Position of the maximum in a sequence with geometric distribution

Margaret Archibald[1]

[1] *The John Knopfmacher Centre for Applicable Analysis and Number Theory, School of Mathematics, University of the Witwatersrand, Private Bag 3, Wits 2050, Johannesburg, South Africa.* `marchibald@maths.wits.ac.za`

As a sequel to [1], the position of the maximum in a geometrically distributed sample is investigated. Samples of length n are considered, where the maximum is required to be in the first d positions. The probability that the maximum occurs in the first d positions is sought for d dependent on n (as opposed to d fixed in [1]). Two scenarios are discussed. The first is when $d = \alpha n$ for $0 < \alpha \leq 1$, where Mellin transforms are used to obtain the asymptotic results. The second is when $1 \leq d = o(n)$.

Keywords: Mellin transforms, generating functions, geometric distribution.

1 Introduction

Consider a word whose letters are natural numbers. Assume that these letters occur independently and with geometric probability. So for $p + q = 1$, each letter j appears in the word with probability pq^{j-1}. We write $Q := q^{-1}$ and $L := \log Q$.

We address the question: "What is the probability that the maximum in a word occurs in the first d positions?" We take words of length n and require $d \leq n$. Previously (in [1]), d was considered fixed. Now, d is allowed to grow with n. First, we assume that d is proportional to n and then we consider the case when d is $o(n)$ (but at least 1). The latter produces the same solutions as when d is fixed (see [1]).

We distinguish between two cases: 'strict' and 'weak'. A 'strict' maximum never recurs, whereas a 'weak' maximum can recur any number of times. These apply separately to the two parts of the word (the first d letters, and the remaining $n - d$ letters). This means that in total, there are four different cases to be dealt with: (strict, strict), (weak, strict), (strict, weak) and (weak, weak), where the first entry refers to the first d letters in the word, and the second entry to the rest of the word. The results in the (strict, strict) case for d fixed will hold for the other scenarios too (i.e., it is not required that d is independent of n). This is because the relevant calculations in [1] still go though when we take the limits as $n \to \infty$. This is not true for the remaining three cases which are dealt with in Section 3.

2 Results

Theorem 1 *The probability that the maximum value in a word of length n is in the first d positions (for $d = \alpha n$) is*

$$\mathit{Max}_{(w,s)}(n) \sim \frac{1}{L}\log\left(\frac{1}{1-\alpha(1-Q^{-1})}\right) + \frac{1}{L}\big(\psi_0(n(1-\alpha p)) - \psi_0(n)\big),$$

$$\mathit{Max}_{(s,w)}(n) \sim \frac{\alpha(Q-1)}{L(1+\alpha(Q-1))}\big(1 + \psi(n(q+p\alpha))\big),$$

$$\mathit{Max}_{(w,w)}(n) \sim \frac{\log(1+\alpha(Q-1))}{L} + \frac{1}{L}\left(\psi_0(n) - \psi_0\Big(n\Big(\frac{q+\alpha p}{q}\Big)\Big)\right),$$

as $n \to \infty$ where $Q := q^{-1}$, $L := \log Q$, and the fluctuations are defined by

$$\psi(x) := \sum_{k\neq 0} \Gamma(1-\chi_k)e^{2k\pi \boldsymbol{i} x}, \qquad \text{for} \quad \mathrm{k} \in \mathbb{Z},$$

and

$$\psi_0(x) := \sum_{k\neq 0} \Gamma(-\chi_k) e^{2k\pi\boldsymbol{i}\log_Q x}, \qquad \text{for} \quad \text{k} \in \mathbb{Z}.$$

Note that we write $\boldsymbol{i}$ rather than the i we use as an index.

Theorem 2 *The probability that the maximum value in a word of length n is in the first d positions, for $1 \le d = o(n)$, is*

$$\mathit{Max}_{(w,s)}(n) \sim \frac{(1-Q^{-1})d}{Ln}\big(1+\psi(n)\big),$$
$$\mathit{Max}_{(s,w)}(n) \sim \frac{(Q-1)d}{Ln}(1+\psi(n)),$$
$$\mathit{Max}_{(w,w)}(n) \sim \frac{(Q-1)d}{Ln}(1+\psi(n)),$$

as $n \to \infty$, where $\psi(x) = \sum_{k\neq 0} \Gamma(1-\chi_k) e^{2k\pi\boldsymbol{i}\log_Q x}$, $k \in \mathbb{Z}$.

3 Suppose $d = \alpha n$

Suppose we now consider $d = \alpha n$ where $0 < \alpha \le 1$. The 'd fixed' results from [1] continue to hold only in the (strict, strict) case. For the other three cases we make use of a technique from complex analysis called the 'Mellin' transform. The rules used below can be found in [2] and [5], among others. The first case is done in greater detail than the other two, as a similar process is used in each case.

3.1 Case (weak, strict), for $d = \alpha n$

For this scenario (finding the probability with which the maximum, k, occurs in the first d letters of a word) 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 $\{1, \cdots, k-1\}$. Now suppose that $d = \alpha n$, for $0 < \alpha \le 1$. The generating function is explained in [1], yielding the following coefficients (with d replaced by αn).

$$f_{(w,s)}(n) := \sum_{k\ge 1} \sum_{i=0}^{\alpha n - 1} (1-q^{k-1})^{i+n-\alpha n} p q^{k-1} (1-q^k)^{\alpha n - 1 - i}. \tag{1}$$

Now note that

$$\sum_{i=0}^{\alpha n-1} \left(\frac{1-q^{k-1}}{1-q^k}\right)^i = \frac{(1-q^k)^{\alpha n} - (1-q^{k-1})^{\alpha n}}{(1-q^k)^{\alpha n-1} q^{k-1}(1-q)}, \tag{2}$$

and thus, since $p = 1-q$ and $(1-a)^n \sim e^{-an}$ for small a,

$$\begin{aligned} f_{(w,s)}(n) &= \sum_{k\ge 1} (1-q^{k-1})^{n(1-\alpha)} \big[(1-q^k)^{\alpha n} - (1-q^{k-1})^{\alpha n}\big] \\ &\sim \sum_{k\ge 1} \Big[e^{-nq^{k-1}(1-\alpha p)} - e^{-nq^{k-1}}\Big]. \end{aligned}$$

We are now in a position to use Mellin transforms to find an approximation. We shift the fundamental strip from $\langle 0, \infty\rangle$ to $\langle -1, 0\rangle$ and define a new function:

$$f_{ws}(x) := \sum_{k\ge 1} \Big[\big(e^{-nq^{k-1}(1-\alpha p)} - 1\big) - \big(e^{-nq^{k-1}} - 1\big)\Big]. \tag{3}$$

Then the Mellin transform of this function is:

$$\begin{aligned} f^*_{ws}(s) &= \sum_{k\ge 1} \Big[\big(q^{k-1}(1-\alpha p)\big)^{-s}\Gamma(s) - (q^{k-1})^{-s}\Gamma(s)\Big] \\ &= \sum_{k\ge 1} q^{s(1-k)}\Gamma(s)\big[(1-\alpha p)^{-s} - 1\big] \\ &= q^s \Gamma(s)\big[(1-\alpha p)^{-s} - 1\big] \sum_{k\ge 1} (q^{-s})^k \\ &= \Gamma(s)\big[(1-\alpha p)^{-s} - 1\big] \frac{1}{1-q^{-s}}, \qquad \text{for } \Re(s) < 0, \end{aligned} \tag{4}$$

where $\Re(s)$ represents the real part of the complex number s. The reason for shifting the fundamental strip is that the transform exists in the intersection of the domain of convergence of the generalised Dirichlet series and the fundamental strip of $f^*(s)$. The intersection $\langle -\infty, 0\rangle \cap \langle 0, \infty\rangle$ is empty, but with the shift we have a final fundamental strip of $\langle -1, 0\rangle$. We choose a value inside this, say $-\frac{1}{2}$, with which to perform our inverse Mellin transform:

$$f_{ws}(x) = \frac{1}{2\pi \boldsymbol{i}} \int\limits_{(-\frac{1}{2})} \Gamma(s)\big[(1-\alpha p)^{-s} - 1\big] \frac{1}{1-q^{-s}} x^{-s} ds. \tag{5}$$

The notation $(-\frac{1}{2})$ under the integral sign means an integral from $-\frac{1}{2} - \boldsymbol{i}\infty$ to $-\frac{1}{2} + \boldsymbol{i}\infty$. This can be approximated by moving the contour to the right (and thus collecting negative residues) since we are interested in x large. The first poles we encounter are the simple pole at $s = 0$ (which would be a double pole except that one cancels with the factor $(1-\alpha p)^{-s} - 1$) and the simple poles at $s = \chi_k := \frac{2k\pi\boldsymbol{i}}{L}$, $k \in \mathbb{Z} \setminus \{0\}$ where $L := \log Q$. The former contributes the main term and the rest contribute the fluctuations which are comparatively extremely small. As $s \to 0$,

$$\Gamma(s) \sim \frac{1}{s},$$

$$(1-\alpha p)^{-s} - 1 \sim 1 - s\log(1-\alpha p) - 1 = -s\log(1-\alpha p),$$

$$\frac{1}{1-q^{-s}} = \frac{1}{1-e^{-s\log q}} \sim \frac{1}{1-(1-s\log q)} = \frac{1}{s\log q},$$

and

$$x^{-s} = e^{-s\log x} \sim 1.$$

Thus the negative residue is

$$-[s^{-1}]\frac{1}{s}\big(-s\log(1-\alpha p)\big)\frac{1}{s\log q} = \frac{\log(1-\alpha p)}{\log q}.$$

We also have simple poles at $s = \chi_k$, for $k \neq 0$. Let $\varepsilon := s - \chi_k$ then expanding around $\varepsilon = 0$ gives

$$\frac{1}{1-q^{-s}} = \frac{1}{1-q^{-\chi_k-\varepsilon}} = \frac{1}{1-q^{-\varepsilon}} = \frac{1}{1-e^{-\varepsilon\log q}} \sim \frac{1}{1-(1-\varepsilon\log q)} = \frac{1}{\varepsilon\log q}.$$

So the negative residues for all non-zero k are

$$\begin{aligned}\sum_{k\neq 0}(-1)[\varepsilon^{-1}]\Gamma(\chi_k)\big[(1-\alpha p)^{-\chi_k} - 1\big]\frac{1}{\varepsilon\log q}x^{-\chi_k} &= \frac{1}{L}\sum_{k\neq 0}\Gamma(\chi_k)\big[(1-\alpha p)^{-\chi_k} - 1\big]x^{-\chi_k} \\ &= \frac{1}{L}\sum_{k\neq 0}\Gamma(\chi_k)\big[e^{-\chi_k\log(1-\alpha p)} - 1\big]e^{-\chi_k\log x} \\ &= \frac{1}{L}\sum_{k\neq 0}\Gamma(\chi_k)\big[e^{-\chi_k\log(x(1-\alpha p))} - e^{-\chi_k\log x}\big] \\ &= \frac{1}{L}\sum_{k\neq 0}\Gamma(-\chi_k)\big[e^{2k\pi\boldsymbol{i}\log_Q(x(1-\alpha p))} - e^{2k\pi\boldsymbol{i}\log_Q x}\big].\end{aligned}$$

Now put these 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 and where $d = \alpha n$ grows with n as $n \to \infty$:

$$\begin{aligned}&\frac{\log(1-\alpha p)}{\log q} + \frac{1}{L}\sum_{k\neq 0}\Gamma(-\chi_k)\big[e^{2k\pi\boldsymbol{i}\log_Q(n(1-\alpha p))} - e^{2k\pi\boldsymbol{i}\log_Q n}\big] \\ &= \frac{\log(1-\alpha(1-Q^{-1}))}{-L} + \frac{1}{L}\big(\psi_0(n(1-\alpha p)) - \psi_0(n)\big)\end{aligned} \tag{6}$$

for $Q = q^{-1}$, $L = \log Q$ and $\psi_0(x)$ as in Theorem 1. Note that if $\alpha = 1$ then $d = n$ and the main term yields a probability of 1, confirming our result on a word with no restrictions.

3.2 Case (strict, weak), for $d = \alpha n$

In this case, we allow the maximum, k, to occur only once in the first d letters, but any number of times in the rest of the word. Again, [1] provides us with the generating function whose coefficients are

$$f_{(s,w)}(n) = \sum_{k\geq 1} \alpha n p q^{k-1}(1-q^{k-1})^{\alpha n-1}(1-q^k)^{n(1-\alpha)}$$
$$\sim \sum_{k\geq 1} \alpha n p q^{k-1} e^{-nq^{k-1}(q+p\alpha)}. \tag{7}$$

If we define the function

$$f_{sw}(x) := \sum_{k\geq 1} \alpha x p q^{k-1} e^{-xq^{k-1}(q+p\alpha)},$$

then the Mellin transform will be

$$f^*_{sw}(s) = \sum_{k\geq 1} \alpha p q^{k-1}(q^{k-1})^{-(s+1)}(q+p\alpha)^{-(s+1)}\Gamma(s+1)$$
$$= \alpha p (q+p\alpha)^{-(s+1)}\Gamma(s+1)\frac{1}{1-q^{-s}}, \qquad \text{for } \Re(s) < 0, \tag{8}$$

and the fundamental strip is the overlap of the interval $(-\infty, 0)$ and the fundamental strip of xe^{-x} which is $\langle -1, \infty\rangle$, i.e., $\langle -1, 0\rangle$. Hence we pick our contour integral from $-\frac{1}{2} - \boldsymbol{i}\infty$ to $-\frac{1}{2} + \boldsymbol{i}\infty$, and perform the inverse Mellin transform to get:

$$f_{sw}(x) = \frac{1}{2\pi \boldsymbol{i}} \int_{(-\frac{1}{2})} \alpha p (q+p\alpha)^{-(s+1)}\Gamma(s+1)\frac{1}{1-q^{-s}} x^{-s} ds. \tag{9}$$

By moving the contour to the right, the first poles we reach are at $s = 0$ and $s = \chi_k$, $k \neq 0$. For the main term, the negative residue is

$$-[s^{-1}]\alpha p (q+p\alpha)^{-1}\Gamma(1)\frac{1}{s \log q} = \frac{\alpha p}{L(q+p\alpha)}.$$

The fluctuations come from the negative residues of the poles at $s = \chi_k$, $k \neq 0$. Let $\varepsilon := s - \chi_k$, then around $\varepsilon = 0$ we get $\frac{1}{1-q^{-s}} \sim \frac{1}{\varepsilon \log q}$, and so these poles contribute:

$$-\sum_{k\neq 0}[\varepsilon^{-1}]\alpha p (q+p\alpha)^{-(\chi_k+1)}\Gamma(\chi_k+1)\frac{1}{\varepsilon \log q} x^{-\chi_k} = \frac{\alpha p}{L(q+p\alpha)} \sum_{k\neq 0}\Gamma(1-\chi_k) e^{2k\pi \boldsymbol{i} \log_Q (x(q+p\alpha))}.$$

This gives a total probability in the (strict, weak) case asymptotic to

$$\frac{\alpha(Q-1)}{L(1+\alpha(Q-1))}\big(1+\psi(n(q+p\alpha))\big), \tag{10}$$

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

(For $\alpha = 1$, the dominant term gives $\frac{p}{L}$, which is the same as the probability of having one winner among n players in a game where each player tosses a coin until a head appears, and the winner is the player who takes the longest to toss a head, see [3].)

3.3 Case (weak, weak), for $d = \alpha n$

Here, the maximum can recur anywhere, having first appeared at least once in the first d letters. From [1], and by (2), we can approximate as in the (w,s) case to get:

$$f_{(w,w)}(n) \sim \sum_{k\geq 1}\big[(e^{-nq^k} - 1) - (e^{-nq^{k-1}(q+\alpha p)} - 1)\big]. \tag{11}$$

We define an exact function in terms of x to be:

$$f_{ww}(x) := \sum_{k\geq 1}\big[(e^{-xq^k} - 1) - (e^{-xq^{k-1}(q+\alpha p)} - 1)\big].$$

The transform of this function is

$$f^*_{ww}(s) = \sum_{k\ge 1}\left[q^{-sk}\Gamma(s) - (q^{k-1})^{-s}(q+\alpha p)^{-s}\Gamma(s)\right]$$
$$= \Gamma(s)\left[1 - q^s(q+\alpha p)^{-s}\right]\frac{1}{q^s - 1}, \tag{12}$$

which exists in the strip $\langle -1, 0\rangle$. We can thus rewrite $f_{ww}(x)$ as a contour integral

$$f_{ww}(x) = \frac{1}{2\pi \boldsymbol{i}} \int_{(-\frac{1}{2})} \Gamma(s)\left[1 - q^s(q+\alpha p)^{-s}\right]\frac{1}{q^s-1}x^{-s}ds. \tag{13}$$

The relevant simple poles occur at $s = 0$ and $s = \chi_k$, $k \neq 0$. The negative residue at $s = 0$ is

$$-[s^{-1}]\frac{1}{s}s\log\Big(\frac{q+\alpha p}{q}\Big)\frac{1}{s\log q} = \frac{1}{L}\log\Big(\frac{q+\alpha p}{q}\Big), \tag{14}$$

which, as in the (weak, strict) case, is one for $\alpha = 1$. For the poles at $s = \chi_k$, let $\varepsilon := s - \chi_k$. Then expanding around $\varepsilon = 0$ gives $\frac{1}{q^s-1} \sim \frac{1}{\varepsilon \log q}$ and so the negative residues are

$$-\sum_{k\neq 0}[\varepsilon^{-1}]\Gamma(\chi_k)\left[1-q^{\chi_k}(q+\alpha p)^{-\chi_k}\right]\frac{1}{\varepsilon\log q}x^{-\chi_k} = \frac{1}{L}\sum_{k\neq 0}\Gamma(-\chi_k)\Big[e^{2k\pi\boldsymbol{i}\log_Q x} - e^{2k\pi\boldsymbol{i}\log_Q(x(\frac{q+\alpha p}{q}))}\Big]. \tag{15}$$

By summing (14) and (15) and replacing x by n the asymptotic result for the (weak, weak) case in Theorem 1 is found.

4 Suppose $1 \le d = o(n)$

Initially, in this section d was considered to be dependent on n according to the relationship $d = \alpha n^\gamma$, where $0 < \gamma < 1$, and $0 < \alpha \le 1$. Mellin transforms were used to obtain the same results as found in the d fixed case (see [1]). However, it was then found (as suggested by a referee) that in fact any d such that $1 \le d = o(n)$ will produce the same results. The explanation is given below.

We show that the results are the same as when d is fixed by referring back to the step in the calculations for d fixed (see [1]) where the $d = \alpha n$ calculations failed. The important stage is when the main term of the probability is given by the expression

$$\sum_{i=0}^{d-1}\sum_{l=0}^{d-1-i}\binom{d-1-i}{l}(-1)^l\frac{Q^{-l}(1-Q^{-1})}{L}\frac{1}{(l+1+N)\binom{N+l}{l}}. \tag{16}$$

This is the (weak, strict) case, but the others are similar. Since N grows like n, we use n instead of N for simplicity. For d fixed, it can be seen that the $l = 0$ term dominates, since each term in the sum on l is of order $O\big(\frac{1}{n^{l+1}}\big)$. For d proportional to n, each term in the inner sum is of order $O(\frac{1}{n})$, so none clearly dominates, and Mellin transforms are required to find the result (see Section 3). But what if $d = \alpha n^\gamma$ for $0 < \gamma < 1$, or even $d = \frac{n}{\log n}$?

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 of order $O\big(\frac{1}{nf^l(n)}\big)$. For $l = 0$, we again have an order of $O(\frac{1}{n})$. In fact, this term can be expressed as $\frac{c}{n}$ where c is a constant. This will dominate all other terms, since even the infinite sum on l (a geometric series) is:

$$\frac{1}{n}\sum_{l=1}^{\infty}(f(n))^{-l} = \frac{1}{n(f(n)-1)} = o\Big(\frac{1}{n}\Big). \tag{17}$$

The results in Theorem 2 follow from the above.

5 Conclusion

Table 1 is a summary of results from this paper. (The d fixed results are from [1]). This table shows only the dominant term for the results, expressed in terms of Q.

Case	(s,s)	(w,s)	(s,w)	(w,w)
$1 \leq d = o(n)$	$\frac{(1-Q^{-1})d}{Ln}$	$\frac{(1-Q^{-1})d}{Ln}$	$\frac{(Q-1)d}{Ln}$	$\frac{(Q-1)d}{Ln}$
$d = \alpha n$	$\frac{(1-Q^{-1})\alpha}{L}$	$\frac{\log(1-\alpha(1-Q^{-1}))}{-L}$	$\frac{\alpha(Q-1)}{L(1+\alpha(Q-1))}$	$\frac{\log(1+\alpha(Q-1))}{L}$

Tab. 1: Table of results for the two catagories

If we consider α small (i.e., close to zero) in the second category, we should get similar solutions to catagory 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 the approximations $\log(1+x) \sim x$ and $\frac{1}{1-x} \sim 1$ as $x \to 0$ ([4]). 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 each α by $\frac{d}{n}$, it can be seen that each of these corresponds to the d fixed case in Table 1 above.

Acknowledgements

I hereby acknowledge the continued support, help and availability of my supervisors Prof. H. Prodinger and Prof. A. Knopfmacher. I would also like to acknowledge the referee for many helpful comments.

References

[1] M. Archibald. Restrictions on the position of the maximum/minimum in a geometrically distributed sample. In *Mathematics and Computer Science III: Algorithms, Trees, Combinatorics and Probabilities*, September 2004.

[2] P. Flajolet and R. Sedgewick. Analytic combinatorics, symbolic combinatorics (chapters I-IX). http://pauillac.inria.fr/algo/flajolet/Publications/books.html, November 2004.

[3] P. Kirschenhofer and H. Prodinger. The number of winners in a discrete geometrically distributed sample. *Annals in Applied Probability*, 6:687–694, 1996.

[4] R. Sedgewick and P. Flajolet. *An Introduction to the Analysis of Algorithms*. Addison-Wesley, 1996.

[5] W. Szpankowski. *Average Case Analysis of Algorithms on Sequences*. John Wiley and Sons, New York, 2001.