

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 5 日現在

機関番号：34416

研究種目：基盤研究(C)

研究期間：2010～2013

課題番号：22540159

研究課題名(和文) 不規則な構造上の離散パターンに関する統計的推測の研究

研究課題名(英文) Statistical inference on discrete patterns on random structures

研究代表者

安芸 重雄 (AKI, Sigeo)

関西大学・システム理工学部・教授

研究者番号：90132696

交付決定額(研究期間全体)：(直接経費) 2,000,000円、(間接経費) 600,000円

研究成果の概要(和文)：主な研究成果は次のとおりである。
向きのないグラフィカル・モデル上の離散パターンに関する厳密な確率分布を導出する方法を提案した。また、可換系列や部分可換系列について、観測された離散パターンのデータから統計的推測ができるような母数モデルを提案し、それらに対する混合測度と系列の逐次生成法を与えた。さらに、クーポン集め問題を用いて置換群上の母数モデルを提案した。このモデルは順位データの解析に有効で、適当なサンプルサイズの標本に基づいて母数の最尤推定や仮説の尤度比検定を行うことができる。

研究成果の概要(英文)：Main results of the research are as follows. A method for deriving exact distributions of numbers of discrete patterns observed on undirected graphical models is proposed. Parametric statistical models are proposed for observed discrete patterns on exchangeable or partially exchangeable sequences of trials. The mixing measures of the exchangeable sequences also obtained. Further, a parametric statistical model on the permutation group is proposed by using the coupon collector's problems. The statistical model is useful for statistical inference of ranking data such as maximum likelihood estimation and likelihood ratio tests.

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：統計数学

1. 研究開始当初の背景

1980年代から急速に発展し始めた、連やパターンなどの離散分布の研究は、90年代以降もマルコフ連鎖埋め込み法や条件付き確率母関数の方法などの新しい方法によって、構造的には、向きをもったグラフィカルモデル上で研究が進み、依存性としては、マルコフ性や高次マルコフ性の下で、新しい結果が次々と得られるようになっていた。当時、それでも扱うことが難しい重要な依存性のモデルとして、可換系列と無向グラフ上のグラフィカルモデルの問題が残っていた。有限マルコフ連鎖埋め込み法も、拡張された条件付き確率母関数の方法も、一度に扱うと複雑になってしまう問題を逐次的(sequential)に扱うことにより処理する方法である。そのためには、自然な「向き」が大きな役割を果たしてきた。可換性のモデルは「系列」になっていて、自然な向きがあるように見えるけれども、可換性により、順序を変えても分布が変わらないという性質から、扱うことは決して易しくはなかった。

2. 研究の目的

上に述べたような背景の下で、厳密な確率分布を実際に計算することができて、それに基づいた統計的推測が可能になるという点を考慮して、次のような目的をもって研究を開始した。

(1) 可換系列上の新しい母数モデルを逐次的に構成し、条件付き確率母関数の方法を使って、複雑な離散パターンに関する厳密な標本分布の導出を目標として研究を行う。

(2) 無向グラフの頂点に確率変数を配したグラフィカル・モデルの上の離散パターンに関する統計量の厳密な標本分布の導出に関しても、適当な「向き」を作り出し、その向きに沿って条件付き確率母関数の方法を使う研究を行う。

3. 研究の方法

(1) 複雑な離散分布の計算を行うためには、数式処理とグラフィックスを統合的に活用する必要があるため、それを可能にすることができるソフトウェアを入手して、離散分布研究に関する手法を開発した。

(2) 確率論、数理統計学、組合せ論など、この研究に関連する書籍を入手して、必要な理論を整理した。

(3) この研究に興味をもつ国内の研究者と積極的に意見交換を行った。

4. 研究成果

研究目的の(1),(2)とも、当初考えていた研究は早期に解決することができ、国際的な学術雑誌に結果を公表することができた。とくに、可換系列上での確率分布の研究に関しては、大変見通しがよくなり、(1)と(2)を合

わせた研究も進んだ。具体的には次のような成果が得られた。

(1) 可換系列上のパターンの分布について、de Finetti の定理との関連を調べた。また、さまざまな連に関連する厳密分布を求めた。さらに、マルコフではない新しいタイプの部分可換性モデルについて研究した。 $\{0,1\}$ -値をとるランダムな2値系列において、2以上の自然数 k を与えたとき、 $"0", "10", "110", \dots, "11\dots10", "11\dots1"$ という長さ k までのブロックをどのように交換しても同時分布が等しいような部分可換列の分布が、オーダー k の2値系列と呼ばれる系列の分布の混合(mixture)になるという結論を得た。この結果を基に、オーダー k の2値系列の混合モデルとなるような具体的な母数モデルにおいて、連やパターンの待ち時間の分布を求めた。さらに、この標本分布に従う観測値に基づく尤度の計算と最尤推定値の計算が実行できることを示すような具体例を与えた。この結果は〔雑誌論文〕の に公表した。

(2) 今までPolyaの壺からのサンプリングに限られていた、逐次的に生成できる可換性モデルを、任意の周辺分布を与えて生成することに成功した。そして、このモデルに基づいて生成した、可換性をもつ長さ n の確率変数列から生成される経験分布関数が n を大きくしていくときに、あるランダムな分布関数に法則収束することを示した。このランダムな極限分布関数はベータ分布に従う確率変数を用いて明示的に与えられる。さらに、このように生成された可換列の順序統計量の同時分布も得られた。また、可換な離散モデルについても考察し、標本分布の導出に有効な条件付き確率母関数の方法を、可換性のもとで利用できるように拡張を行った。そして、今まで知られていなかった新しい離散可換列の混合測度を求めた。それらは、ベータ分布やディリクレ分布とは異なるものである。この結果は〔雑誌論文〕の に公表した。

(3) 向きをもたないグラフィカル・モデルにおいて、条件付き確率母関数法を利用できるように拡張し、パターンの数の確率分布を求めるためのアルゴリズムを与えた。一般に、離散的で複雑な問題については、「場合の数」を一度に求めるのは非常に困難であるので、向きがあれば向きに沿って1ステップ進む度に、条件付き確率母関数を更新していく方法が用いられてきた。この研究では、向きのないグラフィカル・モデルに対して適当な向きを付けて、条件付き確率母関数法を用いることができるように理論を拡張した。その結果、今までは、部品が円上に配置されたような簡単な連続システムに関する確率分布しか研究されていなかったが、この成果によって、たとえば球面やトーラスのような曲面上に部品が配置されているような連続システムの信頼度の計算ができるようになった。この計算はこの研究によって提案されたアルゴリズムを利用して、数式処理システムで行う

ことができる。したがって、各部品が独立である場合にはもちろんであるが、各部品が可換性をもつ場合にも厳密な確率分布を求めることができる。この結果は〔雑誌論文〕の に公表した。

(4) 一般化された Polya サンプルングの下で、クーポン・コレクター問題の待ち時間の厳密分布を求めた。この結果、 m 次対称群の上の自然な統計的母数モデルを構築することができた。この問題は、数学的には、可換性のモデルを真に含む問題であり、今までの研究に対して新しい応用を与えることになった。観測値が何かの順位であるような、「ランク付け問題」は、数学的には m 次対称群が自然な標本空間になる。しかしながら、これまでは、群の行列表現に基づいた指数型分布族が研究されていた他には、直感的な意味がわかるような統計的モデルがほとんど提案されていなかった。ここでの成果は、このようなランク付け問題に有効である。そのような視点から、いくつかの具体的な統計的モデルにおいて、 m 次対称群上の観測値に基づいた母数の最尤推定と統計的仮説の尤度比検定が適当なサンプル・サイズの場合に実行できることをシミュレーションを使って確かめた。とくに、ランク付け問題における帰無仮説の下で、尤度比検定統計量の分布はカイ 2 乗分布で十分に近似されることがわかった。この結果は〔雑誌論文〕の に公表した。

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 7 件)

Aki, S. and Hirano, K. (2013). Coupon collector's problems with statistical applications to rankings, *Annals of the Institute of Statistical Mathematics*, Vol. 65, 571-587. (査読有)
DOI:10.1007/s10463-012-0382-9

Inoue, K. and Aki, S. (2013). Distributions of numbers of runs and scans on directed acyclic graphs with generation, *Computational Statistics*, Vol. 28, 1133-1150. (査読有)
DOI:10.1007/s00180-012-0349-6

Inoue, K., Aki, S., and Balakrishnan, N. (2013). Generating functions of waiting times and numbers of visits for random walks on graphs, *Methodology and Computing in Applied Probability*, Vol. 15, 349-362.

(査読有) DOI:10.1007/s11009-011-9246-7

Aki, S. (2012). Statistical modeling for discrete patterns in a sequence of exchangeable trials, *Annals of the Institute of Statistical Mathematics*, Vol. 64, 633-655.

(査読有) DOI:10.1007/s10463-011-0325-x

Aki, S. and Inoue, K. (2012). Exact distributions of the number of pattern occurrences in undirected graphical models, Vol. 42, 59-79. (査読有)

<http://www.terrapub.co.jp/journals/jjss/frame/42.html>

Inoue, K. and Aki, S. (2011). Bivariate Fibonacci polynomials of order k with statistical applications, *Annals of the Institute of Statistical Mathematics*, Vol. 63, 197-210. (査読有)

DOI:10.1007/s10463-008-0217-x

Inoue, K., Aki, S. and Hirano, K. (2011). Distributions of simple patterns in some kinds of exchangeable sequences, *Journal of Statistical Planning and Inference*, Vol. 141, 2532-2544. (査読有)

DOI:10.1016/j.jspi.2011.02.002

〔学会発表〕(計 6 件)

安芸重雄, 平野勝臣, 「クーポンコレクター問題と統計的推測」, 日本数学会 2014 年度年会, 2014 年 3 月 17 日, 学習院大学 (東京都)

井上潔司, 安芸重雄, 「非サイクル有向グラフ上の連, スキャンの分布」, 日本数学会 2014 年度年会, 2014 年 3 月 17 日, 学習院大学 (東京都)

安芸重雄, 井上潔司, 「向きのないグラフィカルモデル上の離散パターンの数の分布」, 日本数学会 2013 年度年会 2013 年 3 月 21 日, 京都大学 (京都府)

井上潔司, 安芸重雄, 平野勝臣, 「Distributions of simple patterns in some kinds of exchangeable sequences」, 2012 年度統計関連学会連合大会 2012 年 9 月 11 日, 北海道大学 (北海道)

安芸重雄, 「Exchangeable sequences に基づいた離散分布モデルと統計的推測」, 日本数学会 2012 年度年会, 2012 年 3 月 28 日, 東京理科大学 (東京都)

井上潔司, 安芸重雄, 「可換試行列上の連
の分布」, 日本数学会 2011 年度年会, 2011 年
3 月 22 日, 早稲田大学 (東京都)

〔図書〕(計 0 件)

〔産業財産権〕

出願状況 (計 0 件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

取得状況 (計 0 件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕

ホームページ等

6. 研究組織

(1) 研究代表者

安芸 重雄 (AKI SIGEO)
関西大学・システム理工学部・教授
研究者番号：90132696

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

()

研究者番号：