Journal of Social and Political Sciences

The Asian Institute of Research *Social and Political Sciences* is a peer-reviewed International Journal. The journal covers scholarly articles in the fields of Social and Political Sciences, which include, but not limited to, Anthropology, Government Studies, Political Sciences, Sociology, International Relations, Public Administration, History, Philosophy, Arts, Education, Linguistics, and Cultural Studies. As the journal is Open Access, it ensures high visibility and the increase of citations for all research articles published. The *Journal of Social and Political Sciences* aims to facilitate scholarly work on recent theoretical and practical aspects of Social and Political Sciences.

Benford's Law Application: Case of Elections in Sri Lanka

Agim Kukeli[1], Hettiyadura Shehan Karunaratne[2]

[1] Department of Economics, Entrepreneurship, and Finance, Barleti University, Tirana, Albania
[2] Dillard College of Business Administration, Midwestern State University, Texas, USA

Correspondence: Agim Kukeli, Department of Economics, Entrepreneurship, and Finance, Barleti University, Tirana, Albania. Tel: -. E-mail: dr.akukeli@gmail.com

Abstract

Sri Lanka, like many developing countries has been involved in a circle of allegations of election fraud. Usually these claims are pronounced more by losing parties. This study uses Benford's law, a law of probability distribution of digits, to investigate whether the election fraud claims might have merit. A sample in this study is made of 808 election data. This data comes from the 2010 Presidential election for representatives from three major political parties and from 2010 General Election data. All of the data points were obtained through reliable government sources, two of which are, the Department of Elections website of Sri Lanka, and the National News Paper statistics (2010). The study contrasts the distribution of the first digit of election results against the Benford's Law benchmark. After obtaining the results, we organize the data and find median, mean, mode and standard deviation. The preliminary results showe that the data does not align with Benford's law predictions. In other words, it shows that the data does not follow the law where the mean is larger than the median and there is a positive skewness then it likely follows a Benford's distribution. The distribution of the first digit of actual data for three parties disagrees with Benford's law. This misalignment is more pronounced for the winning party than for the second and third place parties, respectively. We, therefore, look forward to run the data through several critical analyses and observing if there shall be any fraud or manipulation in numbers.

Keywords: Benford's Law, Probability Distribution, Sri Lanka, Elections Fraud

1. Introduction

Allegations of electoral fraud seem a prevalent component of democratic process. Fraudulent elections and disputes about election outcomes have always been existent. In this paper, we illustrate the implementation of Benford's Law to election results in the 2010 Presidential and General Elections of Sri Lanka. An island surrounded by the Indian Ocean, Sri Lanka recorded a total population of 20,810,816 (Department of Census and Statistics Sri Lanka, 2010), with a GDP of $49,570,000,000 USD (Central Bank of Sri Lanka, 2010) and a per capita GDP of $2,400.02 USD (Central Bank of Sri Lanka, 2010). The country is mainly segregated into 25 districts which are organized within 09 provinces.

After gaining independence from the British in 1948, Sri Lanka was ruled by its own government having the United National Party as its major political party. After the inception of several other minority parties, the Sri Lankan Freedom Party (SLFP) was then formed by Hon. S.W.R.D Bandaranayake in 1951. Ever since, the political power in Sri Lanka, and the role of the government has been borne hand in hand by these two parties; the United National Party and the Sri Lanka Freedom Party. In the year 2009, under the governance of the SLFP and the bravery of the Sri Lankan forces, the 30 year civil war between the Sri Lankan government and the Liberation Tigers Tamil Eelam or commonly known as the L.T.T.E came to an end. The country itself suffered losses of a great number of resources due to the war as well as the corruption in government. After May 18th, 2009, which was when the L.T.T.E was destroyed, the president of Sri Lanka decided to have the 2010 presidential election seeking a fresh mandate prior to the expiration of then his term in 2011. The election was announced on 23 November 2009. Nominations were accepted on 17 December 2009, and the election was held on 26 January 2010. Rajapaksa, who was elected president for a 6-year term in November 2005, was the candidate of the ruling United People's Freedom Alliance. General Sarath Fonseka, a former commander of the Sri Lankan Army, was his main opponent in the election. Fonseka had been endorsed by a number of main opposition parties, including the United National Party and the Janatha Vimukthi Peramuna.

President Rajapaksa was able to win re-election, over 57% of all votes cast. Fonseka received over 40%, carrying the Northern and Eastern Provinces of Sri Lanka. The remaining twenty candidates all garnered less than 0.5% of the popular vote (Presidential Elections).

2. Violations

Violence, misuse of state resources, and other violations of election laws have been a part and parcel of Sri Lankan politics. The Sri Lankan police received nearly 975 election-related complaints, and 375 arrests were made (Daily Mirror, 2010). The Centre for Monitoring Election Violence (CMEV) had recorded 809 incidents from 23rd November 2009, when the elections were officially announced, to 25 January 2010. People's Action for Free and Fair Elections (PAFFREL) has recorded 666 incidents between 17 December 2009 and 25 January 2010. The UN, United States, and EU expressed concern over the level of violence. Election monitoring groups stated that this election was the most violent in more than 20 years. Although most of the allegations were pinned on Rajapaksa's supporters, supporters of both main candidates were blamed for the incidents as well. However, government politician Wimal Weerawansa accused the opposition of executing violence against their own supporters and meetings in order to blame Rajapaksa's benefactors. According to the National Polls Observation Center, criminal gangs and deserters from the army were employed to rouse violence.

The government took measures in order to have tight security on the day of the election, deploying over 68,000 policemen to keep the polls safe, with the Sri Lankan Military providing additional support (CMEV). The election monitoring group Centre for Free and Fair Elections deployed 6,500 monitors on Election Day, with the PAFFREL deploying a further 6,000, including 14 foreign monitors. Possible post-election violence was denied by the Sri Lankan Government and claimed it as false rumors, with Foreign Minister Rohitha Bogollagama saying, "I don't think the people of Sri Lanka have time for street protests. It has never happened."

According to Inspector General of Police Mahinda Balasuriya, Election Day was largely peaceful, with a few minor incidents reported while there was no report of any major incidents that occurred during the election. A number of explosions, however, were heard in Jaffna and Vavuniya in the north of the country during the day. There were hundreds of violent incidents during the run-up to the election. Officially, there have been four murders related to the election (BBC, 2010):

-   12 January 2010– Kusumawathie Kuruppuarachchi, a 58-year-old mother was shot dead when gunmen on motorbikes opened fire on a bus carrying Fonseka supporters in Hungama, Hambantota District. (BBC, 2010)
-   16 January 2010– Saman Kumara, a Rajapaksa supporter, was shot dead in a clash between supporters of the two main candidates in Madurankuliya, Puttalam District. (Sunday Times, 2010)
-   18 January 2010– Dhammika Herath, a 33-year-old businessman, was beaten to death whilst pasting

          posters supporting Fonseka in Wariyapola, Kurunegala District. (Palihawadena, 2010)
- 18 January 2010– D.M. Suranga Indrajith, a 27-year-old driver for government minister Jayarathna Herath, was killed in a grenade attack in Vanduragala, Kurunegala District. (Daily Mirror, 2010)

The CMEV has linked a fifth death to the elections:
- 19 January 2010– Lingeswaran, a Fonseka supporter, was found dead on the Kandy-Matale railway line in Warapitiya, Matale District (Saravanamuttu, Jayasekera, & Kalupathirana).

Misuse of State Resources

There were claims of extensive misuse of state property during the election campaign. State-owned institutions paid for numerous advertisements supporting Rajapaksa, while public officials, state owned buildings and vehicles were allegedly used for Rajapaksa's campaign (Abeywickrema). Government television stations gave far-reaching coverage to Rajapaksa's campaign, with little mention of other candidates. Elections Commissioner Dissanayake condemned the government for not preventing the misuse of state resources.

3. Aplication of Law

3.1 Benford's Law

Benford's law, also called the first-digit law, is an observation about the frequency distribution of leading digits in many real-life sets of numerical data. In 1881, Simon Newcomb, an astronomer and mathematician, published the first known article describing what has become known as *Benford's Law in the American Journal of Mathematics.* Newcomb calculated that the probability that a number has any particular non-zero first digit is (Durtchi, Hillison, & Pacini):

Formula 01: $P(d) = Log_{10}(1 + (\frac{1}{d}))$, Where: d is a number 1,2 ….9, and P is the probability

Applying the above formula we are able to determine the probability of a certain number being the first, second, third or the ninth digit of a certain number distribution. Table 1 shows the determined probabilities for all digits 0 through 9 in the first place in any number.

Table 1: Probabilities of the first digit

First Digit	Probability
1	0.301
2	0.176
3	0.125
4	0.097
5	0.079
6	0.067
7	0.058
8	0.051
9	0.046

Source: Authors' own calculation

The above probabilities have been determined by applying the (Formula 01). That is,
    Formula (01) application: For first digit of the number,
    Probability $(D_1 = d_2) = log(1 + (1/d_1))$; $d_1 = (1, 2, 3 …9)$

Where, D_1 represents the first digit of a number.

D_2 represents the second digit of a number, etc.

In order to arrive at a better understanding, we can illustrate the above probabilities in a graph.

Graph 1: Benford's distribution
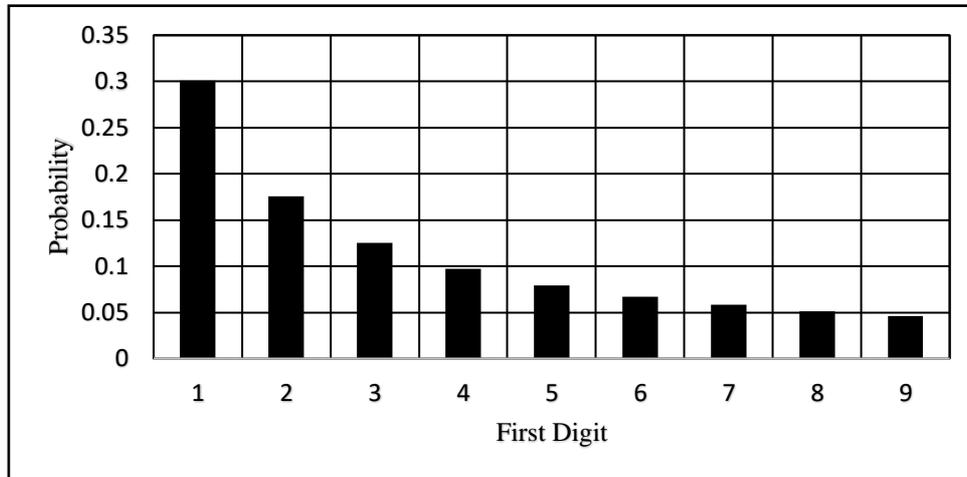
3.2 Benfor's Law Applied to the Elections in Sri Lanka

There are 646 data points that are taken into consideration from the 2010 Presidential Election and 2010 Parliamentary Election. The presidential election consists of 323 data points which were the popular votes obtained by each presidential candidate from each city. In the general election, 326 data points are taken into consideration, which are also the popular votes obtained by each party from each city.

When applying Benford's formula, we also should keep in mind that, if the mean of a particular set of numbers is larger than the median and the skewness value is positive, the data set likely follows the Benford distribution (Durtchi, Hillison, & Pacini, 2004).

3.2.1 *Benford's Law Applied to 2010 Presidential Election*

After being able to gather 323 data points which consist of each candidate's popular votes won by each city, we were able to construct the following table.

Table 2: Data for 2010 Presidential Elections

| Digits	Benford's Law	Candidate 01				Candidate 02			
		000'	0000'	Total	Actual Distribution	000'	0000'	Total	Actual Distribution
One	0.3	1	12	13	0.079754601		47	47	0.29375
Two	0.176	2	30	32	0.196319018		70	70	0.4375
Three	0.124	2	40	42	0.257668712	1	29	30	0.1875
Four	0.096	5	32	37	0.226993865	1	3	4	0.025
Five	0.079	2	15	17	0.104294479		2	2	0.0125
Six	0.066	2	9	11	0.067484663	1	1	2	0.0125
Seven	0.057	1	5	6	0.036809816	1	1	2	0.0125
Eight	0.051	3	1	4	0.024539877	3		3	0.01875
Nine	0.045	1	0	1	0.006134969			0	0
Total				163				160	

The second column in (Table 2) shows Benford's Law distribution, while the "actual distribution" is determined from the election results of both candidates. It is obvious that Benford's Law does not agree with the actual distribution of both candidates. However, we are also able to see that the actual distribution of candidate 02 (the losing candidate) is somewhat closer to the Benford's distribution than the candidate 01's (the winning candidate) actual distribution.

This relationship can be more clearly observed when we plot this data on a graph. The following two graphs consist of the Benford's Distribution line and the Actual Distribution line of each candidate.

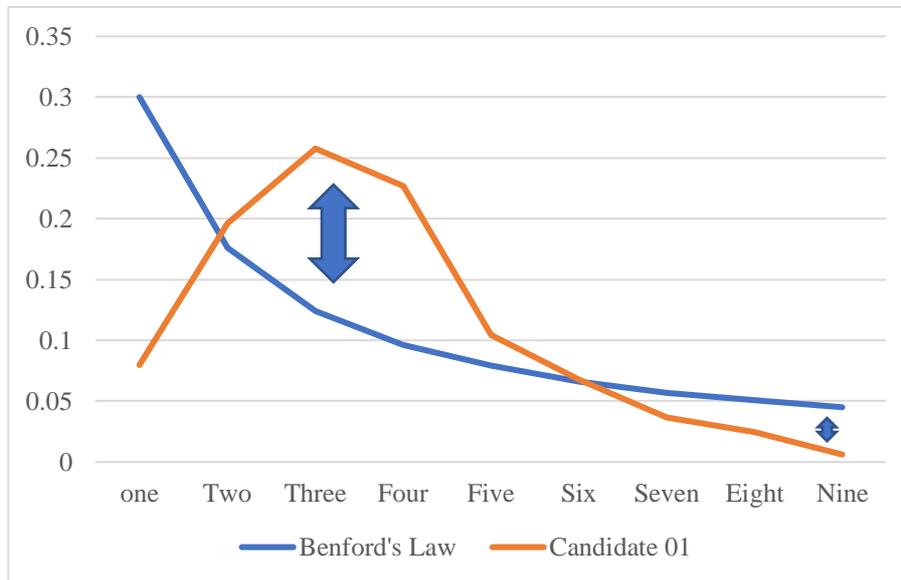Graph 2: Candidate 01, distribution of the 1st digit
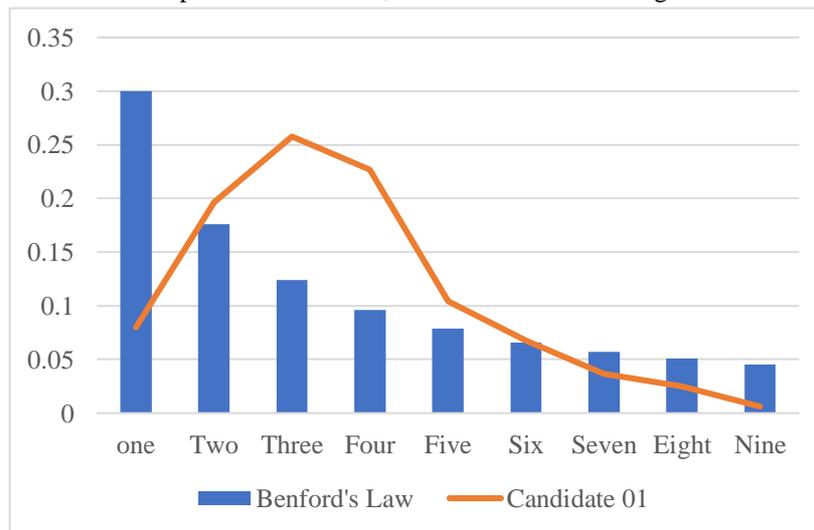
Graph 03: Candidate 02, distribution of the 1st digit
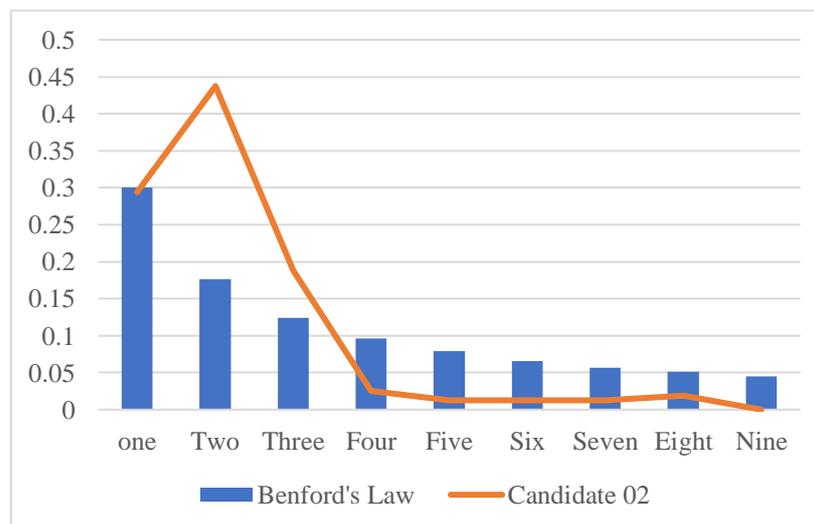
In the above graphs, the blue bars represent the Benford's Law distribution of the first digit, while the orange line represents the actual distribution. Therefore, we can clearly observe the discrepancy between the two relationships. It is also clear how candidate 01's distribution of the first digit exhibits a large dissimilarity between Benford's Law and the first digit being 1 through 5, and the 2nd candidate's distribution shows a much closer relationship between the Benford's Law and the actual distribution. This observation explains itself since, if the candidate 01 who is also the winning party by any chance manipulated any of the numbers it will obviously make an effect in the number of votes of the second candidate since the population or votes counted is a constant factor.

We also can illustrate the following data in a line graph which will mark the discrepancy area. This will help to get a better idea about the difference in the relationship of the distribution data.

Graph 04: Candidate 01, distribution of the 1st digit with discrepancy

Graph 05: Candidate 02, distribution of the 1st digit with discrepancy
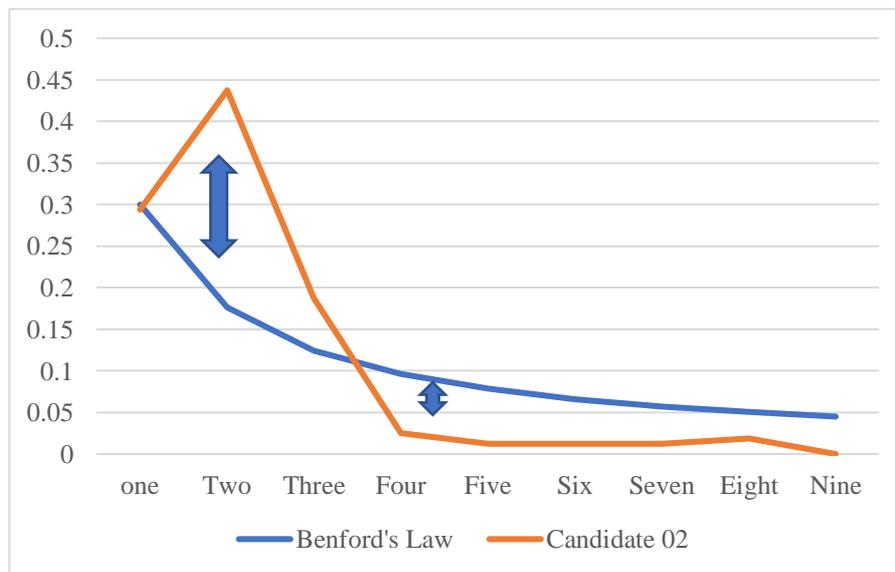
The blue arrows illustrate the discrepancy areas of which we should pay attention to. These gaps tell us that there has been some sort of manipulation in the election result numbers while backing up the claims for election fraud done by critics. These digital analyses can be further conducted by applying the same rule to the 2nd and 3rd digits in the sample distribution. The formulas for the 2nd and 3rd digit distribution probability is as follows.

Formula (02): For second digit of the number,
Probability $(D_2 = d_2) = \sum_{d1=1}^{9} log\ (1 + (1/d_1\ d_2))$; $d_2 = (1, 2, 3 \ldots 0)$
Where, D_1 represents the first digit of a number. D_2 represents the second digit of a number, etc.

3.2.2 *Benford's Law Applied to 2010 General Election*

Table 3: Data for 2010 General Election

Digits	Benf's	UPFA				UNF				DNA				
		000'	0000'	total	UPFA	000'	0000'	Total	UNF	00'	000'	0000'	Total	DNA
One	0.30		21	21	0.13	4	96	100	0.62	1	46	9	56	0.35
Two	0.18	4	58	62	0.38		18	18	0.11	1	32	1	34	0.21
three	0.12	5	35	40	0.25	5	3	8	0.05		14	1	15	0.09
four	0.10	3	18	21	0.13			0	0.00		12		12	0.07
Five	0.08	2	9	11	0.07	2	2	4	0.02	3	16		19	0.12
Six	0.07	1	1	2	0.01	6		6	0.04	4	3		7	0.04
seven	0.06	2	1	3	0.02	6		6	0.04	4	8		12	0.07
eight	0.05	1	0	1	0.01	11		11	0.07	3	3		6	0.04
Nine	0.05	1	0	1	0.01	8		8	0.05	1			1	0.01
Total				162				161					162	

Since we have applied Benford's Law theory to the 2010 Presidential Election data, we will be applying it to the 2010 General Election Data. This contains 485 data points which represents the popular vote won by three political parties in each city. (Table 3) will show Benford's Law distribution and the actual distribution of the first digit of the election results.
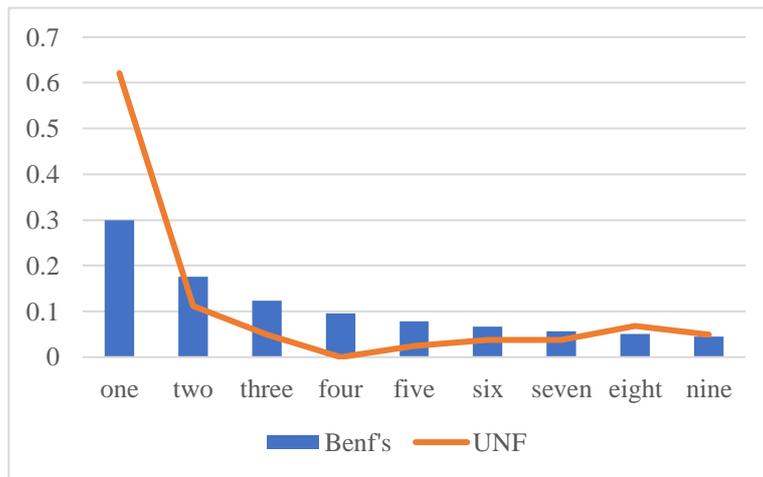
(Table 3) shows us the distribution of the first number digits of the election results of three major parties. We can quite evidently see that the UPFA (the winning party) has no relationship or in other words, the biggest discrepancy between Benford's distribution and its actual distribution. The UNF and the DNA too exhibit a difference in the distribution probability. However, just as the previous time, we can see the least powerful party (DNA) has a much closer relationship with Benford's distribution than the other two parties.

We can further plot the above data on a graph in order to observe the difference between the actual and Benford's law distribution.
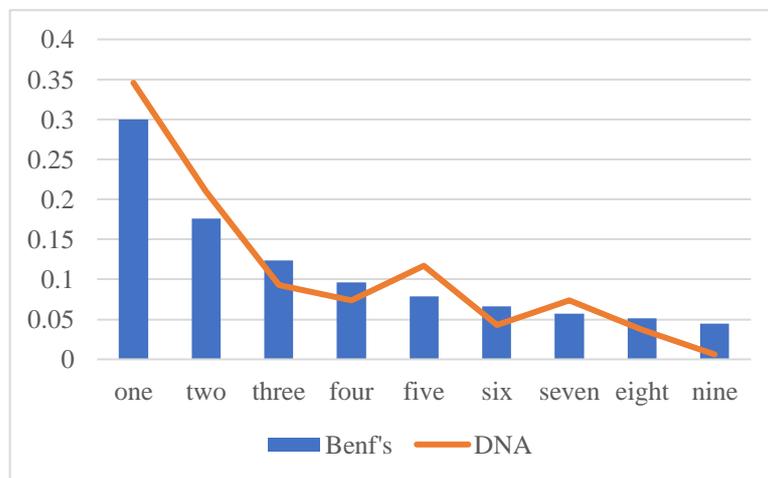
Graph 06: UPFA - distribution on the 1st digit
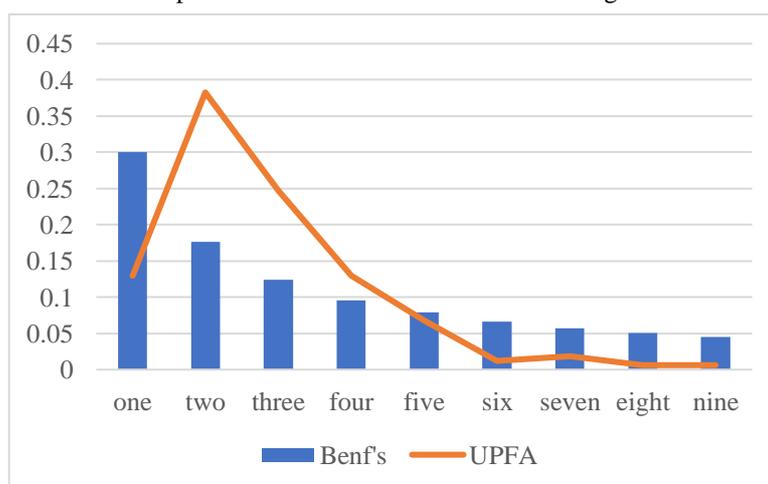
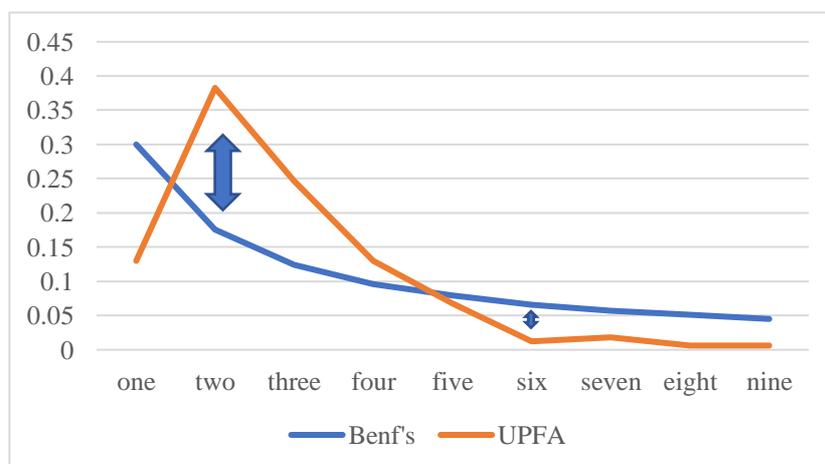Graph 07: UNF - distribution on the 1st digit

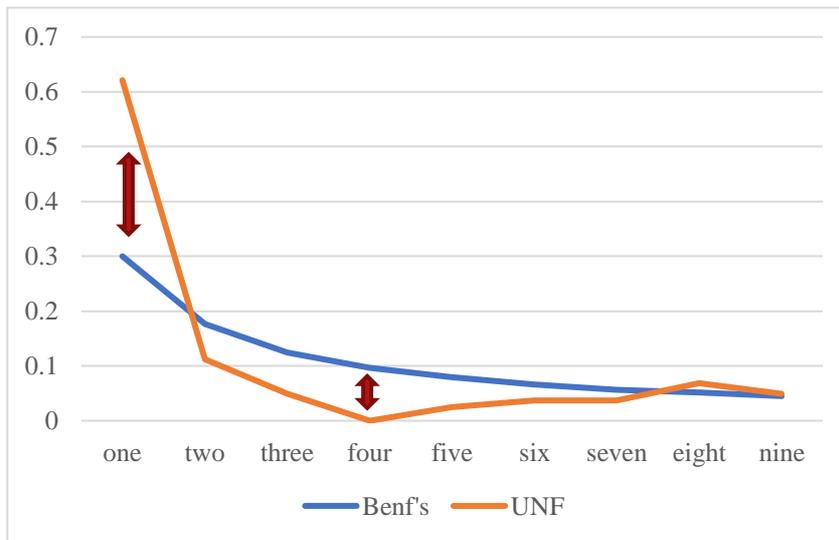Graph 08: DNA - distribution on the 1st digit

The actual distribution of the first digit shows no relationship for the first party, while we can see the distribution probability is closer to Benford's distribution in the third party.
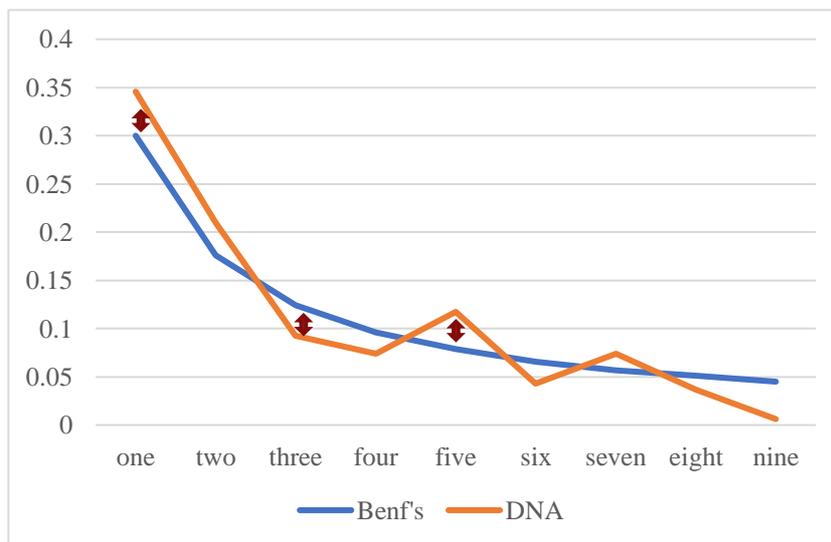
Graph 09: UPFA - distribution of the 1st digit with discrepancy

Graph 10: UNF - distribution of the 1st digit with discrepancy

Graph 11: DNA - distribution of the 1st digit with discrepancy

Notice the gaps between the Benford's distribution line and the actual distribution line in all three parties. The third party shows us a clear cut difference when compared to the first and second party. We have been successful in determining the distribution probability of the first digit of the election results in comparison to Benford's Law. Earlier, we also mentioned that, if the mean of a particular set of numbers is larger than the median and the skewness value is positive, the data set likely follows the Benford distribution (Durtchi, Hillison, & Pacini). Let's see if this agrees with our election data.

Table 4: Statistics for 2010 General Election Data

	UPFA	UNF	DNA
MEAN	27878	13,933	3,868
MEDIAN	27,396	13,553	2,527
ST. DEVIATION	13781	8121	4149
SKEWNESS	0.202101	1.612887	3.329065

Source: Authors' own calculations

(Table 04) shows us that the election results are most likely to follow the Benford's Law distribution. Even though the data in the table agrees to that rule, notice its extremely small margin to its justification point. For example the mean is bigger than the median for UPFA only by 482 points. This difference is hard to spot merely by only looking at the mean, median, and skewness of the data set, but also running digital analyses that will illustrate more precise information.

The relationship between the two distributions points out the fact that there has been some sort of manipulation with the election results. The smallest party (DNA) has the least power or influence in manipulating election results which is why we have been able to observe a close relationship between DNA's actual distribution and Benford's distribution.

4. Conclusions

The vote counts per voting area of the 2010 Presidential and General Election published on January 27th, 2010 and April 21st, 2010 respectively, by the Election Commission of Sri Lanka, show a highly significant excess of the first digit for candidate 01 and Candidate 02 in the Presidential Election, and UPFA, UNF and DNA in the General Election, compared to the expectations either from a uniform Benford's Law or from an empirically derived equivalent of Benford's Law. A possible clue for further investigation is that all the candidates' logarithmic vote count distributions are skewed, especially DNA's vote counts, which are positively skewed by about 3.32 standard errors, except for the UPFA, whose logarithmic vote counts are skewed (positively) by less than one standard error. Any demographic models of Sri Lankan voting patterns will need to either reproduce these statistical characteristics, or else make hypotheses regarding systematic differences in the data.

References

Abeywickrema, M. I. (2010, 01 03). News, Misuse of Public Property and Funds. (M. I. Abeywickrema, Ed.) Retrieved 03 21, 2017, from The Sunday Leader: http://www.thesundayleader.lk/2010/01/03/misuse-of-public-property-and-funds/

BBC. (2010, 01 12). South Asia, Sri Lanka pre-election attack leaves one person dead. Retrieved 03 15, 2017, from BBC News: http://news.bbc.co.uk/2/hi/8454285.stm

BBC. (2010, 01 18). South Asia, Two people killed in Sri Lanka pre-poll attacks. Retrieved 04 01, 2017, from BBC News: http://news.bbc.co.uk/2/hi/south_asia/8465202.stm

Central Bank of Sri Lanka. (2010). Statistics. Retrieved April 02, 2017, from Central Bank of Sri Lanka: http://www.cbsl.gov.lk/

CMEV. (n.d.). Final CMEV Reports. Retrieved 03 20, 2017, from Centre for Monitoring Election Violence: https://cmev.org/

Daily Mirror. (2010, 01 18). News. Retrieved 04 01, 2017, from Daily Mirror: http://www.dailymirror.lk/index.php/news/974-attack-on-minister-jayaratne-heraths-residence-.html

Department of Census and Statistics Sri Lanka. (2010). Population. Retrieved 2017, from Department of Census and Statistics Sri Lanka: http://www.statistics.gov.lk/

Durtchi, C., Hillison, W., & Pacini, C. (2004). The Effcrtive Use of Benford's Law to Assit in Detecting Fraud in Accounting Data. Journal of Forensic Accounting, V (2004), 17-34. Retrieved 02 03, 2017

PAFFREL. (2013). Presidential Election. Retrieved 04 01, 2017, from People's Action for Free and Fair Elections: http://www.paffrel.com

Palihawadena, N. (2010, 01 19). News. Retrieved 04 01, 2017, from The Island: http://www.island.lk/2010/01/19/news25.html

Presidential Elections. (2016). Retrieved 04 02, 2017, from Election Commission of Sri Lanka: http://www.slelections.gov.lk/

Saravanamuttu, P., Jayasekera, S., & Kalupathirana, U. (2010). Presidential Election – 2010, Media Communiqué No. 16. Center of Monitoring Election Violence. Colombo: Center of Monitoring Election Violence. Retrieved 04 02, 2017, from https://cmev.files.wordpress.com/2010/01/16th-media-communique_19_01_2010_english_final.pdf

Sunday Times. (2010, 01 17). News. Retrieved 04 02, 2017, from Sunday Times: http://www.sundaytimes.lk/100117/News/nws_04.html