Monte Carlo Methods Appl. Vol. 15 No. 3 (2009), pp. 241–255

Uniform generation of anonymous and neutral preference profiles for social choice rules

Ömer Eğecioğlu

Abstract. The *Impartial Anonymous and Neutral Culture* (IANC) model of social choice assumes that the names of the voters as well as the identity of the alternatives are immaterial. This models allows for comparison of structural properties social choice rules (SCRs) for large values of the parameters empirically: whether Condorcet winners exist, whether Borda and Condorcet winners are identical, whether Plurality with run-off winners are among Borda winners, for example.

We derive an exact formula for the number of equivalence classes of preference profiles (called roots) in this model. The number of terms in the formula depends only on the number of alternatives m, and not the much larger number of voters n. In IANC, the equivalence classes defining the roots do not have the same size, making their uniform generation for large values of the parameters nontrivial. We show that the Dixon–Wilf algorithm can be adapted to this problem, and describe a symbolic algebra routine that can be used for Monte-Carlo algorithms for the study of various structural properties of SCRs.

Keywords. Social choice, Condorcet, symmetric function, Dixon–Wilf algorithm.

AMS classification. 91B14, 91B12, 68R05, 05E05.

1. Introduction

There are a few basic models used for the analysis of the properties and behaviors of various social choice rules (SCRs) through probabilistic methods designed to generate voter preferences. Among these, *Impartial Culture* (IC) uses preference profiles (which show how each of n voters in an electorate ranks m alternatives as an $m \times n$ matrix) in which each is profile is equally likely. This model has been introduced in social choice literature by Guilbaud [9]. For linearly ordered m alternatives chosen by n voters, IC assumes that each voter independently selects her preference ranking according to a uniform distribution, resulting in a total number of $m!^n$ profiles.

The *Impartial Anonymous Culture* (IAC) model on the other hand, is based on the presentation of voter preferences by anonymous profiles where the names of the voters are neglected. As introduced by Fishburn and Gehrlein [6], an anonymous profile is the representative profile of an anonymous equivalence class (AEC) which is the set of preference profiles that can be generated from each other via permuting only the names of the voters. IAC assumes that each AEC is equally likely. The number of AECs for linearly ordered m alternatives by n voters is given by the binomial coefficient $\binom{n+m!-1}{m!-1}$, as this is a balls-in-boxes type of a counting problem (see Feller [5]).

The *Impartial Anonymous and Neutral Culture* (IANC) model that is studied here, treats voter preferences through a class of preference profiles where not only the names of the voters, but also the names of the alternatives are immaterial. This approach reflects two basic axioms of social choice theory: *Anonymity* and *Neutrality*. Anonymity requires voters to be treated equally whereas neutrality prohibits a SCR from having a built-in bias for or against any one or more alternatives. The equivalence classes in this model are the *anonymous and neutral equivalence classes* (ANECs). The equivalence classes as well as equivalence class representatives are also referred to as *roots*. Thus, each root represents a structurally distinct preference profile under simultaneous fulfillment of anonymity and neutrality axioms. We denote by $R(m, n)$ the number of roots for linearly ordered m alternatives by n voters.

This paper provides a formula for $R(m, n)$, which is then used to provide an algorithm that generates roots from the uniform distribution. This allows for a testbed that can be used to answer various questions about the properties of anonymous and neutral SCRs by the Monte-Carlo method. Properties such as the likelihood of the existence of a Condorcet winner, the probability that the Borda and the Condorcet winners are identical, the probability that Plurality winners are a subset of Borda winners, etc., are among questions that can be empirically answered.

We use ideas from the theory of symmetric functions to obtain a formula for $R(m, n)$ which is a sum of terms where the number of terms depends only on m and not the much larger n. There is a combinatorial explosion in the computation of $R(m, n)$ for large values of m and n, and a simple enumeration of roots is insufficient to select representatives from the uniform distribution. Furthermore, the ANECs do not all have the same size. This makes uniform generation appear somewhat problematic, but we use the Dixon–Wilf algorithm along with the formula for $R(m, n)$ to overcome this problem.

This paper is organized as follows. Section 2 outlines the basic ideas and introduces the notation we need. The first formula for the number of roots $R(m, n)$ appears in Theorem 3.1 of Section 3. By using a result from the theory of symmetric functions (Theorem 3.2), we obtain the simpler expression for $R(m, n)$ given in Theorem 3.3. This immediately yields a number of explicit formulas which we derive for small values of m in Section 4. In Section 5 we outline the Dixon–Wilf algorithm, and indicate its use to generate roots from the uniform distribution. A symbolic algebra routine built on this theory is then described, and sample Mathematica runs are given.

The basic ideas of discrete mathematics, group theory and group actions to the extent used here can be found in Feller [5], Wielandt [13], and Kerber [10]. Ideas related to the symmetric functions can be found in MacDonald [11]. The main reference on SCRs is Moulin [12]. IC and IAC models are presented in more detail in Berg and Lepelley [1] and Gehrlein [7].

2. Preliminaries

In this section, we give a brief outline of the elements of permutations, integer partitions, and group actions on finite sets and their application to the notion of roots. We start with preference profiles, and give an example that demonstrates both type of equivalence classes of preference profiles: AEC, and ANEC.

A *preference profile* is an $m \times n$ matrix which shows how each of the n voters linearly orders m alternatives. We assume that the voters correspond to the columns and the alternatives correspond to the rows of the matrix. As an example, let us consider a case with $n = 4$ voters and two alternatives a_1 and a_2. There are two possible linear preference rankings for two alternatives: a_1 is strictly preferred to a_2, or a_2 is strictly preferred to a_1.

Example 2.1. For $m = 2$ and $n = 4$, $m!^n = 16$ preference profiles are as listed below.

$$x_1 = \begin{array}{|c|c|c|c|}\hline a_1 & a_1 & a_1 & a_1 \\ \hline a_2 & a_2 & a_2 & a_2 \\ \hline \end{array} \quad x_2 = \begin{array}{|c|c|c|c|}\hline a_2 & a_2 & a_2 & a_2 \\ \hline a_1 & a_1 & a_1 & a_1 \\ \hline \end{array} \quad x_3 = \begin{array}{|c|c|c|c|}\hline a_1 & a_1 & a_2 & a_2 \\ \hline a_2 & a_2 & a_1 & a_1 \\ \hline \end{array}$$

$$x_4 = \begin{array}{|c|c|c|c|}\hline a_1 & a_2 & a_1 & a_2 \\ \hline a_2 & a_1 & a_2 & a_1 \\ \hline \end{array} \quad x_5 = \begin{array}{|c|c|c|c|}\hline a_2 & a_2 & a_1 & a_1 \\ \hline a_1 & a_1 & a_2 & a_2 \\ \hline \end{array} \quad x_6 = \begin{array}{|c|c|c|c|}\hline a_2 & a_1 & a_2 & a_1 \\ \hline a_1 & a_2 & a_1 & a_2 \\ \hline \end{array}$$

$$x_7 = \begin{array}{|c|c|c|c|}\hline a_1 & a_2 & a_2 & a_1 \\ \hline a_2 & a_1 & a_1 & a_2 \\ \hline \end{array} \quad x_8 = \begin{array}{|c|c|c|c|}\hline a_2 & a_1 & a_1 & a_2 \\ \hline a_1 & a_2 & a_2 & a_1 \\ \hline \end{array} \quad x_9 = \begin{array}{|c|c|c|c|}\hline a_1 & a_1 & a_1 & a_2 \\ \hline a_2 & a_2 & a_2 & a_1 \\ \hline \end{array}$$

$$x_{10} = \begin{array}{|c|c|c|c|}\hline a_1 & a_1 & a_2 & a_1 \\ \hline a_2 & a_2 & a_1 & a_2 \\ \hline \end{array} \quad x_{11} = \begin{array}{|c|c|c|c|}\hline a_1 & a_2 & a_1 & a_1 \\ \hline a_2 & a_1 & a_2 & a_2 \\ \hline \end{array} \quad x_{12} = \begin{array}{|c|c|c|c|}\hline a_2 & a_1 & a_1 & a_1 \\ \hline a_1 & a_2 & a_2 & a_2 \\ \hline \end{array}$$

$$x_{13} = \begin{array}{|c|c|c|c|}\hline a_1 & a_2 & a_2 & a_2 \\ \hline a_2 & a_1 & a_1 & a_1 \\ \hline \end{array} \quad x_{14} = \begin{array}{|c|c|c|c|}\hline a_2 & a_1 & a_2 & a_2 \\ \hline a_1 & a_2 & a_1 & a_1 \\ \hline \end{array} \quad x_{15} = \begin{array}{|c|c|c|c|}\hline a_2 & a_2 & a_1 & a_2 \\ \hline a_1 & a_1 & a_2 & a_1 \\ \hline \end{array}$$

$$x_{16} = \begin{array}{|c|c|c|c|}\hline a_2 & a_2 & a_2 & a_1 \\ \hline a_1 & a_1 & a_1 & a_2 \\ \hline \end{array}$$

The $\binom{5}{1} = 5$ AECs for this example are

$$\{x_1\},\ \{x_2\},\ \{x_3, x_4, x_5, x_6, x_7, x_8\},\ \{x_9, x_{10}, x_{11}, x_{12}\},\ \{x_{13}, x_{14}, x_{15}, x_{16}\} \quad (2.1)$$

The profiles x_9 and x_{10} are in the same AEC since x_{10} is obtained from x_9 by interchanging the 3-rd and the 4-th columns. Preference profiles x_9 and x_{15} are in different AECs as no permutation of the columns of x_9 will give x_{15}.

When we construct the ANECs, in addition to renaming the columns, there are two possible ways of renaming the alternatives: one leaves the names of the alternatives intact, and the other switches a_1 and a_2. If we apply these operations to the AECs above, we obtain a coarser partition of the 16 preference profiles, giving us three roots (ANECs) $\theta_1, \theta_2, \theta_3$:

$$\begin{aligned}\theta_1 &= \{x_1, x_2\},\\ \theta_2 &= \{x_3, x_4, x_5, x_6, x_7, x_8\},\\ \theta_3 &= \{x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}\}.\end{aligned} \tag{2.2}$$

Therefore $R(2,4) = 3$. In (2.2), preference profiles x_9 and x_{15} are in the same equivalence class since x_{15} is obtained from x_9 by interchanging the 3-rd and the 4-th columns, and simultaneously switching a_1 and a_2.

Let $[n] = \{1, 2, \ldots, n\}$. We denote by $\mathcal{S}_n$ the group of permutations on $[n]$. A group G *acts* on a finite set Ω if each $g \in G$ gives rise to a permutation of the elements of Ω, such a way that the identity element does nothing, while a composition of actions corresponds to the action of the composition. We denote by x^g the image of $x \in \Omega$ under the permutation of Ω induced by g. The subset of Ω

$$\{x^g \mid g \in G\}$$

is called the *orbit* of $x \in \Omega$. A group action splits up Ω into a disjoint union of subsets

$$\Omega = \theta_1 + \theta_2 + \cdots + \theta_R \tag{2.3}$$

where each θ_i is a group orbit and the "+" signifies disjoint union. The θ_i are the equivalence classes under the action of G on Ω where we define $x, y \in \Omega$ to be equivalent iff there exists some $g \in G$ such that $y = x^g$. If $x^g = x$ then x is *fixed* by g. For $g \in G$ let

$$F_g = \{x \in \Omega \mid x^g = x\}$$

denote the set of elements of Ω fixed by g. Consider now a finite group G acting on a set Ω. The number R of equivalence classes can be computed by the formula

$$R = \frac{1}{|G|} \sum_{g \in G} |F_g| \tag{2.4}$$

which is known as the *Frobenius lemma*, or *Burnside lemma*. For detailed information on permutations groups and their actions on finite sets, we refer the reader to Kerber [10], or Wielandt [13].

In the setting of IANC, a preference profile of voters $[n]$ and alternatives $A = \{a_1, a_2, \ldots, a_m\}$ is represented as an $m \times n$ matrix in which each column is a permutation of A. Let $\Omega = \Omega(m,n)$ denote the set of these preference profiles. Evidently, $|\Omega| = (m!)^n$.

Roots can be characterized as the orbits of the action of the product group $\mathcal{S}_n \times \mathcal{S}_m$ on $\Omega = \Omega(m,n)$. $R = R(m,n)$ is then the number of roots. Elements of $\mathcal{S}_n \times \mathcal{S}_m$

are pairs of permutations (σ, τ) with $\sigma \in \mathcal{S}_n$ and $\tau \in \mathcal{S}_m$, where the group operation is componentwise composition of permutations. In the action of $g = (\sigma, \tau)$ on Ω, a preference profile x^g is obtained from the profile x by permuting the columns (voters) according to σ, and simultaneously permuting the alternatives by mapping each a_i to $a_{\tau(i)}$, $i = 1, 2, \ldots, m$. Representing permutations by their cycle factorization, the following example illustrates this action for $n = 4$ voters and alternatives $A = \{a_1, a_2\}$:

Example 2.2.

$$x = \begin{array}{|c|c|c|c|} \hline a_1 & a_1 & a_2 & a_2 \\ \hline a_2 & a_2 & a_1 & a_1 \\ \hline \end{array}$$

$$g = ((13)(24), (1)(2)) \to x^g = \begin{array}{|c|c|c|c|} \hline a_2 & a_2 & a_1 & a_1 \\ \hline a_1 & a_1 & a_2 & a_2 \\ \hline \end{array}$$

$$g = ((13)(24), (12)) \quad \to x^g = \begin{array}{|c|c|c|c|} \hline a_1 & a_1 & a_2 & a_2 \\ \hline a_2 & a_2 & a_1 & a_1 \\ \hline \end{array} = x$$

In particular, $g = ((13)(24), (12))$ fixes x.

$R(m, n)$ can now be computed using the Frobenius lemma (2.4). To state our results we need more notation. A *p*artition λ of an integer n is a weakly decreasing sequence of nonnegative integers $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n)$ with $n = \lambda_1 + \lambda_2 + \cdots + \lambda_n$. Each of the integers $\lambda_i > 0$ is called a *part* of λ. For example $\lambda = (3, 2, 2)$ is a partition of $n = 7$ into three parts. It has two parts of size two and one part of size three. If λ is a partition of n, then this is denoted by $\lambda \vdash n$. Each partition of n has a *type* denoted by the symbol $1^{\alpha_1} 2^{\alpha_2} \cdots n^{\alpha_n}$, which signifies that λ has α_i parts of size i for $1 \leq i \leq n$. For example the type of $\lambda = (3, 2, 2)$ is $1^0 2^2 3^1 4^0 5^0 6^0 7^0$. We can omit the zeros that appear as exponents and write the type of λ as $2^2 3^1$.

A permutation σ of $[n]$ defines a partition of n where the parts of the partition are the cycle lengths in the cycle decomposition of σ. The cycle type of σ is defined as the type of the resulting partition. For example $\sigma = (142)(35)(67)$ has cycle type $2^2 3^1$. For any $\lambda \vdash n$ of type $1^{\alpha_1} 2^{\alpha_2} \cdots n^{\alpha_n}$, define the number

$$z_\lambda = 1^{\alpha_1} 2^{\alpha_2} \cdots n^{\alpha_n} \alpha_1! \alpha_2! \cdots \alpha_n!. \tag{2.5}$$

It is well known that the number of permutations of cycle type $1^{\alpha_1} 2^{\alpha_2} \cdots n^{\alpha_n}$ is given by $z_\lambda^{-1} n!$ where λ is the partition of cycle lengths of σ. For example in the symmetric group $\mathcal{S}_7$, there are

$$\frac{7!}{2^2 3^1 2! 1!} = 210$$

permutations having the same cycle type $2^2 3^1$ as $(142)(35)(67)$. The collection of permutations which have a given cycle type is called a *conjugacy class.* In the group

$\mathcal{S}_n \times \mathcal{S}_m$ conjugacy classes are indexed by a pair of partitions $\lambda \vdash n$, $\mu \vdash m$. If C is the conjugacy class where the cycle types are given by λ and μ, then $|C| = n!m!\, z_\lambda^{-1} z_\mu^{-1}$.

For integers d and n we use the symbol $d\,|n$ to mean that d divides n evenly. For any statement S the indicator function of S is

$$\chi(S) = \begin{cases} 1 & \text{if S is True,} \\ 0 & \text{if S is False.} \end{cases}$$

For partitions λ and μ, GCD(λ) denotes the greatest common divisor (GCD) of the parts of λ, and LCM(μ) denotes the least common multiple (LCM) of the parts of μ.

For an integer k with $0 \leq k \leq x$, extend the definition of the ordinary binomial coefficient $\binom{x}{k}$ to nonintegral values of x by setting

$$\binom{x}{k} = \begin{cases} \frac{x!}{k!(x-k)!} & \text{if } x \text{ is integral,} \\ 0 & \text{otherwise.} \end{cases} \tag{2.6}$$

3. Counting roots

Theorem 3.1. *The number of roots $R(m,n)$ is given by*

$$R(m,n) = \sum_{\lambda \vdash n} \sum_{\mu \vdash m} \chi(\text{LCM}(\mu) \mid \text{GCD}(\lambda))\, z_\lambda^{-1} z_\mu^{-1}\, m!^{\alpha_1+\alpha_2+\cdots+\alpha_n} \tag{3.1}$$

where the type of λ is $1^{\alpha_1}2^{\alpha_2}\cdots n^{\alpha_n}$ and z_λ is as defined in (2.5).

Proof. As we have remarked, the number of roots $R(m,n)$ is given by the number of orbits R in the decomposition (2.3). We first determine the nature of the fixed points of $g \in \mathcal{S}_n \times \mathcal{S}_m$, and then use the Frobenius lemma to prove this theorem. Suppose $g = (\sigma, \tau)$ with the corresponding partitions $\lambda \vdash n$ and $\mu \vdash m$. Suppose the type of λ is $1^{\alpha_1}2^{\alpha_2}\cdots n^{\alpha_n}$. Then

$$|F_g| = \begin{cases} m!^{\alpha_1+\alpha_2+\cdots+\alpha_n} & \text{if } \text{LCM}(\mu) \mid \text{GCD}(\lambda), \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

To prove this claim, suppose $x^g = x$. Let t be the order of τ in $\mathcal{S}_m$. Thus t is the smallest integer such that the permutations $\tau, \tau^2, \ldots, \tau^t$ are all distinct. Consider a cycle c in the cycle decomposition of σ. Without loss of generality, we can assume that $c = (1\ 2\ \cdots\ k)$, and the first column of x is $(a_1, a_2, \ldots, a_m)$. Under the action of g on x, the first column of x is mapped to the second column, the second column to the third, etc., and finally the k-th column is mapped back to the first. At the same time under the action of τ, $(a_1, a_2, \ldots, a_m)$ is mapped to $(a_{\tau(1)}, a_{\tau(2)}, \ldots, a_{\tau(m)})$; $(a_{\tau(1)}, a_{\tau(2)}, \ldots, a_{\tau(m)})$ to $(a_{\tau^2(1)}, a_{\tau^2(2)}, \ldots, a_{\tau^2(m)})$, etc, and finally $(a_{\tau^{t-1}(1)}, a_{\tau^{t-1}(2)}, \ldots, a_{\tau^{t-1}(m)})$ back to $(a_1, a_2, \ldots, a_m)$. It follows that if

x is fixed by g, then the first k columns of x must be made up of a number of repetitions of the block of t columns

$$\begin{array}{ccccc} a_1 & a_{\tau(1)} & a_{\tau^2(1)} & \cdots & a_{\tau^{t-1}(1)} \\ a_2 & a_{\tau(2)} & a_{\tau^2(2)} & \cdots & a_{\tau^{t-1}(2)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_m & a_{\tau(m)} & a_{\tau^2(m)} & \cdots & a_{\tau^{t-1}(m)} \end{array} \tag{3.3}$$

Therefore $t \mid k$. Since this holds for any cycle of σ, t divides the GCD of the cycle lengths of σ, which is GCD(λ). On the other hand, the order of a permutation is the LCM of its cycle lengths, and therefore $t = \text{LCM}(\mu)$. Conversely, if LCM$(\mu) \mid$ GCD(λ) then the above argument shows that $x^g = x$. Hence $x^g = x$ iff LCM$(\mu) \mid$ GCD(λ). The quantity $\alpha_1 + \alpha_2 + \cdots + \alpha_n$ is the total number of cycles of σ. The columns of x permuted by each cycle of σ is determined by a single column of the cycle, which can be picked one of $m!$ ways. It follows that the number of fixed points of

$$|F_g| = \chi(\text{LCM}(\mu) \mid \text{GCD}(\lambda))\, m!^{\alpha_1+\alpha_2+\cdots+\alpha_n}. \tag{3.4}$$

Since $|F_g|$ depends only on the cycle structure of σ and τ, we can make the summation in the Frobenius lemma (2.4) over pairs of partitions that define the conjugacy classes, and multiply the expression in (3.4) by the cardinality of the corresponding conjugacy class. This gives

$$R(m,n) = \frac{1}{n!m!} \sum_{\lambda \vdash n} \sum_{\mu \vdash m} \chi(\text{LCM}(\mu) \mid \text{GCD}(\lambda))\, n!m!\, z_\lambda^{-1} z_\mu^{-1}\, m!^{\alpha_1+\alpha_2+\cdots+\alpha_n}$$

which is (3.1). □

The expression in (3.1) for $R(m,n)$ is a double sum, and the number of terms involved in the summation is the product of the number of partitions of n and the number of partitions of m. Since the number of partitions of an integer grows exponentially, the evaluation of $R(m,n)$ via (3.1) does not look practical.

Suppose however that LCM$(\mu) = d$. Then the contribution of μ to the sum (3.1) can be written as

$$z_\mu^{-1} \sum_{\substack{\lambda \vdash n \\ d\,|\lambda_i, \forall i}} z_\lambda^{-1} m!^{\alpha_1+\cdots+\alpha_n}. \tag{3.5}$$

To be able to use this expression to simplify the number of terms in the computation of $R(m,n)$, we need to evaluate the sum

$$\sum_{\substack{\lambda \vdash n \\ d\,|\lambda_i, \forall i}} z_\lambda^{-1}\, m!^{\alpha_1+\cdots+\alpha_n}. \tag{3.6}$$

Fortunately, we can find a closed form expression for (3.6) by using methods from the theory of symmetric functions.

Theorem 3.2. *For positive integers n, r, d with $d \mid r$,*

$$\sum_{\substack{\lambda \vdash n \\ d \,|\lambda_i, \forall i}} z_\lambda^{-1}\, r^{\alpha_1+\cdots+\alpha_n} = \binom{\frac{n}{d}+\frac{r}{d}-1}{\frac{r}{d}-1} \tag{3.7}$$

where the type of λ is λ is $1^{\alpha_1}2^{\alpha_2}\cdots n^{\alpha_n}$ and the binomial coefficient is defined as in (2.6).

Proof. Let S denote the left hand side of (3.7). Unless d divides each λ_i (and consequently n), S is zero. Otherwise let $\rho_i = \frac{\lambda_i}{d}$ for $i = 1, 2, \ldots, n$. Then $\rho \vdash \frac{n}{d}$. If the type of λ is $1^{\alpha_1}2^{\alpha_2}3^{\alpha_3}\cdots$, then the type of ρ is $1^{\alpha_d}2^{\alpha_{2d}}3^{\alpha_{3d}}\cdots$. Then

$$\begin{aligned} S &= \sum_{\substack{\lambda \vdash n \\ d \,|\lambda_i, \forall i}} \frac{r^{\alpha_d+\alpha_{2d}+\alpha_{3d}+\cdots}}{d^{\alpha_d}(2d)^{\alpha_{2d}}(3d)^{\alpha_{3d}}\cdots \alpha_d!\alpha_{2d}!\alpha_{3d}!\cdots} \\ &= \sum_{\rho \vdash \frac{n}{d}} \frac{r^{\alpha_d+\alpha_{2d}+\alpha_{3d}+\cdots}}{d^{\alpha_d}(2d)^{\alpha_{2d}}(3d)^{\alpha_{3d}}\cdots \alpha_d!\alpha_{2d}!\alpha_{3d}!\cdots} \\ &= \sum_{\rho \vdash \frac{n}{d}} z_\rho^{-1}\left(\frac{r}{d}\right)^{\alpha_d+\alpha_{2d}+\alpha_{3d}+\cdots}. \end{aligned}$$

To evaluate this last expression, we use an identity from the theory of symmetric functions. The n-th power sum p_n and the n-th homogeneous (or complete) symmetric function h_n in the variables $x_1, x_2, \ldots, x_N$ are defined by setting

$$p_n = \sum_{i=1}^{N} x_i^n \tag{3.8}$$

$$h_n = \sum_{1 \le i_1 \le i_2 \le \cdots \le i_n \le N} x_{i_1}x_{i_2}\cdots x_{i_n}. \tag{3.9}$$

For any partition $\lambda = (\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n)$ of n, define

$$p_\lambda = \prod_{i=1}^{n} p_{\lambda_i}. \tag{3.10}$$

It can be shown that (see MacDonald [11])

$$h_n = \sum_{\lambda \vdash n} z_\lambda^{-1}\, p_\lambda.$$

Take the number of variables $N = \frac{r}{d}$ and put each $x_i = 1$. Then each p_{λ_i} evaluates to $\frac{r}{d}$. Therefore p_λ evaluates to $(\frac{r}{d})^{\alpha_1+\alpha_2+\cdots+\alpha_n}$ where the type of λ is $1^{\alpha_1}2^{\alpha_2}\cdots n^{\alpha_n}$.

In our case the partition in question is ρ, and consequently p_ρ evaluates to $(\frac{r}{d})^{\alpha_1+\alpha_{2d}+\alpha_{3d}+\cdots}$. Therefore S is given by n-th homogeneous symmetric function h_n in the variables $x_1, x_2, \ldots, x_{\frac{r}{d}}$ where each variable is set equal to 1. From the definition (3.9), this is a balls-in-boxes type of a count: it is the number of ways of distributing $\frac{n}{d}$ indistinguishable balls into $\frac{r}{d}$ distinguishable boxes (see Feller [5]). This is given by the binomial expression on the right hand side of (3.7). □

Combining the two results we obtain

Theorem 3.3.

$$R(m,n) = \sum_{\mu \vdash m} z_\mu^{-1} \binom{\frac{n}{d} + \frac{m!}{d} - 1}{\frac{m!}{d} - 1} \tag{3.11}$$

where $d = d(\mu) = \mathrm{LCM}(\mu)$, the binomial coefficient is defined as in (2.6), and z_μ is as defined in (2.5).

Note that the summation in (3.11) is over partitions of m only, and is independent of the number of voters n.

4. Explicit formulas

Theorem 3.3 has some immediate implications. We obtain explicit formulas for the number of roots for small values of m as follows.

4.1. n voters and $m = 2$ alternatives

For $m = 2$, the partitions of m are $(1,1)$ and (2) with $\mathrm{LCM}(1,1) = 1$, $\mathrm{LCM}(2) = 2$ and $z_{(1,1)} = z_{(2)} = 2$. Therefore

$$R(2,n) = \frac{1}{2}\binom{n+1}{1} + \frac{1}{2}\binom{\frac{n}{2}}{0} \tag{4.1}$$

This is another way of saying

$$R(2,n) = \begin{cases} \frac{1}{2}n + 1 & \text{if } n \text{ is even,} \\ \frac{1}{2}(n+1) & \text{if } n \text{ is odd.} \end{cases}$$

4.2. n voters and $m = 3$ alternatives

For $m = 3$, there are three partitions $(1,1,1)$, $(2,1)$, and (3) of m with $\mathrm{LCM}(1,1,1) = 1$, $\mathrm{LCM}(2,1) = 2$, $\mathrm{LCM}(3) = 3$, and $z_{(1,1,1)} = 6$, $z_{(2,1)} = 2$, and $z_{(3)} = 3$. Thus

$$R(3,n) = \frac{1}{6}\binom{n+5}{5} + \frac{1}{2}\binom{\frac{n}{2}+2}{2} + \frac{1}{3}\binom{\frac{n}{3}+1}{1}. \tag{4.2}$$

4.3. n voters and $m = 4$ alternatives

The partitions of $m = 4$ are $(1,1,1,1)$, $(2,1,1)$, $(2,2)$, $(3,1)$, (4) with

$$\mathrm{LCM}(1,1,1,1) = 1, \ \ \mathrm{LCM}(2,1,1) = \mathrm{LCM}(2,2) = 2,$$
$$\mathrm{LCM}(3,1) = 3, \ \ \mathrm{LCM}(4) = 4,$$
$$z_{(1,1,1,1)} = 24, \ \ z_{(2,1,1)} = 4, \ \ z_{(2,2)} = 8, \ \ z_{(3,1)} = 3, \ \ z_{(4)} = 4.$$

Therefore

$$R(4,n) = \frac{1}{24}\binom{n+23}{23} + \frac{3}{8}\binom{\frac{n}{2}+11}{11} + \frac{1}{3}\binom{\frac{n}{3}+7}{7} + \frac{1}{4}\binom{\frac{n}{4}+5}{5}.$$

4.4. n and $m!$ relatively prime

In this case, only the term corresponding to the partition is $\mu = (1,1,\ldots,1)$ in the sum (3.11) is nonzero. Since with $z_\mu = m!$ for this partition, an immediate corollary is the following result of Giritligil and Doğan [8]:

Corollary. *When n and $m!$ are relatively prime, the number of roots $R(m,n)$ is given by*

$$R(m,n) = \frac{1}{m!}\binom{n+m!-1}{m!-1}.$$

Remark 4.1. By means of a symbolic algebra package such as Mathematica, we can easily calculate the value of $R(m,n)$ for relatively large values of m and n using the general formula in Theorem 3.3. As examples

$$\begin{aligned} R(5,5) &= 1876255 \\ R(5,10) &= 2049242056940 \\ R(5,20) &= 5908312923863263889174 \\ R(5,30) &= 21465856893663082687992576842 0\,. \end{aligned}$$

5. Dixon–Wilf algorithm and uniform generation of roots

The importance of being able to access the values of $R(m,n)$ for large and unconstrained values of m and n becomes apparent when we try to generate roots from the uniform distribution. The ability to compute the value of $R(m,n)$ together with the Dixon–Wilf algorithm allows us to construct a symbolic package to generate the roots such that each root is produced with probability $1/R(m,n)$.

Suppose in general that a group of permutations G acts on a set Ω. Consider the decomposition of Ω into orbits $\theta_1, \theta_2, \ldots, \theta_R$ as in (2.3). If the number of orbits R is known, then the following procedure, usually referred to as the Dixon–Wilf algorithm (Dixon and Wilf [2]) can be used to generate an orbit θ from the uniform distribution.

5.1. Basic elements of the Dixon–Wilf algorithm

(i) Select a conjugacy class $C \subseteq G$ with probability

$$p_C = \frac{|C||F_g|}{R \cdot |G|}$$

where g is some member of C.

(ii) Select uniformly at random some $x \in F_g$.

(iii) Return the orbit θ that contains x.

The crucial aspect of the Dixon–Wilf algorithm is that it is guaranteed to return an orbit (or the representative x of the orbit) distributed uniformly over the set of all orbits. The task of having to generate an orbit from the uniform distribution is transferred to being able to select a conjugacy class in G with a certain probability as given in the first step of the Dixon–Wilf algorithm, and then being able to pick uniformly a random profile from a fixed point set F_g.

We calculate the number of orbits $R = R(m, n)$ by using Theorem 3.3. In addition, we can calculate the necessary parameters as required in the Dixon–Wilf algorithm such as the size of the conjugacy classes for the product group of IANC.

A conjugacy class $C \subseteq \mathcal{S}_n \times \mathcal{S}_m$ is defined by a pair of partitions $\lambda \vdash n$, $\mu \vdash m$. Suppose the type of λ is $1^{\alpha_1} 2^{\alpha_2} \cdots n^{\alpha_n}$ and $g = (\sigma, \tau)$ is an arbitrary element of C. We have

$$\begin{aligned} |G| &= n!m! \\ |C| &= n!m!\, z_\lambda^{-1} z_\mu^{-1} \\ |F_g| &= \chi(\text{LCM}(\mu) \mid \text{GCD}(\lambda))\, m!^{\alpha_1+\alpha_2+\cdots+\alpha_n} \\ R &= R(m, n). \end{aligned}$$

Therefore we need to pick C with probability

$$p_C = \frac{\chi(\text{LCM}(\mu) \mid \text{GCD}(\lambda))\, z_\lambda^{-1} z_\mu^{-1}\, m!^{\alpha_1+\alpha_2+\cdots+\alpha_n}}{R(m, n)}.$$

Consider the list $\mathcal{L} = \{(\pi_i, f_i) \mid i = 1, 2, \ldots, R\}$ where each π_i is a pair of partitions $\lambda \vdash n$, $\mu \vdash m$ with $\text{LCM}(\mu) \mid \text{GCD}(\lambda)$. For such a pair π_i, the corresponding fraction f_i is defined by

$$f_i = \frac{z_\lambda^{-1} z_\mu^{-1}\, m!^{\alpha_1+\alpha_2+\cdots+\alpha_n}}{R(m, n)}.$$

Thus $f_1, f_2, \ldots, f_R$ is an ordering of the nonzero probabilities p_C of conjugacy classes. Compute the partial sums

$$s_1 = f_1,\ \ s_2 = f_1 + f_2,\ \ s_3 = f_1 + f_2 + f_3,\ \ldots,\ s_R = f_1 + f_2 + \cdots + f_R = 1.$$

If we generate a real number in $x \in [0,1]$ uniformly, the probability that $s_{i-1} < x \leq s_i$ is precisely the probability p_C, where C is the i-th conjugacy class in the ordering of the elements of $\mathcal{L}$. Suppose the selected class is defined by the pair of partitions $\lambda \vdash n$, $\mu \vdash m$. We pick $\sigma \in \mathcal{S}_n$ of type λ by writing the integers $1, 2, \ldots, n$, and grouping them into cycles in the order they appear. For instance for $\lambda = (3, 2, 2)$ this yields $\sigma = (1\,2\,3)(4\,5)(6\,7)$. Similarly we construct $\tau \in \mathcal{S}_m$ by considering the type of μ and grouping the integers $1, 2, \ldots, m$ accordingly in the order they appear. The resulting pair is our element $g \in C$.

Finally, we need to return a preference profile $x \in F_g$, making sure that x is selected uniformly at random from F_g. This is easily accomplished by picking uniformly at random a permutation of $a_1, a_2, \ldots, a_m$ for every cycle of σ, and by placing the permuted alternatives as the smallest indexed column in each cycle. The other columns of the cycle are then filled up by the images of this initial permuted column under the iterates of τ as in (3.3).

We have implemented this idea to generate roots from the uniform distribution as a Mathematica program called $\texttt{GenerateRoot}[\texttt{m}, \texttt{n}]$. The program takes a pair of integers m, n as input and generates an $m \times n$ preference profile x. The resulting x is guaranteed to be distributed over the $R(m, n)$ roots uniformly. This is the surprising application of the Dixon–Wilf algorithm.

Example 5.1. We have run the uniform root generation algorithm k times, for k running from 10 to 10000 in powers of 10 for $m = 2$ and $n = 4$. For each x returned by $\texttt{GenerateRoot}[2, 4]$ we checked whether $x \in \theta_1$, $x \in \theta_2$, or $x \in \theta_3$ (see (2.2) in Section 2). Pr[Hits from orbit θ_1] is the ratio of the number of $x \in \theta_1$ to k. Pr[Hits from orbit θ_2] and Pr[Hits from orbit θ_3] are calculated similarly. Figure 1 shows the resulting computed probabilities. Since there are 3 orbits $\theta_1, \theta_2, \theta_3$ in this case, the actual probability for each is 0.333 . . .

No. of trials k	Pr[Hits from θ_1]	Pr[Hits from θ_2]	Pr[Hits from θ_3]
10	0.2	0.5	0.3
100	0.34	0.29	0.37
1000	0.352	0.348	0.3
10000	0.3235	0.3396	0.3369

Figure 1. Random generation of roots from the uniform distribution with $n = 4$ voters and $m = 2$ alternatives. Each trial is the generation of a root from $\Omega(2, 4)$ by using the Mathematica routine $\texttt{GenerateRoot}[2, 4]$.

After preprocessing, about 7 milliseconds were required to generate a 5 × 20 preference profile, and about 24 milliseconds to generate a 7 × 25 preference profile. The CPU time necessary to generate a random root from the uniform distribution for various values of the input parameters n and m, as well as the measurement details are given in Figure 2.

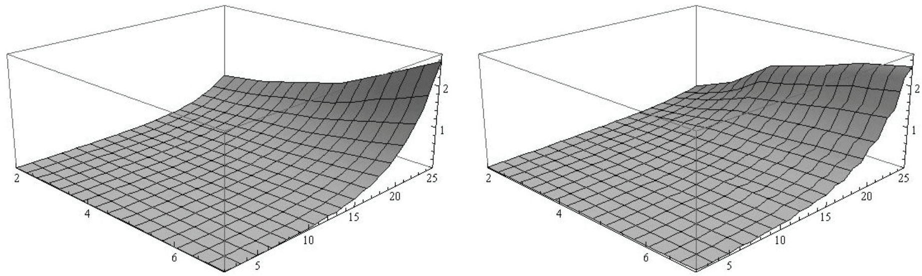

Figure 2. The computer time needed to generate preference profiles in the IANC model as a function of the input integers n and m using GenerateRoot[m, n]. The first graph is the initialization phase of the algorithm for a given n, m pair. This is only executed once for each pair of interest. The second graph is the time required for the generation of 100 preference profiles for the given parameters n, m, after the initialization. The range of values plotted is $2 \leq m \leq 7$ and $2 \leq n \leq 25$. The vertical axis is time in seconds, ranging from 0 to 2.75. The experiments were performed on Mathematica 6.0 running on a desktop with an 2.40 GHz Intel Core 2 CPU and 3.24 GB of RAM.

5.2. Likelihood of Condorcet winner to be a Plurality winner

As an example of applications of IANC and GenerateRoot[m, n], we describe a Monte-Carlo experiment to compute the probability of Condorcet and Plurality Rule's winners to coincide for varying values of m and n.

Recall that an alternative is a *Condorcet-winner* if it is preferred to each other alternative by a majority of voters. However, a *Condorcet Paradox* occurs when social outcome is not transitive, even though the individual preferences are not, due to the conflict in majority wishes.

Plurality Rule simply chooses the alternatives which are most popular as top-ranked candidates in a profile. Consider the likelihood of Condorcet Rule and Plurality Rule choosing the same winner. For simplicity, consider only odd values of n which guarantees that the Condorcet winner, when it exists, is unique. Plurality can choose multiple winners. In this case, we check if any one of the Plurality winners is the Condorcet winner.

The procedure followed for the experiment is as follows: Given n and m, let GenerateRoot[m, n] generate a random root uniformly. If the generated profile x

does not have a Condorcet winner, then we simply generate another root. For each root that does have a Condorcet winner, say a_i, we check and see if a_i is also chosen by Plurality. For this we consider the first row of x and make sure that a_i occurs in this row at least as many times as every a_j, for $1 \leq j \leq m$. The ratio of the number of roots in which the Condorcet winner is also a plurality winner to the total number of roots generated which have Condorcet winners is an approximation to the probability that a Condorcet winner is also a Plurality winner.

A plot of these probabilities for various n and m computed by using $k = 1000$ Condorcet winners for each case, appears in Figure 3.

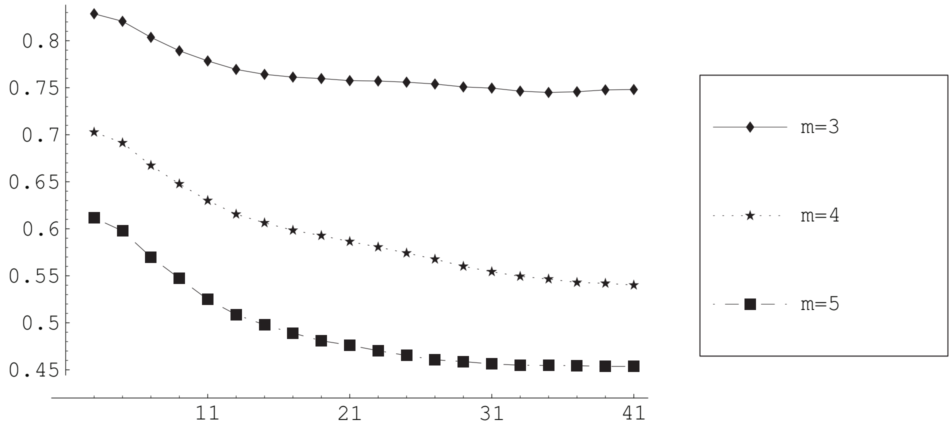

Figure 3. The probability that the Condorcet winner is also a Plurality winner in the IANC model. The raw data has been smoothed out by a 5-term moving-average filter. The horizontal axis is the number of voters n, through odd integers from 3 to 41. The number of samples used is $k = 1000$ per m/n pair.

6. Concluding remarks

Based on two fundamental axioms of social choice, anonymity and neutrality, the IANC model uses root profiles for generating public preferences, where the names of both the voters and the alternatives are ignored.

We derived an efficient formula for their number, and described the the ingredients of a symbolic algebra package for the generation of roots from the uniform distribution by means of the Dixon–Wilf algorithm. In this way, IANC allows for the analysis of the behaviors of anonymous and neutral SCRs with respect to varying number of alternatives and voters by means of Monte-Carlo methods.

Applications of this model and the procedure $\texttt{GenerateRoot}[\texttt{m},\texttt{n}]$ to the study of experimental comparisons for various SCRs is in progress (Eğecioğlu and Giritligil [3], [4]).

Acknowledgments. I would like to thank A. E. Giritligil, who got me interested in this topic during my sabbatical leave in 2004 at Sabanci University, Istanbul, Turkey.

References

1. S. Berg and D. Lepelley, On Probability Models in Voting Theory, *Statistica Neerlandica* **48** (1994), 33–146.
2. J. Dixon and H. Wilf, The random selection of unlabeled graphs, *Journal of Algorithms*, **4** (1983), 205–212.
3. Ö. Eğecioğlu and A. E. Giritligil, Public Preference Structures with Impartial Anonymous and Neutral Culture Model, preprint.
4. Ö. Eğecioğlu and A. E. Giritligil, The Likelihood of Choosing the Borda-winner with Partial Preference Rankings of the Electorate, preprint.
5. W. Feller, *An Introduction to Probability Theory and its Applications,* 3rd Edition, Wiley, New York, 1957.
6. P. C. Fishburn and W. Gehrlein, Condorcet Paradox and Anonymous Preference Profiles, *Public Choice* **26** (1978), 1–18.
7. W. V. Gehrlein, Condorcet's Paradox, the Condorcet Efficiency of the Voting Rules, *Mathematica Japonica* **45** (1997), 173–199.
8. A. E. Giritligil and O. Doğan, An Impossibility Result on Anonymous and Neutral Social Choice Functions, preprint.
9. G. T. Guilbaud, Les théories de l'intéret genéral et le problémelogique de l'agrégation, *Economie Appliquée* **5** (1952), 501–584.
10. A. Kerber, *Algebraic Combinatorics Via Finite Group Actions,* BI & Brockhaus, Mannheim, 1991.
11. I. G. MacDonald, *Symmetric Functions and Hall Polynomials,* Oxford University Press, 1998.
12. H. Moulin, *The Strategy of Social Choice,* North-Holland Publishing Company, New York, 1983.
13. H. Wielandt, *Finite Permutation Groups,* Academic Press, New York, 1964.

Received 19 March, 2008; revised 06 July, 2009

Author information

Ömer Eğecioğlu, Department of Computer Science, University of California, Santa Barbara, CA 93106, USA.
Email: omer@cs.ucsb.edu