STATISTICS OF DOMINO TILINGS ON A RECTANGULAR BOARD

THOTSAPORN "AEK" THANATIPANONDA

ABSTRACT. It is well-known that the Fibonacci sequence, F_n, is the number of ways to cover a 2-by-$(n-1)$ board using only horizontal (H) or vertical (V) 2-by-1 dominoes. The number of ways to tile a rectangular m-by-n board by these dominoes were given in 1961 by Kasteleyn [1] through the evaluation of a determinant. In this paper, we apply an automated method for the mixed moments $E[V^a H^b]$ for fixed non-negative integers a, b on a general m-by-n board. These moments will give information about the distribution of "V-H statistics". This is an implementation of the work of Zeilberger, [3], [4], [2].

1. STATISTICS ON AN 2-BY-n BOARD

In order to demonstrate the method, we first fix the board to be of size 2-by-n. Let V_n be a random variable of the number of vertical dominoes on the tiling of a board of this size.

Let b be a tiling of the 2-by-n board, B_n. We consider the power sum, $S[V^r] := \sum_{b \in B_n} V(b)^r$. The number of possible tilings on the 2-by-n board is $S[V^0] = f_n$, where f_n is the $(n+1)$-th Fibonacci number. The straight moment is $E[V^r] = S[V^r]/S[V^0] = S[V^r]/f_n$.

Ultimately, we are interested in the moment about the mean, $E[(V - \mu)^r]$, and the scaled-moments, $\dfrac{E[(V - \mu)^r]}{E[(V - \mu)^2]^{r/2}}$, which can be used to show the normality distribution of V. We will do all this by gathering data, making conjectures, and then proving them using induction.

Maple package:

All of these methods are efficiently implemented by the program, `Domino.txt`, which is available as a free download at

`http://www.thotsaporn.com/Domino.txt` .

1.1. **Fast Calculations and Conjectures.** The generating function of a random variable V is defined by

$$F_n(v) := \sum_{b \in B_n} v^{V(b)}.$$

For example, on the board of size 2-by-3, $\quad F_3(v) = v^3 + 2v$.

It is a basic fact from probability that $S[V^0] = F_n(1)$. $S[V^r]$ is obtained by applying the operator $(v \dfrac{d}{dv})^r$ to $F_n(v)$, and then substituting $v = 1$.

Next, we define the grand generating function $H(v, t)$ by

$$H(v, t) = \sum_{n=0}^{\infty} F_n(v)t^n.$$

In this problem,

$$H(v, t) = \frac{1}{1 - vt - t^2}, \tag{1.1}$$

which can be derived from a simple recurrence

$$H(v, t) = 1 + vtH(v, t) + t^2 H(v, t).$$

Now, (1.1) allows us to calculate values of $S[V^r]$ (and hence $E[V^r]$) very fast. Recall that

$$\sum_{n=0}^{\infty} S[V^0]t^n = H(1, t) = \frac{1}{1 - t - t^2}.$$

In general,

$$\sum_{n=0}^{\infty} S[V^r]t^n = (v\frac{d}{dv})^r H(v, t)|_{v=1}.$$

Applying the quotient rule from single-variable calculus, we find that

$$\sum_{n=0}^{\infty} S[V^r]t^n = \left[(v\frac{d}{dv})^r \frac{1}{1 - vt - t^2} \right]_{v=1} = \frac{P_r(t)}{(1 - t - t^2)^{r+1}},$$

where $P_r(t)$ is a polynomial in t of degree at most $2r$.

In fact, $S[V^r]$ satisfies the recurrence of the form $(N^2 - N - 1)^{r+1}$ and can be written as

$$S[V^r] = A(n)f_n + B(n)f_{n-1},$$

where $A(n)$ and $B(n)$ are polynomials in n of degree at most r.

This enables us to use a computer program to conjecture the formulas by trying to fit the polynomial to the data. Define the golden ratio, $\phi := \frac{1 + \sqrt{5}}{2}$. Then we have the following formulas:

$$f_n = S[V^0] = \frac{1}{\sqrt{5}} \left(\frac{1 + \sqrt{5}}{2} \right)^{n+1} - \frac{1}{\sqrt{5}} \left(\frac{1 - \sqrt{5}}{2} \right)^{n+1},$$

$$\mu_n := E[V] = \frac{1}{f_n} \left[\frac{n}{5}f_n + \frac{2}{5}(n+1)f_{n-1} \right] \approx \frac{n}{5} + \frac{2}{5\phi}(n+1),$$

$$E[V^2] = \frac{1}{f_n} \left[\frac{n(5n + 12)}{25}f_n + \frac{4(n+1)}{25}f_{n-1} \right] \approx \frac{n(5n + 12)}{25} + \frac{4(n+1)}{25\phi},$$

$$E[V^3] = \frac{1}{f_n} \left[\frac{n(n^2 + 12n + 16)}{25}f_n + \frac{2(n+1)(n^2 + 2n - 4)}{25}f_{n-1} \right]$$

$$\approx \frac{n(n^2 + 12n + 16)}{25} + \frac{2(n+1)(n^2 + 2n - 4)}{25\phi},$$

$$\dots$$

These computations were done entirely by a computer program written in Maple. To get the same result, try `ConjMoV(3,n); `.

From these calculations, we can make a general conjecture of $E[V^r]$ (using a polynomial ansatz i.e. polynomial in n with polynomial coefficients in r) starting from the leading terms (and after simplifying ϕ):

Conjecture 1

$E[V^r]$

$$\approx \frac{n^r}{5^{r/2}}\left[1 + \frac{r(2r + \sqrt{5} - 3)}{\sqrt{5}n} + \frac{r(r-1)(6r^2 - 32r + 6\sqrt{5}r + 37 - 9\sqrt{5})}{15n^2}\right.$$

$$+ \frac{2\sqrt{5}r(r-1)(r-2)(2r^3 + 3\sqrt{5}r^2 - 23r^2 - 16\sqrt{5}r + 77r + 16\sqrt{5} - 73)}{75n^3} +$$

$$\frac{r(r-1)(r-2)(r-3)(6r^4 - 120r^3 + 12\sqrt{5}r^3 - 138\sqrt{5}r^2 + 800r^2 - 2100r + 432\sqrt{5}r - 393\sqrt{5} + 1849)}{225n^4}$$

$$\left.+ \ldots\right],$$

This conjecture was also made by a computer program. For example, try `BigConjMoV(5,n,r);` to get the fifth largest term of the formula.

We use the conjectures of straight moments to calculate the moments about the mean via $E[(V - \mu)^r] = \sum_{i=0}^{r}(-1)^i \binom{r}{i}\mu^i E[V^{r-i}]$.

$$E[(V - \mu)^0] = 1,$$

$$E[(V - \mu)^1] = 0,$$

$$E[(V - \mu)^2] = E[X^2] - E[X]^2 = \frac{4\sqrt{5}n + 4\sqrt{5} - 8}{25},$$

$$E[(V - \mu)^3] = E[X^3] - 3E[X^2]E[X] + 2E[X]^3 = \frac{8\sqrt{5}n}{125} + \frac{8\sqrt{5}}{125} - \frac{48}{125},$$

$$E[(V - \mu)^4] = \frac{48n^2}{125} + \frac{96n}{125} - \frac{272\sqrt{5}n}{625} + \frac{16}{25} - \frac{272\sqrt{5}}{625},$$

$$E[(V - \mu)^5] = \frac{64n^2}{125} + \frac{128n}{125} - \frac{736\sqrt{5}n}{625} + \frac{9856}{3125} - \frac{736\sqrt{5}}{625},$$

$$\ldots$$

For more of these conjectures try, for example, `ConjMoMeanV(7,n);`.

This data leads us to conjecture general formulas of moments about the mean, which we will prove formally in the next section.

Conjecture 2

$$E[(V - \mu)^{2r}] = \left(\frac{2}{5\sqrt{5}}\right)^r \frac{(2r)!n^r}{r!} + \text{ smaller terms,}$$

$$E[(V - \mu)^{2r+1}] = \frac{2}{15}\left(\frac{2}{5\sqrt{5}}\right)^r \frac{(2r+1)!n^r}{(r-1)!} + \text{ smaller terms.}$$

1.2. **The Moments and Asymptotic Distribution.** Conjectures are nice, but we still need to prove them. The proof is elementary in the sense that we set up some recurrences and then apply induction on them. In the end, we use these results to conclude the asymptotic normality distribution of V.

The generating function $F_n(v)$ satisfies the recurrence

$$F_n(v) = vF_{n-1}(v) + F_{n-2}(v) \quad , n \geq 2,$$

where

$$F_0(v) = 1, \quad F_1(v) = v.$$

The centralized probability generating function of $F_n(v)$ is

$$G_n(v) := \sum_i p(i)v^{i-\mu} = \frac{1}{f_n v^\mu} F_n(v).$$

The recurrence of $F_n(v)$ translates to

$$G_n(v) = v\frac{f_{n-1}G_{n-1}(v)}{f_n v^{\mu_n - \mu_{n-1}}} + \frac{f_{n-2}G_{n-2}(v)}{f_n v^{\mu_n - \mu_{n-2}}} \quad , n \geq 2, \tag{1.2}$$

where

$$G_0(v) = 1, \quad G_1(v) = 1.$$

We will use this recurrence to set up some relations for the proof. But before diving into the proof, we need to discuss some necessary probability background.

Probability Background Interim

Definition 1.1 (Exponential moment generating function).

$$\phi(t) := E[e^{tX}] = \begin{cases} \sum_x e^{tx} p(x) & \text{if } x \text{ is discrete} \\ \int_{-\infty}^{\infty} e^{tx} f(x)dx & \text{if } x \text{ is continuous.} \end{cases}$$

For the discrete case:

$$m_n := E[X^n] = \sum_x x^n p(x) = \phi^n(0).$$

We then have that

$$\phi(t) = E[e^{tX}] = \sum_x \sum_n \frac{t^n x^n}{n!} p(x) = \sum_n \frac{t^n}{n!} \sum_x x^n p(x) = \sum_n \frac{t^n}{n!} m_n.$$

For the standard normal distribution, $\phi(t) = e^{\frac{t^2}{2}} = \sum_r \frac{t^{2r}}{r!2^r}$ implies that $m_{2r} = \frac{(2r)!}{r!2^r}$ and $m_{2r+1} = 0$.

Now we can return to the main theorem.

Theorem 1.2. *Let $E_r(n) = \dfrac{E[(V-\mu)^r]}{r!}$. Then*

$$E_{2r}(n) = \left(\frac{2}{5\sqrt{5}}\right)^r \frac{n^r}{r!} + \text{ smaller terms,}$$

$$E_{2r+1}(n) = \frac{2}{15}\left(\frac{2}{5\sqrt{5}}\right)^r \frac{n^r}{(r-1)!} + \text{ smaller terms.}$$

Proof. We see that

$$G_n(e^t) = \sum_i p(i)e^{t(i-\mu)} = \phi(t), \quad \text{where the random variable } X = i - \mu.$$

Now define the Maclaurin series of $G_n(e^t)$ by

$$G_n(e^t) = \sum_r E_r(n)t^r.$$

By properties of probability generating functions mentioned earlier, $E_r(n) = \dfrac{E[(V-\mu)^r]}{r!}$.

The recurrence (1.2) becomes

$$G_n(e^t) = e^t \frac{f_{n-1}G_{n-1}(e^t)}{f_n e^{t(\mu_n - \mu_{n-1})}} + \frac{f_{n-2}G_{n-2}(e^t)}{f_n e^{t(\mu_n - \mu_{n-2})}} \quad , n \geq 2. \tag{1.3}$$

The series expansion of relation (1.3) leads to

$$G_n(e^t) = \frac{f_{n-1}}{f_n}\left[\sum_{r=0}^{\infty} \frac{(1-\mu_n+\mu_{n-1})^r t^r}{r!}\right] \cdot \left[\sum_{r=0}^{\infty} E_r(n-1)t^r\right]$$

$$+ \frac{f_{n-2}}{f_n}\left[\sum_{r=0}^{\infty} \frac{(-\mu_n+\mu_{n-2})^r t^r}{r!}\right] \cdot \left[\sum_{r=0}^{\infty} E_r(n-2)t^r\right].$$

By comparing coefficients of t^r, we obtain the relation

$$E_r(n) - \frac{f_{n-1}}{f_n}E_r(n-1) - \frac{f_{n-2}}{f_n}E_r(n-2)$$

$$= \frac{f_{n-1}}{f_n}a_n E_{r-1}(n-1) + \frac{f_{n-2}}{f_n}b_n E_{r-1}(n-2)$$

$$+ \frac{f_{n-1}}{f_n}\frac{a_n^2}{2!}E_{r-2}(n-1) + \frac{f_{n-2}}{f_n}\frac{b_n^2}{2!}E_{r-2}(n-2)$$

$$+ \frac{f_{n-1}}{f_n}\frac{a_n^3}{3!}E_{r-3}(n-1) + \frac{f_{n-2}}{f_n}\frac{b_n^3}{3!}E_{r-3}(n-2) + \ldots.$$

where $a_n = 1 - \mu_n + \mu_{n-1} \approx 1 - \dfrac{1}{\sqrt{5}}$ and $b_n = -\mu_n + \mu_{n-2} \approx -\dfrac{2}{\sqrt{5}}$.

We make a remark that $\dfrac{b_n}{a_n} \approx \dfrac{1+\sqrt{5}}{2}$.

With this new relation, we can prove the theorem by simply applying an induction on r. The base cases when $r = 0$ and $r = 1$ are obtained from the relation above, along with the values of $E_0(n), E_1(n)$ when $n = 1, 2, 3$ (for example). The induction step can be shown as follows.

Case 1 (Even):

Left hand side:

$$E_{2r}(n) - \frac{f_{n-1}}{f_n} E_{2r}(n-1) - \frac{f_{n-2}}{f_n} E_{2r}(n-2)$$

$$= \left(\frac{f_{n-1}}{f_n} + 2 \frac{f_{n-2}}{f_n} \right) \left(\frac{2}{5\sqrt{5}} \right)^r \frac{n^{r-1}}{(r-1)!} + \text{ smaller terms.}$$

Right hand side:

$$\frac{f_{n-1}}{f_n} a_n E_{2r-1}(n-1) + \frac{f_{n-2}}{f_n} b_n E_{2r-1}(n-2) +$$

$$\frac{f_{n-1}}{f_n} \frac{a_n^2}{2} E_{2r-2}(n-1) + \frac{f_{n-2}}{f_n} \frac{b_n^2}{2} E_{2r-2}(n-2) + \text{ smaller terms}$$

$$= \left(\frac{f_{n-1}}{f_n} \frac{a_n^2}{2} + \frac{f_{n-2}}{f_n} \frac{b_n^2}{2} \right) \left(\frac{2}{5\sqrt{5}} \right)^{r-1} \frac{n^{r-1}}{(r-1)!} + \text{ smaller terms.}$$

The fact that

$$\frac{f_{n-1}}{f_n} a_n + \frac{f_{n-2}}{f_n} b_n = 0, \quad \text{for all } n$$

and

$$\frac{2}{5\sqrt{5}} \left(\frac{f_{n-1}}{f_n} + 2 \frac{f_{n-2}}{f_n} \right) = \frac{f_{n-1}}{f_n} \frac{a_n^2}{2} + \frac{f_{n-2}}{f_n} \frac{b_n^2}{2}, \quad \text{for all } n$$

make both sides equal.

Case2 (Odd):

Left hand side:

$$E_{2r+1}(n) - \frac{f_{n-1}}{f_n} E_{2r+1}(n-1) - \frac{f_{n-2}}{f_n} E_{2r+1}(n-2)$$

$$= \left(\frac{f_{n-1}}{f_n} + 2 \frac{f_{n-2}}{f_n} \right) \frac{2}{15} \left(\frac{2}{5\sqrt{5}} \right)^r \frac{rn^{r-1}}{(r-1)!} + \text{ smaller terms}$$

Right hand side:

$$\frac{f_{n-1}}{f_n}a_n E_{2r}(n-1) + \frac{f_{n-2}}{f_n}b_n E_{2r}(n-2)+$$

$$\frac{f_{n-1}}{f_n}\frac{a_n^2}{2}E_{2r-1}(n-1) + \frac{f_{n-2}}{f_n}\frac{b_n^2}{2}E_{2r-1}(n-2)+$$

$$\frac{f_{n-1}}{f_n}\frac{a_n^3}{6}E_{2r-2}(n-1) + \frac{f_{n-2}}{f_n}\frac{b_n^3}{6}E_{2r-2}(n-2) + \text{ smaller terms}$$

$$= -\left(\frac{f_{n-1}}{f_n}a_n + 2\frac{f_{n-2}}{f_n}b_n\right)\left(\frac{2}{5\sqrt{5}}\right)^r \frac{n^{r-1}}{(r-1)!}$$

$$+ \left(\frac{f_{n-1}}{f_n}\frac{a_n^2}{2} + \frac{f_{n-2}}{f_n}\frac{b_n^2}{2}\right)\frac{2}{15}\left(\frac{2}{5\sqrt{5}}\right)^{r-1}\frac{n^{r-1}}{(r-2)!}$$

$$+ \left(\frac{f_{n-1}}{f_n}\frac{a_n^3}{6} + \frac{f_{n-2}}{f_n}\frac{b_n^3}{6}\right)\left(\frac{2}{5\sqrt{5}}\right)^{r-1}\frac{n^{r-1}}{(r-1)!} + \text{ smaller terms}$$

After some calculations, the computer program verifies that they are equal . □

Next, we show the asymptotic normality of V which follows directly from Theorem 1.

Corollary 1.3. *The distribution of numbers of V on a tiling of a 2-by-n board is asymptotically normal.*

Proof. To show the normality of V, we show that $\dfrac{V-\mu}{\sigma_V} \sim N(0,1)$ as $n \to \infty$. We verify, as $n \to \infty$, $m_{2r} = \dfrac{(2r)!}{2^r r!}$ and $m_{2r+1} = 0$, for every r.

$$\lim_{n\to\infty} m_{2r} = \lim_{n\to\infty}\frac{E[(V-\mu)^{2r}]}{E[(V-\mu)^2]^r} = \lim_{n\to\infty}\frac{\left(\frac{2}{5\sqrt{5}}\right)^r \frac{n^r}{r!}(2r)!}{\left(\frac{4n}{5\sqrt{5}}\right)^r} = \frac{(2r)!}{2^r r!},$$

and

$$\lim_{n\to\infty} m_{2r+1} = \lim_{n\to\infty}\frac{E[(V-\mu)^{2r+1}]}{E[(V-\mu)^2]^{(r+1/2)}} = \lim_{n\to\infty}\frac{\frac{2}{15}\left(\frac{2}{5\sqrt{5}}\right)^r \frac{n^r}{(r-1)!}(2r+1)!}{\left(\frac{4n}{5\sqrt{5}}\right)^{r+1/2}} = 0.$$

□

Remark The conjectures of the straight moments $E[V^r]$ can be shown by the same method but with fewer calculations.

2. STATISTICS ON AN m-BY-n BOARD

Let's turn our attention to the more general statistics $E[V^a H^b]$ on a more general m-by-n board when m is fixed and n is symbolic. The guess-and-check method still works here. We will briefly discuss it. We first consider the number of possible ways to put dominoes on the board. For this purpose, we define

$$S_{(a,b)}(m,n) := \sum_B V^a H^b.$$

The data of $S_{(0,0)}$ from the program are

$m:n$	1	2	3	4	5	6	7	8	9	10	11
1	0	1	0	1	0	1	0	1	0	1	0
2	1	2	3	5	8	13	21	34	55	89	144
3	0	3	0	11	0	41	0	153	0	571	0
4	1	5	11	36	95	281	781	2245	6336	18061	51205
5	0	8	0	95	0	1183	0	14824	0	185921	0
6	1	13	41	281	1183	6728	31529	167089	817991	4213133	21001799

The recurrence relations of each row are as follows:

$$m = 2: \ S(2,n) = S(2,n-1) + S(2,n-2), \quad \text{(Fibonacci)}.$$
$$m = 3: \ S(3,n) = 4S(3,n-2) - S(3,n-4),$$
$$m = 4: \ S(4,n) = S(4,n-1) + 5S(4,n-2) + S(4,n-3) - S(4,n-4).$$

All of these relations (and more) are not hard to show. The recurrence will satisfy a linear system with constant coefficients which comes from solving a system of recurrence equations. Hence, all we need to do is to verify the recurrence up to some numeric values.

2.1. The Grand Generating Function of an m-by-n Board with a Fixed m.

We make a generalization of the grand generating function from a 2-by-n board size to an m-by-n one.

$$m = 2 : H(v,h,t) = \frac{1}{1 - vt - h^2t^2},$$
$$m = 3 : H(v,h,t) = \frac{1 - h^3t^2}{1 - 2h(h^2 + v^2)t^2 + h^6t^4},$$
$$m = 4 : H(v,h,t) = \frac{1 - h^4t^2}{1 - v^2t - h^2(3v^2 + 2h^2)t^2 - h^4v^2t^3 + h^8t^4},$$
$$\cdots$$

For fixed numbers m, the recurrences of $S_{(a,b)}(m,n)$ also follow from this grand generating function. For example:

$S_{(a,b)}(2,n)$ where $a + b = t$, $t \geq 0$ satisfies the recurrence equation

$$(N^2 - N - 1)^{t+1} = 0.$$

$S_{(a,b)}(3,n)$ where $a + b = t$, $t \geq 0$ satisfies the recurrence equation

$$(N^4 - 4N^2 + 1)^{t+1} = 0.$$

$S_{(a,b)}(4,n)$ where $a + b = t$, $t \geq 0$ satisfies the recurrence equation

$$(N^4 - N^3 - 5N^2 - N + 1)^{t+1} = 0.$$

$\cdots$

The calculation of higher moments to show normality are more difficult (although it could be done). However, it is known that the sequence of random variables X_n, whose grand generating function is rational in v, h and t, is asymptotically normal, see [4].

In conclusion, the results and method presented here contain many interesting features, which we outline below.

(1) This is an automated method to generate conjectures and a semi-automated method to prove them. The method is totally elementary and involves only basic knowledge of probability.
(2) This is a classical implementation of symbolic computation, which not only shows the normality of the distribution but also gives more terms of the moments of interest.
(3) It supplies many new tiling identities which could be of interest to the bijectors out there.

APPENDICES

Appendix A. About the program

The process of gathering the data, making conjectures and proving them by induction have been implemented in the program `Domino.txt`. We enumerate the main commands below. Please refer to the author's web site for complete details of the program.

MoV(r, n)

    Input: the r-th moment the length of the board n.

    Output: the r-th moment of V on 2-by-n board.

    Try: seq(MoV$(1, n)$, $n = 1..10$);

ConjMoV(r, n)

    Input: non-negative integer r, symbolic n.

    Output: the conjectured formula of the r-th moment of V in term of n.

    Try: ConjMoV$(2, n)$;

BigConjMoV(k, n, r)

    Input: non-negative integer k, symbolic n and r.

    Output: the k-th biggest term of Moment V-formula as variables in n and r.

    Try: BigConjMoV$(2, n, r)$;

GenVm(m, n, v)

    Input: numeric m, n and symbolic v.

    Output: generating function of v (vertical tile) from the board of size m-by-n computed from empirical data.

    Try: seq(GenVm$(4, n, v)$, $n = 1..10$);

THE FIBONACCI QUARTERLY

References

[1] P. W. Kasteleyn, *The statistics of dimers on a lattice: I. The number of dimer arrangements in a quadratic lattice*, Physica **27** (1961), 1209-1225.

[2] D. Zeilberger, *Automatic Count Tilings*, The Personal Journal of Shalosh B. Ekhad and Doron Zeilberger, Jan. 20, 2006.

[3] D. Zeilberger, *The Automatic Central Limit Theorems Generator (and Much More!)*, Advances in Combinatorial Mathematics: Proceedings of the Waterloo Workshop in Computer Algebra 2008 in honor of Georgy P. Egorychev", chapter 8, pp. 165-174.

[4] D. Zeilberger, *Automated Derivation of Limiting Distributions Of Combinatorial Random Variables Whose Generating Functions are Rational*, The Personal Journal of Shalosh B. Ekhad and Doron Zeilberger, Dec. 24, 2016.

MSC2010: 11B39, 05-04, 05A16

Science Division, Mahidol University International College, Nakornpathom, Thailand

E-mail address: thotsaporn@gmail.com