

第 19 章: 実証研究の実施

Jeffrey Wooldridge (2016).

Introductory Econometrics: A Modern Approach
6rd ed. Cengage Learning.

資料作成: 田中鮎夢

2025-04-04

19-1 問題の提起

19-1 問題の提起

成功する実証分析の基礎は、データで答えられる具体的な問題を提起することから始まる。

明確な目標がなければ、データ収集は非効率的になり、重要な変数を見逃したり、不適切なサンプルを選択したりする可能性がある。

問題を考える際にはデータの入手可能性やプロジェクトの期間を考慮する必要があるが、問題の具体性が最も重要である。

研究テーマの選択

研究テーマは様々な経済分野から生まれる：

- ▶ 労働経済学：教育の質、数学/科学の履修、身体的外見に基づく賃金格差
- ▶ 公共財政：税金やサービスが経済活動に与える影響
- ▶ 教育経済学：支出が成績に与える影響、学校選択の影響
- ▶ マクロ経済学：集計時系列間の関係

研究論文を探すためのリソース

- ▶ Google Scholar <https://scholar.google.co.jp/>
- ▶ Journal of Economic Literature (JEL) 分類システム
- ▶ EconLit などのインターネットサービス *今はあまり使われない
- ▶ Social Sciences Citation Index *今はあまり使われない

テーマ選択時の重要な考慮事項

- ▶ 問題は広範な政策的含意ではなく、局所的な含意を持つこともある
- ▶ 標準的なマクロ経済集計データを用いた独創的研究は、短期間のプロジェクトでは困難である場合が多い
- ▶ 教育収益率のような既に研究されている問題でも、新しいデータや斬新な計量経済学的アプローチによって新たな知見が得られる

適切に定式化された研究問題は具体的であるべきだ。

- ▶ 「犯罪に関する論文を書いている」というよりも、「米国の都市における地域警察活動が犯罪率に与える影響を研究している」のように具体的にすべきである。

19-2 文献レビュー

19-2 文献レビュー

実証論文はたとえ短くても、関連文献のレビューを含むべきである。

オンライン検索サービスにより、関連研究を見つけることが容易になった。

検索の際は、キーワード検索では表示されない可能性のある関連トピックも考慮する必要がある。

例えば、アルコールの影響に関する研究は、薬物使用の影響に関する研究に示唆を与えることがある。

文献レビューの組み込み方

研究者によって文献レビューの組み込み方は異なる：

- ▶ 「文献レビュー」と題した独立したセクションとして
- ▶ 序論の一部として
- ▶ 広範なレビューの場合は独自のセクションが必要となる

19-3 データ収集

19-3a 適切なデータセットの決定

データには様々な形態がある

- ▶ クロスセクション、時系列、プールドクロスセクション、パネルデータ。

選択は研究問題と利用可能なリソースに依存することが多い。

個人や家族レベルの問題では、調査による単一のクロスセクションが一般的だが、十分な制御変数があれば因果関係の確立は困難である。

パネルデータ

パネルデータセット（同一単位の複数の観測値を時間経過とともに収集）は観測されない効果をコントロールするのに役立つが、個人に関して入手するのは難しい。

法人単位（学校、都市、州）のパネルデータは、これらの主体が存続する傾向があり、政府機関が定期的にそのような情報を収集するため、より容易に入手できる。

19-3b データの入力と保存

データは印刷形式または電子形式で入手できる。

電子データの場合、テキスト・ファイルが計量経済学ソフトウェアでの使用に最も柔軟性を提供する。

データ編成は重要である：

- ▶ 時系列の場合：最も古い期間を最初にした時系列順
- ▶ プールドクロスセクションの場合：時間帯識別子を付けて時間帯ごとにグループ化
- ▶ パネルデータの場合：各クロスセクション単位の観測値を時系列順に隣接させる

データの保存

データ入力オプション:

- ▶ テキストエディタを使用して生データファイルを作成
- ▶ データ操作と整理のためのスプレッドシート
- ▶ 計量経済学ソフトウェアへの直接入力

田中注: 現在では、Excel にデータを入力して、Excel もしくは CSV 形式で保存し、R や Stata などの統計ソフトで読み込むことが一般的である。

19-3c データの検査、クリーニング、要約

データ構造を理解することは非常に重要である。主な考慮事項は以下の通り：

- ▶ 欠損値のコード化方法（できれば非数値文字を使用）
- ▶ 変数の性質（二値、順序、測定単位）
- ▶ 金銭的価値の場合、数字が名目値か実質値か
- ▶ 変数が変換されているかどうか

要約統計量を通じたエラーの検出は不可欠である。予期しない最小値/最大値、平均値、標準偏差を発見することで、コーディングエラーが明らかになる場合がある。

時系列データの場合、季節調整と適切な順序付けを理解することが重要である。

19-4 計量経済学的分析

19-4 計量経済学的分析

テーマを確立し適切なデータを収集した後、計量経済学的アプローチを決定する。

最小二乗法（OLS）を使用する場合でも、以下の主要な仮定に対処することによりその適切性を正当化する必要がある：

- ▶ 誤差項が説明変数と相関していないかどうか
- ▶ 内生性の潜在的源泉（測定誤差、同時性）
- ▶ 関数形式の決定（対数、レベル、交互作用）

注意

避けるべき一般的な間違い：

- ▶ カテゴリー変数の生の数値コード（例：職業コード）を含める
- ▶ 順序変数を量的意味を持つものとして扱う

クロスセクション分析では、分散不均一性を考慮し、適切な場合には頑健な統計を使用する。

時系列アプリケーションでは、レベルと差分、トレンド、季節性、およびラグ構造に関する追加の注意が必要である。

モデルの定式化と感度分析

モデルの誤特定の可能性については、誤特定分析を実施して可能なバイアスの方向を推定する。

操作変数を使用する場合は、その妥当性を慎重に正当化する。

優れた実証論文には感度分析が含まれる

- ▶ 以下のようなモデルの合理的な修正に対して結論が頑健であるかをテストする：
- ▶ 代替変数定義
- ▶ 潜在的な外れ値の除外
- ▶ 異なる推定方法

パネルデータ

パネルデータでは、複数のアプローチが利用可能：

- ▶ プールド OLS（時間不変の観測されない要素をコントロールしない）
- ▶ ランダム効果（観測されない効果が条件付き平均値がゼロの場合に系列相関を修正）
- ▶ ラグ付き従属変数の使用
- ▶ 固定効果変換（観測されない効果を除去）

データマイニング（有意な結果が見つかるまで様々なモデルを検索する慣行）には注意が必要である。ほとんどすべての研究者が異なる定式化を試すが、このプロセスを明示的に認め、結果を代表的に報告することが重要である。

19-5 実証論文の執筆

19-5a 序論

序論では目的と重要性を確立し、通常は文献レビューも含む。効果的な序論には以下が含まれる：

- ▶ 興味深いパターンを示す簡単な統計
- ▶ これらのパターンに影響を与える要因をどのように検討するかの説明
- ▶ 読者を引き込むための主な発見の簡潔な要約

19-5b 概念的（または理論的）フレームワーク

このセクションでは、研究問題に答えるための一般的なアプローチを説明する。以下が含まれる場合がある：

- ▶ フォーマルな経済理論
- ▶ 概念的問題に関する直感的な議論
- ▶ コントロールが必要な要因と関心のあるメカニズムの説明

19-5c 計量経済学的モデルと推定方法

推定するモデルを表す方程式を含める。モデル（母集団の関係）と推定方法を区別する。例えば：

$$\text{colGPA} = \beta_0 + \beta_1 \text{alcohol} + \beta_2 \text{hsGPA} + \beta_3 \text{SAT} + \beta_4 \text{female} + u$$

このモデルは母集団の関係を表すため、「ハット」や特定の値を含まない点に注意。

モデルを指定した後、適切な推定方法とその正当性について議論する。

関数形式

関数形式を慎重に検討する：

- ▶ 広範囲の変数に対する対数変換
- ▶ ゼロになる可能性のある変数の適切な処理
- ▶ 率、パーセンテージ、比率の扱い

2SLS などの高度な方法については、操作変数選択の慎重な正当化を提供する。

19-5d データ

他者が作業を再現できるように、データソースを包括的に説明する。以下を含める：

- ▶ 参考文献にすべての公開データソース
- ▶ 各変数の測定単位
- ▶ 変数定義の表
- ▶ 要約統計（最小値、最大値、平均値、標準偏差）
- ▶ サンプルサイズと対象期間

19-5e 結果

前述のモデルの推定値を提示し、複雑なものに進む前に簡単な定式化から始める場合もある。

複数の変数を持つ複数のモデルの場合は、方程式ではなく表を使用する。

解釈と有意性に焦点を当てる：

- ▶ 係数は予想された符号を持つか
- ▶ 統計的に有意か
- ▶ 主要な効果の大きさはどの程度か
- ▶ 経済的有意性と統計的有意性を区別する

関連する仮説検定を提示し、定式化変更に対する結果の感度について議論する。

19-5f 結論

主な発見を簡潔に要約し、注意事項について議論し、さらなる研究の方向性を示唆する。

読者がまず最初にこのセクションを確認して論文全体を読むかどうかを決めることを想定して、このセクションを執筆する。

19-5g スタイルのヒント

- ▶ 記述的なタイトルを付け、あまり長くしない
- ▶ 名前、所属、および可能であれば要約を含む表紙を付ける
- ▶ 方程式は中央揃えにし、連続して番号を付ける
- ▶ 著者と年で論文を参照
- ▶ 表は変数を明確に定義して明瞭に書式設定する
- ▶ 小数点以下の桁数を制限する
- ▶ 必要に応じてデータと補足資料を付録にまとめる

まとめ

成功する実証分析には、問題の定式化、文献レビュー、データ収集、計量経済学的定式化、および明確な発表に細心の注意を払うことが必要である。

いかなる研究の質も、このプロセス全体を通じて投入された注意と努力に最終的に依存する。