} \prod_{i=\alpha+1}^{r} \left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\Bigg|_{u=1} \\
&= \prod_{i=\alpha+1}^{r} \left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\Bigg|_{u=1} \cdot \frac{\partial^2}{\partial u^2} \sum_{i=\alpha+1}^{r} \log\left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\Bigg|_{u=1} \\
&\quad + \frac{\partial}{\partial u} \prod_{i=\alpha+1}^{r} \left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\Bigg|_{u=1} \cdot \frac{\partial}{\partial u} \sum_{i=\alpha+1}^{r} \log\left(1 + \frac{zux_i}{1 - zX_{[i,r]}}\right)\Bigg|_{u=1} \\
&= \frac{1}{1 - zX_{[\alpha+1,r]}} \left( - \sum_{i=\alpha+1}^{r} \frac{z^2 x_i^2}{(1 - zX_{[i+1,r]})^2} \right) \\
&\quad + \frac{1}{1 - zX_{[\alpha+1,r]}} \left( \sum_{i=\alpha+1}^{r} \frac{zx_i}{1 - zX_{[i+1,r]}} \right)^2 \\
&= \frac{1}{1 - zX_{[\alpha+1,r]}} \left( \left( \sum_{i=\alpha+1}^{r} \frac{zx_i}{1 - zX_{[i+1,r]}} \right)^2 - \sum_{i=\alpha+1}^{r} \frac{z^2 x_i^2}{(1 - zX_{[i+1,r]})^2} \right) \\
&= 2 \sum_{i=\alpha+1}^{r} \sum_{k=\alpha+1}^{i-1} \Bigg( \frac{x_i x_k}{(X_{[\alpha+1,r]} - X_{[i+1,r]})(X_{[\alpha+1,r]} - X_{[k+1,r]})(1 - zX_{[\alpha+1,r]})} \\
&\qquad + \frac{x_i x_k}{(X_{[k+1,r]} - X_{[i+1,r]})(X_{[\alpha+1,r]} - X_{[i+1,r]})(1 - zX_{[i+1,r]})} \\
&\qquad - \frac{x_i x_k}{(X_{[k+1,r]} - X_{[i+1,r]})(X_{[\alpha+1,r]} - X_{[k+1,r]})(1 - zX_{[k+1,r]})} \Bigg) \\
&= 2 \sum_{i=\alpha+1}^{r} \sum_{k=\alpha+1}^{i-1} \Bigg( \frac{x_i x_k}{X_{[\alpha+1,i]} X_{[\alpha+1,k]} (1 - zX_{[\alpha+1,r]})} \\
&\qquad + \frac{x_i x_k}{X_{[k+1,i]} X_{[\alpha+1,i]} (1 - zX_{[i+1,r]})} \\
&\qquad - \frac{x_i x_k}{X_{[k+1,i]} X_{[\alpha+1,k]} (1 - zX_{[k+1,r]})} \Bigg).
\end{aligned}$$

Now that the second derivatives have been found, we can proceed to the shuffle products and the Laplace transforms. First, equation (18.22) is a shuffle product between the ordinary generating functions

$$\begin{aligned}
\frac{\partial^2}{\partial u^2} f\Big|_{u=1} = 2 \sum_{i=1}^{\alpha-1} \sum_{k=1}^{i-1} \Bigg( &\frac{x_i x_k}{X_{[i,\alpha-1]} X_{[k,\alpha-1]} (1 - zX_{[1,\alpha-1]})} - \frac{x_i x_k}{X_{[k,i-1]} X_{[i,\alpha-1]} (1 - zX_{[1,i-1]})} \\
&+ \frac{x_i x_k}{X_{[k,i-1]} X_{[k,\alpha-1]} (1 - zX_{[1,k-1]})} \Bigg),
\end{aligned}$$

and

$$g\Big|_{u=1} = \frac{1}{1 - zX_{[\alpha+1,r]}}.$$

These transform to the exponential generating functions

$$2\sum_{i=1}^{\alpha-1}\sum_{k=1}^{i-1}\Bigg(\frac{x_ix_k}{X_{[i,\alpha-1]}X_{[k,\alpha-1]}}e^{zX_{[1,\alpha-1]}}-\frac{x_ix_k}{X_{[k,i-1]}X_{[i,\alpha-1]}}e^{zX_{[1,i-1]}}$$
$$+\frac{x_ix_k}{X_{[k,i-1]}X_{[k,\alpha-1]}}e^{zX_{[1,k-1]}}\Bigg),$$

and

$$e^{zX_{[\alpha+1,r]}}$$

respectively. Now an ordinary product suffices, and we have

$$2\sum_{i=1}^{\alpha-1}\sum_{k=1}^{i-1}\Bigg(\frac{x_ix_k}{X_{[i,\alpha-1]}X_{[k,\alpha-1]}}e^{z(X_{[1,\alpha-1]}+X_{[\alpha+1,r]})}-\frac{x_ix_k}{X_{[k,i-1]}X_{[i,\alpha-1]}}e^{z(X_{[1,i-1]}+X_{[\alpha+1,r]})}$$
$$+\frac{x_ix_k}{X_{[k,i-1]}X_{[k,\alpha-1]}}e^{z(X_{[1,k-1]}+X_{[\alpha+1,r]})}\Bigg),$$

which is transformed back to the ordinary generating function

$$2\sum_{i=1}^{\alpha-1}\sum_{k=1}^{i-1}\Bigg(\frac{x_ix_k}{X_{[i,\alpha-1]}X_{[k,\alpha-1]}\big(1-z(X_{[1,\alpha-1]}+X_{[\alpha+1,r]})\big)}$$
$$-\frac{x_ix_k}{X_{[k,i-1]}X_{[i,\alpha-1]}\big(1-z(X_{[1,i-1]}+X_{[\alpha+1,r]})\big)}$$
$$+\frac{x_ix_k}{X_{[k,i-1]}X_{[k,\alpha-1]}\big(1-z(X_{[1,k-1]}+X_{[\alpha+1,r]})\big)}\Bigg).$$

For the equation (18.24) it is a similar process, and we have a shuffle product between ordinary generating functions

$$\frac{\partial^2}{\partial u^2}g\Big|_{u=1}=2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\Bigg(\frac{x_ix_k}{X_{[\alpha+1,i]}X_{[\alpha+1,k]}(1-zX_{[\alpha+1,r]})}$$
$$+\frac{x_ix_k}{X_{[k+1,i]}X_{[\alpha+1,i]}(1-zX_{[i+1,r]})}$$
$$-\frac{x_ix_k}{X_{[k+1,i]}X_{[\alpha+1,k]}(1-zX_{[k+1,r]})}\Bigg),$$

and

$$\frac{1}{1-zX_{[1,\alpha-1]}}.$$

Making use of the Laplace transform we can write these first as exponential generating functions

$$2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\Bigg(\frac{x_ix_k}{X_{[\alpha+1,i]}X_{[\alpha+1,k]}}e^{zX_{[\alpha+1,r]}}+\frac{x_ix_k}{X_{[k+1,i]}X_{[\alpha+1,i]}}e^{zX_{[i+1,r]}}$$
$$-\frac{x_ix_k}{X_{[k+1,i]}X_{[\alpha+1,k]}}e^{zX_{[k+1,r]}}\Bigg),$$

and

$$e^{zX_{[1,\alpha-1]}},$$

and then multiply them to get

$$2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\left(\frac{x_ix_k}{X_{[\alpha+1,i]}X_{[\alpha+1,k]}}e^{z(X_{[\alpha+1,r]}+X_{[1,\alpha-1]})}+\frac{x_ix_k}{X_{[k+1,i]}X_{[\alpha+1,i]}}e^{z(X_{[i+1,r]}+X_{[1,\alpha-1]})}\right.$$
$$\left.-\frac{x_ix_k}{X_{[k+1,i]}X_{[\alpha+1,k]}}e^{z(X_{[k+1,r]}+X_{[1,\alpha-1]})}\right).$$

This can be written back in ordinary generating function form as

$$2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\left(\frac{x_ix_k}{X_{[\alpha+1,i]}X_{[\alpha+1,k]}(1-z(X_{[\alpha+1,r]}+X_{[1,\alpha-1]}))}\right.$$
$$+\frac{x_ix_k}{X_{[k+1,i]}X_{[\alpha+1,i]}(1-z(X_{[i+1,r]}+X_{[1,\alpha-1]}))}$$
$$\left.-\frac{x_ix_k}{X_{[k+1,i]}X_{[\alpha+1,k]}(1-z(X_{[k+1,r]}+X_{[1,\alpha-1]}))}\right).$$

Lastly, for equation (18.23)

$$\begin{aligned}
&2\frac{\partial}{\partial u}f\Big|_{u=1} ш \frac{\partial}{\partial u}g\Big|_{u=1}\\
&=2\left[\frac{1}{1-zX_{[1,\alpha-1]}}\sum_{i=1}^{\alpha-1}\frac{zx_i}{1-zX_{[1,i-1]}}\right] ш \left[\frac{1}{1-zX_{[\alpha+1,r]}}\sum_{i=\alpha+1}^{r}\frac{zx_i}{1-zX_{[i+1,r]}}\right]\\
&=2\left[\sum_{i=1}^{\alpha-1}\left(\frac{x_i}{X_{[i,\alpha-1]}}\frac{1}{1-zX_{[1,\alpha-1]}}-\frac{x_i}{X_{[i,\alpha-1]}}\frac{1}{1-zX_{[1,i-1]}}\right)\right]\\
&\quad ш \left[\sum_{i=\alpha+1}^{r}\left(\frac{x_i}{X_{[\alpha+1,i]}}\frac{1}{1-zX_{[\alpha+1,r]}}-\frac{x_i}{X_{[\alpha+1,i]}}\frac{1}{1-zX_{[i+1,r]}}\right)\right]
\end{aligned}$$

which transforms to

$$\begin{aligned}
&2\sum_{i=1}^{\alpha-1}\left(\frac{x_i}{X_{[i,\alpha-1]}}e^{zX_{[1,\alpha-1]}}-\frac{x_i}{X_{[i,\alpha-1]}}e^{zX_{[1,i-1]}}\right)\\
&\qquad\cdot\sum_{j=\alpha+1}^{r}\left(\frac{x_j}{X_{[\alpha+1,j]}}e^{zX_{[\alpha+1,r]}}-\frac{x_j}{X_{[\alpha+1,j]}}e^{zX_{[j+1,r]}}\right)\\
&=2\sum_{i=1}^{\alpha-1}\sum_{j=\alpha+1}^{r}\left(\frac{x_ix_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}}e^{z(X_{[1,\alpha-1]}+X_{[\alpha+1,r]})}-\frac{x_ix_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}}e^{z(X_{[1,\alpha-1]}+X_{[j+1,r]})}\right.\\
&\qquad\left.-\frac{x_ix_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}}e^{z(X_{[1,i-1]}+X_{[\alpha+1,r]})}+\frac{x_ix_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}}e^{z(X_{[1,i-1]}+X_{[j+1,r]})}\right),
\end{aligned}$$

and transforms back to

$$2\sum_{i=1}^{\alpha-1}\sum_{j=\alpha+1}^{r}\left(\frac{x_ix_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}\big(1-z(X_{[1,\alpha-1]}+X_{[\alpha+1,r]})\big)}\right.$$

$$
\begin{aligned}
&- \frac{x_i x_j}{X_{[i,\alpha-1]} X_{[\alpha+1,j]} \big(1 - z(X_{[1,\alpha-1]} + X_{[j+1,r]})\big)} \\
&- \frac{x_i x_j}{X_{[i,\alpha-1]} X_{[\alpha+1,j]} \big(1 - z(X_{[1,i-1]} + X_{[\alpha+1,r]})\big)} \\
&+ \left. \frac{x_i x_j}{X_{[i,\alpha-1]} X_{[\alpha+1,j]} \big(1 - z(X_{[1,i-1]} + X_{[j+1,r]})\big)} \right).
\end{aligned}
$$

We have now looked at all three terms in the shuffle product which give us everything for the cost function except the factor $\frac{zx_\alpha}{1-zX}$ which we include below. This means that we must decompose once again into partial fractions in order to find the coefficient of $z^n$. We get:

$$
\begin{aligned}
&\frac{\partial^2}{\partial u^2} f(z, 1, x_1, \dots, x_r) \\
&= 2 \sum_{i=1}^{\alpha-1} \sum_{k=1}^{i-1} \left( \frac{z x_\alpha x_i x_k}{X_{[i,\alpha-1]} X_{[k,\alpha-1]} (1 - zX)(1 - z(X_{[1,\alpha-1]} + X_{[\alpha+1,r]}))} \right. \\
&\qquad - \frac{z x_\alpha x_i x_k}{X_{[k,i-1]} X_{[i,\alpha-1]} (1 - zX)(1 - z(X_{[1,i-1]} + X_{[\alpha+1,r]}))} \\
&\qquad \left. + \frac{z x_\alpha x_i x_k}{X_{[k,i-1]} X_{[k,\alpha-1]} (1 - zX)(1 - z(X_{[1,k-1]} + X_{[\alpha+1,r]}))} \right) \\
&\quad + 2 \sum_{i=\alpha+1}^{r} \sum_{k=\alpha+1}^{i-1} \left( \frac{z x_\alpha x_i x_k}{X_{[\alpha+1,i]} X_{[\alpha+1,k]} (1 - zX)(1 - z(X_{[\alpha+1,r]} + X_{[1,\alpha-1]}))} \right. \\
&\qquad + \frac{z x_\alpha x_i x_k}{X_{[k+1,i]} X_{[\alpha+1,i]} (1 - zX)(1 - z(X_{[i+1,r]} + X_{[1,\alpha-1]}))} \\
&\qquad \left. - \frac{z x_\alpha x_i x_k}{X_{[k+1,i]} X_{[\alpha+1,k]} (1 - zX)(1 - z(X_{[k+1,r]} + X_{[1,\alpha-1]}))} \right) \\
&\quad + 2 \sum_{i=1}^{\alpha-1} \sum_{j=\alpha+1}^{r} \left( \frac{z x_\alpha x_i x_j}{X_{[i,\alpha-1]} X_{[\alpha+1,j]} (1 - zX)(1 - z(X_{[1,\alpha-1]} + X_{[\alpha+1,r]}))} \right. \\
&\qquad - \frac{z x_\alpha x_i x_j}{X_{[i,\alpha-1]} X_{[\alpha+1,j]} (1 - zX)(1 - z(X_{[1,\alpha-1]} + X_{[j+1,r]}))} \\
&\qquad - \frac{z x_\alpha x_i x_j}{X_{[i,\alpha-1]} X_{[\alpha+1,j]} (1 - zX)(1 - z(X_{[1,i-1]} + X_{[\alpha+1,r]}))} \\
&\qquad \left. + \frac{z x_\alpha x_i x_j}{X_{[i,\alpha-1]} X_{[\alpha+1,j]} (1 - zX)(1 - z(X_{[1,i-1]} + X_{[j+1,r]}))} \right) \\
&= 2 \sum_{i=1}^{\alpha-1} \sum_{k=1}^{i-1} \left( \frac{x_\alpha x_i x_k}{X_{[i,\alpha-1]} X_{[k,\alpha-1]} (X - (X_{[1,\alpha-1]} + X_{[\alpha+1,r]}))(1 - zX)} \right. \\
&\quad + \frac{x_\alpha x_i x_k}{X_{[i,\alpha-1]} X_{[k,\alpha-1]} ((X_{[1,\alpha-1]} + X_{[\alpha+1,r]}) - X)(1 - z(X_{[1,\alpha-1]} + X_{[\alpha+1,r]}))} \\
&\quad - \frac{x_\alpha x_i x_k}{X_{[k,i-1]} X_{[i,\alpha-1]} (X - (X_{[1,i-1]} + X_{[\alpha+1,r]}))(1 - zX)} \\
&\quad - \frac{x_\alpha x_i x_k}{X_{[k,i-1]} X_{[i,\alpha-1]} ((X_{[1,i-1]} + X_{[\alpha+1,r]}) - X)(1 - z(X_{[1,i-1]} + X_{[\alpha+1,r]}))}
\end{aligned}
$$

$$
\begin{aligned}
&\quad + \frac{x_\alpha x_i x_k}{X_{[k,i-1]}X_{[k,\alpha-1]}(X-(X_{[1,k-1]}+X_{[\alpha+1,r]}))(1-zX)} \\
&\quad \left. + \frac{x_\alpha x_i x_k}{(X_{[k,i-1]}X_{[k,\alpha-1]}((X_{[1,k-1]}+X_{[\alpha+1,r]})-X)(1-z(X_{[1,k-1]}+X_{[\alpha+1,r]}))}\right) \\
&+ 2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\left(\frac{x_\alpha x_i x_k}{X_{[\alpha+1,i]}X_{[\alpha+1,k]}(X-(X_{[\alpha+1,r]}+X_{[1,\alpha-1]}))(1-zX))}\right. \\
&\quad + \frac{x_\alpha x_i x_k}{X_{[\alpha+1,i]}X_{[\alpha+1,k]}((X_{[\alpha+1,r]}+X_{[1,\alpha-1]})-X)(1-z(X_{[\alpha+1,r]}+X_{[1,\alpha-1]}))} \\
&\quad + \frac{x_\alpha x_i x_k}{X_{[k+1,i]}X_{[\alpha+1,i]}(X-(X_{[i+1,r]}+X_{[1,\alpha-1]}))(1-zX)} \\
&\quad + \frac{x_\alpha x_i x_k}{X_{[k+1,i]}X_{[\alpha+1,i]}((X_{[i+1,r]}+X_{[1,\alpha-1]})-X)(1-z(X_{[i+1,r]}+X_{[1,\alpha-1]}))} \\
&\quad - \frac{x_\alpha x_i x_k}{X_{[k+1,i]}X_{[\alpha+1,k]}(X-(X_{[k+1,r]}+X_{[1,\alpha-1]}))(1-zX)} \\
&\quad \left. - \frac{x_\alpha x_i x_k}{X_{[k+1,i]}X_{[\alpha+1,k]}((X_{[k+1,r]}+X_{[1,\alpha-1]})-X)(1-z(X_{[k+1,r]}+X_{[1,\alpha-1]}))}\right) \\
&+ 2\sum_{i=1}^{\alpha-1}\sum_{j=\alpha+1}^{r}\left(\frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}(X-(X_{[1,\alpha-1]}+X_{[\alpha+1,r]}))(1-zX)}\right. \\
&\quad + \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}((X_{[1,\alpha-1]}+X_{[\alpha+1,r]})-X)(1-z(X_{[1,\alpha-1]}+X_{[\alpha+1,r]}))} \\
&\quad - \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}(X-(X_{[1,\alpha-1]}+X_{[j+1,r]}))(1-zX)} \\
&\quad - \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}((X_{[1,\alpha-1]}+X_{[j+1,r]})-X)(1-z(X_{[1,\alpha-1]}+X_{[j+1,r]}))} \\
&\quad - \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}(X-(X_{[1,i-1]}+X_{[\alpha+1,r]}))(1-zX)} \\
&\quad - \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}((X_{[1,i-1]}+X_{[\alpha+1,r]})-X)(1-z(X_{[1,i-1]}+X_{[\alpha+1,r]}))} \\
&\quad + \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}(X-(X_{[1,i-1]}+X_{[j+1,r]}))(1-zX)} \\
&\quad \left. + \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}((X_{[1,i-1]}+X_{[j+1,r]})-X)(1-z(X_{[1,i-1]}+X_{[j+1,r]}))}\right) \\
&= 2\sum_{i=1}^{\alpha-1}\sum_{k=1}^{i-1}\left(\frac{x_i x_k}{X_{[i,\alpha-1]}X_{[k,\alpha-1]}(1-zX)}\right. \\
&\qquad\qquad - \frac{x_i x_k}{X_{[i,\alpha-1]}X_{[k,\alpha-1]}(1-z(X_{[1,\alpha-1]}+X_{[\alpha+1,r]}))} \\
&\qquad\qquad - \frac{x_\alpha x_i x_k}{X_{[k,i-1]}X_{[i,\alpha-1]}X_{[i,\alpha]}(1-zX)} \\
&\qquad\qquad + \frac{x_\alpha x_i x_k}{X_{[k,i-1]}X_{[i,\alpha-1]}X_{[i,\alpha]}(1-z(X_{[1,i-1]}+X_{[\alpha+1,r]}))} \\
&\qquad\qquad + \frac{x_\alpha x_i x_k}{X_{[k,i-1]}X_{[k,\alpha-1]}X_{[k,\alpha]}(1-zX)}
\end{aligned}
$$

$$
\begin{aligned}
&\qquad - \frac{x_\alpha x_i x_k}{(X_{[k,i-1]}X_{[k,\alpha-1]}X_{[k,\alpha]}(1 - z(X_{[1,k-1]} + X_{[\alpha+1,r]}))}\Bigg) \\
&+ 2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\Bigg(\frac{x_i x_k}{X_{[\alpha+1,i]}X_{[\alpha+1,k]}(1 - zX))} \\
&\qquad - \frac{x_i x_k}{X_{[\alpha+1,i]}X_{[\alpha+1,k]}(1 - z(X_{[\alpha+1,r]} + X_{[1,\alpha-1]}))} \\
&\qquad + \frac{x_\alpha x_i x_k}{X_{[k+1,i]}X_{[\alpha+1,i]}X_{[\alpha,i]}(1 - zX)} \\
&\qquad - \frac{x_\alpha x_i x_k}{X_{[k+1,i]}X_{[\alpha+1,i]}X_{[\alpha,i]}(1 - z(X_{[i+1,r]} + X_{[1,\alpha-1]}))} \\
&\qquad - \frac{x_\alpha x_i x_k}{X_{[k+1,i]}X_{[\alpha+1,k]}X_{[\alpha,k]}(1 - zX)} \\
&\qquad + \frac{x_\alpha x_i x_k}{X_{[k+1,i]}X_{[\alpha+1,k]}X_{[\alpha,k]}(1 - z(X_{[k+1,r]} + X_{[1,\alpha-1]}))}\Bigg) \\
&+ 2\sum_{i=1}^{\alpha-1}\sum_{j=\alpha+1}^{r}\Bigg(\frac{x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}(1 - zX)} \\
&\qquad - \frac{x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}(1 - z(X_{[1,\alpha-1]} + X_{[\alpha+1,r]}))} \\
&\qquad - \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}X_{[\alpha,j]}(1 - zX)} \\
&\qquad + \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}X_{[\alpha,j]}(1 - z(X_{[1,\alpha-1]} + X_{[j+1,r]}))} \\
&\qquad - \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}X_{[i,\alpha]}(1 - zX)} \\
&\qquad + \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}X_{[i,\alpha]}(1 - z(X_{[1,i-1]} + X_{[\alpha+1,r]}))} \\
&\qquad + \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}X_{[i,j]}(1 - zX)} \\
&\qquad - \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}X_{[i,j]}(1 - z(X_{[1,i-1]} + X_{[j+1,r]}))}\Bigg).
\end{aligned}
$$

We can write the coefficient of $z^n$ explicitly as

$$
\begin{aligned}
&[z^n]\frac{\partial^2}{\partial u^2}f(z,1,x_1,\ldots,x_r) \\
&= 2\sum_{i=1}^{\alpha-1}\sum_{k=1}^{i-1}\Bigg(\frac{x_i x_k}{X_{[i,\alpha-1]}X_{[k,\alpha-1]}}X^n - \frac{x_i x_k}{X_{[i,\alpha-1]}X_{[k,\alpha-1]}}(X_{[1,\alpha-1]} + X_{[\alpha+1,r]})^n \\
&\qquad - \frac{x_\alpha x_i x_k}{X_{[k,i-1]}X_{[i,\alpha-1]}X_{[i,\alpha]}}X^n + \frac{x_\alpha x_i x_k}{X_{[k,i-1]}X_{[i,\alpha-1]}X_{[i,\alpha]}}(X_{[1,i-1]} + X_{[\alpha+1,r]})^n \\
&\qquad + \frac{x_\alpha x_i x_k}{X_{[k,i-1]}X_{[k,\alpha-1]}X_{[k,\alpha]}}X^n - \frac{x_\alpha x_i x_k}{X_{[k,i-1]}X_{[k,\alpha-1]}X_{[k,\alpha]}}(X_{[1,k-1]} + X_{[\alpha+1,r]})^n\Bigg) \\
&+ 2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\Bigg(\frac{x_i x_k}{X_{[\alpha+1,i]}X_{[\alpha+1,k]}}X^n - \frac{x_i x_k}{X_{[\alpha+1,i]}X_{[\alpha+1,k]}}(X_{[\alpha+1,r]} + X_{[1,\alpha-1]})^n
\end{aligned}
$$

$$
\begin{aligned}
&\quad + \frac{x_\alpha x_i x_k}{X_{[k+1,i]}X_{[\alpha+1,i]}X_{[\alpha,i]}}X^n - \frac{x_\alpha x_i x_k}{X_{[k+1,i]}X_{[\alpha+1,i]}X_{[\alpha,i]}}(X_{[i+1,r]}+X_{[1,\alpha-1]})^n \\
&\quad - \frac{x_\alpha x_i x_k}{X_{[k+1,i]}X_{[\alpha+1,k]}X_{[\alpha,k]}}X^n + \frac{x_\alpha x_i x_k}{X_{[k+1,i]}X_{[\alpha+1,k]}X_{[\alpha,k]}}(X_{[k+1,r]}+X_{[1,\alpha-1]})^n\Bigg) \\
&+ 2\sum_{i=1}^{\alpha-1}\sum_{j=\alpha+1}^{r}\Bigg(\frac{x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}}X^n - \frac{x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}}(X_{[1,\alpha-1]}+X_{[\alpha+1,r]})^n \\
&\quad - \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}X_{[\alpha,j]}}X^n + \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}X_{[\alpha,j]}}(X_{[1,\alpha-1]}+X_{[j+1,r]})^n \\
&\quad - \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}X_{[i,\alpha]}}X^n + \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}X_{[i,\alpha]}}(X_{[1,i-1]}+X_{[\alpha+1,r]})^n \\
&\quad + \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}X_{[i,j]}}X^n - \frac{x_\alpha x_i x_j}{X_{[i,\alpha-1]}X_{[\alpha+1,j]}X_{[i,j]}}(X_{[1,i-1]}+X_{[j+1,r]})^n\Bigg).
\end{aligned}
$$

Now the coefficient of $x_1^{n_1}\cdots x_r^{n_r}$ needs to be found. The process is exactly the same as for the expectation (refer to the way (18.12) and (18.13) were dealt with), hence (see equation (18.21))

$$
\begin{aligned}
\mathbb{V}_\alpha^{\mathrm{m}} &= [z^n x_1^{n_1}\cdots x_r^{n_r}]\frac{\partial^2}{\partial u^2}f(z,1,x_1,\ldots,x_r) + [z^n x_1^{n_1}\cdots x_r^{n_r}]\frac{\partial}{\partial u}f(z,1,x_1,\ldots,x_r) \\
&\quad - \left([z^n x_1^{n_1}\cdots x_r^{n_r}]\frac{\partial}{\partial u}f(z,1,x_1,\ldots,x_r)\right)^2 \\
&= 2\sum_{i=1}^{\alpha-1}\sum_{k=1}^{i-1}\left(\frac{n_i n_k}{N_{[i,\alpha-1]}N_{[k,\alpha-1]}} - \frac{n_\alpha n_i n_k}{N_{[k,i-1]}N_{[i,\alpha-1]}N_{[i,\alpha]}} + \frac{n_\alpha n_i n_k}{N_{[k,i-1]}N_{[k,\alpha-1]}N_{[k,\alpha]}}\right) \\
&\quad + 2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\left(\frac{n_i n_k}{N_{[\alpha+1,i]}N_{[\alpha+1,k]}} + \frac{n_\alpha n_i n_k}{N_{[k+1,i]}N_{[\alpha+1,i]}N_{[\alpha,i]}} - \frac{n_\alpha n_i n_k}{N_{[k+1,i]}N_{[\alpha+1,k]}N_{[\alpha,k]}}\right) \\
&\quad + 2\sum_{i=1}^{\alpha-1}\sum_{j=\alpha+1}^{r}\Bigg(\frac{n_i n_j}{N_{[i,\alpha-1]}N_{[\alpha+1,j]}} - \frac{n_\alpha n_i n_j}{N_{[i,\alpha-1]}N_{[\alpha+1,j]}N_{[\alpha,j]}} \\
&\qquad\qquad - \frac{n_\alpha n_i n_j}{N_{[i,\alpha-1]}N_{[\alpha+1,j]}N_{[i,\alpha]}} + \frac{n_\alpha n_i n_j}{N_{[i,\alpha-1]}N_{[\alpha+1,j]}N_{[i,j]}}\Bigg) \\
&\quad + \sum_{i=1}^{\alpha-1}\frac{n_i}{N_{[i,\alpha]}} + \sum_{i=\alpha+1}^{r}\frac{n_i}{N_{[\alpha,i]}} \\
&\quad - \left(\sum_{i=1}^{\alpha-1}\frac{n_i}{N_{[i,\alpha]}}\right)^2 - 2\sum_{i=1}^{\alpha-1}\frac{n_i}{N_{[i,\alpha]}}\sum_{i=\alpha+1}^{r}\frac{n_i}{N_{[\alpha,i]}} - \left(\sum_{i=\alpha+1}^{r}\frac{n_i}{N_{[\alpha,i]}}\right)^2. \qquad (18.25)
\end{aligned}
$$

This is an explicit form of the variance for the cost of finding an arbitrary key $\alpha$ in a binary search tree without equal keys formed from a sequence with equal keys. The proof of Theorem 18.3 is thus complete. ■

## 18.6 Variance – probability model

We have (from equations (18.2) and (18.17))

$$
\begin{aligned}
f(z,u,x_1,\ldots,x_r) &= \left[\prod_{i=1}^{\alpha-1}\left(1+\frac{zux_i}{1-zX_{[1,i]}}\right) ш \prod_{j=\alpha+1}^{r}\left(1+\frac{zux_j}{1-zX_{[j,r]}}\right)\right]\cdot\frac{zx_\alpha}{1-zX}\\
&= \sum_{w\in A^*} z^{|w|}u^{a(w)}x_1^{|w|_1}x_2^{|w|_2}\cdots x_r^{|w|_r},
\end{aligned}
$$

where $A=\{1,\ldots,r\}$ etc. as before. For the variance, we need

$$
\begin{aligned}
\mathbb{V}_\alpha^{\mathrm{p}} = {} & [z^n]\frac{\partial^2}{\partial^2 u}f(z,u,p_1,\ldots,p_r)\Big|_{u=1} + [z^n]\frac{\partial}{\partial u}f(z,u,p_1,\ldots,p_r)\Big|_{u=1}\\
& - \left([z^n]\frac{\partial}{\partial u}f(z,u,p_1,\ldots,p_r)\Big|_{u=1}\right)^2,
\end{aligned}
\tag{18.26}
$$

where $p_i$ is the probability of letter $i$ appearing in the sequence. Now, for $P=P_{[1,r]}$,

$$
\begin{aligned}
&[z^n]\frac{\partial^2}{\partial u^2}f(z,1,p_1,\ldots,p_r)\\
&= 2\sum_{i=1}^{\alpha-1}\sum_{k=1}^{i-1}\Bigg(\frac{p_ip_k}{P_{[i,\alpha-1]}P_{[k,\alpha-1]}}P^n - \frac{p_ip_k}{P_{[i,\alpha-1]}P_{[k,\alpha-1]}}(P_{[1,\alpha-1]}+P_{[\alpha+1,r]})^n\\
&\qquad - \frac{p_\alpha p_ip_k}{P_{[k,i-1]}P_{[i,\alpha-1]}P_{[i,\alpha]}}P^n + \frac{p_\alpha p_ip_k}{P_{[k,i-1]}P_{[i,\alpha-1]}P_{[i,\alpha]}}(P_{[1,i-1]}+P_{[\alpha+1,r]})^n\\
&\qquad + \frac{p_\alpha p_ip_k}{P_{[k,i-1]}P_{[k,\alpha-1]}P_{[k,\alpha]}}P^n - \frac{p_\alpha p_ip_k}{P_{[k,i-1]}P_{[k,\alpha-1]}P_{[k,\alpha]}}(P_{[1,k-1]}+P_{[\alpha+1,r]})^n\Bigg)\\
&+ 2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\Bigg(\frac{p_ip_k}{P_{[\alpha+1,i]}P_{[\alpha+1,k]}}P^n - \frac{p_ip_k}{P_{[\alpha+1,i]}P_{[\alpha+1,k]}}(P_{[\alpha+1,r]}+P_{[1,\alpha-1]})^n\\
&\qquad + \frac{p_\alpha p_ip_k}{P_{[k+1,i]}P_{[\alpha+1,i]}P_{[\alpha,i]}}P^n - \frac{p_\alpha p_ip_k}{P_{[k+1,i]}P_{[\alpha+1,i]}P_{[\alpha,i]}}(P_{[i+1,r]}+P_{[1,\alpha-1]})^n\\
&\qquad - \frac{p_\alpha p_ip_k}{P_{[k+1,i]}P_{[\alpha+1,k]}P_{[\alpha,k]}}P^n + \frac{p_\alpha p_ip_k}{P_{[k+1,i]}P_{[\alpha+1,k]}P_{[\alpha,k]}}(P_{[k+1,r]}+P_{[1,\alpha-1]})^n\Bigg)\\
&+ 2\sum_{i=1}^{\alpha-1}\sum_{j=\alpha+1}^{r}\Bigg(\frac{p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}}P^n - \frac{p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}}(P_{[1,\alpha-1]}+P_{[\alpha+1,r]})^n\\
&\qquad - \frac{p_\alpha p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[\alpha,j]}}P^n + \frac{p_\alpha p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[\alpha,j]}}(P_{[1,\alpha-1]}+P_{[j+1,r]})^n\\
&\qquad - \frac{p_\alpha p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[i,\alpha]}}P^n + \frac{p_\alpha p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[i,\alpha]}}(P_{[1,i-1]}+P_{[\alpha+1,r]})^n\\
&\qquad + \frac{p_\alpha p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[i,j]}}P^n - \frac{p_\alpha p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[i,j]}}(P_{[1,i-1]}+P_{[j+1,r]})^n\Bigg).
\end{aligned}
$$

Now recall that $P=P_{[1,r]}=1$, and that $P_{[i,j]}<1$ for all other allowable values of $i$ and $j$. Thus, every second term above can be written in the form $b\cdot a^n$ where

$a < 1$, and $b$ is independent of $n$. These terms will consequently not contribute when taking the limit as $n \to \infty$. Then

$$\begin{aligned}
&[z^n]\frac{\partial^2}{\partial u^2} f(z,1,p_1,\ldots,p_r)\\
&\sim 2\sum_{i=1}^{\alpha-1}\sum_{k=1}^{i-1}\left(\frac{p_ip_k}{P_{[i,\alpha-1]}P_{[k,\alpha-1]}} - \frac{p_\alpha p_ip_k}{P_{[k,i-1]}P_{[i,\alpha-1]}P_{[i,\alpha]}} + \frac{p_\alpha p_ip_k}{P_{[k,i-1]}P_{[k,\alpha-1]}P_{[k,\alpha]}}\right)\\
&\quad + 2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\left(\frac{p_ip_k}{P_{[\alpha+1,i]}P_{[\alpha+1,k]}} + \frac{p_\alpha p_ip_k}{P_{[k+1,i]}P_{[\alpha+1,i]}P_{[\alpha,i]}} - \frac{p_\alpha p_ip_k}{P_{[k+1,i]}P_{[\alpha+1,k]}P_{[\alpha,k]}}\right)\\
&\quad + 2\sum_{i=1}^{\alpha-1}\sum_{j=\alpha+1}^{r}\left(\frac{p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}} - \frac{p_\alpha p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[\alpha,j]}}\right.\\
&\qquad\qquad\left. - \frac{p_\alpha p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[i,\alpha]}} + \frac{p_\alpha p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[i,j]}}\right).
\end{aligned}$$

For the full variance expression, substituting (18.18) into (18.26) gives

$$\begin{aligned}
\mathbb{V}_\alpha^{\mathrm{p}} &\sim 2\sum_{i=1}^{\alpha-1}\sum_{k=1}^{i-1}\left(\frac{p_ip_k}{P_{[i,\alpha-1]}P_{[k,\alpha-1]}} - \frac{p_\alpha p_ip_k}{P_{[k,i-1]}P_{[i,\alpha-1]}P_{[i,\alpha]}} + \frac{p_\alpha p_ip_k}{P_{[k,i-1]}P_{[k,\alpha-1]}P_{[k,\alpha]}}\right)\\
&\quad + 2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\left(\frac{p_ip_k}{P_{[\alpha+1,i]}P_{[\alpha+1,k]}} + \frac{p_\alpha p_ip_k}{P_{[k+1,i]}P_{[\alpha+1,i]}P_{[\alpha,i]}} - \frac{p_\alpha p_ip_k}{P_{[k+1,i]}P_{[\alpha+1,k]}P_{[\alpha,k]}}\right)\\
&\quad + 2\sum_{i=1}^{\alpha-1}\sum_{j=\alpha+1}^{r}\left(\frac{p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}} - \frac{p_\alpha p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[\alpha,j]}}\right.\\
&\qquad\qquad\left. - \frac{p_\alpha p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[i,\alpha]}} + \frac{p_\alpha p_ip_j}{P_{[i,\alpha-1]}P_{[\alpha+1,j]}P_{[i,j]}}\right)\\
&\quad + \sum_{i=1}^{\alpha-1}\frac{p_i}{P_{[i,\alpha]}} + \sum_{i=\alpha+1}^{r}\frac{p_i}{P_{[\alpha,i]}}\\
&\quad - \left(\sum_{i=1}^{\alpha-1}\frac{p_i}{P_{[i,\alpha]}}\right)^2 - 2\sum_{i=1}^{\alpha-1}\frac{p_i}{P_{[i,\alpha]}}\sum_{i=\alpha+1}^{r}\frac{p_i}{P_{[\alpha,i]}} - \left(\sum_{i=\alpha+1}^{r}\frac{p_i}{P_{[\alpha,i]}}\right)^2. \qquad (18.27)
\end{aligned}$$

Theorem 18.4 is thus proved. ■

## 18.7 Variance – all keys distinct

We now insist that only one of each letter occurs (i.e., $n_i = 1, \ \forall\ i = 1, \ldots, r$) and then compare to the result for the case where there are no equal keys. In [36, page 249], this variance is given as asymptotic to $2H_n$. We have (where 'd' stands for distinct)

$$\begin{aligned}\mathbb{V}_\alpha^{\mathrm{d}} = 2\sum_{i=1}^{\alpha-1}\sum_{k=1}^{i-1}\Bigg(\frac{1}{(\alpha-i)(\alpha-k)} - \frac{1}{(i-k)(\alpha-i)(\alpha-i+1)} \\ + \frac{1}{(i-k)(\alpha-k)(\alpha-k+1)}\Bigg) \qquad (18.28)\end{aligned}$$

$$\begin{aligned}+ 2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\Bigg(\frac{1}{(i-\alpha)(k-\alpha)} + \frac{1}{(i-k)(i-\alpha)(i-\alpha+1)} \\ - \frac{1}{(i-k)(k-\alpha)(k-\alpha+1)}\Bigg) \qquad (18.29)\end{aligned}$$

$$\begin{aligned}+ 2\sum_{i=1}^{\alpha-1}\sum_{j=\alpha+1}^{r}\Bigg(\frac{1}{(\alpha-i)(j-\alpha)} - \frac{1}{(\alpha-i)(j-\alpha)(j-\alpha+1)} \qquad (18.30) \\ - \frac{1}{(\alpha-i)(j-\alpha)(\alpha-i+1)} + \frac{1}{(\alpha-i)(j-\alpha)(j-i+1)}\Bigg)\end{aligned}$$

$$+ \sum_{i=1}^{\alpha-1}\frac{1}{(\alpha-i+1)} + \sum_{j=\alpha+1}^{r}\frac{1}{(j-\alpha+1)} \qquad (18.31)$$

$$\begin{aligned}- \left(\sum_{i=1}^{\alpha-1}\frac{1}{(\alpha-i+1)}\right)^2 - 2\sum_{i=1}^{\alpha-1}\frac{1}{(\alpha-i+1)}\sum_{j=\alpha+1}^{r}\frac{1}{(j-\alpha+1)} \\ - \left(\sum_{j=\alpha+1}^{r}\frac{1}{(j-\alpha+1)}\right)^2. \qquad (18.32)\end{aligned}$$

We would like to write this more simply, and for that we use harmonic numbers. From [12, page 280], we have that

$$\sum_{1\le j\le k\le n}\frac{1}{jk} = \frac{1}{2}\big(H_n^2 + H_n^{(2)}\big),$$

where $H_n := \sum_{k=1}^{n}\frac{1}{k}$, and $H_n^{(2)} := \sum_{k=1}^{n}\frac{1}{k^2}$. This is because the left-hand side of the equation represents all possible combinations of products in the denominator of the numbers $1,\dots,n$, each product appearing exactly once. As for the right-hand side, squaring a harmonic number will also give such products, but the terms where $j \neq k$ appear twice. For this reason, we must add the second-order harmonic number (representing the terms where $j = k$). Thus we have every term appearing exactly twice and the factor of a half reduces this to what we want. By a similar explanation we can express

$$\sum_{1\le j<k\le n}\frac{1}{jk} = \frac{1}{2}\big(H_n^2 - H_n^{(2)}\big), \qquad (18.33)$$

which is the same idea but without the terms where $j = k$. The strict inequality on the left-hand side of the equation balances the negative sign on the right-hand side. In this case we subtract the second-order harmonic number, which gets rid of the terms where $j = k$ in $H_n^2$ (each appears once). This leaves us with all the terms where $j \neq k$, but each term appears twice, hence the factor of a half.

We use (18.33) to simplify the first line of the variance in the distinct case. We can rewrite (18.28) as:

$$\begin{aligned}
&2\sum_{i=1}^{\alpha-1}\sum_{k=1}^{i-1}\Bigg(\frac{1}{(\alpha-i)(\alpha-k)} - \frac{1}{(i-k)(\alpha-i)(\alpha-i+1)} \\
&\qquad\qquad + \frac{1}{(i-k)(\alpha-k)(\alpha-k+1)}\Bigg) \\
&= 2\sum_{i=1}^{\alpha-1}\sum_{k=1}^{i-1}\frac{1}{(\alpha-i+1)(\alpha-k+1)} \\
&= 2\sum_{2\leq j<l\leq\alpha}\frac{1}{jl} \\
&= \big(H_\alpha - 1\big)^2 - \big(H_\alpha^{(2)} - 1\big).
\end{aligned}$$

This represents all terms of the form $\frac{1}{jl}$ where $j$ and $l$ run from 2 to $\alpha$, and $j \neq l$. Each term will appear twice. Using the same idea as in [12], we can think of this as the square of the harmonic number $H_\alpha$, without the 1. Again we have a correction factor and this time we want to exclude terms where $j = l$, so we subtract the second-order harmonic number. The factor of two in the second-last line is cancelled with the factor of a half introduced after squaring $H_\alpha$.

The expression (18.29) can be written as:

$$\begin{aligned}
&2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\Bigg(\frac{1}{(i-\alpha)(k-\alpha)} + \frac{1}{(i-k)(i-\alpha)(i-\alpha+1)} \\
&\qquad\qquad - \frac{1}{(i-k)(k-\alpha)(k-\alpha+1)}\Bigg) \\
&= 2\sum_{i=\alpha+1}^{r}\sum_{k=\alpha+1}^{i-1}\frac{1}{(i-\alpha+1)(k-\alpha+1)} \\
&= \big(H_{r-\alpha+1} - 1\big)^2 - \big(H_{r-\alpha+1}^{(2)} - 1\big).
\end{aligned}$$

Again, choosing a few values shows that the second-last line depicts twice all non-equal combinations from 2 to $r - \alpha + 1$, as is shown in the last line by squaring the harmonic number less one, and removing those terms we do not want which are where the factor is repeated.

Merging the terms and applying a partial fraction decomposition to (18.30), produces

$$\begin{aligned}
&2\sum_{i=1}^{\alpha-1}\sum_{j=\alpha+1}^{r}\Bigg(\frac{1}{(\alpha-i)(j-\alpha)}-\frac{1}{(\alpha-i)(j-\alpha)(j-\alpha+1)}\\
&\qquad\qquad -\frac{1}{(\alpha-i)(j-\alpha)(\alpha-i+1)}+\frac{1}{(\alpha-i)(j-\alpha)(j-i+1)}\Bigg)\\
&=2\sum_{i=1}^{\alpha-1}\sum_{j=\alpha+1}^{r}\left(\frac{1}{(\alpha-i+1)(j-i+1)}+\frac{1}{(j-\alpha+1)(j-i+1)}\right)\\
&=2\sum_{k=2}^{\alpha}\frac{1}{k}\big(H_{r-\alpha+k}-H_k\big)+2\sum_{k=2}^{r-(\alpha-1)}\frac{1}{k}\big(H_{k+\alpha-1}-H_k\big)\\
&=2\sum_{k=2}^{\alpha}\frac{1}{k}H_{r-\alpha+k}+2\sum_{k=2}^{r-(\alpha-1)}\frac{1}{k}H_{k+\alpha-1}-2\sum_{k=2}^{\alpha}\frac{1}{k}H_k-2\sum_{k=2}^{r-(\alpha-1)}\frac{1}{k}H_k. \qquad (18.34)
\end{aligned}$$

To deal with this, we refer to two identities: In [12, page 280] we find the identity

$$\sum_{k=1}^{n}\frac{1}{k}H_k=\frac{1}{2}\big(H_n^2+H_n^{(2)}\big), \qquad (18.35)$$

and the reciprocity law in Lemma 2 in [19, page 116], namely

$$\begin{aligned}
\sum_{k=1}^{j}\frac{H_{n+k-j}}{k}+\sum_{k=1}^{n+1-j}\frac{H_{j+k-1}}{k}&=\frac{1}{2}\big(H_j^2+H_j^{(2)}\big)+\frac{1}{2}\big(H_{n+1-j}^2+H_{n+1-j}^{(2)}\big)\\
&\quad+H_jH_{n+1-j}+\frac{1}{j(n+1-j)}\\
&\quad+\frac{n+1}{j(n+1-j)}\big(H_n-H_j-H_{n+1-j}\big),
\end{aligned}$$

takes care of the other two terms. Replacing $n$ and $j$ by $r$ and $\alpha$ respectively; subtracting the relevant $k=1$ term and multiplying everything by 2 simplifies (18.34) (and thus (18.30)) to:

$$\begin{aligned}
&2\sum_{k=2}^{\alpha}\frac{1}{k}H_{r-\alpha+k}+2\sum_{k=2}^{r-(\alpha-1)}\frac{1}{k}H_{k+\alpha-1}-2\sum_{k=2}^{\alpha}\frac{1}{k}H_k-2\sum_{k=2}^{r-(\alpha-1)}\frac{1}{k}H_k\\
&=2\Bigg[\frac{1}{2}\Big(H_\alpha^2+H_\alpha^{(2)}\Big)+\frac{1}{2}\Big(H_{r+1-\alpha}^2+H_{r+1-\alpha}^{(2)}\Big)\\
&\quad+H_\alpha H_{r+1-\alpha}+\frac{1}{\alpha(r+1-\alpha)}\\
&\quad+\frac{r+1}{\alpha(r+1-\alpha)}\big(H_r-H_\alpha-H_{r+1-\alpha}\big)-H_{r-\alpha+1}-H_\alpha\Bigg]\\
&\quad-2\left[\frac{1}{2}\Big(H_\alpha^2+H_\alpha^{(2)}\Big)-1\right]-2\left[\frac{1}{2}\Big(H_{r-(\alpha-1)}^2+H_{r-(\alpha-1)}^{(2)}\Big)-1\right]
\end{aligned}$$

$$
\begin{aligned}
&= H_\alpha^2 + H_\alpha^{(2)} + H_{r+1-\alpha}^2 + H_{r+1-\alpha}^{(2)} + 2H_\alpha H_{r+1-\alpha} + 2\frac{1}{\alpha(r+1-\alpha)} \\
&\quad + 2\frac{r+1}{\alpha(r+1-\alpha)}\big(H_r - H_\alpha - H_{r+1-\alpha}\big) - 2H_{r-\alpha+1} - 2H_\alpha \\
&\quad - H_\alpha^2 - H_\alpha^{(2)} + 2 - H_{r-(\alpha-1)}^2 - H_{r-(\alpha-1)}^{(2)} + 2 \\
&= 2H_\alpha H_{r+1-\alpha} - 2H_{r-\alpha+1} - 2H_\alpha + 4 + 2\frac{1}{\alpha(r+1-\alpha)} \\
&\quad + 2\frac{r+1}{\alpha(r+1-\alpha)}\big(H_r - H_\alpha - H_{r+1-\alpha}\big).
\end{aligned}
$$

The expectation (18.31) and negative expectation squared (18.32) are easier. We get

$$
\begin{aligned}
&\sum_{i=1}^{\alpha-1} \frac{1}{(\alpha-i+1)} + \sum_{j=\alpha+1}^{r} \frac{1}{(j-\alpha+1)} \\
&= \sum_{i=2}^{\alpha} \frac{1}{i} + \sum_{j=2}^{r-\alpha+1} \frac{1}{j} \\
&= H_\alpha - 1 + H_{r-\alpha+1} - 1,
\end{aligned}
$$

and

$$
\begin{aligned}
&-\Bigg(\sum_{i=1}^{\alpha-1} \frac{1}{(\alpha-i+1)}\Bigg)^2 - 2\sum_{i=1}^{\alpha-1} \frac{1}{(\alpha-i+1)} \sum_{j=\alpha+1}^{r} \frac{1}{(j-\alpha+1)} - \Bigg(\sum_{j=\alpha+1}^{r} \frac{1}{(j-\alpha+1)}\Bigg)^2 \\
&= -\Bigg(\sum_{i=2}^{\alpha} \frac{1}{i}\Bigg)^2 - 2\sum_{i=2}^{\alpha} \frac{1}{i} \sum_{j=2}^{r-\alpha+1} \frac{1}{j} - \Bigg(\sum_{j=2}^{r-\alpha+1} \frac{1}{j}\Bigg)^2 \\
&= -(H_\alpha - 1)^2 - 2(H_\alpha - 1)(H_{r-\alpha+1} - 1) - (H_{r-\alpha+1} - 1)^2,
\end{aligned}
$$

respectively.

We can put all of these results together to get:

$$
\begin{aligned}
\mathbb{V}_\alpha^{\mathrm{d}} &= \big(H_\alpha - 1\big)^2 - \big(H_\alpha^{(2)} - 1\big) \\
&\quad + \big(H_{r-\alpha+1} - 1\big)^2 - \big(H_{r-\alpha+1}^{(2)} - 1\big) \\
&\quad + 2H_\alpha H_{r+1-\alpha} - 2H_{r-\alpha+1} - 2H_\alpha + 4 + 2\frac{1}{\alpha(r+1-\alpha)} \\
&\quad + 2\frac{r+1}{\alpha(r+1-\alpha)}\big(H_r - H_\alpha - H_{r+1-\alpha}\big) \\
&\quad + H_\alpha - 1 + H_{r-\alpha+1} - 1 \\
&\quad - (H_\alpha - 1)^2 - 2(H_\alpha - 1)(H_{r-\alpha+1} - 1) - (H_{r-\alpha+1} - 1)^2 \\
&= H_\alpha^2 - 2H_\alpha + 1 - H_\alpha^{(2)} + 1 \\
&\quad + H_{r-\alpha+1}^2 - 2H_{r-\alpha+1} + 1 - H_{r-\alpha+1}^{(2)} + 1
\end{aligned}
$$

$$
\begin{aligned}
&+ 2H_\alpha H_{r+1-\alpha} - 2H_{r-\alpha+1} - 2H_\alpha + 4 + 2\frac{1}{\alpha(r+1-\alpha)} \\
&+ 2\frac{r+1}{\alpha(r+1-\alpha)}\big(H_r - H_\alpha - H_{r+1-\alpha}\big) \\
&+ H_\alpha + H_{r-\alpha+1} - 2 \\
&- H_\alpha^2 + 2H_\alpha - 1 - 2H_\alpha H_{r-\alpha+1} + 2H_\alpha + 2H_{r-\alpha+1} - 2 \\
&- H_{r-\alpha+1}^2 + 2H_{r-\alpha+1} - 1 \\
= {}& 2 + H_\alpha + H_{r-\alpha+1} - H_\alpha^{(2)} - H_{r-\alpha+1}^{(2)} + \frac{2}{\alpha(r+1-\alpha)} \\
&+ \frac{2(r+1)}{\alpha(r+1-\alpha)}\big(H_r - H_\alpha - H_{r+1-\alpha}\big).
\end{aligned}
$$

Recall that from [36, page 249], the variance on average should grow like $2H_r$ as $r \to \infty$ (recall $r = n$ for the distinct case). To get the average variance, we sum on $\alpha$ and divide by $r$. Note the symmetry between $\alpha$ and $r - \alpha + 1$ which aids simplification.

$$
\begin{aligned}
\sum_{\alpha=1}^{r} \frac{1}{r}\mathbb{V}_\alpha^{\mathrm{d}} &= \sum_{\alpha=1}^{r} \frac{1}{r}\Bigg[2 + H_\alpha + H_{r-\alpha+1} - H_\alpha^{(2)} - H_{r-\alpha+1}^{(2)} + \frac{2}{\alpha(r+1-\alpha)} \\
&\qquad + \frac{2(r+1)}{\alpha(r+1-\alpha)}\big(H_r - H_\alpha - H_{r+1-\alpha}\big)\Bigg] \\
&= 2 + \frac{2}{r}\sum_{\alpha=1}^{r} H_\alpha - \frac{2}{r}\sum_{\alpha=1}^{r} H_\alpha^{(2)} + \frac{2}{r}\sum_{\alpha=1}^{r}\frac{1}{\alpha(r+1-\alpha)} \\
&\quad + \frac{2(r+1)}{r}\sum_{\alpha=1}^{r}\frac{1}{\alpha(r-\alpha+1)}\big(H_r - H_\alpha - H_{r+1-\alpha}\big).
\end{aligned}
$$

We now need (see [12, 23])

$$
\sum_{k=1}^{n} H_k = (n+1)H_n - n,
$$

$$
\sum_{k=1}^{n} H_k^{(2)} = (n+1)H_n^{(2)} - H_n,
$$

and

$$
\sum_{k=1}^{n} \frac{1}{n-k+1} H_k = H_{n+1}^2 - H_{n+1}^{(2)},
$$

as well as equation (18.35). Also, we note that by partial fractions:

$$
\frac{1}{\alpha(r-\alpha+1)} = \frac{1}{r+1}\left(\frac{1}{\alpha} + \frac{1}{r-\alpha+1}\right).
$$

Thus we can write

$$
\sum_{\alpha=1}^{r} \frac{1}{r}\mathbb{V}_\alpha^{\mathrm{d}} = 2 + \frac{2}{r}\big((r+1)H_r - r\big) - \frac{2}{r}\big((r+1)H_r^{(2)} - H_r\big)
$$

$$
\begin{aligned}
&+\frac{2}{r(r+1)}\sum_{\alpha=1}^{r}\left(\frac{1}{\alpha}+\frac{1}{r-\alpha+1}\right)\\
&+\frac{2}{r}\sum_{\alpha=1}^{r}\left(\frac{1}{\alpha}+\frac{1}{r-\alpha+1}\right)\left(H_r-H_\alpha-H_{r+1-\alpha}\right)\\
=\;&2+\frac{2}{r}(r+1)H_r-\frac{2}{r}r-\frac{2}{r}(r+1)H_r^{(2)}+\frac{2}{r}H_r\\
&+\frac{2}{r(r+1)}\sum_{\alpha=1}^{r}\frac{1}{\alpha}+\frac{2}{r(r+1)}\sum_{\alpha=1}^{r}\frac{1}{r-\alpha+1}\\
&+\frac{2}{r}\sum_{\alpha=1}^{r}\frac{1}{\alpha}H_r-\frac{2}{r}\sum_{\alpha=1}^{r}\frac{1}{\alpha}H_\alpha-\frac{2}{r}\sum_{\alpha=1}^{r}\frac{1}{\alpha}H_{r+1-\alpha}\\
&+\frac{2}{r}\sum_{\alpha=1}^{r}\frac{1}{r-\alpha+1}H_r-\frac{2}{r}\sum_{\alpha=1}^{r}\frac{1}{r-\alpha+1}H_\alpha\\
&-\frac{2}{r}\sum_{\alpha=1}^{r}\frac{1}{r-\alpha+1}H_{r+1-\alpha}\\
=\;&2+2H_r+\frac{2}{r}H_r-2-\frac{2}{r}H_r^{(2)}r-\frac{2}{r}H_r^{(2)}+\frac{2}{r}H_r\\
&+\frac{2}{r(r+1)}H_r+\frac{2}{r(r+1)}H_r\\
&+\frac{2}{r}H_r^2-\frac{1}{r}\left(H_r^2+H_r^{(2)}\right)-\frac{2}{r}\sum_{\alpha=1}^{r}\frac{1}{r-\alpha+1}H_\alpha\\
&+\frac{2}{r}H_r^2-\frac{2}{r}\left(H_{r+1}^2-H_{r+1}^{(2)}\right)\\
&-\frac{1}{r}\left(H_r^2+H_r^{(2)}\right)\\
=\;&\left(2+\frac{4}{r}+\frac{4}{r(r+1)}\right)H_r-\left(2+\frac{4}{r}\right)H_r^{(2)}\\
&+\frac{2}{r}H_r^2-\frac{4}{r}H_{r+1}^2+\frac{4}{r}H_{r+1}^{(2)}.
\end{aligned}
$$

As $r\to\infty$, the dominant term is the expected $2H_r$. We can use the asymptotics $H_n\sim\log n$ and $H_n^{(2)}\sim\frac{\pi^2}{6}$ from [36] to see this. This expression thus grows like the logarithmic function, and a plot on Mathematica confirms this:

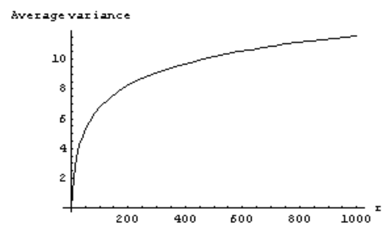


Figure 18.3: Plot of the average variance.

# Chapter 19

# Conclusion

This third and final part of the thesis was concerned with sequences drawn from finite alphabets. If the number of times each letter appeared in the sequence was known, then an exact formula was obtained in terms of these values which make up a multiset. If, however, this was not known and instead a sequence was formed from an alphabet with fixed probabilities attached to each letter, then approximate results were obtained (which become more accurate as the length of the sequence increases).

The average depth of a certain key or node was obtained, by first examining the left-going depth of the smallest letter in a sequence and then examining the right-going depth of the largest letter in a sequence. Then these ideas were merged to get an average depth of

$$\mathbb{E}^{\mathrm{m}}_{\alpha} = \sum_{i=1}^{\alpha-1} \frac{n_i}{N_{[i,\alpha]}} + \sum_{i=\alpha+1}^{r} \frac{n_i}{N_{[\alpha,i]}},$$

using the multiset model, and

$$\mathbb{E}^{\mathrm{p}}_{\alpha} \sim \sum_{i=1}^{\alpha-1} \frac{p_i}{P_{[i,\alpha]}} + \sum_{i=\alpha+1}^{r} \frac{p_i}{P_{[\alpha,i]}},$$

as $n \to \infty$, with the probability model.

For each of the three parameters, the results for the two different models were found to be the same asymptotically as the length of the word went to infinity, as explained in Section 16.3.

# Part IV

# Appendix

# Appendix A

# Fourier series for distinct values

The Fourier series (3.19) we want to simplify can be written as $\delta_V(x) = \sum_{k\neq 0} a_k e^{2\pi \boldsymbol{i} kx}$ where

$$a_k = \frac{2}{L}\Gamma(-\chi_k)\left[\frac{\psi(-\chi_k)+\gamma}{L} - g(\chi_k)\right] - \frac{1}{L^2}\sum_{j\neq 0,\neq k}\Gamma(-\chi_j)\Gamma(-\chi_{k-j}),$$

with $g(x) = -\sum_{l\geq 1}\binom{x}{l}\frac{1}{Q^l-1}$. We consult [32] to do this, and start by using the formula $\Gamma(-x+l)(-1)^l = (x-l+1)\cdots(x-1)x\Gamma(-x)$ to rewrite

$$\Gamma(-\chi_k)g(\chi_k) = -\sum_{l\geq 1}\frac{(-1)^l\Gamma(l-\chi_k)}{l!Q^l-1},$$

so that we have

$$a_k = \frac{2}{L}\Gamma(-\chi_k)\Big[\frac{\psi(-\chi_k)+\gamma}{L}\Big] + \frac{2}{L}\sum_{l\geq 1}\frac{(-1)^l\Gamma(l-\chi_k)}{l!Q^l-1} - \frac{1}{L^2}\sum_{j\neq 0,\neq k}\Gamma(-\chi_j)\Gamma(-\chi_{k-j}).$$

We now consider the function [17]

$$F(z) = L\frac{\Gamma(z)\Gamma(-\chi_k-z)}{e^{Lz}-1},$$

with integral

$$I_1 = \frac{1}{2\pi \boldsymbol{i}}\int_{\frac{1}{2}-\boldsymbol{i}\infty}^{\frac{1}{2}+\boldsymbol{i}\infty} F(z)dz.$$

This function is chosen because of the residues produced when the contour of integration is shifted. We evaluate this integral twice, by shifting the contour first left and then right. We start by shifting the line left to $\Re(z) = -\frac{1}{2}$. Simple poles occur at $z = -\chi_j$ for all $j \in \mathbb{Z}\backslash\{0\}$, with a double pole at $z = 0$.

$$\text{Res}(F, 0) = -\gamma\Gamma(-\chi_k) - \frac{L}{2}\Gamma(-\chi_k) - \Gamma(-\chi_k)\psi(-\chi_k),$$

$$\operatorname{Res}(F, -\chi_k) = -\Gamma(-\chi_k)\psi(-\chi_k) + \frac{L}{2}\Gamma(-\chi_k) - \gamma\Gamma(-\chi_k),$$
$$\operatorname{Res}(F, -\chi_j) = \Gamma(-\chi_k)\Gamma(-\chi_k + \chi_j), \qquad \forall j \neq 0, \neq k.$$

Thus

$$I_1 = \frac{1}{2\pi \boldsymbol{i}} \int_{-\frac{1}{2}-\boldsymbol{i}\infty}^{-\frac{1}{2}+\boldsymbol{i}\infty} F(z)dz - 2\Gamma(-\chi_k)(\gamma + \psi(-\chi_k) + \sum_{j\neq 0, \neq k} \Gamma(-\chi_k)\Gamma(-\chi_k + \chi_j),$$

and we use $\frac{1}{e^{Lz}-1} = -1 - \frac{1}{e^{-Lz}-1}$ and a change of variable $z := z + \chi_k$ to get

$$2I_1 = -LI_2 - 2\Gamma(-\chi_k)(\gamma + \psi(-\chi_k)) + \sum_{j\neq 0,\neq k} \Gamma(-\chi_k)\Gamma(-\chi_k + \chi_j), \tag{1.1}$$

where $I_2$ is an integral of Mellin-Barnes type [40, page 286ff]

$$I_2 = \frac{1}{2\pi \boldsymbol{i}} \int_{-\frac{1}{2}-\boldsymbol{i}\infty}^{-\frac{1}{2}+\boldsymbol{i}\infty} \Gamma(z)\Gamma(-\chi_k - z)dz = \frac{1}{2\pi \boldsymbol{i}} \int_{-\frac{1}{2}-\boldsymbol{i}\infty}^{-\frac{1}{2}+\boldsymbol{i}\infty} \Gamma(z - \chi_k)\Gamma(-z)dz.$$

To evaluate $I_2$ we shift the contour line to the right to get negative residues. The poles we consider are at $z = \chi_k$, a simple pole with residue $-\Gamma(-\chi_k)$ and at $z = l, l \in \mathbb{N}_0$, with residues $\sum\limits_{l\geq 0} \frac{(-1)^l}{l!}\Gamma(l - \chi_k)$. So

$$\begin{aligned} I_2 &= -\Gamma(-\chi_k) + \sum_{l\geq 0} \frac{(-1)^l}{l!}\Gamma(l - \chi_k) \\ &= -\Gamma(-\chi_k) + \Gamma(-\chi_k) \sum_{l\geq 0} \binom{\chi_k}{l} \\ &= \Gamma(-\chi_k)\Big(e^{2\pi \boldsymbol{i} k \log_Q 2} - 1\Big). \end{aligned}$$

On the other hand, if we write $I_1 = \frac{1}{2\pi \boldsymbol{i}} \int_{\frac{1}{2}-\boldsymbol{i}\infty}^{\frac{1}{2}+\boldsymbol{i}\infty} L\frac{\Gamma(-\chi_k+z)\Gamma(-z)}{e^{Lz}-1}dz$ and shift the contour of integration to the right, we collect the negative residues at $l = 1, 2, 3, \ldots$ as

$$I_1 = L \sum_{l\geq 1} \frac{(-1)^l\Gamma(l - \chi_k)}{l!Q^l - 1}. \tag{1.2}$$

Since we now have two expressions for $I_1$, which must be equal, we can combine (1.1) and (1.2), and cancel all terms except $I_2$, leaving us with

$$\delta_V(x) = -\frac{1}{L} \sum_{k\neq 0} \Gamma(-\chi_k)\Big(e^{2\pi \boldsymbol{i} k \log_Q 2} - 1\Big)e^{2\pi \boldsymbol{i} k x} = \delta_E(x + \log_Q 2) - \delta_E(x),$$

which, for $Q = 2$, is $\delta_E(x + 1) - \delta_E(x)$, which is zero since $\delta_E(x)$ has period 1 [32].

# Appendix B

# Mellin transforms for $d$ fixed

Below are the Mellin transform equivalents for the (weak, strict) and (weak, weak) cases for the maximum when $d$ is fixed. The previous calculations made use of Rice's method. Here the same results are obtained using the Mellin transform, as discussed in Chapters 12 and 13 which deal with $d = \alpha n$ and $d = \alpha n^\gamma$. The initial manipulations remain the same, until the coefficient of the generating functions is sought.

## B.1 The (weak, strict) case.

As discussed above, we keep the same generating function and begin after the coefficient of $z^n$ has been found (see equation(8.8)).

$$
\begin{aligned}
[z^n]F_M^{(w,s)}(z) &= \sum_{k\geq 1}\sum_{i=0}^{d-1}(1-q^{k-1})^{i+n-d}pq^{k-1}(1-q^k)^{d-1-i} \\
&= \sum_{k\geq 1}(1-q^{k-1})^{n-d}pq^{k-1}(1-q^k)^{d-1}\sum_{i=0}^{d-1}\left(\frac{1-q^{k-1}}{1-q^k}\right)^i.
\end{aligned}
$$

This geometric series on $i$ can be simplified as:

$$
\begin{aligned}
\sum_{i=0}^{d-1}\left(\frac{1-q^{k-1}}{1-q^k}\right)^i &= \frac{1-\left(\frac{1-q^{k-1}}{1-q^k}\right)^d}{1-\left(\frac{1-q^{k-1}}{1-q^k}\right)} \\
&= \frac{(1-q^k)^d-(1-q^{k-1})^d}{(1-q^k)^d}\frac{1-q^k}{1-q^k-(1-q^{k-1})} \\
&= \frac{(1-q^k)^d-(1-q^{k-1})^d}{(1-q^k)^{d-1}q^{k-1}(1-q)}.
\end{aligned}
$$

Since $p = 1-q$, we have the following cancellations

$$\begin{aligned}
[z^n]F_M^{(w,s)}(z) &= \sum_{k\geq 1}(1-q^{k-1})^{n-d}pq^{k-1}(1-q^k)^{d-1}\frac{(1-q^k)^d-(1-q^{k-1})^d}{(1-q^k)^{d-1}q^{k-1}p}\\
&= \sum_{k\geq 1}(1-q^{k-1})^{n-d}\big((1-q^k)^d-(1-q^{k-1})^d\big)\\
&= \sum_{k\geq 1}(1-q^{k-1})^{n-d}(1-q^k)^d-(1-q^{k-1})^n\\
&\sim \sum_{k\geq 1}e^{-q^{k-1}(n-d)}e^{-q^kd}-e^{-q^{k-1}n}\\
&= \sum_{k\geq 1}e^{-q^{k-1}n+q^{k-1}d(1-q)}-e^{-q^{k-1}n}\\
&= \sum_{k\geq 1}e^{-q^{k-1}n}\big(e^{q^{k-1}dp}-1\big).
\end{aligned}$$

By defining the function $f(x) := \sum_{k\geq 1} e^{-q^{k-1}x}\big(e^{q^{k-1}dp}-1\big)$, we can now use the Mellin transform. The Mellin transform of this function exists in the fundamental strip $\langle 0,\infty\rangle$, and can be found using the 'linearity' and 'scaling' rules in [11, page 576].

$$f^*(s) = \sum_{k\geq 1}\big(e^{q^{k-1}dp}-1\big)(q^{k-1})^{-s}\Gamma(s).$$

In order to sum on $k$, we must expand the exponent as a series, which will bring all $k$'s into the first level power. Then the $j=0$ term of the new sum is one, so that our expression simplifies and we can sum on $k$ as a geometric series.

$$\begin{aligned}
f^*(s) &= \sum_{k\geq 1}\left(\sum_{j\geq 0}\frac{(q^{k-1}dp)^j}{j!}-1\right)(q^{k-1})^{-s}\Gamma(s)\\
&= \sum_{k\geq 1}\sum_{j\geq 1}\frac{(q^{k-1}dp)^j}{j!}(q^{k-1})^{-s}\Gamma(s)\\
&= \sum_{j\geq 1}\frac{(q^{-1}dp)^j}{j!}q^s\Gamma(s)\sum_{k\geq 1}q^{jk}(q^k)^{-s}\\
&= \sum_{j\geq 1}\frac{(q^{-1}dp)^j}{j!}q^s\Gamma(s)\frac{q^{j-s}}{1-q^{j-s}}, \qquad \text{for } j > s\\
&= \sum_{j\geq 1}\frac{(dp)^j}{j!}\Gamma(s)\frac{1}{1-q^{j-s}}.
\end{aligned}$$

The convergence of the series restricts the strip in which $f^*(s)$ exists. The new strip will be $\langle 0,\infty\rangle \cap \langle -\infty, j\rangle = \langle 0, j\rangle$. We must thus pick an $x$-value between 0 and $j$, and since $j \leq 1$, we choose the value $\frac{1}{2}$. This means we can define our inverse transform as

$$f(x) = \frac{1}{2\pi\boldsymbol{i}}\sum_{j\geq 1}\frac{(dp)^j}{j!}\int_{(\frac{1}{2})}\Gamma(s)\frac{1}{1-q^{j-s}}x^{-s}ds.$$

For $x$ large (recall $x = n$, and we are interested in words of length $n$), we move the contour right and collect negative residues. Possible poles in this integrand occur at $s = 0, -1, -2, \ldots$ (from $\Gamma(s)$) and at $j - s = 0$ and $j - s = \chi_k$, $k \neq 0$ $\big($from $\frac{1}{1-q^{j-s}}\big)$. Moving right from $\frac{1}{2}$, the first poles we encounter are at $s = j = 1$ and $1 - s = \chi_k$. The first of these gives the dominant term and the others give the fluctuations (which we called $\delta(n)$ in equation (7.5)). For the dominant term we expand around $s = 1$, replacing $j$ with 1:

$$\frac{1}{1-q^{1-s}} = \frac{1}{1-e^{(1-s)\log q}} \sim \frac{1}{1-(1+(1-s)\log q)} = \frac{1}{(s-1)\log q},$$

and thus the negative residue is

$$\begin{aligned}
&-[(s-1)^{-1}]dp\Gamma(1)\frac{1}{(s-1)\log q}x^{-1}\\
&= -dp\frac{1}{\log q}x^{-1}\\
&= \frac{dp}{x\log Q}\\
&= \frac{dp}{nL},
\end{aligned}$$

as before. For the fluctuating terms, we let $\varepsilon := -1 + s + \chi_k$ and expand around (or evaluate at) $\varepsilon = 0$.

$$\Gamma(s) = \Gamma(1-\chi_k+\varepsilon) = \Gamma(1-\chi_k),$$

$$\frac{1}{1-q^{1-s}} = \frac{1}{1-q^{-\varepsilon+\chi_k}} = \frac{1}{1-q^{-\varepsilon}} \sim \frac{1}{1-(1-\varepsilon\log q)} = \frac{1}{\varepsilon\log q},$$

and

$$x^{-s} = x^{-\varepsilon-1+\chi_k} = x^{\chi_k-1}.$$

So for the fluctuations, we have

$$\begin{aligned}
\sum_{k\neq 0}\Big(-[\varepsilon^{-1}]dp\Gamma(1-\chi_k)\frac{1}{\varepsilon\log q}x^{\chi_k-1}\Big) &= \frac{dp}{xL}\sum_{k\neq 0}\Gamma(1-\chi_k)x^{\chi_k}\\
&= \frac{dp}{xL}\sum_{k\neq 0}\Gamma(1-\chi_k)e^{\chi_k\log x}\\
&= \frac{dp}{xL}\sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q x}.
\end{aligned}$$

Finally, replacing $x$ by $n$, we can confirm the same result as when Rice's method was used, namely that the probability that the maximum in a word occurs in the first $d$ position in the (weak,strict) case is

$$\frac{dp}{nL} + \frac{dp}{nL}\sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n} = \frac{d(1-Q^{-1})}{nL}\big(1+\delta(n)\big),$$

where $\delta(x) = \sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q x}$, as in (7.5).

## B.2 The (weak, weak) case.

Again we start by taking the coefficient of $z^n$ in the relevant generating function. This time, we take line (8.18) from the original version, which has already had some simplifications done, to give

$$\begin{aligned}
[z^n]F_M^{(w,w)}(z) &= \sum_{k\geq 1}\sum_{i=0}^{d-1}(1-q^{k-1})^i pq^{k-1}(1-q^k)^{n-1-i} \\
&= \sum_{k\geq 1} pq^{k-1}(1-q^k)^{n-1}\sum_{i=0}^{d-1}\left(\frac{1-q^{k-1}}{1-q^k}\right)^i \\
&= \sum_{k\geq 1} pq^{k-1}(1-q^k)^{n-1}\frac{(1-q^k)^d-(1-q^{k-1})^d}{(1-q^k)^{d-1}q^{k-1}p} \qquad \text{(as above)} \\
&= \sum_{k\geq 1}(1-q^k)^{n-d}\big((1-q^k)^d-(1-q^{k-1})^d\big) \\
&= \sum_{k\geq 1}(1-q^k)^n-(1-q^k)^{n-d}(1-q^{k-1})^d \\
&\sim \sum_{k\geq 1} e^{-q^k n}-e^{-q^k(n-d)}e^{-q^{k-1}d} \\
&= \sum_{k\geq 1} e^{-q^k n}-e^{-q^k n+q^k d-q^{k-1}d} \\
&= \sum_{k\geq 1} e^{-q^k n}\big(1-e^{-(1-q)q^{k-1}d}\big) \\
&= \sum_{k\geq 1} e^{-q^k n}\big(1-e^{-pq^{k-1}d}\big).
\end{aligned}$$

Now let $f(x) := \sum_{k\geq 1} e^{-q^k x}\big(1-e^{-pq^{k-1}d}\big)$, and our transform will exist in the fundamental strip $\langle 0,\infty\rangle$. We have

$$f^*(s) = \sum_{k\geq 1}\big(1-e^{-pq^{k-1}d}\big)(q^k)^{-s}\Gamma(s),$$

and again expand the exponential to get

$$\begin{aligned}
f^*(s) &= \sum_{k\geq 1}\Big(1-\sum_{j\geq 0}\frac{(-pq^{k-1}d)^j}{j!}\Big)(q^k)^{-s}\Gamma(s) \\
&= -\sum_{j\geq 1}\frac{(-pq^{-1}d)^j}{j!}\Gamma(s)\sum_{k\geq 1}q^{k(j-s)} \\
&= -\sum_{j\geq 1}\frac{(-pq^{-1}d)^j}{j!}\Gamma(s)\frac{q^{j-s}}{1-q^{j-s}}, \qquad \text{for } s<j \\
&= -\sum_{j\geq 1}\frac{(-pd)^j}{j!}q^{-s}\Gamma(s)\frac{1}{1-q^{j-s}}.
\end{aligned}$$

Once again the $s < j$ restriction means that our new strip is $\langle 0, j\rangle$, so our inverse transform becomes

$$f(x) = \frac{-1}{2\pi \boldsymbol{i}} \sum_{j\geq 1} \frac{(-pd)^j}{j!} \int_{(\frac{1}{2})} q^{-s}\Gamma(s)\frac{1}{1-q^{j-s}} x^{-s} ds.$$

As in the case above, the pole at $s = j = 1$ gives the dominant term in the result, with

$$\frac{1}{1-q^{1-s}} \sim \frac{1}{(s-1)\log q},$$

in the negative residue (let $j = 1$)

$$\begin{aligned}
&-[(s-1)^{-1}] - (-pd)q^{-1}\Gamma(1)\frac{1}{(s-1)\log q}x^{-1} \\
&= -pdq^{-1}\frac{1}{\log q}x^{-1} \\
&= \frac{pd}{nq\log Q} \\
&= \frac{(Q-1)d}{nL}.
\end{aligned}$$

For the fluctuations, we look at the poles at $1 - s = \chi_k$, for $k \neq 0$. At this stage, we define $\varepsilon := \chi_k + s - 1$, and expand around (or evaluate at) $\varepsilon = 0$.

$$q^{-s} = q^{\chi_k - 1 - \varepsilon} = q^{\chi_k - 1},$$

$$\Gamma(s) = \Gamma(\varepsilon - \chi_k + 1) = \Gamma(1-\chi_k),$$

$$\frac{1}{1-q^{1-s}} = \frac{1}{1-q^{\chi_k-\varepsilon}} = \frac{1}{1-q^{-\varepsilon}} \sim \frac{1}{1-(1-\varepsilon\log q)} = \frac{1}{\varepsilon\log q},$$

and

$$x^{-s} = x^{\chi_k - 1 - \varepsilon} = x^{\chi_k - 1}.$$

This means that the negative residues from the non-dominant poles are

$$\begin{aligned}
\sum_{k\neq 0}(-1)[\varepsilon^{-1}](-1)(-pd)q^{\chi_k-1}\Gamma(1-\chi_k)\frac{1}{\varepsilon\log q}x^{\chi_k-1} &= \frac{pd}{Lqx}\sum_{k\neq 0} q^{\chi_k}\Gamma(1-\chi_k)x^{\chi_k} \\
&= \frac{pd}{Lqx}\sum_{k\neq 0}\Gamma(1-\chi_k)x^{\chi_k} \\
&= \frac{pd}{Lqx}\sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q x},
\end{aligned}$$

and thus in total the probability is asymptotic to

$$\frac{(Q-1)d}{nL} + \frac{(Q-1)d}{nL}\sum_{k\neq 0}\Gamma(1-\chi_k)e^{2k\pi\boldsymbol{i}\log_Q n} = \frac{(Q-1)d}{nL}\big(1+\delta(n)\big),$$

as $n \to \infty$, for $\delta(x)$ as in (7.5). This corresponds to the result obtained using Rice's method in Theorem 8.4.

# Appendix C

# Proving the (weak, strict) case for $d = n$

The following calculations support the fact that our solution is 1 in the (weak, strict) case of the maximum if $d = n$. Two methods are used.

## C.1 Method 1

We examine the (weak, strict) case of the maximum (the same proof is true for the (weak, weak) case where the result is the same if $d = n$). In this case, if $d = n$ then there is no 'strict' restriction, and so we expect a probability of 1 as there will definitely be a maximum which could recur (i.e., there are no restrictions on the word). If we put $d = n$ into equation (8.8), we have the following argument: Let

$$f(n) := [z^n]F_M^{(w,s)}(z) = \sum_{k\geq 1}\sum_{i=0}^{n-1}(1-q^{k-1})^i pq^{k-1}(1-q^k)^{n-1-i}.$$

In order to show that $f(n) = 1$, we consider the expression $f(n) - f(n-1)$, which turns out to be 0. This tells us that $f(n)$ is a constant and we can find this constant by letting $n = 1$.

$$f(1) = \sum_{k\geq 1}\sum_{i=0}^{1-1}(1-q^{k-1})^i pq^{k-1}(1-q^k)^{1-1-i} = \sum_{k\geq 1} pq^{k-1} = 1,$$

as we are working with a geometric probability distribution of the natural numbers. What remains is to show that $f(n) - f(n-1) = 0$.

$$
\begin{aligned}
&f(n) - f(n-1) \\
&= \sum_{k\geq 1}\sum_{i=0}^{n-1}(1-q^{k-1})^i pq^{k-1}(1-q^k)^{n-1-i} - \sum_{k\geq 1}\sum_{i=0}^{n-2}(1-q^{k-1})^i pq^{k-1}(1-q^k)^{n-2-i} \\
&= \sum_{k\geq 1} pq^{k-1}\Big[(1-q^{k-1})^{n-1} + \sum_{i=0}^{n-2}\left[(1-q^{k-1})^i(1-q^k)^{n-1-i} - (1-q^{k-1})^i(1-q^k)^{n-2-i}\right]\Big] \\
&= \sum_{k\geq 1} pq^{k-1}\left[(1-q^{k-1})^{n-1} + \sum_{i=0}^{n-2}\left[(1-q^k)^{n-1}\left(\frac{1-q^{k-1}}{1-q^k}\right)^i - (1-q^k)^{n-2}\left(\frac{1-q^{k-1}}{1-q^k}\right)^i\right]\right] \\
&= \sum_{k\geq 1} pq^{k-1}\left[(1-q^{k-1})^{n-1} + (1-q^k)^{n-2}(1-q^k-1)\sum_{i=0}^{n-2}\left(\frac{1-q^{k-1}}{1-q^k}\right)^i\right] \\
&= \sum_{k\geq 1} pq^{k-1}\left[(1-q^{k-1})^{n-1} - q^k(1-q^k)^{n-2}\left(\frac{1-\left(\frac{1-q^{k-1}}{1-q^k}\right)^{n-1}}{1-\frac{1-q^{k-1}}{1-q^k}}\right)\right] \\
&= \sum_{k\geq 1} pq^{k-1}\left[(1-q^{k-1})^{n-1} - q^k(1-q^k)^{n-2}\left(\frac{\frac{(1-q^k)^{n-1}-(1-q^{k-1})^{n-1}}{(1-q^k)^{n-1}}}{\frac{1-q^k-1+q^{k-1}}{1-q^k}}\right)\right] \\
&= \sum_{k\geq 1} pq^{k-1}\left[(1-q^{k-1})^{n-1} - q^k(1-q^k)^{n-2}\left(\frac{(1-q^k)^{n-1}-(1-q^{k-1})^{n-1}}{(1-q^k)^{n-2}(-q^k+q^{k-1})}\right)\right] \\
&= \sum_{k\geq 1} pq^{k-1}\left[(1-q^{k-1})^{n-1} - \frac{q^k(1-q^k)^{n-1}}{q^{k-1}(1-q)} + \frac{q^k(1-q^{k-1})^{n-1}}{q^{k-1}(1-q)}\right] \\
&= \sum_{k\geq 1}\left[(1-q)q^{k-1}(1-q^{k-1})^{n-1} - q^k(1-q^k)^{n-1} + q^k(1-q^{k-1})^{n-1}\right] \qquad (p=1-q) \\
&= \sum_{k\geq 1}\left[q^{k-1}(1-q^{k-1})^{n-1} - q^k(1-q^{k-1})^{n-1} - q^k(1-q^k)^{n-1} + q^k(1-q^{k-1})^{n-1}\right] \\
&= \sum_{k\geq 1}\left[q^{k-1}(1-q^{k-1})^{n-1} - q^k(1-q^k)^{n-1}\right],
\end{aligned}
$$

which is a telescoping series in which all terms cancel except the very first one which is zero, i.e., $q^{k-1}(1-q^{k-1})^{n-1}|_{k=1} = 0$.

## C.2 Method 2

Again, we examine the (weak, strict) case of the maximum where, if $d = n$, there are no restrictions on the word and so we expect a probability of 1. If we put $d = n$ into equation (8.8), we have the following argument:

$$f(n) := [z^n]F_M^{(w,s)}(z)$$

$$
\begin{aligned}
&= \sum_{k\geq 1}\sum_{i=0}^{n-1}(1-q^{k-1})^i pq^{k-1}(1-q^k)^{n-1-i} \\
&= \sum_{k\geq 1} pq^{k-1}(1-q^k)^{n-1}\sum_{i=0}^{n-1}\left(\frac{1-q^{k-1}}{1-q^k}\right)^i \\
&= \sum_{k\geq 1} pq^{k-1}(1-q^k)^{n-1}\frac{1-\left(\frac{1-q^{k-1}}{1-q^k}\right)^n}{1-\frac{1-q^{k-1}}{1-q^k}} \\
&= \sum_{k\geq 1} pq^{k-1}(1-q^k)^{n-1}\frac{\frac{(1-q^k)^n-(1-q^{k-1})^n}{(1-q^k)^n}}{\frac{1-q^k-(1-q^{k-1})}{1-q^k}} \\
&= \sum_{k\geq 1} pq^{k-1}(1-q^k)^{n-1}\frac{(1-q^k)^n-(1-q^{k-1})^n}{(1-q^k)^{n-1}q^{k-1}(1-q)} \\
&= \sum_{k\geq 1}[(1-q^k)^n-(1-q^{k-1})^n] \\
&= \sum_{k\geq 1}\left[\big((1-q^k)^n-1\big)-\big((1-q^{k-1})^n-1\big)\right] \\
&= -\big((1-q^{1-1})^n-1\big) \\
&= 1,
\end{aligned}
$$

since it is a telescoping series (adding $-1$ twice to take care of convergence).

# Bibliography


[1] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Dover, New York, 1970.

[2] M. Archibald, A. Knopfmacher, and H. Prodinger. The number of distinct values in a geometrically distributed sample. (To appear in a special volume of *European Journal of Combinatorics* devoted to Eurocomb '03.).

[3] Z-D. Bai, H-K. Hwang, and W-Q. Liang. Normal approximations of the number of records in geometrically distributed random variables. *Random Structures and Algorithms*, 13:319–334, 1998.

[4] W.H. Burge. An analysis of binary search trees formed from sequences of nondistinct keys. *Journal of the Association for Computing Machinary*, 23(3):451–454, 1976.

[5] L. Devroye. A limit theory for random skip lists. *Annals of Applied Probability*, 2:597–609, 1992.

[6] B. Eisenberg, G. Stengle, and G. Strang. The asymptotic probability of a tie for first place. *The Annals of Applied Probability*, 3(3):731–745, 1993.

[7] P. Flajolet, X. Gourdon, and P. Dumas. Mellin transforms and asymptotics: Harmonic sums. *Theoretical Computer Science*, 144:3–58, 1995.

[8] P. Flajolet, P. Grabner, P. Kirschenhofer, H. Prodinger, and R.F. Tichy. Mellin transforms and asymptotics: Digital sums. *Theoretical Computer Science*, 123:291–314, 1994.

[9] P. Flajolet and G.N. Martin. Probabilistic counting algorithms for data base applications. *Journal of Computer and System Sciences*, 31:182–209, 1985.

[10] P. Flajolet and R. Sedgewick. Mellin transforms and asymptotics: Finite differences and Rice's integrals. *Theoretical Computer Science*, 144:101–124, 1995.

[11] P. Flajolet and R. Sedgewick. Analytic combinatorics, symbolic combinatorics. http://pauillac.inria.fr/algo/flajolet/Publications/books.html, November 2004.

[12] R.L. Graham, D.E. Knuth, and O. Patashnik. *Concrete Mathematics*. Addison-Wesley, 1994.

[13] P. Hitczenko and G. Louchard. Distinctness of compositions of an integer: a probabilistic analysis. *Random Structures and Algorithms*, 19(3-4):407–437, 2001.

[14] R. Kemp. Binary search trees constructed from nondistinct keys with/without specified probabilities. *Theoretical Computer Science*, 156:39–70, 1996.

[15] P. Kirschenhofer and H. Prodinger. b-tries: a paradigm for the application of numbertheoretic methods in the analysis of algorithms. In *Contributions to General Algebra*, volume 6, pages 141–153. Hölder–Pichler–Tempsky–Teubner, 1988.

[16] P. Kirschenhofer and H. Prodinger. On the analysis of probabilistic counting. *Lecture Notes in Mathematics*, 1452:117–120, 1990. (E. Hlawka, R.F. Tichy eds.).

[17] P. Kirschenhofer and H. Prodinger. A result in order statistics related to probabilistic counting. *Computing*, 51:15–27, 1993.

[18] P. Kirschenhofer and H. Prodinger. The number of winners in a discrete geometrically distributed sample. *Annals in Applied Probability*, 6:687–694, 1996.

[19] P. Kirschenhofer and H. Prodinger. Comparisons in Hoare's Find algorithm. *Combinatorics, Probability, and Computing*, 7:111–120, 1998.

[20] P. Kirschenhofer, H. Prodinger, and W. Szpankowski. Multidimensional digital searching and some new parameters in tries. *International Journal of Foundations of Computer Science*, 4:69–84, 1993.

[21] A. Knopfmacher and H. Prodinger. Combinatorics of geometrically distributed random variables: Value and position of the $r$th left-to-right maximum. *Discrete Mathematics*, 226:255–267, 2001.

[22] A. Knopfmacher and N. Robbins. Compositions with parts constrained by the leading summand. (To appear in *Ars Combinatoria*).

[23] D.E. Knuth. *The Art of Computer Programming*, volume 1. Addison-Wesley, 3rd edition, 1997.

[24] G. Louchard. The number of distinct part sizes of some multiplicity in compositions of an integer. A probabilistic analysis. *Discrete Mathematics and Theoretical Computer Science*, DRW'03, Paris, full version:155–170, 2003.

[25] G. Louchard and H. Prodinger. Ascending runs of sequences of geometrically distributed random variables: A probabilistic analysis. *Theoretical Computer Science*, 304:59–86, 2003.

[26] G. Louchard, H. Prodinger, and M.D. Ward. The number of distinct values of some multiplicity in sequences of geometrically distributed random variables. (To appear in a special volume of *Discrete Mathematics and Theoretical Computer Science* devoted to AofA'05.).

[27] T. Papadakis, I. Munro, and P. Poblete. Average search and update costs in skip lists. *BIT*, 32:316–332, 1992.

[28] H. Prodinger. Combinatorial problems of geometrically distributed random variables and applications in computer science. In V. Strehl and R. König, editors, *Publications de l'IRMA (Straßbourg)*, volume 30, pages 87–95, 1993.

[29] H. Prodinger. Combinatorics of geometrically distributed random variables: Left-to-right maxima. *Discrete Mathematics*, 153:253–270, 1996.

[30] H. Prodinger. Combinatorics of geometrically distributed random variables: New $q$-tangent and $q$-secant numbers. *International Journal of Mathematics and Mathematical Sciences*, 24, 2000.

[31] H. Prodinger. Combinatorics of geometrically distributed random variables: Inversions and a parameter of Knuth. *Annals of Combinatorics*, 5:241–250, 2001.

[32] H. Prodinger. Compositions and Patricia tries: No fluctuations in the variance! *Proceedings of the 6th Workshop on ALENEX and the 1st workshop on ANALCO, L. Arge, G. Italiano, R. Sedgewick (editors). SIAM.*, pages 211–215, 2004.

[33] W. Pugh. Skip lists: a probabilistic alternative to balanced trees. *Communications of the ACM*, 33:668–676, 1990.

[34] L. Råde. Problem e3436. *American Mathematical Monthly*, 99:366, 1991.

[35] R. Sedgewick. Quicksort with equal keys. *SIAM Journal on Computing*, 6(2):240–267, 1977.

[36] R. Sedgewick and P. Flajolet. *An Introduction to the Analysis of Algorithms.* Addison-Wesley, 1996.

[37] J. Spieß. Some identities involving harmonic numbers. *Mathematics of Computation*, 55(192):839–863, 1990.

[38] W. Szpankowski. *Average Case Analysis of Algorithms on Sequences.* John Wiley and Sons, New York, 2001.

[39] W. Szpankowski and V. Rego. Yet another application of a binomial recurrence: order statistics. *Computing*, 43:401 – 410, 1990.

[40] E.T. Whittaker and G.N. Watson. *A Course in Modern Analysis.* Cambridge University Press, fourth edition, 1927. reprinted 1973.