Random Maps and Permutations

Kent E. Morrison

Department of Mathematics
California Polytechnic State University
San Luis Obispo, CA 93407
kmorriso@calpoly.edu

14 April 1998
Revised: 29 November 2003

1 Random Permutations

Much of this section deals with the cycles of a random permutation in S_n. Let C_k, $1 \leq k \leq n$, be the random variable that counts the number of k-cycles in a permutation. Note that C_1 is the number of fixed points.

Theorem 1 *The expected number of k-cycles is $1/k$.*

Proof Write C_k as the sum of indicator random variables $\mathbf{1}_\gamma$, for γ a k-cycle. This means that $\mathbf{1}_\gamma(\pi)$ is 1 if γ is a cycle of π and 0 otherwise. Then $E(C_k) = \sum_\gamma E(\mathbf{1}_\gamma)$. To determine $E(\mathbf{1}_\gamma)$ we count the number of permutations having γ as a cycle. That number is $(n-k)!$. Thus, $E(\mathbf{1}_\gamma) = (n-k)!/n!$. Now, the number of possible γ is $n(n-1)\cdots(n-k+1)/k$, since a k-cycle is an ordered selection of k elements from n in which any of the k elements can be put first. Thus, $E(C_k) = (n(n-1)\cdots(n-k+1)/k)((n-k)!/n!) = 1/k$. □

Corollary 2 *The expected number of cycles is the nth harmonic number $H_n = 1 + 1/2 + \cdots + 1/n$.*

Proof The total number of cycles is the random variable $\sum C_k$ and its expected value is $\sum E(C_k)$. □

An alternative proof of the corollary was outlined in the exercises fall quarter. It uses the generating function

$$f(z) = \sum_{k=1}^{n} \begin{bmatrix} n \\ k \end{bmatrix} z^k \tag{1}$$

Then the expected number of cycles is $f'(1)/n!$, and the derivative can be calculated using properties of the Stirling numbers.

Next we consider the distribution of C_i, that is we calculate the probability that a random permutation has k i-cycles. First we will deal with the case that $k = 0$ and i is arbitrary. For example, the case that $k = 0$ and $i = 1$ gives the probability of a derangement. The standard way to count derangements is to use the principal of inclusion-exclusion, and that is the way to deal with this more general problem. Let $d_i(n)$ be the number of permutations in S_n with no i-cycles. (Hence, $d_1(n)$ is the number of derangements.)

Theorem 3 *The number of permutations in S_n with no i-cycles is*

$$d_i(n) = n! \sum_{j=0}^{\lfloor n/i \rfloor} (-1)^j \frac{1}{j! i^j}.$$

Proof For each i-cycle γ define

$$A_\gamma = \{\pi \in S_n | \gamma \text{ is a cycle of } \pi\}.$$

The set of permutations with no i-cycles is

$$S_n \setminus \bigcup_{\gamma} A_\gamma.$$

By PIE we have

$$|S_n \setminus \bigcup_{\gamma} A_\gamma| = \sum_{J \subset i\text{-cycles}} (-1)^{|J|} |A_J|$$

where

$$A_J = \bigcap_{\gamma \in J} A_\gamma.$$

Because A_J is empty unless J consists of disjoint cycles, we only need to sum over the subsets J where $|J| \leq n/i$, equivalently for $|J| \leq \lfloor n/i \rfloor$. For a subset J consisting of disjoint cycles and $|J| = j$ we have

$$|A_J| = (n - ji)!$$

To count the number of such J we make an ordered selection of ji elements from n elements. This can be done in $n^{\underline{ji}}$ ways. The first i form a cycle but since any of its i elements can be first we divide by a factor of i. The second group of i elements form a cycle; again we divide by i. We do this for the j groups, which gives a factor of i^j in the denominator. But since the order of the j disjoint cycles is unimportant, we also divide by $j!$. Therefore, the number of subsets of size j is

$$\frac{n^{\underline{ji}}}{j! i^j}$$

and so

$$\sum_J (-1)^{|J|}|A_J| = \sum_{j=0}^{\lfloor n/i \rfloor} (-1)^j \frac{n^{\underline{ji}}(n-ji)!}{j!i^j} \tag{2}$$

$$= n! \sum_{j=0}^{\lfloor n/i \rfloor} (-1)^j \frac{1}{j!i^j} \tag{3}$$

The last step follows from $n^{\underline{ji}}(n-ji)! = n!$. □

Now we are in a position to find the probability of exactly k i-cycles in a random permutation.

Theorem 4 *In S_n the number of permutations with exactly k i-cycles is*

$$\frac{n!}{k!i^k} \sum_{j=0}^{\lfloor n/i-k \rfloor} (-1)^j \frac{1}{j!i^j}.$$

Proof All such permutations are constructed by selecting k i-cycles and then selecting a permutation of the remaining elements with no i-cycles. The first selection of the cycles can be done in $n^{\underline{ki}}/(k!i^k)$ ways, and the second selection can be done in $d_i(n-ki)$ ways. From the previous theorem we know the value for $d_i(n-ki)$. Multiplying the two together gives

$$\frac{n^{\underline{ki}}}{k!i^k} d_i(n-ki) = \frac{n^{\underline{ki}}}{k!i^k}(n-ki)! \sum_{j=0}^{\lfloor (n-ki)/i \rfloor} (-1)^j \frac{1}{j!i^j} \tag{4}$$

$$= \frac{n!}{k!i^k} \sum_{j=0}^{\lfloor n/i-k \rfloor} (-1)^j \frac{1}{j!i^j} \tag{5}$$

□

Corollary 5 *In a random permutation the probability of exactly k i-cycles is*

$$P(C_i = k) = \frac{1}{k!i^k} \sum_{j=0}^{\lfloor n/i-k \rfloor} (-1)^j \frac{1}{j!i^j}.$$

Proof Just divide the number from the theorem by $n!$. □

Corollary 6

$$\sum_{k=0}^{\lfloor n/i \rfloor} \frac{1}{k!i^k} \sum_{j=0}^{\lfloor n/i-k \rfloor} (-1)^j \frac{1}{j!i^j} = 1$$

$$\sum_{k=1}^{\lfloor n/i \rfloor} \frac{1}{(k-1)!i^k} \sum_{j=0}^{\lfloor n/i-k \rfloor} (-1)^j \frac{1}{j!i^j} = 1/i$$

Proof The first identity is that the probabilities sum to 1. The second is that the expected number of i-cycles is $1/i$. □

Corollary 7 *As $n \to \infty$ the distribution of the number of i-cycles approaches a Poisson distribution with parameter $\lambda = 1/i$.*

Proof We have

$$\begin{aligned}
\lim_{n\to\infty} P(C_i = k) &= \lim_{n\to\infty} \frac{1}{k!i^k} \sum_{j=0}^{\lfloor n/i-k \rfloor} (-1)^j \frac{1}{j!i^j} && (6)\\
&= \frac{1}{k!i^k} \sum_{j=0}^{\infty} (-1)^j \frac{1}{j!i^j} && (7)\\
&= \frac{1}{k!i^k} e^{-1/i} && (8)\\
&= e^{-1/i} \frac{(1/i)^k}{k!}. && (9)
\end{aligned}$$

□

Consider the random vector $C = (C_1, C_2, \ldots, C_n)$ describing the complete cycle structure of a permutation, which is the same as describing its conjugacy class in S_n. Each of the components in asymptotically Poisson, but C has exactly the same distribution as the random vector $Z = (Z_1, Z_2, \ldots, Z_n)$ conditioned on the weighted sum $\sum_{i=0}^{n} iZ_i = n$, where Z_i is Poisson with mean $1/i$. This is *not* an asymptotic result, but one that holds for each n. These results can be found in [1].

Theorem 8 *(Cauchy's Formula) If $a = (a_1, a_2, \ldots, a_n) \in \mathbf{N}^n$ and $\sum a_i = n$, then the number of permutations in S_n with a_i i-cycles is*

$$\frac{n!}{\prod a_i! i^{a_i}}.$$

Proof The cycle structure specifies a form

$$(x) \cdots (x)(xx) \cdots (xx) \cdots$$

with a_1 1-cycles, etc. There are $n!$ ways to place the elements $1, 2, \ldots, n$ but each $\pi \in S_n$ with this cycle structure will occur $\prod a_i! i^{a_i}$ times. □

Corollary 9 *(The Law of Cycle Structures) If $a = (a_1, a_2, \ldots, a_n) \in \mathbf{N}^n$ and $\sum a_i = n$, then*

$$P(C = a) = \prod_{i=1}^{n} \frac{1}{a_i!} (1/i)^{a_i}.$$

Proof This follows immediately from Cauchy's Formula. □

Corollary 10

$$\sum_{a\in\mathbf{N}^n,\ \sum ia_i=n}\prod_i \frac{1}{a_i!}(1/i)^{a_i}=1.$$

Proof Sum the probabilities to get 1. □

Theorem 11 *Suppose Z_i is Poisson with parameter $1/i$ and that Z_i, $1\le i\le n$, are independent. Define $T_n=\sum iZ_i$. If $a=(a_1,a_2,\dots,a_n)\in\mathbf{N}^n$, then*

$$P(Z=a|T_n=n)=\prod_{i=1}^{n}\frac{1}{a_i!}(1/i)^{a_i}.$$

Proof The definition of conditional probability gives

$$P(Z=a|T_n=n)=\frac{P(Z=a)}{P(T_n=n)} \tag{10}$$

The independence of the Z_i implies that

$$\begin{aligned} P(Z=a) &= \prod_{i=1}^{n}P(Z_i=a_i)\\ &= \prod_{i=1}^{n}e^{-1/i}(1/i)^{a_i}\frac{1}{a_i!}. \end{aligned}$$

The denominator $P(T_n=n)$ is the sum over all $a\in\mathbf{N}^n$, with $\sum ia_i=n$, of the probability that $Z=a$. Since the Z_i are independent,

$$\begin{aligned} P(T_n=n) &= \sum_{a\in\mathbf{N}^n,\ \sum ia_i=n}\prod_i e^{-1/i}(1/i)^{a_i}\frac{1}{a_i!}\\ &= \prod_i e^{-1/i} \end{aligned}$$

The second line follows using the corollary above on the sum of the probabilities. Then

$$P(Z=a|T_n=n)=\prod_i(1/i)^{a_i}\frac{1}{a_i!}. \tag{11}$$

□

Now we turn to some results from the point of view of an element of $\{1,2,\dots,n\}$ when a random permutation acts on it. We may as well assume the element is 1.

Theorem 12 *The probability that the cycle containing 1 has length k is $1/n$. That is, the length of the cycle containing 1 is equiprobably distributed on the integers from 1 to n.*

Proof We count the permutations that have 1 contained in a cycle of length k. There are $(n-1)^{\underline{k-1}}$ cycles of length k containing 1, since we simply have to choose $k-1$ distinct elements from $n-1$ possibilites to fill up the cycle. There are $(n-k)!$ remaining possibilities for the rest of the permutation. The product of these two numbers is $(n-1)!$. Hence, the probability we seek is $(n-1)!/n! = 1/n$. □

Corollary 13 *The expected length of the cycle containing 1 is* $(n+1)/2$.

Proof The cycle lengths range from 1 to n and each is equally probable. □

Look at all the cycles of all the permutations in S_n. We know that there are $n!H_n$ cycles from Corollary 1. The total length of all these cycles is $n!n$ because the subtotal for each permutation is n. The average length of these cycles is

$$\frac{n!n}{n!H_n} = \frac{n}{H_n} \tag{12}$$

which is approximately $n/\log n$. From this point of view the average cycle length is much smaller than from the element's point of view. (This is reminiscent of the paradox of average class size. The student's average class size can be much larger than the college's average class size. Average family size is another example. The paradox is explained because large classes are counted once for each student in the class from the student point of view but only once from the college point of view.)

Exercise 14 Show the probability that 1 and 2 are in the same cycle is 1/2. And show that the probability that 1,2, and 3 are in the same cycle is 1/3.

Proposition 15 *Assume* $m \le n$. *The probability that* $1, 2, \ldots, m$ *are in the same cycle is* $1/m$.

Proof First we will count the permutations that have $1, 2, \ldots, m$ in the same k-cycle for a fixed value of k. Let's put 1 as the first element of the k-cycle. Then there are $(k-1)^{\underline{m-1}}$ choices for placing $2, \ldots, m$ in the cycle. The remaining elements can be placed in $(n-m)!$ independent ways. The product $(n-m)!(k-1)^{\underline{m-1}}$ is the number of permutations with $1, 2, \ldots, m$ in the same k-cycle. Now we sum over k to get the number of permutations with $1, 2, \ldots, m$ in the same cycle. Then we divide by $n!$ to get the probability that they are in the same cycle. Let $P_{n,m}$ denote this probability. Hence,

$$P_{n,m} = \frac{(n-m)!}{n!} \sum_{k=m}^{n} (k-1)^{\underline{m-1}} = \frac{1}{n^{\underline{m}}} \sum_{k=m}^{n} (k-1)^{\underline{m-1}}. \tag{13}$$

Some routine algebra shows that

$$P_{n+1,m} = \frac{n+1-m}{n+1} P_{n,m} + \frac{1}{n+1}. \tag{14}$$

We know that $P_{m,m} = 1/m$ and so we only have to check that $1/m$ is a solution to the recurrence equation. This is easy to do. □

2 Random Maps

We consider all maps from one finite set to another, paying particular attention to the maps from one set to itself. Let $\underline{n}$ denote the finite set $\{1, 2, \ldots, n\}$. Let $M_{n,m}$ denote the set of all maps from $\underline{n}$ to $\underline{m}$ with the equiprobable measure and let M_n denote the self-maps from $\underline{n}$ to itself. (The self-maps are also called *endo-functions* by Cameron.) There are m^n maps in $M_{n,m}$.

First consider the size of the image as an integer valued random variable on $M_{n,m}$.

Theorem 16 *The probability that the image size is k is*

$$\frac{k!}{m^n}\binom{m}{k}\left\{ n \atop k \right\}.$$

Proof There are $\binom{m}{k}$ possible image sets. There are $\left\{ n \atop k \right\}$ possible partitions for the inverse images of the k image points. There are $k!$ ways to assign map the inverse image sets to the image. □

Corollary 17

$$\sum_{k=1}^{m}\frac{k!}{m^n}\binom{m}{k}\left\{ n \atop k \right\} = 1$$

Theorem 18 *The expected image size is*

$$m(1-\left(\frac{m-1}{m}\right)^n).$$

Proof Let Y_j be the random variable with value 1 if j is in the image and 0 otherwise. Then $Y = Y_1 + \cdots + Y_m$ is the image size and $E(Y) = E(Y_1) + \cdots + E(Y_m)$. The expected value of Y_j is the probability that j is in the image, which is $1 - P(j \notin \text{ image})$, and the probability that j is not in the image is $\left(\frac{m-1}{m}\right)^n$. □

Corollary 19

$$\sum_{k=1}^{m} k\frac{k!}{m^n}\binom{m}{k}\left\{ n \atop k \right\} = m(1-\left(\frac{m-1}{m}\right)^n)$$

Exercise 20 Determine the variance for the image size.

Solution (Thanks to Robert Sawyer, via e-mail, for pointing out that an earlier solution was incorrect because it assumed that the Y_i were independent and for supplying the correct solution below.)

For independent (more generally, uncorrelated) random variables the variance of the sum is the sum of the variances, but in general the variance of a sum is the sum of all the the pairwise covariances. Recall, that the covariance of random variables X_1 and X_2 is $\text{cov}\,(X_1, X_2) = E((X_1 - E(X_1))E(X_2 - E(X_2))$, which is also $E(X_1X_2) - E(X_1)E(X_2)$. Then one can show that

$$\begin{aligned}
\text{var}\,(Y) &= \text{var}\,(\sum_i Y_i) \\
&= \sum_i \text{var}\,(Y_i) + \sum_{i \neq j} \text{cov}\,(Y_i, Y_j) \\
&= m\,\text{var}\,(Y_1) + m(m-1)\,\text{cov}\,(Y_1, Y_2)
\end{aligned}$$

The last line comes from the fact that the variances of the Y_i are the same as are the covariances of the Y_i and Y_j for $i \neq j$. Now

$$\text{var}\,(Y_1) = E(Y_i^2) - E(Y_i)^2 = \left(\frac{m-1}{m}\right)^n - \left(\frac{m-1}{m}\right)^{2n}$$

but for $\text{cov}\,(Y_1, Y_2)$ a little trick makes it easier. Use the fact that

$$\text{cov}\,(Y_1, Y_2) = \text{cov}\,((1 - Y_1), (1 - Y_2))$$

(proof left for the reader). Then $E((1 - Y_1)(1 - Y_2)$ is the probability that $Y_1 = 0$ and $Y_2 = 0$, which is the probability that both 1 and 2 are not in the image, and this probability is $\left(\frac{m-2}{m}\right)^n$. Thus,

$$\text{cov}\,(Y_1, Y_2) = \left(\frac{m-2}{m}\right)^n - \left(\frac{m-1}{m}\right)^{2n}$$

Putting these results together we get the variance of the image size

$$\text{var}\,(Y) = m\left\{\left(\frac{m-1}{m}\right)^n - \left(\frac{m-1}{m}\right)^{2n}\right\} + m(m-1)\left\{\left(\frac{m-2}{m}\right)^n - \left(\frac{m-1}{m}\right)^{2n}\right\}$$

Exercise 21 Determine the limit of $1 - \left(\frac{m-1}{m}\right)^n$ as $m, n \to \infty$ with $m/n = r$ fixed. This will be the asymptotic *proportional image size* of a random map.

Solution

$$\begin{aligned}
\lim \left(\frac{m-1}{m}\right)^n &= \lim \left(\frac{m-1}{m}\right)^{m/r} \\
&= \left(\lim \left(\frac{m-1}{m}\right)^m\right)^{1/r} \\
&= e^{-1/r}.
\end{aligned}$$

The asymptotic proportional image size is $1 - e^{-1/r}$. □

Consider the random variable Y/m, which is the proportional image size. The variance of Y/m is $m^{-2}\text{var}\,(Y)$. In the expression for var (Y) in Exercise 21, the second term coming from the covariances is negative. Therefore

$$\text{var}\,(Y) < m\left\{\left(\frac{m-1}{m}\right)^n - \left(\frac{m-1}{m}\right)^{2n}\right\}$$

and so $m^{-2}\text{var}\,(Y) \to 0$ as m and n go to infinity with a fixed ratio. Thus the probability distribution of the proportional image size becomes more and more concentrated at $1-e^{-1/r}$.

Example 22 The birthday paradox involves a random map from n people to $m = 365$ birthdays. (Disregard February 29 and assume each birthday is equally likely.) The probability that none of the people have the same birthday is the probability that the image size is n, which is

$$\frac{n!}{m^n}\binom{m}{n}\left\{ n \atop n \right\} = \frac{m^{\underline{n}}}{m^n}. \tag{15}$$

As is well-known, this probability is less than 1/2 when $n \geq 23$.

Question 23 There are m coupons that a collector is seeking to acquire. He buys them one at a time, sight unseen. How many purchases are expected before he has them all?

Exercise 24 What is the expected number of purchases for the coupon collector when $m = 2$?

Exercise 25 What is the probability that the coupon collector has all m coupons after n purchases? For small values of m, say between 2 and 10, determine the smallest n so that this probability is greater than 1/2 or greater than 0.9.

Define the random variable F_i on $M_{n,m}$ to be the size of the inverse image of i (the *fiber over* i). Let $F = (F_1, \ldots, F_m)$ be the random vector of all fiber sizes. Then $F_1 + \cdots + F_m = n$.

Proposition 26 *For* $0 \leq k \leq n$,

$$P(F_i = k) = \frac{\binom{n}{k}(m-1)^{n-k}}{m^n}.$$

Proof Immediate. □

Theorem 27 *Let* $s = (s_1, \ldots, s_m) \in \mathbf{N}^m$ *such that* $\sum s_i = m$. *Then*

$$P(F = s) = \frac{\binom{n}{s_1, s_2, \ldots, s_m}}{m^n}.$$

Proof The multinomial coefficient in the numerator is the number of ways to select s_i elements to comprise the fiber over i. □

Example 28 The California lottery and other state lotteries can have multiple winners because tickets can be sold with the same numbers chosen. If we assume that each possible choice of numbers is equally likely (an assumption that is not borne out in practice because people prefer certain combinations over others), then we have a random map from n tickets to $m = \binom{51}{6} = 18,009,460$ choices of the numbers. Suppose that 20 million tickets are sold and that you have a winning ticket. What is the probability that you are the only winner? What is the expected number of winning tickets? Can you answer these questions for general n and m? (Answers: the probability you are the only winner is $((m-1)/m)^{n-1}$. If you are a winner the expected number of additional winners is $(n-1)/m$. The expected number of winners is n/m.)

Let the random variable G_i be the number of fibers of size i. Then $\sum iG_i = n$ and the probability law for the vector random variable $G = (G_1, \ldots, G_n)$ is analagous to the law for the number of cycles for random permutations and the derivation runs along the same lines. First we need the number of maps with no fibers of size i.

Theorem 29 *The number of maps in $M_{n,m}$ having no fibers of size i is*

$$\sum_{j=0}^{\lfloor n/i \rfloor} (-1)^j \binom{n}{\underbrace{i, i, \ldots, i}_{j}} \binom{m}{j} m^{n-ji}.$$

Proof We use the Principle of Inclusion-Exclusion. Let γ be a subset of size i in $\mathbf{N}_n$, that is, a possible fiber of size i. Let $A_\gamma = \{f \in M_{n,m} | \gamma \text{ is a fiber of } f\}$. The maps we seek to count are the complement of the union of the A_γ. By PIE we have

$$|M_{n,m} \setminus \bigcup_\gamma A_\gamma| = \sum_{J \subset\ i\text{-sets}} (-1)^{|J|} |A_J|$$

where

$$A_J = \bigcap_{\gamma \in J} A_\gamma.$$

Note that A_J is empty unless the elements of J are disjoint i-sets. Suppose that $J = \{\gamma_1, \ldots, \gamma_j\}$ where the γ_k are disjoint. Then

$$|A_J| = \binom{m}{j} j! m^{n-ji}.$$

The number of such J whose elements are j disjoint i-sets is

$$\frac{1}{j!} \binom{n}{\underbrace{i, i, \ldots, i}_{j}}$$

Thus,

$$\begin{aligned}\sum_J (-1)^{|J|}|A_J| &= \sum_{j=0}^{\lfloor n/i\rfloor} (-1)^j \binom{m}{j} j! m^{n-ji} & (16)\\ &= \sum_{j=0}^{\lfloor n/i\rfloor} (-1)^j \binom{n}{\underbrace{i,i,\ldots,i}_{j}} \binom{m}{j} m^{n-ji}. & (17)\end{aligned}$$

Theorem 30 *The number of maps in $M_{n,m}$ with k fibers of size i is*

$$\binom{m}{k}\binom{n}{\underbrace{i,i,\ldots,i}_{k}} k! \sum_{j=0}^{\lfloor n/i-k\rfloor} (-1)^j \binom{n-ki}{\underbrace{i,i,\ldots,i}_{j}} \binom{m}{j} m^{n-ki-ji}.$$

Proof Pick k points in $\mathbf{N}_m$. Pick k i-sets in $\mathbf{N}_n$ to be the fibers of the points. Choose an assignment of the i-sets to the points. The rest is equivalent to a map from $\mathbf{N}_{n-ki}$ to $\mathbf{N}_m$ having no fibers of size i. □

For self-maps of a set to itself there is a richer structure because of the possibility of iterating a map. This gives fixed points and periodic points and lots of probabilistic questions about them. Consider the number of fixed points of a random map in M_n. This is random variable which is the sum of indicator random variables, one for each i, whose value is 1 if i is a fixed point and 0 otherwise. The following is easy to prove.

Theorem 31 *The expected number of fixed points is 1.*
The probability that the number of fixed points is k is

$$n^{-n}\binom{n}{k}(n-1)^{n-k}.$$

Corollary 32

$$\sum_{k=0}^{n} \binom{n}{k}(n-1)^{n-k} = n^n$$

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

Proof The first is equivalent to the sum of the probabilities of the number of fixed points being 1. The second is equivalent to the expected value being 1. □

Exercise 33 Find the variance for the number of fixed points. Answer: $(n-1)/n$.

Associated to a map ϕ in M_n is a directed graph with vertex set $\underline{n}$ and an edge going from i to j if $\phi(i) = j$. This graph breaks up into connected components. Each component consists of a cycle with trees attached. When ϕ is a permutation the components are just the cycles and there are no attached trees, so we can regard the components as a natural generalization of the cycles of a permutation. As ϕ is repeatedly iterated (composed with itself) the image of ϕ^m eventually settles down and does not change and this image is the union of the cycles in the associated graph. Let us call this set the *core* of ϕ. The restriction of ϕ to its core is a permutation on the core.

Theorem 34 *The expected size of the core is*

$$\sum_{k=1}^{n} \frac{n^{\underline{k}}}{n^k}.$$

Proof Write the core size as a sum of indicator random variables $\sum_i^n X_i$ with $X_i = 1$ if i is in the core of the map and 0 otherwise. Then, $E(\sum_i^n X_i) = \sum_i^n E(X_i)$, but the X_i are identically distributed. Therefore, $E(\sum_i^n X_i) = nE(X_1)$.

The probability that 1 is in a k-cycle is

$$\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right)\cdots\left(\frac{n-k+1}{n}\right)\left(\frac{1}{n}\right)$$

which is equal to $n^{\underline{k}}/n^{k+1}$. Summing this over k from 1 to n, we get

$$E(X_1) = \sum_{k=1}^{n} \frac{n^{\underline{k}}}{n^{k+1}}.$$

The expected core size is n times this, completing the proof. □

Question 35 What is the asymptotic behavior of the expected core size as n goes to infinity?

Theorem 36 *The expected core size is asymptotic to* $\sqrt{\frac{\pi n}{2}}$.

Proof From the previous theorem we have the expected core size. Then,

$$\sum_{k=1}^{n} \frac{n^{\underline{k}}}{n^k} = \sum_{k=1}^{n} \frac{n!}{(n-k)!n^k} \tag{18}$$

$$= n! \sum_{k=0}^{n-1} \frac{1}{k!n^{n-k}} \tag{19}$$

$$= \frac{n!}{n^n} \sum_{k=0}^{n-1} \frac{n^k}{k!} \tag{20}$$

For the sum we notice that $e^{-n}\sum_{k=0}^{n-1} n^k/k!$ is probability that a Poisson random variable with parameter n has value less than n. However, such a random variable has the same distribution as a sum of n independent Poisson random variables with parameter 1. The Central Limit Theorem shows that the distribution of the average of a sum of n independent Poisson random variables with parameter 1 approaches a normal distribution with mean 1. Our random variable is just the sum or n times the average and so the probability that it is less than n has a limit of 1/2. Therefore, $\sum_{k=0}^{n-1} n^k/k!$ is asymptotic to $e^n/2$. By Stirling's Formula

$$n! \sim \sqrt{2\pi n}(\frac{n}{e})^n$$

and combining these asymptotics gives us the result. □

Let C_k be the random variable that counts the number of k-cycles of a random map. Thus, C_k counts the number of components consisting of a k-cycle with zero or more attached trees.

Theorem 37 *The expected number of k-cycles is*

$$\frac{1}{k}\frac{n^{\underline{k}}}{n^k}.$$

As $n \to \infty$ the expected number of k-cycles goes to $1/k$.

Proof For γ a k-cycle let $\mathbf{1}_\gamma$ be the indicator random variable that takes on the value 1 if γ is a cycle of the random map and 0 otherwise. Then $C_k = \sum_\gamma \mathbf{1}_\gamma$ and $E(C_k) = \sum_\gamma E(\mathbf{1}_\gamma)$. The number of maps that have γ as a k-cycle is n^{n-k} since each of the elements not in γ can be mapped anywhere. Thus, $E(\mathbf{1}_\gamma) = n^{n-k}/n^n = n^{-k}$, while the number of k-cycles is $n^{\underline{k}}/k$. The product of these is the expected value of C_k. The limit as n goes to infinity is straightforward keeping in mind that k is fixed. □

Question 38 Does the distribution of C_k become Poisson with mean $1/k$ as $n \to \infty$? I suspect that is the case but have not worked it out except for $k = 1$.

Theorem 39 *As $n \to \infty$ the distribution of C_1 approaches the distribution of a Poisson random variable with mean 1.*

Proof The number of maps with j fixed points is $\binom{n}{j}(n-1)^{n-j}$ since we choose a j-set of fixed points and then map each of the remaining points to anything but themselves. Dividing by n^n we get

$$P(C_1 = j) = \binom{n}{j}(n-1)^{n-j}n^{-n} \tag{21}$$

$$= \frac{n!}{j!(n-j)!}(n-1)^{-j}\left(\frac{n-1}{n}\right)^n \tag{22}$$

But

$$\lim_{n\to\infty} \frac{n!}{(n-j)!}(n-1)^{-j} = 1 \text{ and } \lim_{n\to\infty} \left(\frac{n-1}{n}\right)^n = \frac{1}{e}$$

and so

$$\lim_{n\to\infty} P(C_1 = j) = \frac{1}{j!}\frac{1}{e}. \tag{23}$$

□

Theorem 40 *The expected number of components is*

$$\frac{n!}{n^n}\sum_{k=0}^{n-1}\frac{1}{n-k}\frac{n^{n-k}}{k!}.$$

Proof The number of components (which is the same as the number of cycles) is the random variable $C = \sum C_k$, and therefore

$$\begin{aligned} E(C) &= \sum_{k=1}^{n}\frac{1}{k}\frac{n^{\underline{k}}}{n^k} &(24)\\ &= \frac{n!}{n^n}\sum_{k=1}^{n}\frac{1}{k}\frac{n^{n-k}}{(n-k)!} &(25)\\ &= \frac{n!}{n^n}\sum_{k=0}^{n-1}\frac{1}{n-k}\frac{n^{n-k}}{k!} &(26)\end{aligned}$$

where the last step is re-indexing with k in place of $n-k$. □

Question 41 What is the asymptotic nature of $E(C)$ as $n \to \infty$?

Notice that the expression is quite similar to that for the core size, but there is an extra wrinkle that causes difficulty. One may proceed heuristically to conjecture the first order asymptotics as follows. The expected core size is asymptotic to $\sqrt{\pi n/2}$ and we know that for a random permutation on an n-set the expected number of cycles is asymptotic to $\log n$. So we proceed under the assumption that a random map is like a random permutation on its core, and so it should have about $\log\sqrt{\pi n/2}$ cycles. But $\log\sqrt{\pi n/2} = \frac{1}{2}(\log(\pi n) - \log 2)$, which is asymptotic to $\frac{1}{2}\log(\pi n) = \frac{1}{2}(\log \pi + \log n)$, which is asymptotic to $(\log n)/2$. (n.b. This is in conflict with the value mentioned without proof in [3] which is $\log n$.)

Here is some numerical evidence for the conjecture.

n	$E(C)$	$\log n$	$E(C)/\log n$
500	3.761	6.215	0.629
1000	4.102	6.908	0.594
10000	5.245	9.210	0.569

Question 42 What is the probability distribution for the size of the components?

Question 43 What is the expected size of the components?

We will focus on a particular element, say 1, and consider what happens to it with the selection of a random map. The orbit of i under the map f is the set of iterates of i, namely $\{i, f(i), f(f(i)), \ldots\}$.

Theorem 44 *The probability that the orbit of* 1 *has size* k *is*

$$\frac{k}{n^k}(n-1)^{\underline{k-1}}.$$

Proof The orbit of 1 must be a set $\{1, x_1, x_2, \ldots, x_{k-1}\}$ of distinct elements and then x_k must be one of the k elements in the orbit set. Thus, there are $n-1$ choices for x_1, $n-2$ choices for x_2, etc. and $n-k-1$ choices for x_{k-1}. Finally, there are k choices for x_k. The remaining $n-k$ elements can be mapped to any of the n elements. The number of maps having 1 in an orbit of size k is then $(n-1)^{\underline{k-1}} k n^{n-k}$. Dividing this by n^n gives the result. □

Theorem 45 *The probability that the orbit of* 1 *has size* k *and the unique cycle in the orbit has size* j *is*

$$\frac{(n-1)^{\underline{k-1}}}{n^k}.$$

Note that this is independent of j*.*

Proof Again we count the maps with this property. The orbit of 1 must be $\{1, x_2, x_3, \ldots, x_{k-1}\}$ and $x_k = x_{k-j}$. The remaining $n-k$ elements can be mapped arbitrarily. There are $(n-1)(n-2)\cdots(n-k+1)n^{n-k}$ such maps. Dividing by n^n gives us the probability. □

Corollary 46 *The probability that the cycle in the orbit of 1 has size* j *is*

$$\sum_{k=j}^{n} \frac{(n-1)^{\underline{k-1}}}{n^k} = \sum_{k=j}^{n} \frac{(n-1)!}{(n-k)!} \frac{1}{n^k}.$$

Proof Sum over k, realizing that the orbit size must be at least as large as the cycle size. □

Corollary 47 *The probability that 1 is* l *steps from the cycle is*

$$\sum_{j=1}^{n-l} \frac{(n-1)!}{(n-j-l)! n^{j+l}}.$$

Proof We sum over j from 1 to $n-l$ the probability that the orbit of 1 has size $j+l$ and the cycle has size l. □

Corollary 48 *The expected number of steps before 1 (or any element) reaches the cycle in its component is*

$$\sum_{l=1}^{n-1} l \sum_{j=1}^{n-l} \frac{(n-1)!}{(n-j-l)!n^{j+l}}.$$

Proof Obvious from the previous corollary. □

Question 49 What is asymptotic expected number of steps to the cycle as $n \to \infty$?

The questions about components are interesting for random graphs, too. See, for example, Appendix A of [3], which refers to [2].

References

[1] R. Arratia, A. D. Barbour, and S. Tavaré, *Random combinatorial structures and prime factorizations*, Notices Amer. Math. Soc. **44**:903-910, 1997.

[2] B. Bollobas, *Random Graphs*, Academic Press, New York, 1985.

[3] T. Hogg and B. A. Huberman, *Artificial intelligence and large scale computation*, Phys. Reports **156**:227-310, 1987.