Random walks and voting theory

Nicolas Vieille*

May 2002

Abstract

Voters' preferences depend on the available information. Following Case-Based Decision Theory, we assume that this information is processed additively. We prove that the collective preferences deduced from the individual ones through majority vote cannot be arbitrary, as soon as a winning quota is required. The proof is based on a new result on random walks.

*Département Finance et Economie, HEC, 1, rue de la Libération, 78 351 Jouy-en-Josas, France. e-mail: vieille@hec.fr

1 Introduction

The purpose of this note is to point out a somewhat unexpected relation between random walks and some aspects of voting theory. Our model is the following. Voters' preferences are influenced by the information that gets available prior to the election day. This information may take the form of a collection of facts, arguments or cases that are brought forth. We here assume that this information is processed in an additive manner by each voter. To be specific, for each voter i, each case c and each alternative x, there is a number $w_i(x,c)$ that measures the support that case c lends to x: if the available information is the set D, voter i prefers alternative x to alternative y iff

$$\sum_{c\in D} w_i(x,c) > \sum_{c\in C} w_i(y,c).$$

This class of decision procedures has been introduced into economics and axiomatized by Gilboa and Schmeidler (2001) following some work in AI and psychology.

Our interest lies in understanding whether these assumptions on individual preferences can be tested at the aggregate level: does the collective preference relation derived from individual preferences through qualified majority voting exhibit any specific pattern? or is it purely arbitrary? In other words, knowing how the population of voters reacts facing some evidence, can we infer anything regarding its reaction to additionnal information ?

Let X be the (finite) set of alternatives and $\mathcal{C}$ the (finite) universe of all possible cases, and let $q \geq 1/2$ be the quota. A society π is defined by a (finite) set N of voters and by the vector $(w_i(x,c))_{i\in N, c\in C, x\in X}$ of individual preferences. Collective preferences are defined using majority vote with quota q: given a subset $D \subset \mathcal{C}$ and two alternatives x and y, x (resp. y) is weakly preferred to y (resp. to x) by society if no more than q of the voters strictly prefers y to x (resp. x to y) given D. Thus, a given society induces a map that associates to any non-empty subset D of $\mathcal{C}$ the above reflexive binary relation over X.

The case $q = 1/2$ is analysed in Gilboa and Vieille (2002). It is shown that majority voting may be unpredictable in the sense that any such map may be the outcome of majority voting. We here prove that this result is specific to the case $q = 1/2$ and that, for higher quotas, binary relations arising from majority voting have some structure. Put somewhat loosely, collective preferences between any two alternatives cannot be reversed by adding an extra piece of information, provided much information is already available.

The paper is organized as follows. Section 2 is devoted to the model and to the statement of the result. The proof of the main result is given in Section

3, which emphasizes the connection to random walk theory and sampling. The proof is based on a new result on random walks, the proof of which is given in Section 4.

2 Model and Result

2.1 Individual and collective preferences

We let two finite sets X and $\mathcal{C}$ be given. The set X is the set of alternatives that are being considered. Most of the paper will focus on the case $|X| = 2$. The set $\mathcal{C}$ is the universe of all facts (or cases, stories, etc.) that may be (publicly) known. A population π is described by the set N of voters, together with the individual preferences over X. Preferences of a voter $i \in N$ over X depends on the available information, that may be any non-empty subset of $\mathcal{C}$. We make the assumption that information is processed additively. Specifically, the preferences of voter i are characterized by a function $w_i : X \times \mathcal{C} \to \mathbb{R}$, with the interpretation that, given information $D \subset \mathcal{C}$, voter i prefers alternative x over alternative y if $\sum_{c \in D} w_i(x, c) > \sum_{c \in D} w_i(y, c)$. A population π is described by the set N of individuals, together with the collection $(w_i)_{i \in N}$ of preferences.

Let $\mathcal{P}_*(\mathcal{C})$ be the set of non-empty subsets of $\mathcal{C}$ and $\mathcal{R}$ be the set of (complete) reflexive binary relations over X. Let a population $\pi = (N, (w_i))$ be given. Individual preferences are aggregated using majority voting with quota q. Fix $D \in \mathcal{P}_*(\mathcal{C})$ and let x, y be any two alternatives. Alternative x is preferred to alternative y given D, written $x \succsim_D^\pi y$, if

$$\left| \left\{ i \in N : \sum_{c \in \mathcal{C}} w_i(x, c) > \sum_{c \in \mathcal{C}} w_i(x, c) \right\} \right| \geq qN. \tag{1}$$

Thus, π induces a map $M_\pi : \mathcal{P}_*(\mathcal{C}) \to \mathcal{R}$ where $M_\pi(D) = \succsim_D^\pi$.

2.2 Results

The imposed structure on individual preferences implies much correlation between the preferences of voter i given various informations. As an illustration, the following holds, given two disjoints sets D_1 and D_2. If voter i prefers x to y given either D_1 or D_2, he still prefers x to y given both D_1 and D_2. Our main focus is in understanding whether any correlation still exists at the collective level.

The following result, due to Gilboa and Vieille (2002) shows that this correlation may entirely vanish at the aggregate level, if $q = \frac{1}{2}$. When $\mathcal{C}$ is a singleton, the statement below is a slight generalization of a result due to McGarvey (1953).

Theorem 1 *Let $q = 1/2$. For each $M : \mathcal{P}_*(\mathcal{C}) \to \mathcal{R}$, there exists a population π such that $M_\pi = M$.*

Our main point is to show that Theorem 1 does not extend to $q > \frac{1}{2}$. We limit ourselves to two alternatives x and y. This case allows for several simplifications. First, individual preferences $w_i : X \times \mathcal{C} \to R$ are equivalently described by $w_i(c) := w_i(x, c) - w_i(y, c)$, so that i prefers x to y given D if $\sum_{c \in D} w_i(c) > 0$. Next, an element $\succsim$ of $\mathcal{R}$ can be identified to the set of winning alternatives, i.e. to x, to y or to $\{x, y\}$ if respectively $x \succ y$, $y \succ x$ or neither of the two holds.

Theorem 2 *Let $q > 1/2$. Let $M : \mathcal{P}_*(\mathcal{C}) \to \mathcal{R}$ be defined by $M(C) = x$ if $|C|$ is even and $M(C) = y$ if $|C|$ is odd. If $|\mathcal{C}|$ is large enough, there is no population π such that $M_\pi = M$.*

This result may be paraphrased by saying that collective preferences M_π, whenever anonymous, cannot be easily reversed as soon as much information is already available. The statement may be strengthened in many respects, as will be clear from the proof in Section 3. The present one has the merit of simplicity.

The formula (1) need not always be the sensible way to define society's preferences in the presence of quotas. In many instances, e.g. when constitutional amendments are being considered, one of the alternatives if a *statu quo* while the other is the reform being considered for implementation, so that the two alternatives do not have symmetric roles. The *statu quo* is preferred to the reform iff the reform fails to attract a fraction of at least q of the voters. It is shown in Gilboa and Vieille (2002) that Theorem 1 extends to arbitrary $q \geq 1/2$ with this modified definition of collective preferences.

The proof of Theorem 2 is based on an auxiliary result on random walks that we present next. Recall that a random walk is a sequence $(S_n)_n$ of random variables with iid one-step increments $S_{n+1} - S_n, n \in \mathbf{N}$.

Proposition 3 *For every $\varepsilon > 0$, there exists N_0 such that the following holds. For each random walk (S_n), and each $N \geq N_0$, one has*

$$\frac{1}{N} \sum_{m=N^2}^{N^2+N-1} \mathbf{P}(S_{2m} \geq 0 > S_{2m+1}) < \varepsilon. \tag{2}$$

We comment briefly on this result. Note first that no integrability condition (mean, variance) is assumed on (S_n). The values N and N^2 that appear in (2) are somewhat arbitrary, what matters is that the number of terms in the summation be small compared to the index of the first term.

We next explore the relationship of Proposition 3 to the Central Limit Theorem. Let $(S_n)_n$ be a random walk, such that the increment $X_n :=$

$S_n - S_{n-1}$ ($S_0 = 0$) has a variance. By definition, the sequence $(X_n)_{n\geq 1}$ is iid. Assume for convenience that $\mathsf{E}[X_n^2] = 1$ and set $m := \mathsf{E}[X_n]$. By the central limit theorem, $(S_n - nm/\sqrt{n})_n$ converges in law to the standard normal distribution $N(0, 1)$.

If, say, $m > 0$, (S_n) converges in probability to $+\infty$, hence $\mathsf{P}(S_n \geq 0 > S_{n+1}) = \mathsf{P}(S_n \geq 0$ and $X_{n+1} < -S_n)$ converges to zero. If now $m = 0$, S_n, when positive, is typically of the order $\sqrt{n}$. In particular, $\mathsf{P}(S_n \geq n^{1/3} | S_n \geq 0)$ converges to one as n goes to infinity. Since $\mathsf{P}(X_{n+1} < -n^{1/3})$ converges to zero, $\mathsf{P}(S_n \geq 0$ and $X_{n+1} < -S_n)$ converges to zero. Thus, for each integrable random walk, $\mathsf{P}(S_n \geq 0 > S_{n+1})$ converges to zero as n goes to infinity.

However, the convergence is not uniform, as shown by the example below. Fix $N \in \mathsf{N}$ and let $(X_n)_n$ be an iid sequence with $\mathsf{P}_N(X_n = N) = \frac{1}{N}$ and $\mathsf{P}_N(X_n = -1) = 1 - \frac{1}{N}$. Plainly, the event $\{S_N \geq 0 > S_{N+1}\}$ coincides with the event $\{S_N = 0, X_{N+1} = -1\} = \{X_{N+1} = -1, X_n = N$ for exactly one $n \in \{1, ..., N\}\}$. Therefore,

$$\mathsf{P}_N(S_N \geq 0 > S_{N+1}) = (1 - \frac{1}{N})^{N+1}$$

which converges to $\frac{1}{e}$ as N goes to infinity.

This example shows that $\mathsf{P}(S_n \geq 0 > S_{n+1})$ does not converge to zero, uniformly w.r.t. the random walk. In that respect, the statement in Proposition 3 is optimal.

2.3 An example

We here partially analyze an incomplete example, in order to illustrate why considering large amounts of information is helpful. Let $M : \mathcal{P}_*(\mathcal{C}) \to \mathcal{R}$ be given, such that $M(C) = x$ (resp. $M(C) = y$) whenever $|C| = 1$ (resp. $|C| = 2$). Let $\pi = (N, (w_i))$ be a (hypothesized) population such that $M_\pi(C) = M(C)$ for $|C| \leq 2$.

For $C \in \mathcal{P}_*(\mathcal{C})$, let $N(C)$ be the set of voters $i \in N$ that strictly prefer x to y given C. Since individual preferences are additive in cases, $N(\{c_1\}) \cap N(\{c_2\}) \subseteq N(\{c_1, c_2\})$. Thus, both $N(\{c_1\})$ and $N(\{c_2\})$ contain at least q of the population while $N(\{c_1\}) \cap N(\{c_2\})$ contains at most $1 - q$ of the population. Thus, for each $c_1 \neq c_2$, the two sets $N(\{c_1\})$ and $N(\{c_2\})$ should be fairly different. If $\mathcal{C}$ is large, this means that the set N contains many large fairly different subsets of voters. Specifically, identify the different voters to mutually disjoint subintervals of length $\frac{1}{|N|}$ of $[0, 1]$. Then, all sets $(N(\{c\}))_{c\in\mathcal{C}}$ are subsets of $[0, 1]$ of length exceeding q, while the length of any pairwise intersection does not exceed $1 - q$. Despite sounding problematic, this is feasible, provided the quota q is not too large.

Indeed, let $q \in [1/2, 3/5]$, and let $\mathcal{C}$ be fixed. Let the set N of voters be the set of functions $i : \mathcal{C} \to \{0, 1, 2, 3, 4\}$. The preferences of voter i are given by $w_i(c) = 1$ if $i(c) \in \{0, 1, 2\}$ and $w_i(c) = -3$ if $i(c) \in \{3, 4\}$. The set $N(\{c\})$ coincides with the set of functions i such that $i(c) \in \{0, 1, 2\}$, hence contains $3/5$ of N, while, for $c_1 \neq c_2$, the set $N(\{c_1, c_2\})$ reduces to $N(\{c_1\}) \cap N(\{c_2\})$ which contains $9/25 < 2/5$ of N.

Note that in this example, preferences of voters are not correlated across cases: the preference given a set C of cases of a (randomly selected) voter gives no information on her preferences given some case $c' \notin C$.

The construction here relies on the idea that it is easier to reverse a voter's preferences by adding a single case if little information has been accumulated so far. A contrario, the basic insight of the proof to come is that this becomes very difficult if a substantial amount of information has been piled.

3 Votes with quota

Let $\varepsilon \in (0, \frac{2q-1}{4})$. We let N_0 be given by Proposition 3.

3.1 Cyclic populations and random walks

Definition 4 A population (N, w) is said to be *cyclic* if there is a one-to-one function $v : \mathcal{C} \to \mathrm{R}$ such that given any permutation σ of C, there is a unique

$i \in N$ such that $w_i = v \circ \sigma$.

Plainly, if (N, w) is a cyclic population, then $|N| = |\mathcal{C}|!$ .We label cases from 1 to $|\mathcal{C}|$. Let ι be a randomly selected voter. For $n \leq |\mathcal{C}|$, we set $X_n := w_\iota(n)$. Plainly, the random vector $(X_1, ..., X_{|\mathcal{C}|})$ is a randomly ordered list of the elements of the set $v(\mathcal{C})$.

Lemma 5 Let an integer $K \geq N_0$ and a set $\mathcal{C}$ be given such that $|\mathcal{C}| \geq \left(2K^2 + 2K - 1\right)^2 / 2\varepsilon$. Let π be a cyclic population. One has

$$\frac{1}{K} \sum_{m=K^2}^{K^2+K-1} \mathsf{P}(X_1 + ... + X_{2m} \geq 0 > X_1 + ... + X_{2m+1}) < 2q - 1.$$

Proof. We rephrase the problem using the following auxiliary experiment. Sample $|\mathcal{C}|$ elements $\mathsf{c}_1, ..., \mathsf{c}_{|\mathcal{C}|}$ from the set $\mathcal{C}$, and let $Y_l := w_1(\mathsf{c}_l)$ be the weight assigned by the first voter given the lth sampled item.

Let Q denote the law of $\left(\mathsf{c}_1, ..., \mathsf{c}_{|\mathcal{C}|}\right)$ when sampling is done without replacement. In that case, $\left(\mathsf{c}_1, ..., \mathsf{c}_{|\mathcal{C}|}\right)$ is a random permutation of the elements of $\mathcal{C}$. Hence, the vector $(Y_1, ..., Y_{|\mathcal{C}|})$ is a randomly ordered list of the elements of $v(\mathcal{C})$, i.e., the law of $(Y_1, ..., Y_{|\mathcal{C}|})$ coincides with the law of $(X_1, ..., X_{|\mathcal{C}|})$. Thus, for each m,

$$\mathsf{P}(X_1 + ... + X_{2m} \geq 0 > X_1 + ... + X_{2m+1}) = \mathsf{Q}(Y_1 + ... + Y_{2m} \geq 0 > Y_1 + ... + Y_{2m+1}) \tag{3}$$

Let Q_1 denote the law of $\left(\mathsf{c}_1, ..., \mathsf{c}_{|\mathcal{C}|}\right)$ when sampling is done with replacement. Plainly, the variables $(Y_1, ..., Y_{|\mathcal{C}|})$ are iid in that case. By the choice of K,

$$\frac{1}{K} \sum_{m=K^2}^{K^2+K-1} \mathsf{Q}_1(Y_1 + ... + Y_{2m} \geq 0 > Y_1 + ... + Y_{2m+1}) < \varepsilon. \tag{4}$$

To conclude, we prove that the laws of $(Y_1, ..., Y_{N_1})$ under the two distributions Q and Q_1 are close, where $N_1 = 2K^2 + 2K - 1$.

Sampling without replacement may be viewed as sampling with replacement, conditional on sampled items being all distinct:

$$\mathsf{Q}((\mathsf{c}_1, ..., \mathsf{c}_{N_1}) = (c_1, ..., c_{N_1})) = \mathsf{Q}_1((\mathsf{c}_1, ..., \mathsf{c}_{N_1}) = (c_1, ..., c_{N_1}) | \mathsf{c}_i \neq \mathsf{c}_j \text{ for each } i \neq j).$$

For each pair (i, j) with $i \neq j$, $\mathsf{Q}_1(\mathsf{c}_i = \mathsf{c}_j) = \frac{1}{|C|}$. Therefore, for any event A that depends only upon $(\mathsf{c}_1, ..., \mathsf{c}_{N_1})$, one has

$$\begin{aligned} |\mathsf{Q}(A) - \mathsf{Q}_1(A)| &\leq \mathsf{Q}_1(\mathsf{c}_i = \mathsf{c}_j \text{ for some } i \neq j \leq N_1) \\ &\leq \frac{N_1(N_1 - 1)}{2} \times \frac{1}{|\mathcal{C}|} \leq \varepsilon \end{aligned} \tag{5}$$

The result follows, by (3), (4) and (5). ■

3.2 Proof of Theorem 2

We now prove Theorem 2. We argue by contradiction, and let $\pi = (N, (w_i)_{i \in N})$ be an hypothesized population such that $M_\pi = M$. We proceed in three steps. We first prove that w_i may be assumed to be one-to-one for each $i \in N$. Next, we enlarge the population π to obtain a population π' that is a disjoint union of cyclic populations. We conclude by using Lemma 5.

Step 1 For each $i \in N$, let $w'_i : \mathcal{C} \to R$ be a one-to-one function such that $\sup_c |w_i(c) - w'_i(c)| < \frac{1}{|C|} \min \left\{ \left| \sum_{i \in C} w_i \right| \right\}$, where the infimum is taken over all C such that $\sum_{i \in C} w_i \neq 0$. Let $\pi' = (N, (w'_i))$ be the population obtained by changing the weights to w'_i. Plainly, $\sum_{i \in C} w'_i(c) > 0$ (resp. < 0) whenever $\sum_{i \in C} w_i(c) > 0$ (resp. < 0). Therefore, for each D.

$$x \succsim^{\pi}_{D} y \ \text{ implies } x \succsim^{\pi'}_{D} y.$$

Thus, the two binary relations $\succsim^{\pi}_{D}$ and $\succsim^{\pi'}_{D}$ coincide for each D. We assume below that w_i is one-to-one for each $i \in N$.

Step 2 Set $\overline{N} := N \times \Sigma$, where Σ is the set of permutations of $\mathcal{C}$. For $(i, \sigma) \in \overline{N}$, we set $\overline{w}_{(i,\sigma)} := w_i \circ \sigma$. In other words, the population $\overline{\pi} = (\overline{N}, \overline{w})$ is obtained from π by adding to each voter $i \in N$ all permutations of w_i. Since $M(C) = M_\pi(C)$ depends only on $|C|$, one has $M_{\overline{\pi}} = M$. Note that, for each $i \in N$, the subpopulation of $\overline{\pi}$ with set of voters $\{i\} \times \mathsf{P}$ is a cyclic population.

Step 3: Conclusion

Let $(\iota, \sigma) \in \overline{N}$ be a randomly selected voter. For $n \leq |\mathcal{C}|$, set $X_n := \overline{w}_{(\iota,\sigma)}(n)$. For each fixed $i \in N$, one has, by Lemma 5

$$\frac{1}{K} \sum_{m=K^2}^{K^2+K-1} \mathsf{P}(X_1 + ... + X_{2m} \geq 0 > X_1 + ... + X_{2m+1} | \iota = i) < 2q - 1. \quad (6)$$

Multiplying (6) by $\mathsf{P}(\iota = i)$ and summing over i yields

$$\frac{1}{K} \sum_{m=K^2}^{K^2+K-1} \mathsf{P}(X_1 + ... + X_{2m} \geq 0 > X_1 + ... + X_{2m+1}) < 2q - 1. \quad (7)$$

On the other hand, since $M_{\overline{\pi}} = M$, one has $\mathsf{P}(X_1 + ... + X_{2m} \geq 0) \geq q$ and $\mathsf{P}(X_1 + ... + X_{2m+1} < 0) \geq q$ for each m, hence $\mathsf{P}(X_1 + ... + X_{2m} \geq 0 > X_1 + ... + X_{2m+1}) \geq 2q - 1$- a contradiction to (7).

4 On random walks

Recall that a random walk is a sequence (S_n) whose increments $(S_{n+1} - S_n)_{n \geq 0}$ are iid. We set $X_n := S_n - S_{n-1}$; S_n is usually seen as the position at time n of a particle that moves from date $n-1$ and n of an amount of X_n. The law of the random walk (S_n) is determined by the law of X_1.

We shall use an alternative construction of the random walk. Define μ^+ (resp. $-\mu^-$) to be the law of X_1 conditioned on $X_1 \geq 0$ (resp. $X_1 < 0$) and set $p = \mathsf{P}(X_1 \geq 0), q = \mathsf{P}(X_1 < 0)$. Let (Y_n) (resp. $(Z_n)_n$) be a sequence of iid variables with law μ^+ and μ^- respectively. Beware that both Y_n and Z_n are nonnegative variables. Set $S_k^+ = \sum_{l=1}^k Y_l$, and $S_k^- = \sum_{l=1}^k Z_l$. Then the sequence (S_n) can be described as follows. At time n: choose a direction $d_n \in \{-1, +1\}$ with respective probabilities p and q; if $d_n = 1$, move up by an amount Y_n; if $d_n = -1$, move downwards by the amount Z_n. Thus,

$$S_n = \sum_{l=1} \{Y_l \mathbf{1}_{d_l=1} + Z_l \mathbf{1}_{d_l=-1}\}.$$

Let $\kappa_n = |\{l \leq n : d_l = +1\}|$ be the number of upward moves up to n. Plainly, κ_n has a binomial distribution $B(n, p)$, and S_n is distributed as $S_{\kappa_n}^+ - S_{n-\kappa_n}^-$. Moreover, ,

$$\begin{aligned}
\mathsf{P}(S_n \geq 0 > S_{n+1}) &= \mathsf{P}(S_n \geq 0, d_{n+1} = -1, S_n + Z_{n+1} < 0) \\
&= q\mathsf{P}(S_n \geq 0 > S_n + Z_{n+1}) \\
&= q\mathsf{P}(S_\kappa^+ - S_{n-\kappa}^- \geq 0 > S_\kappa^+ - S_{n+1-\kappa}^-) \\
&= q\mathsf{P}(S_{n-\kappa}^- \leq S_\kappa^+ < S_{n+1-\kappa}^-). \qquad (8)
\end{aligned}$$

Lemma 6 One has $\mathsf{P}(S_n \geq 0 > S_{n+1}) \leq q \sup_{k\in\{1,...,n\}} C_n^k p^k q^{n-k}$.

Proof. We shall prove that, for each $(y_1, ..., y_n) \in \mathsf{R}_+^n$,

$$\mathsf{P}(S^-_{n-\kappa_n} \leq S^+_{\kappa_n} < S^-_{n+1-\kappa_n} | (Y_1, ..., Y_n) = (y_1, ..., y_n)) \leq \sup_{k\in\{1,...,n\}} C_n^k p^k q^{n-k}.$$

The conclusion will follow by integration over the possible values of $(Y_1, ..., Y_n)$, using (8). Since both sequences (S_l^-) and (S_l^+) are non-decreasing, the sets (indexed by k) $\left\{S^-_{n-k} \leq S^+_k < S^-_{n+1-k}\right\}$ are disjoint, thus,

$$\sum_{k=1} \mathsf{P}(S^-_{n-k} \leq y_1 + ... + y_k < S^-_{n+1-k}) \leq 1.$$

Next,

$$\begin{aligned}
&\mathsf{P}(S^-_{n-\kappa_n} \leq S^+_{\kappa_n} < S^-_{n+1-\kappa_n} | (y_1, ..., y_n)) \\
&\quad = \sum_{k=1} \mathsf{P}(\kappa_n = k) \mathsf{P}(S^-_{n-k} \leq y_1 + ... + y_k < S^-_{n+1-k}) \\
&\quad \leq \sup_{k\in\{1,...,n\}} \mathsf{P}(\kappa_n = k) = \sup_{k\in\{1,...,n\}} C_n^k p^k q^{n-k}.
\end{aligned}$$

■

We proceed to the proof of Proposition 3. The argument goes as follows. Let M be large, and (S_n) be a given random walk. We prove that if p $(= \mathsf{P}(S_1 \geq 0)$ is of the order of $\frac{1}{M}$, the events $\{S_{2m} \geq 0 > S_{2m+1}\}$, where m

ranges over $M^2, ..., M^2 + M - 1$, are approximately disjoint. If the number of these events exceeds $\frac{1}{\varepsilon}$, their average probability can not exceed ε. On the other hand, if p is much greater than $\frac{1}{M}$, the probability of each of the events $\{S_{2m} \geq 0 > S_{2m+1}\}$ is close to zero, hence also the average of these. We now provide details.

By Stirling's formula, there is $\alpha > 1$ such that

$$\frac{1}{\alpha} \leq \frac{n!}{(\frac{n}{e})^n \sqrt{2\pi n}} \leq \alpha \text{ for each } n, \tag{9}$$

(α may be chosen to be small, e.g., $\alpha = 2$ suffices).

Let $\varepsilon > 0$, and $M \geq \max(\frac{1}{\varepsilon}, \frac{2}{\varepsilon}(\frac{\alpha^3 e^2}{\varepsilon\sqrt{\pi}} + 1))$ be given. Let $(S_n)_{n \geq 0}$ be an arbitrary random walk. We prove below that (2) holds.

Case 1: $q \leq \varepsilon$

By Lemma 6, $\mathsf{P}(S_{2m} \geq 0 > S_{2m+1}) \leq q$ for each m. Hence the result is obvious in that case.

Case 2: $M^2 p \geq \frac{\alpha^3 e^2}{\varepsilon\sqrt{\pi}} + 1$ and $q \geq \varepsilon$.

Let an even $m \geq 2M^2$ be given. By Lemma 6,

$$\mathsf{P} := \mathsf{P}(S_m \geq 0 > S_{m+1}) \leq q \sup_k C_m^k p^k q^{m-k}. \tag{10}$$

By Feller (section VI.3), the values $C_m^k p^k q^{m+1-k}$ first increase with k then decrease, the maximum being reached for k such that $(m+1)p-1 < k \leq (m+1)p$. Note that $mq < m-k \leq (m+1)q$.

By (9),

$$C_m^k \leq \alpha^3 \frac{m^m}{k^k(m-k)^{m-k}} \sqrt{\frac{m}{2\pi k(m-k)}}.$$

Next, note that

$$\left(\frac{mp}{k}\right)^k = \left(1+\frac{mp-k}{k}\right)^k \leq \exp\left\{k\ln(1+\frac{1}{k})\right\} \leq e.$$

For the same reason, $\left(\frac{mq}{m-k}\right)^{m-k} \leq e$. Hence, by (10), $\mathsf{P} \leq \alpha^3 e^2 \sqrt{\frac{m}{2\pi k(m-k)}}$. Finally, observe that $\sqrt{\frac{m}{k(m-k)}} \leq \sqrt{2}\max\left(\frac{1}{\sqrt{k}}, \frac{1}{\sqrt{m-k}}\right)$. Therefore

$$\mathsf{P} \leq \frac{\alpha^3 e^2}{\sqrt{\pi}} \max\left(\frac{1}{\sqrt{k}}, \frac{1}{\sqrt{m-k}}\right) \leq \varepsilon,$$

where the second inequality follows by the choice of M. The result follows by averaging over m.

Case 3: $Mp \leq \frac{\varepsilon}{2}$.

Plainly, if $S_{2m_1+1} < 0 \leq S_{2m_2}$ for some $m_2 > m_1$, then $d_m = -1$ for some $m > m_1$. Therefore,

$$\mathsf{P}(S_{2m_1+1} < 0, S_{2m_2} \geq 0 \text{ for some } m_2 > m_1) \leq Mp$$

We next apply the inequality

$$\sum_{i=1} \mathsf{P}(A_i) \leq \mathsf{P}(\cup_{i=1}^n A_i) + \sum_{i=1}^{n-1} \mathsf{P}(A_i \cap (\cup_{j=i+1}^n A_j)$$

with the family of events $A_m = \{S_{2m} \geq 0 > S_{2m+1}\}$ to get

$$\sum_{m=M^2}^{M^2+M-1} \mathsf{P}(S_{2m} \geq 0 > S_{2m+1}) \leq 1 + \sum_{m=M^2}^{M^2+M-1} Mp$$
$$\leq 1 + M^2 p.$$

The result follows by dividing by M.

Acknowledgements: I wish to thank Itzhak Gilboa for suggesting the problem, for comments and for the analysis in Section 1.3.

References

[1] Feller, W (1968) An introduction to Probability Theory and its applications, John Wiley and Sons

[2] Gilboa, I and Schmeidler, D. (2001) Case-Based Decision Theory, Cambridge University Press

[3] Gilboa, I. and Vieille, N. (2002) Majority vote following a debate, mimeo

[4] McGarvey, D. C. (1953), "A Theorem on the construction of voting paradoxes", Econometrica, 21: 608-610