

# ランダムウォークにおける歩行長の分布を用いた 2値系列の統計的検定

坂倉 忠和<sup>†</sup> 林 彬<sup>†</sup>

ある2値系列の性質を調べるためにしばしばベルヌーイ試行が引き合いに出されるが、その検定の1つとしてランダムウォークにおける歩行長を用いた検定がある。しかしながら、従来の歩行長を用いた検定は連テストや組合せテストといったような他の検定とは異なる結果を示すと報告されている。そこで本論文では歩行長を用いた統計的検定について検討した。その考察より平均歩行長による検定では分布の相違をとらえられないことを指摘した。また歩行長の分布に注目し、この分布を用いた頻度検定を行った。歩行長の分布に関する頻度検定では符号連テストおよび組合せテストと同様の結果が得られた。

## Statistical Test for Binary Sequences Based on the Distribution of Random Walk Length

TADAKAZU SAKAKURA<sup>†</sup> and AKIRA HAYASHI<sup>†</sup>

When we investigate properties of binary sequences, they are often compared with the sequence generated from Bernoulli trial. Average random walk length, which is the expected duration of the game in the classical ruin problems, has been used as statistics. However, results of test with average walk length are different from those of other tests, e.g. binary run length tests and combination tests. In this paper, we investigated the test with average walk length and point out some inherent problems. Also, we focused on distribution of walk length and did  $\chi^2$  test. Results of our new tests agreed with those of other tests.

### 1. はじめに

与えられた系列はその性質を調べるために、しばしば確率モデルから生成される系列と比較される。最も簡単でかつよく知られる確率モデルはベルヌーイ試行である。ベルヌーイ試行は与えられた2値系列が2値乱数と見なせるかどうかを判断するために用いられている。

2値系列とベルヌーイ試行の比較の例として Kozak らの研究<sup>1)</sup>があげられる。Kozak らはロジスティック写像

$$\tau_b(x) = bx(1-x), x \in [0, 1], \quad (1)$$

に対してある  $x_0$  を初期値とする系列  $\{x_n = \tau_b(x_{n-1})\}$  の性質を調べた。ただし、 $0 < b \leq 4$  である。このために Kozak らはしきい値関数

$$\sigma_c(x) = \begin{cases} 1, & x \geq c \\ 0, & x < c \end{cases}, \quad (2)$$

によって変換された2値系列とベルヌーイ試行による2値系列をランダムウォークにおける歩行長という統計量を用いて比較した。以後、ベルヌーイ試行から得られる2値系列をベルヌーイ系列と呼ぶことにする。この比較は2値系列とベルヌーイ系列の隔たりが著しく大きいならば、 $\tau_b(\cdot)$  は真にカオス的<sup>2)</sup>でないという考え方に基づく。Kozak らが文献1)で用いたしきい値  $c$  は、 $Pr\{\sigma_c(\tau_b(x)) = 0\} = Pr\{\sigma_c(\tau_b(x)) = 1\} = 0.5$  となる点と、区間  $[b^2(1-b/4)/b, b/4]$  における中点である。その結果、 $b = 4.0$  のときのみ  $\tau_b(\cdot)$  は真にカオス的であるとした。

一方、香田らは<sup>3)</sup>ランダムウォークにおける歩行長を用いて実数値系列の乱雑さを2値系列の乱雑さによって測る方法を提案した。すなわち種々のしきい値により2値系列を生成し、それらがすべてベルヌーイ系列に近いと見なされたとき元の実数値系列を実数値型の擬似乱数と見なす方法である。香田らは実数値型

<sup>†</sup> 金沢工業大学  
Kanazawa Institute of Technology

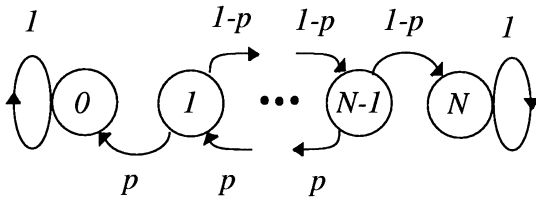


図 1 ランダムウォークの状態推移図

Fig. 1 State transition diagram for random walk.

の擬似乱数に近い系列をロジスティック写像から生成することはできないが、種々のパラメータ  $b$  に対して適当なしきい値  $c$  を選べばベルヌーイ系列に近い 2 値系列が得られると結論を下した。また、香田らは<sup>4)</sup> ロジスティック写像から得られた 2 値系列とベルヌーイ系列の比較を符号連テスト、および組合せテストによって行った。その結果、 $b = 4.0, c = 0.5$  のパラメータの組から生成される 2 値系列だけがベルヌーイ系列に近いとした。

そこで本論文では文献 4) とは異なる結論を導いたランダムウォークによる検定法<sup>3)</sup>を検討し、その改良版ともいえる歩行長の分布を用いた検定を示す。本論文は 5 章から構成されており、まず 2 章では、2 値系列とベルヌーイ系列の比較の例を示す。3 章では擬似乱数の検定として用いられたランダムウォークテストを考察し、その問題点を指摘する。さらに、4 章では歩行長の分布に関する理論式を示し、実験値と比較している。5 章でまとめと今後の課題を述べる。

## 2. ベルヌーイ試行との比較

### 2.1 平均歩行長による検定<sup>3)</sup>

まず、2 値系列  $\{B_n\}$  によるランダムウォークを定式化しよう。図 1 のように状態集合を  $\{0, 1, \dots, N-1, N\}$  とし、両端の状態 0 と  $N$  を吸収状態、残りの状態  $\{1, 2, \dots, N-1\}$  を一時的状態とする。このランダムウォークでは、時刻  $n$  において歩行者がある一時的状態  $l$  に存在するとき、時刻  $n+1$  において、状態  $l-1$  または状態  $l+1$  に移動するものとする。以後、 $\{B_n\}$  がベルヌーイ系列である場合を考えることにする。この場合には歩行者は確率  $p$  で状態  $l-1$  に移動し、確率  $1-p$  で状態  $l+1$  に移動するものと定義する。この状態  $l$  から出発して吸収されるまでの時間を  $l$  における歩行長と呼び、 $T(l; \theta)$  と表す。さらにその平均を  $l$  における平均歩行長と呼び、 $\langle T(l; \theta) \rangle$  と表すことにする。ここで、

$$\theta = \frac{1-p}{p}, \tag{3}$$

である。 $\langle T(l; \theta) \rangle$  は古典的の破産問題におけるゲームの

表 1 ランダムウォークテストの検定結果

Table 1 Results for tests using the average walk length.

$b$	$c$	$l=1$	$l=2$	$l=3$	$l=4$
3.7	0.3	1.00	1.00	1.00	1.00
3.7	0.9	1.00	1.00	1.00	1.00
3.8	0.2	1.00	1.00	1.00	1.00
3.99	0.2	0.99	0.99	1.00	1.01
4.0	0.5	0.99	0.99	0.99	1.00
$b$	$c$	$l=5$	$l=6$	$l=7$	$l=8$
3.7	0.3	1.00	1.00	1.00	1.00
3.7	0.9	1.00	1.00	1.00	1.00
3.8	0.2	1.00	1.00	1.00	1.00
3.99	0.2	1.01	1.00	0.99	1.00
4.0	0.5	1.00	0.99	0.99	1.00
$b$	$c$	$l=9$	$l=10$	$l=11$	$l=12$
3.7	0.3	1.00	1.00	1.02	1.00
3.7	0.9	1.00	1.00	1.00	1.00
3.8	0.2	1.00	1.00	1.00	1.00
3.99	0.2	1.00	0.99	1.00	1.00
4.0	0.5	1.00	1.01	1.00	1.02

期待継続時間<sup>5)</sup>、

$$\langle T(l; \theta) \rangle = \frac{N\theta^N(1-\theta^{-l})(1+\theta)}{(1-\theta^N)(1-\theta)} + \frac{l(1-\theta^N)(1+\theta)}{(1-\theta^N)(1-\theta)}, \theta \neq 1, \tag{4}$$

$$\langle T(l; 1) \rangle = l(N-l), \tag{5}$$

として知られている。

香田らは、あらかじめ

$$\hat{\theta}(b, c) = \frac{Pr\{\sigma_c(\tau_b(x)) = 1\}}{Pr\{\sigma_c(\tau_b(x)) = 0\}}, \tag{6}$$

をシミュレーションによって求めておき、実験値  $\hat{T}(l; \hat{\theta}(b, c))$  と  $\langle T(l; \theta) \rangle$  の比  $\xi(l; \theta)$  を求めた。ただし、 $\theta = \hat{\theta}(b, c)$  であり、また式 (6) の右辺の確率は経験確率である。そして  $0.95 \leq \xi(l; \theta) \leq 1.05$  であれば、ロジスティック写像  $\tau_b$  から得られる 2 値系列はベルヌーイ系列に近いとした。表 1 は、香田らが行ったランダムウォークテストを再現したシミュレーション結果であり、テストに合格したとされるパラメータの一部について  $\xi(l; \theta)$  の値を示している。ただしここでは  $p = Pr\{0\}$  としているが、文献 3) では  $p$  と  $1-p$  が逆の対称なランダムウォークを用いているので、表 1 における  $l=i$  の欄と文献 3) における  $l=13-i$  の欄が対応している。ただし、 $i=1, \dots, 12$  である。その結果、ロジスティック写像のパラメータ  $b$  に対して、適当なしきい値  $c$  を選ぶことによって、ベルヌーイ系列に近い 2 値系列を生成できるとした。

### 2.2 香田と柿本の考察<sup>4)</sup>

香田らは、式 (1), (2) から得られる 2 値系列とベルヌーイ系列の近さを測るために符号連テストと組合せテストを行った。ここで、符号連テストとは系列中

表2 香田らの符号連テストの検定結果<sup>4)</sup>

Table 2 Results of run tests using a logistic map and a threshold function.

b	c					
	0.3	0.4	0.5	0.6	0.7	0.8
3.62		317	194	190	1,043	286
3.7	153	551	802	1,847	6,088	1,663
3.8	269	257	552	1,623	3,926	1,739
3.9	380	250	922	1,754	2,816	2,020
3.99	409	316	167	333	1,108	897
3.999	8.29	28.5	4.61	3.54	27.8	121
4.0	468	164	0.71	208	839	721

に同じ値が続く長さについて検定する方法である。また、組合せテストとは与えられた系列のある長さの組に分けて、それらの組に出現する0または1の数の発生頻度を検定するものである。符号連テスト、および組合せテストでは、各カテゴリについて相対2乗誤差を求め、その総和 $\chi^2$ を求めた。また、 $\chi^2$ と $\chi_0^2$ 分布の対応する値 $\chi_0^2$ の比

$$\eta = \frac{\chi^2}{\chi_0^2}, \quad (7)$$

を定義し、 $\eta < 1$ であれば、与えられた2値系列がベルヌーイ系列に近いとした。表2は、符号連テストの検定結果を示している。また、組合せテストの検定結果は、符号連テストと同様な傾向が現れているので、ここでは省略する。これらの検定結果より、 $b = 4.0$ 、 $c = 0.5$ 以外のパラメータの組により生成される2値系列は、ベルヌーイ系列に近いとはいえないとした。

### 3. ランダムウォークテストに対する検討

#### 3.1 ランダムウォークテストの追試実験

上で見たように $(b, c) = (4.0, 0.5)$ 以外のパラメータの組からは、ベルヌーイ系列に近いとされる2値系列を生成することができない。しかし2.1節で示されているとおり、ランダムウォークテストでは符号連テスト、および組合せテストで不合格であったパラメータの組についても合格としている。もちろん、各検定がとらえることのできる性質はそれぞれ異なるが、ランダムウォークテストに不明な点があるので、それらを考察していくことにする。

ところでMersenne Twister<sup>6)</sup>は性質の良い擬似乱数を生成することで知られている。このMersenne Twisterを用いたランダムウォークテストからどのような結果が得られるかをシミュレーションによって調べた。Mersenne TwisterについてはC言語のライブラリが公開されており、表3はそのライブラリを利用して実数値系列を生成し、式(2)によって変換された2

表3 Mersenne Twisterによるランダムウォークテストの検定結果

Table 3 Results of random-walk tests with sequences generated from Mersenne Twister's.

c	l = 1	l = 2	l = 3	l = 4
0.2	1.00	1.00	1.00	1.00
0.3	0.99	1.00	1.00	1.00
0.4	0.99	1.00	1.00	1.00
0.5	0.99	1.00	1.00	1.00
0.6	0.98	1.00	0.99	1.00
0.7	1.00	1.01	1.01	1.01
0.8	1.01	1.00	1.01	1.00
0.9	1.00	1.00	1.00	1.00
c	l = 5	l = 6	l = 7	l = 8
0.2	1.00	1.00	1.00	1.00
0.3	1.01	1.01	1.00	1.00
0.4	1.00	1.00	1.00	1.01
0.5	1.00	1.00	1.00	1.00
0.6	1.00	1.00	1.00	1.00
0.7	1.00	0.99	1.00	1.00
0.8	1.00	1.00	1.00	1.00
0.9	1.00	1.00	1.00	1.00
c	l = 9	l = 10	l = 11	l = 12
0.2	1.00	1.00	1.00	1.00
0.3	1.00	1.00	1.00	1.01
0.4	0.99	1.00	1.00	1.02
0.5	1.00	0.99	1.00	1.02
0.6	1.01	1.00	1.01	0.99
0.7	1.00	1.00	1.00	1.00
0.8	1.00	1.00	1.00	1.00
0.9	1.00	1.00	1.00	1.00

値系列に対して2.1節に述べたような検定の結果を示している。表3よりMersenne Twisterによって生成された2値系列を用いた場合、 $0.98 \leq \xi(l; \theta) \leq 1.02$ の範囲内に存在することが分かる。

一方、表1で示したパラメータから生成された2値系列の場合、明らかにベルヌーイ系列と見せせないにもかかわらず $0.99 \leq \xi(l; \theta) \leq 1.02$ の範囲内に存在する。このことより、平均歩行長を用いた検定には問題があることが示唆される。

次に、これまで $N = 13$ のランダムウォークを行ってきたが、このランダムウォークの状態数が十分であるかどうか不明である。そこで、これを明らかにするために $N = 25$ のランダムウォークテストを行った。表4は、 $b = 3.7$ 、 $c = 0.3$ としたときの $N = 25$ のランダムウォークテストの結果を示している。この表より、すべての初期状態 $l$ について $0.98 < \xi(l; \theta) < 1.02$ となっていることが分かる。さらに非常に極端な例として、 $N = 101$ 、 $N = 1001$ という状態数で同様のランダムウォークを行ったが、これらについてもすべての初期状態 $l$ について $0.98 < \xi(l; \theta) < 1.02$ となった。これらのシミュレーション結果より、 $N = 13$ で

表 4  $N = 25$  のランダムウォークテストの検定結果  
( $b, c$ ) = (3.7, 0.3)

Table 4 Results of random walk tests with  $N = 25$ , using a pair of parameters ( $b, c$ ) = (3.7, 0.3).

$l$	1	2	3	4	5	6
$\xi(l; \theta)$	1.00	1.00	1.00	1.00	1.00	1.00
$l$	7	8	9	10	11	12
$\xi(l; \theta)$	1.00	1.00	1.00	1.00	1.00	1.00
$l$	13	14	15	16	17	18
$\xi(l; \theta)$	1.00	1.00	1.00	1.00	1.00	1.00
$l$	19	20	21	22	23	24
$\xi(l; \theta)$	1.00	1.00	1.00	1.00	1.00	1.00

Relative Frequency

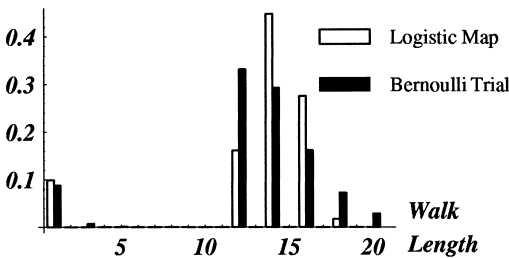


図 2  $N = 13, l = 1$  における歩行長の頻度分布  
( $b, c$ ) = (3.7, 0.3)

Fig. 2 Frequency distribution of the walk length in the case of  $N = 13, l = 12$  and ( $b, c$ ) = (3.7, 0.3).

Relative Frequency

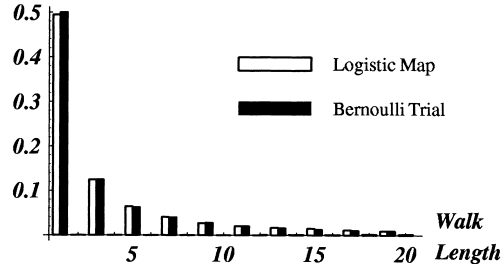


図 3  $N = 13, l = 1$  における歩行長の頻度分布  
( $b, c$ ) = (4.0, 0.5)

Fig. 3 Frequency distribution of the walk length in the case of  $N = 13, l = 12$  and ( $b, c$ ) = (4.0, 0.5).

十分であることが分かる .

3.2 平均歩行長による検定の問題点

前節では、ランダムウォークテストでの不明な点、すなわち検定における棄却域と状態数に検討を加え、検定結果に重大な影響を及ぼすことがないことを明らかにした . そこで次に、ランダムウォークで得られる歩行長がどのような分布になるかを調べた . 図 2、図 3 は、式 (1)、(2) から得られる 2 値系列を用いて、ランダムウォークを行った場合の歩行長の分布を表し

ている . ここではパラメータの組 ( $b, c$ ) をそれぞれ、( $b, c$ ) = (3.7, 0.3)、( $b, c$ ) = (4.0, 0.5) としている . また、比較のために次章で説明するベルヌーイ系列による歩行長の分布も示している . 図 3 より、 $b = 4.0, c = 0.5$  における歩行長の分布はおおよそベルヌーイのそれと一致していることが分かる . それに対して  $b = 3.7, c = 0.3$  の場合は一致していない . これは平均歩行長  $\langle T(l; \theta) \rangle$  を用いているためにこのような現象が起こったと考えられる . このことから歩行長の分布に着目すればよいと考えられる .

4. 歩行長の分布に関する頻度検定

4.1 歩行長の分布

前章で示したように、従来のランダムウォークテストでは、 $l$  における平均歩行長によって検定を行っているために分布の相違をとらえることができない . そこで、歩行長の分布を用いて頻度検定を行うことにする . ある一時的状態  $l$  から出発して  $n$  回目の歩行で吸収状態 0 に吸収される確率は

$$u_{l,n}(p) = \frac{2^n}{N} p^{\frac{n+l}{2}} (1-p)^{\frac{n-l}{2}} \times \sum_{\nu=1}^{N-1} \cos^{n-1} \frac{\nu\pi}{N} \cdot \sin \frac{l\nu\pi}{N} \cdot \sin \frac{\nu\pi}{N}, \quad (8)$$

となることが知られている<sup>5)</sup> . したがって、 $n$  回目の歩行で吸収状態 0 または  $N$  に吸収される確率は

$$w_B(p; l, n) = u_{l,n}(p) + u_{N-l,n}(1-p), \quad (9)$$

である .

4.2 シミュレーション結果

歩行長の分布に関する頻度検定を行うために、次のようなランダムウォークに関するシミュレーションを行った .  $N = 13$  として、すべての一時的状態を初期状態としてシミュレーションを行い、25,000 回吸収されるまで繰返しを行った . また、実験値を  $\chi^2$  検定によって評価するために次の手順で各カテゴリに分類し、統計量  $\chi^2_p$  を求めた .

- (1) 式 (9) により、 $n = 1, \dots, M$  について期待頻度  $L \cdot w_B(p; l, n)$  を求める . ただし、 $L$  はシミュレーションにおける繰返し回数 (本論文では、 $L = 25,000$ ) であり、 $M$  は

$$\sum_{n=1}^M w_B(p; l, n) > 0.99, \quad (10)$$

を満たす最小の整数である .

- (2) すべてのカテゴリに対する期待頻度が、10 以上になるように 10 未満のカテゴリを 1 つのカテゴリにまとめる . その操作の後でさえも、10

未満となる場合には 10 以上のカテゴリの中で最も小さいカテゴリを足し、カテゴリの数を 1 つ減らす。ここで、得られたカテゴリに対する期待頻度を  $\bar{f}_i$  とする。ただし、 $i = 1, \dots, M'$  であり、カテゴリ  $M'$  は期待頻度が 10 未満であるようなカテゴリをまとめて作成したカテゴリとする。

(3) 手順 (1) において考慮していない期待頻度  $0.01 \cdot L$  をカテゴリ  $M'$  に加え、新たにカテゴリ  $M'$  の期待頻度とする。

(4) 最終的に、 $M' = \nu + 1$  個のカテゴリに分類できたとする、

$$\chi^2_\nu = \sum_{i=1}^{\nu+1} \frac{(f_i - \bar{f}_i)^2}{\bar{f}_i}, \quad (11)$$

によって評価値  $\chi^2_\nu$  を求める。ただし、 $f_i$  はカテゴリ  $i$  に対応するシミュレーションで得られた頻度である。

$\chi^2_\nu$  と自由度  $\nu$  の  $\chi^2$  分布の 95% 点  $\chi^2_{0.05}$  との比

$$\eta(l; \theta) = \frac{\chi^2_\nu}{\chi^2_{0.05}}, \quad (12)$$

を定義して、この比  $\eta(l; \theta)$  によって評価を行う。 $\eta(l; \theta) < 1$  であれば、得られた 2 値系列がベルヌーイ系列に近いとする。

表 5 は、各初期状態  $l$  についての  $\eta(l; \theta)$  を表している。この表より、明らかに  $b = 4.0, c = 0.5$  以外のパラメータの組から生成される 2 値系列とベルヌーイ系列とは隔たりがあることが分かる。また、 $b = 4.0, c = 0.5$  については、すべての初期状態  $l$  で  $\eta(l; \theta) < 1$  となっている。この検定結果より、 $b = 4.0, c = 0.5$  のパラメータの組から生成される 2 値系列はベルヌーイ系列に十分近いといえる。一方、表 6 では符号連続テストに不合格となった  $b = 3.9$  についての結果を示している。明らかに、歩行長の分布に関する頻度検定においても不合格となっていることが分かる。また、 $b = 3.9$  以外のパラメータの組についても同様の結果が得られたが、それらについては省略することにする。以上に見るように、平均歩行長を用いた検定で区別できなかったパラメータの組を歩行長の分布を用いることによって識別することができた。さらに、これらの結果は符号連続テストや組合せテストの結果に一致する。すなわち、 $(b, c) = (4.0, 0.5)$  だけがベルヌーイ系列と見なせる。

また、比較のために従来の擬似乱数生成器を用いた場合の歩行長の分布も調べた。従来の擬似乱数生成器として、線形合同法<sup>7)</sup>

表 5 歩行長の分布に関する頻度検定の結果

Table 5 Results of frequency tests with the walk length.

<i>b</i>	<i>c</i>	<i>l</i> = 1	<i>l</i> = 2	<i>l</i> = 3	<i>l</i> = 4
3.7	0.3	550	712	758	880
3.7	0.9	1,291	807	592	255
3.8	0.2	258	306	367	355
3.99	0.2	22.5	29.3	32.4	29.5
4.0	0.5	0.71	0.79	0.88	0.84
<i>b</i>	<i>c</i>	<i>l</i> = 5	<i>l</i> = 6	<i>l</i> = 7	<i>l</i> = 8
3.7	0.3	986	1,024	1,016	1,037
3.7	0.9	1,192	950	940	736
3.8	0.2	391	370	340	273
3.99	0.2	27.9	28.6	32.3	31.9
4.0	0.5	0.80	0.88	0.90	0.77
<i>b</i>	<i>c</i>	<i>l</i> = 9	<i>l</i> = 10	<i>l</i> = 11	<i>l</i> = 12
3.7	0.3	907	529	326	119
3.7	0.9	1,291	807	592	255
3.8	0.2	213	121	62.6	9.30
3.99	0.2	40.8	41.1	42.5	18.4
4.0	0.5	0.98	0.83	0.85	0.70

表 6  $b = 3.9$  における歩行長の分布に関する頻度検定の結果

Table 6 Results of frequency tests with the walk length in the case of  $b = 3.9$ .

<i>c</i>	<i>l</i> = 1	<i>l</i> = 2	<i>l</i> = 3	<i>l</i> = 4
0.2	46.9	59.4	37.1	42.3
0.3	317	500	446	467
0.4	188	248	296	236
0.5	77.4	83.3	91.2	76.4
0.6	272	146	90.4	73.5
0.7	609	354	232	178
0.8	1,109	991	791	583
0.9	2,064	1,348	1,209	1,183
<i>c</i>	<i>l</i> = 5	<i>l</i> = 6	<i>l</i> = 7	<i>l</i> = 8
0.2	41.9	41.7	53.1	76.8
0.3	595	644	633	704
0.4	232	245	266	339
0.5	65.4	51.8	39.7	28.5
0.6	67.1	65.5	65.3	61.8
0.7	155	144	147	165
0.8	475	414	392	377
0.9	1,045	963	880	727
<i>c</i>	<i>l</i> = 9	<i>l</i> = 10	<i>l</i> = 11	<i>l</i> = 12
0.2	139	84.3	162	430
0.3	834	748	700	1,244
0.4	351	401	449	710
0.5	23.6	15.1	12.3	15.6
0.6	60.2	52.3	42.3	29.8
0.7	201	250	301	169
0.8	412	510	788	407
0.9	633	595	707	431

$$X_{n+1} = 314, 159, 269X_n \text{ mod } 2^{31} - 1, \quad (13)$$

と Mersenne Twister を用いてシミュレーションを行った。線形合同法に関しては式 (13) によって整数の系列を生成し、法で割ることにより区間  $[0, 1)$  上の実数値系列を生成した。表 7、表 8 はそれぞれ、線形合

表 7 線形合同法による歩行長の分布に関する頻度検定の結果  
Table 7 Results of frequency tests with the walk length using a linear congruential method.

$c$	$l = 1$	$l = 2$	$l = 3$	$l = 4$
0.2	0.77	0.57	0.50	0.71
0.3	0.74	0.96	0.68	0.66
0.4	0.70	0.55	0.99	0.80
0.5	0.89	0.78	0.91	0.84
0.6	0.73	0.72	0.85	0.71
0.7	0.72	0.68	0.79	0.48
0.8	0.46	0.51	0.83	0.49
0.9	1.38	0.66	0.42	1.07
$c$	$l = 5$	$l = 6$	$l = 7$	$l = 8$
0.2	0.52	0.92	0.76	0.80
0.3	0.70	0.45	0.81	0.72
0.4	0.55	1.09	0.94	0.98
0.5	1.01	0.85	0.76	0.93
0.6	0.67	0.73	0.91	0.72
0.7	0.46	1.08	0.79	0.52
0.8	0.75	0.53	0.52	0.64
0.9	0.26	0.43	0.24	0.52
$c$	$l = 9$	$l = 10$	$l = 11$	$l = 12$
0.2	1.33	1.17	0.70	0.92
0.3	0.81	0.69	0.55	0.55
0.4	0.74	0.80	0.84	0.90
0.5	0.81	0.84	0.72	0.57
0.6	0.63	0.62	0.77	0.67
0.7	0.55	0.67	0.54	0.56
0.8	0.17	0.57	0.72	0.34
0.9	0.27	0.08	0.48	1.19

表 8 Mersenne Twister 法による歩行長の分布に関する頻度検定の結果

Table 8 Results of frequency tests with the walk length using Mersenne Twister.

$c$	$l = 1$	$l = 2$	$l = 3$	$l = 4$
0.2	0.45	0.42	0.59	0.51
0.3	0.59	0.52	0.49	0.63
0.4	0.65	0.56	0.61	0.99
0.5	0.69	0.85	0.59	0.71
0.6	0.79	1.05	0.68	0.71
0.7	0.62	0.56	0.95	0.47
0.8	0.60	0.75	0.67	0.67
0.9	0.47	0.42	1.08	0.32
$c$	$l = 5$	$l = 6$	$l = 7$	$l = 8$
0.2	0.53	0.41	0.27	0.81
0.3	0.92	0.82	0.57	0.80
0.4	0.74	0.80	0.65	0.69
0.5	0.73	0.90	0.94	0.86
0.6	0.76	0.82	0.92	0.62
0.7	0.89	0.77	0.85	0.57
0.8	0.64	0.63	0.78	0.19
0.9	0.28	1.21	0.28	0.77
$c$	$l = 9$	$l = 10$	$l = 11$	$l = 12$
0.2	0.59	0.25	0.91	1.09
0.3	0.79	0.52	0.73	0.54
0.4	0.63	0.81	0.80	0.56
0.5	0.82	0.70	0.79	0.88
0.6	0.87	0.81	0.70	0.57
0.7	0.70	0.82	0.78	0.26
0.8	0.40	0.71	0.26	1.20
0.9	0.50	0.76	0.40	0.82

同法, Mersenne Twister による歩行長の分布に関する頻度検定の結果である。これらの表より, 両者とも  $\eta(l; \theta) > 1$  となるパラメータの組が存在していることが分かる。これは検定のときに  $\chi^2$  分布の 95% 点を用いているので, 5% 程度は範囲外が現れるのは妥当であると考えられる。

ところで歩行長の分布を用いた検定が他の検定法に比べて優れているかどうかは興味深い問題である。しかしながら, 何をもちて優劣を決めるかという問題やそれらをどのように定量化するかという問題は非常に難しい。一方, Knuth<sup>7)</sup> が指摘しているように与えられた 2 値系列がいくつかの検定法に対して合格したとしても別の検定法において不合格にならない保証はない。このことはその結果だけから評価できるような最良の統計的検定法が知られていないことを意味する。したがって, 確からしさを高めるためにより多くの検定法を用いて検討する必要がある, 歩行長の分布を用いた検定もその 1 つであると考えられる。

### 5. む す び

本論文ではある 2 値系列とベルヌーイ系列の比較の際に用いられる統計的検定の 1 つであるランダム

ウォークにおける歩行長を用いた検定法に注目した。平均歩行長を用いた場合は分布が異なる場合でも検定に合格する場合がある。このことに鑑み, 歩行長の分布を用いた検定を示した。数値例として, ロジスティック写像としきい値から生成される 2 値系列を取り上げ, 統計的検定を行った。その結果, 平均歩行長を用いる検定ではベルヌーイ系列ではないものを必ずしも棄却できなかったが, 歩行長の分布を用いる検定では棄却しえた。さらにそれらの結果は符号連テストや組合せテストの結果に一致する。他の検定法とともに検定の確からしさをより高めるためには, 平均歩行長ではなく歩行長の分布が用いられるべきである。

ところで応用の立場では対象となる 2 値系列に対して種々の検定を行い, それらの結果を総合的に評価してその 2 値系列が利用可能であるか, すなわち 2 値乱数と見なせるかどうかを判断する。そこでは用いられた検定が何らかの尺度によって同等に厳しくあることが望まれる。その 1 つとして検定での有意水準がその尺度にあたるがそれだけでは十分であるとはいえない。たとえば同じ有意水準を用いたとしても, ランダムウォークの状態数  $N$  が異なればその判別する能力

は異なることが容易に推測される。一方、状態数  $N$  が増えた場合には検定に必要とされる系列の長さは必然的に長くなる。したがって検定の能力と検定の効率を考えてランダムウォークの状態数を設定する必要があると考えられる。これらについては今後の課題とする。

### 参 考 文 献

- 1) Kozak, J.J., Musho, M.K. and Hatlee, M.D.: Chaos, Periodic Chaos, and the Random-Walk Problem, *Phys. Rev. Lett.*, Vol.49, No.25, pp.1801-1804 (1982).
- 2) Grossman, S. and Thomae, S.: Invariant distributions and stationary correlation functions of one dimensional discrete processes, *Z. Naturforsch.*, Vol.32a, pp.1353-1363 (1977).
- 3) 香田 徹, 緒方栄次: ベルヌイ試行とカオス, 電子情報通信学会論文誌 A, Vol.J68-A, No.2, pp.146-152 (1985).
- 4) 香田 徹, 柿本厚志: 擬似乱数とカオス, 情報処理学会論文誌, Vol.27, No.3, pp.289-296 (1986).
- 5) Feller, W. (著), 河田(訳): 確率論とその応用 I(下), pp.431-466, 紀伊国屋書店 (1974).
- 6) Matsumoto, M. and Nishimura, T.: Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator, *ACM Trans. on Modeling and Computer Simulation*, Vol.8, No.1, pp.3-30 (1998).
- 7) Knuth, D.E.(著), 渋谷(訳): 準数値算法/乱

数, pp.40-74, サイエンス社 (1981).

(平成 13 年 7 月 18 日受付)

(平成 14 年 3 月 14 日採録)



坂倉 忠和(正会員)

昭和 48 年生。平成 8 年近畿大学理工学部経営工学科卒業。平成 10 年金沢工業大学大学院工学研究科修士課程修了。平成 13 年同大学院工学研究科博士課程単位取得後退学。

現在、金沢工業大学助手。離散力学系におけるカオスに興味を持つ。電子情報通信学会会員



林 彬(正会員)

昭和 16 年生。昭和 39 年金沢大学工学部電気工学科卒業。昭和 48 年ミネソタ大学大学院修士課程修了。昭和 51 年ハワイ大学大学院博士課程修了。昭和 39 年東芝中央研究所

入社。昭和 45 年金沢工業大学講師, 昭和 52 年同大学情報工学科教授。平成 12 年電子工学科教授。Ph.D. 音響工学, 通信理論, 情報理論, 暗号理論, ニューロ計算に興味を持つ。電子情報通信学会, 情報理論とその応用学会, IEEE 各会員。