Moments of Matching Statistics

Catherine Yan

Department of Mathematics
Texas A&M University
College Station, TX 77843, USA
joint with Niraj Khare and Rudolph Lorentz

AMS Southeastern Sectional Meeting, October 2015

[Chern, Diaconis, Kane, Rhoades 2014] Closed Expressions for Averages of Set Partition Statistics

Theorem (CDKR)

For a family of combinatorial statistics, the moments have simple closed expressions as linear combinations of shifted Bell numbers, where the coefficients are polynomials in n.

Bell number B_n: number of partitions of a set of size n.
combinatorial statistics: number of blocks, k-crossings, k-nestings, dimension exponents, occurrence of patterns, etc.
Expression:

$$\sum_{\lambda \in \Pi(n)} f^k(\lambda) = \sum_{I \leq j \leq K} Q_j(n) B_{n+j}.$$

What happens for matchings?

matchings: partitions of $[2m]$ in which every block has size 2.
Objective: closed formula for moments of combinatorial statistics on matchings $\mathcal{M}(2m)$

What happens for matchings?

matchings: partitions of $[2m]$ in which every block has size 2.
Objective: closed formula for moments of combinatorial statistics on matchings $\mathcal{M}(2m)$

Theorem (Khare, Lorentz, Y)

For a family of combinatorial statistics, the moments have simple closed expressions as linear combinations of double factorials $T_{2m} = (2m-1)!! = (2m-1)(2m-3)\cdots 3\cdot 1$, *with constant coefficients.*

What happens for matchings?

matchings: partitions of $[2m]$ in which every block has size 2.
Objective: closed formula for moments of combinatorial statistics on matchings $\mathcal{M}(2m)$

Theorem (Khare, Lorentz, Y)

For a family of combinatorial statistics, the moments have simple closed expressions as linear combinations of double factorials $T_{2m} = (2m-1)!! = (2m-1)(2m-3)\cdots 3 \cdot 1$, *with constant coefficients.*

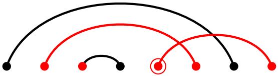
- what kind of (combinatorial) statistics
- general linear combination formula
- How does combinatorial structures help

Definition

1. *A pattern* $\underline{P} := (P, A(P), C(P))$ *of length* k *is a partial matching* P *on* $[k]$ *with a set of arcs* $A(P)$ *and a set of vertices* $C(P) \subseteq [k-1]$.
2. *An occurrence of a pattern* $\underline{P}$ *of length* k *in* $M \in \mathcal{M}_{2m}$ *is a tuple* $s := (t_1, t_2, \cdots, t_k)$ *with* $t_i \in [2m]$ *such that*
   1. $t_1 < t_2 < \cdots < t_k$.
   2. (t_i, t_j) *is an arc of* M *if* $(i, j) \in A(P)$.
   3. $t_{i+1} = t_i + 1$ *whenever* $i \in C(P)$.

   Write $s \in_{\underline{P}} M$ *if* s *is an occurrence of* $\underline{P}$ *in* M.

An occurence of pattern P of length 5 with $A(P) = \{(1,4),(3,5)\}$ and $C(P) = \{3\}$.



Family of Statistics

Simple statistic: a pattern $\underline{P}$ of length k and a valuation polynomial $Q \in \mathbb{Q}[y_1, y_2, \cdots, y_k, n]$,
If $M \in \mathcal{M}_{2m}$ and $s = (x_1, x_2, \cdots, x_k) \in_{\underline{P}} M$, then

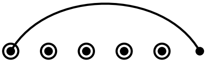
$$f(M) = f_{\underline{P},Q}(M) := \sum_{s \in_{\underline{P}} M} Q(s, m).$$

degree of f:= length of P + degree of Q

General statistic: a finite linear combination of simple statistics.

Example–patterns

- Arcs of fixed length



- crossings and nestings

- left-neighboring crossings/nestings

Examples (con't)

- dimension exponents $d(\lambda) = \sum_{i=1}^{m}(M_i - m_i + 1) - 2m$.
  $A(P) = \{1, 2\}$, $C(P) = \emptyset$ and $Q(y_1, y_2, n) = y_2 - y_1 - 1$.
- Blocks of consecutive vertices $\{i, i+1\}$
  $A(P) = \{1, 2\}$ and $C(P) = \{1\}$, $Q = 1$.

Not include: the length of longest arc, size of maximal crossings/nestings, ...

First moment –simple statistic

For any statistic f, define

$$M(f, 2m) := \sum_{M \in \mathcal{M}_{2m}} f(M).$$

For simple statistic $f_{\underline{P},Q}$ of degree N, let $\ell = |A(\underline{P})|$ and $c = |C(\underline{P})|$. We have

Theorem

$$M(f_{\underline{P},Q}, 2m) = P(m)T_{2(m-\ell)}$$

where $P(x)$ is a polynomial of degree no more than $N - c$. Equivalently,

$$M(f_{\underline{P},Q}, 2m) = \begin{cases} 0 & m < \ell \\ \sum_{-\ell \le i \le N-\ell-c} c_i T_{2(m+i)} & m \ge \ell \end{cases} \quad (1)$$

with constants c_i.

First moment– general case

Theorem

For any statistic f of degree N, there is an integer $L \leq N/2$ such that

$$M(f, 2m) = R(m)T_{2(m-L)} = \sum_{-L \leq i \leq N} d_i T_{2(m+i)} \qquad (m \geq L) \quad (2)$$

where $R(x)$ are polynomials of degree no more than $N + L$.

Corollary

Let f be a simple statistic with pattern $\underline{P}$ and the valuation function $Q = 1$. Then

$$M(f, 2m) = T_{2(m-\ell)} \binom{2m-c}{k-c}.$$

Higher moments

Theorem (CDKR)

Let $\mathcal{S}$ be the set of all statistics thought of as functions $f : \cup_m \mathcal{M}_{2m} \to \mathbb{Q}$. Then $\mathcal{S}$ is closed under the operations of pointwise scaling, addition and multiplication. Thus, if f_1, f_2 $\in \mathcal{S}$ and $a \in \mathbb{Q}$, then there exist matching statistics g_a, g_+ and g_* so that for all matching M,

$$\begin{aligned} af_1(M) &= g_a(M), \\ f_1(M) + f_2(M) &= g_+(M), \\ f_1(M)f_2(M) &= g_*(M). \end{aligned}$$

Furthermore, $d(g_a) \leq d(f_1)$, $d(g_+) \leq \max\{d(f_1), d(f_2)\}$ *and* $d(g_*) \leq d(f_1) + d(f_2)$.

Combinatorially, product of f_1 and f_2 can be computed by considering all the ways to *merge* two patterns.

General formula

Theorem

For any statistic f of degree N and positive integer r, we have

$$M(f^r, 2m) = \sum_{I \leq i \leq J} d_i T_{2(m+i)} \qquad \text{whenever } m \geq |I| \tag{3}$$

where I and J are constants bounded by $I \geq -\frac{rN}{2}$ and $J \leq rN$.

Special form for simple patterns

If f is the occurence of a simple pattern with no isolated vertices, i.e., $\ell = k/2$, $C(P) = \emptyset$ and $Q = 1$.

Theorem

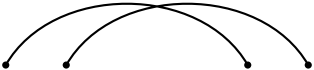
For $m \geq \ell$, the r-th moment can be expressed as

$$M(f^r, 2m) = \sum_{i=0}^{(r-1)\ell} c_i^{(r)} \binom{2m}{2(\ell+i)} T_{2(m-\ell-i)}. \tag{4}$$

Note: $\binom{a}{b} = 0$ if $a < b$, and $T_{2k} = 0$ if $k < 0$. Hence for $m = \ell, \ell+1, \ldots, \ell r$, Eqs.(4) is a triangular system, which gives a linear recurrence for the coefficients.

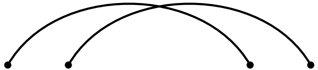
Example: 2-crossings

Let f be the number of 2-crossings, so $\ell = 2$. Let $r = 2$.



Example: 2-crossings

Let f be the number of 2-crossings, so $\ell = 2$. Let $r = 2$.

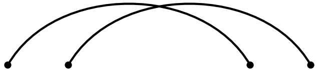


$$M(f^2, 2m) = c_0 \binom{2m}{4} T_{2m-4} + c_1 \binom{2m}{6} T_{2m-6} + c_2 \binom{2m}{8} T_{2m-8}.$$

Data:
If $m = 2$, $M(f^2, 4) = 1$ gives $c_0 = 1$.
If $m = 3$, $M(f^2, 6) = 27$ gives $c_1 = 12$.
If $m = 4$, $M(f^2, 8) = 616$ gives $c_2 = 70$.

Example: 2-crossings

Let f be the number of 2-crossings, so $\ell = 2$. Let $r = 2$.



$$M(f^2, 2m) = c_0 \binom{2m}{4} T_{2m-4} + c_1 \binom{2m}{6} T_{2m-6} + c_2 \binom{2m}{8} T_{2m-8}.$$

Data:
If $m = 2$, $M(f^2, 4) = 1$ gives $c_0 = 1$.
If $m = 3$, $M(f^2, 6) = 27$ gives $c_1 = 12$.
If $m = 4$, $M(f^2, 8) = 616$ gives $c_2 = 70$.

Theorem

The second moment of k-crossings equals the second moment of k-nestings.

Simple patterns II: $Q = 1$ but $C(P) \neq \emptyset$

Note: for $2m - c \geq 0$,

$$\binom{2m-c}{2\ell-c} T_{2(m-\ell)} = \begin{cases} P(m)T_{2m-c} & \text{if } c \text{ is even} \\ Q(m)T_{2m-c+1} & \text{if } c \text{ is odd,} \end{cases} \tag{5}$$

where $P(x)$ is a polynomial of degree $\ell - \frac{c}{2}$, and $Q(x)$ is a polynomial of degree $\ell - \frac{c+1}{2}$.
Let ℓ be the number of arcs in P. Hence

Theorem

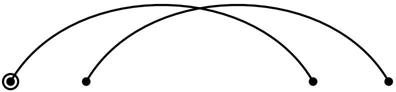
For any positive integer r and $m \geq r(\ell-1)/2$, there is a closed formula

$$M(f^r, 2m) = \sum_{I \leq i \leq J} d_j T_{2(m+j)},$$

where I and J are constants such that $I \geq -r(\ell-1)/2$ and $J \leq (r-1)\ell + 1$.

Example: 2-crossings with left neighboring vertices

Consider the pattern $\underline{P}$ with $A(P) = \{(1,3),(2,4)\}$ and $C(P) = \{1\}$.



$$M((f_P)^2, 2m) = -\frac{1}{6}T_{2(m-1)} + \frac{1}{4}T_{2m} - \frac{1}{6}T_{2(m+1)} + \frac{1}{36}T_{2(m+2)}.$$

$$\begin{aligned} M((f_P)^3, 2m) &= \frac{1}{4}T_{2(m-1)} - \frac{5}{24}T_{2m} + \frac{11}{120}T_{2(m+1)} \\ &\quad -\frac{1}{24}T_{2(m+2)} + \frac{1}{216}T_{2(m+3)}. \end{aligned}$$

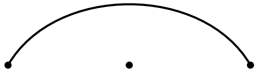
Example: dimension exponent

$$d(M) = -m + \sum_{i=1}^{m} (M_i - m_i).$$

It has $A(P) = \{1, 2\}$, $C(P) = \emptyset$, and $Q(y_1, y_2, m) = y_2 - y_1 - 1$.

Proposition

$d(M)$ also counts the number of occurrence of the pattern T of length 3 with $A(T) = \{(1, 3)\}$ and $C(T) = \emptyset$.



Thus we have the case that $C(P) = \emptyset$ and $Q = 1$.

Theorem

For any positive m *and* r,

$$M(d(M)^r, 2m) = \sum_{j=0}^{2r} d_j T_{2(m+j)}$$

for some constants d_j.

For example,

$$M(d(M), 2m) = \frac{1}{2} T_{2m} - T_{2(m+1)} + \frac{1}{6} T_{2(m+2)}.$$

and

$$M(d(M)^2, 2m) = \frac{1}{4} T_{2m} - \frac{8}{3} T_{2(m+1)} + \frac{5}{2} T_{2(m+2)} - \frac{8}{15} T_{2(m+3)} + \frac{1}{36} T_{2(m+4)}$$