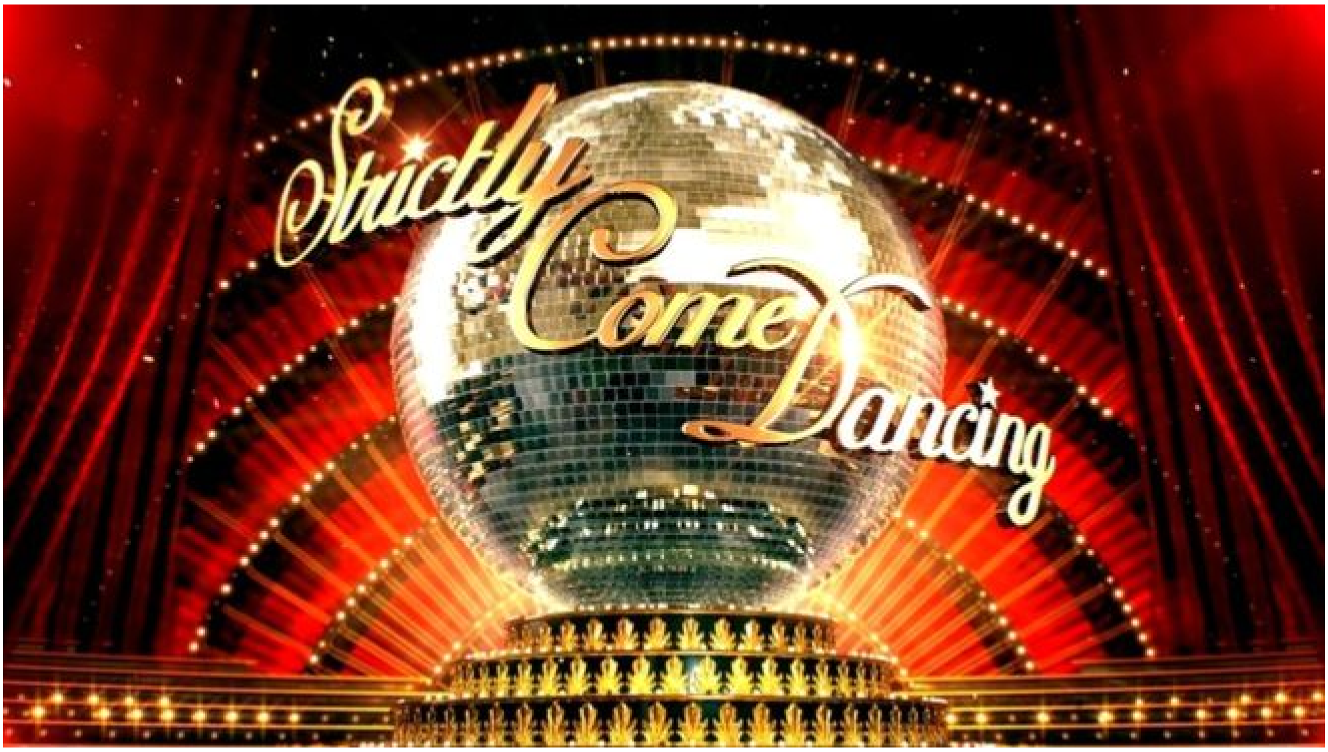
Norman Biggs – Strictly not dancing

Norman Biggs is Emeritus Professor of Mathematics at LSE. He is the author of 13 books and over 100 papers on Mathematics, and has also published in the fields of Numismatics and Metrology. He has been Librarian and General Secretary of the London Mathematical Society, and Vice-President of the British Society for the History of Mathematics. Follow him on Twitter: @norman_biggs.

This note addresses some questions arising from the popular UK television programme Strictly Come Dancing, whose format has been sold to more than 50 countries. The programme's makers have devised a complex voting algorithm, and the inputs are only partially revealed, so it is a challenge to discover how much information can be retrieved by mathematical analysis. In outline, the voting procedure is as follows: first each of the four judges publicly award scores out of 10, and these are converted into points, albeit by a rather peculiar method. Then the public is invited to vote, and the result is also converted into points, and combined with the judges' points to produce a final ranking. However, the public points are not revealed, only the identity of the two lowest contestants in the final ranking.

Here it is shown that in some circumstances the revealed data can indeed provide a great deal of information about the public vote. It is also shown that the peculiar method of producing the judges' points, which might at first sight be thought to provide them with extra voting power, can actually make it less likely that their preferences will prevail. Note: the calculations presented below have been done by hand, and are not guaranteed. However, any minor errors would not affect the main conclusions.



BBC's "Strictly Come Dancing" title sequence

1. The Show

Ostensibly, the aim of the show is to decide who is the best dancer among a group of contestants. There are four professional judges, who award each contestant a mark out 10, based on their performance. However, this is not the only criterion; after the judges have delivered their verdicts, the public is invited to vote, and their votes, together with the judges scores, determine the outcome. So it is also a popularity contest.

In fact, the objective of the show's makers (the British Broadcasting Corporation, aka the BBC) is not to find the best dancer, or the most popular one. It is to maximise the number of people who watch the show. The BBC wishes to ensure that Strictly Come Dancing attracts more viewers than a similar show, made by their rival network (Independent Television, aka ITV) and broadcast at around the same time. This particular game has been in play for many years, and some aspects of the show can only be properly understood in that context.

2. The problem

It is helpful to use some mathematical notation in order to describe exactly how the voting algorithm works. An example is given in the next section.

Suppose there are n contestants. The judges' scores produce a total mark out of 40 for each contestant, and these marks are arranged in numerical order. The contestants with the top mark are awarded n points, those with next highest mark $n-1$ points, and so on. A feature of this procedure is that, when ties occur, the total number of points awarded is increased. For example, in a recent show with 6 contestants the judges awarded 6 + 5 + 5 + 4 + 3 + 2 = 25 points, rather than 6 + 5 + 4 + 3 + 2 + 1 = 21. But with the public votes the numbers involved are so large that it is safe to assume that ties do not occur, and the result can be converted into a list of points in the form $n, n-1, \ldots, 2, 1$ in the usual way.

Denote by jX and pX the number of points awarded to contestant X by the judges and the public respectively. Then the final *rank* of X is ($jX + pX, pX$). This provides a complete linear ordering of the contestants according to the rule that (s, t) is superior to (u, v) if either $s > u$ or $s = u$ and $t > v$. The crucial problem is that neither the combined ranking nor the public ranking are disclosed. The only information provided to the public is the identity of the two contestants who occupy the lowest positions: we shall refer to them as the *chosen* ones. (In the show, these unfortunates are asked to compete in a 'Dance-Off', and the judges then eliminate the one whom they consider to be worse.)

The question that arises naturally is: how much information about the public voting can be deduced from the limited facts available? The public can assign points to the n contestants in $n!$ ways, and we should like to know how many of them are *consistent*; that is, how many produce a combined ranking in which the two chosen ones actually receive the lowest ranks. As we shall see, in the later stages of the competition, when the number n is small (and the public tends to put less weight on the judges' comments), significant conclusions can be made.

3. The example

On 5 December 2015 the judges' scores and the resulting points were as follows:

X	J	G	Ke	Ka	H	A
Score	39	36	36	35	34	31
jx	6	5	5	4	3	2

In the judges' view, the two chosen contestants should have been H and A. There was, therefore, much gnashing of judicial teeth when it was declared that, after taking account of the public points, the two chosen ones were in fact G and H. Clearly the public did not agree with the judges.

Our question is: how many of the 6! = 720 possible orderings of the public votes are consistent with the declared outcome? Consider first the three contestants A, G, and H. Because G and H were chosen, but A was not, it follows that

$$p_G + 5 \leq p_A + 2 \qquad \text{and} \qquad p_H + 3 \leq p_A + 2.$$

We also know that pA, pG, pH are distinct numbers in the range from 1 to 6. In fact one of them must be 1, since if any of pJ, pKe, pKa were equal to 1, then that contestant would be chosen instead of G, because their rank would be inferior to G's rank of least (7,2). At this stage we can conclude that there are just 11 possibilities:

p_A :	4	5	5	6	6	6
p_G :	1	1	2	1	2	3
p_H :	2, 3	2, 3, 4	1	2, 3, 4	1	1

For each possibility there are potentially 6 ways of assigning the remaining three points-values to the other three contestants, but they are not all consistent. Denote the number of consistent assignments when $pA = a, pG = b, pH = c$,by $N(abc)$; this depends effectively only on b and c, because those numbers determine the ranks of the chosen ones G and H, and only assignments which provide all of J, Ke, and Ka with superior ranks are allowed.

For example, N(514) = N(614) = 1. In both cases, H has rank (7; 4) and one of pJ, pKe, pKa must be 2. If pKe= 2 then Ke has inferior rank (7,2), and if pKa = 2 or 3 then Ka has inferior rank (6; 2) or (7; 3). Hence in both cases there is only one possibility: pJ = 2, pKe = 3, and pKa = 6 or 5 respectively. A similar line of argument leads to the results

$$N(412) = N(512) = N(612) = 6, \quad N(413) = N(513) = N(613) = 4,$$

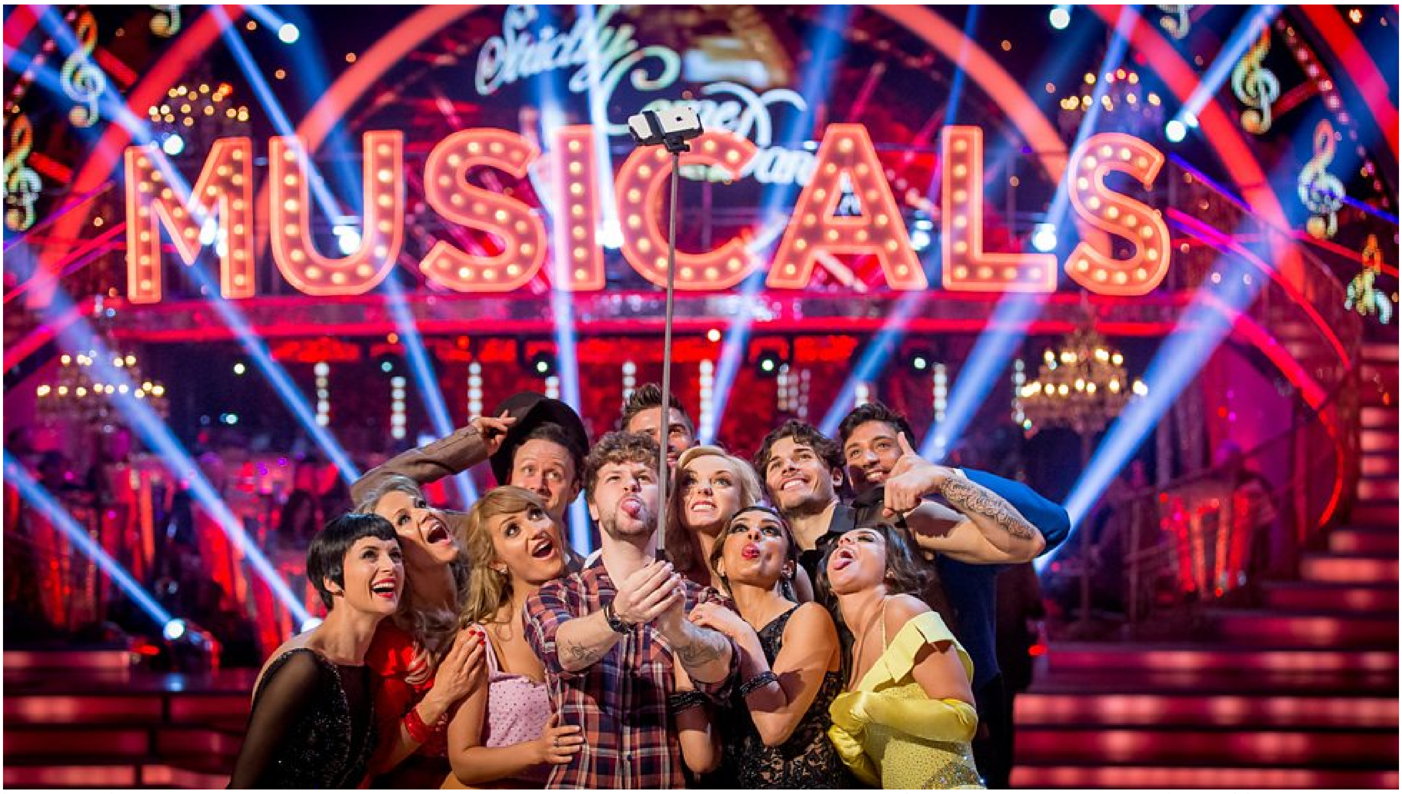
$$N(514) = N(614) = 1, \quad N(521) = N(621) = 6, \quad N(631) = 0.$$

Thus just 44 of the 720 possible orderings are consistent with the declared outcome.

On the assumption that the 44 possibilities are equally likely, we can make an estimate eX of the public's support for contestant X by taking the average of the points scored by X over the 44 possibilities. The answers (approximately) are:

$$e_J = 3.95, \quad e_G = 1.27, \quad e_{Ke} = 4.00, \quad e_{Ka} = 4.52, \quad e_H = 2.09, \quad e_A = 5.16.$$

It must be stressed that this is simply the best guess that we can make, on the basis of the information revealed to us. The truth is out there, but the BBC do not reveal it.



"Strictly Come Dancing" 2015 contestants during musicals week (picture courtesy of BBC)

4. What if…?

The peculiar rule for translating the judges scores into points has already been noted. A more usual rule is that if $m$contestants have the same score and each receives l points, then those with the next highest score receive $l-m$ points. In our example, this would result in the points awarded being 6, 5, 5, 3, 2, 1. In other words, the last three contestants would each receive one point fewer than before. (Note that the total is still greater than 21, but only by one point.)

So we can reasonably ask: if the alternative rule had been used, would the outcome have been different? This entails working out the ranking that results when each of the 44 possible public assignments is combined with the new set of judges points. For example, for any of the six assignments with abc= 412, the old and new calculations are as follows:

contestant	J	G	Ke	Ka	H	A
old judges points	6	5	5	4	3	2
public points		1			2	4
old rank		$(6,1)$			$(5,2)$	$(6,4)$

contestant	J	G	Ke	Ka	H	A
new judges points	6	5	5	3	2	1
public points		1			2	4
new rank		$(6,1)$			$(4,2)$	$(5,4)$

So, in these six cases, the alternative rule would have resulted in H and A being chosen, as the judges wished. The same conclusion holds for the six assignments with abc= 521. It follows that the judges would have been more likely to prevail if their scores had been converted into points in the more usual way.

December 18th, 2015 | LSE comment | 2 Comments