IT・エレクトロニクスステーション－９

「因果グラフと因果効果」

総合情報学部　情報科学科　山本英二研究室

Keywords: 因果グラフ，推定可能性，リスク，統計モデル

「研究目的・応用の可能性」 喫煙による肺ガンリスクの上昇等の議論におけるデータに基づく統計的因果推論には，背景の知識を用いてデータから因果モデルを構築していく立場と，前提となる因果モデルの下で関心のある因果指標をデータから推定する立場がある．近年，後者の立場に立ち，観察研究で得られたデータから因果指標を推定出来るか，どう推定するか，の研究が発展してきている．因果モデルを前提として，因果効果の推定に関心がある．この因果効果が推定可能となるためには，因果モデルがある基準を満たせば良い．

例として４つの２値変量 X:喫煙 Y:肺ガン U:遺伝子型 Z:肺中タール とし, 因果ダイアグラムは次のグラフで示されるとする．

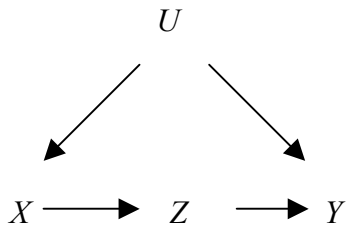

図１　因果ダイアグラム

見かけの喫煙・肺ガンの関連が遺伝子型を示す体質によって説明できる可能性を持つというモデルである．各変量間の矢線が（確率的）因果関係である．因果効果を対象集団に介入により暴露させたときの反応と暴露させないときの反応の差，比を因果効果と定義する．実際には同一集団に暴露と非暴露の状況はあり得ないので，反事実モデルといわれる．観察集団全体に介入して強制的に $X = x$ とすることは因果ダイアグラムで X へ入る矢線を落とすことに相当する．因果効果は因果ダイアグラムがバックドア基準，フロントドア基準を持つとき，(X,Y,U) 又は (X,Y,Z) の周辺分布，又は条件付き分布を与えることで同定可能となることが知られている．

因果構造を因果グラフとして表現することで，視覚的に構造が掴め，求めたい因果効果の推定を行うための条件が理論的に与えられる．医学，教育，社会調査等の統計科学，また人工知能などの情報科学の分野で使われてきている方法であるが，マーケティングなどの商業分野での活用が考えられる．

連絡先 TEL&FAX:086-256-9656　Mail:eiji@mis.ous.ac.jp　URL:http://zeus.mis.ous.ac.jp