

ライフコース疫学研究における統計的因果探索のためのアルゴリズム

ここ数十年で、出生前・幼少期といった人生の早い時期での生物学的・社会的要因が、成人後の多くの疾病の病因として重要な役割を果たすという認識が広まった。このような研究分野を、ライフコース疫学と呼ぶ。Kuh et al(2003)は、ライフコース疫学を以下のように定義している。ライフコース疫学とは「胎児期, 幼少期, 思春期, 青年期, その後の成人期での物理的, 社会的曝露による成人疾病リスクへの長期的な影響を研究する分野」としている。[1]

生物学的, 社会的変数の複雑な相互関係を経時的に研究するためには, ライフコースにわたる縦断的な情報が必要となる。その上で, あらかじめ変数間の関係性を Directed Acyclic Graph (DAG) を用いて特定し, ジョイントモデルや構造方程式モデルなどによりその関連性を定量的に評価することが多い。一方で, ライフコースという非常に複雑な関係性を扱う以上, 変数間の関係性を事前に特定することにはしばしば困難が生じる。そのため, 変数を絞り込み仮説を単純化しすぎずに, 候補となる変数間の関係を特定のアルゴリズムの下で探索し, 変数間の複雑な因果関係に関する新たな知見を得たいというニーズは多い。このような統計的因果探索は, 因果推論における重要な1分野であり, これまでも多くの研究がある。

本抄読会では, ライフコース疫学における統計的因果探索のためのアルゴリズムとして Petersen らにより発表された論文を紹介する。[2]このアルゴリズムは, 既存の因果グラフ探索アルゴリズムを経時データに拡張したものとなっている。本論文をきっかけとして, 統計的因果探索に関する理解を深め, 修士論文につなげていきたいと考えている。

参考文献

- [1]Kuh D, Ben-Shlomo Y, Lynch J, et al. Life course epidemiology. Journal of epidemiology and community health 2003;57.10:778-83.
- [2]Petersen AH, Osler M, Ekstrom CT. Data-driven model building for life-course epidemiology. Am J Epidemiol 2021;190.9:1898-907.