Combinatorial statistics

Given a set of combinatorial objects $\mathcal{A}$, a **combinatorial statistic** is an integer given to every element of the set.

In other words, it is a function $\mathcal{A} \to \mathbb{Z}_{\geq 0}$.

Example. Let $\mathcal{S}$ be the set of subsets of $\{1, 2, 3\}$.
The cardinality of a set is a combinatorial statistic on $\mathcal{S}$.

$$\big|\emptyset\big| = 0 \quad \big|\{1\}\big| = 1 \quad \big|\{2\}\big| = 1 \quad \big|\{3\}\big| = 1$$
$$\big|\{1,2\}\big| = 2 \quad \big|\{1,3\}\big| = 2 \quad \big|\{2,3\}\big| = 2 \quad \big|\{1,2,3\}\big| = 3$$

Combinatorial statistics provide a *refinement* of counting.

less information $\longleftarrow\!\!-\!\!-\!\!-\!\!-\!\!\longrightarrow$ *more information*

counting

8

statistics

0	1	2	3
1	3	3	1

complete enumeration

$\emptyset \quad \{1\} \quad \{2\} \quad \{3\}$

$\{1,2\}\{1,3\}\{2,3\}\{1,2,3\}$

Statistics and Permutations

Questions involving combinatorial statistics:

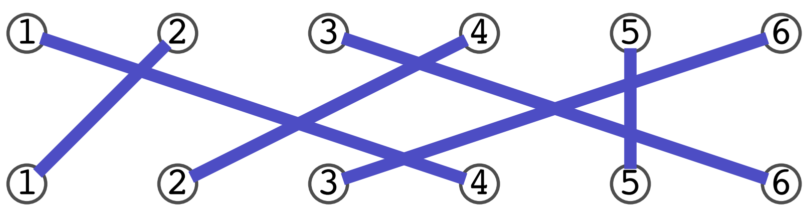
- What is the *distribution* of the statistics?
- What is the *average size* of an object in the set?
- Which statistics have the same distribution?
  - Insight into their structure.
  - Provides non-trivial bijections in the set?

A especially rich playground involves *permutation statistics*.

Representations of permutations

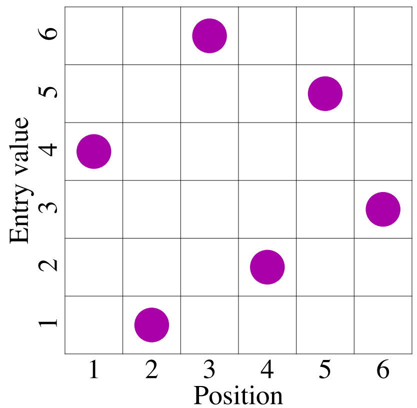
One-line notation: $\pi = 4\,1\,6\,2\,5\,3$ Cycle notation: $\pi = (1\,4\,2)(3\,6)(5)$

String diagram:



(only two crossings at a time)

Matrix-like diagram:

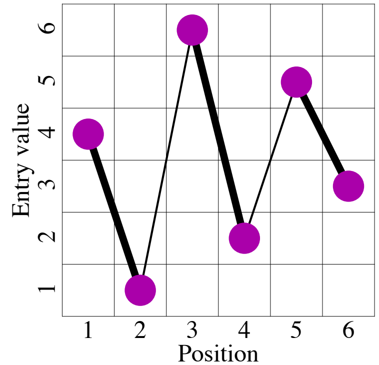


Descent statistic

Definition: Let $\pi = \pi_1\pi_2\cdots\pi_n$ be a permutation.

A **descent** is a position i such that $\pi_i > \pi_{i+1}$.

Define $\mathrm{des}(\pi)$ to be the **number of descents** in π.



Example. When $\pi = 416253$, $\mathrm{des}(\pi) = 3$ since $4 \searrow 1$, $6 \searrow 2$, $5 \searrow 3$.

Question: How many n-permutations have d descents?

$\mathrm{des}(12) = 0$ $\mathrm{des}(123) = \_\_\_$ $\mathrm{des}(213) = \_\_\_$ $\mathrm{des}(312) = \_\_\_$

$\mathrm{des}(21) = 1$ $\mathrm{des}(132) = \_\_\_$ $\mathrm{des}(231) = \_\_\_$ $\mathrm{des}(321) = \_\_\_$

$n \backslash d$	0	1	2	3	4
1	1				
2	1	1			
3	1	4	1		
4	1	11	11	1	
5	1	26	66	26	1

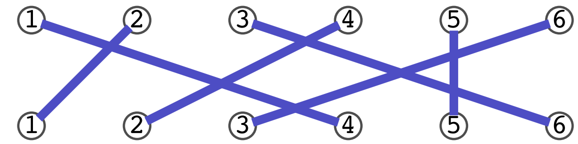
What are the possible values for $\mathrm{des}(\pi)$?

Note the symmetry. If π has d descents, its reverse $\hat{\pi}$ has _______ descents.

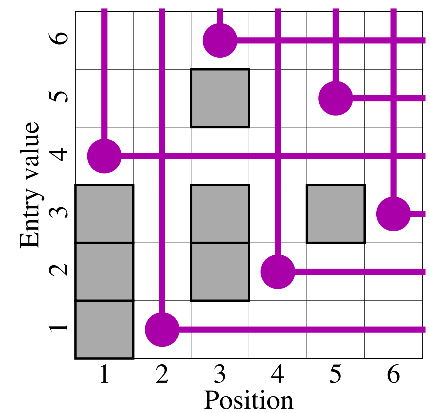
These are the **Eulerian numbers**.

Inversion statistic

Definition: Let $\pi = \pi_1 \pi_2 \cdots \pi_n$ be a permutation. An **inversion** is a pair $i < j$ such that $\pi_i > \pi_j$.



Define $\mathrm{inv}(\pi)$ as the **number of inversions** in π.



Example. When $\pi = 416253$, $\mathrm{inv}(\pi) = 7$ since $4 > 1$, $4 > 2$, $4 > 3$, $6 > 2$, $6 > 5$, $6 > 3$, $5 > 3$.

In a string diagram $\mathrm{inv}(\pi) =$ number of crossings.

In a matrix diagram $\mathrm{inv}(\pi)$, draw *Rothe diagram*:

$\mathrm{inv}(12) = 0$ $\quad\mathrm{inv}(123) =$ ___ $\quad\mathrm{inv}(213) =$ ___ $\quad\mathrm{inv}(312) =$ ___

$\mathrm{inv}(21) = 1$ $\quad\mathrm{inv}(132) =$ ___ $\quad\mathrm{inv}(231) =$ ___ $\quad\mathrm{inv}(321) =$ ___

$n \backslash i$	0	1	2	3	4	5	6
1	1						
2	1	1					
3	1	2	2	1			
4	1	3	5	6	5	3	1

What are the possible values for $\mathrm{inv}(\pi)$?

The inversion number is a good way to count how "far away" a permutation is from the identity.

Major index

Definition: Let $\pi = \pi_1 \pi_2 \cdots \pi_n$ be a permutation.

Define maj(π), the **major index** of π, to be sum of the descents of π.

[Named after Major Percy MacMahon. (British army, early 1900's)]

Example. When $\pi = 416253$, maj$(\pi) = 9$ since the descents of π are in positions 1, 3, and 5.

maj(12) = 0 maj(123) = ___ maj(213) = ___ maj(312) = ___
maj(21) = 1 maj(132) = ___ maj(231) = ___ maj(321) = ___

$n \backslash m$	0	1	2	3	4	5	6
1	1						
2	1	1					
3	1	2	2	1			
4	1	3	5	6	5	3	1

What are the possible values for maj(π)?

The distribution of maj(π)
IS THE SAME AS
the distribution of inv(π)!

A statistic that has the same distribution as inv is called **Mahonian**.

q-analogs

Definition: A q-**analog** of a number c is an expression $f(q)$ such that $\lim_{q\to 1} f(q) = c$.

Example. $\dfrac{1-q^n}{1-q} = \left(1 + q + q^2 + \cdots + q^{n-2} + q^{n-1}\right)$ is a q-analog of n because $\lim_{q\to 1} \frac{1-q^n}{1-q} = n$.

We write $[n]_q = \frac{1-q^n}{1-q}$.

q-analogs work hand in hand with combinatorial statistics.

If stat is a combinatorial statistic on a set S ($\mathrm{stat} : S \mapsto \mathbb{N}$), then $\displaystyle\sum_{s\in S} q^{\mathrm{stat}(s)}$ is a q-analog of $|S|$ because

$$\lim_{q\to 1} \sum_{s\in S} q^{\mathrm{stat}(s)} = \sum_{s\in S} 1^{\mathrm{stat}(s)} = \sum_{s\in S} 1 = |S|$$

Inversion statistics

Question: What is the generating function $\sum_{\pi \in S_n} q^{\mathrm{inv}(\pi)}$?

n	$\sum_{\pi \in S_n} q^{\mathrm{inv}(\pi)}$	
1	$1q^0$	$= 1$
2	$1q^0 + 1q^1$	$= (1+q)$
3	$1q^0 + 2q^1 + 2q^2 + 1q^3$	$= (1+q+q^2)(1+q)$
4	$1q^0 + 3q^1 + 5q^2 + 6q^3 + 5q^4 + 3q^5 + 1q^6$	$=$

Conjecture: $\sum_{\pi \in S_n} q^{\mathrm{inv}(\pi)} = [n]_q \cdots [1]_q =: [n]_q!$, the q-**factorial**.

Claim: This equation makes sense when $q = 1$.

Inversion Statistics

Theorem: $\sum_{\pi \in S_n} q^{\mathrm{inv}(\pi)} = [n]_q!$

Proof. There exists a bijection

$$\left\{ \begin{array}{c} \text{permutations} \\ \pi \in S_n \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \text{lists } (a_1, \ldots, a_n) \\ \text{where } 0 \le a_i \le n-i \end{array} \right\}.$$

Given a permutation π, create its **inversion table**. Define a_i to be the number of entries j to the left of i that are smaller than i. Then $\mathrm{inv}(\pi) = a_1 + a_2 + \cdots + a_n$.

Example. The inversion table of $\pi = 4\,3\,1\,5\,2$ is $(3, 2, 0, 1, 0)$.

$$\begin{aligned} \sum_{\pi \in S_n} q^{\mathrm{inv}(\pi)} &= \sum_{a_1=0}^{n-1} \sum_{a_2=0}^{n-2} \cdots \sum_{a_n=0}^{0} q^{a_1+a_2+\cdots+a_n} \\ &= \left(\sum_{a_1=0}^{n-1} q^{a_1} \right) \left(\sum_{a_2=0}^{n-2} q^{a_2} \right) \cdots \left(\sum_{a_n=0}^{0} q^{a_n} \right) \\ &= \quad [n]_q \qquad [n-1]_q \quad \cdots \quad [1]_q \quad = \quad [n]_q! \end{aligned}$$

Notes

We said that inv and maj are equidistributed. Two possible proofs:

- Find a bijection $f : S_n \to S_n$ such that $\mathrm{maj}(\pi) = \mathrm{inv}(f(\pi))$.
- Or prove $\displaystyle\sum_{\pi \in S_n} q^{\mathrm{inv}(\pi)} = \sum_{\pi \in S_n} q^{\mathrm{maj}(\pi)}$.

With a q-analog of factorials, we can define a q-analog of binomial coefficients. Define

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{[n]_q!}{[k]_q![n-k]_q!}.$$

These *polynomials* are called the **q-binomial coefficients** or **Gaussian polynomials**.

- $\lim_{q\to 1} \begin{bmatrix} n \\ k \end{bmatrix}_q = \binom{n}{k}$.
- They are indeed polynomials.
- Example. $\begin{bmatrix} 4 \\ 2 \end{bmatrix}_q = 1 + q + 2q^2 + q^3 + q^4$

Combinatorial interpretations of q-binomial coefficients!

Combinatorial interpretations of q-binomial coefficients

Consider set $S_{k,n-k}$ of permutations of the multiset $\{1^k, 2^{n-k}\}$.
Define $\mathrm{inv}(\pi) = |\{i < j \colon \pi(i) > \pi(j)\}|$.

Example. $\pi = 1\,1\,2\,2\,1\,2\,1\,1\,2\,2$ is a permutation of $\{1^5, 2^5\}$.
Then $\mathrm{inv}(\pi) = 0+0+3+3+0+2+0+0+0+0 = 8$.

Then $\sum_{\pi \in S_{k,n-k}} q^{\mathrm{inv}(\pi)} = \begin{bmatrix} n \\ k \end{bmatrix}_q$. (Note $|S_{k,n-k}| = \binom{n}{k}$.)

This is a refinement of these permutations in terms of inversions.

Consider the set $\mathcal{P}$ of lattice paths from $(0,0)$ to (a,b).
Let $\mathrm{area}(P)$ be the area above a path P.
Then $\sum_{P \in \mathcal{P}} q^{\mathrm{area}(P)} = \begin{bmatrix} a+b \\ a \end{bmatrix}_q$. (Note $|\mathcal{P}| = \binom{a+b}{a}$.)

This can also be used to give a q-analog of the Catalan numbers.

There's always more to learn!!!

References :

Miklós Bóna. Combinatorics of Permutations, CRC, 2004.

T. Kyle Petersen. Two-sided Eulerian numbers via balls in boxes. http://arxiv.org/abs/1209.6273

The Combinatorial Statistic Finder. http://findstat.org/