The Fairest Voting System Is One Where No One Actually Votes!

Wagner L. Truppel

June 1, 2007

The one-dimensional case

Consider a situation where N *players* have to select one of M possible *alternatives*. Imagine further that each of these alternatives, A_m ($1 \leq m \leq M$), is assigned a *score* $s(A_m)$ in a **linear** scale, say, from 0 to 1, $0 \leq s(A_m) \leq 1$.

If each player P_n ($1 \leq n \leq N$) has a preference in the same scale, p_n, then we can define a measure of *unhappiness* for each player — when the result of the selection process is r — by

$$u(P_n, r) \equiv |p_n - s(r)| \, .$$

The *total* unhappiness resulting from selecting alternative r is, then, the sum of the individual unhappiness measures:

$$u(r) \equiv \sum_{n=1}^{N} u(P_n, r) = \sum_{n=1}^{N} |p_n - s(r)| \, .$$

Clearly, the fairest selection process is one which minimizes the total unhappiness measure, that is, the optimal selection — however it is accomplished — is the alternative A_k such that

$$u(A_k) \leq u(A_m), \qquad 1 \leq m \leq M \, .$$

An approximately equivalent definition of unhappiness, which behaves the same way but which is a differentiable function of its arguments, is

$$u(P_n, r) \equiv [\, p_n - s(r) \,]^2 \, ,$$

in which case the total unhappiness is given by

$$u(r) \equiv \sum_{n=1}^{N} u(P_n, r) = \sum_{n=1}^{N} [\, p_n - s(r) \,]^2 \, .$$

Using this new definition, it's easy to predict the score of the selection that minimizes the total measure of unhappiness. If r^* is the selection that minimizes the total measure of unhappiness, which may or may not actually be one of the alternatives, then

$$\frac{du(r)}{dr}\Big|_{r=r^*} = -2\frac{ds(r)}{dr}\Big|_{r=r^*} \sum_{n=1}^{N} [\, p_n - s(r^*)\,] = 0 \quad \Rightarrow \quad s(r^*) = \frac{1}{N}\sum_{n=1}^{N} p_n \,,$$

the simple average preference among all players. That this does provide a minimum is seen from looking at the second derivative,

$$\frac{d^2u(r)}{dr^2} = -2\frac{d^2s(r)}{dr^2} \sum_{n=1}^{N} [\, p_n - s(r)\,] + 2N\, [\, \frac{ds(r)}{dr}\,]^2 \,.$$

Since the first term vanishes when computed at $r = r^*$, the second derivative is clearly positive when computed at $r = r^*$, indicating a minimum:

$$\frac{d^2u(r)}{dr^2}\Big|_{r=r^*} = 2N\, \Big[\, \frac{ds(r)}{dr}\Big|_{r=r^*} \Big]^2 > 0 \,.$$

The alternative which minimizes the total measure of unhappiness, among those actually available, is then that whose score is closest to $s(r^*)$.

This allows a selection to be made without any actual voting by any of the players. All that's required is the knowledge of each player's preference, in the form of his or her score.

The results above, of course, assume that all players are sincere in their score assignments. What if one or more players decide to be insincere, in order to influence the resulting selection? If each player claims to have a different preference, p_n', then the selected alternative r'^* will have a score given by

$$s(r'^*) = \frac{1}{N}\sum_{n=1}^{N} p_n' \,.$$

This differs from the sincere selection by the amount

$$s(r'^*) - s(r^*) = \frac{1}{N}\sum_{n=1}^{N} (p_n' - p_n) \,.$$

In particular, if only one player, say, player a, chooses a score for himself that's different from his true score, the difference between the scores of the resulting selections is

$$s(r'^*) - s(r^*) = \frac{1}{N}(p_a' - p_a) \,.$$

Given that the score scale ranges from 0 to 1, the largest bias that a single player can introduce into the selection mechanism is, then, merely $1/N$.

The multi-dimensional case

Now imagine a similar situation, but where each alternative is a point in a D-dimensional space, that is, where each alternative is assigned D separate scores. In this case, each player also has to position himself or herself in that space, and thus also assigns to himself or herself D different preferences.

The unhappiness measure of a single player, P_n, when the selected result is r, is still a function of the distance between the player's preference and the resulting selection's score, and the most natural and mathematically convenient way to define it is simply to use the Euclidean distance between the points representing the player and the resulting selection, in the D-dimensional space of the scores:

$$u(P_n, r) \equiv \sum_{d=1}^{D} [\, p_{nd} - s_d(r)\,]^2 \,,$$

where p_{nd} is the score or preference player P_n assigns himself or herself in the d-th axis of the score space, and $s_d(r)$ is the score in the d-th axis associated with the selected alternative r.

As before, the total measure of unhappiness is the sum of the individual measures, over all players:

$$u(r) \equiv \sum_{n=1}^{N} u(P_n, r) = \sum_{n=1}^{N} \sum_{d=1}^{D} [\, p_{nd} - s_d(r)\,]^2 \,,$$

The first derivative is now

$$\frac{du(r)}{dr} = -2 \sum_{n=1}^{N} \sum_{d=1}^{D} [\, p_{nd} - s_d(r)\,] \,\frac{ds_d(r)}{dr} = -2 \sum_{d=1}^{D} \Big[\, \sum_{n=1}^{N} p_{nd} - N s_d(r) \Big] \,\frac{ds_d(r)}{dr} \,.$$

The requirement that it vanishes at the position of the minimum has a solution similar to the one-dimensional case, namely, the average of all players' preferences, in each score axis:

$$s_d(r^*) = \frac{1}{N} \sum_{n=1}^{N} p_{nd} \,.$$

It is possible to show that this solution does correspond to the *global* minimum. In fact, for Euclidean distances, there is always *only one* minimum.

The result of the one-dimensional case is, therefore, applicable in any number of dimensions: the score of the optimal selection is the simple average of all the players preferences, along each of the score axes.

It's easy to see that the largest bias a single player could introduce in the score of the selected alternative is $1/N$ per axis. ■