科学研究費助成事業　　研究成果報告書

平成 ２６ 年　 ６ 月　 ９ 日現在

機関番号：１６３０１
研究種目：若手研究(B)
研究期間：2010～2013
課題番号：２２７４００７５

研究課題名（和文）統計的・代数的視点からの擬似乱数高性能化の研究

研究課題名（英文）Improvement of pseudorandom number generators from viewpoint of algebra and statistics

研究代表者
原本　博史（Haramoto, Hiroshi）

愛媛大学・教育学部・講師

研究者番号：４０５１１３２４

交付決定額（研究期間全体）：（直接経費）　　2,900,000円、（間接経費）　　870,000円

研究成果の概要（和文）：(1)擬似乱数の統計的検定において、疑わしい偏りを検出した際に明確な偏りを検出できるまでサンプルサイズを増やしながら検定を行う繰り返し検定を実装した。(2)擬似乱数の高速ジャンプ法について、sliding window algorithmとKaratsuba乗算を用いた方法を実装し、そのプログラムを公開した。(3)擬似乱数の非統計的検定に関してMRGの各ビットごとの検定を行うため、MacWilliams恒等式の一般化を定式化し、検定を行った。

研究成果の概要（英文）：(1) We proposes an adaptive modification of statistical tests, in particular TestU01. This procedure automatically increases the sample size, and tests the PRNG again. It stops when the p-value falls in a clearly rejectable range. (2) We have implemented the codes of the jumping-ahead for MT19937 and WELL19937, in C and C++ languages. These are distributed from a homepage. (3) we report a method to compute the weight distribution of the second to the sixth lowest bits of several PRNGs by using MacWilliams identity.

研究分野：数物系科学

科研費の分科・細目：数学・数学一般(含確率論・統計数学)

キーワード：擬似乱数　代数学　応用数学　統計数学

１．研究開始当初の背景
計算機の高速化・並列化に伴いシミュレーションが大規模化し、大量の擬似乱数を消費するようになっている。そのため擬似乱数生成法に対する高速化・高品質化の要請は一層強まっている。

(1) 擬似乱数の評価法として、一般的に統計的検定が用いられる。NIST や TestU01 などの統計的検定ソフト群では、予め決められたサンプルサイズのサンプルを複数個用意し、ある検定法によりそれぞれの p 値を求め、それらの分布の一様性をカイ二乗検定するという「二重検定」が用いられている。しかし、多くの場合分布の漸近的近似を用いて求めるため、よい擬似乱数にも関わらず近似誤差により生じた偏りにより棄却されてしまうという問題点を持つ。
ところが擬似乱数の利用者にとっては「検定ソフトの使いやすさ」がむしろ重要であり、上記の問題は無視されている現状である。

(2) 計算機の並列化により、擬似乱数の初期値割り当て問題の重要さが増している。複数の計算機で同一擬似乱数生成法を用いたシミュレーションを行う際、擬似乱数列の重複利用があると結果に重大な偏りを含む危険がある。この危険を避けるためには、一台の計算機が使用する擬似乱数よりも大きな自然数 J を取り、一台目の計算機に割り当てる状態の J ステップ先の状態を二台目、さらにその J ステップ先の状態を三台目に、という初期値割り当てを行えばよい。
これまでジャンプの研究は、小さい状態集合や単純な漸化式からなる擬似乱数生成法に関するものが中心であった。そのためメルセンヌツイスター法や WELL など、信頼性の高い超長周期の生成法に対してはこれまでの研究成果は利用不可能であった。

２．研究の目的
以上の状況に対して、本研究では(1)危険を伴う二重検定に対する解決策として「繰り返し検定法」を用いることでより正確で明快な検定結果を提示できるアルゴリズムの開発、(2)先行して研究した sliding window algorithm とは異なるさらに高速な線形擬似乱数生成法のジャンプ計算法の完成を目指した。

以上の研究成果は実行可能なプログラムコードとして擬似乱数研究関連のホームページから公開し、必ずしも数学に詳しくない利用者にとっても利用しやすい形で公表することを目的とした。

また、関連する研究として、有限体・有限環上の線形漸化式を用いた擬似乱数生成法に対する非統計的な検定法を研究した。

(3)MacWilliams 恒等式を用いた擬似乱数の非統計的検定により、特定の擬似乱数に関して、統計的検定では到達できない精度(10 の -100 乗の精度)での偏りを検出する手段の確率と、実際の擬似乱数に対する検定を行う。例えば、ラグ付きフィボナッチ法による擬似乱数生成法に関してビットごとの安全／危険なサンプルサイズを計算することで、下位ビットの捨て去り改良がどの程度有効であるかを不安定な統計的検定によらず示すことを目的とした。

３．研究の方法
(1)擬似乱数の繰り返し検定に関する研究
C 言語による擬似乱数検定ライブラリ群 TestU01 を用いて、繰り返し検定の実装と代表的な擬似乱数生成法に関する検定、および既存の検定結果との比較と考察を行う。繰り返し検定の手法は、
①サンプルサイズ n のサンプルに対して検定を行い p 値を計算する
②p 値が 10 の -10 乗以下ならば、擬似乱数は「危険」と判定し棄却する
③p 値が 0.1 以下なら、倍のサンプルサイズ 2n で前のサンプルとは独立なものを生成し①に戻る
④一定回数ループ①から③を抜けなければ「危険」と判断して棄却、それ以外は「安全」と判断して終了する
この手法によって、一度疑わしい偏りを検出したら徹底して追及し、10 の -10 乗程度の精度で偏りの検出によって「安全」か「危険」かの判断を目指す。

(2)擬似乱数の高速ジャンプの研究
メルセンヌツイスターに関する高速ジャンプ計算法を NTL(Number Theory Library)を用いて実装する。状態空間の次元を d とするとき、Karatsuba の多項式乗算の計算量は d の 1.59 乗のオーダーであることに注目し、形式的冪級数を用いて「次状態の計算」を「多項式の乗算」に対応させることで効率的な計算が可能となる。先行研究である sliding window algorithm と比較して高速な計算を実現しており、擬似乱数の超長周期化が望まれるこの分野の発展に大きな寄与が見込める。そこでまず、いくつかの擬似乱数生成法について、計算アルゴリズムの発見とその計算量評価をする。
また使いやすいプログラムコードを公表することで利用しやすい形で公表する。

(3)MacWilliams 恒等式を用いた検定法
先行研究で、擬似乱数の出力の偏りが引き続く未来の出力にどう影響するかを調べるため、出力の関係式のなす空間で数え上げを行い、符号理論に現れる MacWillams 恒等式を用いて出力のパターンの数え上げを可能にした。この方法により、ある種の擬似乱数の最下位１ビットの 0-1 分布を完全に求めている。この方法を拡張し、ラグ付きフィボナッ

チ生成法に代表される擬似乱数生成法 MRG について任意のビット毎にその 0-1 分布を計算する。これまでの 2 元体上のベクトル空間上での MacWilliams 恒等式を拡張し、2 のべき乗を法とする整数環の剰余環上の MacWilliams 恒等式を用いて数え上げを行うのが要点である。数え上げ部分は C 言語および NTL を用いて高速に計算し、多項式計算の部分は Mathemacita の組み込み関数を用いて行う。

４．研究成果

(1) 擬似乱数の統計的検定のプログラムに関して、TestU01 で広く用いられている Crush(約 100 種の統計を一斉に行うテスト群プログラム)を改良し、繰り返し検定を行うプログラムを作成した。一部検定法は検定途中でソートを行う等しているため、サンプルサイズを大きくした際にプログラムが停止する恐れがある。これらの検証を行いテストを行った結果、これまでは 10 の-2 乗程度で棄却されていた結果を 10 の-10 乗程度で棄却できるようになった。

LCG(31 ビット擬似乱数生成法)に Hamming weight 検定を行ったときの p 値のオーダー

標本数	10^6	2x10^6	4x10^6	8x10^6
1 回目	10^-2	10^-8	10^-7	<10^-15
2 回目	10^-2	10^-3	10^-10	-

プログラムコードの公開に関して最終的な調整を行っており、近日中に擬似乱数関係のホームページに掲載を予定している。

(2) 擬似乱数のジャンプ法に関して、メルセンヌツイスター法に関するプログラムを整備し、擬似乱数関係のホームページに掲載した。特に利用者からのコード改良に応え、更新を行っている。一部結果は並列計算機用にメルセンヌツイスターのパラメータを複数生成するコードに引用されている。
擬似乱数生成法 WELL に関するプログラムに関しては、検証の結果 sliding window algorithm のみ実装を行った。Karatsuba 乗算を行ったあとの結果に対応する多項式から、擬似乱数の状態空間の元を復元する作業に d の 2 乗のオーダーの計算時間がかかっており、NTL による計算高速化の部分が生かせないことによる。

(3) ラグ付きフィボナッチ生成法に関するビットごとの安全／危険なサンプルサイズの計算により、以下のような結果を得た。ラグ付きフィボナッチ生成法の最下位１ビットを捨て去る改良法に関しては古くから用いられていたが、数値的にこの改良の効果を示すことができた。また、状態空間の次元と引き続く未来１ビットの安全／危険なサンプルサイズを計算することにより、ビットがあがるごとに安全／危険なサンプルサイズが何倍になるかがほぼ計算できることが予想された。これらの結果に関して論文を作成し、査読付き論文誌に受理された。

漸化式 x_{i+31}=x_{i+28}+x_i による擬似乱数生成法の安全／危険なサンプルサイズ

ビット	1	2	3	4
安全	2566	10293	41195	164806
危険	12248	49113	196568	786390

漸化式 x_{i+100}=-x_{i+63}+x_i による擬似乱数生成法の安全／危険なサンプルサイズ

ビット	1	2	3	4
安全	115728	462951	1851846	7407423
危険	482844	1931532	7726286	30905312

さらに MRG の項数を増やした際のサンプルサイズの変化を、3 項から 11 項までの MRG について計算した。ビットの捨て去り方によっては前のビットのサンプルサイズよりも小さいサンプルサイズが得られる場合があり、必ずしも分布が改善できるとは限らないことが判明した。

9 項 MRG の危険なサンプルサイズ

ビット	1	2	3	4
倍率	-	256	14.1	0.216

11 項 MRG の危険なサンプルサイズ

ビット	1	2	3	4
倍率	-	1024	19.5	0.233

(一つ前のビットの危険なサンプルサイズに対して、何倍になったかを示している)

重み数え上げ多項式は状態空間の次元とほぼ同じ次数の多項式であり、高品質の擬似乱数については多項式の展開が困難になる。そこで重み数え上げ多項式を 0 の重みに対応する最も次数の高い項とその次に高い項のみを取り出し、形式的に MacWillimas 恒等式の反転公式を適用した。下の表のように計算誤差は 10 の-10 程度と極めて小さく、サンプルサイズの計算に関して無視できることが実験的に得られた。これにより 6 ビット目までが限界だった分布の計算が 9 ビット目程度まで可能となった。

漸化式 x_{i+31}=x_{i+28}+x_i に関する計算誤差

ビット	1	2	3	4
誤差	10^-8	10^-9	10^-10	10^-11

2014 年 1 月にモンテカルロ法・準モンテカルロ法の研究集会を広島大学で開催し、東京大学大学院博士課程学生 2 名を招聘した。MacWilliams 恒等式を用いた準乱数の評価方法に関する最新成果を基に、擬似乱数の評価法に関する非統計的検定の方法を改良し、非零係数が全て 1 の場合の MRG については、任意ビットの検定が可能となった。特に十分高

いビットについては、次数や項数によらず、約4倍で安全／危険なサンプルサイズが増加することがわかった。
以上の結果を 2014 年 4 月に開催されたモンテカルロ法・準モンテカルロ法の国際研究集会 MCQMC2014 において口頭発表した。今後、同研究集会報告集への査読付き論文を作成し、投稿を予定している。

５．主な発表論文等
（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 2 件）
(1) Hiroshi Haramoto, Makoto Matsumoto, Takuji Nishimura, Yuki Otsuka, “A non-empirical test on the second to the sixth least significant bits of pseudorandom number generators,” Monte Carlo and Quasi-Monte Carlo Methods 2012, 417-426, 2013 年 (査読あり)

(2) 原本博史, 松本眞, 西村拓士, 大塚祐樹, MacWilliams 恒等式による擬似乱数列の下位ビットの分布計算, 日本応用数理学会 2012 年度年会講演予稿集, 43-44, 2012 年 (査読なし)

〔学会発表〕（計 5 件）
(1) Hiroshi Haramoto, Makoto Matsumoto, “An approximation of the weight distribution of the n-th bits of pseudorandom number generators,” Eleventh International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing (MCQMC2014), 2014 年 4 月 9 日, ルーヴェン・ベルギー

(2) 原本博史, 松本眞, 西村拓士, 大塚祐樹, MacWilliams 恒等式による擬似乱数の非統計的検定, Workshop on Galois point and related topics, 2012 年 9 月 16 日, 山形大学

(3) 原本博史, 松本眞, 西村拓士, 大塚祐樹, MacWilliams 恒等式による擬似乱数列の下位ビットの分布計算, 日本応用数理学会 2012 年年会, 2012 年 8 月 29 日, 稚内市

(4) Hiroshi Haramoto,”A non-empirical test on the second to the sixth lowest bits of pseudorandom number generators,” Workshop for Quasi-Monte Carlo and Pseudo Random Number Generation, 2012 年 6 月 13 日, 東京大学

(5) Hiroshi Haramoto, Makoto Matsumoto, Takuji Nishimura, Yuki Otsuka, “A non-empirical test on the second to the sixth least significant bits of pseudorandom number generators,” Tenth International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing (MCQMC 2012), 2012 年 2 月 13 日, シドニー・オーストラリア

〔図書〕（計 4 件）
(1) 河東泰之他(編), 原本博史(共著), 線形代数, サイエンス社, 2013 年, 121-139

(2) 河東泰之他(編), 原本博史(共著), 線形代数問題集, サイエンス社, 2013 年, 70-89

(3) 河東泰之他(編), 原本博史(共著), 基礎数学, サイエンス社, 2012 年, 206-220

(4) 河東泰之他(編), 原本博史(共著), 基礎数学問題集, サイエンス社, 2012 年, 82-93

〔その他〕
ホームページ等
http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/JUMP/index.html

６．研究組織
(1)研究代表者
原本 博史 （HARAMOTO, Hiroshi）
愛媛大学・教育学部・講師
研究者番号：40511324

(2)研究分担者
なし

(3)連携研究者
なし