

令和 元 年 6 月 10 日現在

機関番号：34419

研究種目：基盤研究(C)（一般）

研究期間：2015～2018

課題番号：15K00057

研究課題名（和文）2×2分割表における因果効果の正確検定に関する研究

研究課題名（英文）Exact tests of causal effects for a two-by-two contingency table

研究代表者

千葉 康敬 (CHIBA, Yasutaka)

近畿大学・医学部附属病院・准教授

研究者番号：80362474

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：2値アウトカムを比較する2群間ランダム化試験で用いられる正確検定手法は、全員について治療効果がないというsharpな因果帰無仮説のためのものである。sharpな因果帰無仮説の棄却は、因果リスクが2群間で等しいというweakな因果帰無仮説の棄却を意味しない。よって、sharpな因果帰無仮説では平均因果効果が0ではないことを示せない。

本研究では、weakな因果帰無仮説に対する統計的仮説検定手法を議論した。この手法を、Fisherの正確検定の自然な拡張として提示したのをはじめ、層別解析、アウトカムが順序カテゴリカル変数の場合、ベイズ流アプローチへの拡張などを行った。

研究成果の学術的意義や社会的意義

本研究で提示した方法を用いることにより、ランダム化試験においてweakな因果帰無仮説に対する統計的仮説検定を、正確な方法として行うことができる。また、この検定方法に対応する信頼区間も構成することができる。これらにより、平均因果効果の推測をより正確に行うことができる。

将来的には、薬剤評価のための臨床試験などに、本研究で提示した方法あるいは類似の改良型が用いられることになる可能性がある。そうなれば、臨床試験における薬剤評価の方法が変わることになる。

研究成果の概要（英文）：In randomized trials in which two treatment groups are compared with a binary outcome, exact tests are statistical hypothesis tests for the sharp causal null hypothesis that there is no treatment effect for all individuals. The rejection of sharp causal null hypothesis does not mean that the weak causal null hypothesis, which is a hypothesis that causal risks are equal between two treatment groups, is rejected. Therefore, we cannot prove that the average treatment effect is not zero by the sharp causal null hypothesis.

In this research, I discussed methods of statistical hypothesis tests for the weak causal null hypothesis. I proposed such a method as a natural extension of Fisher's exact test. Furthermore, I extended the method to several situations such as stratified analysis, the case of an ordinal outcome, and Bayesian approach.

研究分野：生物統計学

キーワード：因果推論

様式 C-19、F-19-1、Z-19、CK-19（共通）

1. 研究開始当初の背景

2 値アウトカムを比較する 2 群間ランダム化試験においては、Fisher の正確検定 [1] が主要な統計的仮説検定手法の 1 つである。しかし、Fisher の正確検定で検定しているのは、治療の因果効果が全員について等しいという sharp な因果帰無仮説のためのものであって、因果リスクが 2 群間で等しいという weak な因果帰無仮説のためのものではない [2]。「sharp な因果帰無仮説が成立すれば weak な因果帰無仮説も成立する」という命題は正しいが、その裏は一般に正しくない。したがって、Fisher の正確検定で帰無仮説が棄却されても、それは平均因果効果が 0 ではないことを意味しない。よって、weak な因果帰無仮説に対する統計的仮説検定を用いるべきなのであるが、そのような手法は開発されていない。

同様に、weak な因果帰無仮説のための検定手法に対応する信頼区間の構成法を提示されていない。つまり、平均因果効果のための信頼区間の構成法は提示されていない。

2. 研究の目的

(1) weak な因果帰無仮説のための正確な統計的仮説検定手法とこれに対応する信頼区間の構成法を開発する。

本研究の主題であると同時に根幹をなす部分である。

(2) 既存の代表的な 2 つの正確検定手法（Fisher の正確検定 [1] と Barnard の正確検定 [3]）が weak な因果帰無仮説のための正確な統計的仮説検定手法ではないことを示す。

(3) 研究の目的 (1) で開発した手法を、サンプルサイズ設計法、層別ランダム化、アウトカムが順序カテゴリカル変数の場合、ベイズ流アプローチ、予後因子・予測因子の検討に拡張する。

3. 研究の方法

(1) 研究の目的 (1) について

主要なアイデアは、潜在的結果変数のペアである response type [4] を用いて統計的仮説検定手法とこれに対応する信頼区間を考えることである。2 値アウトカムを比較する 2 群間ランダム化試験においては 4 つの response type が存在する。各 response type に属する対象者の人数を未知パラメータ（攪乱パラメータ）として扱うことになる。

(2) 研究の目的 (2) について

Barnard の正確検定 [3] に基づく信頼区間は、因果効果の推定値の正確な信頼区間にはなっていないと予想される。このことを、これまで研究してきた nonparametric bounds（因果効果の仮定によらない最狭存在範囲）を利用することで示す。

Fisher の正確検定 [1] が sharp な因果帰無仮説のためのものであって weak な因果帰無仮説のためのものではないことはすでに知られている [2] が、これを研究の目的 (1) で提示する方法の枠組みで議論する。

(3) 研究の目的 (3) について

① サンプルサイズ設計法

単群試験における二項分布に基づく正確な方法と同様に、サンプルサイズを計算するための方法よりもむしろ検出力を計算するための方法を検討する。

② 層別ランダム化への拡張

研究の目的 (1) で提示する単純ランダム化の下での方法を層別したサブグループごとに適用し、それを標準化等の方法を用いて結合する。

③ アウトカムが順序カテゴリカル変数の場合への拡張

効果の指標として、Agresti [5] が定義した relative treatment effect の response type バージョンを考える。

④ ベイズ流アプローチへの拡張

帰無値を与える代わりに、事前分布を想定する。

⑤ 予後因子・予測因子の検討

予後因子か予測因子かを調べたい 2 値の因子のサブグループ間の各 response type に属する人数の割合の違いに着目する。

4. 研究成果

(1) 研究の目的 (1) について

Weak な因果帰無仮説のための正確な統計的仮説検定手法を、多変量超幾何分布に基づく方法として提示した（雑誌論文①）。ランダム化に基づく方法に分類される。併せて、この検定手法に対応する正確な信頼区間の構成、非劣性試験への拡張も容易に行えることも示した。提示した方法では、信頼区間が必ず nonparametric bounds の範囲内に存在するという利点がある。提案した方法は、統一的なアプローチとして実際のランダム化研究に適用可能である。

雑誌論文①の投稿中に、同じくランダム化に基づく方法として、競合する正確な信頼区間の構成法 [6] が提案されたことを知った。そこで、この方法と本研究で提示した方法の比

較を、実データに基づいて行った。その結果、競合する方法[6]が過度に保守的になることがあることを例示することができた（雑誌論文④）。

(2) 研究の目的(2)について

Barnard の正確検定に対応する信頼区間である Santner-Snell 信頼区間が因果効果の信頼区間とはなり得ないことを、実データを用いて例示した（雑誌論文②）。

4 つのうちの特定の 2 つの response type に属する人数が 0 人であるという特殊な状況の下では、因果帰無仮説は sharp な因果帰無仮説となる。この特殊な状況の下では、研究の目的(1)で提示した方法が Fisher の正確検定に一致することを示した（雑誌論文①）。逆に言うと、研究の目的(1)で提示した方法は、Fisher の正確検定の weak な因果帰無仮説への自然な拡張とみなすことができる。

(3) 研究の目的(3)について

① サンプルサイズ設計法

有意水準、各群イベント発生割合とサンプルサイズを固定しときの検出力を計算するための方法を提示した（雑誌論文⑤）。原理（考え方）自体はシンプルで難しくないが、現実的に適用不可能なほど計算に時間がかかるという難点がある。

② 層別ランダム化への拡張

研究の目的(1)で提示した単純ランダム化の下での方法を、層別ランダム化の場合に拡張した（雑誌論文⑥）。上記「(a)サンプルサイズ設計法」と同様、原理（考え方）自体はシンプルで難しくないが、現実的に適用不可能なほど計算に時間がかかるという難点がある。計算（時間とメモリ）に限界があるために、sharp な因果帰無仮説に対してであるが、層別しない場合、既存の方法[7]、提案法をシミュレーションで比較した。結果は、提案法の検出力が最も高いことを示した。Weak な因果帰無仮説については、実際例で上記 3 つの方法を比較した。Sharp な因果帰無仮説に対するシミュレーションと同様、提案法の検出力が最も高かった。

③ アウトカムが順序カテゴリカル変数の場合への拡張

効果の指標として、Agresti [5]が定義した relative treatment effect の response type バージョンを考えた。研究の目的(1)で提示した方法の自然な拡張として定式化することができた（雑誌論文⑦）。研究の目的(1)で提示した方法と同様、信頼区間が必ず nonparametric bounds の範囲内に存在するという利点がある。ここでは、正確な統計的仮説検定と信頼区間に加えて、nonparametric bounds を導出するための数値計算法も提示した。

④ ベイズ流アプローチへの拡張

上記「(c)アウトカムが順序カテゴリカル変数の場合への拡張」をベイズ流アプローチの枠組みで議論した（雑誌論文⑨）。これまでに提示してきた一連の方法と同様、事後分布が必ず nonparametric bounds の範囲内に存在するという利点がある。その一方で、各 response type に属する対象者の割合について事前分布を設定しなければならない。この割合は一般に識別不可能であるために、現実的に、無情報事前分布以外の事前分布を設定するのは困難である。

⑤ 予後因子・予測因子の検討

研究の目的(1)で提示した方法を、臨床試験で注目されることがある予後因子（受けた治療にかかわらずこの因子があると予後が良いという因子）と予測因子（この因子があるかないかで治療効果が変わるという因子）の推測法に拡張した（雑誌論文⑩）。特筆すべきは、4 つの response type と 3 つの結果（因子のあるグループとないグループでの差がプラスかマイナスか 0）の 12 の組み合わせから、その因子が予後因子であるか予測因子であるか、あるいはどちらでもないかを判断するという新しい概念を導入したことである。既存の予後因子・予測因子との関係も考察した。既存の予後因子・予測因子を response type で説明することにより、既存の予後因子・予測因子では、特定の条件の下でのみ予後因子と予測因子の推測を適切に行うことができることがわかった。それに対して、提案した方法では特定の条件を必要としないため、柔軟にデータに適用することができる。ただし、アウトカムが 2 値変数のときにしか適用できないという弱点がある。

<引用文献>

1. Fisher RA. *The Design of Experiments*, 8th ed. Oliver and Boyd: Edinburgh, 1966.
2. Greenland S. On the logical justification of conditional tests for two-by-two contingency tables. *American Statistician* 1992; 45: 248–251.
3. Barnard GA. A new test for 2×2 tables. *Nature* 1945; 156: 177.
4. Hernán MA, Robins JM. *Causal Inference*. Chapman and Hall/CRC: Boca Raton, 2019.
5. Agresti A. *Analysis of Ordinal Categorical Data*, 2nd ed. John Wiley and Sons: New Jersey, 2010.
6. Rigdon J, Hudgens M. Randomization inference for treatment effects on a binary outcome. *Statistics in Medicine* 2015; 34: 924–935.

7. Jung S.-H. Stratified Fisher's exact test and its sample size calculation. *Biometrical Journal* 2014; 56: 129-140.

5. 主な発表論文等

〔雑誌論文〕（計 11 件）

- ① Chiba Y. Inference on covariate effect types for treatment effectiveness in a randomized trial with a binary outcome. *Clinical Trials* 2019; 16: 237-245（査読有）
DOI: 10.1177/1740774519828301
- ② Chiba Y. An inverse probability weighting method for estimating the net benefit in survival analyses in observational studies. *Epidemiology, Biostatistics and Public Health* 2018; 15: e12993（査読有）
DOI: 10.2427/12993
- ③ Chiba Y. Bayesian inference of causal effects for an ordinal outcome in randomized trials. *Journal of Causal Inference* 2018; 6: 20170019（査読有）
DOI: 10.1515/jci-2017-0019
- ④ Chiba Y. Causal measures for prognostic and predictive biomarkers. *Open Journal of Statistics (Special Issue on Biostatistics)* 2018; 8: 241-248（査読有）
DOI: 10.4236/ojs.2018.82014
- ⑤ Chiba Y. Sharp nonparametric bounds and randomization inference for treatment effects on an ordinal outcome. *Statistics in Medicine* 2017; 36: 3966-3975（査読有）
DOI: 10.1002/sim.7400
- ⑥ Chiba Y. Stratified exact tests for the weak causal null hypothesis in randomized trials with a binary outcome. *Biometrical Journal* 2017; 59: 986-997（査読有）
DOI: 10.1002/bimj.201600085
- ⑦ Chiba Y. Sample size calculation of exact tests for the weak causal null hypothesis in randomized trials with a binary outcome. *Open Journal of Statistics (Special Issue on Statistical Hypothesis Testing)* 2016; 6: 766-776（査読有）
DOI: 10.4236/ojs.2016.65063
- ⑧ Chiba Y. A note on exact confidence interval for causal effects on a binary outcome in randomized trials. *Statistics in Medicine* 2016; 35: 1739-1741（査読有）
DOI: 10.1002/sim.6826
- ⑨ Chiba Y. Causal inference in randomized trials: a shift from the sharp causal null hypothesis to the weak causal null hypothesis. *Journal of Biometrics and Biostatistics* 2016; 7: 288（査読有）
DOI: 10.4172/2155-6180.1000288
- ⑩ Chiba Y. A note on exact tests and confidence intervals for two-by-two contingency tables in randomized trials. *International Journal of Clinical Research and Trials* 2015; 1: 102（査読有）
<http://dx.doi.org/10.15344/ijcrt/2015/102>
- ⑪ Chiba Y. Exact tests for the weak causal null hypothesis on a binary outcome in randomized trials. *Journal of Biometrics and Biostatistics* 2015; 6: 244（査読有）
DOI: 10.4172/2155-6180.1000244

〔学会発表〕（計 7 件）

- ① Chiba Y. Bayesian inference of causal effects in randomized trials with a binary outcome. The 29th International Biometric Conference, 2018 (Barcelona in Spain)
- ② 千葉康敬. 因果推論で分割表正確検定を考える. リスク解析戦略研究センター研究集会「因果推論の基礎」 2017（東京）
- ③ 千葉康敬. 研究目的とデータ取得方法に応じた統計解析. 第 112 回医学物理学学会 2016（教育講演）（沖縄）
- ④ Chiba Y. Stratified exact test and confidence interval for causal effects on a binary outcome based on principal stratification. The 2016 Joint Statistical Meeting, 2016 (Chicago in U.S.A.)
- ⑤ Chiba Y. Exact tests for the weak causal null hypothesis on a binary outcome in randomized trials. East Asia Regional Biometric Conference 2015（招待セッション）(Fukuoka in Japan)
- ⑥ 千葉康敬. 統計的仮説検定における注意点. 第 110 回医学物理学学会 2015（教育講演）（北海道）
- ⑦ 千葉康敬. 非劣性試験における 2×2 分割表の正確検定. 統計関連学会連合大会 2015（岡山）

〔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：
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：
国内外の別：

〔その他〕
ホームページ等
<http://researchmap.jp/read0069849/>

6. 研究組織

(1) 研究分担者

研究分担者氏名：
ローマ字氏名：
所属研究機関名：
部局名：
職名：
研究者番号（8 桁）：

(2) 研究協力者

研究協力者氏名：
ローマ字氏名：

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。