AA Quantitative analysis notes

数理分析　― 演繹的

1. パズルを見つける
2. 前提条件を設定する
3. 理論を考える　理論＝前提条件の束

数量分析（＝計量分析）― 帰納的、実証分析

1. 理論から仮説を引き出す（作業仮説、正しければ理論は正しい）（対抗理論も同時に提示する）
2. 仮説検証する　（統計的に）（対抗理論を反証する）

よい理論の3条件

1. 反証可能である理論
2. N＝K問題がない
   1. Nは説明されるべき事例の数
   2. Kは説明の数
   3. なんでも説明できる→何も説明できていない。事例ごとに細分化していない、抽象性、普遍性、一般性を持たせるべきである。
3. トートロジーでない
   1. A statement in which you say the same thing twice in different words when unnecessary. Eg they spoke in turn, one after another.
   2. An explanation that fails to rule out other explanations
   3. A screenshot of a cell phone

      Description automatically generatedNo additional information, a restatement of known information.

予測と効果分析のための基礎

理論・仮説・仮説検証

良い理論

計量政治の流れ

1. 理論＝前提条件から構成。から、仮説を引き出す。
   1. 国の産業化の進展に伴い政治参加の度合いは大きくなる。
2. 仮説を作業仮説（検証可能）に変換する。仮説と作業仮説は同様と考えてもいい。
   1. 製造業に従事している労働力の割合が大きいほど、国政選挙での投票率が高い。その学問の常識の範囲内に指数を定め、観察可能なデータに変換する。
3. 理論の妥当性をテストするために、観察可能なデータを使って作業仮説を検証する。

Causal and empirical theory: If X happens, Y will follow. X -> Y   
Dependent variable (Y), Independent variable (X)

Why do we need theory? -> to uncover the underlying mechanism that governs phenomena

良い理論の条件

1. シンプル。
   1. しばしば条件つきで表現。他の条件が一定ならば、などなど。
   2. Difference between a model and a theory: Theories are generalized and simple, models are specifically designed to describe a certain application of a theory in reality.
   3. デゥベルジェの法則：小選挙区制では、有効政党数は2に収斂する。
   4. 軍事費を増強すると、その国家は戦争をする。
   5. 一人当たり収入が高い国ほど民主的な国家である。
   6. 民主的な国家は戦争しない傾向がある。(democratic peace theory)
2. Predictive accuracy
   1. 入手可能なデータを使う範囲では、理論は正確な予測をする必要がある。
   2. 一連の独立変数が、かなりの程度、従属変数を説明できるなら、it has strong explanatory power（寄与率、R^2決定係数） 一般的に15％程度もあれば良い。
3. Importance
   1. 新たな理論は従来の様々な理論に対して影響を与え、何らかの新たな知識を付与すべき。
4. Elegance
   1. A theory satisfying all three previously outlined conditions are elegant.
   2. Robert Putman, *Making Democracy Work*, 1993
5. 意外性Beautiful surprise
   1. 常識や期待をひっくり返すような説明とリサーチ結果
   2. 国境警備を強化すればするほど不法移民が増える。
   3. 非営利・非政府活動（=community activities）に人々が参加すればするほど、地方政治の効率がよくなる
   4. 民主主義国同士がより多く戦争する傾向にある。

理論と仮説の関係

* 理論：特定のトピックに関する一般化の束（前提条件の束）  
  a set of generalizations about a topic)  
  XとYという変数間の因果関係を示す。仮説よりも抽象的。
  + 人間の容姿は遺伝する
* 仮説：理論から引き出された、特定の変数に関する陳述。理論よりも具体的で焦点が絞られている。「もしこの理論が正しければ・・のはず」
  + 背の高い両親の子供は背が高い
  + 作業仮説
    - 母親が高ければ、子供も背が高い
    - 父親が背が高ければ、子供もせが高い。

回帰分析の二つの目的

* 予測のための回帰分析（観察研究―実験ではなく観察されたデータを用いる）
  + 重回帰分析・ロジスティック回帰分析
  + → Machine Learning
* 効果分析のための回帰分析（実験研究）
  + セレクションバイアス
  + R C T（ランダム化比較実験）
  + 自然実験・サーベイ実験
  + 傾向スコアマッチング
  + 回帰不連続デザイン（R D D）

コントロール変数

* 主要な独立変数（X）以外に影響を与える要因
  + 主要な独立変数は仮説から導かれている。X母親の身長、父親の身長
  + コントロール変数（睡眠時間、栄養）
* 独立変数＝X＋コントロール変数（Z）
* 主要な独立変数は一つとは限らない
* コントロール変数によって主要な独立変数と従属変数の間の相関が家いることも（人種と投票率、Z＝社会経済的地位、教育、住居地域）

作業化（operationalization）

* 人間の知性―＞IQ
* 政治家の選挙動員量→選挙費用
* 都会度→人口密度
* 国家の民主化度→FreedomHouse

分析単位

* 個人・市町村・都道府県・国・企業などの組織・年・政策・その他
* 分析単にが作業仮説における変数の作業かを左右する
* 作業仮説で使われる全ての変数の分析単位は同一でなければならない。
* 分析単位で次第で、一つの仮説から複数の作業仮説を作れる。

作業化で明らかにすべき事項：

独立変数

従属変数

コントロール変数

分析単位

例

理論1：衆議院候補者が選挙費用を使うほど、選挙は成功する

独立変数：選挙費用―＞候補者が総務省に提出する選挙運動資金データ

従属変数：成功する→候補者の得票率、当選結果

分析単位：衆議院候補者

作業仮説1：

衆議院候補者が総務省に提出する選挙運動資金データの額が大きいほどその候補者の得票率は大きい（当選する傾向がある）

理論2：都市は田舎よりも犯罪率が高い

独立変数：都市化度

従属変数：犯罪率

分析単位：州、都道府県、市町村など

作業仮説2：都市化度の高い州（都道府県、市町村）は、都市化度の低い州より（都道府県、市町村）犯罪率が高い

理論3：G N Pをコントロールすると、共産国家の方が資本主義国家よりも多くの軍事費を使う

独立変数：政治制度

従属変数：軍事費支出

コントロール変数：G N P

分析単位：国

仮説3：G N Pの影響を排除すると、政治制度が資本主義の国より共産主義の国方が軍事支出額が大きい。

理論4：正統をコントロールすると、中絶問題に関連する議員の投票は、議員の宗教と教育に関連している

独立変数：宗教と教育　→2つある。これに対応して作業仮説も2つ

従属変数：中絶問題に関連する投票

コントロール変数：政党

分析単位：議員

作業仮説4-1:　議員がキリスト教を信仰していると、所属政党に関わらず、その議員は中絶に関する法案には反対票を投じる

作業仮説4-2：議員が大卒だと、所属政党にかかわらず、その議員は中絶に関する法案には反対票を投じる。

エコロジカルな誤謬

* 集計データだけに基づいて個人の行動に関して誤った判断を下すこと。  
  例）カウンティのアフロアメリカンの割合と得票率。  
  強い相関からアフロアメリカンが投票したと結論

理論：収入が少ない有権者ほど、共産党に投票する

* 分析単位が都道府県の場合
  + 独立変数「収入」  
    作業化：都道府県別平均所得
  + 従属変数「共産党に投票」  
    作業化：参院選挙で都道府県別に共産党が得た票数
  + コントロール変数：  
    共産党への投票に影響を与えると思われる変数
* 分析単位が都道府県だとエコロジカルな誤謬の可能性を完全に排除できない（集計データのみに基づく）
* 分析単位が有権者個人
  + 収入―＞世論調査で聞いた年間所得
  + 共産党に投票―＞前回の参院選挙でどの政党に投票？
  + コントロールは一緒
* エコロジカルな誤謬は避けられるが、有権者が世論で嘘をつく可能性が否定できない。

変数測定 (Measurement) の問題と解決

* コンセプト・・・理論、仮説における観察できない変数
* メジャー・・・・作業仮説における「観察できる」変数
* メジャーはコンセプトの一部でしかない。
* 理論レベルのコンセプトAとコンセプトBの関係は観察できない
  + メジャーAとメジャーBの関係のみ観察できる
  + この関係がコンセプト同士の関係を反映していると言えるのは次の2つの前提条件が成立するとき
    - コンセプトAとメジャーAが関係あり
    - コンセプトBとメジャーBが関係あり
* 解決策
  + 複数のメジャーを統合した変数(a pooled measure) を作って、コンセプトに近づける。