

講義全体のまとめ

労働経済学 2

川田恵介

再現可能性 (透明性)

- 分析: 圧倒的に複雑な事象（社会、経済、顧客等々）を考えるために、整理された順序よい考察
- 今日の”分析”におけるキーワード
 - なぜそのような結論に至ったのか、極力第三者が検証できるようにする
 - ≠ 予言、論破
- 実務において、より重要？

重要概念

- 研究対象: 知りたいけど観察できない
- 分析対象: 観察できるけどどうでもいい
- 仮定: 議論の透明性を確保

仮定

- 分析の前提
 - 複雑な現象（社会）に対して、なんの前提もない分析は存在しない
- 必要な仮定を最初に列挙する
- “暗黙の前提”を極力避ける
 - 第三者にとってわかりにくい
- 「仮定がどの程度問題があるか？」に議論を明確に集中できる
- 数理モデル、定量的データ分析においては、比較的容易

例: フェルミ推定

- 問題: “西武池袋線の一車両に対して、ピンポン玉を何個詰め込めるか”
- 仮定: 車両、ピンポン玉の堆積
- 結論: 車両の堆積/ピンポンだまの堆積

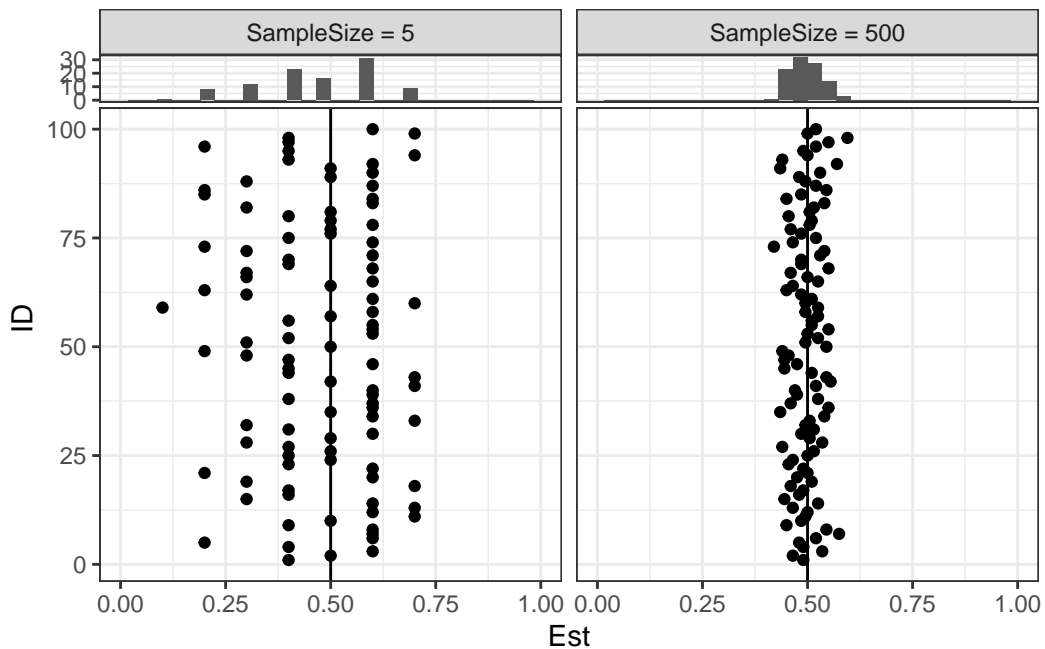
近似的仮定

- “完全に正しい” 仮定のみで、有意な結論は導くことは難しい
 - フェルミ推定: 正しい堆積とは?
- 近似的仮定: 検証可能なほぼ正しい仮定
 - “ほぼ” はケースバイケース

信頼区間

- 一般にデータにおける平均 \neq 社会における平均
 - 人によってデータが異なる = 平均が異なる
- 近似的仮定: データの平均値は、正規分布に従う
 - データが十分に大きければ、近似的に成立
- 近似的信頼区間 = データから計算される、一定確率で、真の平均を含む区間
 - ランダムサンプリング・データが十分に大きければ、信頼できる
 - 平均以外にも幅広く応用できる

例



解釈のための仮定

- トレードオフ
 - より多くの含意
 - より多く (厳しい) の仮定
- 「データの収集を頑張る」だけでは正当化しづらい仮定を導入せざるえない

メカニズムの理解のための仮定

- 経済学の目標: Action(政策決定、企業戦略決定、家計の意思決定) の支援
- 「差の原因」を理解することが重要
 - 物価が上がった原因は、国内需要の増加?、輸入品価格の増加?
 - 発熱の原因は、運動?、ウィルス?

因果推論

- D が異なる二つのグループ間で差があるかどうかは、近似的仮定 (ランダムサンプリング・サンプルサイズが十分に大きい) のみで検証可能
- 差の原因は D というためには？
 - 「 D がランダムに決定されている」 仮定が必要
 - 実験できないケースは、家庭への入念な議論が必要

因果推論の問題

- 色々限界がある
 - かなりざっくりした因果効果しかわからない
- D が Y に影響を与える経路は、論じられない
- Y は観察できる必要がある

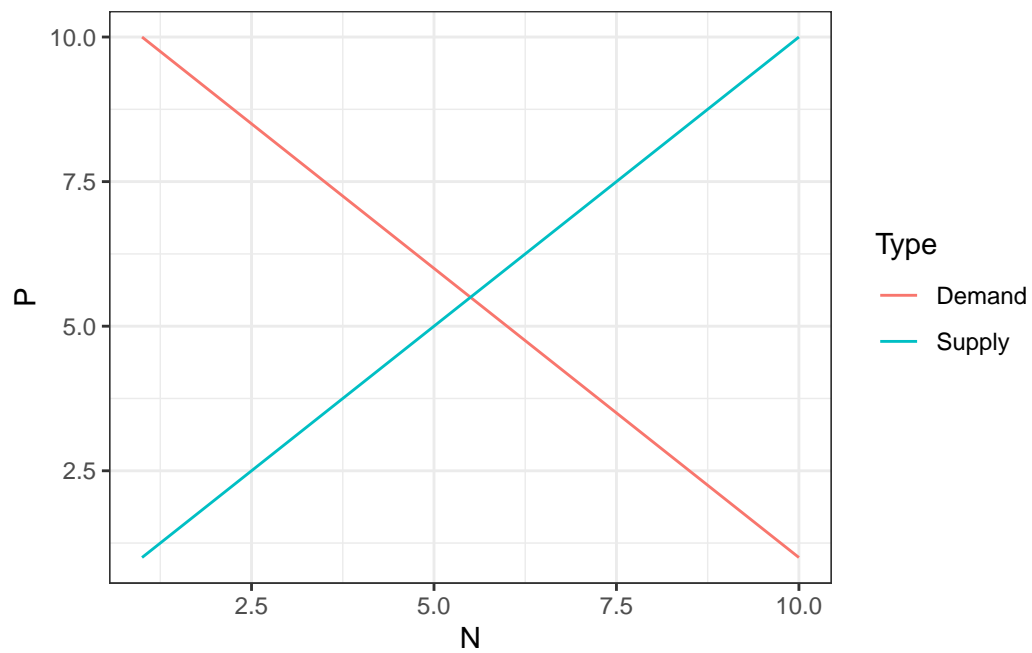
経済理論

- より難しい研究課題について、論点整理を提供
 - 統計的に推論できる変数と観察できない変数を接続する仮定
 - 観察できる変数について、解釈できるグループわけ

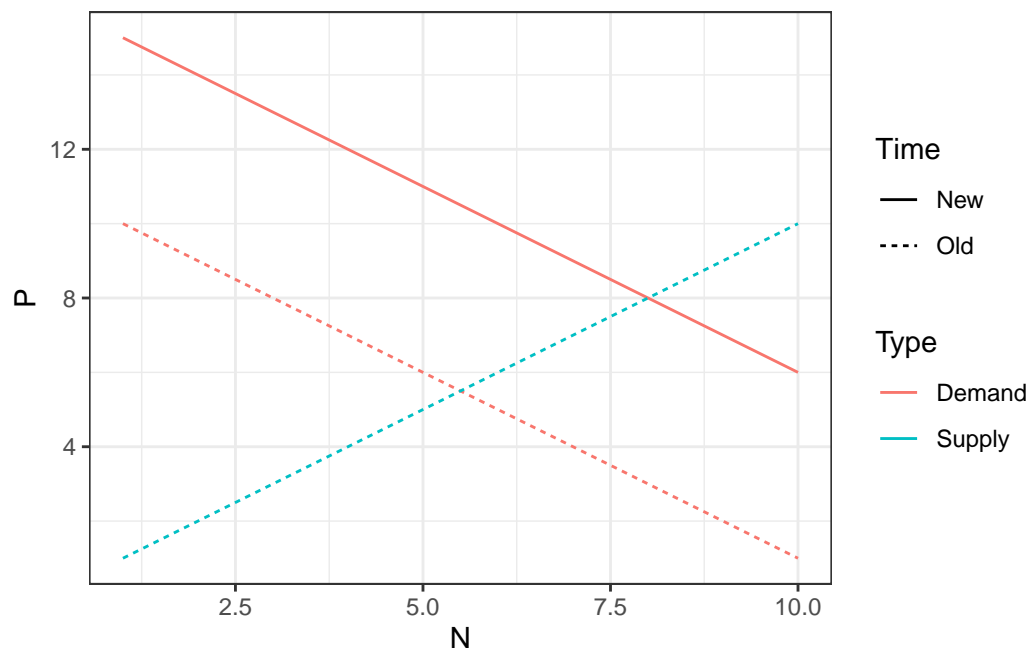
需求フレームワーク

- 価格と数量の変動を、観察できない需要と供給に紐付け
- 労働経済学においては、他の観察できない変数、独占力、への紐付けも重要

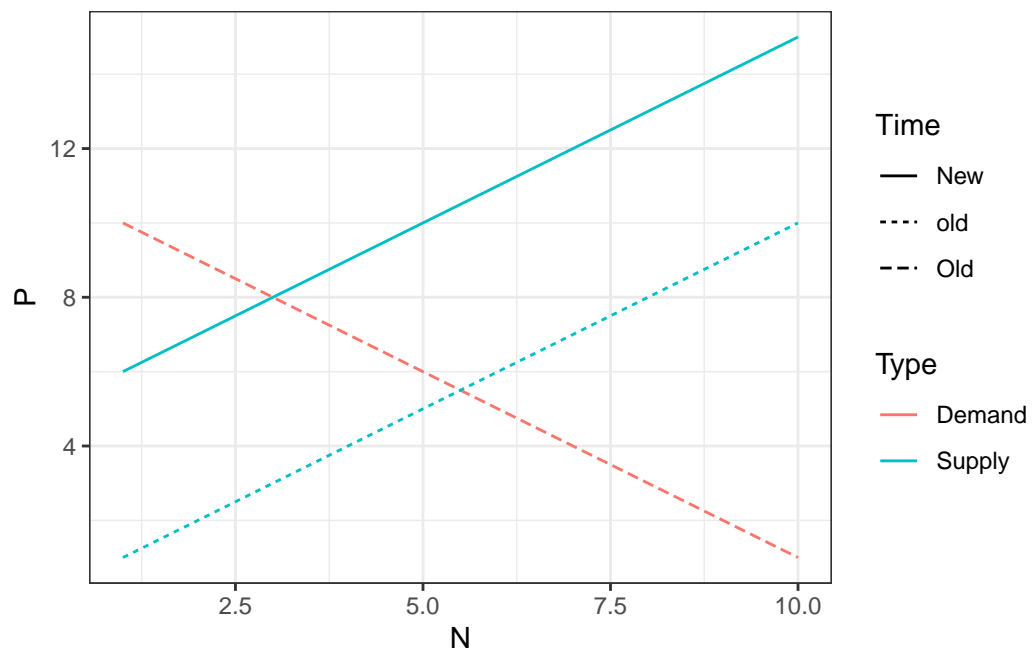
相場決定



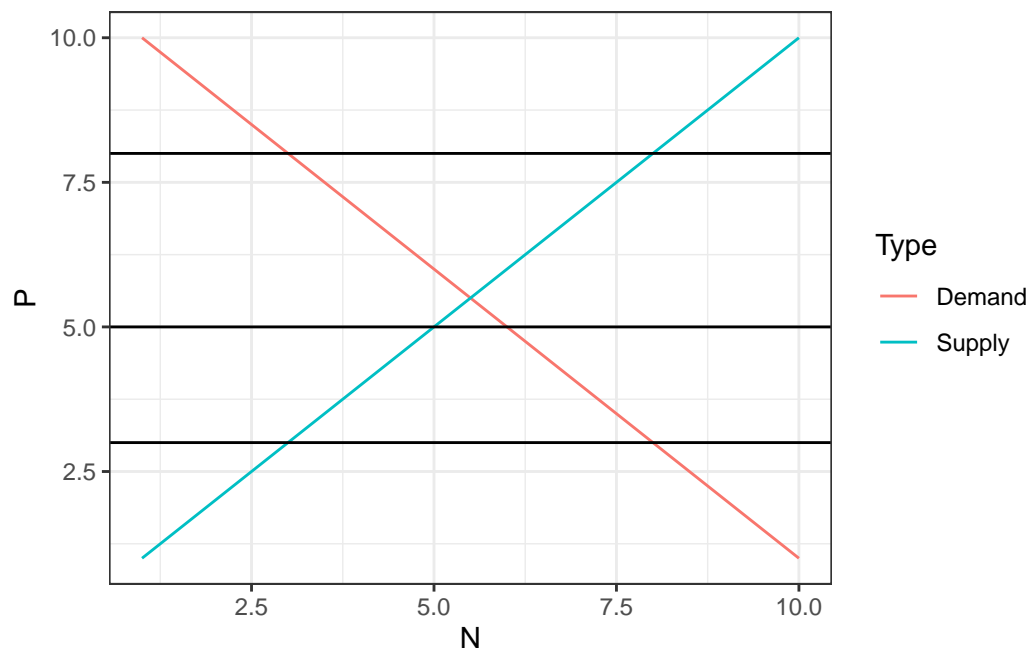
相場決定: 需要主導



相場決定: 供給主導



相場決定: 価格規制



他の例

- Dependency ratio: 少子高齢化に起因する社会問題への対策として、出生率向上政策は有効
 - 若年・高齢者はともに従属人口
- 最適失業率: 求人を増やすことの社会的費用を考えると？

データ整備

- 分析結果を不透明化させる大きな要因
- “分析” と比べて、必要なコード量が多く、バグの温床

tidyverse

- dplyr パッケージを同梱: データ整備について豊富な関数を提供
- 特に重要な関数
 - select: 変数の限定
 - mutate: 変数の作成
 - filter: 事例の限定

元データ

```
# A tibble: 4 x 3
      X     Y     Z
  <dbl> <dbl> <dbl>
1     0     0   NA
2     1     4   NA
3     2     2   NA
4    NA    10    1
```

- NA: 欠損値
 - 処理が難しい

変数の作成

```
# A tibble: 4 x 4
```

	X	Y	Z	YX
	<dbl>	<dbl>	<dbl>	<dbl>
1	0	0	NA	0
2	1	4	NA	5
3	2	2	NA	4
4	NA	10	1	NA

変数の除去

- あまりにも欠損値がある変数は除去するのも手

```
# A tibble: 4 x 3
      X     Y   YX
  <dbl> <dbl> <dbl>
1     0     0     0
2     1     4     5
3     2     2     4
4    NA    10    NA
```

同じ書き方

- あまりにも欠損値がある変数は除去するのも手

事例の除去

- 欠損値がある事例は除去するのも手

```
# A tibble: 2 x 3
      X     Y   YX
  <dbl> <dbl> <dbl>
1     1     4     5
2     2     2     4
```

他の例

- 欠損値がない事例のみ残す

```
# A tibble: 2 x 3
      X     Y   YX
  <dbl> <dbl> <dbl>
1     1     4     5
```


2 2 2 4

まとめ

- 分析成果を上司、実務家、社会に公表した後にミスが発覚すると焦るので、丁寧な議論・分析の徹底を！！！！