Polya's Theory of Counting

Example 1 A disc lies in a plane. Its centre is fixed but it is free to rotate. It has been divided into n sectors of angle $2\pi/n$. Each sector is to be coloured Red or Blue. How many different colourings are there? One could argue for 2^n. On the other hand, what if we only distinguish colourings which cannot be obtained from one another by a rotation. For example if $n = 4$ and the sectors are numbered 0,1,2,3 in clockwise order around the disc, then there are only 6 ways of colouring the disc – 4R, 4B, 3R1B, 1R3B, RRBB and RBRB.

Example 2 Now consider an $n \times n$ "chessboard" where $n \geq 2$. Here we colour the squares Red and Blue and two colourings are different only if one cannot be obtained from another by a rotation or a reflection. For $n = 2$ there are 6 colourings.

The general scenario that we consider is as follows: We have a set X which will stand for the set of colourings when transformations are not allowed. (In example 1, $|X| = 2^n$ and in example 2, $|X| = 2^{n^2}$). In addition there is a set G of permutations of X. This set will have a group structure: Given two members $g_1, g_2 \in G$ we can define their composition $g_1 \circ g_2$ by $g_1 \circ g_2(x) = g_1(g_2(x))$ for $x \in X$. We require that G is *closed* under composiiton i.e. $g_1 \circ g_2 \in G$ if $g_1, g_2 \in G$. We also have the following:

A1 The *identity* permutation $1_X \in G$.

A2 $(g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3)$ (Composition is associative).

A3 The inverse permutation $g^{-1} \in G$ for every $g \in G$.

(A set G with a binary relation $\circ$ which satisfies **A1,A2,A3** is called a **Group**).

In example 1 $D = \{0, 1, 2, \ldots, n-1\}$, $X = 2^D$ and the group is $G_1 = \{e_0, e_1, \ldots, e_{n-1}\}$ where $e_j * x = x + j \ mod \ n$ stands for rotation by $2j\pi/n$.

In example 2, $X = 2^{[n]^2}$. We number the squares 1,2,3,4 in clockwise order starting at the upper left and represent X as a sequence from $\{r, b\}^4$ where for example rrbr means colour 1,2,4 Red and 3 Blue. $G_2 = \{e, a, b, c, p, q, r, s\}$ is in a sense independent of n. e, a, b, c represent a rotation through $0, 90, 180, 270$ degrees respectively. p, q represent reflections in the vertical and horizontal and r, s represent reflections in the diagonals 1,3 and 2,4 respectively. Now check the following table:

	rrrr	brrr	rbrr	rrbr	rrrb	bbrr	rbbr	rrbb	brrb	rbrb	brbr	bbbr	bbrb	brbb	rbbb	bbbb
e	rrrr	brrr	rbrr	rrbr	rrrb	bbrr	rbbr	rrbb	brrb	rbrb	brbr	bbbr	bbrb	brbb	rbbb	bbbb
a	rrrr	rbrr	rrbr	rrrb	brrr	rbbr	rrbb	brrb	bbrr	brbr	rbrb	rbbb	bbbr	bbrb	brbb	bbbb
b	rrrr	rrbr	rrrb	brrr	rbrr	rrbb	brrb	bbrr	rbbr	rbrb	brbr	brbb	rbbb	bbbr	bbrb	bbbb
c	rrrr	rrrb	brrr	rbrr	rrbr	brrb	bbrr	rbbr	rrbb	brbr	rbrb	bbrb	brbb	rbbb	bbbr	bbbb
p	rrrr	rbrr	brrr	rrrb	rrbr	bbrr	brrb	rrbb	rbbr	brbr	rbrb	bbrb	bbbr	brbb	brbb	bbbb
q	rrrr	rrrb	rrbr	rbrr	brrr	rrbb	rbbr	bbrr	brrb	brbr	rbrb	rbbb	brbb	bbrb	bbbr	bbbb
r	rrrr	brrr	rrrb	rrbr	rbrr	brrb	rrbb	rbbr	bbrr	rbrb	brbr	brbb	bbrb	bbbr	rbbb	bbbb
s	rrrr	rrbr	rbrr	brrr	rrrb	rbbr	bbrr	brrb	rrbb	rbrb	brbr	bbbr	rbbb	brbb	bbrb	bbbb

From now on we will write $g * x$ in place of $g(x)$.

Orbits: If $x \in X$ then its orbit $O_x = \{y \in X : \exists g \in G \text{ such that } g * x = y\}$.

Lemma 1. *The orbits partition X.*

Proof $x = 1_X * x$ and so $x \in O_x$ and so $X = \bigcup_{x \in X} O_x$.

Suppose now that $O_x \cap O_y \neq \emptyset$ i.e. $\exists g_1, g_2$ such that $g_1 * x = g_2 * y$. But then for any $g \in G$ we have

$$g * x = (g \circ (g_1^{-1} \circ g_2)) * y \in O_y$$

and so $O_x \subseteq O_y$. Similarly $O_y \subseteq O_x$. Thus $O_x = O_y$ whenever $O_x \cap O_y \neq \emptyset$. $\square$

The two problems we started with are of the following form: Given a set X and a group of permutations *acting* on X, compute the number of orbits i.e. distinct colourings.

A subset H of G is called a *sub-group* of G if it satisfies *axioms* **A1,A2,A3** (with G replaced by H).

The *stabilizer* S_x of the element x is $\{g : \ g * x = x\}$. It is a sub-group of G.

Lemma 2.
If $x \in X$ then $|O_x|\,|S_x| = |G|$.

Proof Fix $x \in X$ and define an equivalence relation $\sim$ on G by

$$g_1 \sim g_2 \text{ if } g_1 * x = g_2 * x.$$

Let the equivalence classes be $A_1, A_2, \ldots, A_m$. We first argue that

$$|A_i| = |S_x| \qquad i = 1, 2, \ldots, m. \tag{1}$$

Fix i and $g \in A_i$. Then

$$h \in A_i \leftrightarrow g * x = h * x. \leftrightarrow (g^{-1} \circ h) * x = x \leftrightarrow (g^{-1} \circ h) \in S_x. \leftrightarrow h \in g \circ S_x$$

where $g \circ S_x = \{g \circ \sigma : \ \sigma \in S_x\}$. Thus $|A_i| = |g \circ S_x|$. But $|g \circ S_x| = |S_x|$ since if $\sigma_1, \sigma_2 \in S_x$ and $g \circ \sigma_1 = g \circ \sigma_2$ then $g^{-1} \circ (g \circ \sigma_1) = (g^{-1} \circ g) \circ \sigma_1 = \sigma_1 = g^{-1} \circ (g \circ \sigma_2) = \sigma_2$. This proves (1).

Finally, $m = |O_x|$ since there is a distinct equivalence class for each distinct $g * x$. $\square$

Examples:

In example 1 with $n = 4$ we have

x	O_x	S_x
rrrr	$\{rrrr\}$	G
brrr	{brrr,rbrr,rrbr,rrrb}	$\{e_0\}$
rbrr	{brrr,rbrr,rrbr,rrrb}	$\{e_0\}$
rrbr	{brrr,rbrr,rrbr,rrrb}	$\{e_0\}$
rrrb	{brrr,rbrr,rrbr,rrrb}	$\{e_0\}$
bbrr	{bbrr,rbbr,rrbb,brrb}	$\{e_0\}$
rbbr	{bbrr,rbbr,rrbb,brrb}	$\{e_0\}$
rrbb	{bbrr,rbbr,rrbb,brrb}	$\{e_0\}$
brrb	{bbrr,rbbr,rrbb,brrb}	$\{e_0\}$
rbrb	{rbrb,brbr}	$\{e_0, e_2\}$
brbr	{rbrb,brbr}	$\{e_0, e_2\}$
bbbr	{bbbr,rbbb,brbb,bbrb}	$\{e_0\}$
bbrb	{bbbr,rbbb,brbb,bbrb}	$\{e_0\}$
brbb	{bbbr,rbbb,brbb,bbrb}	$\{e_0\}$
rbbb	{bbbr,rbbb,brbb,bbrb}	$\{e_0\}$
bbbb	{bbbb}	G

In example 2 we have

x	O_x	S_x
rrrr	$\{e\}$	G
brrr	{brrr,rbrr,rrbr,rrrb}	{e,r}
rbrr	{brrr,rbrr,rrbr,rrrb}	{e,s}
rrbr	{brrr,rbrr,rrbr,rrrb}	{e,r}
rrrb	{brrr,rbrr,rrbr,rrrb}	{e,s}
bbrr	{bbrr,rbbr,rrbb,brrb}	{e,p}
rbbr	{bbrr,rbbr,rrbb,brrb}	{e,q}
rrbb	{bbrr,rbbr,rrbb,brrb}	{e,p}
brrb	{bbrr,rbbr,rrbb,brrb}	{e,q}
rbrb	{rbrb,brbr}	{e,b,r,s}
brbr	{rbrb,brbr}	{e,b,r,s}
bbbr	{bbbr,rbbb,brbb,bbrb}	{e,s}
bbrb	{bbbr,rbbb,brbb,bbrb}	{e,r}
brbb	{bbbr,rbbb,brbb,bbrb}	{e,s}
rbbb	{bbbr,rbbb,brbb,bbrb}	{e,r}
bbbb	{e}	G

Let $\nu_{X,G}$ denote the number of orbits.

Theorem 1.

$$\nu_{X,G} = \frac{1}{|G|} \sum_{x \in X} |S_x|.$$

Proof

$$\begin{aligned} \nu_{X,G} &= \sum_{x \in X} \frac{1}{|O_x|} \\ &= \sum_{x \in X} \frac{|S_x|}{|G|}, \end{aligned}$$

from Lemma 2. □

Thus in example 1 we have

$$\nu_{X,G} = \frac{1}{4}(4+1+1+1+1+1+1+1+1+1+2+2+1+1+1+1+4) = 6.$$

In example 2 we have

$$\nu_{X,G} = \frac{1}{8}(8+2+2+2+2+2+2+2+2+4+4+2+2+2+2+8) = 6.$$

Theorem 1 is hard to use if $|X|$ is large, even if $|G|$ is small. For $g \in G$ let $Fix(g) = \{x \in X : g * x = x\}$.

Theorem 2. *(Frobenius, Burnside)*

$$\nu_{X,G} = \frac{1}{|G|} \sum_{g \in G} |Fix(g)|.$$

Proof Let $A(x,g) = 1_{g*x=x}$. Then

$$\begin{aligned} \nu_{X,G} &= \frac{1}{|G|} \sum_{x \in X} |S_x| \\ &= \frac{1}{|G|} \sum_{x \in X} \sum_{g \in G} A(x,g) \\ &= \frac{1}{|G|} \sum_{g \in G} \sum_{x \in X} A(x,g) \\ &= \frac{1}{|G|} \sum_{g \in G} |Fix(g)|. \end{aligned}$$

□

Let us consider example 1 with $n = 6$. We compute

g	e_0	e_1	e_2	e_3	e_4	e_5
$\|Fix(g)\|$	64	2	4	8	4	2

Applying Theorem 2 we obtain

$$\nu_{X,G} = \frac{1}{6}(64+2+4+8+4+2) = 14.$$

Cycles of a permutation

Let $\pi : D \to D$ be a permutation of the finite set D. Consider the digraph $\Gamma_\pi = (D, A)$ where $A = \{(i, \pi(i)) : i \in D\}$. Γ_π is a collection of vertex disjoint cycles. Each $x \in D$ being on a unique cycle. Here a cycle can consist of a loop i.e. when $\pi(x) = x$.

Example: $D = [10]$.

i	1	2	3	4	5	6	7	8	9	10
$\pi(i)$	6	2	7	10	3	8	9	1	5	4

The cycles are $(1, 6, 8), (2), (3, 7, 9, 5), (4, 10)$.

In general consider the sequence $i, \pi(i), \pi^2(i), \ldots,$. Since D is finite, there exists a first pair $k < \ell$ such that $\pi^k(i) = \pi^\ell(i)$. Now we must have $k = 0$, since otherwise putting $x = \pi^{k-1}(i) \neq y = \pi^{\ell-1}(i)$ we see that $\pi(x) = \pi(y)$, contradicting the fact that π is a permutation. So i lies on the cycle $C = (i, \pi(i), \pi^2(i), \ldots, \pi^{k-1}(i), i)$. If j is not a vertex of C then $\pi(j)$ is not on C and so we can repeat the argument to show that the rest of D is partitioned into cycles.

Example 1 First consider $e_0, e_1, \ldots, e_{n-1}$ as permutations of D. The cycles of e_0 are $(1), (2), \ldots, (n)$. Now suppose that $0 < m < n$. Let $a_m = gcd(m, n)$ and $k_m = n/a_m$. The cycle C_i of e_m containing the element i is is $(i, i+m, i+2m, \ldots, i+(k_m-1)m)$ since n is a divisor $k_m m$ and not a divisor of $k'm$ for $k' < k_m$. In total, the cycles of e_m are $C_0, C_1, \ldots, C_{a_m-1}$. This is because they are disjoint and together contain n elements. (If $i + rm = i' + r'm \mod n$ then $(r - r')m + (i - i') = \ell n$. But $|i - i'| < a_m$ and so dividing by a_m we see that we must have $i = i'$.)

Next observe that if colouring x is fixed by e_m then elements on the same cycle C_i must be coloured the same. Suppose for example that the colour of $i + bm$ is different from the colour of $i + (b+1)m$, say Red versus Blue. Then in $e_m(x)$ the colour of $i + (b+1)m$ will be Red and so $e_m(x) \neq x$. Conversely, if elements on the same cycle of e_m have the same colour then in $x \in Fix(e_m)$. This property is not peculiar to this example, as we will see.

Thus in this example we see that $|Fix(e_m)| = 2^{a_m}$ and then applying Theorem 2 we see that

$$\nu_{X,G} = \frac{1}{n} \sum_{m=0}^{n-1} 2^{gcd(m,n)}.$$

Example 2 It is straightforward to check that when n is even, we have

g	e	a	b	c	p	q	r	s
$\lvert Fix(g)\rvert$	2^{n^2}	$2^{n^2/4}$	$2^{n^2/2}$	$2^{n^2/4}$	$2^{n^2/2}$	$2^{n^2/2}$	$2^{n(n+1)/2}$	$2^{n(n+1)/2}$

For example, if we divide the chessboard into 4 $n/2 \times n/2$ sub-squares, numbered 1,2,3,4 then a colouring is in $Fix(a)$ iff each of these 4 sub-squares have colourings which are rotations of the colouring in square 1.

1 The pattern inventory

We now extend the above analysis to answer questions like: How many *distinct* ways are there to colour an 8×8 chessboard with 32 white squares and 32 black squares?

The scenrio now consists of a set D (*Domain*, a set C (colours) and $X = \{x : D \to C\}$ is the set of colourings of D with the colour set C. G is now a group of permutations of D.

We see first how to extend each permutation of D to a permutation of X. Suppose that $x \in X$ and $g \in G$ then we define $g * x$ by

$$g * x(d) = x(g^{-1}(d)) \qquad \text{for all } d \in D.$$

Explanation: The colour of d is the colour of the element $g^{-1}(d)$ which is mapped to it by g.

Consider Example 1 with $n = 4$. Suppose that $g = e_1$ i.e. rotate clockwise by $\pi/2$ and $x(1) = b, x(2) = b, x(3) = r, x(4) = r$. Then for example

$$g * x(1) = x(g^{-1}(1)) = x(4) = r, \text{ as before.}$$

Now associate a **weight** w_c with each $c \in C$. If $x \in X$ then

$$W(x) = \prod_{d \in D} w_{x(d)}.$$

Thus, if in Example 1 we let $w(r) = R$ and $w(b) = B$ and take $x(1) = b, x(2) = b, x(3) = r, x(4) = r$ then we will write $W(x) = B^2R^2$.

For $S \subseteq X$ we define the **inventory** of S to be

$$W(S) = \sum_{x \in S} W(x).$$

The probelm we discuss now is to compute the **pattern inventory** $PI = W(S^*)$ where S^* contains one member of each orbit of X under G.

For example, in the case of Example 2, with $n = 2$, we gt

$$PI = R^4 + R^3B + 2R^2B^2 + RB^3 + B^4.$$

To see that the definition of PI makes sense we need to prove

Lemma 3. *If x, y are in the same orbit of X then $W(x) = W(y)$.*

Proof Suppose that $g * x = y$. Then

$$\begin{aligned} W(y) &= \prod_{d \in D} w_{y(d)} \\ &= \prod_{d \in D} w_{g*x(d)} \\ &= \prod_{d \in D} w_{x(g^{-1}(d))} && (2) \\ &= \prod_{d \in D} w_{x(d))} && (3) \\ &= W(x) \end{aligned}$$

Note, that we can go from (2) to (3) because as d runs over D, $g^{-1}(d)$ also runs over d. $\square$

Let $\Delta = |D|$. If $g \in G$ has k_i cycles of length i then we define

$$ct(g) = x_1^{k_1} x_2^{k_2} \cdots x_\Delta^{k_\Delta}.$$

The **Cycle Index Polynomial** of G, C_G is then defined to be

$$C_G(x_1, x_2, \dots, x_\Delta) = \frac{1}{|G|} \sum_{g \in G} ct(g).$$

In Example 2 with $n = 2$ we have

g	e	a	b	c	p	q	r	s
$ct(g)$	x_1^4	x_4	x_2^2	x_4	x_2^2	x_2^2	$x_1^2 x_2$	$x_1^2 x_2$

and so

$$C_G(x_1, x_2, x_3, x_4) = \frac{1}{8}(x_1^4 + 3x_2^2 + 2x_1^2 x_2 + 2x_4).$$

In Example 2 with $n = 3$ we have

g	e	a	b	c	p	q	r	s
$ct(g)$	x_1^9	$x_1 x_4^2$	$x_1 x_2^4$	$x_1 x_4^2$	$x_1^3 x_2^3$	$x_1^3 x_2^3$	$x_1^3 x_2^3$	$x_1^3 x_2^3$

and so

$$C_G(x_1, x_2, x_3, x_4) = \frac{1}{8}(x_1^9 + x_1 x_2^4 + 4x_1^3 x_2^3 + 2x_1 x_4^2).$$

Theorem 3. *(***Polya***)*

$$PI = C_G\left(\sum_{c \in C} w_c, \sum_{c \in C} w_c^2, \dots, \sum_{c \in C} w_c^\Delta\right).$$

In Example 2, we replace x_1 by $R + B$, x_2 by $R^2 + B^2$ and so on. When $n = 2$ this gives

$$\begin{aligned} PI &= \frac{1}{8}((R+B)^4 + 3(R^2+B^2) + 2(R+B)^2(R^2+B^2) + 2(R^4+B^4)) \\ &= R^4 + R^3 B + 2R^2 B^2 + RB^3 + B^4. \end{aligned}$$

Putting $R = B = 1$ gives the number of distinct colourings. Note also the formula for PI tells us that there are 2 distinct colourings using 2 reds and 2 Blues.

Proof of Polya's Theorem
Let $X = X_1 \cup X_2 \cup \dots \cup X_m$ be the equivalence clases of X under the relation

$$x \sim y \text{ iff } W(x) = W(y).$$

By Lemma 3, $g * x \sim x$ for all $x \in X, g \in G$ and so we can think of G acting on each X_i individually i.e. we use the fact that $x \in X_i$ implies $g * x \in X_i$ for all $i \in [m], g \in G$. We use the notation $g^{(i)} \in G^{(i)}$ when we restrict attention to X_i. Let m_i denote the number of orbits $\nu_{X_i, G^{(i)}}$. Then

$$\begin{aligned} PI &= \sum_{i=1}^m m_i W_i \\ &= \sum_{i=1}^m W_i \left(\frac{1}{|G|} \sum_{g \in G} |Fix(g^{(i)})|\right) \qquad \text{by Theorem 2} \\ &= \frac{1}{|G|} \sum_{g \in G} \sum_{i=1}^m |Fix(g^{(i)})| W_i \\ &= \frac{1}{|G|} \sum_{g \in G} W(Fix(g)) \end{aligned} \tag{4}$$

Note that (4) follows from $Fix(g) = \bigcup_{i=1}^{m} Fix(g^{(i)})$ since $x \in Fix(g^{(i)})$ iff $x \in W_i$ and $g * x = x$.

Suppose now that $ct(g) = x_1^{k_1} x_2^{k_2} \cdots x_\Delta^{k_\Delta}$ as above. Then we claim that

$$W(Fix(g)) = \left(\sum_{c \in C} w_c \right)^{k_1} \left(\sum_{c \in C} w_c^2 \right)^{k_2} \cdots \left(\sum_{c \in C} w_c^\Delta \right)^{k_\Delta}. \tag{5}$$

Substituting (5) into (4) yileds the theorem.

To verify (5) we use the fact that if $x \in Fix(g)$, then the elements of a cycle of g must be given the same colour. A cycle of length i will then contribute a factor $\sum_{c \in C} w_c^i$ where the term w_c^i comes from the choice of colour c for every element of the cycle. $\square$