Dice games and Arrow's theorem

Lorenzo Traldi
Department of Mathematics
Lafayette College, Easton, PA 18042

Abstract

We observe that non-transitivity is not the most general way that Arrow's impossibility theorem is reflected in dice games.

Non-transitivity in dice games is a well-known phenomenon. It was described in several *Scientific American* columns in the 1970s, beginning with [3], and has since become a popular topic, discussed in articles like [4, 5] and textbooks on elementary mathematics and probability. In many discussions, non-transitivity of dice is mentioned in connection with voting paradoxes like Arrow's impossibility theorem [1], which states that transitivity and several other seemingly natural conditions cannot all be satisfied by any mechanism that might be used to determine the outcome of elections involving three or more candidates. Our purpose here is to observe that non-transitivity is not actually the most general way that dice games illustrate Arrow's theorem: dice games connected with elections involving three or more candidates (or more generally, dice games connected with elections in which voters express three or more levels of preference) violate independence of irrelevant alternatives (IIA), but only dice games connected with elections involving four or more candidates can be non-transitive.

If m is a positive integer then an *m-sided generalized die with integer labels* is simply a list $X = (x_1, ..., x_m)$ of integers; for convenience' sake we presume that $x_1 \leq x_2 \leq ... \leq x_m$. For integers $a \leq b$ let $D(a, b)$ denote the collection of all such dice with $a \leq x_1 \leq ... \leq x_m \leq b$. If $X = (x_1, ..., x_m), Y = (y_1, .., y_n) \in D(a, b)$ then we call $|\{(i, j)|x_i > y_j\}| - |\{(i, j)|x_i < y_j\}|$ the *win-loss difference of X against Y*. X is *stronger* than Y if this difference is positive, and Y is stronger if the difference is negative; if the win-loss difference is 0 then X and Y are *tied*. This relation reflects the natural game in which a "roll" consists of picking one of the x_i at random with probability $1/m$, and picking one of the y_j at random with probability $1/n$; X wins that roll if $x_i > y_j$. Note that if $b - a = d - c$ then $D(a, b)$ is isomorphic to $D(c, d)$ under the map $(x_1, ..., x_m) \mapsto (x_1+c-a, ..., x_m+c-a)$.

Arrow's theorem [1] applies to elections in which each voter expresses a *preference order* of the candidates. A preference order is a reflexive, transitive relation which is complete (no two candidates are incomparable); a voter may express identical levels of preference for some candidates. Observe that a preference order does not specify how strongly a voter prefers one candidate over another; see [1, 2] for discussions of the difference between ordinal and cardinal measures of preference.

The preference order of an individual voter is equivalently expressed by assigning each candidate the rating $r = |\{\text{candidates whom that voter does not strictly prefer to that candidate}\}|$. The ratings assigned to a given candidate by the various voters determine an element of $D(1, c)$, where c is the number of candidates. If we allow voters to decide to abstain from assessing certain candidates, the dice may not all be the same size. Although the *stronger* relation on $D(1, c)$ is a natural way to judge a dice game, it is rather unnatural as a vote-counting scheme: candidate X is stronger than candidate Y if, when one among the m voters who assessed X is chosen randomly with probability $1/m$ and one among the n voters who assessed Y is chosen randomly with probability $1/n$, it is more likely than not that the first voter's rating of candidate X is at least as large as the second voter's rating of candidate Y.

A crucial feature of the representation of elections with dice is commonly neglected. When we restrict our attention to an election involving a subset of the original set of candidates, the ballots pertinent to this "sub-election" are obtained by restricting the voters' preference orders to that subset; in general, restricting the preference orders will change the dice associated with the candidates.

Example 1. Consider the *Condorcet triple*.

voter	1	2	3
rating for candidate 1	3	1	2
rating for candidate 2	2	3	1
rating for candidate 3	1	2	3

If we represent this election with dice then every candidate is represented by the same die $(1, 2, 3)$, so naturally no candidate is stronger than any other. However, if we restrict our attention to the election involving only candidates 1 and 2 then candidate 1 wins because voters 1 and

3 both favor candidate 1 over candidate 2, so that in this election candidate 2 is represented by the die $(1, 1, 2)$ and candidate 1 is represented by the stronger die $(1, 2, 2)$. Consequently, when applied to the Condorcet triple the vote-counting scheme based on dice games violates IIA: removing candidate 3 from consideration changes the assessment of whether or not candidate 1 is stronger than candidate 2.

Example 2. Here is an example which appears in several places in [2].

voter	1	2	3	4	5
rating for candidate 1	3	3	3	1	1
rating for candidate 2	2	2	2	3	3
rating for candidate 3	1	1	1	2	2

If this election is represented as a dice game, then candidate 2 is represented by the strongest die, with candidate 1 next and candidate 3 last. However IIA is violated because in the election involving only candidates 1 and 2, candidate 2 is represented by the die $(1, 1, 1, 2, 2)$ and candidate 1 is represented by the stronger die $(1, 1, 2, 2, 2)$.

In each of these examples the *stronger* relation is transitive – it is trivial in the first example, and a strict linear order in the second. It turns out that all dice games representing three-candidate elections are transitive.

Theorem. *If* $a \in \mathbb{Z}$ *then* stronger *is a transitive relation on* $D(a, a+2)$.

Proof. Suppose $X = (x_1, ..., x_m) \in D(a, a+2)$ and let m_a, m_{a+1} and m_{a+2} denote (respectively) the numbers of labels x_i which are equal to $a, a+1$ and $a+2$. Define the *strength* of X as follows.

$$str(X) = \frac{m_{a+2} - m_a}{m_a + 2m_{a+1} + m_{a+2}}$$

If $X = (x_1, ..., x_m), Y = (y_1, ..., y_n) \in D(a, a+2)$ then we claim that X is stronger than Y if and only if $str(X) > str(Y)$. To see why, observe that if n_a, n_{a+1} and n_{a+2} are (respectively) the numbers of labels y_j which equal $a, a+1$ and $a+2$ then

$$\begin{aligned} & 2 \cdot |\{(i,j)|x_i > y_j\}| - 2 \cdot |\{(i,j)|x_i < y_j\}| \\ = \quad & 2m_{a+2}(n_a + n_{a+1}) + 2m_{a+1}(n_a - n_{a+2}) - 2m_a(n_{a+1} + n_{a+2}) \\ = \quad & (m_a + 2m_{a+1} + m_{a+2})(n_a + 2n_{a+1} + n_{a+2})(str(X) - str(Y)), \end{aligned}$$

so $str(X) - str(Y)$ and the win-loss difference of X against Y are both positive, both negative or both zero.

The usual ordering of real numbers is transitive, so it follows that *stronger* is a transitive relation on $D(a, a+2)$. Q. E. D.

Observe that the sum (or equivalently, the mean) of labels does not determine *stronger* for elements of $D(a, a+2)$. For instance, suppose X is a 50-sided die with 39 labels equal to 0 and 11 labels equal to 1 and Y is a 50-sided die with 10 labels equal to -1, 20 labels equal to 0 and 20 labels equal to 1. Then the labels of X have a greater sum than do those of Y, but Y is a stronger die than X. We leave it to the reader interested in the interrelationships among vote-counting schemes to verify that if the dice X and Y reflect 50 voters' ratings of two of the candidates in an election, then either of these two candidates could be preferred over the other by a plurality. Similarly, Example 2 above can be modified in such a way that candidate 2 defeats candidate 1 by a plurality in their pairwise contest, without changing the dice representing candidates 1 and 2 in the three-way contest.

To summarize: if the voters in an election express three levels of preference, then the only way the dice games corresponding to this election and its "sub-elections" can illustrate Arrow's theorem is by violating IIA. If there are more than three levels of preference, then there may be non-transitivity too; for instance the dice $(1, 1, 4, 4)$, $(1, 3, 3, 3)$, and $(2, 2, 2, 4)$ form a cycle in $D(1, 4)$.

References

[1] K. J. Arrow, *Social Choice and Individual Values*, 2nd ed., J. Wiley and Sons, New York, 1963.

[2] K. J. Arrow, A. K. Sen and K. Suzumura (editors), *Handbook of Social Choice and Welfare*, vol. 1, Elsevier Science, Amsterdam, 2002.

[3] M. Gardner, *Mathematical games: the paradox of the non-transitive dice and the elusive principle of indifference*, Scientific American **223** (1970), 110-114.

[4] R. P. Savage, Jr., *The paradox of nontransitive dice*, Amer. Math. Monthly **101** (1994), 429-436.

[5] R. L. Tenney and C. C. Foster, *Non-transitive dominance*, Math. Mag. **49** (1976), 115-120.