標準的なゲームプレイヤにとって自然に見える疑似乱数列の生成法

野村 久光　　テンシリリックン シラ　　池田 心
北陸先端科学技術大学院大学
{s1210043, temsiririrkkul, kokolo} @jaist.ac.jp

疑似乱数生成の研究は古くからあり，偏りのなさや周期の長さ，生成速度などの改良が進められてきた．Mersenne twister など最近の手法は数学的な意味で真の乱数に十分近いと言え，確率的最適化やモンテカルロ法などさまざまに応用されている．テレビゲームでも疑似乱数が必要になることは多く，例えばすごろくではサイコロの目をコンピュータが決めなければならない．このとき，出た目およびその系列によっては，プレイヤはそのサイコロの目が自分に都合の悪いようにコンピュータに操作されていると感じる．本稿では，数学的な意味で良い乱数と，標準的なゲームプレイヤにとっての自然な乱数は異なるという仮定をおき，どのような特徴を持たせれば自然に“見える”乱数が作れるのかを考察，実装する．被験者実験の結果，標準的な乱数よりも自然に見え，またすごろくで使ったときの不満が小さい乱数列を生成できていることを確認した．

Generation of “Natural” Pseudorandom Numbers from the Standard Player’s View

Hisamitsu Nomura,　Sila Temsiririrkkul,　Kokolo Ikeda
Japan Advanced Institute of Science and Technology

The generation of pseudorandom numbers is a well-studied problem, and many results have been obtained in terms of the random distribution quality, the length of the generation period or the speed of the generator. Pseudorandom numbers generated by algorithms like Mersenne twister are close to real random numbers in a mathematical meaning, and they are widely used in many applications such as optimization or Monte-Carlo methods. Pseudorandom numbers are also used in video games, for example, to simulate a virtual dice, but it is possible that, for some sequences of numbers, the player will believe that the dice is biased against him. In this article, we investigate the possibility that randomness in a mathematical meaning is not necessarily the best way to produce a sequence of numbers that look random to the average player. We show with subject experiments that some specially crafted sequences of numbers will look more random to the players than a mathematical pseudorandom sequence, and that such sequences lead to less complaints about the virtual dice in a Japanese game of sugoroku.

1. はじめに

コンピュータの中で特定の確率分布に従って数を取り出すことを疑似乱数生成と呼び，シミュレーション・確率的最適化法・モンテカルロ法・強化学習法など情報学の多くの分野で必須の部品となっている．線型合同法などの古典的な乱数生成法は多くの欠点を持つが[1]，近年，Mersenne twister(MT)法[2]など多くの優れた生成法が開発され，上記のような用途では，「偏りのなさ」「周期の長さ」「生成の速度」などの要求を十分満たすものが簡単に利用できるようになっている．

コンピュータゲームでも，麻雀牌を混ぜる・トランプを配る・サイコロを振るなどの作業はすべてコンピュータが肩代わりするため，多くのゲームで疑似乱数が用いられる．これらをコンピュータプレイヤ側に有利になるように操作する（いわゆる“ズル”をする）ことは簡単であり，実際多くのソフトで行われてきた．『桃太郎電鉄』等パーティ要素の強いすごろくゲームや『ぎゅわんぶらあ自己中心派』など個性の強いキャラクタの出る麻雀ゲームなどではこれらは楽しさの演出のためにほぼ明示的に行われているため，プレイヤもそれを許容できる場合が多い．

一方，（特に，強い思考ルーチンを開発者が作れないために）暗黙的に行われる乱数操作は好まれず，そのようなゲームは評価が下がる傾向にある．ところが，思考ルーチンがハンデを必要とし

ないくらいに強い場合でも，本当の乱数に近い『数学的な意味で優れた』乱数（例えばMT法）を使うことがプレイヤの満足に最善とは言えない場合もある．そうではなく，本当の乱数に近いと『プレイヤが思える』乱数を使う必要があるはずだ，というのが本研究の着眼点である．どのようなゲームやプレイヤの層を対象にするかによっても事情は異なるが，本研究では，すごろくゲームの平均的なプレイヤを対象に，自然と思えるサイコロの出目の生成法を考案することを目的とする．

2. 関連事項・関連研究

『カルドセプト』はカード対戦要素とすごろく要素を持つ有名なテレビゲームのシリーズであるが，このゲームについてはしばしばユーザ側から「サイコロの目がユーザに不利なように操作されている」という意見・報告が疑惑ではなく確信として出されている（例えば[3]ではレビュアの1/3以上がそのように述べている）．面白いことに，これに対してメーカー側も公式見解として「サイコロの目は操作していない，普通の疑似乱数を使っている」と10年以上前から述べているのだが[4]，ユーザは納得していない．

・ 一部の人がしつこく批判しているだけ
・ メーカー側が嘘の説明をしている
・ 予期していないバグによるもの

という解釈もありうるが，我々は正しい疑似乱数が数学が得意でない人間にとっては乱数に見えないという仮説も有力であると考える．

対象が自然であるかの判断は，評価者の対象への理解に大きく依存する．例えばアマチュア用に手加減をする将棋プログラムの研究では，２つの棋譜に対する人間らしさの評価がプロ棋士とアマチュアでは逆転しうることが確認された[5]．著者の仲道氏は「アマチュアのためのプログラムなら，アマチュアにとって自然に見えるほうが望ましい」と述べており，我々もその考え方に賛同する．つまり同様に，数学者にとって自然な乱数よりも，一般的なプレイヤにとって自然に見える乱数のほうがゲームソフトでは必要になると考える．

実際，人間はしばしば確率的な事象に対して誤った判断を下す．これは広くは認知バイアスと呼ばれ[6]，個々にはギャンブラーの誤謬，クラスターの錯覚などと分類されている．自然と思ってもらうためには，どのような勘違いをしやすいかということを考慮にいれる必要がある．認知バイアスの例については次章で述べる．

3. 認知バイアス

認知バイアスとは，「ある対象を評価する際に，自身の利害や希望に沿った方向に歪められたり，対象の目立ちやすい特徴に引きずられたりして，ほかの特徴についての評価が歪められる現象」を指す[7]．

人間の情報処理能力には限界がある．そのため，ある程度の誤差を容認し，素早い知覚・判断を行い，用いる情報量を節約している．これをヒューリスティクス（無意識下の簡便な思考法）という．これは，判断にいたる時間は短いが，必ずしも正しいとはいえない判断結果を示す．そこで，認知バイアスが生じる[8]．

その性質上，認知バイアスは非常に多種多様である．本章では，本研究とかかわりのある可能性のあるものについていくつか紹介する．

【確証バイアス】

人は現在持っている信念，理論，仮説に一致する情報を求め，潜在的に反証となる証拠の収集を避ける基本傾向を持つ．数多くの被験者実験によって，被験者は関連のある証拠を積極的に探すように求められた時，自分の仮説を論駁するのではなく確証するのに都合の良い方略を採用するため，一般的規則を発見するのに失敗する[9]．

ゲームプレイ中の人間は，一たびサイコロが操作されていると仮説を立ててしまうと，それ以降のプレイでその仮説に都合の良い部分ばかり目がいき，ますます確信を深めてしまう可能性がある．

【バンドワゴン効果】

人間は，他人の行動につられたり，他人の意見に影響される傾向がある．これを心理学では，「同調」と呼ぶ．特に，日本人は同調意識が強く，「これが流行っている」とテレビや雑誌に紹介されただけで商品を買ってしまったりする．自分の好みで選ぶのでなく，周囲と同じ趣向であることに安心感を得るのである[10]．場合によっては，ゲームのレビューなどでも同様の現象が生じている可能性はある．

【小数の法則】

統計学の大数の法則(law of large numbers)をもじってつけられたものである．「大数の法則」は母集合から抽出される標本の大きさが大きくなるにつれ，その標本平均は母集団平均に近づくと

いう統計法則であるのに対し，「小数の法則」は母集団からのランダムに抽出したどんなに少ない標本でも，母集団の性質を表している“はずだ”という人々の思いこみのことである．

例えば長い乱数列中の数の出現頻度は全体では一様であるが，どの部分を抽出しても一様であるべきだと思ってしまいがちである．実際には，狭い領域では，ある程度の偏りは見られる．

【ギャンブラーの誤謬】

例えば，表と裏の出る確率が50%ずつのコインを振って，4回続けて表が出ると，「次はそろそろ裏が出るだろう」と考えることは，一般人にはありがちである．4回も表の続く確率は，1/2×1/2×1/2×1/2＝1/16であり，次も表が出る確率は，1/32で100回のうち3回程度の確率である．しかし，何回続けて表が出ようとも，次に表の出る確率は50%である．表が続くと次に表の出る確率を本来の50%から過小評価してしまう．このような考えを「ギャンブラーの誤謬」と呼ぶ．

【クラスターの錯覚】

クラスターの錯覚とは，ランダムに起こるべきある出来事がまとまって起こったとき，それをランダムでないと錯覚してしまうことを指す．たとえば，コイン投げで表が続けて4回出たら，多くの人は驚くであろう．しかし，20回連続して投げた場合，表が続けて4回出る確率は50%である．[11]

【制御幻想】

実際には制御不可能であるものを，自分が制御できると思い込んでしまい，成功確率を高く錯覚して行動してしまうことを「制御幻想」と呼ぶ．テレビゲーム上のサイコロと，実際にサイコロを振るすごろくでは，異なる実験結果が出る可能性は高い．

【ランダム系列の誤認知】

人間が考えるランダムさは理論的なランダムさより入れ替わりが激しくなっている傾向がある．ポール・バカンの実験（1960年）では，70名の大学生に，「偏りのないコインを偏りのない方法で300回投げた時に表と裏がどのような配列で出るかを考えて欲しい」と指示した．この系列には表と裏のあいだで平均して150回の交代が起こるはずだ．ところが，学生の90%は，交代が頻繁すぎる配列を作った（平均175回）．つまり，二つの結果の間の交代が，ランダムな配列で実際に起こるよりも頻繁に起こるものと期待するということである[12]．

【効用・プロスペクト理論】

確率を伴う事象における意思決定には，意思決定者の嗜好や認知バイアスが強く反映する．期待効用仮説は主に，将来の利得に対する喜びを利得の量に対して非線形な効用関数で表し，その利得が得られる確率で期待値を取ることで，人の行動をモデル化したものである．

プロスペクト理論はさらに，期待効用仮説では説明できない人間の行動をうまく説明するために，人が確率に対しても非線形な偏りを持った認知をすることを見出し，モデル化するものである．例えば多くの人にとって 1%で起きる事象と 0.1%で起きる事象はどちらも「非常にまれ」であって，10倍も違うように実感することは難しい．

本項目は認知バイアスそのものではないが，認知バイアスを明示的に扱う理論として稀な例の扱いが重要になる本研究とも関わりが深い．

4. 研究のアプローチ

本研究の全体的な構想としては大きく4つの段階を想定しており，現在は2つめまで完了している．3つめは比較的小規模な将来の目標，4つめは将来的なアイデアである．

1) 認知バイアスの分類に基づき，主にアンケート調査を用いて，乱数・確率について多くの人に共通する誤解・偏り・思い込みの傾向を調査する．
2) その認識の偏りをなんらかの意味で再現するような乱数列の生成法を考案・実装，「自然さ」を評価する．中には，誤解はあっても不自然さの原因にはならないものもあることに注意する．
3) 適用対象をすごろくに限った場合，止まりたくないマスに止まるような頻度を，乱数列の生成とは別に調整する手法を考案・実装し，実際のすごろくで用いて評価する．具体的には，10回中3回止まるような確率の場合は，実際に10回中2～4回止まらせる（1回以下や5回以上にしない）などの方法をとる．
4) 3)までは，多くのユーザに共通する特徴を利用するが，実際には認識の偏りや自然さの判断基準は人により大きく異なる．特定のゲームの熟練者や，数学的素養のあるプレイヤから見れば，

3)のようにある意味「数学ができない人」向けの調整はかえって不自然になる可能性が高い．そこで，ゲーム実施中に，各ユーザの傾向をオンラインで分析して，その人に特化した乱数列・サイコロ制御法を適用する．

5. アンケート調査

前章段階1で示した目標のため，確率の問題，特にすごろくの問題，および乱数列に関するアンケート調査を行った．アンケート調査は2回に分けて行われ，前者はゲームが好きな情報系大学院生12人，後者は同16人を対象としている．うち10人は共通しており，両者を合算することはしない．前者は無償の予備実験，後者は評価を目的に含むため有償の被験者実験として行った．本章では，得られた結果の中で興味深いものをいくつか紹介するが，必ずしも実験が行われた順番や設問の順番通りでないことに注意されたい．

5. 1.　落とし穴のあるすごろく

本稿ではすごろくを対象ゲームとして用いる．2章で挙げた『カルドセプト』や，海外で有名な『モノポリー』などはすごろくの類型あるいは発展形と捉えることができる．これらすごろくゲームには大抵の場合，プレイヤにとって好ましい効果のあるマス，好ましくない効果のあるマスがあり，プレイヤはサイコロの目に一喜一憂する．ところが，なじみ深いゲームである一方で「あるマスに止まる確率はいくらか」という数学的知識を持たずにプレイすることは珍しくない．

このことを確かめるため，一回目実験では，図１のような単純なすごろくを被験者に見せ，９マス目にある落とし穴にはまる確率を予測してもらった．サイコロが公正なものである場合正しい確率は約28%である．しかし，12人中10人は20%以下の確率を答えた．2割以下だと想定している落とし穴に3割近い割合で落ちたとしたら，「サイコロが操作されている」と思ったとしても仕方ないであろう．この結果は2章の仮説の裏付けとなっている．

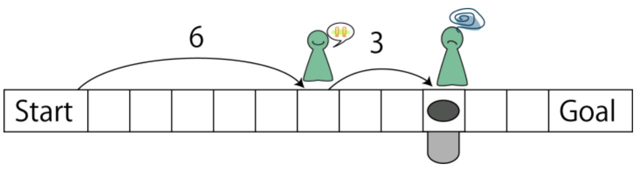

図１　落とし穴のあるすごろく

5. 2.　事象の起こりやすさ

すごろくに限らず，確率の問題というのは苦手な人が多い．二回目実験では，どのような問題を特に苦手にしているかを確認するため，いくつかの比較的想像しやすい日常的な事象について，その起こりやすさを(a)2%以下，(b)2〜8%，(c)8〜25%，(d)25〜75%，(e)75〜92%，(f)92%〜98%，(g)98%以上，の7段階で予想してもらった．１問あたり20秒ほどしか時間を与えず，直感で答えることを促している．項目は以下の通りである．正解を付す．

Q1.　コイントス10回で，表が8回出る（b）
Q2.　コイントス40回で，表が20回出る（c）
Q3.　10面サイが6面サイより大きい目を出す（d）
Q4.　40人のクラスで誕生日が同じ2人がいる（e）
Q5.　60人のクラスで誕生日が同じ2人がいる（g）
Q6.　確率30%の罠に10回中6回以上はまる（b）
Q7.　確率30%の罠に11回中1回以下だけはまる（c）
Q8.　確率1%の罠に100回中3回はまる（b）
Q9.　図1で6マス先に罠があった場合（d）

平均的に成績が良かったのはQ1,Q2,Q3,Q9で，半分以上の人が正解，間違っても一段階であった．特にQ9の結果は5.1節の結果と比べると意外であった．6マス先に落ちる確率は約36%なので9マス先（約28%）よりは高いが，16人中12人が25%以上と答えていることから，近い罠は比較的高く見積もる（遠い罠を過小評価する）ようなバイアスがあるのかもしれない．

Q4,Q5は90%，99%程度が実際の値であるが，Q4でも，またQ5ですら半数が25%以下（a〜c）と答えている．Q5で92%以上(f，g)と答えたのは16人中3人に過ぎない．被験者の多くはゲーム好きのはずだが，誕生日のパラドックスとして知られるこの有名な問題を答えられないということは，そういう層でも必ずしも確率に関して十分な素養を持たないということを示唆している．

Q6,Q8 は過大評価が多く（Q6では9人，Q8では12人がcまたはd），Q7は過小評価が多い（12人がbまたはa）．Q6は，10回中6回というのが“中間的な”事象，Q7は，10回中1回というのが“目立つ，極端な”事象に見えている可能性がある．これは次節の乱数生成にも関わることで，目立つ特徴は不自然さに関係する．Q8については，1%という小さい確率を，心理的確率としては比較的高く見積もってしまうというプロスペクト理論で説明できるか

もしれない．

5．3． 乱数の“手動”生成

4章アプローチ2で述べたように，本研究ではまずすごろくとは関係なしに，単に系列として自然に見える乱数列を作成することをサブゴールとする．そこで，一回目実験・二回目実験では，被験者に1～6の数字からなる長さ100の乱数列を「サイコロを振るつもりで」書いてもらった．実際にはサイコロ等の使用は認めていない．

2人の被験者の先頭から40文字分を示す．

・4525143326144641355542665654121422351611

・1523645326413253412156362436152342615243

前者には，2連続・3連続で同じ文字が含まれるような部分が多く含まれているが，後者には2連続の部分すら含まれておらず，数学的には明らかに乱数とは言えない特徴を持ってしまっている．これはこの被験者の誤りあるいは偏りと言え，この被験者に自然に見える乱数を作るなら，目の連続は好ましくないという予想ができる．

【乱数系列の特徴量】

本研究では，疑似乱数系列に「同じ目が2連続する部分の数」といった15の特徴量を導入する．その上で理論値と実際の値を比較し，偏りを検出し，その偏りに沿うように系列を生成することを目指す．用いた特徴量と長さ100の場合のおよその理論値は次の通りである．

F1：全体で出た目の回数のχ^2値．(5.0)

F2～F5：系列の約4分の1（長さ100の場合には1-30番目，24-53番目，48-77番目，70-99番目）のそれぞれにおける出た目の回数のχ^2値．(5.0)

F6：偶数と奇数が並ぶ部分の数．(49.5)

F7：同じ目が2連続する部分の数．(16.5)

F8：同じ目が3連続する部分の数．(2.7)

F9：同じ目が4連続する部分の数．(0.45)

F10：XXYY，XYXY，XYYX など，2つの目が2つずつ登場する部分の数．ツーペア．(6.7)

F11：XXYYYなど，2つの目が2つと3つ登場する部分の数．フルハウス．(1.5)

F12：XYXXなど，両端を含め4つ中3つが同じ目である部分の数．(4.5)

F13：XYXZXなど，5つ中3つの場合．(5.6)

F14：XYXZXXなど，6つ中4つの場合．(1.8)

F15：XXYXZWXなど，7つ中4つの場合．(2.5)

【被験者が生成した系列の特徴】

これら特徴量を用いて，2回目実験16人分を分析した結果を表1にまとめる．F3-F5についてはF2とほぼ同じ結果なので省略した．

F1，F2 は出た目のバラツキを表す．実際にはある目があまり出ない・比較的出るといったことは稀ではないが，被験者は比較的均等に目を配置している．

F6から，偶数→奇数，奇数→偶数の変更を多めに含めていることが分かる．そのほうが「変わった＝乱数である」という感じを持ちやすいのであろう．これは3章で述べたポール・バカンの実験と同様の結果である．

F7，F8，F9は目の連続に関するものだが，全てにおいて，上位8人の平均ですら，理論値よりも小さいことが見て取れる．3章で述べたクラスターの錯覚に関係する．実は，この生成実験後のアンケートで“目が3連続することは何個くらいあるか”という問いには，半数以上が2あるいは3と正しい答えをしている．頭ではそのくらいあると思っていても，短時間で乱数を書けと言われると，“乱数＝散らばった数”といった思い込みからか，連続を忌避する傾向があるようである．

F10からF15は，3連続・4連続ほどではないが，ある目がある部分の中に頻繁に登場するようなパターンである．F7-F9と同様に，上位平均ですら理論値よりも低く，被験者たちにとって「無意識あるいは意識的に，忌避したくなるパターン」であることが見て取れる．

表1　被験者の生成した乱数の特徴

特徴量	理論値	平均	上位平均	下位平均
F1	5.0	2.4	3.3	1.5
F2（～F5）	5.0	2.2	3.0	1.3
F6	49.5	55.6	62.0	49.3
F7	16.5	10.9	15.9	6.0
F8	2.7	0.8	1.5	0.1
F9	0.45	0.06	0.1	0.0
F10	6.7	3.3	5.3	1.3
F11	1.5	0.4	0.8	0.0
F12	4.5	1.4	2.8	0.0
F13	5.6	1.8	2.8	0.9
F14	1.8	0.3	0.5	0.0
F15	2.5	0.5	1.0	0.0

以上のように，被験者によって作成された系列は理論値と比較すると大きく偏った特徴を持つ．これは自然に見える乱数生成のために重要であると考える．一方で，冒頭に挙げたように個人差が

大きいことにも注意すべきである．上位平均と下位平均の差は多くの項目でかなり開いており，またこの表にはないが，一部の人はF7-F15が全て比較的高い（乱数に近い）ことも分かっている．従って，将来的には，全体的な傾向のみならず，アプローチ4にあるように個人的な傾向に沿うように調整をすることが有望であろう．

6. 疑似乱数の生成と評価

5.3節の結果に基づき，我々は「被験者が書いた乱数系列と同じような特徴・偏りを持つ系列のほうが自然に見えるはずである」という仮定を置き，そのような系列を作成することにした．以下に説明のために用いる記号を説明する．

- S: 疑似乱数の系列.
- $f_i(S)$: 系列S の，特徴量 Fi の値
- $[\alpha_i, \beta_i]$: 特徴量 Fi の望ましい範囲
- $Err_i(x)$: 範囲からの逸脱量.
  - $x < \alpha_i$ なら $\alpha_i - x$,
  - $\beta_i < x$ なら $x - \beta_i$,
  - $\alpha_i < x < \beta_i$ なら 0
- γ_i : 逸脱に対する重み

このとき，系列Sに対する最小化したい評価値を

$$Err(S) = \Sigma_i (\gamma_i Err_i(f_i(S)))$$

と定義する．

実験では長さ50の系列を作成して用いることにする．パラメータα_i, β_i, γ_iは，一回目実験の際に得られたもの（表１とは異なるが傾向は同じ）を参考に，表2の通り定めた．F6は理論値より多めに，他は理論値より低めの範囲となっている．

表2　系列評価のパラメータ

特徴量	理論値	下限α	上限β	重みγ
F1	5.0	2	5	3.0
F2（～F5）	5.0	2	5	3.0
F6	24.7	27	30	1.0
F7	8.3	5	8	1.0
F8	1.3	0	1	3.0
F9	0.2	0	0	10.0
F10	3.4	1	3	4.0
F11	0.9	0	0	4.0
F12	2.2	0	1	4.0
F13	2.8	1	2	4.0
F14	0.9	0	0	4.0
F15	1.2	0	0	4.0

標準の疑似乱数生成器を用いて長さ50の系列を300用意したところ，その最良評価値は 2.7，最悪評価値は 239.0 ，平均は52.1であった．最良・最悪のものを示す．

最良のもの(2.7)：
5216623263554331656343121661532335664265315
4342315

最悪のもの(239.0)：
3144325553355455465624644556154356445454256
6656544

最良のものはある意味「不自然なほど」F7～F15の値が小さいわけだが，人間には自然に感じる可能性が高いと考える．逆に最悪のものは同じ目が連続あるいは偏って出ているが，これもまぎれもなく標準の疑似乱数生成器から出力されたものである．

【最小化アルゴリズム】

Err(S)が小さくなるような系列Sを求めるために，本稿の実験では局所探索アルゴリズムを用いた．以下にその概略を示す．

1. Sを標準の疑似乱数で初期化する．
2. 系列の1箇所をランダムに変更しS’とする
3. Err(S’) < Err(S) なら SをS’で置き換える
4. Err(S)=0になるか1000回に達するまで2.～3.を繰り返す

なお，当初はSimulated Annealingを用いる予定であったが，今回の設定では局所探索でも容易にErr(S)=0となったため実装しなかった．特徴量を増やしたり，上限下限の制約をきつくする場合には，それらの手法が必要になるだろう．最適化にかかった時間は今回の設定と標準的なPCで0.02秒程度であり，すごろくの用途には全く問題ない．

6.1. 自然さの評価

Err(S)=0 になるように最適化した疑似乱数系列（調整乱数と呼ぶ）が実際に自然に見えるかを確認するため，二回目実験で被験者実験を行った．

【評価対象】

まず，疑似乱数系列として，以下のＡ，Ｂ，Ｃの群を用意する．

Ａ（標準乱数）：無作為に作成した標準の疑似乱数系列

Ｂ（下位乱数）：300作成した標準の疑似乱数系列をErr()によって評価し，悪い方から20％にあたる60系列を選んだもの

C（調整乱数）：Err(S)=0 となる最適化を行ったもの

これらから（異なる）48つずつを選び，被験者16人に3つずつ割り当てる．

【評価方法】

被験者は，9つのWindowsプログラムを順番に実行する．プログラムが実行されると，１秒間に１文字ずつ，現在のサイコロの目と過去５回分がアラビア数字で表示される．被験者は同じプログラムを2回実行する（同じ系列を2回見る）ことを求められ，その後，以下の指標のうち一つで回答する．

1. 偏りのある乱数に違いない
2. おそらく偏りのある乱数である
3. 分からない
4. おそらく通常の疑似乱数である
5. 通常の疑似乱数に違いない

9つのプログラムにはそれぞれＡＢＣいずれかの群に属する系列が割り当てられており，全ての被験者が異なる系列を見る．9つのプログラムの実行順は定められており，ランダムに，かつ同じ群に属する系列を連続して見ることがないように並べられているが，被験者には群の数や順番は知らされない．

【評価結果】

Ａ，Ｂ，Ｃそれぞれについて，回答1～5が何回あったかを表3にまとめる．

Ｂ群は48回中40回で評価1,2つまり偏りのある乱数であると判断されている．ゲームプレイ中，標準的な乱数が使われていれば，5回に1回程度はこのような乱数系列を経験するわけであり，乱数の自然さへの疑いは生じやすいと言える．

Ｃ群は半数以上の回数評価4,5つまり通常の疑似乱数であると「誤解させる」ことができている．Err()の値が大きいものを選んだＢ群と0にしたＣ群でこれだけはっきりした評価の違いが出ていることから，我々が設計したErr()が少なくとも総合的には効果をあげていることが分かる．ただし，例えば15個の特徴量が全て必要なのか，あるいは他に重要な要素がないのか，という問いにはまだ答えられない．さらなる検討と，より詳細かつ大規模な実験が必要になるであろう．

Ａ群の評価はＢとＣの中間的なものとなっている．中にはErr()が2.7のものも238.0のものもあるように，「人間にとって自然に見えやすい」ものから「不自然に見えやすい」ものまでさまざまであることからこのような結果になるのは自然であろう．

表3　自然さの評価結果

群	１点	２点	３点	４点	５点	平均
Ａ（標準）	11	**19**	2	**11**	5	2.58
Ｂ（下位）	**17**	**23**	2	5	1	1.96
Ｃ（調整）	7	9	4	**23**	5	3.21

6.2.　すごろくにおける不満度の評価

第二回実験ではさらに，作成したＢ群・Ｃ群の疑似乱数系列を用いて，被験者に簡単な一人すごろくをプレイしてもらうことで，どのような場合に不満を感じるのかを調べた．

【実験方法】

図2にプログラムのキャプチャ画面を示す．被験者がプログラムを起動すると，以下の手順で実験が行われる．

1. スタート地点（左端）にコマが置かれる
2. サイコロを振るボタンを押す
3. 分かれ道がなければ，自動でそれだけ進む
4. 分かれ道があれば，上，右，下のどのルートを選ぶか，ボタンで決定する
5. 赤い穴（罠）のあるマスに落ちると負け，緑のマス（右端）に到達すると勝ち．
6. 上記を15ゲーム，2セット行う．15ゲームが終わると，サイコロが変わる旨が表示される．

なお，Ｂ群・Ｃ群の系列は被験者ごとに異なり，また半数が前半Ｂ群後半Ｃ群，半数がその逆の順番で系列を用いるが，被験者には知らされない．

本ゲームでは，30回中に罠に落ちた回数が少なかった16人中上位3位までに1000円分の賞品を渡す旨を事前告知し，インセンティブとした．最適戦略は自明ではないので，プレイヤは試行錯誤しながらルートを選択することになる．

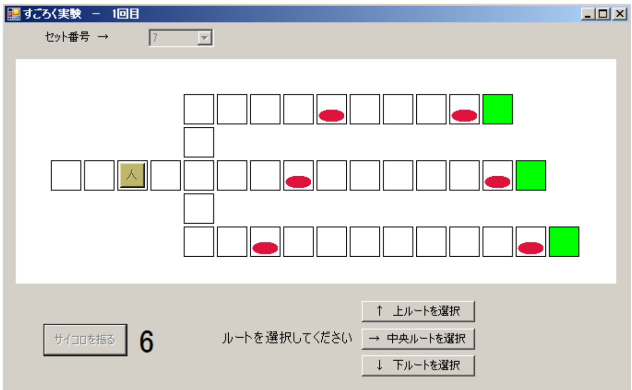

図2：　すごろく実験プログラムの画面

【実験結果1：落ちる回数の予測】

実験後に，実際に罠に落ちた回数と，「15ゲームあたり平均何回落ちるようなすごろくか」の予想値を記入してもらった．予想回数は3回から8回までさまざまだったが，概ね実際に多く落ちた人（最大16回）は多い回数を予想し，あまり落ちなかった人（最小7回）は少ない回数を予想するという結果が出た．その相関係数は0.54と比較的高く，“限られた結果に引きずられて一般的な傾向を予想する”というしばしば指摘される現象が生じている．

【実験結果2：自然さの評価】

実験後に，6.1節と同様の基準で，前半15ゲーム分と後半15ゲーム分のサイコロの自然さを評価してもらった．Ｂ群の平均点は2.44点，Ｃ群の平均点は3.50点で，1ポイント強の差がついた．これは表3にある結果と傾向としては同じである．

前節の実験のほうが全体的に少し点数が低いのは，過去5回分を含め6回分の出目を参照できることで不自然さがより目立ったためではないかと考えている．

【実験結果3：不満度・不信感】

実験後に，前半後半それぞれについて，“プログラムに「サイコロの目を調整して罠に落とされた」と思いますか？”という質問で「１．全く思わない」から「５．強く思う」までの５段階評価をしてもらった．Ｂ群の平均点は3.19，Ｃ群の平均点は2.88で，若干Ｃ群のほうが不満が小さいという結果が得られたが，残念ながら統計的に有意な水準ではない．

罠に落ちた回数と，罠に落とされたと思うかの評価の間には相関があり，Ｂ群で 0.31，Ｃ群で 0.55の相関係数となった．つまり，罠に落ちた人ほど，それがサイコロ調整のせいだと思っていることになる．従って，不満度をより解消するためには，単に自然に見える乱数を使うだけでなく，4章のアプローチ3で述べたように，罠に落ちる回数や落ち方（連続しては落とさないなど）自体を調整してやる必要があるだろう．

7. まとめと今後の課題

本研究では，数学的な意味で良い疑似乱数が，標準的なゲームプレイヤにとっては良いものではない可能性を指摘し，自然に見える疑似乱数系列の生成法を提案，評価した．同じ目が続くなど目立つ部分を少なくすることで，乱数そのものを見せた場合の自然さ，あるいはすごろくで使った場合の自然さともに，大きく改善することが分かった．

一方で，止まりたくないマスがあるようなすごろくでは，そのマスに実際に止まったかどうかが「プログラムに目を操作された」という不信感に強く繋がることが分かり，提案手法だけでは不満度の抑制には不十分であった．今後は，そのようなマスに止まるかどうかも含め調整を行うことで，より不自然さ・不信を感じさせないサイコロの目が生成できると考えている．

謝辞

本研究の一部は，科学研究費補助金 基盤Ｂ研究「ミスを犯す人間らしいゲームAIの研究」の助成を得て行われた．

参考文献

[1] Donald E. Knuth, The Art of Computer Programming, Vol 2, Third Edition, Addison-Wesley, 1997

[2] 松本，擬似乱数ユーザーの方へ-Mersenne Twister法開発者より，日本統計学会誌, J35(2), pp. 165-180 , 2006

[3] カルドセプトR レビュー，DS評価サイト，http://ndsmk2.net/3ds/title.php?title=269

[4] 大宮ソフト開発者jin氏, カルドセプトにおけるダイス目についての一考察，http://www.omiyasoft.com/jnote07.html, 1999

[5] 仲道，伊藤，機械学習を用いた棋力の調整方法の提案と認知科学的評価，30回GI研究会，2013

[6] M.H. Bazerman, D.A. Moore, 行動意志決定論, 白桃書房，2011

[7] 友野典男, 行動経済学 経済は「感情」で動いている，光文社, 2006

[8] Thomas Gilovich, 守 一雄, 守 秀子, 人間この信じやすきもの —迷信・誤信はどうして生まれるか, 新曜社, 1993

[9] J.St.B.T.エバンス, 中島実, 思考情報処理のバイアス 思考心理からのアプローチ，信山社出版，1995

[10] 樺 旦純, ダマされる人・ダマされない人—心のワナにかからないための心理学 , PHP 研究所, 2002

[11]The Skeptic's Dictionary 日本語版 URL:http://www.genpaku.org/skepticj/clustering.html

[12] スコット・ブラウス, 浦谷 計子, 判断力 −判断と意思決定のメカニズム−，日本経済新聞出版社 , 2012