Abstract

In this talk, we show an interesting fact that a quarter of paths of random walks of any length \( n \geq 2 \) have two maximums. Moreover, it holds that the asymptotic distribution of number of maximums of paths obeys the geometric distribution of parameter \( 1/2 \). Furthermore, we provide the generating function for counting number of paths jointly with number \( k \) of maximums and number \( l \) of minimums. This result comes from the careful consideration of paths with restricted width by using the combinatorial symbolic method. We can extend to other topics on number of zero returns, zero touches, etc. for paths of random walks. As by-product, we obtain many combinatorial identities.

1 Distribution of number of maximums in random walks

Dyck paths are characterized as sequences of numbers \( x = (x_0, x_1, \ldots, x_{2n}) \) satisfying the conditions:

\[
x_0 = x_{2n} = 0, \quad x_j \geq 0, \quad |x_{j+1} - x_j| = 1 \quad \text{for} \ 0 \leq j \leq 2n - 1.
\]

It is common that Dyck paths of length \( 2n \) correspond one-to-one to binary trees having \( n \) internal nodes, and can be enumerated by Catalan numbers.

Let \( \mathcal{D} \) be the class of Dyck paths, the size of which is defined by length. Then, we can symbolically express the class as follows:

\[
\mathcal{D} = \phi + (\mathcal{D} \searrow) \times \mathcal{D},
\]

or,

\[
\mathcal{D} = \phi + (\mathcal{D} \searrow) + (\mathcal{D} \searrow)^2 + \ldots,
\]

where \( \nearrow \) denotes an ascent step, and \( \searrow \) denotes a descent step.
From the symbolic expression, we deduce the generation function for Dyck paths:
\[ D(z) = 1 + z^2 D(z)^2 = \frac{1}{1 - z^2 D(z)}. \] (1.4)

Then, we have the explicit expression of generator \( D(z) \) as
\[ D(z) = \frac{1 - \sqrt{1 - 4z^2}}{2z^2}, \] (1.5)
and its coefficients are Catalan numbers:
\[ c_n = [z^{2n}] D(z) = \frac{1}{2n + 1} \binom{2n + 1}{n}, \] (1.6)
where the coefficient extractor \([z^n]f(z)\) gives the coefficient of \(z^n\) in \(f(z)\).

Now, we are ready to express the generating function to enumerate the number of maximums in a path of random walks by following the symbolic guide of [1]. The path is decomposed by three phases, that is, the initial part that attains the first maximum, the repeating part of maximums, and the final part that is under the maximum. The initial part that attains the first maximum is depicted in Figure 1, and the repeating part of maximums in Figure 2. And, the final part that is under the maximum has the same generating function as the initial part.

Thus, we obtain BGF \( M(z, u) \) marking the maximums with \( u \) to enumerate the number of maximums in random walks as follows:
\[ M(z, u) = \frac{1}{1 - zD(z)} \cdot \frac{u}{1 - u z^2 D(z)} \cdot \frac{1}{1 - zD(z)}. \] (1.7)

Here, we note that the number of maximum is the number of related bridges plus one.
By applying the equation (1.5), we have
\[ M(z, u) = \frac{2uz^2}{(1-2z)((1-u)(1-\sqrt{1-4z^2})+2uz^2)}. \quad (1.8) \]

Here, it can be easily checked that
\[ M(z, 1) = \frac{1}{1-2z}. \quad (1.9) \]

This corresponds to the fact that there are \(2^n\) paths of length \(n\) in total.

Now, we rewrite the \(M(z, u)\) such as
\[ M(z, u) = \frac{1}{(1-zD(z))^2} \cdot \frac{u}{1-uz^2 D(z)} = R(z) \cdot u D(z, u), \quad (1.10) \]
where
\[ R(z) = \frac{1}{(1-zD(z))^2}; \quad (1.11) \]
and
\[ D(z, u) = \frac{1}{1-uz^2 D(z)}. \quad (1.12) \]

By the BGF \(M(z, u)\), we can compute the value of coefficients of term \(z^n u^k\).

Figure 3 is the table of \(m(n, k) = [z^n u^k]M(z, u)\).

Let us denote the coefficient polynomial of \(z^n\) of \(M(z, u)\) by
\[ \mu_n(u) = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor + 1} m(n, k) u^k, \quad (1.13) \]
and show the initial part of expansion of \(M(z, u)\) as follows:
\[
M(z, u) = \sum_{n=0}^{\infty} \mu_n(u) z^n \\
= u + 2uz + (3u + u^2)z^2 + (6u + 2u^2)z^3 + (11u + 4u^2 + u^3)z^4 \\
+ (22u + 8u^2 + 2u^3)z^5 + (42u + 16u^2 + 5u^3 + u^4)z^6 \\
+ (84u + 32u^2 + 10u^3 + 2u^4)z^7 + (163u + 64u^2 + 22u^3 + 6u^4 + u^5)z^8 + \cdots
\]

Here, we notice the interesting relations:
\[ \mu_{2i+1}(u) = 2\mu_{2i}(u) \quad \text{or} \quad m(2l + 1, k) = 2m(2l, k), \quad (1.14) \]
\[ [u^2] \mu_n(u) = m(n, 2) = 2^{n-2}. \quad (1.15) \]

That is, among paths of random walks, a quarter have two maximums for any length \(n\).

Moreover, we can deduce the important recursion of coefficients, that is,
\[ [u^l] \mu_{2i}(u) = [u^{l-1}] \mu_{2i}(u) - [u^{l-2}] \mu_{2i-2}(u) \quad (1.16) \]
\[
m(n, k) = \left[ z^n u^k \right] M(z, u)
\]

<table>
<thead>
<tr>
<th>n ( \backslash ) k</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>22</td>
<td>8</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>42</td>
<td>16</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>84</td>
<td>32</td>
<td>10</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>163</td>
<td>64</td>
<td>22</td>
<td>6</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>326</td>
<td>128</td>
<td>44</td>
<td>12</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>638</td>
<td>256</td>
<td>93</td>
<td>29</td>
<td>7</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1276</td>
<td>512</td>
<td>186</td>
<td>58</td>
<td>14</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>2510</td>
<td>1024</td>
<td>386</td>
<td>130</td>
<td>37</td>
<td>8</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>5020</td>
<td>2048</td>
<td>772</td>
<td>260</td>
<td>74</td>
<td>16</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>9908</td>
<td>4096</td>
<td>1586</td>
<td>562</td>
<td>176</td>
<td>46</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>19816</td>
<td>8192</td>
<td>3172</td>
<td>1124</td>
<td>352</td>
<td>92</td>
<td>18</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>39203</td>
<td>16384</td>
<td>6476</td>
<td>2380</td>
<td>794</td>
<td>232</td>
<td>56</td>
<td>10</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>78406</td>
<td>32768</td>
<td>12952</td>
<td>4760</td>
<td>1588</td>
<td>464</td>
<td>112</td>
<td>20</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>155382</td>
<td>65536</td>
<td>26333</td>
<td>9949</td>
<td>3473</td>
<td>1093</td>
<td>299</td>
<td>67</td>
<td>11</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 3: Distribution of number of random walks of length \( n \) with \( k \) maximums.

By this relation, as depicted in Figure 4, we can calculate the whole columns from the first and second columns.

Now, we analyze \( R(z) \) of (1.11) that is the generating function of number of random walks with single maximum.

\[
R(z) = \frac{1}{(1 - zD(z))^2} = \sum_{n=0}^{\infty} r(n) z^n
\]

The coefficients \( r(n) = [z^n] R(z) = [z^n] M(z, u) = [u^1] \mu_n(u) = m(n, 1) \) are shown as the column \( k = 1 \) in the Figure 3, and expressed as follows:

\[
r(n) = [z^n] R(z) = [z^n] \frac{1}{(1 - zD(z))^2}
\]

\[
r(n) = 2^n \left\{ 1 - \sum_{i=0}^{\left\lfloor \frac{n-2}{2} \right\rfloor} \frac{1}{2i+2} \frac{1}{2i+1} \left( \begin{array}{c} 2i+1 \\ i \end{array} \right) \right\}.
\]
For simplicity, by restricting to $n = 2l$, we have

$$r(2l) = 2^{2l} \left( 1 - \sum_{i=0}^{l-1} \frac{1}{2^{2i+2}} \cdot \frac{1}{2l + 1} \binom{2l + 1}{i} \right) = 2^{2l-1} \cdot \frac{1}{2} \binom{2l}{l}. \quad (1.19)$$

Thought the analysis of OGF $D(z, u)$ in (1.12), we can calculate $m(2l, k)$ by obtaining first the coefficient polynomial $\lambda_k(z)$ of $u^k$ in BGF $M(z, u)$, and next by evaluating the coefficient of $z^{2l}$ in $\lambda_k(z)$. And we conclude the following formula:

$$[z^{2l} u^k] M(z, u)$$

$$= \begin{cases} 
  r(2l) = 2^{2l-1} \cdot \frac{1}{2} \binom{2l}{l}, & \text{if } k = 1, \\
  \sum_{j=0}^{l-k+1} \frac{k - 1}{2l - k + 1 - 2j} \binom{2l - k + 1 - 2j}{l - j} \cdot \left( 2^{2j-1} + \frac{1}{2} \binom{2j}{j} \right), & \text{if } 2 \leq k \leq l + 1. 
\end{cases} \quad (1.20)$$

When $k = 2$, we can show from the second line of the above equality (1.20),

$$\sum_{j=0}^{l-1} \frac{1}{2l - 2j - 1} \binom{2l - 2j - 1}{l - j} \cdot \left( 2^{2j-1} + \frac{1}{2} \binom{2j}{j} \right) = 2^{2l-2}. \quad (1.21)$$
This proves the equality (1.15).

Furthermore, we can calculate \( m(2l, k) \) by obtaining first the coefficient polynomial \( \mu_{2l}(u) \) of \( z^{2l} \) in BGF \( M(z, u) \), and next by evaluating the coefficient of \( u^k \) in \( \mu_{2l}(u) \). And we conclude the following formula:

\[
[u^k][z^{2l}]M(z, u) = \begin{cases} 
1 & \text{if } k = 1, \\
\sum_{j=0}^{l-k+1} \frac{(2j+1)(2j+k-1)}{2l-k+1} \binom{2l-k+1}{l-j-k+1} & \text{if } 2 \leq k \leq l+1.
\end{cases}
\] (1.22)

When \( k = 2 \), we can also show from the second line of the above equality (1.22),

\[
\sum_{j=0}^{l-1} \frac{(2j+1)^2}{2l-1} \binom{2l-1}{l+j} = 2^{2l-2}
\] (1.23)

By the way, we have another method to evaluate the value of \( m(n, k) \). Let us explain the method in what follows. By the equations (1.19) and (1.15), we have all information on the first and second columns of Figure 4, and by (1.14), we can construct any size of table as Figure 3. Thus,

\[
m(2l, 1) = \frac{1}{2} \cdot 2^{2l} + \frac{1}{2} \binom{2l}{l},
\]

\[
m(2l, 2) = \frac{1}{4} \cdot 2^{2l}.
\]

The recursion (1.16) can be rewrite as

\[
m(2l, k) = m(2l, k - 1) - m(2(l - 1), k - 2).
\] (1.24)

Then, we deduce the sequence of formulas as follows:

\[
m(2l, 3) = \frac{1}{8} \cdot 2^{2l} - \frac{1}{2} \binom{2l - 1}{l - 1},
\]

\[
m(2l, 4) = \frac{1}{16} \cdot 2^{2l} - \frac{1}{2} \binom{2l - 1}{l - 1},
\]

\[
m(2l, 5) = \frac{1}{32} \cdot 2^{2l} - \frac{1}{2} \binom{2l - 1}{l - 1} + \frac{1}{2} \binom{2(l - 2)}{l - 2},
\]

\[\vdots\]

In general, we obtain for \( k \geq 2 \) the explicit form,

\[
m(2l, k) = \frac{1}{2^k} \cdot 2^{2l} - \sum_{i=0}^{\lfloor k/2 \rfloor} (-1)^i \cdot \frac{1}{2} \binom{2(l - i - 1)}{l - i - 1} \binom{k - 3 - i}{i}.
\] (1.25)
**Theorem 1.1** The probability \( p_n(k) \) of the event that symmetric random walks of length \( n \) starting 0 has \( k \) maximums asymptotically obeys the geometric distribution with parameter \( 1/2 \):

\[
\lim_{n \to \infty} p_n(k) = \lim_{n \to \infty} \frac{m(n, k)}{2^n} = \frac{1}{2^k}.
\]

**(1.26)**

*Proof:* For even \( n \), we have already show the formula (1.25). It is easy to check that the second term in (1.25) is order \( o(2^n) \), and the theorem is established. For odd \( n \), we can derive the same result by using the relation (1.14). \( \square \)

## 2 Dyck path with height constraint

To study the joint distribution of maximums and minimums among paths of random walks, we need the knowledge of Dyck paths with height constraint (see Figure 5).

![Figure 5: Dyck path with height constraint](image)

The generator \( D(z, s) \) of Dyck paths with height at most \( s \geq 0 \), that are also called as Dyck paths with height constraint \( s \), is recursively defined by the rule:

\[
D(z, s) = \frac{1}{1 - z^2 D(z, s - 1)}, \quad D(z, 0) = 1.
\]

**(2.1)**

Solving this difference equation, we have the explicit formula of \( D(z, s) \) for \( s \geq 0 \) as follows:

\[
D(z, s) = \frac{f(z, s)}{f(z, s + 1)},
\]

**(2.2)**

where

\[
f(z, s) = \sum_{i=0}^{\lfloor \frac{s}{2} \rfloor} (-1)^i \binom{s - i}{i} z^i = z^s U_s \left( \frac{1}{2z} \right).
\]

**(2.3)**

Here, \( U_s(x) \) is the second kind of Chebyshev polynomial that is defined as

\[
U_s(\cos \theta) = \frac{\sin (s + 1) \theta}{\sin \theta}.
\]

## 3 Boundary restricted Pascal triangle

Here, we consider the Pascal triangle but constrained in interval \([-t, s], (s, t \geq 0)\) (see Figure 6).
To give the OGF counting the number of paths starting the origin to the position $-t \leq i \leq s$ in the level $n$, we first consider the case when $i = s$. The situation is depicted in Figure 7. Thus, the OGF for number of random walks between $-t$ and $s$ ending at the upper boundary $s$ is obtained by using generators of Dyck paths with restricted hight as follows:

$$A(z, s, t) = (D(z, t)z)(D(z, t + 1)z)(D(z, t + 2)z)\cdots(D(z, t + s))$$

$$= z^s \frac{f(z, t)}{f(z, t + s + 1)}, \quad (3.1)$$

Finally, considering the way of constructing the triangle, we can express the OGF $A(z, s, t, i)$ counting the number of paths starting 0 to the position $i$ in the level $n$ as follows: for $0 \leq i \leq s$,

$$A(z, s, t, i) = \frac{f(z, s - i)}{z^{s-i}} A(z, s, t)$$

$$= z^i \frac{f(z, s - i)f(z, t)}{f(z, s + t + 1)}, \quad (3.2)$$

and for $-t \leq i \leq 0$,

$$A(z, s, t, i) = A(z, t, s, -i). \quad (3.3)$$
For example, for the case when $s = 4, t = 2$, we have

\[
A(z; 4; 2; 4) = z^4 + 5z^6 + 20z^8 + 74z^{10} + 264z^{12} + \ldots
\]

\[
A(z; 4; 2; 3) = z^3 + 5z^5 + 20z^7 + 74z^9 + 264z^{11} + \ldots
\]

\[
A(z; 4; 2; 2) = z^2 + 4z^4 + 15z^6 + 54z^8 + 190z^{10} + \ldots
\]

\[
A(z; 4; 2; 1) = z + 3z^3 + 10z^5 + 34z^7 + 116z^9 + 396z^{11} + \ldots
\]

\[
A(z; 4; 2; 0) = 1 + 2z^2 + 6z^4 + 19z^6 + 62z^8 + 206z^{10} + \ldots
\]

\[
A(z; 4; 2; -1) = z + 3z^3 + 9z^5 + 28z^7 + 90z^9 + 296z^{11} + \ldots
\]

\[
A(z; 4; 2; -2) = z^2 + 3z^4 + 9z^6 + 28z^8 + 90z^{10} + 296z^{12} + \ldots
\]

4 Joint distribution of number of maximums and minimums in random walks

Using the results of previous sections, we can express the generating function $G(z; u; v)$, where $z$ conveys the length information of paths, $u$ is used for the number of maximums, and $v$ is used for the number of minimums. Due to page restriction, we only show the table for the case $n = 30$ (see Figure 8). It is interesting to notice that $g(n, k, l) = [u^k v^l] [z^n] G(z; u; v)$ is not necessarily monotonically decreasing for fixed $l$ and $n$ when $k$ increases.

5 Combinatorial identities related to random walks

Here, we give a list of the interesting combinatorial identities that we have encountered during the studies of random walks. The first five equations express different decompositions of Catalan Numbers (1.6): $c_m = \frac{1}{m+1} \binom{2m+1}{m} = \frac{1}{m+1} \binom{2m}{m}$.

\[
\sum_{k=0}^{m} \binom{k}{2m-k} \binom{2m-2k}{m} = c_m. \quad (5.1)
\]
\[
\frac{1}{m+1} \sum_{k=0}^{m} 2^k \binom{2m-k}{m} = c_m. 
\] (5.2)

\[
\frac{m}{m+1} \sum_{k=0}^{m} \frac{1}{2m-k} \binom{2m-k}{m} = c_m. 
\] (5.3)

\[
\frac{m+1}{m} \sum_{k=0}^{m} \frac{1}{2m-k} \binom{2m-k}{m+1} = c_m. 
\] (5.4)

\[
\frac{m+1}{2m+1} \sum_{k=0}^{m} \frac{1}{2m+1-k} \binom{2m+1-k}{m+1} = c_m. 
\] (5.5)

The next three equations are different weighted sums of Catalan numbers resulting the value 1.

\[
\sum_{i=0}^{\infty} \frac{1}{2^{i+1}} \cdot c_i = 1. 
\] (5.6)

\[
\sum_{i=0}^{\infty} \frac{2i+1}{i+2} \cdot \frac{1}{2^{i+1}} \cdot c_i = 1. 
\] (5.7)

\[
\frac{1}{\log 2} \sum_{i=1}^{\infty} \frac{i+1}{i} \cdot \frac{1}{2^{i+1}} \cdot c_i = 1. 
\] (5.8)

The following equation results in the power of 2 as the equations (1.21) and (1.23).

\[
\sum_{k=0}^{m} 2^k \binom{2m-k}{m} = 2^{2m}. 
\] (5.9)

**Acknowledgment**

The first author would like to thank National Insititute of Information and Communications Technology (NICT) of Japan for the financial support to participate in ITA2015.

**References**