

マッチング、傾向スコアマッチング、**matching weight**

東京大学 大学院医学系研究科 生物統計学分野

篠崎 智大

マッチングは共変量調整を前提とするデータ解析の有用な前処理である一方で、実のところその働き方は難解である。例えば、データ生成過程において関連のある変数同士（曝露と共変量など）を標本において必ず無相関にできる性質は、直感的には「妥当な」比較を可能にするように思えるが、この正当性を簡便な因果推論ツールの一つ、因果ダイアグラムで説明するのは骨の折れる仕事である。また、マッチされた層を解析で考慮（するのが常套手段だが、必ずしも）しなければいけないかどうかについて、サンプリングデザインを始め、仮定するモデルや変数の組み合わせによって微妙な議論が引き出される。マッチングの単純でない標本サンプリングは標的集団の周辺指標の推測を難しくする。それでも、共変量そのものによる古典的なマッチングは統計モデルへの依存度（**model dependence**）を低下させる点で擁護できる手法である；共変量バランスのとれた標本においては、そうでない標本に比べて推定値のモデル依存度が小さくなる。

残念ながら、傾向スコアによるマッチングにはこの擁護論が適用できない。特に、**King and Nielsen (2016)** は、傾向スコアマッチング基準を強める（よりタイトに標本を選択する）ことで共変量バランスがかえって容易に悪化し得る点を指摘し、このバイアス・分散トレードオフの欠落を傾向スコアパラドクスと呼んでいる。さらに、共変量マッチングと異なり、傾向スコアマッチング後の解析では、マッチングを考慮してもしなくても推定量の漸近分布の複雑さを正しく評価できない。しかし、本稿に紹介する **matching weight** によって、**Li and Greene (2013)** は傾向スコアマッチングによる単純な推定量の **estimand** を重み付け解析、すなわち基本的なセミパラメトリック推測の枠組みによって推定できることを示し、エレガントに後者の問題を解消した。傾向スコアマッチング後の単純比較または共変量調整がパラドクスの影響下にあるのに対し、標本全体を解析に寄与させる **matching weight** はそれ自体パラドクスとは関係なく、さらにセミパラメトリック有効な推定量は二重頑健性の意味でモデル誤特定へのプロテクションをかけることもできる。重み付け解析では一般的にネックとなる推定効率についても **matching weight** は優れた特性を持つことを、特に **g**-推定量との関係性から議論しよう。

文献

King, G. and Nielsen, R. (2016). *Why propensity scores should not be used for matching*. Copy at <https://j.mp/1sexgVw> Export BibTex Tagged XML Download Paper, 481.

Li, L. and Greene, T. (2013). A weighting analogue to pair matching in propensity score analysis. *International Journal of Biostatistics*, **9**, 215–234.