

피보나치수와 벤포드법칙에 대한 탐색적 접근

장대흥¹

¹부경대학교 수리과학부 통계학전공

(2009년 7월 접수, 2009년 8월 채택)

요약

피보나치수열의 첫 숫자수열이 벤포드법칙을 따름은 알려진 사실이다. 이러한 피보나치수열을 확장하여 임의의 두 개의 자연수를 정하고 재귀식 $a_{n+2} = a_{n+1} + a_n$ 을 만족하는 수열을 만들었을 때 이 수열의 첫 숫자수열이 벤포드법칙을 만족하는지를 확인하고 이러한 수열의 첫 숫자수열의 구조를 탐색적 자료분석의 입장에서 살펴보았다.

주요용어: 피보나치수, 루카스수, 벤포드법칙.

1. 서론

피보나치수는 다음과 같이 재귀적으로 정의되는 수열이다.

$$\begin{aligned} F(1) &= 1, \\ F(2) &= 1, \\ F(n+2) &= F(n+1) + F(n), \quad n \geq 1. \end{aligned}$$

물론 $F(0) = 0$ 부터 시작할 수도 있다. 그러면 피보나치수는 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, ...로 나타난다. 이러한 피보나치수의 첫 숫자(first digits, leading digits)수열 1, 1, 2, 3, 5, 8, 1, 2, 3, 5, 8, 1, ...이 벤포드법칙을 따름은 알려진 사실이다. Cenessa (2003)는 피보나치수의 첫 숫자수열이 왜 벤포드법칙을 따르는지를 통계열역학 이론을 이용하여 설명하였다. 피보나치수와 비슷한 수열이 있는데 이 수열이 루카스수이다. 루카스수는 다음과 같이 재귀적으로 정의되는 수열이다.

$$\begin{aligned} L(1) &= 1, \\ L(2) &= 3, \\ L(n+2) &= L(n+1) + L(n), \quad n \geq 1. \end{aligned}$$

물론 $L(0) = 2$ 로부터 시작할 수도 있다. 그러면 루카스수는 1, 3, 4, 7, 11, 18, 29, 47, 76, 123, ...로 나타난다. Giles (2007)는 자료의 원형성(circularity) 성질에 불변(invariant)인 Kuiper검정을 사용하여 이러한 루카스수의 첫 숫자수열 1, 3, 4, 7, 1, 1, 2, 4, 7, 1, ...이 벤포드법칙을 따름을 증명하였다.

Atalay (2004) 책에 보면 다음과 같은 단락이 나온다.

“임의의 한 쌍으로 시작한다고 가정하면, 첫째와 둘째 항을 넣어 셋째 항을 얻고 둘째와 셋째 항을 넣어 넷째 항을 얻고, 이런 식으로 계속하여 1과 1의 쌍으로 시작하는 피보나치의 본래 수열과 흡사한 수

¹(608-737) 부산시 남구 대연3동 599-1 부경대학교 수리과학부 통계학전공, 교수. E-mail: dhjang@pknu.ac.kr

열을 만들 수 있다. 그리고 여기에서 우리는, 어떤 수열에서도 연속하는 쌍들의 비율은 궁극적으로 마법의 1.618034...로 수렴한다는 사실을 발견하게 된다. 개인의 생일을 예로 한번 실험해 보자. 내 생일인 6월 10일은 6, 10 혹은 10, 6으로 나타낼 수 있는데, 비율을 재빨리 계산하는데 특히 편하다. 6, 10, 16, 26, 42, 68, 110, ... 이렇게 수열을 만들어 계산해 보면 5항이나 6항 이후로 비율은 1.618로 정리된다. 조지 워싱턴의 생일인 2월 22일로 수열을 만들어 보면 다음과 같다. 2, 22, 24, 46, 70, 116, 186, 302, 488, 790, 1278, ... 이 경우 최소한 10항 이상은 나열해야, 소수 여섯 번째 자리까지 정밀하게 표현하였을 때 이 비율이 나타난다.”

위에서 언급한 수열은 다음과 같이 재귀적으로 정의되는 수열이다. 여기서 a 와 b 는 임의의 자연수라 하자. 이 수열은 피보나치수와 루카스수를 특수한 경우로서 포함한다.

$$\begin{aligned} a_1 &= a, \\ a_2 &= b, \\ a_{n+2} &= a_{n+1} + a_n, \quad \forall n \geq 1. \end{aligned}$$

이러한 수열을 피보나치계수열이라 칭하자. 이산수학에서 재귀식 $t_n = At_{n-1} + Bt_{n-2}$ 에 대한 해는 $t_n = c_1r_1^n + c_2r_2^n$ 이 된다. 여기서, c_1 과 c_2 는 상수이고 r_1 과 r_2 는 특성방정식 $r^2 - Ar - B = 0$ 에 대한 두 개의 근들이다. 그러므로 피보나치계수열에서 재귀식 $a_{n+2} = a_{n+1} + a_n$ 에 대한 특성방정식은 $\phi^2 - \phi - 1 = 0$ 이 되고 이 해는 $\phi = (1 \pm \sqrt{5})/2$ 이 된다. 이 해 중 $\phi = (1 + \sqrt{5})/2$ 이 황금비(golden ratio)가 된다. 피보나치계수열에서 연속하는 쌍들의 비율의 극한값이 황금비이다. 즉

$$\lim_{n \rightarrow \infty} \frac{a_{n+1}}{a_n} = \phi \simeq 1.618034$$

이다.

다음 절에서 탐색적자료분석을 통하여 이러한 피보나치계수열의 첫 숫자수열이 벤포드법칙을 따름을 보이고 이러한 피보나치계수열의 첫 숫자수열의 구조를 알아보았다.

2. 피보나치계수열과 벤포드법칙

벤포드법칙이란 인간이 사회생활을 하면서 발생시키는 숫자들의 첫 숫자가 각각 1, 2, 3, 4, 5, 6, 7, 8, 9가 될 확률은 균등분포처럼 각각 1/9이 되지 못하고 $\log_{10}(1 + 1/i)$ ($i = 1, 2, 3, \dots, 9$)이 되어 제일 큰 것은 1일 때 0.301, 제일 작은 것은 9일 때 0.046이 되고 1에서 9 순으로 발생 확률이 작아진다는 법칙이다.

그림 2.1은 a 와 b 를 1에서 365까지의 자연수 중 임의로 두 수를 의사난수발생기를 이용하여 뽑아 구한 4가지 피보나치계수열의 첫 숫자수열(각각 1,450개의 첫 숫자수열)에 대한 막대그래프이다. 4가지 피보나치계수열의 첫 숫자수열 모두 벤포드법칙을 따름을 확인할 수 있다.

4가지 피보나치계수열의 첫 숫자수열(각각 1,450개의 첫 숫자수열)에 대하여 벤포드법칙을 따르는 지를 검정하기 위하여 카이제곱검정 및 Kuiper검정을 행하니 표 2.1과 같았다. 두 검정 모두에서 4가지 피보나치계수열의 첫 숫자수열 모두 벤포드법칙을 만족함을 알 수 있다.

그림 2.2는 a 와 b 가 아주 큰 차이가 나거나 같이 크거나 작은 4가지 경우 피보나치계수열의 첫 숫자수열(각각 1,450개의 첫 숫자수열)에 대한 막대그래프이다. 4가지 피보나치계수열의 첫 숫자수열 모두 벤포드법칙을 따름을 확인할 수 있다.

a 와 b 가 아주 큰 차이가 나거나 같이 크거나 작은 4가지 경우 피보나치계수열의 첫 숫자수열(각각 1,450개의 첫 숫자수열)에 대하여 벤포드법칙을 따르는 지를 검정하기 위하여 카이제곱검정 및

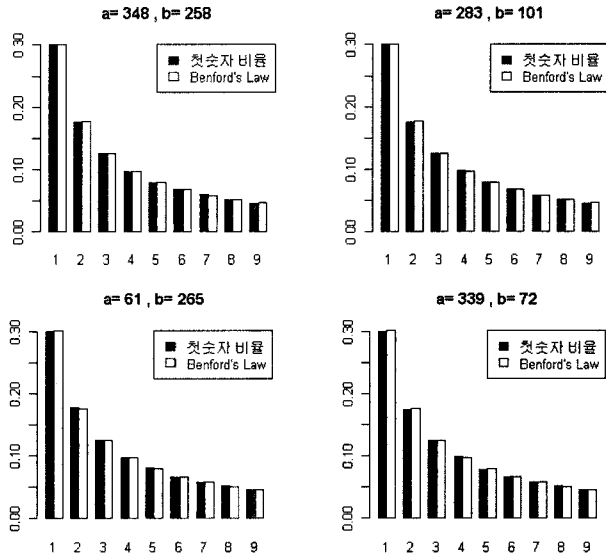


그림 2.1. 피보나치계수열의 첫 숫자수열에 대한 막대그래프와 벤포드법칙(I)

표 2.1. 4가지 피보나치계수열의 첫 숫자수열에 대한 검정통계량(I)

피보나치계수열	카이제곱 검정통계량	Kuiper 검정통계량
$a = 348, b = 258$	0.0569	0.0704
$a = 283, b = 101$	0.0236	0.0348
$a = 61, b = 265$	0.0904	0.0577
$a = 339, b = 72$	0.1090	0.0880

카이제곱 임계값: 15.51(자유도: 8, 유의수준: 5%), Kuiper 임계값: 1.747(유의수준: 5%)

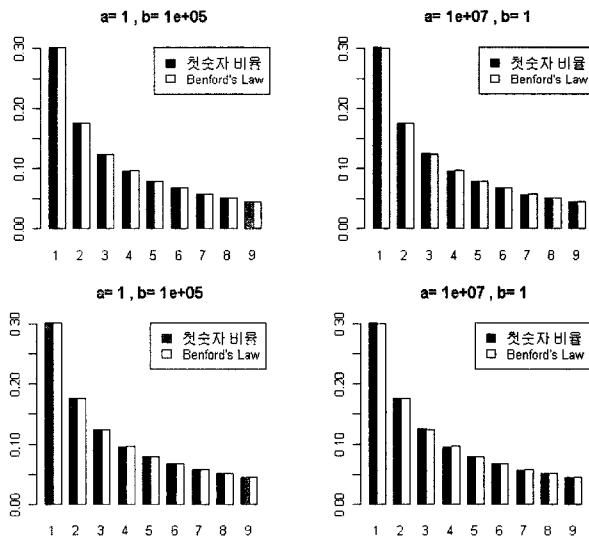


그림 2.2. 피보나치계수열의 첫 숫자수열에 대한 막대그래프와 벤포드법칙(II)

표 2.2. 4가지 피보나치계수열의 첫 숫자수열에 대한 검정통계량(II)

피보나치계수열	카이제곱 검정통계량	Kuiper 검정통계량
$a = 1, b = 10^5$	0.0491	0.0700
$a = 10^7, b = 1$	0.0993	0.1579
$a = 10^7, b = 10^5$	0.0949	0.1716
$a = 1, b = 1$	0.0300	0.0436

카이제곱 임계값: 15.51(자유도: 8, 유의수준: 5%), Kuiper 임계값: 1.747(유의수준: 5%)

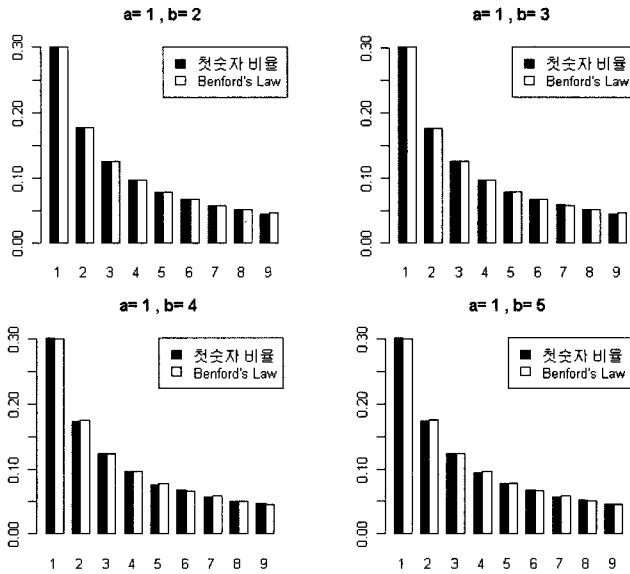


그림 2.3. 피보나치계수열의 첫 숫자수열에 대한 막대그래프와 벤포드법칙(III)

표 2.3. 4가지 피보나치계수열의 첫 숫자수열에 대한 검정통계량(III)

피보나치계수열	카이제곱 검정통계량	Kuiper 검정통계량
$a=1, b=2$	0.0161	0.0225
$a=1, b=3$	0.0310	0.0355
$a=1, b=4$	0.1264	0.0596
$a=1, b=5$	0.0749	0.0795

카이제곱 임계값: 15.51(자유도: 8, 유의수준: 5%), Kuiper 임계값: 1.747(유의수준: 5%)

Kuiper검정을 행하니 다음 표 2.2와 같았다. 두 검정 모두에서 4가지 피보나치계수열의 첫 숫자수열 모두 벤포드법칙을 만족함을 알 수 있다.

그림 2.3은 $a = 1$ 이고 b 가 각각 2, 3, 4, 5인 경우 피보나치계수열의 첫 숫자수열(각각 1,450개의 첫 숫자수열)에 대한 막대그래프이다. 4가지 피보나치계수열의 첫 숫자수열 모두 벤포드법칙을 따름을 확인할 수 있다.

$a = 1$ 이고 b 가 각각 2, 3, 4, 5인 경우 피보나치계수열의 첫 숫자수열(각각 1,450개의 첫 숫자수열)에 대하여 벤포드법칙을 따르는 지를 검정하기 위하여 카이제곱검정 및 Kuiper검정을 행하니 표 2.3과 같았다. 두 검정 모두에서 4가지 피보나치계수열의 첫 숫자수열 모두 벤포드법칙을 만족함을 알 수 있다.

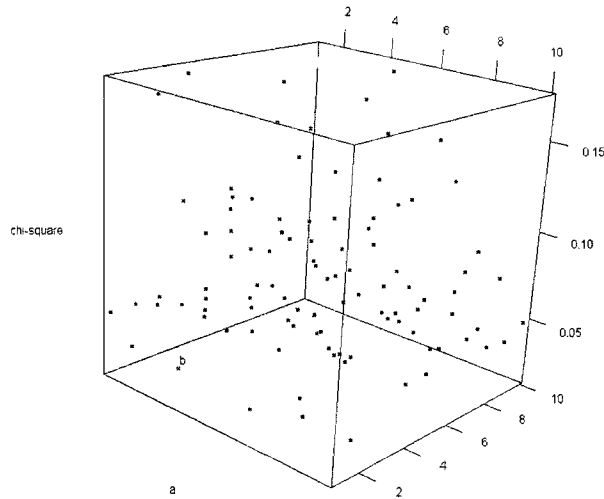


그림 2.4. $a = 1 \sim 10, b = 1 \sim 10$ 의 조합 100가지 경우 피보나치계수열의 첫 숫자수열에 대한 카이제곱검정통계량들

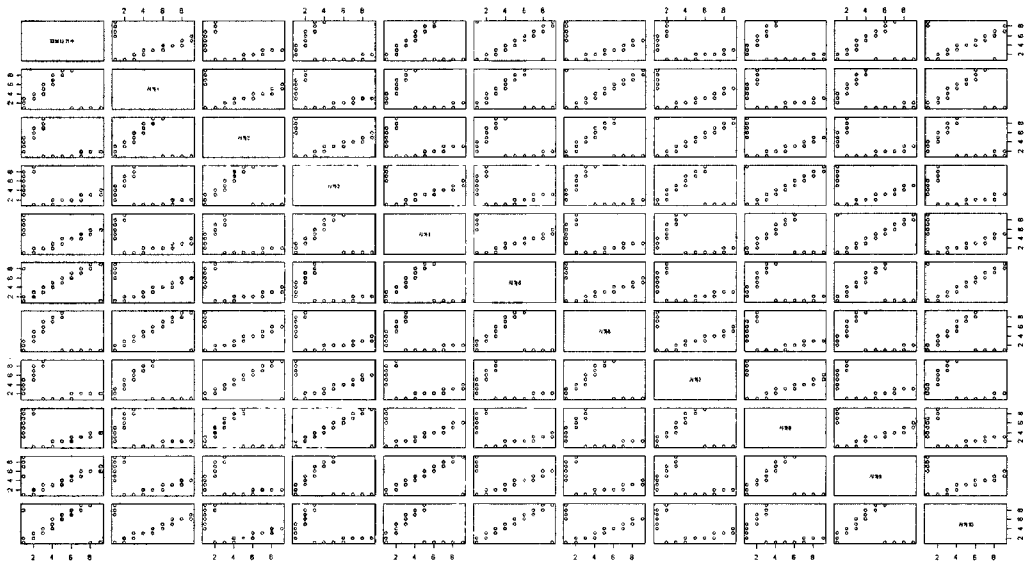


그림 2.5. 피보나치수의 첫 숫자수열과 시차수열에 대한 산점도행렬

$a = 1 \sim 10, b = 1 \sim 10$ 의 조합 100가지 경우 각각에 대하여 피보나치계수열의 첫 숫자수열이 벤포드법칙을 따르는 지를 검증하기 위하여 카이제곱검정통계량들을 구하여 보니 그림 2.4와 같았다. 모두 카이제곱 임계값인 15.51(자유도: 8, 유의수준: 5%)보다 훨씬 작은 값들임을 알 수 있다.

피보나치계수열의 첫 숫자수열에 대한 구조를 알기 위하여 피보나치계수열 중 세 가지 수열(피보나치수, 루카스수, $a = 1, b = 365$ 인 피보나치계수열)을 대상으로 첫 숫자수열과 첫 숫자수열의 시차(lag)수열(시차 1에서 시차 10까지)을 구한 후 산점도행렬을 각각 그리면 그림 2.5~2.7과 같다. 세 가지 첫 숫

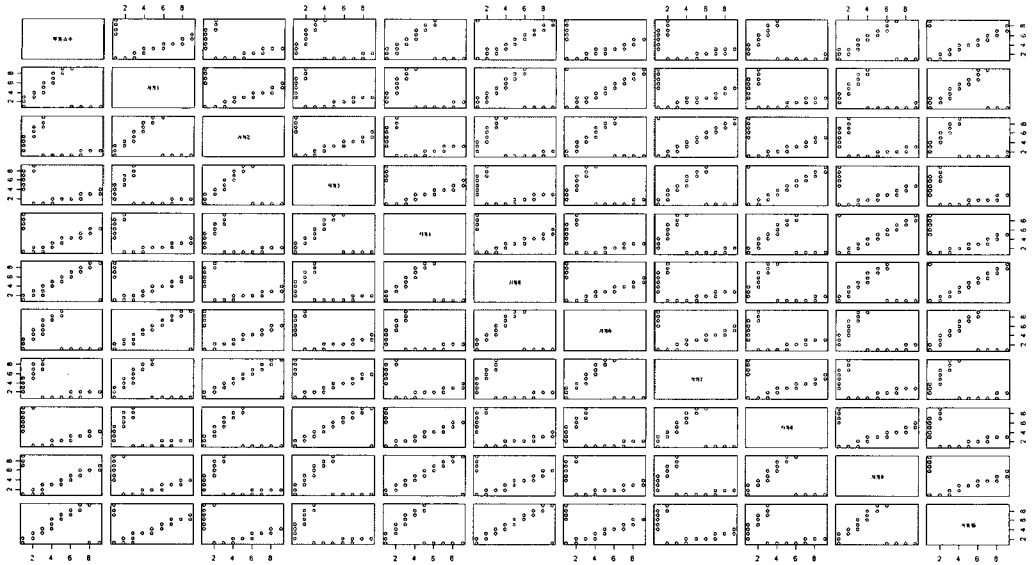


그림 2.6. 루카스수의 첫 숫자수열과 시차수열에 대한 산점도행렬

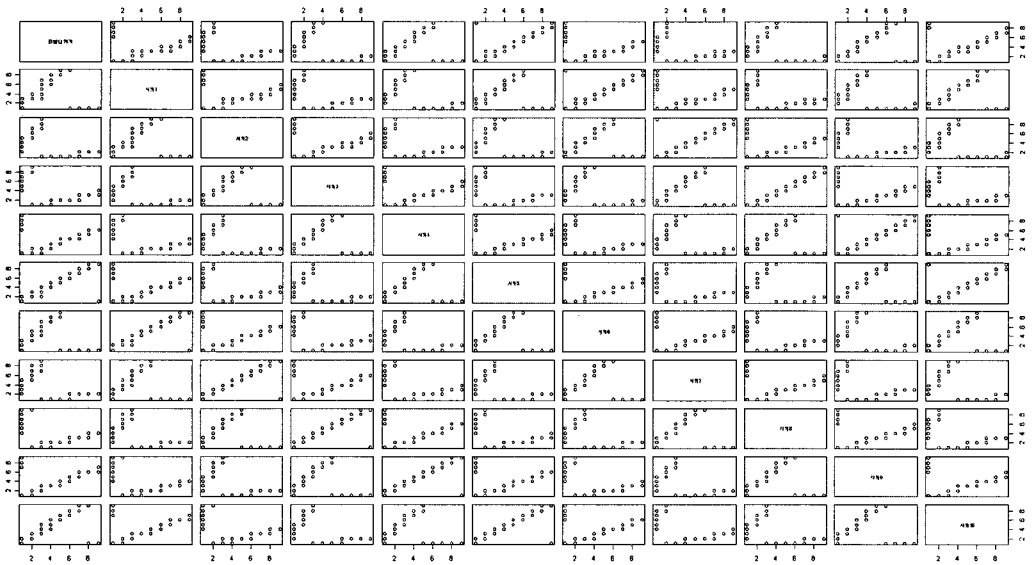


그림 2.7. $a = 1, b = 365$ 인 피보나치계수열의 첫 숫자수열과 시차수열에 대한 산점도행렬

자수열 사이에 통계적 유사성(statistical similarity)이 존재함을 알 수 있다. 또한 각 첫 숫자수열은 특정한 고유 패턴을 갖는 결정론적인 구조를 갖고 있음을 알 수 있다. 예로 피보나치수의 첫 숫자수열과 첫 숫자수열의 시차 1 수열 사이에는 9-1 패턴은 있어도 1-9 패턴은 존재하지 않는다. 각 산점도행렬의 비대각선을 따라 패턴이 아주 유사하거나 같음을 알 수 있다. 예로 첫 숫자수열과 첫 숫자수열의 시차 1

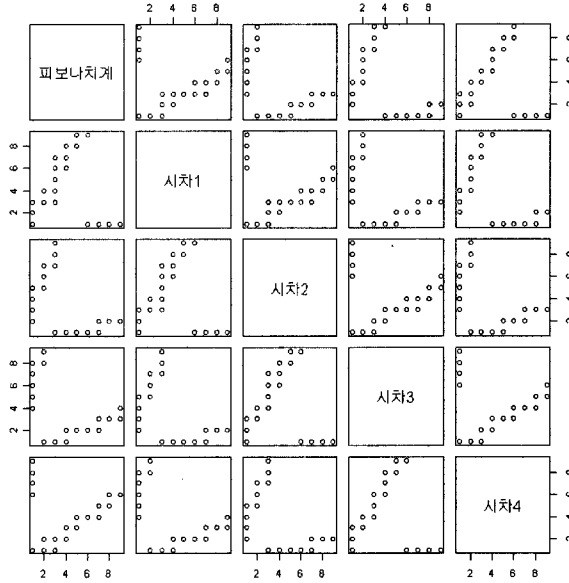


그림 2.8. $a = 1, b = 365$ 인 피보나치계수열의 첫 숫자수열과 시차수열(시차 1에서 시차 4까지)에 대한 산점도행렬

수열, 첫 숫자수열의 시차 1 수열과 시차 2 수열, ..., 첫 숫자수열의 시차 9 수열과 시차 10 수열 사이의 패턴이 아주 유사하거나 같음을 알 수 있다.

산점도행렬을 좀 더 자세히 보기 위하여 $a = 1, b = 365$ 인 피보나치계수열의 첫 숫자수열과 시차수열(시차 1에서 시차 4까지)에 대한 산점도행렬을 그려보면 그림 2.8과 같다. $a = 1, b = 365$ 인 피보나치계수열의 첫 숫자수열이 특정한 고유 패턴을 갖는 결정론적인 구조를 갖고 있음을 알 수 있고 산점도행렬의 비대각선을 따라 패턴이 아주 유사하거나 같음을 알 수 있다. 즉, 첫 숫자수열과 첫 숫자수열의 시차 1 수열, 첫 숫자수열의 시차 1 수열과 시차 2 수열, 첫 숫자수열의 시차 2 수열과 시차 3 수열, 첫 숫자수열의 시차 3 수열과 시차 4 수열 사이의 패턴이 아주 유사하거나 같음을 알 수 있고 첫 숫자수열과 첫 숫자수열의 시차 2 수열, 첫 숫자수열의 시차 1 수열과 시차 3 수열, 첫 숫자수열의 시차 2 수열과 시차 4 수열의 패턴이 같음을 알 수 있다. 또한 첫 숫자수열과 첫 숫자수열의 시차 3 수열, 첫 숫자수열의 시차 1 수열과 시차 4 수열의 패턴이 같음을 알 수 있다.

다른 시각에서 피보나치계수열의 첫 숫자수열에 대한 구조를 알기 위하여 앞에서 언급한 세 가지 수열(피보나치수, 루카스수, $a = 1, b = 365$ 인 피보나치계수열)을 대상으로 첫 숫자수열과 첫 숫자수열의 시차수열을 이용하여 평행좌표그림을 그리면(호트림(jittering)을 적용함) 그림 2.9~2.11과 같다. 우리는 다시 한 번 세 가지 첫 숫자수열 사이에 통계적 상사성이 존재함을 확인할 수 있다. 또한 각 첫 숫자수열은 특정한 고유 패턴을 갖는 결정론적인 구조를 갖고 있음도 확인할 수 있다. 아울러 첫 숫자수열과 첫 숫자수열의 시차수열 사이에는 어떤 $i-j$ 패턴($i, j = 1, 2, \dots, 9$)이 존재하며 이 패턴의 강도는 어느 정도인지를 알 수 있다.

$a = 1, b = 365$ 인 피보나치계수열의 첫 숫자수열에서 1항에서 100항까지를 대상으로 꺾은선그래프를 그려보면 다음 그림 2.12와 같다. 특정한 고유 패턴을 갖고 있음(작은 주기가 4나 5를 갖는 패턴이고 큰 주기도 4나 5를 갖는 패턴)을 알 수 있다. 보나치수나 루카스수에 대해서도 각각 꺾은선그래프를 그려

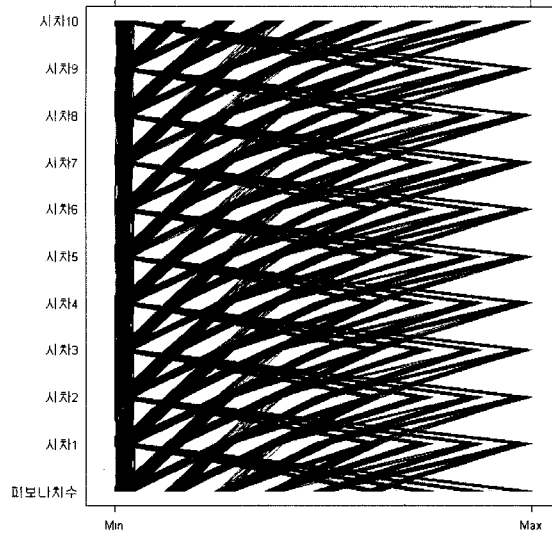


그림 2.9. 피보나치수의 첫 숫자수열과 시차수열에 대한 평행좌표그림

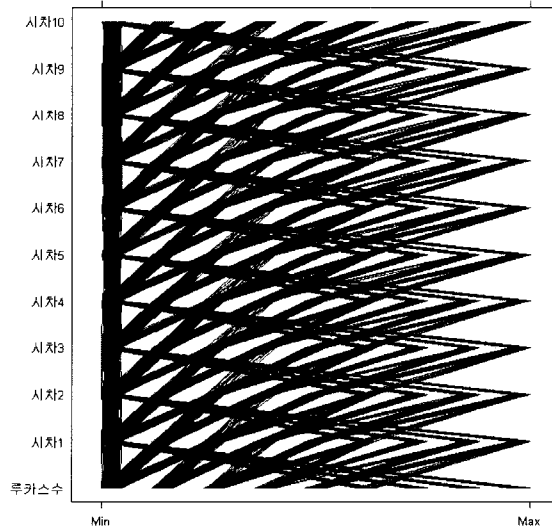


그림 2.10. 루카스수의 첫 숫자수열과 시차수열에 대한 평행좌표그림

보면 피보나치수나 루카스수도 특정한 고유 패턴을 갖고 있음(작은 주기가 4나 5를 갖는 패턴이고 큰 주기도 4나 5를 갖는 패턴)을 알 수 있다.

$a = 1, b = 365$ 인 피보나치계수열의 첫 숫자수열에 대하여 자기상관함수그림을 그려보면 그림 2.13과 같다. 시차가 1,000을 넘어도 자기상관이 유의한 패턴을 가지고 있음을 알 수 있다. 마찬가지로 피보나치수나 루카스수에 대해서도 각각 자기상관함수그림을 그려보면 피보나치수나 루카스수도 시차가 1,000을 넘어도 자기상관이 유의한 패턴을 갖고 있음을 확인할 수 있다.

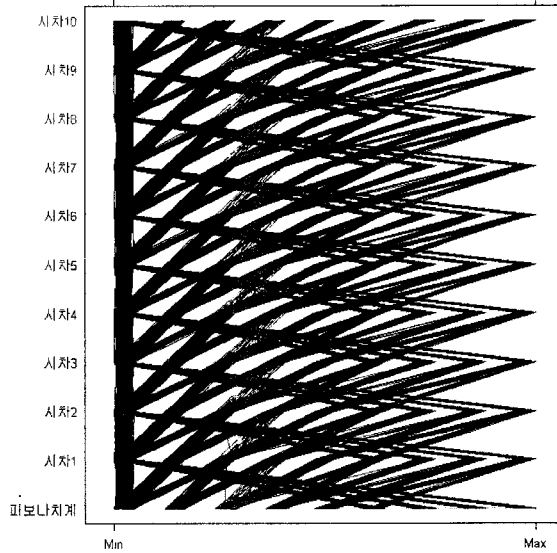


그림 2.11. $a = 1, b = 365$ 인 피보나치계수열의 첫 숫자수열과 시차수열에 대한 평행좌표그림

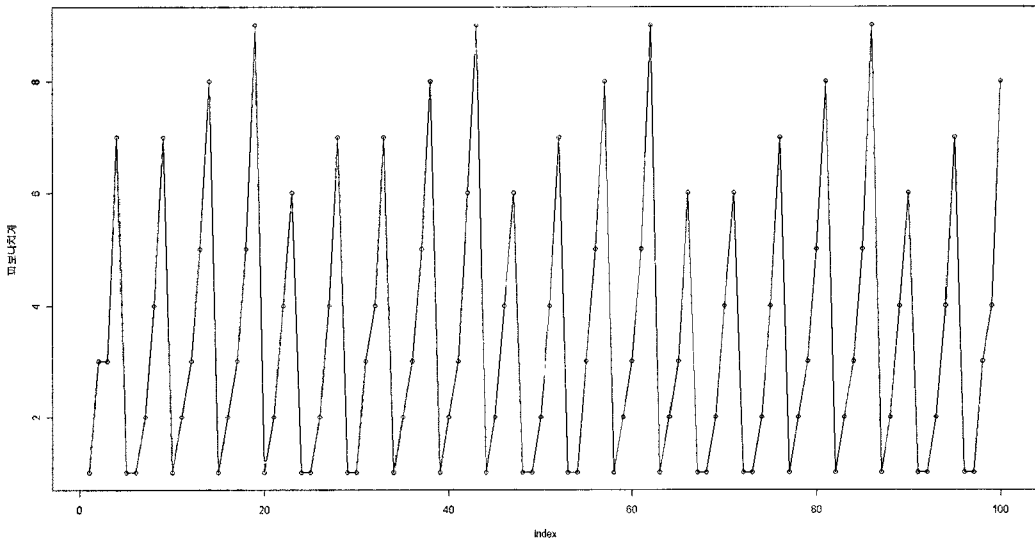


그림 2.12. $a = 1, b = 365$ 인 피보나치계수열의 첫 숫자수열에 대한 꺾은선그래프

3. 결론

임의의 두 개의 자연수를 정하고 재귀식 $a_{n+2} = a_{n+1} + a_n$ 을 만족하는 수열을 만들었을 때 이 수열의 첫 숫자수열이 벤포드법칙을 만족하는 지를 확인하고 이러한 수열의 첫 숫자수열의 구조를 탐색적 자료분석의 입장에서 살펴보았다. 이러한 피보나치계수열의 첫 숫자수열이 왜 벤포드법칙을 만족하는 지,

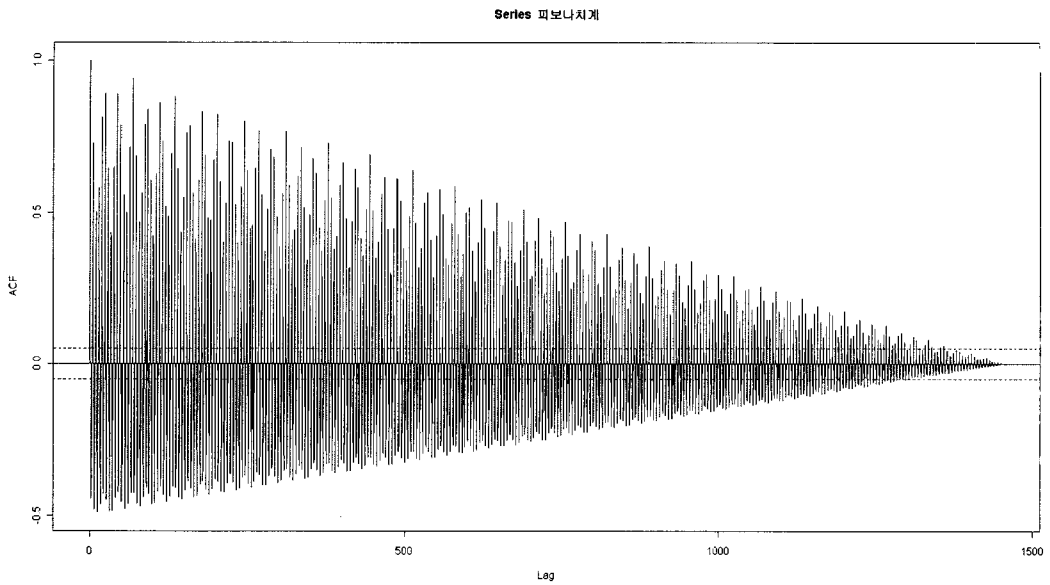


그림 2.13. $a = 1, b = 365$ 인 피보나치계수열의 첫 숫자수열에 대한 자기상관함수그림

피보나치계수열의 첫 숫자수열이 왜 특정 패턴을 갖는 지를 이론적으로 밝히는 작업이 추후 작업이 될 수 있다.

참고문헌

- Atalay, B. (2004). <다빈치의 유산-숨겨진 과학과 인간의 신비> (원 제목: *Math & The Mona Lisa*), 채은진 옮김, 도서출판 말글빛남.
- Canessa, E. (2003). Theory of analogous force on number sets, *Physica A: Statistical Mechanics and its Applications*, **328**, 44–52.
- Giles, D. E. (2007). Benford's law and naturally occurring prices in certain ebaY auctions, *Applied Economics Letters*, **14**, 157–161.

Exploratory Approach for Fibonacci Numbers and Benford's Law

Dae-Heung Jang¹

¹Division of Mathematical Sciences, Pukyong National University

(Received July 2009; accepted August 2009)

Abstract

We know that the first digits sequence of fibonacci numbers obey Benford's law. For the sequence in which the first two numbers are the arbitrary integers and the recurrence relation $a_{n+2} = a_{n+1} + a_n$ is satisfied, we can find that the first digits sequence of this sequence obey Benford's law. Also, we can find the structure of the first digits sequence of this sequence with the exploratory data analysis tools.

Keywords: Fibonacci number, Lucas number, Benford's law.

¹Professor, Division of Mathematical Sciences, Pukyong National University, 599-1 Daeyeon-dong, Nam-gu, Busan 608-737, Korea. E-mail: dhjang@pknu.ac.kr