

令和 4 年 6 月 7 日現在

機関番号：17102

研究種目：基盤研究(C)（一般）

研究期間：2019～2021

課題番号：19K06856

研究課題名（和文）ノンパラメトリック検定の拡張によるデータ解析の有効範囲の拡大

研究課題名（英文）Enlargement of the scope of data analysis by extending nonparametric tests

研究代表者

粕谷 英一（Kasuya, Eiiti）

九州大学・理学研究院・准教授

研究者番号：00161050

交付決定額（研究期間全体）：（直接経費） 2,400,000円

研究成果の概要（和文）：広い範囲のデータ解析において、ノンパラメトリックな統計的検定が広く使われてきた。ノンパラメトリックな検定について、実際にはデータの分布の仮定があるにもかかわらず、分布について全く条件がないといった、誤解があった。対応のない2処理間の位置の比較である、Mann-WhitneyのU検定（Wilcoxon順位和検定）は、母集団分散が異なる状況に対応していないが、その条件を緩めて分散が異なっても適用可能な検定が提案されている。これらのノンパラメトリックな検定が、現実のデータ解析の状況をよくカバーしているかを検討し、カバーされているケースとされていないケースおよびその特徴を明らかにした。

研究成果の学術的意義や社会的意義

データ解析の広い範囲において、母集団分布が既存のよく使用される確率分布（代表的には正規分布など）にしたがうことは想定し難く、その場合、ノンパラメトリックな検定が広く使われている。だが、ノンパラメトリックな検定といえどもどんな場合にでも使えるわけではない。これまでのノンパラメトリックな検定でカバーされている範囲とされていない範囲を検討し、明らかにした。

研究成果の概要（英文）：For a wide range of data analysis, nonparametric statistical tests have been applied. These tests make a kind of assumptions on the distribution of the data. However, this has been often overlooked. Mann-Whitney U test (i.e., Wilcoxon rank sum test), which is a test for comparison of locations between 2 treatments in unmatched data, does not allow unequal variances. Several nonparametric tests that allowed unequal variances were postulated. Whether these nonparametric tests cover actual problems in data analysis is examined. The region covered by these methods and that not covered by them are clarified. Factors affecting this result is also clarified.

研究分野：生態学

キーワード：データ解析 ノンパラメトリック検定 ノンパラメトリック法 確率分布 不等分散 データの分布
高次のモーメント

1. 研究開始当初の背景

ノンパラメトリックな統計的検定は、データが由来する母集団分布として、等分散の正規分布や二項分布などの単純な理論的な確率分布を想定し難い時に、さまざま分野におけるデータ解析に対して広く使われてきた (Siegel & Castellan Jr., 1988, *Nonparametric Statistics for the Behavioral Sciences*, Second edition, McGraw-Hill)。ノンパラメトリックな統計的検定の代表的な方法には Mann-Whitney の U 検定 (Wilcoxon の順位和検定)、Kruskal-Wallis の検定 (順位分散分析と呼ばれる検定の1つ)、Wilcoxon の符号化順位検定、Kendall の順位相関、Spearman の順位相関などがある。

ノンパラメトリックな統計的検定が特定の分布を仮定していないことは、広い範囲のデータへの適用上の大きな利点であるが、一方では、データの分布に関していかなる仮定もおいておらず、データの分布についてはなんら注意することなく用いることが可能であるという誤解も招いてきた。実際にはデータの分布について何の仮定もしていないわけではなく、たとえば、ノンパラメトリックな統計的検定の代表とも言える、Mann-Whitney の U 検定 (すなわち Wilcoxon の順位和検定) を例にとると、比較される分布が定数だけずれていること、言い換えると、分布の形が同じであることが想定されている (ずれがゼロである時が帰無仮説の場合にあたることになる) (柳川, 1995, ノンパラメトリック法、培風館など)。この仮定は、location-shift model あるいは shift model と呼ばれることがある。

しかし、ノンパラメトリックな統計的検定はデータの分布について何の仮定も置いていないという誤解から、この分布に関する仮定に明瞭に反する形で使われることが少なくなかった。実際のデータあるいはその現象への知識から、分散が等しいことが想定しがたいときに、むしろ好んで意図的に (Student の t 検定のような等分散を仮定した統計的検定ではなく)、Mann-Whitney の U 検定 (Wilcoxon の順位和検定) のようなノンパラメトリックな検定が使われることも多いことが指摘されている (たとえば、Kasuya, 2001, *Animal Behaviour*, 61: 1247-1249)。

観察されたデータから、分散が等しくないことが高い確度で言えることは、決して少なくはない。そのための現実のデータからの必要性もあり、分散が等しくないことに対応したノンパラメトリックな統計的検定が提案されている。Mann-Whitney の U 検定のような対応のない2処理の位置の比較についての、Fligner-Policello の検定 (M. A. Fligner & G. E. Policello, *Journal of the American Statistical Association*, 76: 162-168, 1981) や Brunner-Munzel の検定 (*Biometrical Journal*, 42: 17-25, 2000) がそのような方法の主要な例である。これらは論文題名に Behrens-Fisher 問題が含まれていることから、正規分布の場合の Welch の検定 (Welch, B. L. *Biometrika* 29: 350-362, 1938 および Welch, B. L. *Biometrika* 34: 28-35, 1947) などに対応するものと考えられることができる。

2. 研究の目的

実際のデータ解析上の必要性を考えると、分散の違い以外の分布の違いがあるときに、位置の違いを検定しようとしている例は多い。そこから、これらのいわば改良型の Mann-Whitney の U 検定 (Wilcoxon の順位和検定) やそれに類似した方法によりカバーされておらず、新たな方法が必要とされる状況があるという可能性が高く、その部分がどのような場合なのかを明らかにするとともに、その部分に対応した方法を見出すことが本研究の目的である。

ノンパラメトリックな統計的検定は、さまざまな構造のデータに対して用いられるが、本研究では、とくに2処理の位置パラメータの比較でデータ点間に対応のない場合をおもに扱った。この場合は単純であるとともに、そこで得られた結果は他の状況に拡張が容易であると考えられる。2処理の比較でデータ点間に対応のない場合への集中は、たとえば母集団分布が、独立な2サンプルの t 検定との相対的な検出力にどのような影響を与えるかの分析のように、検定法の比較においてこれまで使われてきており、有力な戦略である。

3. 研究の方法

ノンパラメトリックな統計的検定では、2処理の比較でデータ点間に対応のない場合とは、代表的には Mann-Whitney の U 検定 (Wilcoxon の順位和検定) が使われてきたケースに当たる。このケースでは、対応がないのでブロックはない。なお、このケースは、2水準のカテゴリカルなデータを説明変数とし、量的な変数を目的変数とした、回帰であるとも見ることができる。

本課題では、おもにコンピューターを用いた数値計算により、統計的検定における、2種類の誤りの率を計算した。まず第1種の誤り (type I error) の率であるが、これは帰無仮説が正しい時に誤って検定において有意な結果を得て帰無仮説を棄却してしまう率である。この率が宣言値 (すなわち有意水準) よりも高いことはその検定方法の妥当性が失われていることを示し、もう1つの誤りである第2種の誤りの率を検討するまでもなく、当該検定方法は使われるべきではないことを意味する。

第2種の誤り(type II error)は、対立仮説側が正しい時に帰無仮説を棄却することができず、帰無仮説を受容してしまうことを指す。対立仮説が正しい時にこの誤りを犯さない、すなわち有意な結果を得て正しく帰無仮説を棄却できる確率が検出力である。数値計算の結果と、あわせて、解析的な検討による結果をも用いた。数値計算にはR上でプログラムを作成して行なった。

想定した状況の広い範囲において第1種の誤り(type I error)の率がインフレーションを起こしていない時には、必要に応じて第2種の誤り(type II error)の率をも求めた。これはその状況における当該検定方法の検出力の大きさを求めていることにあたる。なお、Mann-WhitneyのU検定(Wilcoxonの順位和検定)などのノンパラメトリックな統計的検定では、尺度水準としては、順序あるいはそれ以上であればよいが、量的な変数すなわち間隔尺度や比率尺度のデータを用いた。したがって、平均値も計算可能である。誤りの率の計算においては、1つの条件において十万回の検定(二十万個のサンプルということになる)を行うことを基本とした。誤りの率の計算では、有意水準0.05を採用したが、可能な限り、たとえば十万回の検定の個々の検定での有意確率を記録したので、他の有意水準での結果もすぐに集計可能であった。

使用した検定方法は、Mann-WhitneyのU検定(Wilcoxon順位和検定)、中央値検定(メジアン検定)、Fligner-Policelloの検定、Brunner-Munzelの検定、Studentのt検定、Welchのt検定などであり、必要があれば、いったんデータの値ではなくその順位を使ってStudentのt検定およびWelchのt検定を行う方法も検討した。Studentのt検定、Welchのt検定などは一種の比較対象としてのものであり、おもな検討の対象はノンパラメトリックな検定方法であった。また、サンプルサイズが大きい場合について、尤度比検定など最尤法に基づく検定も適用した。

4. 研究成果

まず、母集団の分散が異なる状況での、第1種の誤りの率を求めた。母集団の分散が異なる状況は、理論的な主要な確率分布(正規分布、三角分布、ロジスティック分布、一様分布など)において尺度パラメーターを変更した複数の場合の比較と、ある確率分布を位置パラメーターのまわりで拡大あるいは縮小した分布と元の分布の比較の両方で行った。後者は、ヒストグラムを位置パラメーターの値を揃えておいて横方向に伸ばすあるいは縮める操作をすることにあたる。これらの状況でもlocation-shift modelは成立していない。

先行する研究においても示されているように、Mann-WhitneyのU検定(Wilcoxonの順位和検定)は、広い範囲で第1種の誤りの率がインフレーションを示し、宣言値よりも実質的に大きかった。さらに、サンプルサイズと母集団分散の関係によっては、第1種の誤りの率は過小な値をとった(このときには検出力の低下を招いていることを第2種の誤りの率の計算により確認した)。このようなMann-WhitneyのU検定(Wilcoxon順位和検定)の挙動は、Studentのt検定を不等分散の状況で適用したときの結果と定性的によく類似していた。

ノンパラメトリックな統計的検定の中で、標本中央値に対する大小関係という二値的な情報しか使っておらず、きわめて頑健であろうと期待される中央値検定でさえも、母集団分散に大きな違いがあるときには、Mann-WhitneyのU検定(Wilcoxonの順位和検定)と同様の、第1種の誤りの率のインフレーションが観察された。ただし、全体的には、中央値検定の第1種の誤りの率の程度はMann-WhitneyのU検定(Wilcoxon順位和検定)よりも小さかった。その意味では、比較すると、相対的には中央値検定の方がより頑健であると言えるかもしれない。だが、検定の計算的な操作として極めて単純な中央値検定と言えども、不等分散下での第1種の誤りの率がインフレーションと無縁ではないことは銘記しておくべきと考えられる。中央値検定が第1種の誤りの率のインフレーションを示すのは、大小関係の基準として標本中央値を使っており、標本中央値はサンプリングエラーの影響を受けるためであろうと考えられる。この点で、中央値検定は、同様に見えるかもしれないが、対応がある場合に使われるノンパラメトリックな統計的検定である符号検定とは異なっていると考えられる。

Fligner-Policelloの検定およびBrunner-Munzelの検定は、Behrens-Fisher問題に対応したものと考えられ、この分散が異なる状況では、一般に第1種の誤りの率のインフレーションを示さず良好な結果を示した。両検定の間の違いは全般的には小さかったが、異なる場合には、Fligner-Policelloの検定の方がBrunner-Munzelの検定に比べて、第1種の誤りの率が高い値を示すことが多かった。Behrens-Fisher問題が想定している、正規分布あるいはそれに類似した分布で、分散だけが異なるときには、両検定は全般的に良い結果をもたらしたと考えられる。一方で、正規分布とは三次や四次のモーメントが大きく異なる分布では、両検定でも第1種の誤りの率が実質的に宣言値を超える場合も見られた。

次に分散だけでなく、それ以外の点でも母集団分布が異なる場合について、同様に第1種の誤りの率を検討したところ、Mann-WhitneyのU検定(Wilcoxon順位和検定)だけでなく、Fligner-Policelloの検定やBrunner-Munzelの検定においても第1種の誤りの率が宣言値との比較で実質的に高くなる場合が見られた。分布の三次や四次のモーメントがその結果に影響していることが示唆された。

分散以外の指標においても母集団分布が異なる場合、サンプルサイズがある程度大きく、得られたデータから、それぞれの母集団の確率分布が十分正確にわかるのであれば、最尤法に基づいて位置パラメーターを母数にとって検定することは、第1種の誤りの率の点でも効果的だった。母集

団分布の形に大きな違いがあっても、尤度比検定などは良好な結果を与えた。したがって、サンプルサイズがあまり小さくなく、既存の研究での知識から母集団分布が正確にはわからないケースが問題となる。現実にノンパラメトリックな統計的検定が適用されているのは、そのようなケースが多く、またデータ解析全体を見渡してみてもそういったケースの頻度は決して低くない。

サンプルサイズが小さいか中程度の場合を想定して、Max test など複数の検定方法を同じデータに適用した結果を統合する検定方法は有効な場合があることが示唆された。また、位置パラメータの種類の影響についても検討を加えた。

本課題の成果として日本生態学会全国大会において、発表を行なった。なお、新型コロナウイルス感染症の蔓延に伴う規制のため、当初に想定していた学会等での発表のうちのその他の部分はやむを得ず中止した。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 粕谷英一・阿部真人
2. 発表標題 ノンパラメトリック検定が平均値や中央値の差の検定に使えないとき
3. 学会等名 日本生態学会
4. 発表年 2021年

1. 発表者名 粕谷英一
2. 発表標題 一般化線形モデルによるデータ解析：交互作用項の存在は回帰係数の意味を変える
3. 学会等名 日本生態学会
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------