Discrete Mathematics and Theoretical Computer Science DMTCS vol. **9**, 2007, 137–170

A combinatorial and probabilistic study of initial and end heights of descents in samples of geometrically distributed random variables and in permutations

Guy Louchard[1] and Helmut Prodinger[2] †

[1] *Université Libre de Bruxelles, Département d'Informatique, CP 212, Boulevard du Triomphe, B-1050 Bruxelles, Belgium.* louchard@ulb.ac.be

[2] *Stellenbosch University, Department of Mathematics, 7602 Stellenbosch, South Africa.* hproding@sun.ac.za

received May 16, 2007, *accepted July 30, 2007*.

In words, generated by independent geometrically distributed random variables, we study the lth descent, which is, roughly speaking, the lth occurrence of a neighbouring pair ab with $a > b$. The value a is called the initial height, and b the end height. We study these two random variables (and some similar ones) by combinatorial and probabilistic tools.

We find in all instances a generating function $\Psi(v,u)$, where the coefficient of $v^j u^i$ refers to the jth descent (ascent), and i to the initial (end) height. From this, various conclusions can be drawn, in particular expected values.

In the probabilistic part, a Markov chain model is used, which allows to get explicit expressions for the heights of the *second* descent. In principle, one could go further, but the complexity of the results forbids it.

This is extended to permutations of a large number of elements.

Methods from q-analysis are used to simplify the expressions. This is the reason that we confine ourselves to the geometric distribution only. For general discrete distributions, no such tools are available.

1 Introduction

Let X be a random variable (RV), distributed according to the geometric distribution with parameter p ($\mathsf{geom}(p)$): $\mathbb{P}(X = k) = pq^{k-1}$, with $q = 1 - p$. We consider a sequence $X_1X_2 \dots X_n$ of independent RVs. We also speak about *words* $a_1 \dots a_n$; there is some interest in combinatorial parameters of such words, generated by independent geometric random variables.

The two most prominent examples are *Skip lists* and *Probabilistic counting* (8; 5; 4).

In this paper we continue the study of *descents*.

†This material is based upon work supported by the National Research Foundation under grant number 2053748

In a word w_1abw_2 we say that ab is the lth descent, if $a > b$ (strict model) or $a \geq b$ (weak model), and the initial word w_1a has $l-1$ descents. Furthermore, we refer to a as the *initial height* and to b as the *end height*. Equivalently, we use the notions initial value and end value.

In (6), these random variables were studied, but only for $l = 1$, i.e., the *first descent*. Here, we are able to deal with the general case.

This paper uses a generating functions approach and a probabilistic approach. The results complement each other, but are not disjoint. Of course, both are very useful and interesting.

The generating function approach is as follows: First, we construct a generating function $F(z, v, u)$ in 3 variables, z, v, u, where z marks the length of the word, v the number of descents, and u the last letter of the word. In other words,

$$F(z, v, u) = \sum_{n,j \geq 1,\ l \geq 0} \mathbb{P}[\text{a word of length } n \text{ has } l \text{ descents and last letter } j] z^n v^l u^j.$$

Occasionally, it is clearer to write $F(u)$ only.

Once this is achieved, we construct a new generating function (G for initial height, H for end height), by attaching a descent (which is a simple substitution, since the variable u "remembers" the last letter) and an arbitrary rest. In this way, we have a generating function, where the variable u no longer codes the last letter, but the initial height (resp. end height) of the lth descent.

The quantities that we get for $l = 1$ coincide (of course!) with the older paper; however, they come out in different forms. To show formally that they are the same, one uses identities from q-analysis, such as *Heine's transformation formula.* This was demonstrated extensively in (6).

We also consider the analogous questions for *ascents;* the motivation is that, in (7), the *last* descent was studied. In the reversed word, the last descent becomes the *first* ascent, and now we have developed the machinery to deal in general with the lth ascent.

The probabilistic approach works as follows. We start from an *infinite* sequence of geometric random variables, with parameter p ($\mathsf{geom}(p)$). First, we consider the successive descents as a Markov chain, related to initial and end values of each descent. Next, we use this Markov chain to obtain the distribution of initial and end values of first and second descents. Then, the first moments of the first and second descents initial values are analyzed by intensive use of some combinatorial identities. Next, we obtain the asymptotic distribution of the number of descents initial values in some interval. Finally, starting from n $\mathsf{geom}(p)$ RV, as $q \to 1$, we can derive the asymptotic properties of first and second descents, in a large permutation.

In this part, only the strict model and the descents will be considered.

The explicit forms of the distribution of the descents become very complicated when going from first to second etc. descent. However, a stationary distribution, which is very simple, is rapidly approached. It is given by

$$\frac{1+q}{q} pq^{i-1}(1-q^{i-1})$$

for the initial height, and by

$$\frac{1+q}{q} pq^{2i-1}$$

for the end height. There is an intuitive explanation of them: The first one is the conditional probability that we have a pair ij, given that it is a descent, and the second one that we have a pair ji, again given that it is a descent.

Several useful combinatorial identities, derived from *Heine's formula*, are given in the Appendix.

We will need notation from q-analysis; the most important ones are $(x)_n := (1-x)(1-qx)\dots(1-xq^{n-1})$, and $\begin{bmatrix} n \\ k \end{bmatrix} = \frac{(q)_n}{(q)_k(q)_{n-k}}$ (Gaussian coefficients). The relevant formulæ can be found in (1).

Although one could think about performing the present analysis for general discrete distributions,we refrain from doing so. The reason is this: Methods from q-analysis are used to simplify the expressions, and this is very special to the geometric distribution. In general, no such tools are available.

Part I
Combinatorial Analysis

2 Descents: the weak model

Let $f_i(u)$ be the generating function with $[z^n u^j]f_i(u)$ is the probability that a word of length $n \geq 1$ has i descents, and that the last letter is j. (Only the dependency on the variable u has been made explicit.)

Here is the recursion for $i \geq 1$.

$$f_i(u) = \frac{puz}{1-qu} f_{i-1}(1) - \frac{puz}{1-qu} f_{i-1}(uq) + \frac{puz}{1-qu} f_i(uq).$$

This is a typical application of the *adding-a-new slice technique*: The variable u keeps track of the last letter in the word. So, if we replace u^m by

$$\sum_{1 \leq j \leq m} pq^{j-1}u^j = pu\frac{1-(qu)^m}{1-qu},$$

it means that we gain one descent, and if we replace it by

$$\sum_{j>m} pq^{j-1}u^j = pu\frac{(qu)^m}{1-qu},$$

we stay with the same number of descents. Translating this principle to the generating functions $f_i(u)$, gives the indicated recursion. We will use this principle in various places again.

Now let

$$F(u) = F(z, v, u) = \sum_{i \geq 0} f_i(u)v^i.$$

Then we get, by summing up,

$$F(u) - f_0(u) = \frac{puvz}{1-qu}F(1) - \frac{puvz}{1-qu}F(uq) + \frac{puz}{1-qu}F(uq) - \frac{puz}{1-qu}f_0(uq).$$

But

$$f_0(u) = \frac{puz}{1-qu} + \frac{puz}{1-qu} f_0(uq)$$

and so

$$F(u) = \frac{puz}{1-qu} + \frac{puvz}{1-qu} F(1) + \frac{puz(1-v)}{1-qu} F(uq).$$

This functional equation can be solved by iterating it: (Technically, this iteration leads to an equation involving $F(u)$ and $F(1)$; after plugging in $u = 1$, the quantity $F(1)$ can be found, and thus in the next step $F(u)$.)[(i)]

$$F(z, v, u) = \frac{\sum_{k\ge1} \frac{(puz)^k (1-v)^k q^{\binom{k}{2}}}{(qu)_k}}{1 - v \sum_{k\ge0} \frac{(pz)^k (1-v)^k q^{\binom{k}{2}}}{(q)_k}}. \tag{2.1}$$

Now we turn to the end heights. We want to obtain the generating function

$$H(z, v, u) = \sum_{n,j,i\ge1} \mathbb{P}[\text{in a word of length } n, \text{ the } j\text{th descent has end height } i] z^n v^j u^i.$$

As indicated already, it can be obtained from $F(z, v, u)$ by substitution: we make one down step, record its height with the u-variable, ($u^j \to pu + pqu^2 + \cdots + pq^{j-1}u^j$) and then attach anything ($1/(1-z)$). This gives the generating function

$$H(z, v, u) = \frac{puvz}{1-qu} [F(z, v, 1) - F(z, v, uq)] \frac{1}{1-z} + \frac{pzu}{1-qu}.$$

The next step is to look at the behaviour for $n \to \infty$. Intuitively, it is quite clear, there should be a limit, since what happens at the lth descent must become independent of letters very far to the right. Indeed, we see that there is a simple pole at $z = 1$, and the generating function

$$\Psi(v, u) = \frac{puv}{1-qu} [F(1, v, 1) - F(1, v, uq)], \tag{2.2}$$

is obtained from $H(z, v, u)$ by dropping the factor $1/(1-z)$, and the irrelevant additive term and replacing $z = 1$. Thus, considering u and v as parameters, there is a dominant simple pole, and its residue (that we call $\Psi(v, u)$) corresponds to the the limit $n \to \infty$, or, what amounts to the same, to an *infinite word*. One could perhaps avoid dealing with n and z and deal with an infinite word directly, as we will do in the second part of this paper, but we would lose information in that way that we can have for free. – The coefficient of v^l in $\Psi(v, u)$ is a probability generating function in the variable u alone. Indeed, it is easy to check that $\Psi(v, 1) = \frac{v}{1-v}$. Explicit expressions become very messy, but at least the instance $l = 1$ (which was studied in (6)) is manageable:

$$\begin{aligned} [v^1]\Psi(v, u) &= \frac{pu}{1-qu} [F(1, 0, 1) - F(1, 0, uq)] \\ &= \frac{pu}{1-qu} \left[\sum_{k\ge1} \frac{p^k q^{\binom{k}{2}}}{(q)_k} - \sum_{k\ge1} \frac{(pqu)^k q^{\binom{k}{2}}}{(q^2u)_k} \right]. \end{aligned}$$

(i) As a referee has pointed out, such functional equations are not always as "smooth" as this one.

From a probability generating function, all moments (of the underlying random variable) can be derived by differentiations, followed by $u = 1$. Let us just do this for the average (first moment):

$$\begin{aligned}
\mathbb{E}[\text{first descent}] &\sim \frac{1}{p} - \frac{\partial}{\partial u}\sum_{k\ge1}\frac{(pqu)^k q^{\binom{k}{2}}}{(q^2u)_k}\Big|_{u=1}\\
&\sim \frac{1}{p} - \sum_{k\ge1}\frac{k(pq)^k q^{\binom{k}{2}}}{(q^2)_k} - \sum_{k\ge1}\frac{(pq)^k q^{\binom{k}{2}}}{(q^2)_k}\sum_{i=2}^{k+1}\frac{q^i}{1-q^i}\\
&\sim \frac{1}{p} - \sum_{k\ge2}\frac{(k-1)p^k q^{\binom{k}{2}}}{(q)_k} - \sum_{k\ge2}\frac{p^k q^{\binom{k}{2}}}{(q)_k}\sum_{i=2}^{k}\frac{q^i}{1-q^i}.
\end{aligned}\tag{2.3}$$

Note that this average was given as

$$\frac{1}{p} - \sum_{h\ge0}(h+1)pq^{2h+1}(-p)_h \tag{2.4}$$

in (6).

For the limiting behaviour of this, as $q \to 1$, we should consider $\lim_{q\to1}(1-q)\mathbb{E}[\text{first descent}]$. It turns out that this tends to a constant, and thus the expectation of the end height of the first descent goes to infinity as $\frac{\text{constant}}{1-q}$:

$$1 - \sum_{k\ge1}\frac{1}{k!}\sum_{i=2}^{k}\frac{1}{i} = 1 - \sum_{k\ge1}\frac{H_k}{k!} + \sum_{k\ge1}\frac{1}{k!} = e - \sum_{k\ge1}\frac{H_k}{k!}.$$

A similar computation can be done for *general* l:

$$\begin{aligned}
\frac{\partial}{\partial u}\Psi(v,u)\big|_{u=1} &= \frac{v}{p}[F(1,v,1) - F(1,v,q)] - v\frac{\partial}{\partial u}F(1,v,uq)\big|_{u=1}\\
&= \frac{v}{p}\frac{1}{1 - v\sum_{k\ge0}\frac{p^k(1-v)^k q^{\binom{k}{2}}}{(q)_k}}\left[\sum_{k\ge1}\frac{p^k(1-v)^k q^{\binom{k}{2}}}{(q)_k} - \sum_{k\ge1}\frac{(pq)^k(1-v)^k q^{\binom{k}{2}}}{(q^2)_k}\right]\\
&\quad - v\frac{1}{1 - v\sum_{k\ge0}\frac{p^k(1-v)^k q^{\binom{k}{2}}}{(q)_k}}\left[\sum_{k\ge1}\frac{k(pq)^k(1-v)^k q^{\binom{k}{2}}}{(q^2)_k} - \sum_{k\ge1}\frac{(pq)^k(1-v)^k q^{\binom{k}{2}}}{(q^2)_k}\sum_{j=2}^{k+1}\frac{q^j}{1-q^j}\right].
\end{aligned}$$

And now

$$\begin{aligned}
\lim_{q\to1}(1-q)\frac{\partial}{\partial u}\Psi(v,u)\big|_{u=1} &= \frac{v}{1-ve^{1-v}}\left[e^{1-v} - 1 - \frac{1}{1-v}\left(e^{1-v} - 1 - (1-v)\right)\right]\\
&\quad + \frac{v}{1-ve^{1-v}}\sum_{k\ge1}\frac{(1-v)^k(H_{k+1}-1)}{(k+1)!}
\end{aligned}$$

$$= -1 + \frac{1}{1-ve^{1-v}} - \frac{v}{1-ve^{1-v}} \sum_{k\ge1} \frac{H_k(1-v)^{k-1}}{k!}.$$

(The coefficient of v^1 is the previous expression).

Now we turn to the *initial heights.* The same approach applies, but the substitution is even simpler. We get

$$G(z,u,v) = vz[F(z,v,u) - F(z,v,uq)]\frac{1}{1-z} + \frac{pzu}{1-qu}.$$

The limit for $n \to \infty$ leads then to the generating function

$$\Phi(v,u) = v[F(1,v,u) - F(1,v,uq)]. \tag{2.5}$$

Let us compute again the instance $l = 1$; the coefficient of v^1 is particularly simple:

$$[v^1]\Phi(v,u) = F(1,0,u) - F(1,0,uq).$$

From this, we find that the average is asymptotic to

$$\frac{\partial}{\partial u}\Bigg[\sum_{k\ge1}\frac{(pu)^k q^{\binom{k}{2}}}{(qu)_k} - \sum_{k\ge1}\frac{(pqu)^k q^{\binom{k}{2}}}{(q^2u)_k}\Bigg]\Bigg|_{u=1}. \tag{2.6}$$

This checks with the expression

$$\sum_{h\ge0}(h+1)pq^{h+1}(-p)_h - \sum_{h\ge0}(2h+1)pq^{h+1}(-p)_h, \tag{2.7}$$

given in (6).

For the limit $\lim_{q\to1}(1-q)\mathbb{E}$ we get the generating function

$$\frac{1}{v(1-ve^{1-v})} - \frac{1}{v(1-v)} - \frac{v}{1-ve^{1-v}}\sum_{k\ge1}\frac{H_k(1-v)^{k-1}}{k!}.$$

The coefficient of v in this is

$$e - 1 - \sum_{k\ge1}\frac{H_k}{k!}.$$

3 Descents: the strict model

Computations are similar; we only give the key steps. The functional equation is

$$F(u) = \frac{puz}{1-qu} + \frac{puvz}{1-qu}F(1) + \frac{pz(1-v)}{q(1-qu)}F(uq).$$

Thus

$$F(z,v,u) = u\sum_{k\ge1}\frac{(pz)^k(1-v)^{k-1}}{(uq)_k}\Big(1 + vF(z,v,1)\Big) = \frac{u\sum_{k\ge1}\frac{(pz)^k(1-v)^k}{(uq)_k}}{1 - v\sum_{k\ge0}\frac{(pz)^k(1-v)^k}{(q)_k}}. \tag{3.1}$$

Also,

$$H(z,v,u)=\left[\frac{puvz}{1-qu}F(z,v,1)-\frac{pvz}{q(1-qu)}F(z,v,uq)\right]\frac{1}{1-z}+\frac{pzu}{1-qu}.$$

Hence

$$\Psi(v,u)=\frac{puv}{1-qu}F(1,v,1)-\frac{pv}{q(1-qu)}F(1,v,uq). \tag{3.2}$$

And

$$\begin{aligned}[v^1]\Psi(v,u)&=\frac{pu}{1-qu}F(1,0,1)-\frac{p}{q(1-qu)}F(1,0,uq)\\&=\frac{pu}{1-qu}\sum_{k\ge1}\frac{p^k}{(q)_k}-\frac{pu}{1-qu}\sum_{k\ge1}\frac{p^k}{(uq^2)_k}.\end{aligned}$$

The expectation can be obtained by differentiation, followed by $u=1$:

$$\frac{1}{p}-\sum_{k\ge1}\frac{p^k}{(q)_k}\sum_{i=2}^{k}\frac{q^i}{1-q^i}. \tag{3.3}$$

The formula given in (6) is

$$\frac{1}{p}-\sum_{h\ge1}\frac{hq^{2h-1}}{(pq)_h}. \tag{3.4}$$

Now, let us compute the limit of $(1-q)\mathbb{E}$, $q\to1$:

$$1-\sum_{k\ge1}\frac{H_k-1}{k!}=e-\sum_{k\ge1}\frac{H_k}{k!}.$$

And in general:

$$\frac{1}{1-ve^{1-v}}-\frac{1}{1-v}-\frac{v}{1-ve^{1-v}}\sum_{k\ge2}\frac{H_k(1-v)^{k-1}}{k!}.$$

Now, for the initial heights, we must consider

$$G(z,v,u)=vz[F(z,v,u)-\frac{1}{q}F(z,v,qu)]\frac{1}{1-z}+\frac{pz}{1-qu},$$

and

$$\Phi(v,u)=v[F(1,v,u)-\tfrac{1}{q}F(1,v,qu)]. \tag{3.5}$$

Furthermore, to look at the first descent,

$$[v^1]\Phi(v,u)=F(1,0,u)-\tfrac{1}{q}F(1,0,qu)=u\sum_{k\ge1}\frac{p^k}{(qu)_k}-u\sum_{k\ge1}\frac{p^k}{(q^2u)_k}.$$

From this, the average, obtained by differentiation, is

$$\begin{aligned}
&1+\sum_{k\ge1}\frac{p^k}{(q)_k}\sum_{i=1}^{k}\frac{q^i}{1-q^i}-\sum_{k\ge1}\frac{p^k}{(q^2)_k}\sum_{i=2}^{k+1}\frac{q^i}{1-q^i}\\
&=1+\sum_{k\ge1}\frac{p^k}{(q)_k}\sum_{i=1}^{k}\frac{q^i}{1-q^i}-\sum_{k\ge2}\frac{p^k}{(q)_k}\sum_{i=2}^{k}\frac{q^i}{1-q^i}\\
&=1+\frac{q}{p}\sum_{k\ge1}\frac{p^k}{(q)_k}\\
&=1+\frac{q}{p}\Big(-1+\frac{1}{(p)_\infty}\Big). \qquad (3.6)
\end{aligned}$$

The last simplification was by (A.6). The version given in (6) is

$$1+\sum_{h\ge1}\frac{hpq^{h-1}}{(pq)_h}-\sum_{h\ge1}\frac{hpq^{2h-1}}{(pq)_h}. \qquad (3.7)$$

The limiting function $\lim_{q\to1}(1-q)\mathbb{E}$ is

$$\frac{v}{1-ve^{1-v}}\Bigg[\sum_{k\ge1}\frac{(1-v)^kH_k}{k!}-\sum_{k\ge1}\frac{(1-v)^k(H_{k+1}-1)}{(k+1)!}\Bigg].$$

4 Ascents: the strict model

First, we consider the case where only $a<b$ is an ascent.

Again, the treatment is very similar to before, so we only give the key steps.

Let $f_i(u)$ be the generating function with $[z^nu^j]f_i(u)$ is the probability that a word of length $n\ge1$ has i ascents, and that the last letter is j.

Here is the recursion for $i\ge1$:

$$f_i(u)=\frac{puz}{1-qu}f_i(1)-\frac{puz}{1-qu}f_i(uq)+\frac{puz}{1-qu}f_{i-1}(uq).$$

The explanation is similar to before: The variable u keeps track of the last letter, and thus, if we replace u^m by

$$\sum_{1\le j\le m}pq^{j-1}u^m=pu\frac{1-(qu)^m}{1-qu},$$

we stay with the same number of ascents, but if we replace u^m by

$$\sum_{j>m}pq^{j-1}u^m=pu\frac{(qu)^m}{1-qu},$$

we gain one descent.

Now let

$$F(u) = F(z, v, u) = \sum_{i \geq 0} f_i(u) v^i.$$

Then we get

$$F(u) = \frac{puz}{1 - qu} + \frac{puz}{1 - qu} F(1) + \frac{puz(v-1)}{1 - qu} F(uq).$$

This functional equation can be solved by iterating it:

$$F(z, v, u) = \frac{\sum_{k \geq 1} \frac{(puz)^k (v-1)^{k-1} q^{\binom{k}{2}}}{(qu)_k}}{1 - \sum_{k \geq 1} \frac{(pz)^k (v-1)^{k-1} q^{\binom{k}{2}}}{(q)_k}}. \tag{4.1}$$

Further,

$$H(z, v, u) = \frac{puzv}{1 - qu} F(z, v, uq) \frac{1}{1 - z} + \frac{puz}{1 - qu},$$

and

$$\Psi(v, u) = \frac{puv}{1 - qu} F(1, v, uq). \tag{4.2}$$

The case of the first accent, i.e., the coefficient of v^1, is thus

$$[v^1]\Psi(v, u) = \frac{pu}{1 - qu} F(1, 0, uq) = \frac{pu}{1 - qu} \frac{\sum_{k \geq 1} \frac{(pqu)^k (-1)^{k-1} q^{\binom{k}{2}}}{(q^2 u)_k}}{1 - \sum_{k \geq 1} \frac{p^k (-1)^{k-1} q^{\binom{k}{2}}}{(q)_k}} = \frac{\sum_{k \geq 2} \frac{(pu)^k (-1)^k q^{\binom{k}{2}}}{(qu)_k}}{\sum_{k \geq 0} \frac{p^k (-1)^k q^{\binom{k}{2}}}{(q)_k}}.$$

Notice that

$$\sum_{k \geq 0} \frac{p^k (-1)^k q^{\binom{k}{2}}}{(q)_k} = (p)_\infty.$$

And therefore the average of the end height of the first ascent tends to

$$\frac{1}{(p)_\infty} \sum_{k \geq 2} \frac{k p^k (-1)^k q^{\binom{k}{2}}}{(q)_k} + \frac{1}{(p)_\infty} \sum_{k \geq 2} \frac{p^k (-1)^k q^{\binom{k}{2}}}{(q)_k} \sum_{i=1}^{k} \frac{q^i}{1 - q^i}. \tag{4.3}$$

This quantity appears in (7) in the form

$$\frac{q}{p} + \frac{1}{(p)_\infty} \sum_{h \geq 0} (h+2) p q^{2h+1} (p)_h. \tag{4.4}$$

Furthermore,

$$G(z, v, u) = vzF(z, v, uq) \frac{1}{1 - z} + \frac{puz}{1 - qu},$$

and

$$\Phi(v, u) = vF(1, v, uq). \tag{4.5}$$

Also,

$$[v^1]\Phi(v,u) = F(1,0,uq) = \frac{1}{(p)_\infty}\sum_{k\ge1}\frac{(pqu)^k(-1)^{k-1}q^{\binom{k}{2}}}{(q^2u)_k}.$$

And therefore the average of the initial height of the first ascent tends to

$$\frac{1}{(p)_\infty}\sum_{k\ge2}\frac{(k-1)p^k(-1)^kq^{\binom{k}{2}}}{(q)_k} + \frac{1}{(p)_\infty}\sum_{k\ge2}\frac{p^k(-1)^kq^{\binom{k}{2}}}{(q)_k}\sum_{i=2}^{k}\frac{q^i}{1-q^i}. \tag{4.6}$$

The version given in the paper (6) is

$$\frac{1}{(p)_\infty}\sum_{h\ge0}(h+1)pq^{2h+1}(p)_h. \tag{4.7}$$

5 Ascents: the weak model

For completeness, we only collect the relevant formulæ here. This brief section is for reference only, and can be skipped, as the methods are the same as before.

$$F(u) = \frac{puz}{1-qu} + \frac{puz}{1-qu}F(1) + \frac{pz(v-1)}{q(1-qu)}F(uq),$$

$$F(z,v,u) = \frac{u\sum_{k\ge1}\frac{(pz)^k(v-1)^{k-1}}{(qu)_k}}{1-\sum_{k\ge1}\frac{(pz)^k(v-1)^{k-1}}{(q)_k}}.$$

$$H(z,v,u) = \frac{pzv}{q(1-qu)}F(z,v,uq)\frac{1}{1-z} + \frac{pz}{1-qu},$$

$$\Psi(v,u) = \frac{pv}{q(1-qu)}F(1,v,uq), \tag{5.1}$$

$$G(z,v,u) = \frac{zv}{q}F(z,v,uq)\frac{1}{1-z} + \frac{pz}{1-qu},$$

$$\Phi(v,u) = \frac{v}{q}F(1,v,uq). \tag{5.2}$$

6 Summary of results from combinatorial analysis

For the reader's convenience, we list where the generating function $\Psi(v,u)$, that we consider to be the most important finding, as it contains all the information, can be found: (Recall that the coefficient of $v^j u^i$ refers to the jth descent (ascent), and the probability that the end height (initial height) is i, in an infinitely long string.)

	Descents (weak)	Descents (strict)	Ascents (strict)	Ascents (weak)
End height	(2.2)	(3.2)	(4.2)	(5.1)
Initial height	(2.5)	(3.5)	(4.5)	(5.2)

The auxiliary functions and the averages can be found in the text.

Part II
Probabilistic Analysis

We will analyze the descents in the strict model with probabilistic tools. We start from an infinite sequence of $\mathsf{geom}(p)$ RVs.

7 Markov chains

In this section, we consider the successive descents as a Markov chain, related to initial and end values of each descent.

Let

$$\begin{aligned}
\pi(i) &:= pq^{i-1},\\
P(i) &:= \sum_{j\geq i}\pi(j) = q^{i-1},\\
I_k &:= \text{ beginning of the } k\text{th descent (initial height)},\\
J_k &:= \text{ end of the } k\text{th descent (end height)},\\
&I_k > J_k,\ I_{k+1}\geq J_k,\ I_k\geq 2.
\end{aligned}$$

By convention, J_0 is the first $\mathsf{geom}(p)$ RV. We have

$$\mathbb{P}[I_2=i_2, J_2=j_2 | I_1=i_1, J_1=j_1] = \sum_{l\geq 0}\ \sum_{j_1\leq k_2\leq\cdots\leq k_l\leq i_2}\pi(k_2)\ldots\pi(k_l)\pi(i_2)\pi(j_2),$$

with the conventions

$$l=0: i_2\equiv j_1,$$

$$l = 1 : i_2 \geq j_1.$$

This is independent of i_1. Set

$$A(a,b,t) := \sum_{a \leq k_1 \leq \cdots \leq k_t \leq b} q^{k_1+\cdots+k_t},$$

and

$$B(a,b) := \sum_{t \geq 0} \left(\frac{p}{q}\right)^t A(a,b,t).$$

Then

$$\mathbb{P}[I_2 = i_2, J_2 = j_2 | J_1 = j_1] = [\![i_2 = j_1]\!]\pi(j_2) + B(j_1, i_2)\pi(i_2)\pi(j_2).$$

But we know that

$$\sum_{0 \leq k_1 \leq \cdots \leq k_t \leq n} q^{k_1+\cdots+k_t} = \frac{1}{(q)_t} \sum_{j=0}^{t} (-1)^j \begin{bmatrix} t \\ j \end{bmatrix} q^{jn+\binom{j+1}{2}} = \begin{bmatrix} n+t \\ t \end{bmatrix}.$$

Recall that $\left[{n \atop m}\right]$ denotes the q-binomial (Gaussian) coefficients.

Therefore

$$\begin{aligned} A(a,b,t) &:= \sum_{a \leq k_1 \leq \cdots \leq k_t \leq b} q^{k_1+\cdots+k_t} \\ &= q^{at} \sum_{0 \leq k_1 \leq \cdots \leq k_t \leq b-a} q^{k_1+\cdots+k_t} \\ &= q^{at} \begin{bmatrix} b-a+t \\ b-a \end{bmatrix}. \end{aligned}$$

We will use

$$\sum_{t \geq 0} \alpha^t \begin{bmatrix} m+t \\ m \end{bmatrix} = \frac{1}{(\alpha)_{m+1}}.$$

Then

$$B(a,b) := \sum_{t \geq 0} \left(\frac{p}{q}\right)^t A(a,b,t) = \frac{(p)_{a-1}}{(p)_b},$$

and

$$\mathbb{P}[I_2 = i_2, J_2 = j_2 | J_1 = j_1] = [\![i_2 = j_1]\!]\pi(j_2) + \frac{(p)_{j_1-1}}{(p)_{i_2}}\pi(i_2)\pi(j_2), \tag{7.1}$$

$$\mathbb{P}[I_2 = i_2 | J_1 = j_1] = [\![i_2 = j_1]\!](1 - q^{i_2-1}) + \frac{(p)_{j_1-1}}{(p)_{i_2}}\pi(i_2)(1 - q^{i_2-1}). \tag{7.2}$$

The transition matrix between I_1 and I_2 is given by

$$\mathbb{P}[I_2 = i_2 | I_1 = i_1] = \sum_{j_1 < i_1} \frac{\pi(j_1)}{1 - P(i_1)} \mathbb{P}[I_2 = i_2 | J_1 = j_1]$$

$$= [\![i_2 < i_1]\!] \frac{\pi(i_2)}{1-P(i_1)}(1-q^{i_2-1}) + \sum_{j_1<i_1, j_1\le i_2} \frac{\pi(j_1)}{1-P(i_1)} \frac{(p)_{j_1-1}}{(p)_{i_2}} \pi(i_2)(1-q^{i_2-1}). \tag{7.3}$$

Let us first check that $\sum_{i_2\ge 2} \mathbb{P}[I_2 = i_2 | I_1 = i_1] = 1$. We have, using (A.8)

$$\begin{aligned}
&\sum_{i_2=2}^{i_1-1} \frac{pq^{i_2-1}}{1-q^{i_1-1}}(1-q^{i_2-1}) + \sum_{i_2\ge 2} \sum_{j_1<i_1, j_1\le i_2} \frac{pq^{j_1-1}}{1-q^{i_1-1}} \frac{(p)_{j_1-1}}{(p)_{i_2}} pq^{i_2-1}(1-q^{i_2-1}) \\
&= \sum_{i_2=2}^{i_1-1} \frac{pq^{i_2-1}}{1-q^{i_1-1}} - \sum_{i_2=2}^{i_1-1} \frac{pq^{i_2-1}}{1-q^{i_1-1}} q^{i_2-1} + \sum_{i_2\ge 2} \frac{p}{1-q^{i_1-1}} \frac{1}{(p)_{i_2}} pq^{i_2-1}(1-q^{i_2-1}) \\
&\quad + \sum_{j_1=2}^{i_1-1} \sum_{i_2\ge j_1} \frac{pq^{j_1-1}}{1-q^{i_1-1}} \frac{(p)_{j_1-1}}{(p)_{i_2}} pq^{i_2-1}(1-q^{i_2-1}) \\
&= \frac{1}{1-q^{i_1-1}}[1-p-q^{i_1-1}] - \sum_{i_2=2}^{i_1-1} \frac{pq^{i_2-1}}{1-q^{i_1-1}} q^{i_2-1} + \frac{p}{1-q^{i_1-1}} + \frac{1}{1-q^{i_1-1}} \sum_{j_1=2}^{i_1-1} pq^{j_1-1} q^{j_1-1} \\
&= 1.
\end{aligned}$$

Now we compute the stationary measure $\varphi(i)$ of this matrix. (see, for instance, Chung (3)). We must have

$$\sum_{i_1} \varphi(i_1) \mathbb{P}[I_2 = i_2 | I_1 = i_1] = \varphi(i_2),$$

which leads to

$$\sum_{i_1>i_2} \varphi(i_1) \frac{pq^{i_2-1}}{1-q^{i_1-1}}(1-q^{i_2-1}) + \sum_{i_1\ge 2} \varphi(i_1) \sum_{j_1<i_1, j_1\le i_2} \frac{pq^{j_1-1}}{1-q^{i_1-1}} \frac{(p)_{j_1-1}}{(p)_{i_2}} pq^{i_2-1}(1-q^{i_2-1}) = \varphi(i_2),$$

or, setting

$$\psi(i) = \frac{\varphi(i)}{pq^{i-1}(1-q^{i-1})},$$

we have

$$\sum_{i_1>i_2} pq^{i_1-1}\psi(i_1) + \sum_{i_1\ge 2} pq^{i_1-1}\psi(i_1) \sum_{j_1<i_1, j_1\le i_2} pq^{j_1-1} \frac{(p)_{j_1-1}}{(p)_{i_2}} = \psi(i_2).$$

After some algebra, we will find that $\psi(i) = \mathsf{constant}$ is a solution of this equation. But it is probabilistically obvious: the stationary distribution is proportional to $pq^{i-1}(1-q^{i-1})$. We have, setting $\psi \equiv 1$,

$$\begin{aligned}
\sum_{i_1>i_2} pq^{i_1-1} + \sum_{i_1\ge 2} pq^{i_1-1} \sum_{j_1<i_1, j_1\le i_2} pq^{j_1-1} \frac{(p)_{j_1-1}}{(p)_{i_2}} &= q^{i_2} + \frac{1}{(p)_{i_2}} \sum_{j_1=1}^{i_2} pq^{j_1-1}(p)_{j_1-1} q^{j_1} \\
&= q^{i_2} + \frac{1}{(p)_{i_2}} pq \sum_{m\ge 0} q^{2m}(p)_m - \frac{1}{(p)_{i_2}} pqq^{2i_2}(p)_{i_2} \sum_{v\ge 0} q^{2v}(pq^{i_2})_v
\end{aligned}$$

$$= q^{i_2} + (pq^{i_2})_\infty + 1 - q^{i_2} - (pq^{i_2})_\infty = 1,$$

as expected. (We used (A.5) twice.)

The stationary distribution is given by

$$f_{\text{stationary}}(i) = \frac{1+q}{q} pq^{i-1}(1-q^{i-1}). \tag{7.4}$$

The stationary distribution generating function is given by

$$G(z) = \frac{p^2(1+q)z^2}{(1-qz)(1-q^2z)},$$

from which the stationary moments are easily derived:

$$\begin{aligned}\mathbb{E}(I) &= \frac{2+q}{1-q^2},\\ \mathbb{E}(I^2) &= \frac{q^3+4q^2+5q+4}{(1-q^2)^2}.\end{aligned}$$

The stationary distribution of the end value is given by

$$\frac{1+q}{q}\sum_{i_1=i_2+1}^{\infty} pq^{i_1-1}(1-q^{i_1-1})pq^{i_2-1}/(1-q^{i_1-1}) = (1+q)pq^{2i_2-2},\ i_2 \ge 1,$$

which sums correctly to 1.

Another transition matrix of interest is given by

$$\mathbb{P}[J_2 = j_2 | J_1 = j_1] = \pi(j_2)[\![j_2 < j_1]\!] + \sum_{i_2 \ge j_1, i_2 > j_2} \frac{(p)_{j_1-1}}{(p)_{i_2}} \pi(i_2)\pi(j_2). \tag{7.5}$$

8 Descent distributions

In this section, we use the Markov chain derived in the previous section to obtain the distribution of initial and end values of first and second descents. For the first descent initial value I_1, for example, with (A.9), we have

$$\begin{aligned}f_1(i_1) := \mathbb{P}[I_1 = i_1] &= \sum_{j_0=1}^{i_1} pq^{j_0-1}\mathbb{P}[I_1 = i_1 | J_0 = j_0]\\ &= pq^{i_1-1}(1-q^{i_1-1}) + \sum_{j_0=1}^{i_1} pq^{j_0-1}\frac{(p)_{j_0-1}}{(p)_{i_1}} pq^{i_1-1}(1-q^{i_1-1})\\ &= pq^{i_1-1}(1-q^{i_1-1})\frac{1}{(p)_{i_1}}.\end{aligned} \tag{8.1}$$

A graph of $f_1(i)$, $q = 0.7$ is given in Figure 1. It is easily checked that

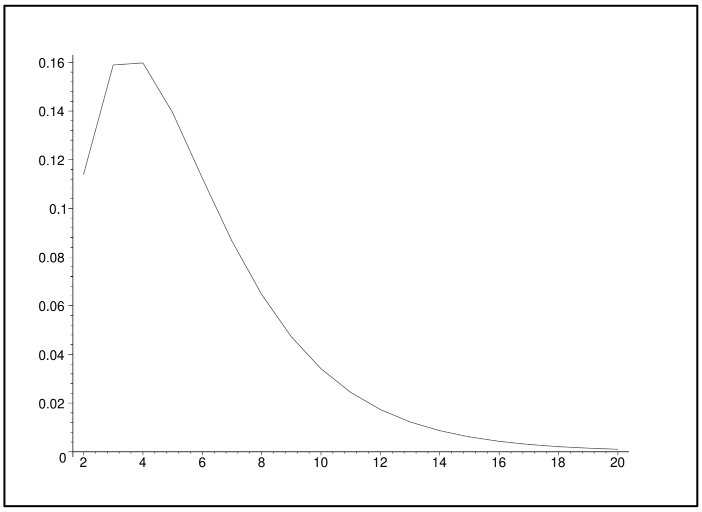


Fig. 1: $f_1(i), q = 0.7$

$$\sum_{i=2}^{\infty} f_1(i) = 1.$$

Indeed

$$\sum_{i \geq 2} pq^{i-1}(1 - q^{i-1}) \frac{1}{(p)_i} = 1, \text{ by (A.4).}$$

A comparison between $f_1(i)$ and the stationary distribution $f_{\text{stationary}}(i)$, $q = 0.7$ is given in Figure 2.

For the end height of the first descent, we have, with (A.9) and (A.10),

$$\begin{aligned}\gamma_1(j_1) := \mathbb{P}[J_1 = j_1] &= \sum_{j_0 > j_1} pq^{j_0-1}pq^{j_1-1} + \sum_{i_1 > j_1} \sum_{j_0=1}^{i_1} pq^{j_0-1} \frac{(p)_{j_0-1}}{(p)_{i_1}} pq^{i_1-1}pq^{j_1-1} \\ &= pq^{j_1-1}\left[\frac{1}{(p)_\infty} - \frac{1}{(p)_{j_1}}\right]. \end{aligned} \tag{8.2}$$

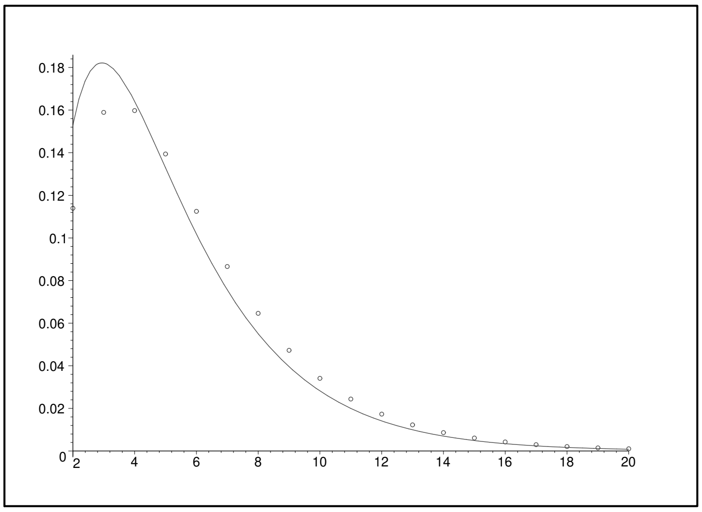
Of course $\sum_{j \geq 1} \gamma_1(j) = 1$, by (A.10).

Note that we also have, with (A.10),

$$\gamma_1(j_1) = \sum_{i_1=j_1+1}^{\infty} f_1(i_1) \frac{\pi(j_1)}{1 - P(i_1)} = \sum_{i_1=j_1+1}^{\infty} pq^{i_1-1}(1-q^{i_1-1}) \frac{1}{(p)_{i_1}} \frac{pq^{j_1-1}}{1 - q^{i_1-1}} = pq^{j_1-1}\left[\frac{1}{(p)_\infty} - \frac{1}{(p)_{j_1}}\right].$$

The distribution of the second descent (I_2, J_2) is given by the square of the Markov matrix, i.e.

$$\gamma_2(i_2, j_2) = \sum_{j_0 \geq 1} pq^{j_0-1} \sum_{j_1 \geq 1} \mathbb{P}[J_1 = j_1 | J_0 = j_0]\mathbb{P}[I_2 = i_2, J_2 = j_2 | J_1 = j_1].$$



$\circ$: $f_1(i)$
line : $f_{\text{stationary}}(i)$

Fig. 2: Comparison between $f_1(i)$ and $f_{\text{stationary}}(i)$, $q = 0.7$

The distribution of the next descents is related to successive powers of the transition matrix.

A compact form for $\gamma_2(i_2, j_2)$ is given below.

Now, the distribution of the initial value of the *second* descent I_2 is given by (we use (A.3), (A.9), (7.3))

$$
\begin{aligned}
f_2(i_2) := \mathbb{P}[I_2 = i_2] &= \sum_{i_1=2}^{\infty} f_1(i_1)\mathbb{P}[I_2 = i_2 | I_1 = i_1] = \sum_{i_1=2}^{\infty} pq^{i_1-1}(1-q^{i_1-1})\frac{1}{(p)_{i_1}}\mathbb{P}[I_2 = i_2 | I_1 = i_1] \\
&= \sum_{i_1 > i_2} pq^{i_1-1}(1-q^{i_1-1})\frac{1}{(p)_{i_1}}\frac{pq^{i_2-1}}{1-q^{i_1-1}}(1-q^{i_2-1}) \\
&+ \sum_{i_1 \ge 2} pq^{i_1-1}(1-q^{i_1-1})\frac{1}{(p)_{i_1}} \sum_{j_1 < i_1, j_1 \le i_2} \frac{pq^{j_1-1}}{1-q^{i_1-1}}\frac{(p)_{j_1-1}}{(p)_{i_2}} pq^{i_2-1}(1-q^{i_2-1}) \\
&= pq^{i_2-1}(1-q^{i_2-1})\frac{1}{(p)_{i_2-1}(1-pq^{i_2-1})}\left[\frac{1}{(pq^{i_2})_\infty} - 1\right] \\
&+ pq^{i_2-1}(1-q^{i_2-1})\frac{1}{(p)_{i_2}}\sum_{j_1=1}^{i_2} pq^{j_1-1}(p)_{j_1-1}\sum_{i_1>j_1} pq^{i_1-1}\frac{1}{(p)_{i_1}} \\
&= pq^{i_2-1}(1-q^{i_2-1})\frac{1}{(p)_{i_2-1}(1-pq^{i_2-1})}\left[\frac{1}{(pq^{i_2})_\infty} - 1\right] \\
&+ pq^{i_2-1}(1-q^{i_2-1})\frac{1}{(p)_{i_2}}\sum_{j_1=1}^{i_2} pq^{j_1-1}\frac{1}{1-pq^{j_1-1}}\left[\frac{1}{(pq^{j_1})_\infty} - 1\right]
\end{aligned}
$$

$$
\begin{aligned}
&= pq^{i_2-1}(1-q^{i_2-1})\frac{1}{(p)_{i_2}}\left[\frac{1}{(pq^{i_2})_\infty}-1\right] \\
&+ pq^{i_2-1}(1-q^{i_2-1})\frac{1}{(p)_{i_2}}\sum_{j_1=1}^{i_2} pq^{j_1-1}\left[\frac{1}{(pq^{j_1-1})_\infty}-\frac{1}{1-pq^{j_1-1}}\right] \\
&= pq^{i_2-1}(1-q^{i_2-1})\left[\frac{1}{(p)_\infty}-\frac{1}{(p)_{i_2}}\right] \\
&+ p^2q^{i_2-1}(1-q^{i_2-1})\frac{1}{(p)_{i_2}}\frac{1}{(p)_\infty}\sum_{j_1=0}^{i_2-1} q^{j_1}(p)_{j_1} - p^2q^{i_2-1}(1-q^{i_2-1})\frac{1}{(p)_{i_2}}\sum_{j_1=0}^{i_2-1}\frac{q^{j_1}}{1-pq^{j_1}} \\
&= pq^{i_2-1}(1-q^{i_2-1})\left[\frac{1}{(p)_\infty}-\frac{1}{(p)_{i_2}}\right] \\
&+ pq^{i_2-1}(1-q^{i_2-1})\frac{1}{(p)_{i_2}}\frac{1}{(p)_\infty}[1-(p)_{i_2}] - p^2q^{i_2-1}(1-q^{i_2-1})\frac{1}{(p)_{i_2}}\sum_{j_1=0}^{i_2-1}\frac{q^{j_1}}{1-pq^{j_1}} \\
&= pq^{i_2-1}(1-q^{i_2-1})\frac{1}{(p)_{i_2}}\left[\frac{1}{(p)_\infty}-1\right] - p^2q^{i_2-1}(1-q^{i_2-1})\frac{1}{(p)_{i_2}}\sum_{j_1=0}^{i_2-1}\frac{q^{j_1}}{1-pq^{j_1}}.
\end{aligned}
$$

Note that this gives an explicit expression for $\gamma_2(i_2, j_2)$:

$$\gamma_2(i_2, j_2) = f_2(i_2)\frac{\pi(j_2)}{1-P(i_2)}.$$

A comparison between $f_2(i)$ and $f_{\text{stationary}}(i)$, $q = 0.7$ is given in Figure 3. The convergence to the stationary distribution is quite fast.

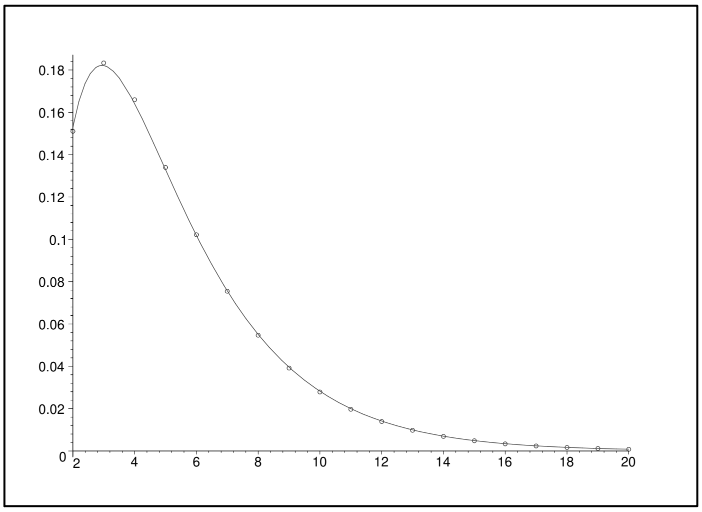
Let us check that $\sum_{i=2}^{\infty} f_2(i) = 1$. We have, using (A.8) and (A.1)

$$
\begin{aligned}
&\sum_{i=2}^{\infty}\left[pq^{i-1}(1-q^{i-1})\frac{1}{(p)_i}\left[\frac{1}{(p)_\infty}-1\right] - p^2q^{i-1}(1-q^{i-1})\frac{1}{(p)_i}\sum_{j_1=0}^{i-1}\frac{q^{j_1}}{1-pq^{j_1}}\right] \\
&= \frac{1}{(p)_\infty} - 1 - \sum_{j_1\ge 1} p^2\frac{q^{j_1}}{1-pq^{j_1}}\sum_{i\ge j_1+1} q^{i-1}(1-q^{i-1})\frac{1}{(p)_i} - p^2\frac{1}{1-p}\sum_{i\ge 2} q^{i-1}(1-q^{i-1})\frac{1}{(p)_i} \\
&= \frac{1}{(p)_\infty} - 1 - \sum_{j_1\ge 1} p^2\frac{q^{j_1}}{1-pq^{j_1}}\frac{q^{j_1}}{p(p)_{j_1}} - p^2\frac{1}{1-p}\frac{1}{p} = 1.
\end{aligned}
$$

9 The moments of descent parameters

The first moments of the first and second descents initial values are analyzed here by intensive use of some combinatorial identities.

We derive the mean of the initial value of the first descent I_1,



$\circ$: $f_2(i)$
line : $f_{\text{stationary}}(i)$

Fig. 3: Comparison between $f_2(i)$ and $f_{\text{stationary}}(i)$, $q = 0.7$

$$\begin{aligned}\mathbb{E}(I_1) &= \sum_{i=2}^{\infty} f_1(i)i = \sum_{i\geq 2} pq^{i-1}(1-q^{i-1})\frac{1}{(p)_i}i \\ &= p\sum_{i\geq 1}\frac{q^{i-1}i}{(pq)_i} - p\sum_{i\geq 1}\frac{q^{2i-1}i}{(pq)_i} + p\sum_{i\geq 1}\frac{q^i}{(p)_{i+1}}(1-q^i) \\ &= p\sum_{i\geq 1}\frac{q^{i-1}i}{(pq)_i} - p\sum_{i\geq 1}\frac{q^{2i-1}i}{(pq)_i} + 1 \text{ by (A.4).}\end{aligned} \tag{9.1}$$

This is Theorem 3 in (6). However, we want to show now independently that this coincides with the closed form obtained earlier as (3.6). This is done in Appendix (B).

More generally higher moments are given by

$$\mathbb{E}(I_1^k) = \sum_{h\geq 1} pq^h(1-q^h)(h+1)^k + \sum_{i_1=2}^{\infty}\sum_{j_0=1}^{i_1} pq^{j_0-1}(p)_{j_0-1}\frac{pq^{i_1-1}}{(p)_{i_1}}(1-q^{i_1-1})i_1^k.$$

Let us look at the modified second moment of I_1 (again we use (A.3), (A.1) for simplifications):

$$
\begin{aligned}
\mathbb{E}(I_1) &= \sum_{i=2}^{\infty} f_1(i)i(i+1)/2 = p\sum_{i\geq 0}\frac{q^i}{(p)_{i+1}}(1-q^i)(i+1)(i+2)/2\\
&= p\sum_{0\leq k\leq h}\sum_{i\geq h}\frac{q^i}{(p)_{i+1}}(1-q^i)\\
&= \frac{1}{q}\sum_{0\leq k\leq h}\frac{q^h}{(pq)_h} - \frac{p}{q}\sum_{0\leq k\leq h}\frac{q^{2h}}{(pq)_h}\\
&= \frac{1}{q}\sum_{k\geq 0}\frac{q^k}{(pq)_k}\sum_{h\geq 0}\frac{q^h}{(pq^{k+1})_h} - \frac{p}{q}\sum_{k\geq 0}\frac{q^{2k}}{(pq)_k}\sum_{h\geq 0}\frac{q^{2h}}{(pq^{k+1})_h}\\
&= \frac{1}{q}\sum_{k\geq 0}\frac{q^k}{(pq)_k}\left[\frac{1}{pq^k(pq^{k+1})_\infty}+1-\frac{1}{pq^k}\right] - \frac{p}{q}\sum_{k\geq 0}\frac{q^{2k}}{(pq)_k}\left[\frac{1}{pq^{2k}}\frac{1}{(pq^{k+1})_\infty}-\frac{1}{pq^{2k}}-\frac{1}{pq^{k-1}}+1\right]\\
&= \frac{1}{q}\sum_{k\geq 0}\frac{1}{(pq)_k}\left[\frac{1}{p(pq^{k+1})_\infty}+q^k-\frac{1}{p}\right] - \frac{1}{q}\sum_{k\geq 0}\frac{1}{(pq)_k}\left[\frac{1}{(pq^{k+1})_\infty}-1-q^{k+1}+pq^{2k}\right]\\
&= \frac{1}{q}\sum_{k\geq 0}\frac{1}{(pq)_k}\left[\frac{1}{p(pq^{k+1})_\infty}+q^k-\frac{1}{p}-\frac{1}{(pq^{k+1})_\infty}+1+q^{k+1}-pq^{2k}\right]\\
&= \frac{1}{q}\sum_{k\geq 0}\frac{1}{(pq)_k}\left[\frac{q}{p(pq^{k+1})_\infty}-\frac{q}{p}+q^k(1+q)-pq^{2k}\right]\\
&= \frac{1}{p}\sum_{k\geq 0}\frac{1}{(pq)_k}\left[\frac{1}{(pq^{k+1})_\infty}-1\right]+\frac{1+q}{q}\sum_{k\geq 0}\frac{q^k}{(pq)_k}-\frac{p}{q}\sum_{k\geq 0}\frac{q^{2k}}{(pq)_k}\\
&= \frac{1}{p}\sum_{k\geq 0}\left[\frac{1}{(pq)_\infty}-\frac{1}{(pq)_k}\right]+\frac{1+q}{q}\left[\frac{1}{p(pq)_\infty}+1-\frac{1}{p}\right]-\frac{p}{q}\left[\frac{1}{p}\frac{1}{(pq)_\infty}-\frac{1}{p}-\frac{q}{p}+1\right]\\
&= \frac{1}{p}\sum_{k\geq 0}\left[\frac{1}{(pq)_\infty}-\frac{1}{(pq)_k}\right]+\frac{2}{p(pq)_\infty}+3-\frac{2}{p}\\
&= \frac{1}{p}\lim_{t\to 1}\left[\frac{1}{(pq)_\infty}\frac{1}{1-t}-\sum_{k\geq 0}\frac{t^k}{(pq)_k}\right]+\frac{2}{p(pq)_\infty}+3-\frac{2}{p}.
\end{aligned}
$$

In Appendix (C), we show that

$$\lim_{t\to 1}\left[\frac{1}{(pq)_\infty}\frac{1}{1-t}-\sum_{k\geq 0}\frac{t^k}{(pq)_k}\right] = \frac{1}{(pq)_\infty}\sum_{k\geq 1}\frac{[1-(p)_k]q^k}{1-q^k}.$$

So finally:

$$\mathbb{E}(I_1) = \frac{1}{p(pq)_\infty}\sum_{k\geq 1}\frac{[1-(p)_k]q^k}{1-q^k}+\frac{2}{p(pq)_\infty}+3-\frac{2}{p}.$$

From this, the variance can be stated as:

$$2\mathbb{E}\mathbb{E}(I_1) - 2\mathbb{E}(I_1) - \mathbb{E}^2(I_1).$$

Now we turn to the second descent. The mean $\mathbb{E}(I_2)$ of the initial value of the second descent is given by (we use (3.6) and (9.1)

$$\begin{aligned}
\mathbb{E}(I_2) = \sum_{i=2}^{\infty} f_2(i)i &= p\left[\frac{1}{(p)_\infty} - 1\right]\sum_{i\ge1} iq^{i-1}(1-q^{i-1})\frac{1}{(p)_i} - \sum_{i\ge1} ip^2q^{i-1}(1-q^{i-1})\frac{1}{(p)_i}\sum_{j=0}^{i-1}\frac{q^j}{1-pq^j}\\
&= p\left[\frac{1}{(p)_\infty} - 1\right]\sum_{i\ge0}(i+1)q^i(1-q^i)\frac{1}{(p)_{i+1}} - p^2\sum_{j\ge0}\frac{q^j}{1-pq^j}\sum_{i\ge j}(i+1)q^i(1-q^i)\frac{1}{(p)_{i+1}}\\
&= \left[\frac{1}{(p)_\infty} - 1\right]\left[1 - \frac{q}{p} + \frac{q}{p(p)_\infty}\right] - p\sum_{j\ge0}\frac{pq^j}{1-pq^j}\sum_{i\ge j}(i+1)q^i(1-q^i)\frac{1}{(p)_{i+1}}\\
&= \left[\frac{1}{(p)_\infty} - 1 - \sum_{j\ge0}\frac{pq^j}{1-pq^j}\right]\left[1 - \frac{q}{p} + \frac{q}{p(p)_\infty}\right] + p\sum_{j\ge0}\frac{pq^j}{1-pq^j}\sum_{0\le i<j}(i+1)q^i(1-q^i)\frac{1}{(p)_{i+1}}.
\end{aligned}$$

The last inner sum is computed in Appendix (D) as S_0. This gives a first expression for the mean:

$$\begin{aligned}
\mathbb{E}(I_2) &= \frac{1}{p}\left[\frac{1}{(p)_\infty} - 1 - \sum_{j\ge0}\frac{pq^j}{1-pq^j}\right]\left[2p - 1 + \frac{q}{(p)_\infty}\right]\\
&+ \frac{1}{p}\sum_{j\ge0}\frac{pq^j}{1-pq^j}\left[2p - 1 + \frac{q}{(p)_{j+1}} - \frac{p(1+q)q^j}{q(p)_{j+1}} + \frac{p^2q^{2j}}{q(p)_j} - \frac{pjq^j}{(p)_{j+1}} + \frac{jp^2q^{2j}}{(p)_{j+1}}\right]\\
&= \frac{1}{p}\left[\frac{1}{(p)_\infty} - 1\right]\left[2p - 1 + \frac{q}{(p)_\infty}\right]\\
&+ \frac{1}{p}\sum_{j\ge0}\frac{pq^j}{1-pq^j}\left[\frac{q}{(p)_{j+1}} - \frac{q}{(p)_\infty} - \frac{p(1+q)q^j}{q(p)_{j+1}} + \frac{p^2q^{2j}}{q(p)_j} - \frac{pjq^j}{(p)_{j+1}} + \frac{jp^2q^{2j}}{(p)_{j+1}}\right]\\
&= \frac{1}{p}\left[\frac{1}{(p)_\infty} - 1\right]\left[2p - 1 + \frac{q}{(p)_\infty}\right]\\
&+ \frac{1}{p}\sum_{j\ge0}\frac{pq^j}{1-pq^j}\left[\frac{q}{(p)_{j+1}} - \frac{q}{(p)_\infty} - \frac{p(1+q)q^j}{q(p)_{j+1}} - \frac{pjq^j}{(p)_{j+1}} + \frac{jp^2q^{2j}}{(p)_{j+1}}\right] + \frac{1}{p}\sum_{j\ge0}\frac{p^3q^{3j}}{q(p)_{j+1}}.
\end{aligned}$$

The last sum can be simplified, using S_3 from Appendix (D).

Hence

$$\mathbb{E}(I_2) = \frac{1}{p}\left[\frac{1}{(p)_\infty} - 1\right]\left[2p - 1 + \frac{q}{(p)_\infty}\right]$$

$$+\frac{1}{p}\sum_{j\geq 0}\frac{pq^j}{1-pq^j}\left[\frac{q}{(p)_{j+1}}-\frac{q}{(p)_\infty}-\frac{p(1+q)q^j}{q(p)_{j+1}}\right]-\frac{1}{p}\sum_{j\geq 0}\frac{jp^2q^{2j}}{(p)_{j+1}}$$
$$+\frac{1}{p}\left[\frac{(q)_3q^2}{(p)_\infty}-q(q-p)(q)_3-(q-p)(q)_3-\frac{p^2(q-p)(1-q^3)}{q}\right].$$

Again, using S_4 from Appendix (D), we obtain

$$\begin{aligned}\mathbb{E}(I_2) &= \frac{1}{p}\left[\frac{1}{(p)_\infty}-1\right]\left[2p-1+\frac{q}{(p)_\infty}\right]\\
&+\frac{1}{p}\sum_{j\geq 0}\frac{pq^j}{1-pq^j}\left[\frac{q}{(p)_{j+1}}-\frac{q}{(p)_\infty}-\frac{p(1+q)q^j}{q(p)_{j+1}}\right]\\
&-\frac{1}{p^2q}\frac{1}{(pq)_\infty}\sum_{k\geq 1}\frac{[1-(p)_k]q^k}{1-q^k}+\frac{1}{p^2}\left[\frac{1}{p(pq)_\infty}+1-\frac{1}{p}\right]-\frac{1}{pq}\left[\frac{1}{p(pq)_\infty}-\frac{1}{p}-\frac{q}{p}+1\right]\\
&+\frac{1}{p}\left[\frac{(q)_3q^2}{(p)_\infty}-q(q-p)(q)_3-(q-p)(q)_3-\frac{p^2(q-p)(1-q^3)}{q}\right].\end{aligned}$$

Although this computation was already quite involved, we could get all cross-moments from (7.3), but only with considerable effort.

10 Markov Chains and Sojourn times

In this section, we obtain the asymptotic distribution of the number of descents initial values in some interval.

10.1 General asymptotic distribution

Let X_i, $i=1,\dots,m$ be an ergodic Markov chain (MC) and A be a subset of states. Assume that the MC is stationary and set $x_i := [\![X_i \in A]\!]$, with $M := \mathbb{E}(x_i)$. The number D of times the MC is in A on $1,\dots,m$ is such that, by standard theory of ergodic Markov chains,

$$\mathbb{E}(D) = mM,$$
$$\mathbb{V}(D) = \mathbb{E}\Bigg[\sum_{i=1}^{m}(x_i-M)^2+2\sum_{i=1}^{m-1}\sum_{j=i+1}^{m}(x_i-M)(x_j-M)\Bigg].$$

If

$$\mathbb{E}[(x_i-M)(x_j-M)]=0,\quad j\geq i+2,$$

then, setting $B:=\mathbb{E}(x_ix_{i+1})$, we obtain

$$\mathbb{V}(D)=mM(1-M)+2(m-1)(B-M^2).$$

But we have a central limit theorem for MC (again, see Chung,(3)). This gives

$$\frac{D-\mathbb{E}(D)}{\sqrt{\mathbb{V}(D)}}\sim\mathcal{N}(0,1),\quad m\to\infty.$$

10.2 Number of descents values in some interval

Here, the states of the MC are couples Y_tY_{t+1} with geometric distribution. Also the MC starts with the stationary distribution $pq^{i-1}pq^{j-1}$, hence $m = n-1$. And the couples Y_tY_{t+1}, Y_uY_{u+1} are independent as soon as $u > t+1$.

If we are interested in the asymptotic (gaussian) distribution of the number of descents initial values in some interval $[\tau, \tau+\Delta]$, we compute

$$M = \sum_{i=\tau}^{\tau+\Delta} pq^{i-1}(1-q^{i-1}),$$
$$B = \sum_{i=\tau+1}^{\tau+\Delta} pq^{i-1} \sum_{j=\tau}^{i-1} pq^{j-1}(1-q^{j-1}).$$

Of course, an explicit formula could be written for M resp. B.

11 Permutations

Starting from n $\mathsf{geom}(p)$ RVs, as $q \to 1$, we can derive the asymptotic properties of first and second descents, in a large permutation. We will consider large size ($n \to \infty$) permutations of $\{1, \dots, n\}$, or n-permutations for short. It is well known that all rank statistics of an n-permutation can be derived from the corresponding ones of a sequence of n $\mathsf{geom}(p)$ RVs as $q \to 1$. But, of course, it is not possible to directly deduce the moments of the beginning (initial height) of the first descent of a n-permutation from the corresponding moments of the geometric variables statistic, as $q \to 1$. However, the asymptotic *distribution* of this RV can be derived as follows. Let $q = 1-\varepsilon$. This gives an asymptotic distribution function for each geom RV K ($\varepsilon \to 0$)

$$1 - q^i \sim F(i) := 1 - e^{-i\varepsilon}.$$

Set $U = F(K)$; U is asymptotically distributed as a $\mathsf{uniform}[0,1]$ RV. Let us consider n $\mathsf{geom}(p)$ RVs $K_1, \dots, K_n$. If we scale the corresponding U variables with n (i.e. multiply by n), take the integer part, we have asymptotically a permutation on n, for large n. More precisely the rank of U_i is the value of the ith element of the permutation. This gives (asymptoticaly, we consider u as a continuous variable)

$$u = 1 - e^{-i\varepsilon}, \quad du = \varepsilon(1-u)di, \quad i = -\ln(1-u)/\varepsilon.$$

Set $g(u) := f_1(i)$ as given by (8.1). Using n geometric RVs instead of an infinite sequence of ones introduces only an exponentially small error.

For instance, this leads for the initial height of the first descent in a large permutation, using Euler-Maclaurin, to

$$\mathbb{E}(I_1) \sim n \int_0^1 g(u)u \frac{du}{\varepsilon(1-u)}, \quad \varepsilon \to 0.$$

We have by simple algebra,

$$g(u) \sim \varepsilon(1-u)u + e^u u(1-e^{-u})\varepsilon(1-u), \quad \varepsilon \to 0$$

and the asymptotic density of I_1/n is given by (we multiply by $di = du/(\varepsilon(1-u))$)

$$h_1(u) = u[1 + e^u(1 - e^{-u})] = ue^u = u\sum_{j=0}^{\infty}\frac{u^j}{j!}.$$

But this has a clear direct probabilistic interpretation: if $U_{j+1} = u$, the probability

$$\mathbb{P}[U_1 \le U_2 \le \cdots \le U_j \le U_{j+1} \ge U_{j+2}] = u\frac{u^j}{j!}. \tag{11.1}$$

Again, as n is large, we can use an infinite summation on j, with exponentially small error terms.

Note that

$$\int_0^1 h_1(u)du = 1,$$

as expected, and

$$\mathbb{E}(I_1) \sim n\int_0^1 h_1(u)udu = n(e-2) := nE_{1,0}, \quad n \to \infty,$$

which conforms to Theorem 9 of (6). All moments can be derived. For instance

$$\mathbb{E}(I_1^2) \sim n^2\int_0^1 h_1(u)u^2du = n^2(6-2e) := n^2E_{2,0},$$
$$\mathbb{E}(I_1^3) \sim n^3E_{3,0}, \text{ with } E_{3,0} = 9e - 24.$$

A general expression can be derived as follows. We have (we drop the second index)

$$E_k = \int_0^1 h_1(u)u^kdu = e - (k+1)E_{k-1}, \quad E_0 = 1.$$

This first order recursion can be solved by iteration:

$$E_k = (k+1)!(-1)^k\left[e\sum_{i=1}^{k}\frac{(-1)^i}{(i+1)!} + 1\right]. \tag{11.2}$$

By convention, $E_{-1} = \int_0^1 e^u du = e - 1$. Note that E_k is a linear function of e, that we will write as

$$E_k := D_{k,0} + D_{k,1}e.$$

We could derive similar expressions for J_1.

The errors terms, as $n \to \infty$ can be derived as follows. The rank R of U_{j+1} is such that $R - (j+2)$ is $\mathsf{Binomial}(n-(j+2), u)$, if U_{j+1} has value u and corresponds to the first descent. Indeed, j values are already below u (on the left) and one on the right. If d values among the $n-(j+2)$ remaining ones are below u, then U_{j+1} possesses the rank $d + j + 2$. So the (conditioned on u, j) moments of R/n are given by the characteristic function

$$F := \left[1 + u\left(e^{\mathbf{i}\theta/n} - 1\right)\right]^{n-(j+2)} e^{\mathbf{i}\theta(j+2)/n},$$

from which we derive

$$\mathbb{E}(R/n|j) \sim u + \frac{-u(u-1)(j+2)}{n},$$
$$\mathbb{E}(R^2/n^2|j) \sim u^2 + \frac{-u(u-1)(5+2j)}{n}.$$

With (11.1), we multiply by $u\frac{u^j}{j!}$, and sum on j, this gives

$$\sum_{j=0}^{\infty} u\frac{u^j}{j!}\mathbb{E}(R/n|j) \sim h_1(u)u + \frac{-e^u u(u-1)(u+2)}{n},$$
$$\sum_{j=0}^{\infty} u\frac{u^j}{j!}\mathbb{E}(R^2/n^2|j) \sim h_1(u)u^2 + \frac{-e^u u^2(u-1)(2u+5)}{n}.$$

Hence, integrating on $u \in [0..1]$, we finally obtain

$$\mathbb{E}(I_1/n) \sim E_{1,0} + \frac{e-2}{n},$$
$$\mathbb{E}(I_1^2/n^2) \sim E_{2,0} + \frac{20-7e}{n}.$$

The first expression fits with Theorem 9 in (6); the second moment wasn't computed in this paper.

As the first and second descent I_1, I_2 are not independent, their joint moments are of some interest. They can be asymptotically computed as follows. First, from (7.3), $\mathbb{P}[I_2 = i_2 | I_1 = i_1]$ leads to

$$\begin{aligned} g_1(u_2, u_1) &= [\![u_2 < u_1]\!]\varepsilon(1-u_2)\frac{u_2}{u_1} + \int_0^{\min(u_1,u_2)} \frac{dv_1}{\varepsilon(1-v_1)}\varepsilon\frac{1-v_1}{u_1}e^{u_2-v_1}\varepsilon u_2(1-u_2) \\ &= [\![u_2 < u_1]\!]\varepsilon(1-u_2)\frac{u_2}{u_1} + e^{u_2}\varepsilon u_2\frac{1-u_2}{u_1}(1-e^{-\min(u_1,u_2)}). \end{aligned}$$

This leads to the Markov kernel

$$h(u_2, u_1) = \frac{g_1(u_2,u_1)}{\varepsilon(1-u_2)} = [\![u_2 < u_1]\!]\frac{u_2}{u_1} + e^{u_2}u_2\frac{1}{u_1}(1-e^{-\min(u_1,u_2)}). \tag{11.3}$$

This can also be derived from the $\mathsf{uniform}[0,1]$ random variables properties.

The joint moment $\mathbb{E}(I_1^k I_2^l)$ is asymptotically given given by $n^2 E_{k,l}$, with

$$\begin{aligned} E_{k,l} &= \int_0^1\int_0^1 h_1(u_1)h(u_2,u_1)u_1^k u_2^l du_1 du_2 \\ &= \int_0^1 u_1 e^{u_1}\int_0^1 u_1^k u_2^l \frac{1}{u_1}\Big[[\![u_2 < u_1]\!]u_2 + e^{u_2}u_2(1-e^{-\min(u_1,u_2)})\Big]du_1 du_2. \end{aligned}$$

The first values of $E_{k,l}$ are given in Table 1.

$k \diagdown l$	0	1	2
0	1	e^2-e-4	$-2e^2-e+18$
1	$e-2$	$-7e/2+10$	$20e-54$
2	$6-2e$	$e^2+32e/3-36$	$-2e^2-235e/3+228$

Tab. 1: $E_{k,l}$

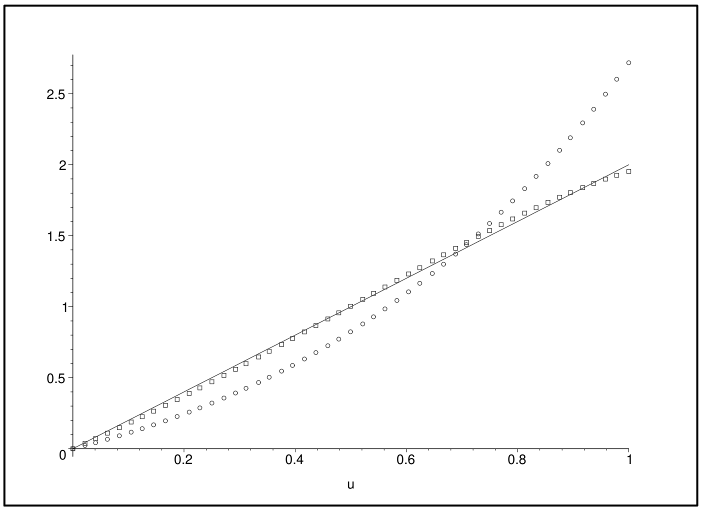
Let us denote by $h_i(u)$ the asymptotic density of I_i/n. For instance,

$$h_2(u) = \int_0^1 h_1(u_1)h(u,u_1)du_1 = u[e^{1+u} - ue^u - e^u] = ue^u[-1-u+e],$$

and similarly for $h_3(u)$. This gives eventually

$$h_3(u) = \int_0^1 h_2(u_2)h(u_3,u_2)du_2 = ue^u[(u+u^2/2) + e(-u-2) + e^2].$$

The convergence to the stationary distribution $h(u) = 2u$ is very fast. It would be interesting to have a precise rate of convergence, but we have not pursued this approach so far. $h(u)$ can be derived from (7.4), or directly by considering two successive RVs U_k. Figure 4 gives $h(u), h_1(u), h_2(u)$



$\circ$: $h_1(u)$
$\square$: $h_2(u)$
line : $h(u)$

Fig. 4: $h(u), h_1(u), h_2(u)$

The difference $h_3(u) - h(u)$ is given in Figure 5.

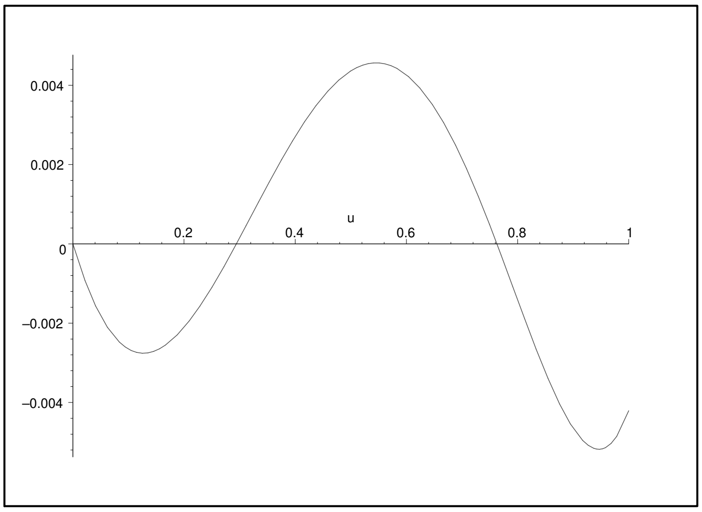


Fig. 5: $h_3(u) - h(u)$

A general expression for $h_i(u)$ can be computed as follows: we observe that $h_i(u)$ is of the form

$$h_i(u) = ue^u \sum_{k=0}^{i-1} e^k P_{i,k}(u),$$

where

$$P_{i,k}(u) = \sum_{l=0}^{i-k-1} P_{i,k,l} u^l, \quad P_{1,0}(u) = 1.$$

But, setting $E_l := D_{l,0} + D_{l,1}e$, from (11.2) and (11.3),

$$\begin{aligned}
h_i(u) &= \int_0^1 h_{i-1}(v) h(u,v) dv \\
&= \int_0^1 h_{i-1}(v) \frac{1}{v} \left[[\![u < v]\!] u + e^u u (1 - e^{-\min(v,u)}) \right] dv \\
&= \int_0^u h_{i-1}(v) \frac{1}{v} e^u u (1 - e^{-v}) dv + \int_u^1 h_{i-1}(v) \frac{1}{v} e^u u dv \\
&= \int_0^1 h_{i-1}(v) \frac{1}{v} e^u u dv - \int_0^u h_{i-1}(v) \frac{1}{v} e^u u e^{-v} dv \\
&= e^u u \left[\sum_{k=0}^{i-2} e^k \int_0^1 P_{i-1,k}(v) e^v dv - \sum_{k=0}^{i-2} e^k \int_0^u P_{i-1,k}(v) dv \right] \\
&= e^u u \left[\sum_{k=0}^{i-2} e^k \sum_{l=0}^{i-k-2} P_{i-1,k,l} [D_{l-1,0} + D_{l-1,1} e] - \sum_{k=0}^{i-2} e^k \sum_{l=0}^{i-k-2} P_{i-1,k,l} \frac{u^{l+1}}{l+1} \right].
\end{aligned}$$

This leads to the following recurrence

$$
\begin{aligned}
P_{i,i-1}(u) &= 1,\\
P_{i,0,0} &= \sum_{l=0}^{i-2} P_{i-1,0,l} D_{l-1,0},\\
P_{i,k,0} &= \sum_{l=0}^{i-k-2} P_{i-1,k,l} D_{l-1,0} + \sum_{l=0}^{i-k-1} P_{i-1,k-1,l} D_{l-1,1}, \quad k = 1,\dots,i-2,\\
P_{i,k,l} &= -P_{i-1,k,l-1}/l, \quad k = 0,\dots,i-2,\ l = 1,\dots,i-k-1.
\end{aligned}
$$

Finally, asymptotically, given I_k, J_k is $\mathsf{uniform}[1 \dots I_k - 1]$.

Let D denote the number of descents initial values in some interval $[n\tau, n(\tau+\Delta)]$. D is asymptotically Gaussian, with

$$
\begin{aligned}
\mathbb{E}(D) &\sim mM,\\
\mathbb{V}(D) &\sim n[M(1-M) + 2(B - M^2]),
\end{aligned}
$$

and

$$
\begin{aligned}
M &= \int_{\tau}^{\tau+\Delta} u du,\\
B &= \int_{\tau}^{\tau+\Delta} du_1 \int_{\tau}^{u_1} u_2 du_2.
\end{aligned}
$$

12 Conclusion

In this paper, we have made a combinatorial and probabilistic study of initial and end heights of first, second, ... descents in samples of geometrically distributed random variables and in permutations. Several other (similar) models can be analyzed with our tools, let us mention: the weak model and/or ascents with the probabilistic approach, k-descents, size d or more descents, (see (2)), last descents, (see (7)), k-ascents with a combinatorial approach (see (2)), etc.

We leave these topics to future work (or to the interested reader), in order to keep the length of the paper within reasonable limits.

Appendix

A Some combinatorial identities

We will use intensively Heine's formula:

$$\sum_{m\geq 0}\frac{(a)_m(b)_m t^m}{(q)_m(c)_m}=\frac{(b)_\infty(at)_\infty}{(c)_\infty(t)_\infty}\sum_{m\geq 0}\frac{(c/b)_m(t)_m b^m}{(q)_m(at)_m}.$$

Several q-combinatorial relations are deduced from Heine.

$$pz\sum_{m\geq 1}\frac{q^{2m}}{(pzq)_m}=\frac{1}{z}\left[\frac{1}{(pqz)_\infty}-1\right]-q. \text{ This is Thm.1 in (6).} \tag{A.1}$$

$$\sum_{m\geq 0}(pz)_m q^m=\frac{1}{pz}[1-(pz)_\infty]. \text{ This is (5) in (6).} \tag{A.2}$$

$$pz\sum_{m\geq 1}\frac{q^m}{(pz)_{m+1}}=\frac{1}{1-pz}\left[\frac{1}{(pqz)_\infty}-1\right]. \text{ This is implicitly used in (6).} \tag{A.3}$$

Heine with $t=q$, $a=0$, $b=q$, $c=pzq$, and (A.2).

$$\sum_{m\geq 1}\frac{pq^m(1-q^m)}{(p)_{m+1}}=1. \text{ Application of (A.3), (A.1).} \tag{A.4}$$

$$\sum_{m\geq 0}(pz)_m q^{2m}=\frac{1}{(pz)^2}\left[1-(pz)_\infty-\frac{1-pz}{q}+\frac{(pz)_\infty}{q}\right]. \text{ This is (6) in (6).} \tag{A.5}$$

$$\sum_{k\geq 1}\frac{p^k}{(q)_k}=-1+\frac{1}{(p)_\infty}. \text{ Heine with } t=p,\ a=0,\ b=q,\ c=q. \tag{A.6}$$

$$\sum_{k\geq 0}\frac{t^k}{(pq)_k}=\frac{(q)_\infty}{(pq)_\infty(t)_\infty}\sum_{m\geq 1}\frac{(p)_m(t)_m q^m}{(q)_m}. \text{ Heine with } a=0,\ b=q,\ c=pq. \tag{A.7}$$

$$(p)_{j-1}\sum_{i=j}^{\infty}\frac{pq^{i-1}(1-q^{i-1})}{(p)_i}=q^{j-1}. \text{ Application of (A.1), (A.3).} \tag{A.8}$$

$$\begin{aligned}\sum_{0\leq j<J}pq^j(p)_j&=p\sum_{j\geq 0}q^j(p)_j-p\sum_{j\geq J}q^j(p)_j\\&=p\frac{1}{p}\Big[1-(p)_\infty\Big]-pq^J(p)_J\frac{1}{pq^J}\Big[1-(pq^J)_\infty\Big]\\&=1-(p)_\infty-(p)_J\Big[1-(pq^J)_\infty\Big]\\&=1-(p)_J. \text{ By (A.2).}\end{aligned} \tag{A.9}$$

$$\begin{aligned}
\sum_{i_1>i_0}\frac{pq^{i_1-1}}{(p)_{i_1}} &= \frac{1}{(p)_{i_0-1}(1-pq^{i_0-1})}\left[\frac{1}{(pq^{i_0})_\infty}-1\right]\\
&= \frac{1}{(p)_{i_0}}\left[\frac{1}{(pq^{i_0})_\infty}-1\right]\\
&= \left[\frac{1}{(p)_\infty}-\frac{1}{(p)_{i_0}}\right]. \text{ By (A.3).}
\end{aligned} \tag{A.10}$$

B Identification of (9.1) and (3.6)

$$\begin{aligned}
\mathbb{E}(I_1) &= p\sum_{i\ge 0}\frac{q^i}{(p)_{i+1}}(1-q^i)(i+1)\\
&= p\sum_{i\ge 0}\sum_{h=0}^{i}\frac{q^i}{(p)_{i+1}}(1-q^i)\\
&= \frac{p}{q}\sum_{h\ge 0}\sum_{i\ge h}\frac{q^i}{(pq)_i}-\frac{p}{q}\sum_{h\ge 0}\sum_{i\ge h}\frac{q^{2i}}{(pq)_i}\\
&= \frac{p}{q}\sum_{h\ge 0}\frac{q^h}{(pq)_h}\sum_{i\ge 0}\frac{q^i}{(pq^{h+1})_i}-\frac{p}{q}\sum_{h\ge 0}\frac{q^{2h}}{(pq)_h}\sum_{i\ge 0}\frac{q^{2i}}{(pq^{h+1})_i}.
\end{aligned}$$

Now use (A.3):

$$\sum_{i\ge 0}\frac{q^i}{(pq^{h+1})_i}=\frac{1}{pq^h(pq^{h+1})_\infty}+1-\frac{1}{pq^h}.$$

We need also this (use(A.1)):

$$\sum_{i\ge 0}\frac{q^{2i}}{(pq^{h+1})_i}=\frac{1}{pq^{2h}}\frac{1}{(pq^{h+1})_\infty}-\frac{1}{pq^{2h}}-\frac{1}{pq^{h-1}}+1.$$

We now plug these two results in:

$$\begin{aligned}
\mathbb{E}(I_1) &= \frac{p}{q}\sum_{h\ge 0}\frac{q^h}{(pq)_h}\sum_{i\ge 0}\frac{q^i}{(pq^{h+1})_i}-\frac{p}{q}\sum_{h\ge 0}\frac{q^{2h}}{(pq)_h}\sum_{i\ge 0}\frac{q^{2i}}{(pq^{h+1})_i}\\
&= \frac{p}{q}\sum_{h\ge 0}\frac{q^h}{(pq)_h}\left[\frac{1}{pq^h(pq^{h+1})_\infty}+1-\frac{1}{pq^h}\right]-\frac{p}{q}\sum_{h\ge 0}\frac{q^{2h}}{(pq)_h}\left[\frac{1}{pq^{2h}}\frac{1}{(pq^{h+1})_\infty}-\frac{1}{pq^{2h}}-\frac{1}{pq^{h-1}}+1\right]\\
&= \frac{p}{q}\sum_{h\ge 0}\frac{1}{(pq)_h}\left[\frac{1}{p(pq^{h+1})_\infty}+q^h-\frac{1}{p}\right]-\frac{p}{q}\sum_{h\ge 0}\frac{1}{(pq)_h}\left[\frac{1}{p(pq^{h+1})_\infty}-\frac{1}{p}-\frac{q^{h+1}}{p}+q^{2h}\right]\\
&= \frac{p}{q}\sum_{h\ge 0}\frac{1}{(pq)_h}\left[q^h+\frac{q^{h+1}}{p}-q^{2h}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{p}{q}\sum_{h\ge0}\frac{1}{(pq)_h}\Big[q^h-q^{2h}\Big]+\sum_{h\ge0}\frac{q^h}{(pq)_h}\\
&= \frac{p}{q}\sum_{h\ge0}\frac{q^h}{(pq)_h}-\frac{p}{q}\sum_{h\ge0}\frac{q^{2h}}{(pq)_h}+\sum_{h\ge0}\frac{q^h}{(pq)_h}\\
&= \frac{1}{q}\sum_{h\ge0}\frac{q^h}{(pq)_h}-\frac{p}{q}\sum_{h\ge0}\frac{q^{2h}}{(pq)_h}\\
&= \frac{1}{q}\Big[\frac{1}{p(pq)_\infty}+1-\frac{1}{p}\Big]-\frac{p}{q}\Big[\frac{1}{p}\frac{1}{(pq)_\infty}-\frac{1}{p}-\frac{q}{p}+1\Big]\\
&= \frac{1}{q}\Big[\frac{q}{p(pq)_\infty}-\frac{q}{p}+2q\Big]\\
&= \frac{1}{p(pq)_\infty}-\frac{1}{p}+2.
\end{aligned}
$$

Since

$$\frac{1}{p(pq)_\infty}-\frac{1}{p}+2=1-\frac{q}{p}+\frac{q}{p(p)_\infty},$$

we established the formula (3.6)

$$\mathbb{E}(I_1)=1+\frac{q}{p}\Big(-1+\frac{1}{(p)_\infty}\Big).$$

C Computation of $\lim_{t\to1}\left[\frac{1}{(pq)_\infty}\frac{1}{1-t}-\sum_{k\ge0}\frac{t^k}{(pq)_k}\right]$

To compute this limit, we use (A.7):

$$
\begin{aligned}
&\lim_{t\to1}\Big[\frac{1}{(pq)_\infty}\frac{1}{1-t}-\sum_{k\ge0}\frac{t^k}{(pq)_k}\Big]\\
&=\lim_{t\to1}\Big[\frac{1}{(pq)_\infty}\frac{1}{1-t}-\frac{(q)_\infty}{(pq)_\infty(t)_\infty}\sum_{m\ge0}\frac{(p)_m(t)_mq^m}{(q)_m}\Big]\\
&=\lim_{t\to1}\Big[\frac{1}{(pq)_\infty}\frac{1}{1-t}-\frac{(q)_\infty}{(pq)_\infty(t)_\infty}-\frac{(q)_\infty}{(pq)_\infty(t)_\infty}\sum_{m\ge1}\frac{(p)_m(t)_mq^m}{(q)_m}\Big]\\
&=\lim_{t\to1}\Big[\frac{1}{(pq)_\infty}\frac{1}{1-t}-\frac{(q)_\infty}{(pq)_\infty(1-t)(qt)_\infty}\Big]-\frac{(q)_\infty}{(pq)_\infty(q)_\infty}\sum_{m\ge1}\frac{(p)_m(q)_{m-1}q^m}{(q)_m}\\
&=\frac{1}{(pq)_\infty}\lim_{t\to1}\Big[\frac{1}{1-t}-\frac{(q)_\infty}{(1-t)(qt)_\infty}\Big]-\frac{1}{(pq)_\infty}\sum_{m\ge1}\frac{(p)_m(q)_{m-1}q^m}{(q)_m}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{(pq)_\infty} \lim_{t\to 1} \frac{(qt)_\infty - (q)_\infty}{(1-t)(qt)_\infty} - \frac{1}{(pq)_\infty} \sum_{m\ge 1} \frac{(p)_m (q)_{m-1} q^m}{(q)_m} \\
&= \frac{1}{(pq)_\infty} \lim_{t\to 1} \frac{\frac{d}{dt}(qt)_\infty\big|_{t=1}(t-1)}{(1-t)(qt)_\infty} - \frac{1}{(pq)_\infty} \sum_{m\ge 1} \frac{(p)_m (q)_{m-1} q^m}{(q)_m} \\
&= -\frac{1}{(pq)_\infty (q)_\infty} \frac{d}{dt}(qt)_\infty\big|_{t=1} - \frac{1}{(pq)_\infty} \sum_{m\ge 1} \frac{(p)_m (q)_{m-1} q^m}{(q)_m} \\
&= \frac{1}{(pq)_\infty} \sum_{k\ge 1} \frac{q^k}{1-q^k} - \frac{1}{(pq)_\infty} \sum_{m\ge 1} \frac{(p)_m q^m}{1-q^m} \\
&= \frac{1}{(pq)_\infty} \sum_{k\ge 1} \frac{[1-(p)_k] q^k}{1-q^k}.
\end{aligned}
$$

D Computation of S_0, S_1, S_2, S_3, S_4

$$
\begin{aligned}
S_0 &:= \sum_{0\le i<j} (i+1) q^i (1-q^i) \frac{1}{(p)_{i+1}} = \sum_{0\le i<j} \sum_{0\le h\le i} q^i (1-q^i) \frac{1}{(p)_{i+1}} \\
&= \sum_{0\le h\le i<j} \frac{q^i}{(p)_{i+1}} - \sum_{0\le h\le i<j} \frac{q^{2i}}{(p)_{i+1}} \\
&= \sum_{0\le h<j} \sum_{0\le i<j-h} \frac{q^{i+h}}{(p)_{i+h+1}} - \sum_{0\le h<j} \sum_{0\le i<j-h} \frac{q^{2i+2h}}{(p)_{i+h+1}} \\
&= \sum_{0\le h<j} \frac{q^h}{(p)_{h+1}} \sum_{0\le i<j-h} \frac{q^i}{(pq^{h+1})_i} - \sum_{0\le h<j} \frac{q^{2h}}{(p)_{h+1}} \sum_{0\le i<j-h} \frac{q^{2i}}{(pq^{h+1})_i}.
\end{aligned}
$$

We do two auxiliary calculation (with (A.3)):

$$
\begin{aligned}
S_1 &:= \sum_{0\le i<I} \frac{q^i}{(pq^{h+1})_i} \\
&= \sum_{i\ge 0} \frac{q^i}{(pq^{h+1})_i} - \sum_{i\ge I} \frac{q^i}{(pq^{h+1})_i} \\
&= \frac{1}{pq^h (pq^{h+1})_\infty} + 1 - \frac{1}{pq^h} - \frac{q^I}{(pq^{h+1})_I} \sum_{i\ge 0} \frac{q^i}{(pq^{h+I+1})_i} \\
&= \frac{1}{pq^h (pq^{h+1})_\infty} + 1 - \frac{1}{pq^h} - \frac{q^I}{(pq^{h+1})_I} \left[\frac{1}{pq^{h+I}(pq^{h+I+1})_\infty} + 1 - \frac{1}{pq^{h+I}} \right] \\
&= 1 - \frac{1}{pq^h} - \frac{q^I}{(pq^{h+1})_I} + \frac{1}{pq^h (pq^{h+1})_I},
\end{aligned}
$$

and with (A.1):

$$\begin{aligned}
S_2 &:= \sum_{0\le i<I} \frac{q^{2i}}{(pq^{h+1})_i}\\
&= \sum_{i\ge 0} \frac{q^{2i}}{(pq^{h+1})_i} - \sum_{i\ge I} \frac{q^{2i}}{(pq^{h+1})_i}\\
&= \frac{1}{pq^{2h}(pq^{h+1})_\infty} - \frac{1}{pq^{2h}} - \frac{1}{pq^{h-1}} + 1 - \frac{q^{2I}}{(pq^{h+1})_I} \sum_{i\ge 0} \frac{q^{2i}}{(pq^{h+I+1})_i}\\
&= \frac{1}{pq^{2h}(pq^{h+1})_\infty} - \frac{1}{pq^{2h}} - \frac{1}{pq^{h-1}} + 1 - \frac{q^{2I}}{(pq^{h+1})_I} \left[\frac{1}{pq^{2h+2I}(pq^{h+I+1})_\infty} - \frac{1}{pq^{2h+2I}} - \frac{1}{pq^{h+I-1}} + 1\right]\\
&= -\frac{1}{pq^{2h}} - \frac{1}{pq^{h-1}} + 1 + \frac{1}{pq^{2h}(pq^{h+1})_I} + \frac{q^I}{pq^{h-1}(pq^{h+1})_I} - \frac{q^{2I}}{(pq^{h+1})_I}.
\end{aligned}$$

Therefore

$$\begin{aligned}
S_0 &= \sum_{0\le h<j} \frac{q^h}{(p)_{h+1}} \sum_{0\le i<j-h} \frac{q^i}{(pq^{h+1})_i} - \sum_{0\le h<j} \frac{q^{2h}}{(p)_{h+1}} \sum_{0\le i<j-h} \frac{q^{2i}}{(pq^{h+1})_i}\\
&= \sum_{0\le h<j} \frac{q^h}{(p)_{h+1}} \left[1 - \frac{1}{pq^h} - \frac{q^{j-h}}{(pq^{h+1})_{j-h}} + \frac{1}{pq^h(pq^{h+1})_{j-h}}\right]\\
&- \sum_{0\le h<j} \frac{q^{2h}}{(p)_{h+1}} \left[-\frac{1}{pq^{2h}} - \frac{1}{pq^{h-1}} + 1 + \frac{1}{pq^{2h}(pq^{h+1})_{j-h}} + \frac{q^{j-h}}{pq^{h-1}(pq^{h+1})_{j-h}} - \frac{q^{2(j-h)}}{(pq^{h+1})_{j-h}}\right]\\
&= \sum_{0\le h<j} \left[\frac{q^h}{(p)_{h+1}} - \frac{1}{p(p)_{h+1}} - \frac{q^j}{(p)_{j+1}} + \frac{1}{p(p)_{j+1}}\right]\\
&- \sum_{0\le h<j} \left[-\frac{1}{p(p)_{h+1}} - \frac{q^{h+1}}{p(p)_{h+1}} + \frac{q^{2h}}{(p)_{h+1}} + \frac{1}{p(p)_{j+1}} + \frac{q^{j+1}}{p(p)_{j+1}} - \frac{q^{2j}}{(p)_{j+1}}\right]\\
&= \sum_{0\le h<j} \left[\frac{q^h}{p(p)_{h+1}} - \frac{q^j}{p(p)_{j+1}} - \frac{q^{2h}}{(p)_{h+1}} + \frac{q^{2j}}{(p)_{j+1}}\right].
\end{aligned}$$

Recalling the expressions for S_1 and S_2 that we just derived, we get

$$\begin{aligned}
S_0 &= \sum_{0\le h<j} \left[\frac{q^h}{pq(pq)_h} - \frac{q^{2h}}{q(pq)_h}\right] - \frac{jq^j}{p(p)_{j+1}} + \frac{jq^{2j}}{(p)_{j+1}}\\
&= \frac{1}{pq}\left[1 - \frac{1}{p} - \frac{q^j}{(pq)_j} + \frac{1}{p(pq)_j}\right] - \frac{1}{q}\left[-\frac{2q}{p} + \frac{1}{p(pq)_j} + \frac{q^{j+1}}{p(pq)_j} - \frac{q^{2j}}{(p)_j}\right] - \frac{jq^j}{p(p)_{j+1}} + \frac{jq^{2j}}{(p)_{j+1}}
\end{aligned}$$

$$= \frac{1}{pq} - \frac{1}{p^2q} - \frac{q^j}{pq(pq)_j} + \frac{1}{p^2q(pq)_j} + \frac{2}{p} - \frac{1}{pq(pq)_j} - \frac{q^j}{p(pq)_j} + \frac{q^{2j}}{q(p)_j} - \frac{jq^j}{p(p)_{j+1}} + \frac{jq^{2j}}{(p)_{j+1}}$$
$$= \frac{2}{p} - \frac{1}{p^2} + \frac{q}{p^2(p)_{j+1}} - \frac{(1+q)q^j}{pq(p)_{j+1}} + \frac{q^{2j}}{q(p)_j} - \frac{jq^j}{p(p)_{j+1}} + \frac{jq^{2j}}{(p)_{j+1}}.$$

Next, we want to compute

$$S_3 := \sum_{j\geq 0} \frac{p^3q^{3j}}{q(p)_{j+1}}.$$

We will use Heine with $t = pq^h$, $a = 0$, $c = q$, $b = q$:

$$\begin{aligned}
\sum_{i\geq 0} \frac{q^{3i}}{(pq^{h+1})_i} &= \frac{(q)_2}{(pq^{h+1})_\infty} \sum_{m\geq 0} \frac{(pq^h)_m (q^3)_m}{(q)_m} q^m \\
&= \frac{(q)_2}{(pq^{h+1})_\infty} \frac{(pq^h)_\infty (q^4)_\infty}{(q)_\infty} \sum_{m\geq 0} \frac{(q)_m}{(q)_m (q^4)_m} (pq^h)^m \\
&= \frac{1-pq^h}{1-q^3} \sum_{m\geq 0} \frac{1}{(q^4)_m} (pq^h)^m \\
&= \frac{(1-pq^h)(q)_3}{(pq^h)^3} \sum_{m\geq 3} \frac{1}{(q)_m} (pq^h)^m \\
&= \frac{(1-pq^h)(q)_3}{(pq^h)^3} \left[\frac{1}{(pq^h)_\infty} - 1 - q^h - \frac{(pq^h)^2}{(q)_2} \right] \\
&= \frac{(q)_3}{(pq^h)^3} \frac{1}{(pq^{h+1})_\infty} - \frac{(1-pq^h)(q)_3}{p^3q^{3h}} - \frac{(1-pq^h)(q)_3}{p^3q^{2h}} - \frac{(1-pq^h)(1-q^3)}{pq^h}.
\end{aligned}$$

From this we derive (set $h = -1$)

$$\begin{aligned}
S_3 := \sum_{j\geq 0} \frac{p^3q^{3j}}{q(p)_{j+1}} &= \frac{p^3}{q} \left[\frac{(q)_3}{(p/q)^3} \frac{1}{(p)_\infty} - \frac{(1-p/q)(q)_3}{p^3q^{-3}} - \frac{(1-p/q)(q)_3}{p^3q^{-2}} - \frac{(1-p/q)(1-q^3)}{p/q} \right] \\
&= \frac{p^3}{q} \left[\frac{(q)_3 q^3}{p^3} \frac{1}{(p)_\infty} - \frac{q^2(q-p)(q)_3}{p^3} - \frac{q(q-p)(q)_3}{p^3} - \frac{(q-p)(1-q^3)}{p} \right] \\
&= \frac{(q)_3 q^2}{(p)_\infty} - q(q-p)(q)_3 - (q-p)(q)_3 - \frac{p^2(q-p)(1-q^3)}{q}.
\end{aligned}$$

Now we compute (we use S_2)

$$S_4 := \sum_{j\geq 0} \frac{jq^{2j}}{(p)_{j+1}} = \sum_{h\geq 0} \sum_{j\geq h} \frac{q^{2j}}{(p)_{j+1}}$$

$$\begin{aligned}
&= \sum_{h\ge 0} \frac{q^{2h}}{(p)_{h+1}} \sum_{j\ge 0} \frac{q^{2j}}{(pq^{h+1})_j} \\
&= \sum_{h\ge 0} \frac{q^{2h}}{(p)_{h+1}} \left[-\frac{1}{pq^{2h}} - \frac{1}{pq^{h-1}} + 1 + \frac{1}{pq^{2h}(pq^{h+1})_\infty} \right] \\
&= -\frac{1}{p} \sum_{h\ge 0} \frac{1}{(p)_{h+1}} \left[1 - \frac{1}{(pq^{h+1})_\infty} \right] - \frac{q}{p} \sum_{h\ge 0} \frac{q^h}{(p)_{h+1}} + \sum_{h\ge 0} \frac{q^{2h}}{(p)_{h+1}} \\
&= \frac{1}{pq} \sum_{h\ge 0} \left[\frac{1}{(pq)_\infty} - \frac{1}{(pq)_h} \right] - \frac{1}{p} \left[\frac{1}{p(pq)_\infty} + 1 - \frac{1}{p} \right] + \frac{1}{q} \left[\frac{1}{p(pq)_\infty} - \frac{1}{p} - \frac{q}{p} + 1 \right] \\
&= \frac{1}{pq} \frac{1}{(pq)_\infty} \sum_{k\ge 1} \frac{[1-(p)_k]q^k}{1-q^k} - \frac{1}{p} \left[\frac{1}{p(pq)_\infty} + 1 - \frac{1}{p} \right] + \frac{1}{q} \left[\frac{1}{p(pq)_\infty} - \frac{1}{p} - \frac{q}{p} + 1 \right],
\end{aligned}$$

Acknowledgements

The insight of two anonymous referees is gratefully acknowledged.

References

[1] G. Andrews. *The Theory of Partitions*, volume 2 of *Encyclopedia of Mathematics and its Applications*. Addison–Wesley, 1976.

[2] C. Brennan and A. Knopfmacher. The distribution of ascents of size d or more in samples of geometric random variables. *Discrete Mathematics and Theoretical Computer Science*, AD:341–350, 2005. 2005 International Conference on Analysis of Algorithms.

[3] K. L. Chung. *Markov Chains with Stationary Transition Probabilities*. Springer-Verlag, 1967.

[4] P. Flajolet and G.N. Martin. Probabilistic counting algorithms for data base applications. *Journal of Computer and System Sciences*, 31:182–209, 1985.

[5] P. Kirschenhofer and H. Prodinger. The path length of random skip lists. *Acta Informatica*, 31:775–792, 1994.

[6] A. Knopfmacher and H. Prodinger. The first descent in samples of geometric random variables and permutations. *Discrete Mathematics and Theoretical Computer Science*, 8:214–234, 2006.

[7] A. Knopfmacher and H. Prodinger. The last descent in samples of geometric random variables and permutations. *Ars Combinatoria*, 83:365–379, 2007.

[8] W. Pugh. Skip lists: A probabilistic alternative to balanced trees. In *Algorithms and Data Structures*, volume 382 of *Lecture Notes in Computer Science*, pages 437–449, 1989.