CYCLES IN RANDOM PERMUTATIONS

WENBO GAO

Abstract. We provide double counting proofs for several results on the statistics of cycles in random permutations, which were discussed by Tao.

Let σ be a permutation, chosen randomly from the uniform distribution on $\mathcal{S}_n$. Let C_k denote the number of cycles in σ of length k. In his 2011 post [1], Terry Tao lists several identities involving C_k, and comments that the identities can be proven by double counting. We provide such proofs, though some appear to require more computation than suggested by Tao. The reader is encouraged to find more elegant ways of counting to produce these results.

Theorem 1. $\mathbb{E}C_k = \frac{1}{k}$.

Proof. Rearrange and use the definition of expectation to obtain an equivalent expression:

$$\sum_{\sigma\in\mathcal{S}_n} kC_k(\sigma) = n!$$

Consider the set $A = \{(a,\sigma) : \sigma \in \mathcal{S}_n,\ a \text{ in a } k\text{-cycle of } \sigma\}$ of pairs consisting of a permutation σ and an element a in a k-cycle of σ. The permutation σ has $C_k(\sigma)$ k-cycles, so $|A| = \sum_\sigma kC_k(\sigma)$. On the other hand, consider fixing the element a. Each permutation in which a is in a k-cycle is obtained by selecting a set B of $k-1$ elements from $[n]\setminus\{a\}$, selecting a cyclic permutation on $B\cup\{a\}$, and selecting an arbitrary permutation on the remaining $n-k$ elements. Thus, for each a, there are

$$\binom{n-1}{k-1}(k-1)!(n-k)! = (n-1)!$$

such permutations. Since there are n possible values of a, we obtain $|A| = n(n-1)! = n!$. □

Theorem 2. *If $jk \le n$, then $\mathbb{E}\binom{C_k}{j} = \frac{1}{k^j j!}$.*

Proof. Rearrange to obtain the equivalent expression

$$\sum_\sigma k^j j!\binom{C_k(\sigma)}{j} = n!$$

Consider the set $A = \{(a_1,\dots,a_j,\sigma) : a_1,\dots,a_j \text{ in distinct } k\text{-cycles of } \sigma\}$. Note that $a_1,\dots,a_j$ are *ordered*, so $(a_2,a_1,\dots,a_j,\sigma)$ is considered distinct from $(a_1,a_2,\dots,a_j,\sigma)$.

We first count A by counting the number of such tuples $(a_1,\dots,a_j)$ for each σ. Each tuple is obtained from an ordered selection $S_1,\dots,S_j$ of k-cycles of σ. The element a_i is then chosen from S_i. Thus, the number of such tuples for σ is $k^j j!\binom{C_k(\sigma)}{j}$. Summing over all σ, we have $|A| = \sum_\sigma k^j j!\binom{C_k(\sigma)}{j}$.

Next, suppose we fix $a_1,\dots,a_j$. To count the number of permutations σ such that $a_1,\dots,a_j$ are in distinct k-cycles, we encode σ as follows. Let s_i be the word representation of the permutation obtained by removing a_i from its cycle, and let s_{j+1} be the word form of the remaining cycles. We encode σ as the concatenated word $s_1s_2\dots s_{j+1}$. This gives a bijection between the desired permutations σ and the permutations on $\mathcal{S}_{n-j}$. Thus, since there are $n(n-1)\dots(n-j+1)$ ways of choosing $a_1,\dots,a_j$, we find that $|A| = n(n-1)\dots(n-j+1)(n-j)! = n!$. □

Date: 2016.

While this double counting proof of Theorem 2 is elegant and simple, it requires us to already know the value of the expectation. Deriving Theorem 2 using algebra is not entirely trivial, and a good exercise in manipulating generating functions.

Theorem 3. *We show that* $\mathbb{E}\binom{C_k}{j} = \frac{1}{k^j j!}$ *without using the final value of the expectation.*

Proof. Consider the combinatorial species of permutations, weighted by the number of k-cycles.

$$F(x,y) = \sum_{n=1}^{\infty} \left(\sum_{\sigma \in \mathcal{S}_n} y^{C_k(\sigma)} \right) \frac{x^n}{n!}$$

By decomposing each permutation into a set of cycles, we find that

$$F(x,y) = \exp\left(-\log(1-x) - \frac{x^k}{k} + y\frac{x^k}{k} \right) = \frac{1}{1-x} \exp((y-1)x^k/k)$$

Let b denote the largest integer such that $kb \leq n$. The coefficient of x^n in $F(x,y)$ is

$$p(y) = \sum_{r=0}^{b} \frac{(y-1)^r}{k^r r!}$$

For $q \leq b$, the coefficient of y^q in $p(y)$ is

$$c_q = \sum_{r=0}^{b} \frac{1}{k^r r!} \binom{r}{q} (-1)^{r-q}$$

Our desired expectation is given by

$$\mathbb{E}\binom{C_k}{j} = \sum_{q=0}^{b} \binom{q}{j} c_q$$

Substitute the expression for c_q and use the binomial identity $\binom{r}{q}\binom{q}{j} = \binom{r}{j}\binom{r-j}{q-j}$ to obtain

$$\begin{aligned} \sum_{q=0}^{b} \binom{q}{j} c_q &= \sum_{q=0}^{b} \sum_{r=0}^{b} \frac{(-1)^{r-q}}{k^r r!} \binom{r}{q} \binom{q}{j} \\ &= \sum_{r=0}^{b} \frac{(-1)^r}{k^r r!} \binom{r}{j} \sum_{q=0}^{b} (-1)^q \binom{r-j}{q-j} \end{aligned}$$

Suppose $s = r - j$, and let $t = q - j$. Observe that

$$\sum_{q=0}^{b} (-1)^q \binom{r-j}{q-j} = (-1)^j \sum_{t=0}^{s} (-1)^t \binom{s}{t} = (-1)^j \cdot \begin{cases} 0 & \text{if } s > 0 \\ 1 & \text{if } s = 0 \end{cases}$$

The last equality follows from the well-known bijection between odd and even subsets of $[s]$.

Thus, every term vanishes except for $r = j$, and we obtain

$$\mathbb{E}\binom{C_k}{j} = \frac{1}{k^j j!}$$

□

Theorem 4. *If* $\sum_{i=1}^{r} j_i k_i \leq n$, *then* $\mathbb{E} \prod_{i=1}^{r} \binom{C_{k_i}}{j_i} = \prod_{i=1}^{r} \frac{1}{k_i^{j_i} j_i!}$.

Proof. Use the same technique as in Theorem 2, but consider tuples $(b_1, \ldots, b_r)$, where b_i is itself a tuple $(a_1, \ldots, a_{j_i})$ with elements from distinct k_i-cycles. □

Theorem 5. *Let* $C = C_1 + \ldots + C_n$. *Then for* $m \in \mathbb{N}$, $\mathbb{E}m^C = \binom{n+m-1}{n}$.

Proof. Rearrange to write this as

$$\sum_{\sigma} m^{C(\sigma)} = \frac{(n+m-1)!}{(m-1)!}$$

Let A be the set of all sets $\{(s_1, 1), \ldots, (s_m, m)\}$, where $s_1, \ldots, s_m$ are permutations on disjoint subsets $S_1, \ldots, S_m$ covering $[n]$. We allow for $S_i = \emptyset$.

We first count A as follows. Taking the union of $s_1, \ldots, s_m$ defines a permutation on $[n]$. Consider all $(s_1, 1), \ldots, (s_m, m)$ which correspond to the same permutation. The labels i in (s_i, i) correspond to a mapping from the cycles of σ to $[m]$. The number of such mappings is $m^{C(\sigma)}$. Hence, we have $|A| = \sum_\sigma m^{C(\sigma)}$.

On the other hand, consider a set B of size $n+m-1$, consisting of n objects $1, \ldots, n$ and $m-1$ dividers $d_1, \ldots, d_{m-1}$. Consider a linear ordering ℓ of B. We treat the objects $d_1, \ldots, d_{m-1}$ as dividers separating ℓ into m (possibly empty) words

$$s_1 d_{i_1} s_2 d_{i_2} \ldots d_{i_{m-1}} s_m$$

We map this to the corresponding set $\{(s_1, 1), \ldots, (s_m, m)\}$ in A. There are $(n+m-1)!$ linear orderings ℓ of B. However, we do not wish to distinguish the labels on the dividers $d_1, \ldots, d_{m-1}$, merely their positions, so we divide by $(m-1)!$ to select one ordering ℓ from each class. Hence, $|A| = \frac{(n+m-1)!}{(m-1)!}$. □

Corollary 6. $\mathbb{E}2^C = n + 1$.

References

[1] Terry Tao. The number of cycles in a random permutation. `https://terrytao.wordpress.com/2011/11/23/the-number-of-cycles-in-a-random-permutation`, 2011.