Three ways to compute accurately the probability of the referendum paradox

Dominique Lepelley [*], Vincent Merlin [†] et Jean-Louis Rouet [‡]

Abstract

In a recent paper published in *MSS*, Wilson and Pritchard (2007) suggest that the limiting probability of the referendum paradox given in Feix *et al.* (2004) is not correct. Using three distinct analytical and complementary methods, we show in this note that this suggestion is to be rejected.

Keywords : voting paradoxes, probability calculations, IAC hypothesis.
JEL Classification : C9, D72

1 Introduction

A referendum paradox occurs in a two candidate election when a candidate wins in a majority of districts (or states) while she gets less votes than her opponent in the whole country (or federal union). Given that such paradoxical situations have been observed in real political elections -the most famous occurrence being the 2000 election of G.W. Bush-, it is of interest to investigate the theoretical likelihood of this paradox. Some results on this question have been obtained by Feix *et al* (2004) in the simplest case where there are equal size districts. The results provided by these authors are based on various probabilistic models, including the usual Impartial Culture (IC) and the Impartial Anonymous Culture (IAC) models. Using both analytical and simulation-based arguments, Feix *et al* (2004) show, among other results,

[*]CERESUR, Faculté de Droit et de Sciences Economiques, Université de la Réunion, 97715 Saint-Denis cedex 9 France. E-mail: dominique.lepelley@univ-reunion.fr

[†]CREM UMR6211, CNRS and Université de Caen, 19 rue Claude Bloch, 14032 Caen cedex, France. E-mail: vincent.merlin@unicaen.fr

[‡]ISTO UMR6113 and MAPMO UMR6628, CNRS and Université d'Orléans, 1A rue de la Férollerie, 45071 Orléans cedex 2. E-mail: jean-louis.rouet@univ-orleans.fr

that the paradox probability under IAC tends to a limit of around 0.165 when the number of districts goes to infinity. In a paper recently published in this journal, Wilson and Pritchard (2007) offer some further IAC results suggesting that the 0.165 limit could be wrong.

The main purpose of the present note is to provide various and new arguments showing that the limiting probabilities given in Feix *et al* (2004) are, definitely, correct. In what follows, $\mathscr{A}$ and $\mathscr{B}$ denote the two candidates in contention, N denotes the number of districts, n is the number of voters in each district. Let $P_X(n,N)$ be the probability of the referendum paradox under the probability model X for N districts and n voters. In Section 2, we show that the values given by Wilson and Pritchard (2007) for $P_{IAC}(\infty,5)$, $P_{IAC}(\infty,7)$ and $P_{IAC}(\infty,9)$ are erroneous and we provide analytical representations for $P_{IAC}(n,5)$, $P_{IAC}(n,7)$ and $P_{IAC}(n,9)$. Another method for computing exactly the values of $P_X(\infty,N)$ is introduced and applied in Section 3. Finally, we show in Section 4 how the Central Limit Theorem can be used to obtain approximated values of $P_X(\infty,N)$ for sufficiently large N.

2 First approach: Exact results for a small number of districts

We only consider in this section the IAC model. Let n_i denote the number of voters voting for $\mathscr{A}$ in district i. A voting situation is defined as a vector $(n_1,n_2,...,n_N)$. As each n_i is an integer taking its values between 0 and n, there are $(n+1)^N$ voting situations to consider and, under IAC, all these situations are assumed to be equally likely. Wilson and Pritchard (2007) describe the situations giving rise to the referendum paradox by the following inequalities (ties are ignored for simplicity):

$n_i \ > \ n/2$ for $1 \leq i \leq k$ ($\mathscr{A}$ wins k districts),

$0 \leq n_i < n/2$ for $k+1 \leq i \leq N$ ($\mathscr{B}$ wins $N-k$ districts),

$\Sigma_i n_i < Nn/2$ ($\mathscr{B}$ wins overall)

for $[N/2]+1 \leq k \leq N-1$. For each k, the above inequalities define a polytope and measuring the probability paradox under IAC is tantamount to count the number of lattice points in the union of these polytopes (multiplied, for each k, by $2C_N^k$ to account for the symmetries of the problem). As emphasized in Wilson and Pritchard (2007) and Lepelley *et al* (2008), the volume of these polytopes can be represented by Ehrhart polynomials and there exist efficient algorithms to compute these volumes, at least when the number of variables (here the number of n_i's) is not too high. The problem with the results given by Wilson and Pritchard (2007) for the referendum paradox does not come from the algorithm they used but from the set of inequalities they consider: they omit to make precise in the first equations ($n_i \ > \ n/2$ for $1 \leq i \leq k$) that the n_i's have to be lower than or equal to n. For $N=3$, this omission has no consequence ($n_1 > n/2,\ n_2 > n/2$ and $n_1+n_2+n_3 < 3n/2$ imply together $n_1 \leq n$ and $n_2 \leq n$) and they find the right result. For $N \geq 5$, however, the results they give are wrong since they consider unduly situations with some n_i's higher than n. Using an algorithm due to Barvinok and presented in Lepelley *et al* (2008), we have been able to derive the Ehrhart polynomials giving the desired probabilities as a function of the number n of voters in each state for N=5, 7 and 9.

For n odd:

$$P_{IAC}(n,5) = \frac{5(n-1)(n+3)(11n^2+22n+27)}{384(n+1)^4}, \tag{1}$$

$$P_{IAC}(n,7) = \frac{(n-1)(n+3)(577n^4+2308n^3+4482n^2+4348n+3325)}{3840(n+1)^6}, \tag{2}$$

$$P_{IAC}(n,9) = \frac{(n-1)(n-3)(1589879n^6+9539274n^5+26892941n^4+43976604n^3 +48525617n^2+34536090n+22876875)}{10321920(n+1)^8}. \tag{3}$$

For n even:

$$P_{IAC}(n,5) = \frac{5n(11n^4-48n^3+52n^2-96n+192)}{384(n+1)^5}, \tag{4}$$

$$P_{IAC}(n,7) = \frac{n(n-2)(3462n^5-6061n^4+1528n^3-22004n^2-11920n-78720)}{23040(n+1)^7}, \tag{5}$$

$$P_{IAC}(n,9) = \frac{1589879n^9-5431860n^8+5531736n^7-9932832n^6+7748496n^5-24816960n^4 +35839744n^3-52190208n^2+209018880n+4915680}{10321920(n+1)^9}. \tag{6}$$

We have consequently $P(\infty,5) = 55/384 = 0.1432$, $P(\infty,7) = 577/3840 = 0.1503$, $P(\infty,9) = 1589879/10321920 = 0.1540$ and these values are perfectly consistent with the limiting probability of 0.165 given in Feix *et al.* (2004).

3 Second approach: Exact results for a large number of voters

We propose here a method to compute the paradox probability, for a binary issue vote ($\mathscr{A}$ or $\bar{\mathscr{A}} = \mathscr{B}$), when the number of voters and the number of elections is large. In that limit it is possible to introduce a distribution function $f(p)$ where p is an *a priori* probability of voting for the proposal. Once p is chosen all voters vote according to p. We have $\int_0^1 f(p)\mathrm{d}p = 1$ and if the election model has no bias $\int_0^1 pf(p)\mathrm{d}p = 1/2$. The IAC case corresponds to an uniform distribution on $p \in [0,1]$. Another famous model, the Impatial Culture asymption (IC) assumes that each voter casts his ballot with probability 1/2 . The IC case is obtained for $f(p) = \delta(p-1/2)$ where $\delta(x)$ is the Dirac distribution.

For convenience we will introduce the shifted distribution function $h(x)$, with $x = 2(p-1/2)$ such that for $x > 0$ the result gives alternative $\mathscr{A}$ and alternative $\mathscr{B}$ otherwise.

For N states (or districts), let us compute the repartition $h^N(z_N)$ of $z_N = \sum_{i=1}^N x_i$, where x_i is picked into $h(x_i)$ for all states. Suppose we already have computed the repartition h^k after the vote of states $i = 1$ to k. The repartition of z_{k+1} is given by the following convolution product (noted *)

$$h^{k+1}(z_{k+1}) = (h^k * h)(z_{k+1}) = \int_{-\infty}^{\infty} h^k(z_{k+1} - x_{k+1})h(x_{k+1})\mathrm{d}x_{k+1} \tag{7}$$

Because h^k itself is given by $h^{k-1} * h$, we get

$$h^{k+1} = \underbrace{h * h * \cdots * h}_{k+1\,\text{times}} \tag{8}$$

and finally, for N states

$$h^N = \underbrace{h * h * \cdots * h}_{N\,\text{times}} \tag{9}$$

Nevertheless, we want to be more precise and be able to distinguish the cases for which states vote for $\mathscr{A}$ or $\mathscr{B}$. Let us introduce $h^N_{N_A,N_B}(z_N)$ the distribution of z_N after N_A states for $\mathscr{A}$ and $N_B = N - N_A$ states for $\mathscr{B}$. These distributions are obtained in the following way. Let us define $h_A(x)$ and $h_B(x)$ the distributions of the votes x in a state, knowing that this state is in favor of $\mathscr{A}$ or $\mathscr{B}$ respectively. In the same way as for h^N given equation (9), the result after N_A states vote for $\mathscr{A}$ and $N_B = N - N_A$ states vote for $\mathscr{B}$ reads

$$h^N_{N_A,N_B} = \underbrace{h_A * h_A * \cdots * h_A}_{N_A\,\text{times}}\,\underbrace{h_B * h_B * \cdots * h_B}_{N_B\,\text{times}} \tag{10}$$

Now the paradox probability $P_X(\infty, N)$, using model X, will be given by all situations with $N_A > N_B$ or $N_A < N_B$ while $z_N < 0$ or $z_N > 0$ respectively :

$$
\begin{aligned}
P_X(\infty,N) &= \left[\sum_{N_A=1}^{N/2-1} C_N^{N_A} p_A^{N_A} p_B^{N-N_A} \int_0^{\infty} h_{N_A,N-N_A}^N(x)\mathrm{d}x \right. \\
&\qquad \left. + \sum_{N_A=N/2+1}^{N-1} C_N^{N_A} p_A^{N_A} p_B^{N-N_A} \int_{-\infty}^{0} h_{N_A,N-N_A}^N(x)\mathrm{d}x \right] \qquad (11)
\end{aligned}
$$

where p_A (p_B) is the probability that a state casts its vote for $\mathscr{A}$ ($\mathscr{B}$). The first term of left member computes the probability that less than half of the states vote for $\mathscr{A}$ although his supporters gather a majority of the national votes. Symmetricaly, the second term consider cases where $\mathscr{A}$ wins in a majority of states but does not gather a majority of voters. Formula (11) is quite general and no assumption has been made on $f(p)$, in particular p_A and p_B could not be equal at this stage.

3.1 IAC case

The simplest case is given for the IAC model where $h(x)$ is uniform on $[-1,1]$. Consequently $h_A(x)$ is a uniform distribution on $[0,1]$ and h_B on $[-1,0]$. In addition we have $h_B(x) = h_A(x+1)$, which drives to

$$
(h_A * h_B)(x) = (h_A * h_A)(x+1) \qquad (12)
$$

and more generally

$$
(\underbrace{h_A * h_A * \cdots * h_A}_{N_A\,\text{times}} * \underbrace{h_B * h_B * \cdots * h_B}_{N_B\,\text{times}})(x) = (\underbrace{h_A * h_A * \cdots * h_A}_{N\,\text{times}})(x+N_B) = h_{N,0}^N(x+N_B) \qquad (13)
$$

Consequently, it is enough to consider the case where all the votes are in favour of $\mathscr{A}$ and to shift the result of N_B to obtain the N_A votes for $\mathscr{A}$ and N_B votes for $\mathscr{B}$ case. Because $h_B(x) = h_A(-x)$, it is enough to consider the paradox on $\mathscr{A}$, and equation (11) reads

$$
\begin{aligned}
P_{IAC}(\infty,N) &= 2\left(\frac{1}{2}\right)^N \sum_{N_A=N/2+1}^{N-1} C_N^{N_A} \int_{-\infty}^{0} h_{N,0}^N(x+N_B)\mathrm{d}x \\
&= 2\left(\frac{1}{2}\right)^N \sum_{N_A=N/2+1}^{N-1} C_N^{N_A} \int_{0}^{N_B} h_{N,0}^N(x)\mathrm{d}x \qquad (14)
\end{aligned}
$$

Note first that as soon as $f(p)$ is symmetric respective to $p = 1/2$, $p_A = p_B = 1/2$ and the two terms of equation (11) are equal. Secondly, as the support of $h_{N,0}^N$ is $[0,N]$ for IAC, it is equivalent to integrate $h_{N,0}^{N}(x+N_B)$ in between $-N_B$ and 0 or to perform the integration of $h_{N,0}^{N}(x)$ in between 0 and N_B.

To illustrate formula (13), $h_{3,0}^{3}(x)$, $h_{5,0}^{5}(x)$ and $h_{7,0}^{7}(x)$ are plotted figures 1. On each interval, $h_{N,0}^{N}$ is given by a polynomial of order $N-1$ and as N is increasing, $h_{N,0}^{N}$ progressively tends to a gaussian distribution. This convergence to the gaussian will be explored in detail in section 4. For $N=7$ (figure 1c), the contribution for three states for $\mathscr{B}$ and the four others for $\mathscr{A}$ is given by the area delimited by the curve between 0 and 3, the contribution for two states in favour of $\mathscr{B}$ and the five others for $\mathscr{A}$ is given by the area delimited by the curve between 0 and 2 and so on. As seen figure 1c, the contribution for a single states for $\mathscr{B}$ while the six others are for $\mathscr{A}$ is negligeable. The weight of the different contributions are displayed on table 2.

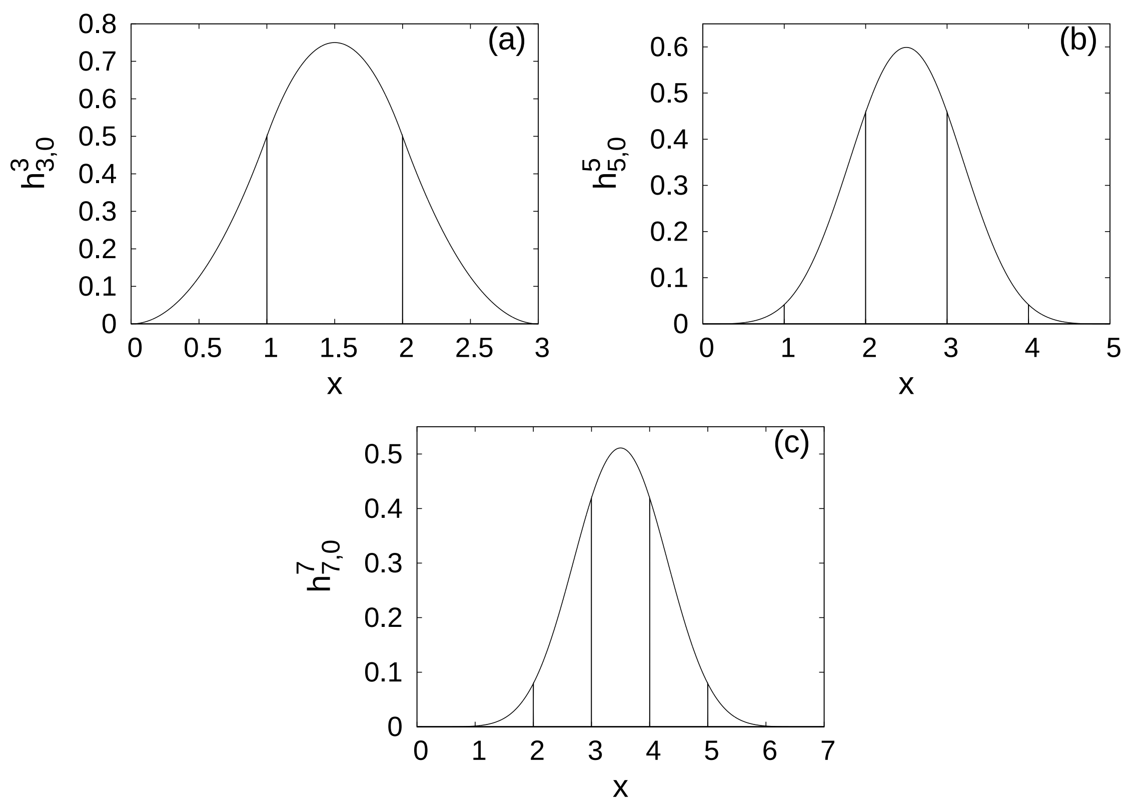

Figure 1: Distribution functions $h_{N,0}^{N}(z_N)$ for (a) $N=3$, (b) $N=5$ and (c) $N=7$.

In the IAC case, the computation of $h_{N,0}^{N}$ is possible but cumbersome as soon as N reaches 5 or more. Nevertheless it is possible to use a symbolic language to obtain the values of $P_{IAC}(\infty,N)$. It has been done here and computed for odd values of N from 3 to 99. The first values are reported table 3, the other are plotted on figure 2 for odd values of N (see IAC ($\alpha=0$) curve). The values computed by Monte Carlo simulations fit nicely with the values

Table 1: IAC case : paradox contribution for the different values of N_A for $N = 7$ states.

N_A	N_B	$C_N^{N_A}$	$\int_0^{N_B} h_{N,0}^N(x)\mathrm{d}x$	paradox contribution
1	6	7	1/5040	1/92160
2	5	21	121/5040	121/30720
3	4	35	82/315	41/576
$P_{IAC}(\infty,7) = 2\times$ sum of last column				577/3840

computed by this method.

Table 2: paradox probability $P_{IAC}(\infty,N)$ as a function of the number of states

N	paradox probability
3	1/8
5	55/384
7	577/3840
9	1589879/10321920
11	20754241/132710400
13	77500241039/77500241039
$\vdots$	$\vdots$
27	$\frac{274943009335014836861029976581279}{1691526988569166523752513536000000}$
$\vdots$	$\vdots$

Remark: the case $N_A = N_B = N/2$ (N even) drives to a second ballot, the probability of which is given by

$$p_{tie}(N) = C_N^{N/2}\left(\frac{1}{2}\right)^N. \qquad (15)$$

3.2 BRIAC case

The Bias Rescaled IAC model (BRIAC) has been introduced by Feix *et al.* [1]. It is defined by an uniform distribution function $f(p)$ for p between $p_{min} \geq 0$ and $p_{max} \leq 1$. The BRIAC is characterized by the parameter $\alpha = E/D$, where $E = (p_{max} + p_{min})/2 - 1/2$ which gives the bias and $D = (p_{max} - p_{min})/2$ which gives the dispersion of p. We will only consider cases where the bias is in favor of $\mathscr{A}$, this other being symmetric. As before, this distribution is

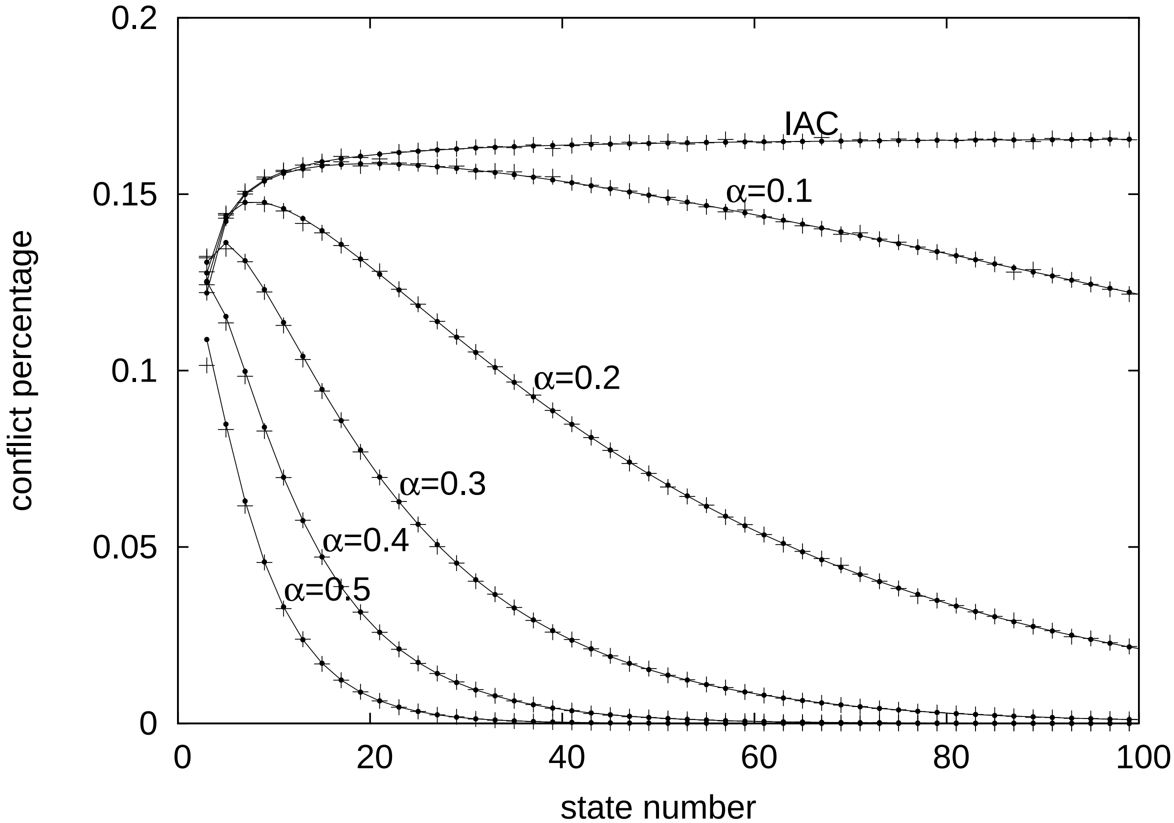

Figure 2: Paradox probabilities as a function of the number of states in the IAC and BRIAC cases. In the IAC case, the values are computed by the convolution method given in section 3.1; in the BRIAC cases, they are computed with the gaussian approximation (method of section 4) for $\alpha = 0.1$ to 0.5. These results are plotted with connected points. The crosses give the results of the Monte Carlo trials (see [1]).

shifted such that for $x < 0$, the state votes for $\mathscr{B}$ while it votes for $\mathscr{A}$ if $x > 0$, with $x = 2(p - 1/2)$. Without loss of generality, we take $x_{max} = 1$ and $-1 < x_{min} < 0$. The paradox probability $P_{BRIAC}(\infty, N)$ could be obtained following the same scenario as before : the distributions $h_A(x)$ and $h_B(x)$ read

$$h_A(x) = \mathrm{rect}(x - \frac{1}{2}), \quad h_B(x) = \frac{1}{-x_{min}} \mathrm{rect}(-\frac{x}{x_{min}} + \frac{1}{2}) \tag{16}$$

with

$$\mathrm{rect}(x) = \begin{cases} 1 & \text{if } |x| < \frac{1}{2} \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

The paradox probability $P_{BRIAC}(\infty, N)$ will be given by equation (11). Because the two distributions $h_A(x)$ and $h_B(x)$ are not symmetric, it is not possible to reduce the computation to the sole determination of $h^N_{N,0}$ as for the IAC case and all $C^{N_A}_N$ cases have to be considered. We give the result for $N=3$ with $\alpha=1/2$. For a $\mathscr{A}\mathscr{A}\mathscr{B}$ result, the contribution is given by

$$\int_{x_{min}}^{0}(h_A*h_A*h_b)(x)\mathrm{d}x=\int_{-1/3}^{0}3(\frac{x^2}{2}+\frac{x}{3}+\frac{1}{18})\mathrm{d}x=\frac{3}{162} \tag{18}$$

and a $\mathscr{B}\mathscr{B}\mathscr{A}$ result, the contribution is given by

$$\int_{0}^{x_{max}}(h_B*h_B*h_A)(x)\mathrm{d}x=\int_{0}^{1/3}1\mathrm{d}x+\int_{1/3}^{2/3}9(-\frac{x^2}{2}+\frac{x}{3}+\frac{1}{18})\mathrm{d}x+\int_{2/3}^{1}9(\frac{x^2}{2}-x+\frac{1}{2})\mathrm{d}x=\frac{2}{3} \tag{19}$$

Taking into account the probability $p_A=3/4$ of $\mathscr{A}$, $p_B=1/4$ of $\mathscr{B}$ and the number of combinaisons equal to C^2_3 in both cases, we have :

$$P_{BRIAC,\alpha=.5}(\infty,N)=C^2_3\left(\frac{3}{4}\right)^2\frac{1}{4}\frac{3}{162}+C^2_3\frac{3}{4}\left(\frac{1}{4}\right)^2\frac{2}{3}=\frac{13}{128}=.1015625 \tag{20}$$

The treatment for the BRIAC case is much more cumbersome than for the IAC case. We will then rely on the next approach to verify the simulation results that were presented in [1].

4 Third approach: Approximative results for a large number of districts

We still consider the limit of an infinite number of voters and an infinite number of elections. As before let us call $h_A(x)$ and $h_B(x)$ the repartitions of the votes in a state that favors $\mathscr{A}$ and $\mathscr{B}$ respectively. Let x_i be random variables with mean m_i and variance σ_i. Then, using the Central Limit Theorem, the distribution of $z_N=\sum_{i=1}^N x_i$ can be approximated by a normal law with mean $m=\sum_{i=1}^N m_i$ and variance $\sigma^2=\sum_{i=1}^N \sigma_i^2$ if N is large enough. Then for N_A states in favor of $\mathscr{A}$ and N_B states in favor of $\mathscr{B}$, in this limit, we have

$$h^N_{N_A,N_B}(z_N)=\frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}\exp(-\frac{(z_N-m)^2}{2\sigma^2}) \tag{21}$$

with

$$m=N_Am_A+N_Bm_B \tag{22}$$

$$\sigma^2=N_A\sigma_A^2+N_B\sigma_B^2 \tag{23}$$

where m_A, σ_A and m_B, σ_B are the mean value and standard deviation of distribution h_A and h_B respectively. Once these expressions are obtained, we can compute the approximation of the paradox probability $\tilde{P}_X(\infty,N)$ which is given by :

$$
\begin{aligned}
\tilde{P}_X(\infty,N) &= \sum_{N_A=1}^{N/2-1} C_N^{N_A} p_A^{N_A} p_B^{N-N_A} \int_0^{\infty} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) \mathrm{d}x \\
&\quad + \sum_{N_A=N/2+1}^{N-1} C_N^{N_A} p_A^{N_A} p_B^{N-N_A} \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) \mathrm{d}x \qquad (24) \\
&= \frac{1}{2} \sum_{N_A=1}^{N/2-1} C_N^{N_A} p_A^{N_A} p_B^{N-N_A} \mathrm{erfc}\left(-\frac{m}{\sqrt{2}\sigma}\right) \\
&\quad + \frac{1}{2} \sum_{N_A=N/2+1}^{N-1} C_N^{N_A} p_A^{N_A} p_B^{N-N_A} \mathrm{erfc}\left(\frac{m}{\sqrt{2}\sigma}\right) \qquad (25)
\end{aligned}
$$

where m and σ depends on N_A, $N_B = N - N_A$, m_A, m_B, σ_A, σ_B as shown by equations (22) and (23) and erfc is the complementary error function defined by :

$$
\mathrm{erfc}(x) = 2 \int_x^{\infty} \frac{1}{\sqrt{\pi}} \exp(-u^2)\, \mathrm{d}u \qquad (26)
$$

4.1 Application to the BRIAC model

We now applied the above described method for a BRIAC case with $\alpha = .5$ and $N = 7$. With $x_{max} = 1$, then $x_{min} = -1/3$ and $m_A = 1/2$, $\sigma_A^2 = 1/12$, $m_B = -1/6$, $\sigma_B^2 = 1/108$, $p_A = x_{max}/(x_{max} - x_{min}) = 3/4$, $p_B = -x_{min}/(x_{max} - x_{min}) = 1/4$. Using these values in equation (25), we get $\tilde{P}_{BRIAC,\alpha=.5}(\infty,7) = 0.0629$. The different contributions are reported table 4. As expected, small values of N_A or N_B gives the smallest contributions and as N increases, they will become negligible. To be complete, the cases $N_A = N$ and $N_A = 0$, for which there is theoretically no paradox, have been computed. Because of the infinite support of the gaussian, the contributions are not zero but these values are not taken into account in formula (25).

The value for $\alpha = .5$ and $N = 3$ is equal to 0.1088 while the exact result computed section 3.2 gives $13/128 = 0.1015625$. Although the first result is valid for N going to infinity, it is already not so far even for this value of N.

Using this asymptotic approach, figure 1 shows the paradox probability as function of the number N of states for BRIAC.

4.2 Application to the IC model

For the IC case, each issue $\mathscr{A}$ or $\mathscr{B}$ has the same probability $p = 1/2$ to occur. Then for a state with n voters, the distribution of the vote is a binomial law. In addition, when the number of

Table 3: BRIAC case : paradox contribution for the different values of N_A, in the case $N = 7$ for $\alpha = .5$.

N_A	N_B	m	σ	$C_N^{N_A}$	paradox contribution
7	0	3.500	0.764	1	$3.065\,10^{-7}$
6	1	2.833	0.714	7	$1.118\,10^{-5}$
5	2	2.167	0.660	21	$1.592\,10^{-4}$
4	3	1.500	0.601	35	$1.086\,10^{-3}$
3	4	0.833	0.536	35	$5.422\,10^{-2}$
2	5	0.167	0.461	21	$7.394\,10^{-3}$
1	6	-0.500	0.373	7	$1.152\,10^{-4}$
0	7	-1.167	0.255	1	$1.402\,10^{-10}$
$\tilde{P}_{BRIAC,\alpha=.5}(\infty,7)$					0.06298850883

voters tends to infinity, this law tends to a normal law of mean $m = n/2$ and variance $\sigma^2 = n/4$. Let us shift and rescale this law by introducing the variable $x = (n-m)/\sigma$; we get

$$h(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right). \tag{27}$$

The two distributions $h_A(x)$ and $h_B(x)$ are truncated normal laws the mean and variance of which are given by $m_A = -m_B = \sqrt{2/\pi}$, $\sigma_A^2 = \sigma_B^2 = 1 - 2/\pi$ respectively. Then, using these values in formula (25), we are able to compute an approximation of the paradox probability in the IC case when n goes to infinity. For $N = 7$, $\tilde{P}_{IC}(\infty,7) = .1912$ and Table 5 gives the contribution of the different terms of the summation (25). Again we should not take into account the values for $N_A = 0$ and $N_A = 7$ for which there is no paradox. For $N = 21$, we get $\tilde{P}_{IC}(\infty,21) = .2012$ which is very close to the asymptotic limit $N \to \infty$ obtained by Monte Carlo method (see reference [1]). Figure 3 gives $\tilde{P}_{IC}(\infty,N)$ function of N.

5 Conclusion

Wilson and Pritchard cast a doubt about the results that were reported in Feix *et al.* [1]. The prime objective of this paper was to check who was right. Indeed, we recovered with different complementary methods all the results of Feix *et al.*

On the top of that, we have enhanced the technics at our disposal to compute exactly and/or estimate accurately the probability of the referendum paradox.

Table 4: IC case : paradox contribution for the different values of N_A, in the case $N = 7$.

N_A	N_B	m	σ	C_N^{NA}	paradox contribution
7	0	5.585	1.595	1	$1.804\,10^{-6}$
6	1	3.989	1.595	7	$3.383\,10^{-4}$
5	2	2.394	1.595	21	$1.094\,10^{-2}$
4	3	0.798	1.595	35	$8.434\,10^{-2}$
3	4	-0.798	1.595	35	$8.434\,10^{-2}$
2	5	-2.394	1.595	21	$1.094\,10^{-2}$
1	6	-3.989	1.595	7	$3.383\,10^{-4}$
0	7	-5.585	1.595	1	$1.804\,10^{-6}$
		$\tilde{P}_{IC}(\infty, 7)$			0.1912441506

The first method, based on the Ehrhart polynomial, gives exactly the probabilities as a function of the number n of voters in each state under the IAC assumption. Its main drawback comes from the fact that its computation cannot be performed when N (the number of states or districts) becomes large.

The second method assumes that n goes to infinity to simplify the computation. It is based upon a probabilistic approach developed in Feix *et al.* [2] which characterizes the voting behavior by a distribution function $f(p)$, where p is the probability to vote for one of the two candidates. It can be performed for simple models such as IAC (here up to $N = 100$) or BRIAC.

Using the Central Limit Theorem, the last method approximates the repartition of the votes for $\mathscr{A}$ and $\mathscr{B}$, by gaussian distributions. It is the most general method we presented, but it only gives an approximation when n and N are large. However, our approximations turn out to be excellent even for small values of N.

We hope that theses techniques could be further developed and applied to tackle other issues in social choice theory.

References

[1] Feix M, D Lepelley, V Merlin and JL Rouet (2004) The probability of paradox in a U.S. presidential type election. *Economic Theory* 23: 227-257.

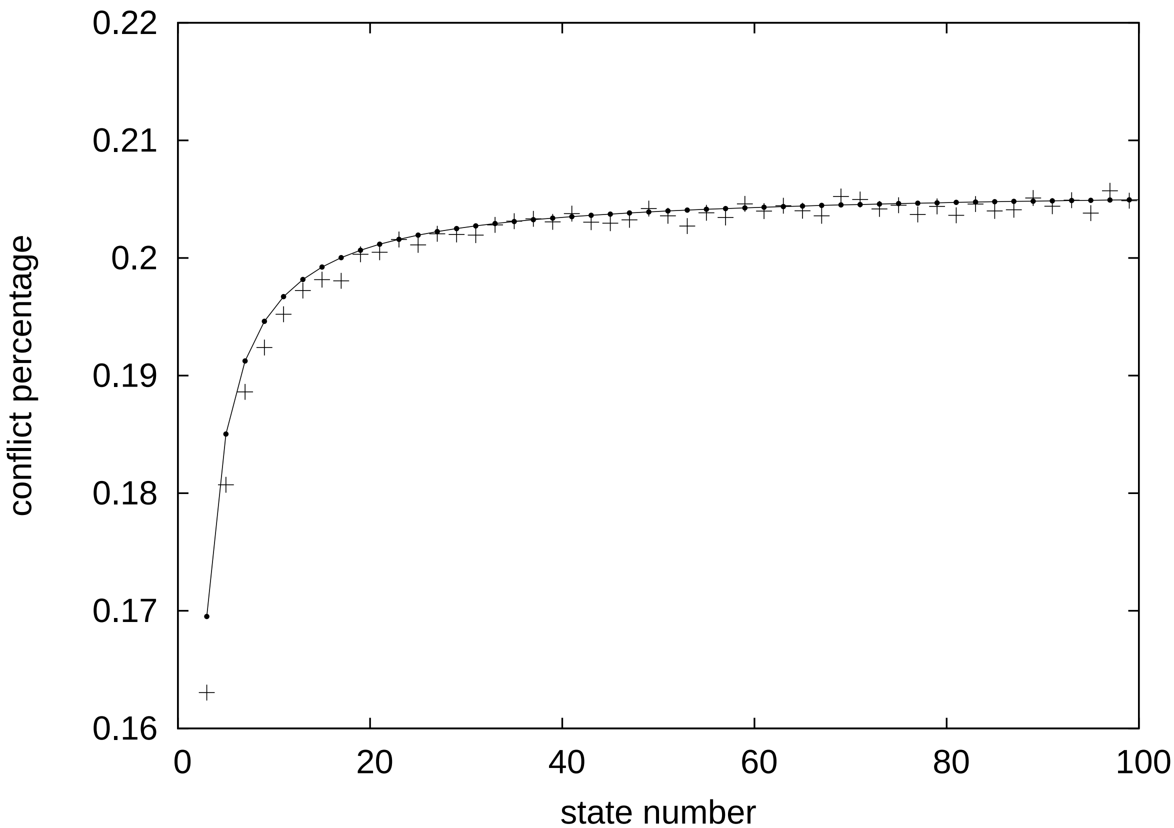

Figure 3: Paradox probability as a function of the number of states in the IC case. The values computed with the approximation method of section 4 are plotted with connected points; the crosses give the results obtained by Monte Carlo trials (see reference [1]).

[2] M.R. Feix, D. Lepelley, V. Merlin, J.L. Rouet : 2009, *On the probability to act in the European Union*, forthcoming in *The Mathematics of Preference, Choice and Order: Essays in Honor of Peter C. Fishburn, Studies in Choice and Welfare, Ed. S. Brams et al. (eds)*, Springer-Verlag Berlin Heidelberg.

[3] Lepelley D, A Louichi and H Smaoui (2008) On Ehrhart polynomials and probability calculations in voting theory. *Social Choice and Welfare* 30: 363-383.

[4] Wilson MC, G Pritchard (2007) Probability calculations under the IAC hypothesis. *Mathematical Social Sciences*.