

Journal of the Korean Data & Information Science Society 掲載論文

「媒介の割合に基づく二値アウトカムでの因果媒介モデルにおける簡単な検出力解析法」

Young Min Kim, John B. Cologne, Harry M. Cullings

“Simple power analysis in causal mediation models for a dichotomous outcome based on the mediation proportion”

Journal of the Korean Data & Information Science Society 2017 (May); 28(3):669–84

(doi: 10.7465/jkdi.2017.28.3.669)

今回の調査で明らかになったこと

複雑な因果モデルについて調べるための統計的方法論の研究が急速に進歩しているが、二値（二項）アウトカム^{*}に対する放射線リスクの媒介を検出するための調査デザインの方法論は、一層の開発を要する分野である。我々は、連続媒介変数を伴う疾患有病率に関する調査デザインに、モンテカルロ・シミュレーション法^{***}とブートストラップ法^{****}を用い、コンピューターを利用する方法で検出力および必要対象者数を算出した。また、シミュレーション調査により、この方法の実用性および HbA1c の血清レベルを潜在的な（すなわち、放射線被曝との因果関係が仮定される）媒介変数とする心疾患の放射線リスクへの適用について説明した。

^{*}二値（二項）アウトカム：「はい」または「いいえ」（例えば、がんなど調査対象の事象が対象者に発生したか否か）など二つの値しか持たないアウトカム。

^{***}モンテカルロ・シミュレーション法：ランダムな過程を示す統計学用語。モンテカルロ・シミュレーションでは、ある方法の統計的特性を調べるため、コンピューター・プログラムを用いランダムに発生させたデータセットによるシミュレーションを行う。この場合「ランダム」とは、観察データに生じる自然変動のことをいい、データを定義するパラメータではない（これらのパラメータとは対象集団における平均被曝線量や放射線影響などであり、ランダムではなく固定である）。

^{****}ブートストラップ法：理論上の情報源に関する仮定（統計的分布など）を追加せず、観察されたデータ値にのみに基づく一連の統計的手順。

解説

従来の疫学調査は標準的な回帰分析（線形回帰^{*}、ロジスティック回帰^{**}または Cox 回帰^{***}）に依存しており、推定される原因（暴露などの説明変数）と効果（がんなどのアウトカム）の間の関連性を分析・推定するに過ぎない。説明変数がひとつ以上の多変量回帰解析では、説明変数の間に相関性があるかもしれず、因果関係上の関連性を特定することはできないことが、実際にはよくある。言い換えれば、従来の回帰解析では調査対象である変数間の複雑な関係をしばしば処理しないままである。

最近、この推論上の「ブラックボックス」を開封する必要性が生じ、説明変数間の

回帰的関連性など複数の回帰的関連性を組み入れることができる複雑な因果モデルが広く利用されるようになった。暴露 X はアウトカム Y および別の説明変数 Z の両方の原因であるという状況は特に興味深い。この場合、X の Y に対する効果の一部（またはおそらく全部）は経路 $X \rightarrow Z \rightarrow Y$ で説明できる可能性があり、Z は「媒介変数」と呼ばれる。この場合、X が及ぼす全効果のうち Z が媒介する部分の占める割合、すなわち「媒介の割合」にしばしば興味注がれる。

このようなモデリングの基礎はいわゆる「パス解析」にあるが、パス解析では標準的な線形回帰モデルを連続説明変数および連続アウトカム変数に利用する。しかし、アウトカム変数が連続変数でない場合、例えば Y が疾患の有病率を示す二値（1 か 0 か、または「はい」か「いいえ」か、のような二項）の変数である場合、単純なパス解析方法では媒介割合の有効な推定値は得られない。このようは調査ではしばしば、X および Z が Y に及ぼす効果を示すモデルとしてロジスティック回帰を用いる。このようなデータを用いて媒介割合を推定し、適切な検出力を得るための調査のデザインの仕方を理解するためには、さらに理論的研究が必要である。

※線形回帰：統計上の回帰関係を最も単純に表すもので、アウトカム変数（血圧値など、連続変数で多数の数値を持つ）が予測変数（放射線量など）と 1 対 1 で直線（線形）の対応関係を示すと推定される。

***ロジスティック回帰：二値（二項）アウトカムに適した回帰法。予測変数とアウトカム変数の間に 1 対 1 の対応はできないが、予測変数とアウトカムが発生する確率の間に 1 対 1 の対応を割り当てることができる。

****Cox 回帰：がん発症年齢などの時間経過に従って事象が発生するアウトカムに適した回帰法。最終的なアウトカムは対象者ががんが発生したか否かなど二値であるが、解析に使用されるアウトカム変数は、事象が発生したという事実だけでなく事象の発生ままでに要した時間などの情報を含むため、より有益な情報である。

1. 調査の目的

今回の目的は、アウトカムが二値の場合の媒介に関わる複雑な因果モデルについて、検出力および必要対象者数の算出が可能となる方法論を展開することである。

2. 調査の方法

我々は、コンピューターを駆使したモンテカルロ・シミュレーション法とブートストラップ法を適用して検出力および必要対象者数を算出し、その手順をシミュレーションにより評価した。

3. 調査の結果

我々は順調に手順を実施し、実用性を証明した。また原爆被爆者から成る放影研成人健康調査（AHS）において、血清の HbA1c レベルの媒介する放射線リスクの割合が 0.33 と考えられる場合には、これが関与する心疾患の放射線による有病率を検出力 80% 以上で実施できるであろうことを示した。

今回の調査の意義

放射線被曝と疾患の関連性の機序をより正確に解明することは重要である。放射線影響に関する科学的理解を深めることへの貢献に加え、媒介変数について介入が可能な場合には、本調査により放射線被曝後の長期的治療に関する臨床上有益な手掛かりが得られるかもしれない。したがって、媒介を組み込んだ複雑な因果モデルを用いて放射線リスクを評価できるような方法の開発および実施は極めて重要である。今回の研究により、放射線リスクに関する放影研調査においてこのような因果モデリングを行う能力が高まった。

放射線影響研究所は、広島・長崎の原爆被爆者および被爆二世を約 70 年にわたり調査してきた。その研究成果は、国連原子放射線影響科学委員会（UNSCEAR）の放射線リスク評価や国際放射線防護委員会（ICRP）の放射線防護基準に関する勧告の主要な科学的根拠とされている。被爆者および被爆二世の調査協力に深甚なる謝意を表明する。

§ *Journal of the Korean Data & Information Science Society* 誌は、韓国データ情報科学学会（Korean Data & Information Science Society）の公式学会誌で、統計学、数学および関連するデータ解析技術が用いられるその他関連分野における原著研究論文や短報、総説、症例報告、ならびにその他招待論文を掲載している。（2015 年のインパクト・ファクター：0.808）