POWER VOTING

Degree Thesis

BY

NIKÉ S. PANTA

BSc Mathematics

Mathematical Analyst Specialisation

Supervisor:

László Varga, assistant lecturer
Department of Probability Theory and Statistics
Eötvös Loránd University, Faculty of Science

Eötvös Loránd University
Faculty of Science
2017

Acknowledgements

First of all I would like to thank my thesis advisor, Laszló Varga, assistant lecturer of the Faculty of Science at Eötvös Lóránd University. Varga was truly dedicated towards helping me and the door to his office was always open in case I needed advice or I had a question about my thesis and research. Without his advice, comments and help the thesis could not have been successfully written.

I would also like to acknowledge dr. Gábor Dániel Nagy, senior lecturer of Faculty of Humanities at University of Szeged. He played an important role as an advisor for modeling different voting systems and his comments and advices are truly appreciated and thanked.

Finally, I would like to thank my family, friends and colleagues for all the support during the final year of my studies. Special thanks to my sister, Kira who never had any doubt in me.
Without all of the people I am surrounded by, this thesis would not have been completed.

Niké S. Panta

Contents

1 Introduction

"We hold these truths to be self evident, that all men are created equal, that they are endowed by their Creator with certain unalienable Rights, that among these are Life, Liberty and the pusrsuit of Happiness."

The Declaration of Independence, 1776

When I was living in the U.S. my first homework for American History Class was to memorize the above quotation. My teachers, Mr. Pagaard and Mr. Home were the first ones to introduce me to how the voting system in the United States of America works. Even back then - at the age of 17 - I was interested in Math, and was already having questions about the system. I found this a truly exciting topic, and was hoping to get more familiar with it in the future.

A few years later, I saw the movie "Swing Vote", which is an American movie showing a case when the election in the U.S. is determined by one man's vote. I found myself sitting and thinking about whether it is mathematically possible, later I understood that my mathematical knowledge is not enough for either understanding nor solving the complex problem.

When one of my university teachers advised me on topics for my thesis, he asked me to read articles from the *understandinguncertainty.org* webpage. After going through numerous essays, one of them really caught my eye. The title of it was *When does a single vote count?*.
I was speechless.
Years after I came home and saw the movie, finally the time had come. I proposed this theme as the field of my thesis, and luckily, my supervisor, Laszlo Varga was also interested in the topic. Finally I had the basic knowledge to write and think about the statistical and probabilistic background of the American voting system in which the form of decision making is power voting.

Power Voting is a well known form of decision making in many organizations. This form of governance is aimed to give different amounts of influence/power to different members throughout decision making. One good example is the election of the President of the United States which is unique in many ways.
The President is elected by the Electoral College of the United States of America, and not by the direct vote of citizens. Each of the 50 states - and Washington D.C. as well- have a pre-assigned number of Electors, based mostly on the population of the state. This

number gives the voting power of each state and therefore can be analyzed as a power voting phenomenon.

As already mentioned, numerous organizations are using this form of governing - such as joint stock companies, economic organizations, the World Bank and the Council of the European Union. In all of these systems, the individual votes are cast as blocks of different sizes.

These voting habits can be seen as weighted voting games, where one can be studying the possible voting outcomes as well as the relation of individual members' to them.

As Leech stated

> *"the power of a member is the same as his/her share of the votes."*

Later on, in his article (Leech [2002]) Leech illustrates the power voting system with a few examples, including the case of an imaginary company, where the shareholders are having the following percents of the shares: 49% , 49% and 2%. He notes, that

> *"it is not useful to describe these figures as shares of the power each has in running the company because if the decision rule requires a simple majority of more than 50 percent of the votes, then any two shareholders are required to support a motion for it to pass (...) and therefore the one with 2 percent has exactly the same power as one with 49 percent. Therefore by considering all possible voting outcomes it becomes clear that each shareholder has equal power despite the disparity in their votes."*

Another famous example is the original Council of the European Economic Community. Between 1958 and 1972 the community had 6 members : France, West Germany, Italy, Belgium, the Netherlands and Luxembourg with their allocated voting power of 4, 4, 4, 2, 2 and 1.
They have used 3 different systems for decision making, one of them required a fixed number of votes (12) for passing an agreement.
After checking all the possible outcomes, one can easily note, that the vote of Luxembourg could never have made any difference, therefore the real voting power of Luxembourg - was 0 instead of $\frac{1}{17}$.

The field of voting interest, power voting and voting power in mathematics and economics have been studied at least for the past 60 years. The work of Shapley and Shubik, Banzhaf, Andrew Gelman, Nicholas R. Miller and Dennis Leech provided a really good

base for my thesis.
The thesis is mostly focusing on the computation of power indices and the estimation of the probability of such event that has never occured- the case of one vote being decisive in the American Presidential Election.

2 Notations and Definitions

Assuming $N = \{1, 2, 3, \dots, n\}$ is the set of n players / members participating in a weighted voting game. Each of the players have a pre-assigned voting weight, which for player i is w_i, $w_i \in \mathbb{R}^+$. The members of a coalition are represented by subset T ($T \subset N$), and the combined voting weight of the members of T is marked by $w(T)$.

$$w(T) = \sum_{i \in T} w_i$$

In all decision making situations, a rule is defined regarding the quote that is needed to pass a regulation. In this thesis, the quote will be illustrated by q, $q \in \mathbb{R}^+$. Coalition T must have a combined voting weight bigger than, or equal to q in order to pass a decision or to come out as a winner.

if $w(T) \geq q \implies$ coalition T wins;

if $w(T) < q \implies$ coalition T loses.

Throughout the thesis, the voting body will be represented by the following notation: $\{q; w_1, \dots, w_n\}$. Defining $q > \frac{w(N)}{2}$ is essential, as in most voting cases a coalition has to have the majority of votes in order to win.

Definition 2.1 (power index) $p \in \mathbb{R}^n$ *defines the ability of each player for determining the outcome of a vote.*

The power index for each player is defined the following way: assuming that coalition T_i is losing without the voting weight of i, but in case i transfers its vote, and becomes a member of the previously defined coalition T_i (noted by $T_i + \{i\}$) is winning. In this case, player i is often referred as a **swing player**, and usually represented and defined as $(T_i, T_i + \{i\})$. When defining the power indices, it is worth to mention, that based on the way a coalition is formed/treated, the power indices vary the following way:

1. Assuming that each possible coalition is treated equally, the power index for player i is defined as the relative frequency of swings for player i.

2. In case coalitions are being formed randomly, the power index for player i is a probability.

Please note, that the two concepts defining the power index are fundamentally different.

Throughout the thesis, the following terms will also be used: Shapley - Shubik-, Banzhaf- and Coleman index. These indices were formulated to measure the powers of players in a voting game.

For understanding the Shapley - Shubik (SS) index, let us take a look at the below example (Wikipedia [2016b]):

Suppose the decision in a voting game is made according to the simple majority rule. The voting body consists 3 players : A, B and C, with the votes: 3, 2 and 1. Due to the previously defined decision making rule $q = 4$. This can easily be represented by using the following notation: $\{4; 3, 2, 1\}$

There are $3! = 6$ possible orders in which the members can vote. In case a voter is the first one to raise the combined voting weight of a coalition to 4 or to above 4, it will be marked with **bold** letters.

A**B**C	A**C**B
B**A**C	BC**A**
C**A**B	CB**A**

Player A is a swinger in 4 cases, B in 1 and C in 1. In this voting game, the power index for player i is the probability of a swing for the player.

For player A the power index is $\frac{4}{6} = \frac{2}{3}$, for player B it is $\frac{1}{6}$ and for player C it is $\frac{1}{6}$.

Definition 2.2 (Shapley - Shubik power index) *The SS power index for player i is the probability of a swing for the player in the given game.*

To analyze the voting of the Council of the European Union the Shapley - Shubik index is used.

Definition 2.3 (swing of player i) *The swing for player i is the number of all possible orders/sequences of the members of T_i and its supplementary set.*

$$N - T_i - \{i\} = t!(n - t - 1)!$$

Supposing t is the number of the members of T_i, and n marks the number of the members of N. Therefore, the number of all swings for player i as the member of the already defined coalition:

$$\sum_{T_i} t!(n - t - 1)!$$

The index ϕ_i is used by Leech. ϕ_i is defined as the proportion of numbers of orderings all players in N.

$$\phi_i = \sum_{T_i} \frac{t!(n - t - 1)!}{n!} \tag{1}$$

The Banzhaf index treats all coalitions as they are being found with equal probability, and the order in which the players are voting does not matter. Just like there was an example for understanding the SS index, now let us see an example using the same voting body for the Banzhaf index (Wikipedia [2016a]).

The groups below are reprezenting the possible coalitions that are having a combined voting weight greater than or equal to 4. In case a player is a swing player for the given coalition it will be marked with a **bold** letter.

AB AC ABC

In total there are 5 swings, A is a swing player in 3 cases, and both B and C in 1 case. Therefore the Banzhaf power index for player A is $\frac{3}{5}$, for player B is $\frac{1}{5}$ and for player C is $\frac{1}{5}$.

Definition 2.4 (Banzhaf power index) *The Banzhaf power index for player i is the quotient of the total number of player i being a swing player and the total swings.*

The Banzhaf index is used when analyzing the voting of the U.S. Electoral College.
The number of swings for player i can be calculated in the following way:

$$\eta_i = \sum_{T_i} 1 \tag{2}$$

β_i' is the number of coalitions that are not having player i as a member. When thinking about the basics of set theory, one can easily see the connection β_i' has with 2^{n-1}. The number of subsets is $N - \{i\}$, therefore β_i' can be also be calculated as:

$$\beta_i' = \sum_{T_i} \frac{1}{2^{n-1}} = \frac{\eta_i}{2^{n-1}} \tag{3}$$

This value is also called the non-normalized Banzhaf index.

In case the total number of swings for all players are used to measure relative voting power among players

$$\beta_i = \frac{\eta_i}{\sum\limits_{i=1}^{n} \eta_i} \tag{4}$$

β_i is called the normalized Banzhaf index. If the value would like to be calculated by using the non-normalized Banzhaf index, the formula looks like this:

$$\beta_i = \frac{\beta_i'}{\sum\limits_{i=1}^{n} \beta_i'} \tag{5}$$

After using both the Shapley - Shubik and the Banzhaf power indices for the same voting game, it is important to note how the power indices vary based on which method is used.

Player	SS power index	Banzhaf power index
A	$\frac{2}{3}$	$\frac{3}{5}$
B	$\frac{1}{6}$	$\frac{1}{5}$
C	$\frac{1}{6}$	$\frac{1}{5}$

As already mentioned, a third type of power index also exists. It is called the Coleman index. For calculating the Coleman index, it is important to know the number of the winning coalitions.

$$\omega = \sum_{S} 1 \text{ for all } S \in N, \text{ supposing } w(S) \geq q$$

Based on the above, Leech noted the following 3 consequences:

- First of all, the power of a voting body to Act:

$$A = \frac{\omega}{2^n} \tag{6}$$

- The power of player i to prevent action:

$$PPA_i = \frac{\eta_i}{\omega} \tag{7}$$

- The power of member i to initiate action:

$$PIA_i = \frac{\eta_i}{2^n - \omega} \tag{8}$$

3 Computation of power indices

Different methods can be found for computing power indices. However, it is important to mention, most of these methods and algorithms are designed to work well only for games with a small amount of players. For n players the run time is estimated to be 2^n, therefore in case a voting game has more than 30 players it takes a long time to compute the power indices.
Leech has proposed to use generating functions for calculating not just the Shapley-Shubik, but also the Banzhaf and Coleman indices of a voting game.
The method itself is linear in n therefore it can be used for calculating the indices for games that are having a larger number of players. It is also important to mention that the method of generating functions demands a lot of storage, therefore it limits the use for really large games.

3.1 Banzhaf index

The aim is to find the number of swings for each and every member, as well as the number of winning coalitions, therefore finding the number of coalitions S - which marks a subset of N with each sizes of $w(S)$- is essential. The number of occurrences of coalitions of each size can be calculated in the following way:

$$f(x) = \prod_{j=1}^{n} (1 + x^{w_j}) \tag{9}$$

Please note, x has no important meaning in itself, it is a coefficient that helps measuring the function.
(7.) can be written as the below:

$$f(x) = \sum_{j=0}^{w(N)} a_j x^j \tag{10}$$

while a_j is the number of coalitions having the total weight of j. It is easily noted, that a_j equals to the number of the subsets of S, therefore $w(S) = j$. All the coefficients can be found numerically, by rewriting (7.) using the correct multiplicators. This will give the following:

$$\begin{aligned} f(x) &= \prod_{j=1}^{n} (1 + x^{w_j}) = [1 + x^{w_1}] \prod_{j=2}^{n} (1 + x^{w_j}) \\ &= [1 + x^{w_1} + x^{w_2} + x^{w_1+w_2}] \prod_{j=3}^{n} (1 + x^{w_j}) = \dots \end{aligned}$$

Instead of using the above formula, the polynomial can be denoted as:

$$1 + a_1^{(r)}x + a_2^{(r)}x^2 + a_3^{(r)}x^3 + \ldots + a_w^{(r)}x^w$$

$f(x)$ can be rewritten in the following way, supposing $r = 1, \ldots, n$ and for all $j : a_j^{(n)} = a_j$:

$$f(x) = \prod_{j=1}^{n}(1 + x^{w_j}) = [1 + a_1^{(1)}x + \ldots + a_w^{(1)}x^w]\prod_{j=2}^{n}(1 + x^{w_j})$$
$$= [1 + a_1^{(2)}x + \ldots + a_w^{(2)}x^w]\prod_{j=3}^{n}(1 + x^{w_j})$$
$$= \ldots = 1 + a_1^{(n)}x + \ldots + a_w^{(n)}x^w$$

Supposing that for all a_j^r, $j \geq w_r$ and $a_j^0 = 0$ for all $j \neq 0$

$$a_j^r = a_j^{r-1} + a_{j-w_r}^{r-1} \tag{11}$$

for all $r = 1, \ldots, n$ while $a_j^n = a_j$. The coefficients, such as a_j can be found after iterating the formula.

3.2 Coleman index

The computation of the Coleman index is quite similar to the computation of the Banzhaf index. Contonuing from $a_j^r = a_j^{r-1} + a_{j-w_r}^{r-1}$, it is important to mention how there is no need to store 2 arrays for containing all data about the coefficient a, and one array is perfectly enough to contain all a related data, if the rule is applied in reverse order. However, it is essential to know the number of winning coalitions,

$$\omega = \sum_{j=q}^{w} a_j$$

which is necessary for finding all a_j, which is needed for calculating the number of swings for all members, which also depends on the weight of the member. Rewriting the formula $f(x)$ with the help of a new coefficient c is possible, where the creation rule for c is the following one :

$$c_j = a_j - c_{j-w_i} \text{ for all } j = 1, \ldots, v \tag{12}$$

$$f(x) = (1 + c_1x + \ldots + c_vx^v)(1 + x^{w_i}) = 1 + a_1x + \ldots + a_wx^w$$

Therefore the number of swings for each member i, η_i is:

$$\eta_i = \sum_{j=q-w_i}^{q-1} c_j$$

3.3 Shapley-Shubik index

For computing the Shapley-Shubik index, the following generating function is used:

$$f(x,y) = \prod_{j=1}^{n}(1 + x^{w_j}y) \tag{13}$$

Arguments x and y are both needed, since this index requires not only the number of the players, but the number of votes for each different sized coalitions. Using the notation of d_{jk} which marks the number of k sized coalitions with the combined voting weight of j. In case of using a matrix for reprezenting the values of d_{jk} for different j and k values the function can be rewritten as:

$$d_{jk}(r) = \begin{cases} d_{jk}(r-1) + d_{j-w_r,k-1}(r-1) & \text{if } j,k \geq 0 \\ 0 & \text{if } j < 0 \text{ or } k < 0 \end{cases}$$

As before, in case of a negative subscript - the coeffitient is zero. Using this rule one is able to generate a matrix with $w+1$ rows and $n+1$ coloumns. The index for each player i is found by calculating the number of swings with k members with votes equal to j, c_{jk}, using a similar rule as before:

$$c_{jk} = d_{jk} - c_{j-w_i,k-1} \tag{14}$$

Using the above formula, ϕ for each i can be rewritten as:

$$\phi_i = \sum_{k=0}^{n-1} \frac{k!(n-1-k)!}{n!} \sum_{j=q-w_i}^{q-1} c_{jk} \tag{15}$$

3.4 Other methods for calculating power indices

The so called Multilinear Extension Approximation Algorithms were introduced for calculating the Non-Normalized Banzhaf and the Shapley-Shubik indices for larger games. This methos uses the central limit theorem in order to get approximations for the already defined (1.) and (3.) equations. The methods have linear complexity and the calculation of the indices for a player depends on the number of players n only in the data input and in the calculation of $w(N)$. Even though these algorithms have been widely used, their accurancy strongly depends on the validity of the normal approximation. In real life, these methods are not reliable as the approximation is not good and consequent computation errors may be substantial because of the failure of the central limit theorem due to the concentration of the voting weights in the hands of a few players.

The algorithms of curse can be modified in order to get a more sustainable outcome. However, the already proposed changes require a search over multiple subsets and the calculations have to be repeated 2^m times (having m as the player with the largest voting weight). The modified algorithms are successful in calculating the indices for larger games, yet the computing time can be incredibly big.

An oceanic game is a game where the number of the players is allowed to be infinitive. Indices for oceanic games were studied by Shapley. Due to the infinite number of players, one can easily see how the previously mentioned methods are not going to work for calculating the power indices. In his article (Leech [2002]) Leech only mentions these kind of games for completeness. He describes, that the calculation of power indices works only, if one is putting the gamers into 2 groups - a group with having a fixed voting weight, and one with very small weights. As the second goes to infinity, the Shapley-Shubik indices can be calculated with the help of the modified Multilinear Extension Approximation Algorithms. In case the first group is having too many players, this method may not work, and modifications are needed. The Banzhaf indices for oceanic games can be calculated by any algorithms that can be used for voting games with finite players.

These methods and cases will not be studied in this thesis.

4 Computation of the Banzhaf and Shapley Shubik indices for different systems

4.1 Writing the program

Using *Project R*, a program was written for computing the Banzhaf and Shapley-Shubik indices using the generating functions that were previously defined.
When introducing the Banzhaf and Shapley-Shubik indices in the thesis, the following example was shown:
The voting body consists of 3 players: A,B and C, with the votes: 3,2 and 1. Due to the simple majority rule $q = 4$.
For calculating the Banzhaf indices of each player, the generating funcion of the example can be written as:

$$f(x) = (1 + x^3)(1 + x^2)(1 + x)$$

Having the previously introduced rule $a_j^{(r)} = a_j^{(r-1)} + a_{j-w_r}^{(r-1)}$, where $r = 1, \ldots, 3$; $j = 1, \ldots, 6$; $q = 4$; $n = 3$ and $w = 6$ the below table can be generated.

Table 1: The evolution of array a

		A	B	C
		w_1=3	w_2=2	w_3=1
j	0	r_1	r_2	r_3
0	1	1	1	**1**
1	0	0	0	**1**
2	0	0	1	**1**
3	0	1	1	**2**
4	0	0	0	**1**
5	0	0	1	**1**
6	0	0	0	**1**

Meaning the generating function is equal to:

$$f(x) = 1 + x + x^2 + 2x^3 + x^4 + x^5 + x^6$$

The number of winning coalitions $\omega = \sum_{j=q}^{w} a_j = a_4 + a_5 + a_6 = 1 + 1 + 1 = 3$. By using the below array the number of swings for each player (η_i). This is found from the number of coalitions with the weight of j ehich do not include i: $c_j = a_j - c_{j-w_i}$ - therefore the η

and β values can be calculated.

Table 2: Array used for calculating the number of swings for each i and j

		A	B	C
j	a_j	3	2	1
0	*1*	1	1	1
1	*1*	1	1	0
2	*1*	1	0	1
3	2	1	1	1
4	*1*	0	1	0
5	*1*	0	0	1
6	*1*	0	0	0

The above table is used to calculate the η and β values for each player.

$$\eta_A = 3, \eta_B = 1 \text{ and } \eta_C = 1$$
$$\beta_A = \frac{3}{5}, \beta_B = \frac{1}{5} \text{ and } \beta_C = \frac{1}{5}$$

The original values and the ones calculated by using the generating function are the same.

4.2 The United States of America

As already mentioned, power voting is a well known form of decision making and one good example for that is the election of the President of the United States of America, since the President is elected by the Electoral College and not by direct vote of citizens. The method is declared by the Constitution.

The numbers of the electoral votes vary based on the population of the state and the values are currently between the ranges of 3 to 55. Therefore the total number of electors in the Electoral College is 538 (please see table later). As the American voting system is using the simple mayority method, the number of electoral votes needed for someone to win the election is 270.

In case both candidates fail to gain 270 votes, the decision goes to Congress. Assuming, that both candidates get 269 - 269 votes, the decision counts as an undecided one, and therefore is most likely to go to Congress. In case of a tie vote, each state's Congressman would have 1 vote, and the majority of the votes are needed for a candidate to win.

Gaining exactly 269 votes is mathematically possible. Assuming that one candidate earns the electoral votes of: California, Florida, Georgia, Illinois, Michigan, Missouri, New York, North Carolina, Ohio, Pennsylvania, Rhode Island and Texas,the total number of electoral votes are 269. Even though the probability of this happening is relatively small, it is still possible.

It is important to note, that throughout my thesis, I will disregard that some of the states can distribute their electoral votes between the 2 candidates (for example Maine). Overall, there is no law that requires the Electors to vote according to the results of the popular vote in their states. Some of the states require their electors to cast their votes according to the popular vote, and some of them allow the elector to go against it - however it is really rare.

The unique system makes it possible for one candidate to win the popular vote (the sum of all votes cast in each state and the District of Columbia) but to fail gaining the 270 electoral votes. This has happened in 2000, when George W. Bush run against Al Gore and won the election by gaining 271 electoral votes but only 48.87 % of the popular votes.

Also, the most recent polls show, that the same situation happened in the Presidential Election of the USA in 2016. Both (Times [2016]) and (CNN [2016]) reports, that Secretory Clinton did win the popular vote (65 845 000), but also failed to gain the 270 electoral votes- she gained only 48.1 % - 232 votes.

It is worth to mention that there are certain states that have a long history for voting for one of the parties, these states are known as the safe states. States that are teetering between parties are known as swing states. The most interesting swing states are probably Ohio and Florida. During the last 4 Presidential Elections, they both have given their electoral votes to the Democrats twice, and to the Republicans twice. It is worth to mention, that from 1904 to 2004, the candidate who won in Ohio won the presidency 24 of 26 times.

November 8, 2016 was the day of the Presidential Election in the United States of America. Based on the surveys, (Politico [2016]) names Colorado, Florida, Iowa, Michigan, Nevada, New Hampshire, North Carolina, Ohio, Pennsylvania, Virginia and Wisconsin as swing states. These states are having the sum of 146 electoral votes.

Table 3: The Distribution of Electoral Votes

State	Electors	State	Electors
Alabama	9	Montana	3
Alaska	3	Nebraska	5
Arizona	11	Nevada	6
Arkansas	6	New Hampshire	4
California	55	New Jersey	14
Colorado	9	New Mexico	5
Connecticut	7	New York	29
Delaware	3	North Carolina	15
District of Columbia	3	North Dakota	3
Florida	29	Ohio	18
Georgia	16	Oklahoma	7
Hawaii	4	Oregon	7
Idaho	4	Pennsylvania	20
Illinois	20	Rhode Island	4
Indiana	11	South Carolina	9
Iowa	6	South Dakota	3
Kansas	6	Tennessee	11
Kentucky	8	Texas	38
Louisiana	8	Utah	6
Maine	4	Vermont	3
Maryland	10	Virginia	13
Massachusetts	11	Washington	12
Michigan	16	West Virginia	5
Minnesota	10	Wisconsin	10
Mississippi	6	Wyoming	3
Missouri	10		

First of all, let us calculate the Banzhaf Index for each state. Using the method of generating functions the indices have been calculated. It is not surprizing, that California got the biggest ($0,114$), while states with only 3 Electoral Votes got the smallest ($0,005$) value.

One could easily suggest, that Pennsylvania (20) would be twice as powerful as Minnesota (10), since it has twice as many Electoral votes. Based on the Banzhaf indices, Pennsyl-

vania is 2.016 times more powerful. The same assumption could be made when thinking about Ohio (18) and Iowa (6). Using the same technique it can be easily calculated that Ohio is $3,024$ times as powerful as Iowa.
As the indices are finally calculated, it is worth to compare the number of electoral votes and the Banzhaf indices of safe states - both Republican and Democratic, and of swing states. This can be seen in the below table:

Table 4: The comparision of the power indices of swing and safe states

	Trump	Clinton	Swing	$\sum$
$\sum$ of Electors	191	201	146	538
Relative frequency of Electoral Votes	0,3550	0,3736	0,3048	1
$\sum$ of Banzhaf indices	0,3509	0,3811	0,2680	1

In his article (Miller [2013]), Miller states, that some see the existing system as a system which favors small states, while others see a large-state advantage. This could be partly because some thinks about the Electoral System as a system which favors small-states, while others think the opposite.
By taking the example of California and New Mexico, the following table can be created:

Table 5: Comparsion of the largest and one of the smallest states

State	E=Electors	P=Population (million)	$Q = \frac{P}{E}$	Banzhaf index
California	55	38,8	0,71	0,1136
New Mexico	5	2,09	0,41	0,0091

Where *E* marks the number of electors of the state and *Q* is shows the amount of population for each elector in the state.
It is important to mention, that the Banzhaf index of California is 12,5 times bigger, than New Mexico's, while the population of it is 18,5 times greater, and it has only 11 times as many electors as New Mexico.
Personally, I belive the above shows that the system - clearly speaking about the indices and the population- favors small-states.

4.3 The original Council of the European Economic Community

The Community was created by the Treaty of Rome in 1957 aiming to bring economic integration between its 6 founding members - France, West Germany, Italy, Belgium, The Netherlands and Luxembourg.
The Council was the body that held the executive and legislative powers of the Community, therefore it was the main decision making body of the ECC. The Council consisted of the ministers of each Country, and the presidency of the Council rotated between the members on a yearly basis.

For decision making the Treaty has defined 3 different majority types, among these are :

- *simple majority* - all members had the same amount of power,
- *qualified majority* - also known as weighted voting,
- *special majorities* - weighted voting based on the financial contribution to the respective Funds (either the European Social Fund or the Development Fund).

Unless declared otherwise, the Treaty also stated that the simple majority rule should be used (Lindberg [1963]). If the qualified majority rule is used (12 out of 17 votes), no single big power could have prevent nor block the work of the Council.

Using the generating functions the below η and β values are calculated:

Table 6: The Distribution of Votes in the ECC - Qualified Majority

Country	Votes	Population (thousand)	η	β
France	4	43 501	10	0,238
West Germany	4	54 064	10	0,238
Italy	4	48 229	10	0,238
Belgium	2	8 806	6	0,143
The Netherlands	2	10 823	6	0,143
Luxembourg	1	303	0	0

If using the qualified majority method, Luxembourg was never a swing player and its one vote could never have made a difference. It is worth to mention that France, West Germany and Italy had the same amount of vote, even though West Germany's population was $1,2$ times bigger than France's and $1,121$ times bigger than Italy's.

In case of using the special majorities rule, the required majority was 67 out of 100, preventing a single big power from blocking a decision.

Table 7: The Distribution of Votes in the ECC - Special Majorities - European Social Fund

Country	**Contribution(%)**	**E.S.F. weight**	η	β
France	32	32	14	0,304
West Germany	32	32	14	0,304
Italy	20	20	6	0,130
Belgium	8.8	8	6	0,130
The Netherlands	7	7	6	0,130
Luxembourg	0.2	1	0	0

Assuming that the Council made a decision using the weight based on contribution to the European Social Fund, France and West Germany were the most powerful countries - both being swing players in 14 cases. Italy's voting weight was more than twice as big as Belgium's and The Netherlands', but they were all equally powerful due to being swing players in 6 cases.
As previously, Luxembourgs' 1 vote never made a difference, therefore the real voting power of it was 0.

Table 8: The Distribution of Votes in the ECC - Special Majorities - Development Fund

Country	**Contribution(%)**	**Dev. weight**	η	β
France	34,41	33	15	0,357
West Germany	34,41	33	15	0,357
Italy	6,88	11	3	0,071
Belgium	12,04	11	3	0,071
The Netherlands	12,04	11	3	0,071
Luxembourg	0,22	1	3	0,071

In case of using the voting weight based on the contribution to the Development Fund, France and West Germany were the most powerful states, meaning they were swing players in 15 cases. Luxemburg has had only 1 vote based on the contribution, however the country itself was as powerful as Italy, Belgium or The Netherlands, as being swing players in 3 cases.

Based on how many times a country was a swing player using different voting methods, the below table is created.

Table 9: The real power of the ECC's members

Country	Qualified maj.	Spec. maj. - E.S.F.	Spec. maj. - D.F.	Sum
France	10	14	15	**39**
West Germany	10	14	15	**39**
Italy	10	6	3	**19**
Belgium	6	6	3	**15**
The Netherlands	6	6	3	**15**
Luxembourg	0	0	3	**3**

It can be easily noted how France and West Germany were the most powerful, while Luxembourg was the least powerful member of the Council of the European Economic Community.

4.4 The Council of the European Union

The Council of the European Union is declared in the Treaty of the European Union. The main function of the Council is to act as one of the two branches of the legislative branch along with the European Parliament. Therefore, the Council can be viewed as the main decision making body of the EU.
The members of the Council are the current presidents of the countries of the EU, meaning currently it has 28 members (EU [2016]).

Based on the matter the Council is voting about, the Council can vote:

- *simple majority* - 15 member states vote in favor
- *double majority* - 55% of all member states, representing at least 65% of the EU population vote in favor
- *unamous vote* - all votes are in favor

The Council is able to vote if the majoity of the members is present.

Table 10: The members of the European Union

Country	Population (thousand)	Country	Population (thousand)
Germany	81 090	Austria	8 582
France	66 352	Bulgaria	7 202
United Kingdom	64 767	Denmark	5 653
Italy	61 438	Finland	5 472
Spain	46 440	Slovakia	5 403
Poland	38 006	Ireland	4 626
Romania	19 861	Croatia	4 225
Netherlands	17 155	Lithuania	2 921
Belgium	11 258	Slovenia	2 063
Greece	10 847	Latvia	1 986
Czech Republic	10 420	Estonia	1 313
Portugal	10 375	Cyprus	847
Hungary	9 856	Luxembourg	563
Sweden	9 790	Malta	429

There is a so called alliance - Visegrad Group, often referred to as the Visegrad Four. It consists of 4 Central European states - Hungary, Poland, the Czech Republic and Slovakia. The aim of the alliance is to align the members' economical, diplomatic and political interest, and allignment of certain moves in the previously mentioned fields.
It is common, that the Visegrad Group makes a coalition before voting in the Council of the European Union, and the sum of the population of the 4 states is 60 343 802.

June 16, 2016 was an important day in the life of the European Union, as 52 $\%$ of UK voters voted in favor of *Brexit*, also known as British exit, Great Britain has decided to leave the EU by March, 2019. Even though the terms of withdrawal have not yet been fully negotiated, the United Kingdom remains a member of the EU.
Due to Brexit, the full population of the EU is estimated to decrease by 64 767 000, and the number of the member countries will fall to 27. Assuming that, the rule of double majority decision making won't be changed, 55% of all member states will drop from $15,4$ to $14,85$, as well as 65% of all EU's population will drop from 330 811 000 to 288 712 450- based on the latest data.

Expecting that the first tier of decision making (55% rule) will remain the same, requiring 15 ministers to vote in favor of the changes, would there be any change regarding the power of each country, and the power of the Visegrad Group?
Assuming that, the first tier of decision making is something we can ignore - as 55% of ministers will vote in favor of a decision, we can now move on to modeling the voting power of all member countries.

Table 11: Voting power of members of the EU - before and after Brexit

Country	**Before**	**After**	**Country**	**Before**	**After**
Germany	0,172	0,200	Austria	0,015	0,018
France	0,135	0,156	Bulgaria	0,013	0,015
United Kingdom	0,132	–	Denmark	0,010	0,012
Italy	0,124	0,143	Finland	0,010	0,011
Spain	0,091	0,103	Slovakia	0,010	0,011
Poland	0,074	0,083	Ireland	0,008	0,010
Romania	0,035	0,042	Croatia	0,007	0,009
Netherlands	0,031	0,037	Lithuania	0,005	0,006
Belgium	0,020	0,024	Slovenia	0,004	0,004
Greece	0,019	0,023	Latvia	0,004	0,004
Czech Republic	0,019	0,022	Estonia	0,002	0,003
Portugal	0,019	0,022	Cyprus	0,001	0,002
Hungary	0,018	0,021	Luxembourg	0,001	0,001
Sweden	0,017	0,020	Malta	0,001	0,001

It is important to mention, that the generating function defined in the used article - Leech [2002] was not properly specified. Therefore, the program that was written for calcualting the power indices was not able to run in case of having a really small or really big q. Also, having 28 players increased the runtime to hours. In order to still have a program that runs within minutes - the Monte Carlo method was used.

It is interesting to see that, Germany's population is almost 8 times bigger than Greece's, the voting power of the Germans' before Brexit is 9,05 times as much as Greeks'. Even though, Malta's population is 189 times smaller than Germany's, the voting power of it is only 172 times smaller.
One would easily think that the withdrawal of the UK would make much more difference

to the voting power of each country than it actually does. The voting power of all but 5 countries have increased after Brexit. If we take a look at the already mentioned connection between Germany and Greece- after Brexit- Germany is 8,1 times more powerful than Greece.
It is also interesting to see, that the voting power of countries with very low population- including Cyprus, Luxemburg and Malta- has not changed.
It is impressive, how the largest country, Germany with a population of 81 million is 166 times as powerful as the smallest country, Malta, with the population of 429 thousand. The quotient of the two populations is 189, which is significally bigger than 166.
Before Brexit, the Visegrad Group is representing $18,24\%$ of the population of the European Union. Using the same data, the 4 countries of the Visegrad Alliance will represent $20,09\%$ of the European Union's population after Brexit.
While calculating the power indices for the European Union, the Visegrad Group will be represented as a coalition.

Table 12: Voting power of the Visegrad Group as a coalition- before and after Brexit

	Before Brexit	**After Brexit**
The Visegrad Group as a coalition	0,126	0,144

Supposing, that the Visegrad Gruop really acts like a coalition, it can be seen as the fourth most decisive "player" before Brexit, and as the third after Brexit.

Let us take a look at the connection (quotient) of the voting power and the population of some countries - 2 really big ones, 3 medium ones and a small one. For this calculation the before Brexit values will be used.

Table 13: Connection between population and voting power

Country	**Population**	**Voting Power**	**Quotient * 10^6**
Germany	81 090	0,172	2,121
Italy	61 438	0,143	2,328
Belgium	11 258	0,020	1,776
Hungary	9 856	0,018	1,826
Ireland	4 626	0,008	1,729
Latvia	1 986	0,004	2,014

One can see, that the really large and the truly small countries are overrepresented by their voting power. This can be seen as the system favoring very big and really small countries over the medium sized ones.

5 The probability of events that have never occured

The probability of a single vote being decisive in the American Presidential Election is quite low. This probability represents one's influence in the Election which is compelling in a democracy.
Throughout this part of the thesis, the focus will be on the American system as well as the Electoral College,which is different from the previous parts. Here, the states will be modeled as individual states/jurisdictions and the following 2 situations will be studied:

1. *weighted voting* which has been already introduced earlier
2. *two- stage voting* in this case the voters are divided into J jurisdictions/districts/states etc. The winner within each jurisdiction is decided by a simple majority rule. The overall winner will be decided by the weighted majority of the winners in the jurisdictions.

The voting power here is defined the following way:

Definition 5.1 (voting power) *The possibility that a given voter can affect the outcome of an election.*

Assuming that v_i is the vote of an individual voter, where $1 \leq i \leq n$ and $v_i = \pm 1$. Vector v is representing the set of all v_i values, therefore $v = (v_1, \dots, v_n)$, aggregating the votes of n individuals into 1 single outcome, let us introduce $R = R(v) = \pm 1$. There is no need to define the probabilitical distribution of R and in this the voting power of player i is equal to the number of coalitions of the other $n-1$ players, where i is a swing player. Assuming that p_i denotes the voting power of player i, it can be calculated as:

$$p_i = P(R = 1 | v_i = 1) - P(R = 1 | v_i = -1) \tag{16}$$

In case $p_i = 0$, player i does not have any influence on the outcome.
As already mentioned, the weighted voting and the two-sage voting systems will both be modeled throughout this part of the thesis.
In case of the simple weighted voting, w_i marks the weight of player i and R can be denoted as: $R(v) = \text{sign}\left(\sum_{i=1}^{n} w_i v_i\right)$. If the voting has two stages, the jurisdictions have their own voting weigth, which will be marked by w_j.
It is important to note that votes and ties on all levels are decided by flipping coins.

Based on Gelman (Gelman et al. [2002]) the most broadly publicated statement regarding two staged voting system is that larger jurisdictions have disproportionate power,

and the power is roughly proportional to $\sqrt{n_j}$. The conclusion is that the Americal Electoral Collage favors large states.
The probabolity of a tie election (TE) can be estimated as:

$$P(TE) \approx \frac{f_V(0)}{n} \tag{17}$$

where V marks the proportions between the two leading candidates of an election, and f_V denotes the density function of V.
As an election has many factors that might influence the outcome of the election, it is almost impossible for a single vote to be decisive. However, it is still possible to affect the outcome and this can be identified with the probability of a tie to a very close approximation.

5.1 Random Voting Model

Assuming, that voters vote by flipping coins -this is the so called random model- the most basic model in power voting literature is found. It is important to mention how this model is not the most accurate in case of modeling the election.

5.1.1 Simple Weighted Voting

Let us introduce the of V_{-i} which is the combined voting weight of all players in N excluding i.

$$V_{-i} = \sum_{k \neq i} w_k v_k$$

Furthermore, defining the sum of square of all voting weights is essential.

$$W^2 = \sum_{k=1}^{n} w_k^2$$

$$p_i = P(|V_{-i}| < w_i) + \frac{1}{2} P(|V_{-i}| = w_i) \tag{18}$$

Based on the specialities of the random voting model, the following values can be also calculated:

$$E(V_{-i}) = 0$$
$$D(V_{-i}) = \sqrt{\sum_{k \neq i} w_k^2} = \sqrt{W^2 - w_i^2}$$

Where the calculation of $D(V_i)$ is explained below:

$$\begin{aligned} D^2(V_{-i}) &= D^2(w_1v_1 + \cdots + w_{i-1}v_{i-1} + w_{i+1}v_{i+1} + \cdots + w_nv_n) \\ &= D^2\left(\sum_{k=1}^{n} w_kv_k - w_iv_i\right) = D^2\left(\sum_{k=1}^{n} w_kv_k\right) + D^2(w_iv_i) - 2cov\left(\sum_{k=1}^{n} w_kv_k, w_iv_i\right) \\ &= \sum_{k=1}^{n} w_k^2D^2v_k + w_i^2D^2v_i - 2w_i^2D^2v_i = \sum_{k=1}^{n} w_k^2 + w_i^2 - 2w_i^2 = W^2 - w_i^2 \end{aligned}$$

as $cov(v_i, v_i) = D^2v_i = Ev_i^2 = 1^2\frac{1}{2} + (-1)^2\frac{1}{2} = 1$.

As the number of voters (players) in the U.S. Presidential Election is enormous, one can easily note how no single voter or small set of voters is dominant, also there is no discrete feature in the weights, therefore the distribution of V_{-i} can be approximated as a normal distribution. Therefore the p_i can be approximated by the following calculation, assuming Φ is the cumulative normal distribution function:

$$\Phi\left(\frac{w_i}{\sqrt{W^2 - w_i^2}}\right) - \Phi\left(\frac{-w_i}{\sqrt{W^2 - w_i^2}}\right)$$

The voting power is almost linearly close to w_i in case $w_i^2 << W^2$. A good example is the Electoral College of the USA. As mentioned in one of the previous sections, the electoral vote varies between 3 and 55 for each state, with the sum of 538. The value of p_i can be calcualted for each and every one of them, and also, they can be compared to the linear approximation: $p_i \approx \sqrt{\frac{2}{\pi}\frac{w_i}{W}}$, which is the probability density function of the normal distribution.

For a better understanding, on how the above p_i approximation is calculated, please see the below:

$$p_i = P(|V_{-i}| < w_i) = P(-w_i < V_{-i} < w_i)$$

which can be exracted the following way:

$$\Phi\left(\frac{w_i}{\sqrt{W^2 - w_i^2}}\right) - \Phi\left(\frac{-w_i}{\sqrt{W^2 - w_i^2}}\right) = 2\Phi\left(\frac{w_i}{\sqrt{W^2 - w_i^2}}\right) - 1$$

by using the central limit theorem. In case $w_i^2 << W^2$ then

$$\left(\frac{w_i}{\sqrt{W^2 - w_i^2}}\right) \approx \frac{w_i}{W}$$

therefore

$$2\Phi\left(\frac{w_i}{\sqrt{W^2 - w_i^2}}\right) \approx 2\Phi\left(\frac{w_i}{W}\right) - 1.$$

Since

$$\Phi(x) \approx \frac{1}{2} + \frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}\left(x + \frac{x^3}{3}\right)$$

from which $\frac{x^3}{3}$ can be left out, therefore the formula can be rewritten as

$$2\Phi\left(\frac{w_i}{W}\right) - 1 \approx 2\left(\frac{1}{2} + \frac{1}{\sqrt{2\pi}}e^{-\frac{{w_i}^2}{W^2}}\frac{w_i}{W}\right) - 1 = 1 + \sqrt{\frac{2}{\pi}}1\frac{w_i}{W} - 1 = \frac{2}{\pi}\frac{w_i}{W}$$

5.1.2 Two-stage Voting

For computing probabilities in a two-stage voting, the compution of the voting weight of the jurisdictions and then of each n_j player in the district is necessary. In case the number of the district is large, and none of them is dominant, the voting power of district j will be noted by w_j.

The probability of player i being decisive in jurisdiction j can be calculated as it follows:

$$P(V_{-i} = 0) + \frac{1}{2}P(|V_{-i}| = 1) = \frac{2}{2^{n_j}} \cdot \begin{cases} \binom{n_j-1}{(n_j-1)/2} & n_j = 2k+1 \\ \binom{n_j-1}{n_j/2} & n_j = 2k, k \in \mathbb{N} \end{cases}$$

The above can be approximated by using the Stirling's approximation $\sqrt{\frac{2}{\pi n_j}}$ in cases where n is not too small. One can easily note, this is a special case of (16.).

For better understanding, please see the Stirling's approximation and some calculations:

$$n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$$

$$\binom{n_j - 1}{\frac{n_j-1}{2}} = \frac{(n_j-1)!}{\left(\left(\frac{n_j-1}{2}\right)!\right)^2} \approx \frac{\frac{(n_j-1)^{n_j-1}}{e^{n_j-1}}\sqrt{2\pi(n_j-1)}}{\frac{\left(\frac{n_j-1}{2}\right)^{2\frac{n_j-1}{2}}}{e^{2\frac{n_j-1}{2}}}2\pi\left(\frac{n_j-1}{2}\right)} = \frac{2^{n_j-1}2}{\sqrt{2\pi}\sqrt{n_j-1}}$$

The advantage of using the random voting model is that all votes inside and outside of a jurisdiction are independent, therefore

$$p_i \approx \frac{w_j}{\sqrt{n_j}}.$$

which let us calculate the voting power of player i from jurisdiction j, which leads us to the conclusion of assuming that the allocation of the weights in a system that has two - stages is corresponding with the square root of the number of voters in a jurisdiction.

5.1.3 Voting power of individuals, and average voting power

Each voting system is unique, and they have different affects on the voting power of individuals in the system, as well as on the average voting power. Based on Gelman - Gelman et al. [2002], below an example can be seen regarding the positive and negative

effects of the random voting model. For this example, the simple majority voting and a two-staged voting system will be used, with 9 voters.
The total effect of the two-stage voting model can be studied by calculating the average voting power of all players. The average voting power is maximized in the random voting model while using the simple majority rule. In contrast to that, it is lower under any other system, meaning the average power of decisiveness is less.
In case of recalculating the voting power using the probability of satisfacition, which is defined as "the voter's preferred candidate winning the election", p_i can be rewritten as:

$$p_i = P(R = +1|v_i = +1) + P(R = -1|v_i = -1) - 1 = 2P(R = 1 \text{ and } v_i = -1) - 1$$
$$= 2P(\text{voter i is satisfied}) - 1$$

The above can be understood as maximizing satisfaction and maximizing average voting power is the same, as only the voters on the winning side will be satisfied.
As already mentioned, Gelman calculated (Gelman et al. [2002], pg. 425) the individual voting power and the average voting power for 9 players in a two-staged, simple majority rule voting system. In this model, he calculated the values in 4 different cases : *no coalitions, a single coalition with 5 players, a single coalition with 3 players*, and *3 coalitions with 3 players each.*

1. no coalitions
   - a player is decisive if the other voters are split 4-4 $\Longrightarrow P(\text{player is decisive}) = \binom{8}{4}2^{-8} = 0,273$
   - average $P(\text{player is decisive}) = 0,273$

2. a single coalition with 5 players
   - a player in the coalition is decisive if others in the coalition are split 2-2 $\Longrightarrow P(\text{player is decisive}) = \binom{4}{2}2^{-4} = 0,375$
   - a player not in the coalition can never be decisive, therefore: $\Longrightarrow P(\text{player is decisive}) = 0$
   - average $P(\text{player is decisive}) = \frac{5}{9} \cdot 0,375 + \frac{4}{9}(0) = 0,208$

3. a single coalition with 3 players
   - a player in the coalition is decisive if others in the coalition are split 1-1 and the coalition is decisive $\Longrightarrow P(\text{player is decisive}) = \frac{1}{2}\frac{50}{64} = 0,391$
   - a player not in the coalition is decisive $\Longrightarrow P(\text{player is decisive}) = \binom{5}{1}2^{-5} = 0,156$
   - average $P(\text{player is decisive}) = \frac{3}{9}(0,391) + \frac{6}{9}(0,156) = 0,234$

4. 3 coalitions with 3 players each
   - a player is decisive if others in the cialition are split 1-1 and the other two coalitions are split 1-1 $\Longrightarrow P(\text{player is decisive}) = \frac{1}{2}\frac{1}{2} = 0,250$
   - average $P(\text{player is decisive}) = 0,250$

The average conclusion is that the average probability of a voter being decisive is less than or equal to the probability of a player being decisive in general.

5.1.4 Empirical failings of the random voting model

As already mentioned, the random voting model is not the most appropiate and utterly unrealistic when it comes to modeling an election. Therefore it is essential to create a more realistic, a more complex model.
The random voting model makes predictions that are not valid when it comes to a real election. For example, the random voting model implies that the preoportional vote margin should gave a mean of 0 and a standard deviation of $0,001$. However, if we take a look at real elections - the case of Al Gore and George W. Bush - it is easily noted, how even a close election has a difference bigger than 1% . In other words, it is truly unrealistic to have an election which is that close.
Based on previous calculations, in case an election is less close than it is assumed by the random voting model, the power of individuals decreases and the results of relative voting power is also affected.
Having a voting district which is having a moderate or large size, the probability of a vote being decisive within district j can be approximated by using the below equation while n_j marks the number of voters within district j:

$$P(\text{a vote is decisive within district j }) \approx \frac{f_j(0)}{n_j}$$

and f_j marks the distribution of the proportional vote differential $V_j = \frac{1}{n}\sum_{i=1}^{n_j} v_i$ within district j.Supposing that f_j has a mean of 0 and a fixed distributional form, then:

$$P(\text{a vote is decisive within district j}) \approx \frac{1}{n_j \cdot sd(V_j)}$$

Thus, when comparing the probability of decisiveness for voters in districts of different sizes n_j the behavior of $sd(V_j)$ as a function of n_j is crucial.

Presidential Election in the U.S. is an already mentioned example of power voting. The relative power of voters in different states vary based on the number of the electors of each state. As already discussed in section (4.2), many people have different thoughts about the size of favored states when it comes to the Election. Based on Gelman, the general approach of most of the publishers *(Shapley, Banzhaf and Mann)* is that the current system favors larger states.

The $\frac{1}{\sqrt{n_j}}$ rule means that the probability of decisiveness of a voting district is proportional to the already mentioned value. It is assumed that the extra power of voters in larger states derives from the assumption that elections in these states are much more likely to be close, and of curse, the assumption is tested in Gelman et al. [2002]. The conclusion is that the rule of $\frac{1}{\sqrt{n_j}}$ is impossible to hold in practice, as it would mean extreme landslides for low n_j or extremely close elections for high n_j.

5.2 Stochastic modeling of voting power

It is already discussed in the thesis, that the random voting model where voters get their votes assigned by flipping coins is not the most accurate one. In his article Gelman (Gelman et al. [2002]) states, that using different methods of assigning votes to players might lead us to having a better approach to a more realistic model. The simplest generalization of random voting models having more exact values than the previously known coin flipping method would be the following one: assuming that the voters are still independent, but instead of with the probability of $\frac{1}{2}$, now with the probability of p of voting as $+1$. Even though this model is not useful, it corrects the mean value. Furthermore, granting each player the opportunity of having a separate p_i and to model the distribution of these values is necessary.
The already declared $\frac{1}{\sqrt{n}}$ rule does not apply anymore, and is replaced by a special form of the probability of a vote being decisive in a jurisdiction, due to the binomial variation from which it derives is minor compared to any realistic variation among the probabilities of p_i.
A dependence structure is also needed to be given for the voters' probabilities p_i. This dependence is built up based on already existing relations between the voters, and one of the most obvious starting point of building up such a relationship would be the geographical or demographical relation.
For understading the basics of the stohastical model, Gelman derives some of the basic results he got from the usage of the Ising model on symmetric trees. Supposing the number of voters is $n = d^k$, which is arrayed in a tree of depth d with k branches at each node, where the variables are equal to ± 1. The leafs are representing the votes, and the k children are independent having the probability of π being different from the parent. In this model, it is also assumed, that the marginal probabilties of $+1$ and -1 are equal.

6 Summary

Throughout this thesis, first the 3 power indices - among these are the Banzhaf, the Coleman and the Shapley-Shubik indices - and the term *voting power* were defined. The basic differences between the Shapley-Shubik and the Banzhaf indices were introduced.
During the following section, the method of generating functions was presented using a simple example and was used for calculating the power indices for different systems, such as the: Presidential Election of the United States of America, the original Council of the European Economic Community and the Council of the European Union.
Later on, based on a Gelman article the calculation of the probability of a player being decisive has been introduced.
It is important to mention, that there are many interesting articles that gave me a better understanding of how power voting works, and that would have been a great base for this thesis.

2016 was a truly interesting year - Leonardo DiCaprio won an Oscar, Apple's iPhone 7 was released, NASA's Juno spacecraft successfully entered Jupiter's orbit and the 31st Summer Olympics was held. Let us also mention two major events that are related to power voting: The Presidential Election of the U.S.A. and Brexit. These two events allowed me to read many up to date articles about how different power voting systems work.

Power voting is a special field in mathematics. Specialists have to create a model that works for estimating voting power, and that is also able to represent the system in such way, that the power of individuals can be calculated with only minor errors. Voting power is important in many ways it helps us to model and study political relations, strategy and fairness, and to calculate the probability of events that have never occured.

In the future, I would be interested in digging into this field much deeper, and to continue my research about the probability of an individual being decisive in the Presidential Election of the United States.

Finally, I would like to close my thesis with the following quote:

> *Researchers sometimes argue that statisticans have little to contribute when few realizations of the process being estimated are observed. We show that this argument is incorrect (...).*

A. Gelman, G. King and W. J. Boscardin, (Gelman et al. [1998])

7 Appendix

List of Tables

References

CNN. It's official: Clinton swamps trump in popular vote. `http://edition.cnn.com/2016/12/21/politics/donald-trump-hillary-clinton-popular-vote-final-count/`, 2016. Accessed: 2016-12-22.

Council EU. Voting system. `http://www.consilium.europa.eu/en/council-eu/voting-system/`, 2016. Accessed: 2016-12-31.

Andrew Gelman, Gary King, and W John Boscardin. Estimating the probability of events that have never occurred: when is your vote decisive? *Journal of the American Statistical Association*, 93 (441):1–9, 1998.

Andrew Gelman, Jonathan N. Katz, and Francis Tuerlinckx. The mathematics and statistics of voting power. *Statistical Science*, pages 420–435, 2002.

Dennis Leech. *Computation of power indices*. Warwick Economic Research Papers, 2002.

Leon N. Lindberg. *The political dynamics of European economic integration*. Stanford University Press, 1963.

Nicholas R. Miller. A priori voting power and the us electoral college. In *Power, Voting, and Voting Power: 30 Years After*, pages 411–442. Springer, 2013.

Politico. Politico. http://www.politico.com/, 2016. Accessed: 2016-12-31.

The New York Times. Clinton wins popular vote by nearly 2.9 million. http://www.nytimes.com/aponline/2016/12/22/us/politics/ap-us-clinton-popular-vote.html, 2016. Accessed: 2016-12-22.

Wikipedia. Banzhaf power index. https://en.wikipedia.org/wiki/Banzhaf_power_index, 2016a. Accessed: 2016-12-22.

Wikipedia. Shapley-shubik power index. https://en.wikipedia.org/wiki/Shapley%E2%80%93Shubik_power_index, 2016b. Accessed: 2016-12-22.