Condorcet Winner - A Statistical Perspective

M. Raghavachari
Decision Sciences and Engineering Systems,
RPI, Troy, NY 12180

M. S. Krishnamoorthy
Department of Computer Science,
RPI Troy, NY 12180

1 Introduction

2 Condorcet Winner: Statistical Formulation and Results for General Case

In this section, we calculate the limiting probability of a Condorset winner when the voters are independent. Let us define

$$
\begin{aligned}
m &= \text{Number of Candidates} \\
n &= \text{Number of Voters}
\end{aligned}
$$

We have $K = m!$ preference rankings of m candidates. Let p_i be the probability that a voter prefers the rank order i, for $i = 1, 2, \cdots, K$. Further, we know that $p_i \geq 0$ and $\sum_{i=1}^{n} p_i = 1$. Let N_i be the number of voters voting for the ith preference ranking, for $i = 1, 2, \cdots, K$. We clearly have $\sum_{i=1}^{K} N_i = n$. Now, it is clear that $(N_1, N_2, \cdots, N_K)$ has a multinomial distribution.

$$
P(N_i = n_i, i = 1, \cdots, K) = \frac{n!}{n_1! n_2! \cdots n_K!} \prod_{i=1}^{K} p_i^{n_i} \tag{1}
$$

where n_i are nonnegative integers with $\sum_{i=1}^{K} n_i = n$.

It is well known that (see, for example, reference), for $i = 1, \cdots, K$, $j = 1, \cdots, K$ and $i \neq j$.

$$\begin{aligned} E[N_i] &= np_i, \\ Var[N_i] &= np_i(1-p_i), \\ Covariance(N_i, N_j) &= -np_ip_j, \\ Correlation(N_i, N_j) &= -\sqrt{\frac{p_ip_j}{(1-p_i)(1-p_j)}} \end{aligned}$$

Denote the m candidates by $C_1, C_2, \cdots, C_m$. Denote C_iYC_j, $i \neq j$ if C_i is preferred to C_j by a voter. Define C_iMC_j, $i \neq j$ if C_i beats C_j by majority rule. i.e., the total number of voters preferring C_i to C_j is more than the total number of voters preferring C_j to C_i. This fact can be mathematically expressed as

$$a_{1,(i,j)}N_1 + a_{2,(i,j)}N_2 + \cdots + a_{K,(i,j)}N_K > 0 \tag{2}$$

where $a_{l,(i,j)}$'s are $\pm$ 1. $a_{l,(i,j)} = 1$, if in the preference ranking voted by the N_l voters, C_iYC_j. Similarly $a_{l,(i,j)} = -1$ if C_jYC_i in the prefernce ranking voted by N_l voters. Clearly $\frac{K}{2}$ of the $a_{l,(i,j)}$'s are 1 and $\frac{K}{2}$ of the $a_{l,(i,j)}$'s are equal to -1. We define, for example, C_1 to be the Condorcet winner if and only if $C_1 \ M \ C_j$ for each $j = 2, \cdots, m$. The probability of C_1 is the Condorcet winner is the probability of the joint event

$$\begin{aligned} a_{1,(1,2)}N_1 + a_{2,(1,2)}N_2 + \cdots + a_{K,(1,2)}N_K &> 0 \\ a_{1,(1,3)}N_1 + a_{2,(1,3)}N_2 + \cdots + a_{K,(1,3)}N_K &> 0 \\ \cdots &> 0 \\ \cdots &> 0 \\ a_{1,(1,m)}N_1 + a_{2,(1,m)}N_2 + \cdots + a_{K,(1,m)}N_K &> 0 \end{aligned} \tag{3}$$

For small values n and m , this probability can be computed using the multinomial distribution (reference). We are interested in finding $\lim_{n\to\infty}$ $P(C_1$ is the Condorcet winner) for a given value of m.

For a given voter v, for $i = 1, \cdots, m$, j= $1, \cdots, m$, $i \neq j$, define

$$\begin{aligned} X(i,j,v) &= 1 \text{ if } C_iYC_j \\ &= -1 \quad \text{if} \quad C_jYC_i \end{aligned} \tag{4}$$

For the voter profile $(N_1, N_2, \cdots, N_K)$, $C_i M C_j$ if $\sum_{v=1}^{n} X(i,j,v) > 0$. The above representation has been used in Bell (reference) and Gehrlein (reference). Note that

$$\begin{aligned}
\sum_{v=1}^{n} X(i,j,v) &= a_{1,(i,j)} N_1 + a_{2,(i,j)} N_2 + \cdots + a_{K,(i,j)} N_K \\
\lambda_{ij} &= E(X(i,j,v)) \\
&= a_{1,(i,j)} p_1 + a_{2,(i,j)} p_2 + \cdots + a_{K,(i,j)} p_K \qquad (5) \\
Var(X(i,j,v)) &= 1 \ - \ \lambda_{ij}^2 \qquad (6)
\end{aligned}$$

Define

$$Z_{ij} \quad = \quad \frac{\sum_{v=1}^{n} (X(i,j,v) - \lambda_{ij}) \sqrt{(n)} \sqrt{(1 - \lambda_{ij}^2)}}{,} i \neq j \qquad (7)$$

We also need to obtain the correlation coefficient between $X(i,j,v)$ and $X(i,l,v)$, $i \neq j$, $i \neq l$ and $j \neq l$. To express correlation, we first define

$$\begin{aligned}
a_{v,(i,j),l} &= 1 \ \text{ if in the preference ranking corresponding to } N_v, \\
&\quad C_i Y C_j \text{ and } C_i Y C_l \\
&\quad \text{or } C_j Y C_l \text{ and } C_l Y C_i \\
&= -1 \ \text{ Otherwise} \qquad (8)
\end{aligned}$$

Covariance matrices are defined by the following equation.

$$Covariance(X(i,j,v), X(i,l,v)) \quad = \quad E[X(i,j,v), X(i,l,v)] - \lambda_{ij} \lambda_{il}$$

The Correlation between $X(i,j,v)$ and $X(i,l,v)$ is given by

$$\begin{aligned}
P_{jl}^{(i)} &= \frac{E(X(i,j,v).X(i,l,v)) \ - \ \lambda_{ij}\lambda_{il}}{\sqrt{(1-\lambda_{ij}^2)(1-\lambda_{il}^2)}} \quad j = 1, \cdots, m, \\
&\qquad\qquad l = 1, \cdots, m \text{ and } j \neq l \\
&= \frac{\sum_{v=1}^{K} a_{v,(i,j),l} N_v - \lambda_{ij}\lambda_{il}}{\sqrt{(1-\lambda_{ij}^2)(1-\lambda_{il}^2)}}
\end{aligned}$$

The probability that $C_i M C_j$ is given by

$$P(\sum_{v=1}^{n} X(i,j,v) > 0) \Longrightarrow P(Z_{ij} > \frac{-\sqrt{n}\lambda_{ij}}{\sqrt{1-\lambda_{ij}^2}}), \text{from 7.}$$

The probability that C_i is the Condorcet winner is therefore given by the joint probability

$$P[Z_{ij} > \frac{-\sqrt{n}\lambda_{ij}}{\sqrt{1-\lambda_{ij}^2}}, j = 1, \cdots, m, \;\; i \neq j \;]. \tag{9}$$

By multivariate central limit theorem applied to independent summands Z_{ij},$j = 1, \cdots, m$, $i \neq j$, the joint distribution of $(Z_{i1}, \cdots, Z_{i,i-1}, Z_{i,i+1}, \cdots, Z_{iK})$ tends as $n \to \infty$ to the multivariate normal distribution with zero mean vector, unit variances and correlation matrix $P_i \;=\; (P_{jl}^{(i)})$. Let us write

$$\delta_{ij} = \begin{cases} -\infty & \text{if } \lambda_{ij} > 0 \\ \infty & \text{if } \lambda_{ij} < 0 \\ 0 & \text{if } \lambda_{ij} = 0 \end{cases}$$

Denote by $L(h_1, h_2, \cdots, h_{m-1}; P)$ the multivariate normal probability

$$P(Z_1^* \geq h_1, Z_2^* \geq h_2, \cdots, Z_{m-1}^* \geq h_{m-1})$$

where $(Z_1^*, Z_2^*, \cdots, Z_{m-1}^*)$ has the multivariate normal distribution with zero mean vector, unit variances and Correlation matrix P. Then we have

$$\lim_{n\to\infty} P[Z_{ij} > -\frac{\sqrt{n}\lambda_{ij}}{\sqrt{1-\lambda_{ij}^2}}, j = 1, \cdots, m, \;\; i \neq j \;]$$

$$= L(\delta_{i1}, \delta_{i2}, \cdots, \delta_{i,i-1}, \delta_{i,i+1}, \cdots, \delta_{iK}; P_i).$$

The limiting probability of a Condorcet winner is then

$$P(\infty, m) \;=\; \sum_{i=1}^{m} L(\delta_{i1}, \delta_{i2}, \cdots, \delta_{i,i-1}, \delta_{i,i+1}, \cdots, \delta_{iK}; P_i).$$

3 Condorcet Winner: Some Special Cases

First we assume that all preference rankings for a voter are equally likely. This implies that $p_i \;\equiv\; 1/K$, for $i = 1, \cdots, K$. For this For this case, symmetry conditions imply that $\lambda_{ij} \;\equiv\; 0$ for $i = 1, \cdots, m$, $j = 1, \cdots, m$ and $i \neq j \Rightarrow \delta_{ij} \;\equiv\; 0$. For this case, it can be verified that $P_i \;\equiv\; P^*$ where P^*

is the $(m-1) \times (m-1)$ equicorrelated matrix with equal correlation value $= 1/3$. The limiting probability of a Condorcet winner is therefore given by

$$P(\infty, m) \quad = \quad L(0,0,\cdots,0;P^*) \tag{10}$$

The expression 10 for general $m \geq 4$ candidates, we believe is new. This expression for $m = 3$ was obtained by Gilband(1952) (reference) as $\frac{3}{4} + \frac{3}{2\pi}\arcsin(\frac{1}{3})$. For $m \geq 4$, several researchers have obtained approximate values for expression 10 based on simulation studies. See for example Gehrlein(1969), Jones et al (1995) (refernces). The expression 10 shows that the limiting probability for m candidates can be obtained by an evaluation of positive orthant probabilities for $(m-1)$ dimensional multivariate normal distributioon with zero mean vector, unit variances and equicorrelated correlation matrix with all correlations equal to 1/3. Several papers in statistical literature deal with such evaluations. See, for example, David and Mallows(1961), Sondhi (1961), Placket (1954), Johnson and Kotz (1972) and Bacon (1963) (references). Bacon(1963) derives explicit expressions for the positive orthont probabilities for a few small values of m for multivariate normal distribution with equicorrelated correlation matrix with all correlations equal to ρ. He also derives a recursive relation to calculate the probabilities for successive values of m. Such expressions involve the evaluation of multiple integerals without known closed form expressions. Denote the expression 10 by $P(m)$. Using the expressions from Bacon (1963) (reference), the values of $P(m)$ for some small values of m are

$m = 3, \quad P(3) = \frac{3}{4} + \frac{3}{2\pi}\arcsin(\frac{1}{3})$

$m = 4, \quad P(4) = \frac{1}{2}[1 + \frac{6}{\pi}\arcsin(\frac{1}{3})]$

$m = 5, \quad P(5) = \frac{5}{16}[1 + \frac{12}{\pi}\arcsin(\frac{1}{3}) + \frac{24}{\pi^2}\int_0^{\frac{1}{3}}\arcsin(\frac{\lambda}{1+2\lambda})\frac{d\lambda}{\sqrt{(1-\lambda^2)}}]$

$m = 6, \quad P(6) = \frac{3}{16}[1 + \frac{20}{\pi}\arcsin(\frac{1}{3}) + \frac{120}{\pi^2}\int_0^{\frac{1}{3}}\arcsin(\frac{\lambda}{1+2\lambda})\frac{d\lambda}{\sqrt{(1-\lambda^2)}}]$

$m = 7, \quad P(7) = \frac{7}{64}[1 + \frac{30}{\pi}\arcsin(\frac{1}{3}) + \frac{36}{\pi^2}\int_0^{\frac{1}{3}}\arcsin(\frac{\lambda}{1+2\lambda})\frac{d\lambda}{\sqrt{(1-\lambda^2)}}$
$\quad + \frac{720}{\pi^3}\int_0^{\frac{1}{3}}\int_0^{\frac{\mu}{1+2\mu}}\arcsin(\frac{\lambda}{1+2\lambda})\frac{d\lambda}{\sqrt{(1-\lambda^2)}}\frac{d\mu}{\sqrt{(1-\mu^2)}}]$

For $m \geq 7$, Bacon provides an approximation believd to be reasonably close to the true value. Figure 1 shows the plot of $P(\infty, m)$ for increasing values of m. It can be seen that $P(\infty, m) \to 0$, as $m \to \infty$ at the rate of $\frac{1}{m}$ (reference May). It is clear that $P(\infty, m)$ monotonically decreases to a limit.

4 Condorcet Winner: Special Case $m = 3$ candidates

May(1971) (reference) discussed the case of 3 candidates to obtain the limiting probability of a Condorcet winner. We drastically improve and give a compact solution utilizing the results of Section 2. Suppose the 6 rankings are

rankings	Number of Voters	Probabilities
$C_1C_2C_3$	N_1	p_1
$C_1C_3C_2$	N_2	p_2
$C_2C_1C_3$	N_3	p_3
$C_2C_3C_1$	N_4	p_4
$C_3C_1C_2$	N_5	p_5
$C_3C_2C_1$	N_6	p_6

It can be verified that

$$\begin{aligned}
\lambda_{12} &= p_1 + p_2 + p_5 - p_3 - p_4 - p_6 \\
\lambda_{13} &= p_1 + p_2 + p_3 - p_4 - p_5 - p_6 \\
\lambda_{21} &= p_3 + p_4 + p_6 - p_1 - p_2 - p_5 \\
\lambda_{23} &= p_1 + p_3 + p_4 - p_2 - p_5 - p_6 \\
\lambda_{31} &= p_4 + p_5 + p_6 - p_1 - p_2 - p_3 \\
\lambda_{32} &= p_2 + p_5 + p_6 - p_1 - p_3 - p_4
\end{aligned}$$

The three correlation matrices P_1, P_2, P_3 are given by:

$$P_1 = \begin{pmatrix} 1 & \frac{(p_1+p_2+p_4+p_6-p_3-p_5)-\lambda_{12}\lambda_{13}}{\sqrt{(1-\lambda_{12}^2)(1-\lambda_{13}^2)}} \\ \frac{(p_1+p_2+p_4+p_6-p_3-p_5)-\lambda_{12}\lambda_{13}}{\sqrt{(1-\lambda_{12}^2)(1-\lambda_{13}^2)}} & 1 \end{pmatrix}$$

The (1,2) entry of P_1 is P_{23}^1.

$$P_2 = \begin{pmatrix} 1 & \frac{(p_2+p_3+p_4+p_5-p_1-p_6)-\lambda_{21}\lambda_{23}}{\sqrt{(1-\lambda_{21}^2)(1-\lambda_{23}^2)}} \\ \frac{(p_2+p_3+p_4+p_5-p_1-p_6)-\lambda_{21}\lambda_{23}}{\sqrt{(1-\lambda_{21}^2)(1-\lambda_{23}^2)}} & 1 \end{pmatrix}$$

The (1,2) entry of P_2 is P^2_{13}.

$$P_3 = \begin{pmatrix} 1 & \frac{(p_1+p_3+p_5+p_6-p_2-p_4)-\lambda_{31}\lambda_{32}}{\sqrt{(1-\lambda_{31}^2)(1-\lambda_{32}^2)}} \\ \frac{(p_1+p_3+p_5+p_6-p_2-p_4)-\lambda_{31}\lambda_{32}}{\sqrt{(1-\lambda_{31}^2)(1-\lambda_{32}^2)}} & 1 \end{pmatrix}$$

The (1,2) entry of P_3 is P^3_{12}. The probability of a Condorcet winner is

$$L(\delta_{12}, \delta_{13}, P^1_{23}) + L(\delta_{21}, \delta_{23}, P_D 13^2) + L(\delta_{31}, \delta_{32}, P^3_{12}) \tag{11}$$

. By considering the 27 possibilities with which $\lambda_{12}, \lambda_{13}$ and λ_{23} can take zero, positive and negative values, we can simplify the expression further. The table 1 shows the 27 possibile cases and the corresponding limiting value of the probability of a Condorcet winner. A set of different values is obtained for varying values of p_i's. It is also to be noted that for a few alternatives the values of 0 and 1 are obtained for the limiting probability of a Condorcet winner. See alternatives 7 and 17 in Table 1

It can be verified that case 7, for example, is reached when $p_1 = p_2 = p_5 = \frac{1}{6}$; $p_3 = \frac{2}{5}$ and $p_6 = \frac{1}{10}$. Case 17 is reached when $p_1 = \frac{3}{22}, p_2 = \frac{3}{22}, p_3 = \frac{9}{44}, p_4 = \frac{13}{44}$, $p_5 = \frac{5}{22}$ and $p_6 = 0$.

Number	λ_{12}	λ_{13}	λ_{23}	Probability of a Winner
1	0	0	0	$\frac{1}{4}+\frac{1}{2\pi}\arcsin(P^1_{23})+\frac{1}{4}+\frac{1}{2\pi}\arcsin(P^2_{13})+\frac{1}{4}+\frac{1}{2\pi}\arcsin(P^3_{12})$
2	0	0	+	$\frac{1}{2}+\frac{1}{4}+\frac{1}{2\pi}\arcsin(P^1_{23})$
3	0	0	-	$\frac{1}{2}+\frac{1}{4}+\frac{1}{2\pi}\arcsin(P^1_{23})$
4	0	+	0	$\frac{1}{2}+\frac{1}{4}+\frac{1}{2\pi}\arcsin(P^2_{13})$
5	0	-	0	$\frac{1}{2}+\frac{1}{4}+\frac{1}{2\pi}\arcsin(P^2_{13})$
6	0	+	-	$\frac{1}{2}$
7	0	+	+	1
8	0	-	+	$\frac{1}{2}$
9	0	-	-	1
10	+	0	+	$\frac{1}{2}$
11	+	+	0	1
12	+	0	-	1
13	+	-	0	$\frac{1}{2}$
14	+	+	+	1
15	+	-	-	1
16	+	+	-	1
17	+	-	+	0
18	+	0	0	$\frac{1}{2}+\frac{1}{4}+\frac{1}{2\pi}\arcsin(P^3_{12})$
19	-	0	0	$\frac{1}{2}+\frac{1}{4}+\frac{1}{2\pi}\arcsin(P^3_{12})$
20	-	0	+	1
21	-	0	-	$\frac{1}{2}$
22	-	+	0	$\frac{1}{2}$
23	-	+	+	1
24	-	+	-	0
25	-	-	0	1
26	-	-	+	1
27	-	-	-	1

Table 1: Condorcet Winner Probabilities