

抄読会

共変量が一部欠測したデータについて

修士課程1年 鎌田悠太

概要

データの欠測は様々な場面で起こりうるものである。どの変数が欠測したのか、どのような理由で欠測したのかを考慮する必要がある、適切に扱わないことで結果にバイアスが起こる可能性がある。欠測が起こる変数として、アウトカムと共変量に分けることができる。アウトカムの欠測については、近年、医学系研究の分野において注目度が上がっており、その取扱い方には十分な考慮が必要であることが認知されている。一方で、共変量の欠測については、介入を行う臨床試験などではなく、介入を行わないような観察研究やデータベース研究において起こることが多い。曝露因子とアウトカムさえわかれば単純な解析を行うことは可能だが、多くの研究において共変量、特に交絡因子となりうる共変量に関して調整を行ったうえで、解析が行われる。しかし、その交絡因子が欠測している場合、交絡の調整は行えない。それにより、結果としてバイアスや、推定精度・検出力の低下の可能性が考えられる。このようにアウトカムの欠測同様に共変量の欠測も重要である。そのような背景から、私の研究テーマとして「共変量の一部欠測したデータ」を考えている。

今回、共変量の欠測に注目して、どのような研究が行われているか、特にデザインによる欠測に注目した論文¹⁾²⁾について紹介し、今後の研究について発表した。

参考文献

- 1) Keogh RH, White IR. Using full-cohort data in nested case-control and case-cohort studies by multiple imputation. *Stat Med.* 2013;32(23):4021-4043.
doi:10.1002/sim.5818
- 2) Keogh RH, Seaman SR, Bartlett JW, Wood AM. Multiple imputation of missing data in nested case-control and case-cohort studies. *Biometrics.* 2018.
doi:10.1111/biom.12910