

実験・準実験 (1)

別所俊一郎

2006年7月12日

Today's attraction

- 実験・準実験とは
- 「理想的な」実験とは
- 実際の実験に付きまとう問題点

実験による因果関係の推定

- 心理学・医学・薬学分野では実験はしばしば因果関係の推定に用いられる
 - － 治験：偽薬（placebo）を用いたりする
- 経済学分野で実験はまだ多くはないものの...
 1. 概念的なレベルで，理想的な実験は実際の因果関係の推定のベンチマークとなりうる
 2. 実際に実験が行われると，その結果は影響力を持つ：実験の強みや弱み，限界を知っておくべき
 3. 外的な環境変化が実験のような状況を作り出すことがある
 - － 自然実験（natural experiment）
 - － 準実験（quasi-experiment）
 - － 「理想的な」実験についての治験を応用できる

実験の評価

- 「政策」の効果の推定
 - 実験群 (treatment group) と対照群 (control group)
- 用いる手法
 - OLS , パネル分析 , 操作変数法を (準) 実験の分析にどう応用するか?
- プログラム評価 (Program Evaluation) への応用
 - 政策・プログラム・介入などの効果の推定
 - 職業訓練 , 奨学金などなど

「理想的な」実験と因果関係の推定

- 「理想的な」実験
 - 無作為割り当て実験：randomized controlled experiment
 - 標本を母集団から無作為抽出
 - 実験群と対照群に無作為に割り当て
 - 結果の差が「政策」の効果
- 理想的な実験の基本的な考え方
 - 「政策」以外には全く同じ2人を連れてきて結果を比べる，
というわけではない．そのような2人は存在しない
 - 無作為に対象を選んで無作為に「政策」を割り当てることで
因果関係が計測できる

無作為割り当てと因果関係

- 「政策」の割り当てが無作為であれば、「政策」の水準は結果に影響する他の要因から独立して分布
 - Missing variable bias が存在しない
 - (例) 職業訓練プログラムを無作為に割り当てたとすれば、就職する確率（結果）に影響する他の要因の分布は、プログラムを受けた人と受けなかった人で同一
- 単回帰モデルで言い直すと...

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

- X_i が無作為に割り当てられていれば、 X_i と u_i の分布は独立
- X_i の条件付きの u_i の分布は X_i に依存せず、 $E[u_i|X_i] = 0$
- OLS の一致性の条件の第 1 が満たされる
- X_i は実験群で政策の水準が全て等しければダミー変数、異なれば線形制約がかかっていることになる

因果効果 Causal effect/Treatment effect

理想的な無作為割当て実験での「政策」 X の結果 Y への因果効果 (Causal effect/Treatment effect) は

$$E[Y|X = x] - E[Y|X = 0]$$

- 条件付き期待値の差
- 実験群と対照群の期待値の差
- Y と X に線形関係を仮定したときには, β_1 は「政策」の水準 X が 1 単位変化したときの因果効果の大きさを表す

Difference Estimator

- 因果効果は期待値の差だから（同時）分布の未知の特性値の1つ
 - X_i がダミー変数で無作為割り当ての実験なら，実験群と対照群の標本平均の差で推定可能
 - X_i が連続変数なら OLS で推定可能： $E[u_i|X_i] = 0$ だから $\hat{\beta}_1$ は一貫性を持つ
 - これらが Difference Estimator
- 実際の実験では $\hat{\beta}_1$ が一貫性を持つとは限らない
 - $\text{corr}(u_i, X_i) \neq 0$ となるような問題が発生

実際の実験でおきうる問題点

- 内的妥当性 internal validity
 - 因果効果の統計的推測が，研究している母集団について妥当
 - 無作為の失敗・割当て手続きの失敗・欠落・実験効果・小標本
 - $\text{corr}(u_i, X_i) \neq 0$ となり，difference estimator (OLS 推定量) が不偏性や一致性を持たない
- 外的妥当性 external validity
 - 統計的推測や結論が，他の母集団へ一般化可能
 - 実験群の代表性・実験の代表性・一般均衡効果・自発的参加

Failure to randomize

- 無作為の割当ての失敗
 - 「政策」の割当てが主体の特性や選好に依存
 - (自発的な「政策」への参加)
 - 実験群と対照群の結果の差は、政策の効果と、無作為でない割当ての効果の両方を反映
 - 「政策」 X_i と誤差項 u_i に相関が発生：割当てを受けるかどうかは誤差項に依存
- (例) 職業訓練
 - 割当てを苗字のアルファベットの前半後半で決めたとすると...
 - 苗字には民族的な傾向があるから、割当てが民族について無作為でなくなる
 - 就業経験・教育などの労働市場での特性が民族について系統的に異なれば、omitted variable が実験群と対照群で系統的に差がつくかもしれない

Failure to follow treatment protocol

- 人々は言われたことを行うとは限らない
 - 「政策」の割当てが無作為であっても、「政策」を実際に受けるかどうかは無作為でないかもしれない
 - 主体が実際に受ける「政策」は、部分的には無作為に、部分的には主体の特性に依存するかも
 - 職業訓練：実験群の人は訓練を受けずに休むかもしれない / 対照群の人はどうにかして訓練を受けてしまうかもしれない
- 手続きの部分的な遵守 (partial compliance)
 - 実験する側が実際に行われている「政策」を記録できるとき：「政策」 X_i と誤差項 u_i に相関が発生．Self-select が起き，OLS 推定量にバイアス
 - 実験する側が実際に行われている「政策」を観測できないとき（服薬コンプライアンスなど）：測定誤差が発生し，OLS 推定量にバイアス

Attrition, Small sample size

Attrition

- 無作為に割当てを行った後に主体が実験から脱落すること
 - － 「政策」と関係のない脱落：病欠など
 - － 「政策」と相関を持つ脱落：実験群と対照群で誤差項の分布が異なるため、「政策」 X_i と誤差項 u_i に相関が発生．一種の selection bias.
 - － (例) 職業訓練：優秀な参加者はすばやく職を見つけ，プログラムが終わるころには優秀でない参加者のみが残る

Small sample size

- 大規模な実験は費用がかかる
- 小標本そのものはバイアスをもたらさないが，推定値が不正確になる（SE が大きい，近似が悪い）

Experimental effects / Hawthorne effects

- 実験に参加しているというだけで行動が変わる
 - 実験に参加しているという興奮や高揚感
- “double blind”（二重盲検法）でホーソン効果は緩和されるかも
 - 被験者も実験者も，その被験者が対照群にいるのか実験群にいるのかを知らない
 - 薬の治験における偽薬（プラセボ）の利用
 - 「プラセボ効果」とはベツモノ
- “double blind”は経済分野では難しい
 - 「政策」の可否が実験者の報酬と関係する場合にはホーソン効果が大きく出る？
 - ホーソン効果の大きさは実験の細部に依存

標本の代表性・「政策」の代表性

- 標本の代表性
 - Population studied と population of interest は似ているか?
 - * 服役囚に対する職業訓練の効果は一般化できるか?
 - 実験への参加が自発的：実験の効果が大きく出やすい
 - 標本選択が無作為でないときには容易に一般化できない
- 政策やプログラムの代表性
 - Program studied と program of interest は似ているか?
 - 小規模でよく管理された実験の質を，大規模にしたときに保てるか?
 - 短期間の実験の効果を長期間に一般化できるか?
 - 質を保つに十分な資金はあるか?

一般均衡効果，資格効果

一般均衡効果

- 小規模で一時的な「実験」が，大規模で恒久的な「政策」になると，経済的な状況・市場状況・政策環境が変化
 - － (例) 職業訓練：雇用者による補完的なプログラムの衰退
 - － (例) 少人数学級：平均的な教師の質の低下

Treatment v.s. eligibility effect

- プログラムの参加が自発的（自分で決められる）状況では，“やる気”のある人が参加してしまう
- 実際の条件を満たす人だけに実験群に参加してもらう
 - － 実験群と対照群の差は，参加資格による差を表すことになってしまう