

Note on Causal Inference, RCT, Natural Experiment, and Matching

Lê Việt Phú

Chúng ta đã học mục đích của đánh giá chính sách là thiết lập quan hệ nhân quả giữa một can thiệp xảy ra và kết quả tiềm năng. Chúng ta chỉ kết luận được về quan hệ nhân quả khi thiết lập được phản chứng (counterfactual) của nhóm đối tượng hưởng lợi. Nếu không xây dựng được phản chứng thì không thể kết luận được về quan hệ nhân quả. Khi này, các khuyến nghị chỉ có thể dừng lại ở mối quan hệ tương quan.

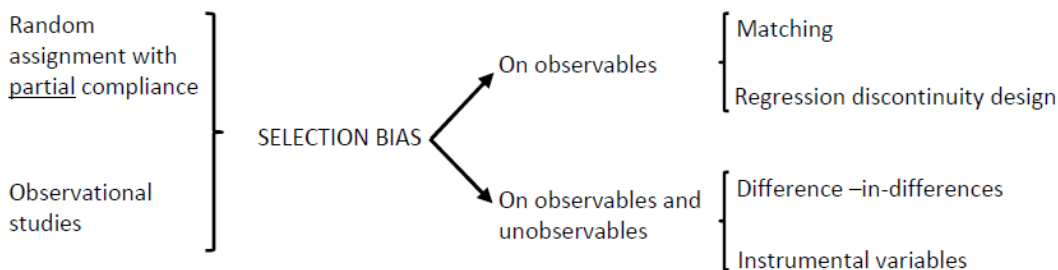
Để xây dựng phản chứng thì học viên đã được làm quen với khung đánh giá kết quả tiềm năng (potential outcome framework) và thiết lập mối quan hệ giữa tác động can thiệp trung bình (ATE) với tác động can thiệp lên đối tượng hưởng lợi (ATT).

$$\begin{aligned} ATE &= \underbrace{E(Y|T=1)}_{\text{Expected income with participation}} - \underbrace{E(Y|T=0)}_{\text{Expected income without participation}} \\ &= \underbrace{E(Y_i^1|T=1) - E(Y_i^0|T=1)}_{ATT} + \underbrace{E(Y_i^0|T=1) - E(Y_i^0|T=0)}_{Bias} \end{aligned}$$

Thiết kế thử nghiệm ngẫu nhiên có kiểm soát triệt tiêu *Bias* bằng cách kiểm soát đặc tính đầu vào của mẫu được đưa vào chương trình thử nghiệm sao cho hai nhóm hưởng lợi và đối chứng hoàn toàn tương đồng, sau đó ngẫu nhiên hóa quá trình can thiệp $Y_T, Y_C \perp T$ để loại trừ vấn đề tự lựa chọn vào nhóm hưởng lợi. Với cỡ mẫu đủ lớn thì khác biệt về kết quả trung bình của nhóm không tham gia và *nhóm nếu như không tham gia nhưng trên thực tế có tham gia (phản chứng)* là tương đồng, do đó Bias bằng không, và tác động của chương trình can thiệp ATT bằng ATE với RCT.

Khi không thể nghiên cứu bằng RCT, chúng ta phải sử dụng dữ liệu điều tra hay quan sát được (observational studies) nhằm thiết lập quan hệ nhân quả. Do dữ liệu này không đảm bảo nhóm hưởng lợi hay nhóm đối chứng hoàn toàn tương đồng, nên Bias luôn luôn tồn tại. Việc đánh giá nguyên nhân gây ra Bias sẽ giúp tìm được phương pháp đánh giá phù hợp.

Random assignment with average $\longrightarrow$ NO SELECTION BIAS $\longrightarrow$ ATT=(Average effect on treated) – (Average effect on non-treated)



Thử nghiệm tự nhiên (natural experiment-NE) là tình huống cho phép đánh giá tác động bằng cách mô phỏng lại một thử nghiệm ngẫu nhiên, từ đó cho phép xây dựng phản chứng hợp lệ (valid counterfactual)¹. Thử nghiệm tự nhiên là tình huống khi một can thiệp xảy ra theo đó có sự phân định ngẫu nhiên của nhóm hưởng lợi và nhóm đối chứng. Tuy nhiên, do không kiểm soát đặc tính mẫu trước khi xảy ra can thiệp nên không đảm bảo hai nhóm tương đồng về mọi mặt. Sự giống nhau giữa RCT và NE là quá trình can thiệp là ngẫu nhiên, do đó không xảy ra quá trình tự lựa chọn mẫu. NE thì không đảm bảo hai nhóm tương đồng như RCT. Để đánh giá tác động can thiệp với tình huống thử nghiệm tự nhiên, chúng ta không thể dùng ATE để tính ATT do tồn tại Bias. Có nhiều cách thức để tính ATT, tùy thuộc vào nguyên nhân gây ra Bias.

Nếu nguyên nhân gây ra *Bias* quan sát được thì chúng ta có tình huống Selection on Observables (SOO). Khi này, chúng ta chỉ cần ước lượng một mô hình hồi quy đa biến có kiểm soát tất cả các yếu tố X có ảnh hưởng đến *Bias*. Do chính sách xảy ra ngẫu nhiên nên T không tương quan với phần dư u và ước lượng của tham số β không chệch và nhất quán. Phương pháp này được gọi là phương pháp điều chỉnh hồi quy (regression adjustment - RA) hoặc phương pháp hàm kiểm soát (control function approach - CFA).

$$Y_i = \beta_0 + \beta * T_i + \delta X_i + u_i$$

Matching là một phương pháp khác của tình huống SOO. Với giả định quan sát được tất cả các nhân tố ảnh hưởng đến xác suất tham gia, chúng ta dùng xác suất tham gia làm tiêu chí để tìm kiếm phản chứng bằng cách ghép các quan sát có xác suất tham gia giống nhau với nhau. Việc ghép cặp dựa trên điểm xu hướng dễ hơn nhiều so với ghép cặp dựa trên tất cả các thuộc tính quan sát được. Đặc biệt với dữ liệu có nhiều biến giải thích, rất khó có thể đảm bảo tất cả các thuộc tính giữa hai nhóm đều tương đồng. Matching là một phương pháp phi tham số (non-parametric) - có nghĩa là nó không dựa vào các tham số để giải thích tác động của can thiệp. Phương pháp phi tham số có lợi thế với phương pháp tham số là nó không phụ thuộc vào một cấu trúc hàm cụ thể nào.

Một đặc điểm lưu ý khi sử dụng NE để đánh giá tác động chính sách là hiệu lực của kết quả. Hiệu lực nội tại của RCT là tốt nhất (tiêu chuẩn vàng trong đánh giá chính sách), và $ATE=ATT$. Tuy nhiên, hiệu lực ngoại vi của RCT thường yếu do chương trình đánh giá được áp dụng vào nhóm mẫu đã lựa chọn kỹ càng, khó ngoại suy từ mẫu nghiên cứu ra quần thể. NE có thể có hiệu lực nội tại tốt hoặc kém, tùy thuộc vào việc ghép được bao nhiêu quan sát trong mẫu nghiên cứu. Hiệu lực ngoại vi của NE tùy thuộc vào tính chất của can thiệp và khả năng ngoại suy từ mẫu ra quần thể. Nếu can thiệp xảy ra ở quy mô rộng và hiệu lực nội tại tốt thì có khả năng nghiên cứu có hiệu lực ngoại vi tốt.

Matching và regression adjustment thực tế rất giống nhau. Chúng đều cần giả định quan sát được các thuộc tính ảnh hưởng đến quá trình tham gia, gọi là giả định độc lập có điều kiện (conditional independence/unconfoundedness assumption) $Y_T, Y_C \perp T | X_i$ với hàm ý rằng trạng thái tham gia chương trình là ngẫu nhiên sau khi đã kiểm soát được tất cả các nhân tố quan sát được trong mô

¹ Sẽ có nhiều tình huống nghiên cứu mà phản chứng xây dựng được là không hợp lệ. Phản chứng không hợp lệ dẫn đến hậu quả tác động ước lượng được bị chệch. Ví dụ đơn giản nhất của phản chứng không hợp lệ là so sánh trước – sau (pre – post analysis) hay so sánh có – không (with and without analysis).

hình. Giả định này cũng hàm ý rằng mô hình không gặp phải vấn đề thiếu biến quan trọng (Omitted Variables Bias). Giả định này khá mạnh và khó có thể đảm bảo trong hầu hết các tình huống. Tuy nhiên, về mặt phương pháp thì RA hoặc Matching đều khá rõ ràng, dễ hiểu, dễ giải thích.

Những thiết kế phức tạp hơn như hồi quy với biến công cụ hay hồi quy gián đoạn nhiều fuzzy regression discontinuity design rất khó giải thích. Tuy nhiên, những phương pháp phức tạp cho phép ước lượng các bài toán với các giả định gần với thực tế hơn. Các tình huống có nguyên nhân của *Bias* không quan sát được (Selection on Unobservables) thường nằm trong nhóm này. Tất nhiên, nếu có dữ liệu quan sát lặp thì có thể dùng một phương pháp khá đơn giản là DiD với đề khử nhân tố không quan sát được không thay đổi theo thời gian.