Assessing Electoral Anomalies in South Korean Presidential Elections: A Benford's Law Approach

Donghyuk Kim[1], Doyoung Kim[2], Woohyuk Choi[3], Yewon Hwang[4]
[1]Department of Electrical Engineering, POSTECH, kimdonghyuk@postech.ac.kr
[2]Department of Mathematics, POSTECH, kdy0920@postech.ac.kr
[3]Department of Mathematics, POSTECH, asaf0121@postech.ac.kr
[4]Department of Electrical Engineering, POSTECH, yewon0510@postech.ac.kr

***Abstract*—This research investigates electoral integrity in South Korea's presidential elections from the 13th to the 20th term, using a comprehensive dataset of vote counts from each polling station and constituency. Employing Benford's Law as a primary analytical tool, the study aims to assess the presence and extent of electoral fraud. Two key areas of focus are explored: firstly, the evolution of democratic engagement among voters over successive elections and its correlation with Benford's Law conformity; and secondly, the prevalence of electoral anomalies in regions with strong political leanings. This study seeks to quantify the likelihood of electoral fraud and offers insights into the patterns of voting behavior and electoral integrity in South Korea, contributing to the discourse on the efficacy of Benford's Law in political science.**

I. Introduction

South Korea, as a democratic nation, conducts presidential elections that are a fundamental act of democracy executed by its citizens. The fairness of these elections serves as a key indicator of the country's democratic vitality.

However, the possibility of electoral manipulation during the electoral process is a significant issue that threatens the foundational mutual trust of democracy. In our study, we aim to explore the potential for electoral fraud by using statistical methods, analyzing election data from the 13th to the 20th presidential elections in South Korea.

The purpose of this research is to assess the integrity of the data from each presidential election using Benford's Law and chi-squared testing. Benford's Law is based on the theory that numbers in 'natural data sets' exhibit a specific pattern, enabling the detection of potential data manipulation or fraud. Specifically, in election data, deviations from Benford's Law in vote counts may indicate the likelihood of electoral fraud.

In this study, we applied Benford's Law to both national and regional data for each election. By analyzing the distribution of the 2nd and 3rd digits, we assessed how closely the data aligned with Benford's Law. This analysis sought to determine whether the vote counts in each election displayed natural distributions or signs of artificial manipulation.

The chi-squared test is a statistical method used to numerically evaluate the difference between observed data and expected distributions. We employed this technique to quantitatively analyze the discrepancies between the expected distributions according to Benford's Law and the actual distributions of the election data, thus identifying any anomalies.

The objective of this research is to demonstrate the effectiveness of statistical methods in evaluating the transparency and fairness of South Korea's presidential elections. Furthermore, this study aims to identify potential manipulations in the electoral process, contributing to the preservation of democratic integrity.

II. Related Works

Benford's Law, first noted by Simon Newcomb and later formalized by Frank Benford, predicts the frequency distribution of the first digits in many datasets [1]. This mathematical principle has found applications in diverse fields, including election forensics. Its premise rests on the observation that in many naturally occurring datasets, the first digit is more likely to be small. These first digit observations are used to detect numerous observations of nature using Benford's Law. [2][3] However, there is a study that shows that using the second digit is much better when applying Benford's Law to election data. In fact, these studies were conducted using election voting data in the United States. [4] Even if it was not the United States, Benford's Law was often used to raise election negativity in other countries.[5] "Despite its widespread use, the application of Benford's Law, particularly in elections, has been met with skepticism and challenges. Researchers have pointed out that while Benford's Law can be a useful tool for detecting anomalies, it is not a definitive indicator of fraud. The complexity of electoral data and the multifaceted nature of voting processes necessitate a cautious approach in interpreting results derived from Benford's analysis [6].

Furthermore, the efficacy of Benford's Law varies depending on the nature of the data and the context of the election. For example, in instances where data manipulation is subtle or sophisticated, Benford's Law may not effectively detect anomalies [7]. Additionally, the assumption that election data naturally conforms to Benford's distribution has been contested, with some studies suggesting that election data may not always follow this pattern due to various socio-political and demographic factors [6].

In the context of regional studies, the application of Benford's Law in politically biased areas presents unique challenges. Political dynamics and voter behavior patterns in these regions may significantly deviate from the expected

distribution, leading to false positives or negatives in fraud detection [8]. This highlights the need for a nuanced approach in applying Benford's Law to election data, considering the specific political and social context of each region.

III. Theoretical Background

A. Historical Background and Discovery of Benford's Law

Benford's Law, alternatively referred to as the First-Digit Law, represents an intriguing statistical principle concerning the frequency distribution of leading digits within a multitude of real-world numerical datasets. It was first noted by Simon Newcomb, an astronomer and mathematician, in 1881. Newcomb observed that the early pages of logarithm tables, which contained lower digits, were more worn than those with higher digits, implying a natural bias towards lower leading digits in scientific calculations.

Frank Benford, a physicist, later rediscovered and formally tested this observation in the 1930s across various datasets. He found a consistent pattern where the digit '1' appeared as the leading digit about 30% of the time, significantly more than expected if digits were uniformly distributed. To illustrate, the digit '1' emerges as the initial digit approximately 30% of the time, whereas larger digits such as '9' appear as the leading digit with lesser frequency, approximately 5% of the time.

B. Mathematical Formulation

Benford's Law states that in many naturally occurring collections of numbers, the leading digit is likely to be small. For example, the number '1' appears as the first digit about 30% of the time, rather than 11.11% as one might expect if all digits were equally likely.

Mathematically, the probability of a digit d (where d is a digit from 1 to 9) being the leading digit is given by:

$$P(d) = \log_{10}\left(1 + \frac{1}{d}\right)$$

This logarithmic distribution skews heavily towards lower digits.

C. Appoaches to prove Benford's Law

Initially, explanations for Benford's Law ranged from the psychological explanations, such as how people perceive and record numbers to the physical explanations, such as inherent properties of certain dataset. These theories, however, lacked a rigorous mathematical basis. Over the years, empirical studies in various fields, from finance to demographics, consistently supported Benford's Law, showing that certain datasets naturally followed this logarithmic distribution of leading digits.

The most significant advancement in understanding Benford's Law came from mathematician Ted Hill in 1995. His work provided a formal and rigorous mathematical proof for the law. Key insights from Hill's proof are the followings. It revealed that numbers in sets conforming to Benford's distribution are typically derived from various other distributions. This was a crucial insight, as it moved the law from an empirical observation to a statistically predictable outcome. Also, Hill demonstrated that when numbers from different distributions (each with its own unique properties) are combined, the resulting dataset tends to follow Benford's Law. This finding highlighted the importance of the mixture of distributions in generating a Benford-like pattern. A critical aspect of Hill's proof was the idea of 'true randomness'. He posited that when numbers from different sources are combined, it results in a distribution of distributions—a phenomenon that reflects a universal, natural randomness.

This comprehensive proof opened the door for the application of Benford's Law in numerous fields, most notably in fraud detection and forensic accounting, where it serves as a tool for identifying anomalies in financial data.

D. Continuing Research and Developments

Subsequent research has extended the application of Benford's Law to second and further digits, though with less pronounced effects compared to the first digit. Despite Hill's proof, there remains an active academic discussion around the boundaries of Benford's Law's applicability, particularly in datasets where its assumptions may not hold.

E. Application in Detecting Fraud and Anomalies

In Accounting and Auditing: Benford's Law has been effectively used in accounting and auditing as a tool for fraud detection. It identifies inconsistencies in financial data where the distribution of leading digits deviates from Benford's expected distribution. Notably, researchers like Nigrini have applied it extensively to accounting numbers, revealing its potential in detecting tax evasion and earnings manipulation.

In Election Forensics: The application of Benford's Law extends to election forensics, particularly useful in analyzing vote counts. This method is based on the distribution of digits in vote counts, referred to as the second-digit Benford's Law (2BL), and helps in detecting fraud or irregularities in election data.

Studies have shown that Benford's Law can be applied to the first and second significant digits of precinct-level vote counts, with tests revealing how closely these distributions match Benford's expectations.

Limitations and Considerations: Benford's Law is not universally applicable and its effectiveness varies depending on the dataset. For instance, certain types of fraud, such as off-the-books transactions, may not be detectable using this method.

It's important to understand the nature of the data and the types of irregularities that Benford's Law can identify. In electoral data, it has been shown that vote counts at the precinct level conform to Benford's Law under specific conditions, but not necessarily at the level of individual voting machines.

IV. Problem and Implementation

A. Data Preprocessing

These data, reflecting the outcomes at each polling station in the 13rd to 20th presidential elections post-1987 democratization in South Korea, were meticulously evaluated to ensure comprehensive analysis. In this process, a crucial step involved the exclusion of candidates who did not meet a predetermined threshold for the number of votes, which was necessary to maintain statistical validity and relevance in the analysis. This threshold was determined based on a combination of statistical norms and practical considerations, such as ensuring a sufficient sample size for each candidate to apply Benford's Law effectively.

Criteria for Exclusion: The exclusion criteria for candidates were rigorously defined to maintain the integrity of the analysis. Candidates with vote counts below a certain threshold were excluded to ensure that the data analyzed was representative of significant electoral trends and not skewed by outliers or minimal data points. This threshold was set by considering the overall distribution of votes and the need for a robust data set that could yield statistically significant results. In fact, since we conducted the 1BL, 2BL, and 3BL tests, we excluded the candidate from the data if more than 50% of the corresponding vote-gathering data did not exceed three digits. Here's an example deciding elimination of third-place candidate with our criteria.

	Ratio of Less-than-3-Digits (%)
Kim Dae-Jung in 13rd	93.9
Jung Ju-Young in 14th	63.4
Lee Han-Dong in 15th	0.1
Kang Ji-Won in 18th	0.0
Shim Sang-Jung in 20th	31.1
Critical	50.0

TABLE I
THE RATIO OF NUMBERS WITH LESS THAN THREE DIGITS FOR 3RD PLACE CANDIDATES

Preparation for Benford's Law Analysis: Before applying Benford's Law, it was necessary to prepare the data in a way that would make the analysis meaningful. This preparation involved aggregating vote counts at different levels – including constituency and national levels – to observe patterns at various scales. The first-digit , second-digit , third-digit distribution of these aggregated vote counts was then analyzed to see if they conformed to the expected distribution as per Benford's Law.

Statistical Techniques for Deeper Analysis: In addition to Benford's Law, a range of statistical techniques were employed to analyze the data thoroughly. These techniques included regression analysis to identify trends over time, chi-square tests to assess the goodness of fit for Benford's Law, and spatial analysis to understand regional voting patterns. The combination of these methods provided a multi-faceted view of the data, allowing for a more nuanced understanding of electoral dynamics in South Korea.

Adjustment to the case of Korea: In case of Korea, there are traditionally clarified dichotomous political ideologies for each presidential-elections [9]. So to adjust our research and our methodology especially in case of South Korea, we labelled each districts as a conservative-camp and a liberal-camp for a particular area that has traditionally been such a ideological-propensity. In addition, we also labelled same ideologies for each candidates. This labelling-process greatly helped our research to make it easier to see the characteristics of each camp for each elections.

B. Methodology

Overview: This study adopts a mixed-methods approach, combining quantitative analysis with statistical inference to investigate the integrity of South Korean presidential elections from the 13th to the 20th term. The methodology is structured to explore two primary objectives: firstly, the application of Benford's Law to detect anomalies in electoral data, and secondly, the examination of voter behavior and regional political biases over time.

Data Collection and Processing: As previously detailed, the data consists of vote counts from each polling station across all constituencies for the 13th to 20th presidential elections in South Korea. The data was subjected to rigorous cleaning and eliminating procedures to ensure consistency and reliability for analysis.

Application of Benford's Law: Benford's Law, a significant tool in forensic accounting and fraud detection, is firstly utilized to analyze the first-digit distribution of the vote counts. The expected distribution of first digits according to Benford's Law is compared with the actual distribution in the dataset. Actually, we got too-high value of chi-square value when we process with First-digit distribution as we can see in the table below.

	1BL Chi-Square Value
Park Geun-Hye	7441.7
Moon Jae-In	3179.5
Critical	15.5

TABLE II
1BL CHI-SQUARE VALUES FOR CANDIDATES IN 18TH ELECTION

So we decided to process our methodology with 2BL, 3BL case Chi-square test, this is supported by the research by Mebane [4].

Statistical Analysis: To supplement the findings from Benford's Law, various statistical tests and methods are employed. This includes chi-square tests for goodness-of-fit to Benford's distribution, regression analysis to observe trends over time, and spatial analysis to examine regional voting patterns. We set significance-level of our test as 0.05, which leads to 95% confidence interval. And degree of freedom(ν) is 9 since ten digit (0 - 9) exists, then we can get $\nu = 10 - 1 = 9$. These methods provide a comprehensive understanding of the data and help validate the findings from Benford's Law analysis.

Comparative Analysis: A comparative approach is taken to analyze the data across different elections. This involves comparing the conformity to Benford's Law over time and across regions with varying political leanings. The aim is to identify any trends or patterns in the data that could signify changes in electoral integrity and democratic engagement in South Korea.

Limitations and Considerations: The methodology acknowledges the limitations inherent in the data and analysis techniques. While Benford's Law is a powerful tool for detecting anomalies, it is not conclusive proof of fraud. The study therefore carefully interprets findings within the broader political and historical context of South Korea.

C. Results

We applied Benford's Law to the presidential elections from the 13th to the 20th, based on the total number of votes and regional vote counts. This involved applying Benford's Law to the second and third digits, with the vote counts at each polling station used as the basic unit. Additionally, Benford's Law is effective only when the number of votes exceeds a certain value, so it was applied to the two or three candidates who received the most votes in each election.

As described in the previous research method, the degrees of freedom in the Chi-squared test is 9, and the confidence interval is 95%, i.e., $\alpha = 0.05$. Therefore, the critical value derived from the Chi-squared distribution table was 16.9.

D. 13th Election

1) National Data of the 13th Election: The results of applying Benford's Law to the second and third digits using vote counts from all polling stations in the 13th national election data are as follows:

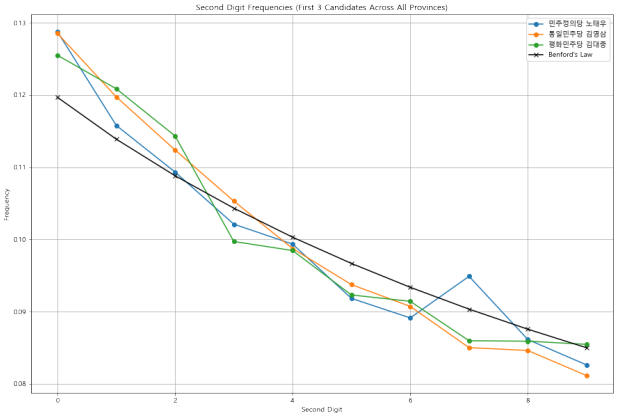

Fig. 1. 13th Election Second Digit Frequencies

The Chi-squared values calculated for candidates Roh Tae-woo, Kim Young-sam, and Kim Dae-jung were:

All values are displayed up to two decimal places. Based on these results, we can make the following interpretations:

1) In the 2nd digit Benford, the Chi-squared values of all three candidates exceeded the critical value, but not significantly, and the values for all three candidates were similar, making it difficult to suspect election manipulation.
2) In the 3rd digit Benford, Roh Tae-woo's Chi-squared value was 18.58, slightly exceeding the critical value of 16.9. Kim Young-sam's value did not exceed the critical value, while Kim Dae-jung's value was 37.55, significantly higher than the other candidates. This shows that Kim Dae-jung's value is relatively abnormal compared to the other candidates, suggesting it as an indicator for possible election manipulation.

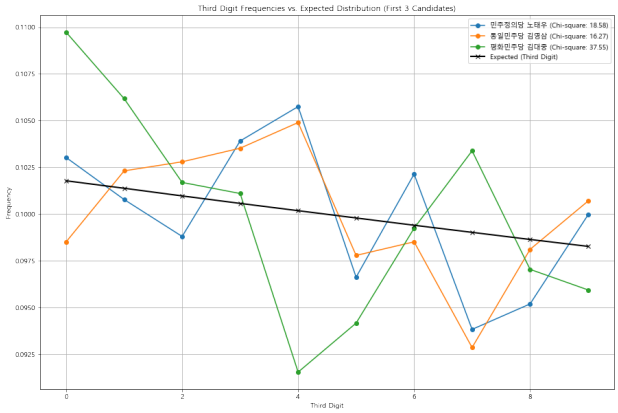

Fig. 2. 13th Election Third Digit Frequencies

Candidate	2nd Digit Benford	3rd Digit Benford
Roh Tae-woo	26.77	18.58
Kim Young-sam	32.48	16.27
Kim Dae-jung	29.89	37.55

TABLE III
13TH ELECTION CHI-SQUARED VALUES FOR CANDIDATES

2) Regional Data of the 13th Election: The following are the results of applying Benford's Law to each region (city/province). Below are the Chi-squared values.

Province	Roh Tae-woo	Kim Young-sam	Kim Dae-jung
Seoul Metropolitan Government	56.60	49.38	121.86
Busan Metropolitan City	11.56	55.07	10.94
Daegu Metropolitan City	9.20	4.39	6.27
Incheon Metropolitan City	22.66	9.80	3.69
Gwangju Metropolitan City	17.46	22.47	14.22
Gyeonggi Province	19.24	12.07	19.02
Gangwon-do	5.52	4.27	7.54
Chungcheongbuk-do	10.51	6.95	8.17
Daejeon Metropolitan City	5.47	7.44	7.02
Chungcheongnam-do	9.30	5.32	11.55
Jeollabuk-do	8.94	36.17	8.97
Jeollanam-do	7.76	52.52	5.64
Gyeongsangbuk-do Province	13.87	10.32	11.26
Gyeongsangnam-do Province	7.48	7.44	10.77
Ulsan Metropolitan City	14.25	6.62	14.32
Jeju Special Self-Governing Province	8.74	9.45	9.10

TABLE IV
CHI-SQUARE VALUE OF 2ND DIGIT BENFORD BY REGION FOR THE 13TH ELECTION

Firstly, in the 2nd digit Benford, regions of note are Seoul, Busan, and Jeollanam-do. Seoul's data showed Chi-squared values of 56.60, 49.38, and 121.86 respectively. Although

Province	Roh Tae-woo	Kim Young-sam	Kim Dae-jung
Seoul Metropolitan Government	18.93	8.94	6.71
Busan Metropolitan City	11.30	4.58	3.62
Daegu Metropolitan City	11.70	4.62	47.50
Incheon Metropolitan City	3.56	4.16	14.80
Gwangju Metropolitan City	33.65	578.67	15.93
Gyeonggi Province	7.32	12.83	33.81
Gangwon Province	11.93	10.59	22.87
North Chungcheong Province	17.46	11.10	3.77
Daejeon Metropolitan City	10.80	11.06	14.47
South Chungcheong Province	6.76	8.12	6.63
North Jeolla Province	34.97	164.95	4.53
South Jeolla Province	42.33	494.04	11.74
North Gyeongsang Province	7.76	14.15	78.81
South Gyeongsang Province	6.88	8.90	50.53
Ulsan Metropolitan City	10.76	14.83	25.94
Jeju Special Self-Governing Province	10.72	8.92	10.75

TABLE V
CHI-SQUARE VALUE OF 3RD DIGIT BENFORD BY REGION FOR THE 13TH ELECTION

all values exceed the critical value, Kim Dae-jung's value is notably higher. This can be observed as an influence from the data of Seoul in the 13th national election, where Kim Dae-jung's Chi-squared value was also abnormal.

In Busan and Jeollanam-do, Chi-squared values for candidates Roh Tae-woo and Kim Dae-jung were within the critical value (Busan Roh Tae-woo: 11.56, Busan Kim Dae-jung: 10.94, Jeollanam-do Roh Tae-woo: 7.76, Jeollanam-do Kim Dae-jung: 5.64). In contrast, Kim Young-sam's values in Busan were 55.07 and in Jeollanam-do were 52.52, significantly exceeding the critical value and showing a substantial difference from the other candidates. This could be interpreted as a result of specific political leanings in certain regions, particularly suggesting potential election manipulation for Kim Young-sam in Busan and Jeollanam-do in the 13th election.

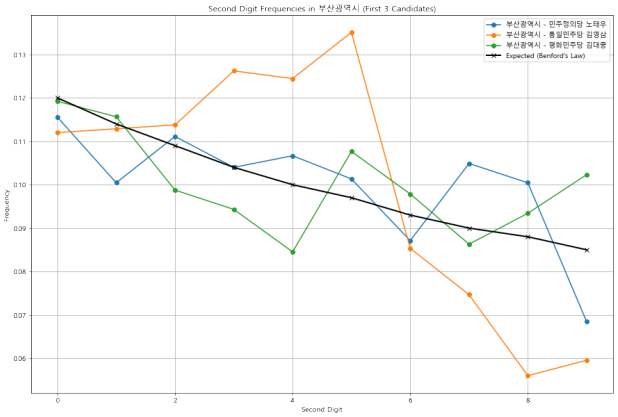

Fig. 3. 13th Election Busan Second Digit Frequencies

Next, in the 3rd digit Benford, notable regions include Daegu, Gwangju, Jeollanam-do, and Gyeongsangbuk-do. In Gwangju and Jeollanam-do, Kim Young-sam's Chi-squared values were significantly high compared to other candidates, with values of 578.67 in Gwangju and 494.04 in Jeollanam-do. Given that the values for other candidates in these regions ranged between 10 and 50, this is clearly abnormal.

In Daegu and Gyeongsangbuk-do, Kim Dae-jung's Chi-squared values were notably high. While the other two candidates' values ranged between 4 and 15, Kim Dae-jung's values were 47.50 in Daegu and 78.81 in Gyeongsangbuk-do, indicating relatively high values.

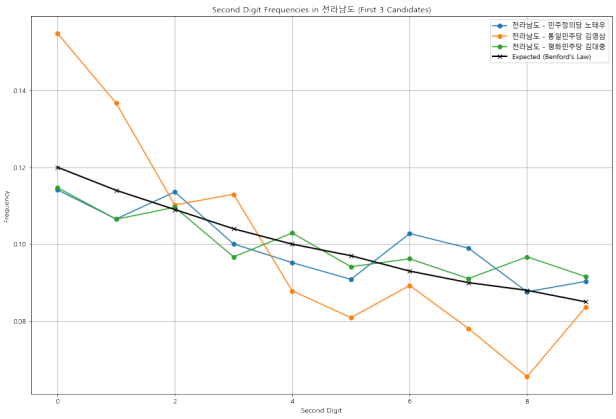

Fig. 4. 13th Election Jeollanam-do Second Digit Frequencies

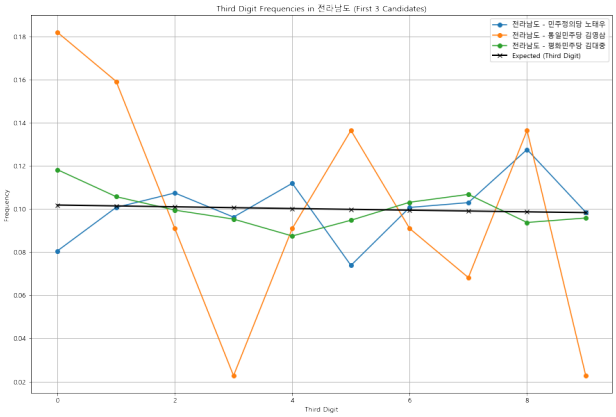

Fig. 5. 13th Election Jeollanam-do Third Digit Frequencies

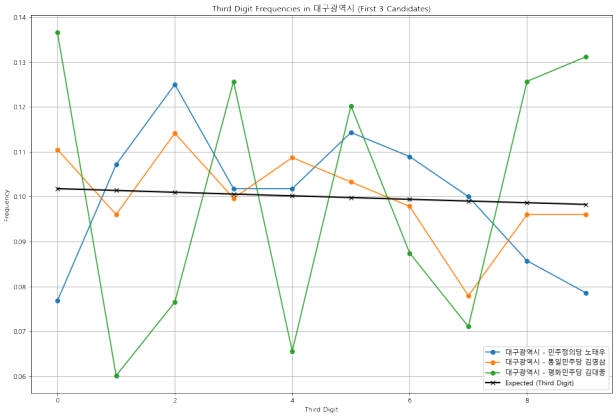

Fig. 6. 13th Election Daegu Third Digit Frequencies

In summary, we observed a slightly abnormal Chi-squared value for Kim Dae-jung in the national data's 3rd digit Benford. In specific regions, the 2nd digit Benford in Busan and Jeollanam-do for Kim Young-sam, and the 3rd digit Benford

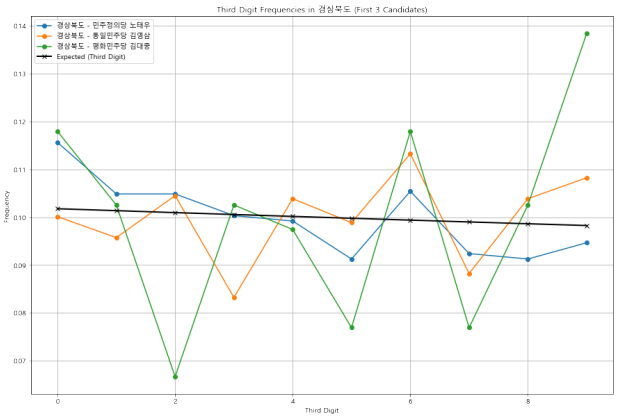

Fig. 7. 13th Election Gyeongsangbuk-do Third Digit Frequencies

in Gwangju and Jeollanam-do for Kim Young-sam, as well as in Daegu and Gyeongsangbuk-do for Kim Dae-jung showed abnormal Chi-squared values. While Roh Tae-woo's results did not suggest election manipulation, the results for the other two candidates did, with stronger suspicion for Kim Young-sam in specific regions and for Kim Dae-jung nationally.

E. 14th Election

1) National Data of the 14th Election: The results of applying Benford's Law to the second and third digits using vote counts from all polling stations in the 14th national election data are as follows:

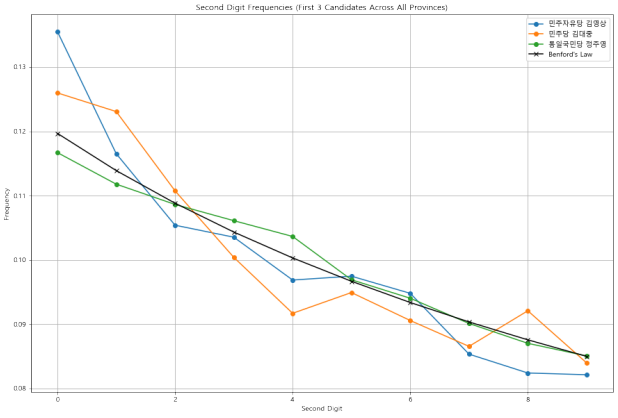

Fig. 8. 14th Election Second Digit Frequencies

The Chi-squared values calculated for candidates Kim Young-sam, Kim Dae-jung, and Chung Ju-young were:

Candidate	2nd Digit Benford	3rd Digit Benford
Kim Young-sam	53.97	11.12
Kim Dae-jung	43.88	5.11
Chung Ju-young	4.62	39.23

TABLE VI
14TH ELECTION CHI-SQUARED VALUES FOR CANDIDATES

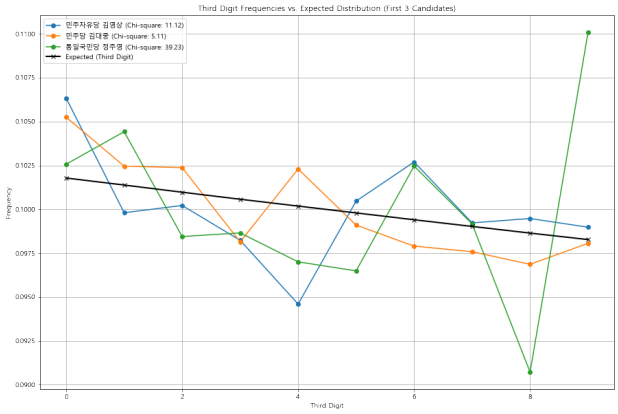

Fig. 9. 14th Election Third Digit Frequencies

First, in the 2nd digit Benford, both Kim Young-sam and Kim Dae-jung's Chi-squared values exceeded the critical value. Second, in the 3rd digit Benford, Chung Ju-young's Chi-squared value exceeded the critical value. Overall, there were no candidates significantly deviating from the critical value, making it difficult to suspect election manipulation.

2) Regional Data of the 14th Election: The following are the results of applying Benford's Law regionally.

Province	Kim Young-sam	Kim Dae-jung	Chung Ju-yung
Seoul Metropolitan Government	213.41	222.94	6.59
Busan Metropolitan City	43.04	35.72	24.52
Daegu Metropolitan City	34.34	5.42	14.62
Incheon Metropolitan City	35.51	19.39	16.08
Gwangju Metropolitan City	4.17	22.88	20.75
Daejeon Metropolitan City	13.28	8.33	21.64
Gyeonggi Province	40.07	28.83	9.29
Gangwon Province	8.19	7.40	6.77
North Chungcheong Province	10.94	3.78	4.52
South Chungcheong Province	20.34	17.22	7.49
North Jeolla Province	13.72	18.76	9.73
South Jeolla Province	21.99	5.82	43.91
North Gyeongsang Province	6.39	11.35	2.76
South Gyeongsang Province	11.49	6.70	7.20
Jeju Special Self-Governing Province	9.42	8.22	6.66

TABLE VII
CHI-SQUARE VALUE OF 2ND DIGIT BENFORD BY REGION FOR THE 14TH ELECTION

Province	Kim Young-sam	Kim Dae-jung	Chung Ju-yung
Seoul Metropolitan Government	14.20	15.38	13.32
Busan Metropolitan City	8.85	9.46	6.45
Daegu Metropolitan City	10.07	4.30	13.48
Incheon Metropolitan City	5.65	4.10	2.63
Gwangju Metropolitan City	131.84	4.59	-
Daejeon Metropolitan City	10.16	2.77	3.03
Gyeonggi Province	7.29	4.55	7.24
Gangwon Province	8.39	10.67	13.66
North Chungcheong Province	10.54	7.55	16.62
South Chungcheong Province	8.85	8.58	4.46
North Jeolla Province	121.90	6.14	583.45
South Jeolla Province	240.25	6.77	886.32
North Gyeongsang Province	11.65	21.14	19.20
South Gyeongsang Province	12.36	24.10	60.32
Jeju Special Self-Governing Province	6.10	16.06	17.93

TABLE VIII
CHI-SQUARE VALUE OF 3RD DIGIT BENFORD BY REGION FOR THE 14TH ELECTION

Firstly, in the 2nd digit Benford, Seoul is a region of note. Seoul's data showed Chi-squared values of 213.41, 222.94, and

6.59 respectively. Both Kim Young-sam and Kim Dae-jung's Chi-squared values significantly exceeded the critical value in the national data of the 14th election, suggesting the influence of Seoul's data and indicating these values are abnormal.

Next, in the 3rd digit Benford, notable regions are Gwangju and Jeollanam-do. In both areas, Kim Young-sam's Chi-squared values were notably high, being 131.84 in Gwangju and 240.25 in Jeollanam-do. Additionally, Chung Ju-young's Chi-squared value in Jeollanam-do was abnormally high at 886.32. Interestingly, similar results were obtained in the previous 13th election, where Kim Young-sam's Chi-squared values in Gwangju and Jeollanam-do were abnormal.

In summary, for the 14th election, slight abnormalities were observed in Kim Young-sam and Kim Dae-jung's Chi-squared values in the 2nd digit Benford. In the 3rd digit Benford, Kim Young-sam's values in Gwangju and Jeollanam-do were highly abnormal, suggesting potential election manipulation in these regions.

F. 15th Election

1) National Data of the 15th Election: The results of applying Benford's Law to the second and third digits using vote counts from all polling stations in the 15th national election data are as follows:

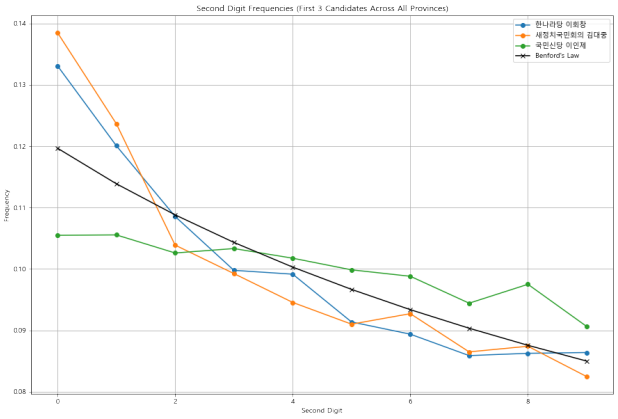

Fig. 10. 15th Election Second Digit Frequencies

The Chi-squared values calculated for candidates Lee Hoi-chang, Kim Dae-jung, and Lee In-je were:

Candidate	2nd Digit Benford	3rd Digit Benford
Lee Hoi-chang	48.51	4.70
Kim Dae-jung	90.37	16.28
Lee In-je	83.51	10.02

TABLE IX
15TH ELECTION CHI-SQUARED VALUES FOR CANDIDATES

First, in the 2nd digit Benford, all three candidates' Chi-squared values exceeded the critical value, with Lee Hoi-chang's value being relatively lower compared to the other two. Second, in the 3rd digit Benford, all candidates showed Chi-squared values below the critical value.

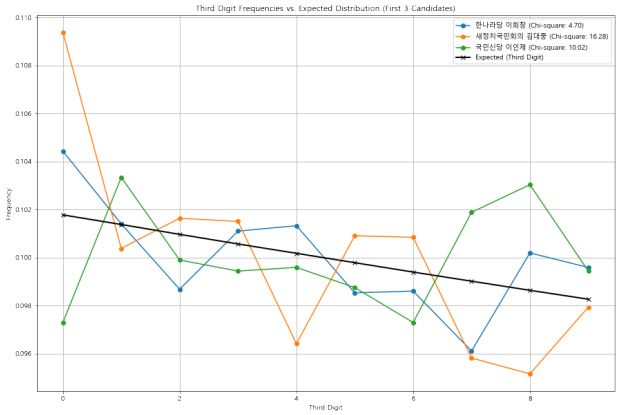

Fig. 11. 15th Election Third Digit Frequencies

2) Regional Data of the 15th Election: The following are the results of applying Benford's Law regionally.

Province	Lee Hoi-chang	Kim Dae-jung	Lee In-je
Seoul Metropolitan Government	102.56	473.81	49.39
Busan Metropolitan City	94.24	15.14	21.42
Daegu Metropolitan City	36.25	15.35	7.91
Incheon Metropolitan City	17.70	8.41	25.02
Gwangju Metropolitan City	24.72	23.51	49.35
Daejeon Metropolitan City	4.93	14.53	12.19
Ulsan Metropolitan City	14.23	8.08	11.25
Gyeonggi Province	10.21	38.31	24.34
Gangwon Province	15.52	9.66	7.67
North Chungcheong Province	5.21	12.80	13.73
South Chungcheong Province	14.04	13.51	4.78
North Jeolla Province	5.22	7.48	13.18
South Jeolla Province	27.12	9.64	60.09
North Gyeongsang Province	17.50	5.56	7.44
South Gyeongsang Province	7.26	12.08	20.07
Jeju Special Self-Governing Province	7.54	16.90	8.21

TABLE X
CHI-SQUARE VALUE OF 2ND DIGIT BENFORD BY REGION FOR THE 15TH ELECTION

Province	Lee Hoi-chang	Kim Dae-jung	Lee In-je
Seoul Metropolitan Government	5.86	10.07	8.04
Busan Metropolitan City	5.87	7.05	12.29
Daegu Metropolitan City	4.53	3.02	11.86
Incheon Metropolitan City	14.08	7.04	7.66
Gwangju Metropolitan City	1127.98	12.63	-
Daejeon Metropolitan City	16.68	12.05	3.99
Ulsan Metropolitan City	11.32	12.87	5.58
Gyeonggi Province	11.07	16.36	7.17
Gangwon Province	6.44	11.94	8.66
North Chungcheong Province	7.59	7.77	11.62
South Chungcheong Province	13.67	8.25	12.67
North Jeolla Province	53.04	9.91	604.33
South Jeolla Province	133.40	6.07	544.76
North Gyeongsang Province	9.41	14.44	8.80
South Gyeongsang Province	2.83	22.64	12.21
Jeju Special Self-Governing Province	16.42	14.98	13.67

TABLE XI
CHI-SQUARE VALUE OF 3RD DIGIT BENFORD BY REGION FOR THE 15TH ELECTION

Firstly, in the 2nd digit Benford, regions of note are Seoul, Busan, Gwangju, and Jeollanam-do. Seoul's data showed Chi-squared values of 102.56, 473.81, and 49.39 respectively. All values exceeded the critical value, especially Kim Dae-jung's, suggesting potential election manipulation. In Busan, Lee In-

je's Chi-squared value exceeded the critical value but was relatively smaller compared to Lee Hoi-chang's value of 94.24. Considering the significant abnormality in Lee Hoi-chang's value in Busan, election manipulation could be suspected.

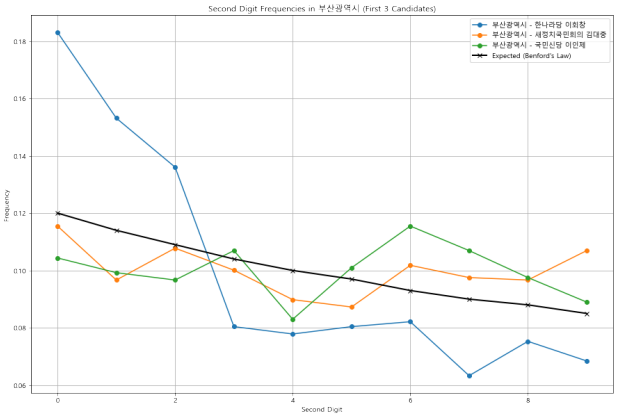

Fig. 12. 15th Election Busan Second Digit Frequencies

In Gwangju and Jeollanam-do, Lee In-je's Chi-squared values were notably high at 49.35 and 60.08 respectively, indicating possible election manipulation in these regions.

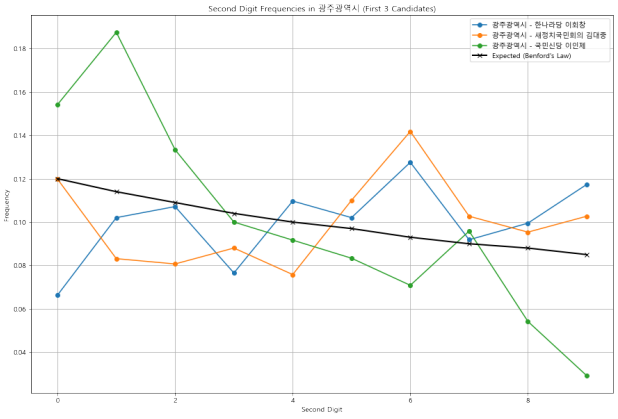

Fig. 13. 15th Election Gwangju Second Digit Frequencies

In the 3rd digit Benford, Gwangju and Jeollanam-do are notable regions. Lee Hoi-chang's Chi-squared values were abnormally high at 35.54 and 125.33 respectively, similar to the pattern observed in the previous 15th election.

Thus, for the 15th election, slight abnormalities were observed nationally in Kim Dae-jung and Lee In-je's Chi-squared values, with regional anomalies suggesting possible election manipulation.

G. 16th Election

1) National Data of the 16th Election: The results of applying Benford's Law to the second and third digits using vote counts from all polling stations in the 16th national election data are as follows:

The Chi-squared values calculated for candidates Lee Hoi-chang, and Roh Moo-hyun were:

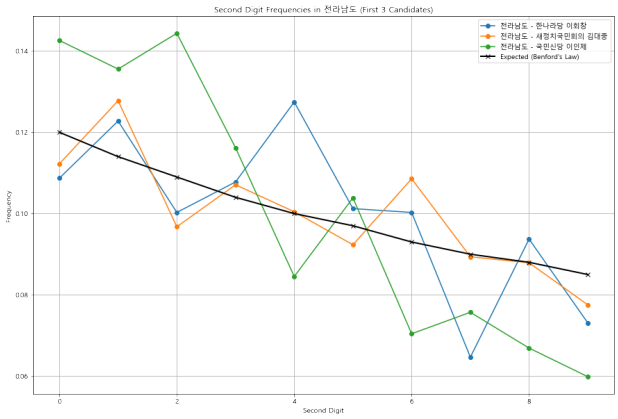

Fig. 14. 15th Election Jeollanam-do Second Digit Frequencies

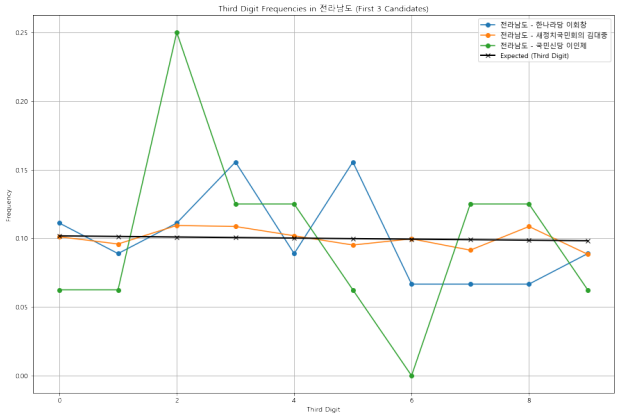

Fig. 15. 15th Election Jeollanam-do Third Digit Frequencies

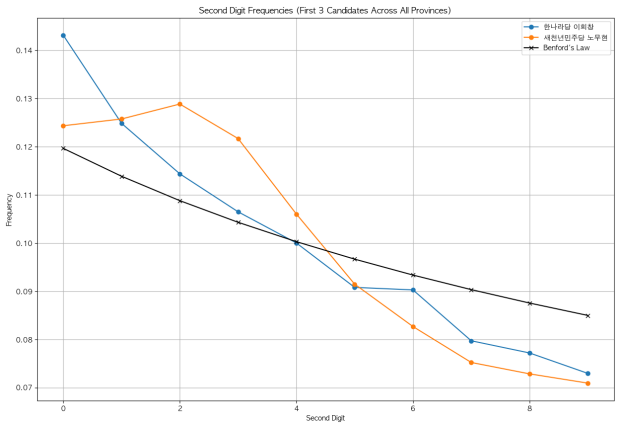

Fig. 16. 16th Election Second Digit Frequencies

Candidate	2nd Digit Benford	3rd Digit Benford
Lee Hoi-chang	185.73	8.27
Roh Moo-hyun	299.70	4.98

TABLE XII
16TH ELECTION CHI-SQUARED VALUES FOR CANDIDATES

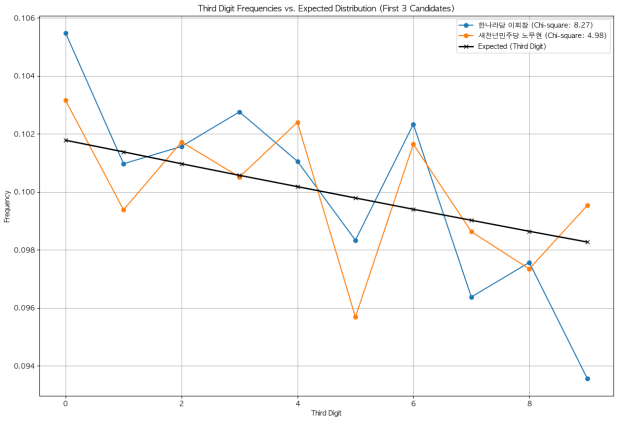

Fig. 17. 16th Election Third Digit Frequencies

First, in the 2nd digit Benford, both candidates' Chi-squared values exceeded the critical value, with Roh Moo-hyun's being higher. Second, in the 3rd digit Benford, both candidates showed Chi-squared values below the critical value.

2) Regional Data of the 16th Election: The following are the results of applying Benford's Law regionally.

Province	Lee Hoi-chang	Roh Moo-hyun
Seoul Metropolitan Government	237.73	571.65
Busan Metropolitan City	55.40	24.26
Daegu Metropolitan City	36.90	16.99
Incheon Metropolitan City	57.67	125.00
Gwangju Metropolitan City	7.24	11.50
Daejeon Metropolitan City	8.45	48.05
Ulsan Metropolitan City	13.21	4.90
Gyeonggi Province	56.29	130.34
Gangwon Province	9.12	13.19
North Chungcheong Province	5.46	25.74
South Chungcheong Province	10.43	9.36
North Jeolla Province	17.73	4.76
South Jeolla Province	8.03	10.61
North Gyeongsang Province	60.31	12.36
South Gyeongsang Province	14.08	13.04
Jeju Special Self-Governing Province	4.32	8.10

TABLE XIII
CHI-SQUARE VALUE OF 2ND DIGIT BENFORD BY REGION FOR THE 16TH ELECTION

Firstly, in the 2nd digit Benford, Chi-squared values varied across regions. As the images indicate, in Seoul, both candidates had high Chi-squared values, with Roh Moo-hyun's being higher. Additionally, in Incheon, Daejeon, and Gyeonggi-do, Roh Moo-hyun's Chi-squared values exceeded the critical value and appeared abnormal. For Lee Hoi-chang, abnormal Chi-squared values were observed in Busan and Gyeongsangbuk-do, exceeding the critical value as well. In the 3rd digit Benford, Gwangju and Jeollanam-do are regions of note. Lee Hoi-chang's Chi-squared values were abnormally high at 35.54 and 125.33, echoing a similar trend from the 15th election.

In conclusion, for the 16th election, both candidates showed high Chi-squared values in specific regions, not sufficient to suspect widespread election manipulation. However, the

Province	Lee Hoi-chang	Roh Moo-hyun
Seoul Metropolitan Government	9.05	9.53
Busan Metropolitan City	1.48	7.69
Daegu Metropolitan City	2.97	11.52
Incheon Metropolitan City	9.59	6.94
Gwangju Metropolitan City	35.54	2.68
Daejeon Metropolitan City	7.14	15.18
Ulsan Metropolitan City	11.82	7.87
Gyeonggi Province	7.57	7.27
Gangwon Province	17.59	9.92
North Chungcheong Province	5.41	19.07
South Chungcheong Province	8.31	4.86
North Jeolla Province	35.59	16.71
South Jeolla Province	125.33	6.37
North Gyeongsang Province	4.25	7.73
South Gyeongsang Province	8.94	4.77
Jeju Special Self-Governing Province	13.10	10.83

TABLE XIV
CHI-SQUARE VALUE OF 3RD DIGIT BENFORD BY REGION FOR THE 16TH ELECTION

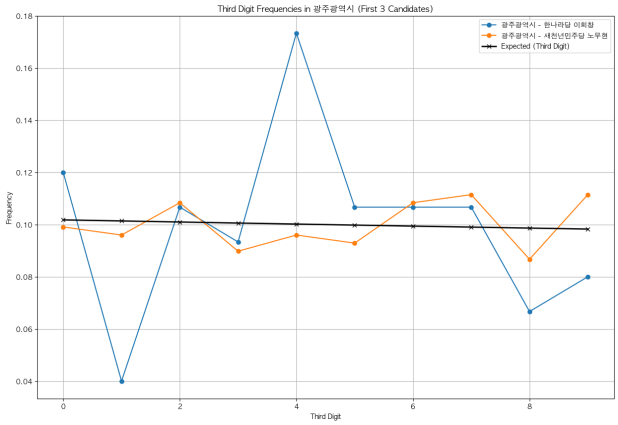

Fig. 18. 16th Election Gwangju Third Digit Frequencies

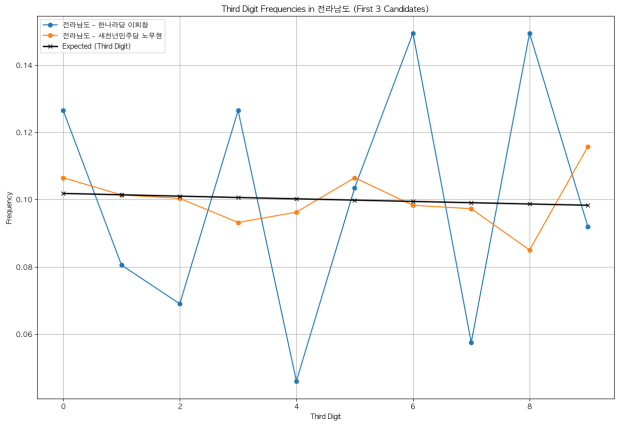

Fig. 19. 16th Election Jeollanam-do Third Digit Frequencies

analysis of the 3rd digit Benford showed interesting results with high Chi-squared values for Lee Hoi-chang in Gwangju and Jeollanam-do, following the trend from the 15th election.

H. 17th Election

1) National Data of the 17th Election: The results of applying Benford's Law to the second and third digits using vote counts from all polling stations in the 17th national election data are as follows:

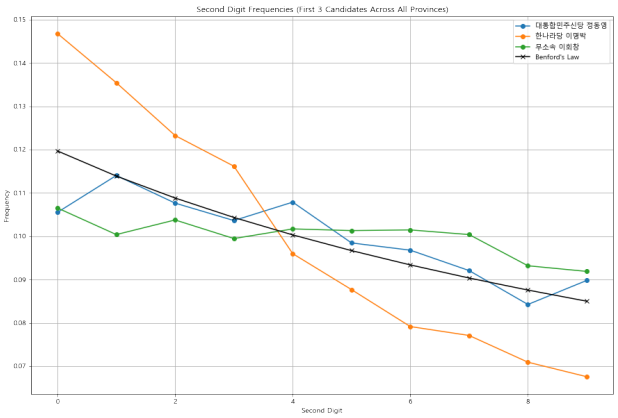

Fig. 20. 17th Election Second Digit Frequencies

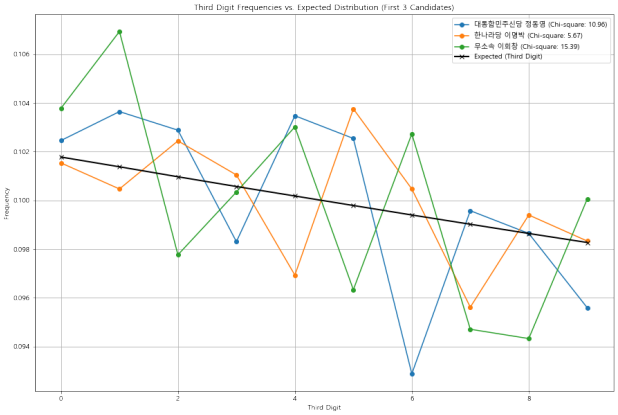

Fig. 21. 17th Election Third Digit Frequencies

The Chi-squared values calculated for candidates Chung Dong-young, Lee Myung-bak, and Lee Hoi-chang were:

Candidate	2nd Digit Benford	3rd Digit Benford
Chung Dong-young	49.74	10.96
Lee Myung-bak	442.52	5.67
Lee Hoi-chang	113.15	15.39

TABLE XV
17TH ELECTION CHI-SQUARED VALUES FOR CANDIDATES

First, in the 2nd digit Benford, all three candidates' Chi-squared values exceeded the critical value, with Lee Myung-bak's being the highest and most abnormal. Second, in the 3rd digit Benford, all candidates showed Chi-squared values below the critical value.

2) Regional Data of the 17th Election: The following are the results of applying Benford's Law regionally.

Province	Chung Dong-young	Lee Myung-bak	Lee Hoi-chang
Seoul Metropolitan Government	26.59	245.09	47.45
Busan Metropolitan City	8.17	62.46	13.63
Daegu Metropolitan City	47.06	21.44	15.66
Incheon Metropolitan City	7.82	65.62	20.65
Gwangju Metropolitan City	9.76	7.95	12.24
Daejeon Metropolitan City	10.16	13.28	16.98
Ulsan Metropolitan City	11.62	22.60	16.24
Gyeonggi Province	35.72	91.96	14.02
Gangwon Province	6.76	14.17	7.85
North Chungcheong Province	21.54	14.19	10.97
South Chungcheong Province	12.61	22.58	7.84
North Jeolla Province	10.99	5.33	7.97
South Jeolla Province	11.52	12.27	10.31
North Gyeongsang Province	25.37	33.90	4.82
South Gyeongsang Province	15.92	35.95	34.32
Jeju Special Self-Governing Province	11.75	8.40	15.03

TABLE XVI
CHI-SQUARE VALUE OF 2ND DIGIT BENFORD BY REGION FOR THE 17TH ELECTION

Province	Chung Dong-young	Lee Myung-bak	Lee Hoi-chang
Seoul Metropolitan Government	2.93	9.26	6.23
Busan Metropolitan City	4.51	14.06	3.04
Daegu Metropolitan City	9.60	8.64	7.59
Incheon Metropolitan City	2.92	6.70	13.39
Gwangju Metropolitan City	7.06	10.77	80.01
Daejeon Metropolitan City	10.69	9.10	9.05
Ulsan Metropolitan City	3.57	4.34	16.99
Gyeonggi Province	14.87	10.37	13.62
Gangwon Province	8.28	9.38	18.71
North Chungcheong Province	13.65	7.98	7.15
South Chungcheong Province	14.15	11.98	13.19
North Jeolla Province	8.84	15.10	147.50
South Jeolla Province	12.31	14.18	356.37
North Gyeongsang Province	21.43	9.33	4.81
South Gyeongsang Province	13.84	18.47	21.47
Jeju Special Self-Governing Province	4.62	5.48	25.30

TABLE XVII
CHI-SQUARE VALUE OF 3RD DIGIT BENFORD BY REGION FOR THE 17TH ELECTION

In the 2nd digit Benford, notable regions are Seoul, Busan, Gyeonggi-do, and Daegu. In Seoul, all candidates exceeded the critical value, but Lee Myung-bak's was particularly high at 245.09. In Busan, only Lee Myung-bak's Chi-squared value exceeded the critical value at 62.46. In Gyeonggi-do, both Chung Dong-young and Lee Myung-bak exceeded the critical value, with Lee Myung-bak's being significantly higher at 91.96. In Daegu, Chung Dong-young's Chi-squared value was notably high at 47.06.

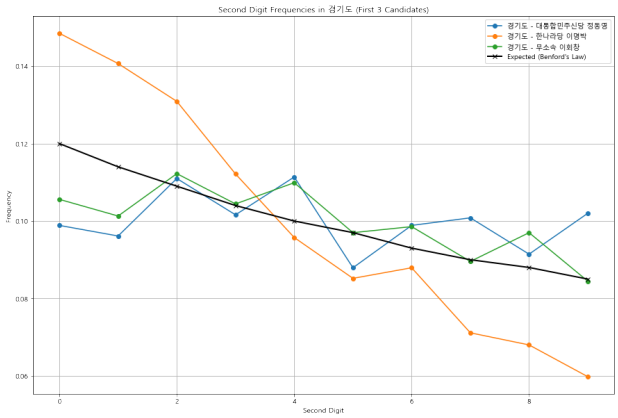

Fig. 22. 17th Election Gyeonggi-do Second Digit Frequencies

In the 3rd digit Benford, Jeollanam-do and Jeollabuk-do

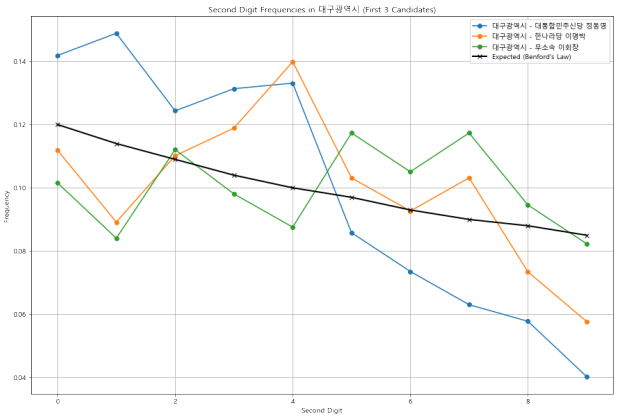

Fig. 23. 17th Election Daegu Second Digit Frequencies

are regions of note. In contrast to other regions, only in Jeolla, Lee Hoi-chang's Chi-squared values significantly exceeded the critical value and the values of other candidates. Specifically, in Jeollanam-do, it was 356.37 and in Jeollabuk-do, 147.50, indicating abnormal voting results for Lee Hoi-chang in these regions.

In conclusion, for the 17th election, Lee Myung-bak's election data appeared abnormal in the 2nd digit Benford analysis, especially in Seoul, Busan, and Gyeonggi-do. Meanwhile, in the 3rd digit Benford, Lee Hoi-chang's data in Jeolla was abnormal.

I. 18th Election

1) National Data of the 18th Election: The results of applying Benford's Law to the second and third digits using vote counts from all polling stations in the 18th national election data are as follows:

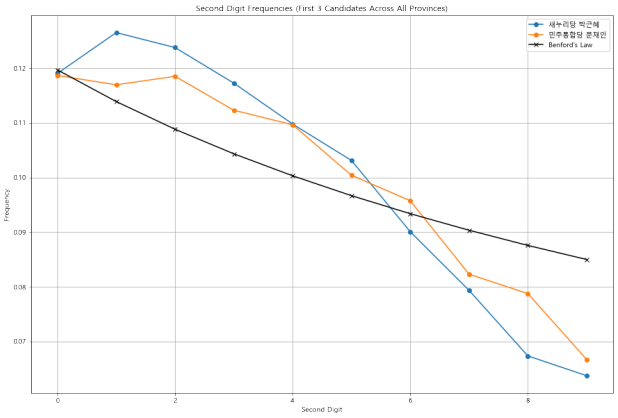

Fig. 24. 18th Election Second Digit Frequencies

The Chi-squared values calculated for candidates Park Geun-hye and Moon Jae-in were:

First, in the 2nd digit Benford, both candidates' Chi-squared values exceeded the critical value, with Park Geun-hye's being relatively higher. Second, in the 3rd digit Benford, both candidates showed Chi-squared values below the critical value.

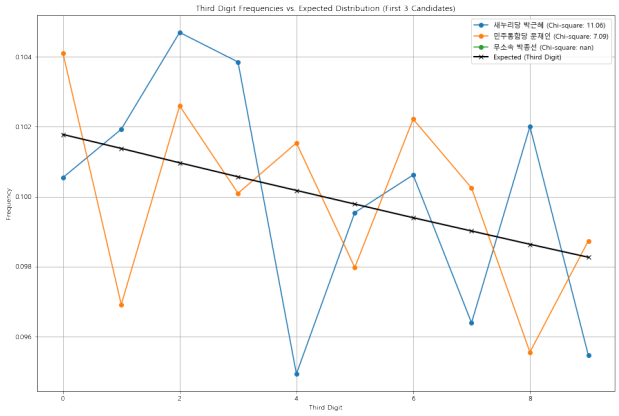

Fig. 25. 18th Election Third Digit Frequencies

Candidate	2nd Digit Benford	3rd Digit Benford
Park Geun-hye	310.95	11.06
Moon Jae-in	142.61	7.09

TABLE XVIII
18TH ELECTION CHI-SQUARED VALUES FOR CANDIDATES

2) Regional Data of the 18th Election: The following are the results of applying Benford's Law regionally.

Province	Park Geun-hye	Moon Jae-in
Seoul Metropolitan Government	272.91	130.77
Busan Metropolitan City	44.20	57.73
Daegu Metropolitan City	29.18	9.05
Incheon Metropolitan City	79.68	42.86
Gwangju Metropolitan City	10.94	6.82
Daejeon Metropolitan City	30.52	15.05
Ulsan Metropolitan City	31.41	5.88
Sejong Special Autonomous City	2.95	14.21
Gyeonggi Province	140.02	89.60
Gangwon Province	14.49	16.49
North Chungcheong Province	33.22	9.71
South Chungcheong Province	20.94	7.54
North Jeolla Province	6.70	2.98
South Jeolla Province	9.86	7.72
North Gyeongsang Province	14.01	5.68
South Gyeongsang Province	4.56	10.63
Jeju Special Self-Governing Province	20.70	4.80

TABLE XIX
CHI-SQUARE VALUE OF 2ND DIGIT BENFORD BY REGION FOR THE 18TH ELECTION

In the 2nd digit Benford, interestingly, Moon Jae-in's Chi-squared values exceeded the critical value only in Seoul, Busan, Incheon, and Gyeonggi-do, and in all but Busan, his values were lower than Park Geun-hye's. This shows how significantly higher Park Geun-hye's values were in comparison, suggesting potential election manipulation for Park Geun-hye.

In the 3rd digit Benford, most Chi-squared values did not exceed the critical value, and even those that did were within 23, making it difficult to suspect election manipulation.

In summary, for the 18th election, Park Geun-hye's election data raised suspicions, especially in the 2nd digit Benford analysis. Although the 3rd digit Benford results were within

Province	Park Geun-hye	Moon Jae-in
Seoul Metropolitan Government	9.36	13.98
Busan Metropolitan City	6.56	4.11
Daegu Metropolitan City	7.74	7.61
Incheon Metropolitan City	15.51	10.02
Gwangju Metropolitan City	19.90	5.45
Daejeon Metropolitan City	5.31	8.26
Ulsan Metropolitan City	10.26	6.05
Sejong Special Autonomous City	1.95	8.04
Gyeonggi Province	12.87	3.09
Gangwon Province	14.24	12.26
North Chungcheong Province	6.16	9.06
South Chungcheong Province	4.02	4.45
North Jeolla Province	10.19	13.07
South Jeolla Province	10.22	6.00
North Gyeongsang Province	8.54	6.38
South Gyeongsang Province	11.95	10.61
Jeju Special Self-Governing Province	7.22	22.23

TABLE XX
CHI-SQUARE VALUE OF 3RD DIGIT BENFORD BY REGION FOR THE 18TH ELECTION

the critical value, the significantly higher Chi-squared values for Park Geun-hye in the 2nd digit analysis suggest potential manipulation.

J. 19th Election

1) National Data of the 19th Election: The results of applying Benford's Law to the second and third digits using vote counts from all polling stations in the 19th national election data are as follows:

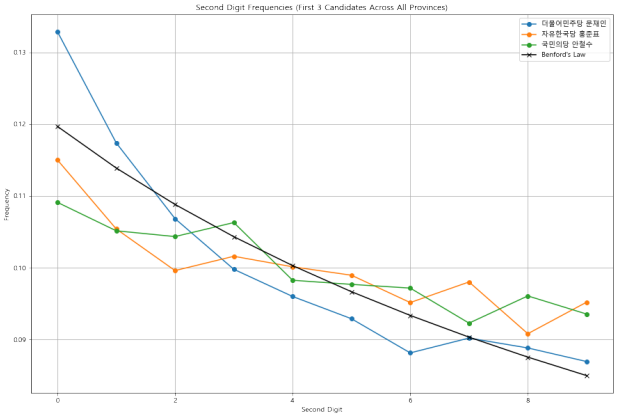

Fig. 26. 19th Election Second Digit Frequencies

The Chi-squared values calculated for candidates Moon Jae-in, Hong Joon-pyo, and Ahn Cheol-soo were:

Candidate	2nd Digit Benford	3rd Digit Benford
Moon Jae-in	43.28	6.39
Hong Joon-pyo	65.56	7.99
Ahn Cheol-soo	65.73	8.69

TABLE XXI
19TH ELECTION CHI-SQUARED VALUES FOR CANDIDATES

First, in the 2nd digit Benford, all three candidates' Chi-squared values exceeded the critical value, but the differences between the candidates were not substantial, making it difficult to suspect election manipulation. Second, in the 3rd digit Benford, all candidates showed Chi-squared values below the critical value.

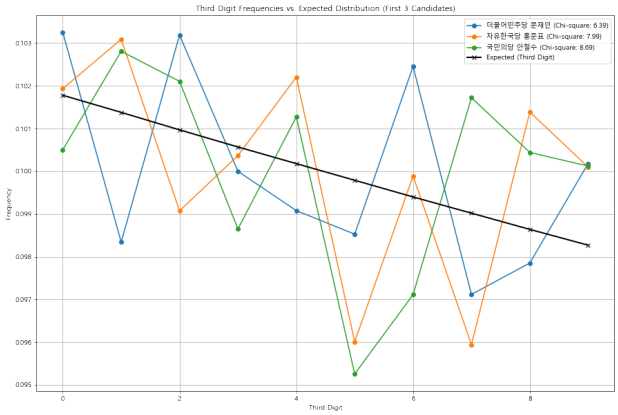

Fig. 27. 19th Election Third Digit Frequencies

2) Regional Data of the 19th Election: The following are the results of applying Benford's Law regionally.

Province	Moon Jae-in	Hong Joon-pyo	Ahn Cheol-soo
Seoul Metropolitan Government	42.47	33.64	19.36
Busan Metropolitan City	8.39	7.91	17.79
Daegu Metropolitan City	12.85	35.35	8.98
Incheon Metropolitan City	23.95	7.59	16.32
Gwangju Metropolitan City	25.13	15.19	11.56
Daejeon Metropolitan City	13.99	19.22	14.46
Ulsan Metropolitan City	2.98	20.98	18.73
Sejong Special Autonomous City	5.30	18.05	14.72
Gyeonggi Province	43.44	33.73	23.90
Gangwon Province	14.17	8.31	6.48
North Chungcheong Province	7.48	14.34	17.17
South Chungcheong Province	5.52	15.65	5.62
North Jeolla Province	8.16	12.03	14.09
South Jeolla Province	18.54	13.13	15.06
North Gyeongsang Province	10.32	4.60	12.83
South Gyeongsang Province	15.92	14.45	6.94
Jeju Special Self-Governing Province	7.10	5.08	9.05

TABLE XXII
CHI-SQUARE VALUE OF 2ND DIGIT BENFORD BY REGION FOR THE 19TH ELECTION

Province	Moon Jae-in	Hong Joon-pyo	Ahn Cheol-soo
Seoul Metropolitan Government	9.81	9.31	12.33
Busan Metropolitan City	12.57	6.54	4.97
Daegu Metropolitan City	12.08	7.94	9.80
Incheon Metropolitan City	12.29	13.12	13.81
Gwangju Metropolitan City	7.05	4167.33	9.80
Daejeon Metropolitan City	4.42	6.25	4.49
Ulsan Metropolitan City	1.73	6.73	3.82
Sejong Special Autonomous City	7.80	10.98	17.02
Gyeonggi Province	1.84	3.75	5.03
Gangwon Province	8.57	8.53	8.76
North Chungcheong Province	4.41	8.54	16.67
South Chungcheong Province	11.62	6.03	25.52
North Jeolla Province	10.75	69.70	6.54
South Jeolla Province	10.96	844.22	5.12
North Gyeongsang Province	17.78	6.33	31.43
South Gyeongsang Province	11.90	8.50	29.42
Jeju Special Self-Governing Province	2.84	5.12	18.31

TABLE XXIII
CHI-SQUARE VALUE OF 3RD DIGIT BENFORD BY REGION FOR THE 19TH ELECTION

In the 2nd digit Benford, notable regions are Seoul, Daegu,

and Gyeonggi-do. In Seoul and Gyeonggi-do, all candidates exceeded the critical value, but the differences in their Chi-squared values were not substantial. In Daegu, only Hong Joon-pyo's Chi-squared value exceeded the critical value at 35.35.

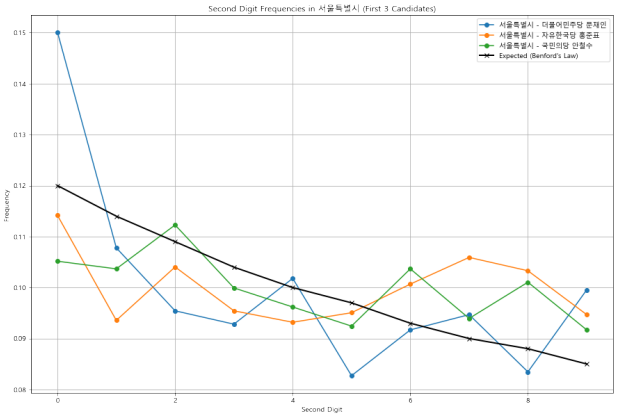

Fig. 28. 19th Election Seoul Second Digit Frequencies

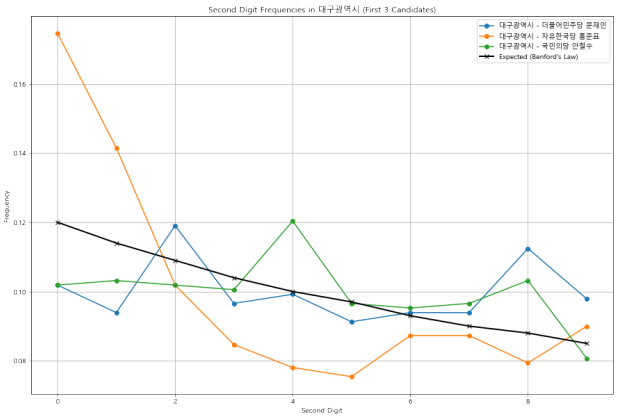

Fig. 29. 19th Election Daegu Second Digit Frequencies

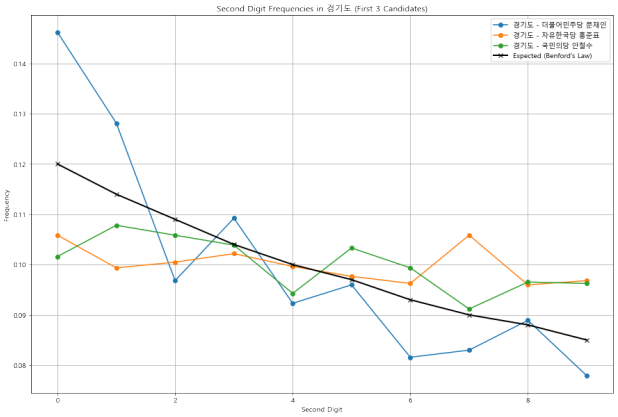

Fig. 30. 19th Election Gyeonggi-do Second Digit Frequencies

In the 3rd digit Benford, Gwangju and Jeollanam-do are regions of note. In these regions, Hong Joon-pyo's Chi-squared values were extremely high at 4167.33 and 844.22, respectively, indicating possible election manipulation for Hong Joon-pyo in these regions.

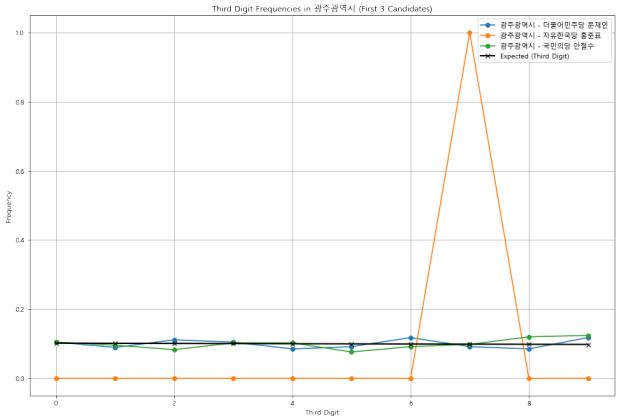

Fig. 31. 19th Election Gwangju Third Digit Frequencies

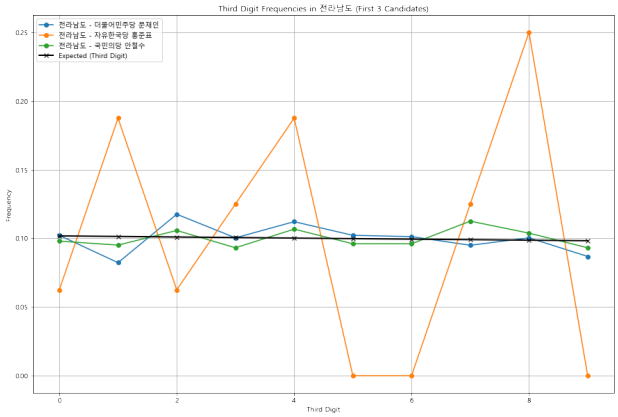

Fig. 32. 19th Election Jeollanam-do Third Digit Frequencies

In conclusion, for the 19th election, there were no substantial clues for election manipulation in the 2nd digit Benford analysis. However, the 3rd digit Benford analysis revealed abnormal results for Hong Joon-pyo in Gwangju and Jeollanam-do, suggesting possible region-specific election manipulation.

K. 20th Election

1) National Data of the 20th Election: The results of applying Benford's Law to the second and third digits using vote counts from all polling stations in the 20th national election data are as follows:

The Chi-squared values calculated for candidates Lee Jae-myung and Yoon Seok-youl were:

In both analyses, Yoon Seok-youl's Chi-squared values exceeded the critical value and were significantly higher than Lee Jae-myung's, suggesting potential election manipulation for Yoon Seok-youl.

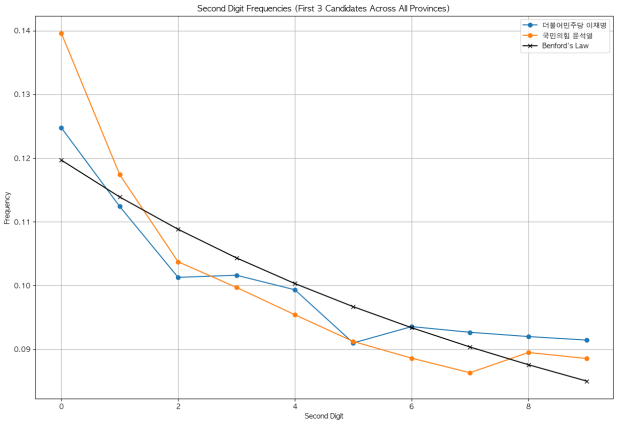

Fig. 33. 20th Election Second Digit Frequencies

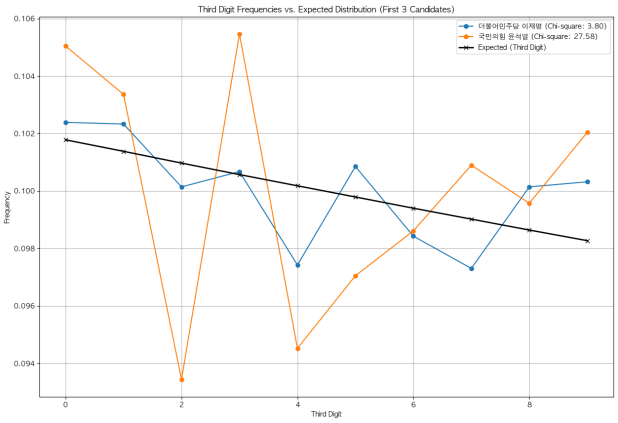

Fig. 34. 20th Election Third Digit Frequencies

Candidate	2nd Digit Benford	3rd Digit Benford
Lee Jae-myung	33.89	3.80
Yoon Seok-youl	87.77	27.58

TABLE XXIV
20TH ELECTION CHI-SQUARED VALUES FOR CANDIDATES

2) Regional Data of the 20th Election: The following are the results of applying Benford's Law regionally.

In the 2nd digit Benford, focus on the Chi-squared values in Seoul, Daejeon, Daegu, and Busan. In Seoul and Daejeon, Lee Jae-myung's values were higher, while in Daegu and Busan, Yoon Seok-youl's were higher, all exceeding the critical value. However, these differences were not substantial enough to suspect election manipulation.

In the 3rd digit Benford, most Chi-squared values did not exceed the critical value, and even those that did were within 24, making it difficult to suspect election manipulation.

In summary, the 20th election showed little evidence of election manipulation, especially when analyzed regionally. However, the national data analysis revealed abnormal Chi-squared values for Yoon Seok-youl, suggesting a slight possibility of election manipulation.

3) Subsubsection Heading Here: Subsubsection text here.

Province	Lee Jae-myung	Yoon Suk-yeol
Seoul Metropolitan Government	42.11	23.85
Busan Metropolitan City	6.74	23.75
Daegu Metropolitan City	19.91	41.17
Incheon Metropolitan City	9.62	7.74
Gwangju Metropolitan City	12.01	5.14
Daejeon Metropolitan City	30.05	6.76
Ulsan Metropolitan City	15.93	18.36
Sejong Special Autonomous City	13.78	9.28
Gyeonggi Province	49.13	49.47
Gangwon Province	8.48	16.55
North Chungcheong Province	10.41	7.92
South Chungcheong Province	8.29	18.92
North Jeolla Province	19.65	3.53
South Jeolla Province	4.22	9.29
North Gyeongsang Province	13.91	17.69
South Gyeongsang Province	8.82	11.00
Jeju Special Self-Governing Province	5.64	9.52

TABLE XXV
CHI-SQUARE VALUE OF 2ND DIGIT BENFORD BY REGION FOR THE 20TH ELECTION

Province	Lee Jae-myung	Yoon Suk-yeol
Seoul Metropolitan Government	6.29	10.19
Busan Metropolitan City	3.05	7.23
Daegu Metropolitan City	7.15	4.00
Incheon Metropolitan City	4.94	8.98
Gwangju Metropolitan City	8.73	9.89
Daejeon Metropolitan City	5.60	10.18
Ulsan Metropolitan City	6.60	12.72
Sejong Special Autonomous City	4.35	10.78
Gyeonggi Province	4.32	9.91
Gangwon Province	4.96	10.97
North Chungcheong Province	8.56	6.86
South Chungcheong Province	7.92	12.96
North Jeolla Province	5.21	23.12
South Jeolla Province	12.99	33.76
North Gyeongsang Province	16.18	13.99
South Gyeongsang Province	17.75	16.80
Jeju Special Self-Governing Province	5.14	4.97

TABLE XXVI
CHI-SQUARE VALUE OF 3RD DIGIT BENFORD BY REGION FOR THE 20TH ELECTION

V. CONCLUSION

In this study, we applied Benford's Law to election data from the 13th to the 20th South Korean presidential elections to explore the possibility of electoral fraud. Benford's Law is based on the theory that numerical distributions follow specific patterns, allowing for the assessment of the integrity of election data. For each election, we applied Benford's Law to both national and regional data and analyzed the results using chi-squared tests.

When comparing the second and third digits, we observed that in most elections, the chi-squared values of the 2nd digit Benford's analysis often exceeded the threshold. However, the 3rd digit Benford's analysis, except in specific instances, did not exceed this threshold, suggesting a stronger indication of potential electoral manipulation in the 3rd digit analysis than in the 2nd.

Observing trends across the elections, some showed significantly high chi-squared values in the 2nd digit Benford's analysis for certain candidates, especially in regional data.

These results imply possible electoral manipulations in specific regions, potentially linked to regional political dynamics or unique occurrences during the election processes.

Interestingly, it has been observed that with the passage of time, the overall data from each election increasingly conforms to 2nd digit Benford's Law. However, it is challenging to definitively conclude whether this trend is a result of the realization of democracy, leading to an increase in vote counts and thus improving the goodness-of-fit with Benford's Law, or if it indicates a possibility of subtle manipulation.

Regarding the 3rd digit Benford's Law, there were no clear indications of improvements in more recent elections, and it was difficult to discern a consistent trend across elections. Unlike the 2nd digit, most elections did not exceed the chi-square threshold for the 3rd digit. However, the appearance of abnormally high chi-square values for certain candidates or in specific regions suggests the possibility of election manipulation.

The regional analysis for each election suggested that localized preferences or political backgrounds for specific candidates might have influenced election outcomes. The consistently high chi-squared values in areas like Gwangju Metropolitan City and the Jeolla provinces suggest a strong regional bias, potentially indicative of electoral manipulation.

This research demonstrates the crucial role of statistical analysis in evaluating the transparency of electoral processes. However, since the analysis was conducted on data where clear evidence of manipulation was not found, the results should be viewed as signals pointing to the need for further investigation, rather than as definitive evidence of electoral fraud.

In conclusion, our study provides significant insights into assessing the transparency and accuracy of election processes through statistical analysis.

This methodology, when applied to election data, should be used alongside broader investigations, and caution is required in its independent interpretation. Moreover, when analyzing potential electoral manipulation, it is essential to consider changing patterns of preferences over time and regional characteristics, necessitating deeper analyses and understanding of each election.

REFERENCES

[1] Benford, F. *The Law of Anomalous Numbers* , Proceedings of the American Philosophical Society Vol.78 No.4 , American Philosophical Society, 1938.

[2] Durtschi, C. , Hillison, W. , Pacini, C. *The effective use of Benford's law to assist in detecting fraud in accounting data*, Journal of Forensic Accounting Vol.5, 2004.

[3] Nigrini, M.J., Mittermaier, L.J., *The Use of Benford's Law as an Aid in Analytical Procedures*, Auditing: A Journal of Practice & Theory, Vol. 16, 1997.

[4] Mebane, W.R. *Election Forensics: The Second-digit Benford's Law Test and Recent American Presidential Elections*, Summer Meeting of the Political Methodology Society, 2006.

[5] Breunig, C., Goerres, A. *Searching for electoral irregularities in an established democracy: Applying Benford's Law tests to Bundestag elections in Unified Germany* , Electoral Studies, Vol.30, Issue 3, 2011.

[6] Deckert, J., Myagkov, M., and Ordeshook, P. C. *Benford's Law and the Detection of Election Fraud* , Political Analysis, Vol.19, Issue 3, 2011.

[7] Morrow, J. *Benford's Law, Families of Distributions and a Test Basis*, CEP Discussion Paper No 1291, August 2014.

[8] Roukema, B.F. *A first-digit anomaly in the 2009 Iranian presidential election* , Journal of Applied Statistics, Vol.41, Issue 1, 2014.

[9] Chae, H., Kim, S. *Conservatives and Progressives in South Korea* , 31 Washington Q. 77, 2017.