

Rで学ぶ 傾向スコア解析入門

@yokkuns: 里 洋平

2011.09.24 第17回R勉強会@東京 (#TokyoR)

AGENDA

- 自己紹介
- 傾向スコア解析
 - 実験出来るデータ
 - 実験出来ないデータ
 - 共変量調整
 - 傾向スコア推定
 - 傾向スコアを用いた調整
- Rによる実行
- 最後に

AGENDA

- 自己紹介
- 傾向スコア解析
 - 実験出来るデータ
 - 実験出来ないデータ
 - 共変量調整
 - 傾向スコア推定
 - 傾向スコアを用いた調整
- Rによる実行
- 最後に

自己紹介

- ID : yokkuns
- 名前 : 里 洋平
- データマイニングエンジニア
- 統計解析 パターン認識
機械学習 データマイニング
NLP 金融工学 などを勉強中



Tokyo.Rの主催者

Google グループ

« [Google グループのホーム](#)

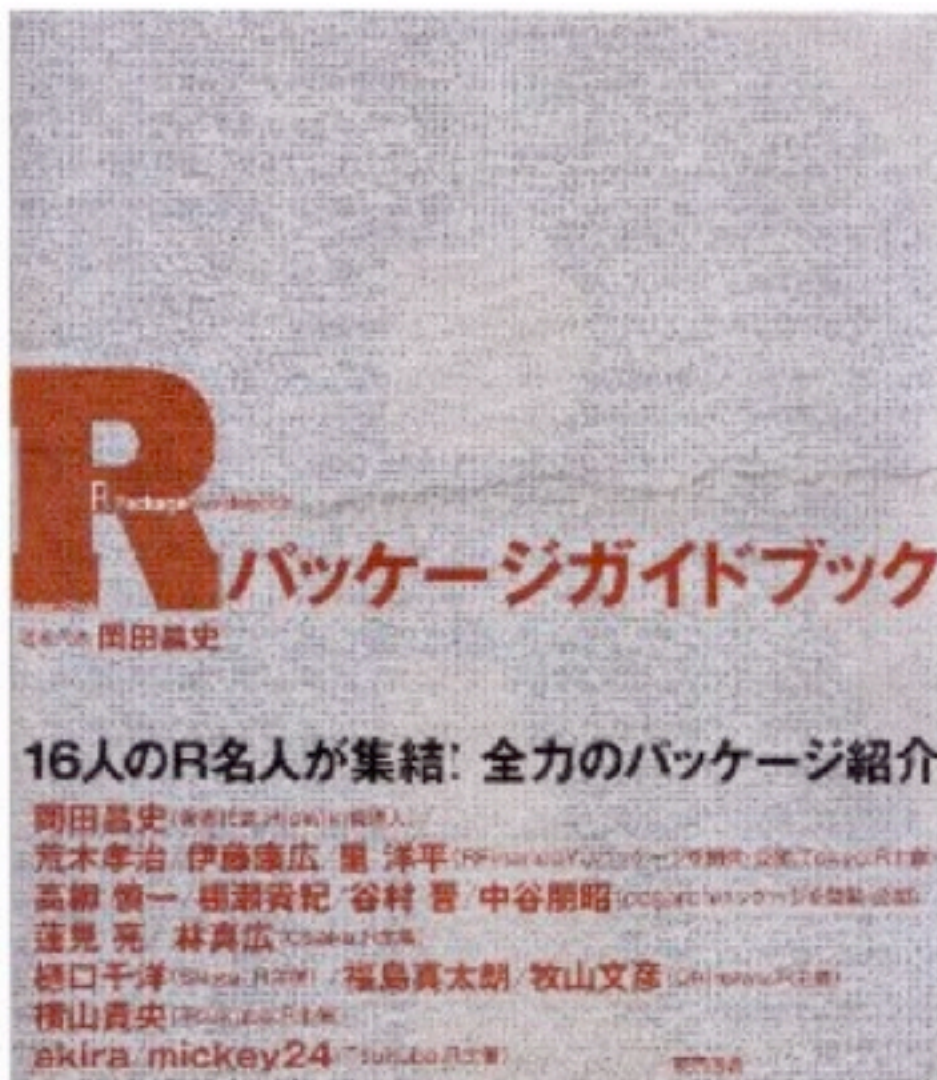


R勉強会@東京 - Tokyo.R

ホーム

ご参加ありがとうございます！

Rパッケージ本執筆



Rパッケージガイドブック [単行本]

[岡田 昌史](#) ☒ (著), [荒木 孝治](#) (その他), [伊藤 康広](#) (その他), [里 洋平](#) (その他), [高柳 慎一](#) (その他), [棚瀬 貴紀](#) (その他), [谷村 晋](#) (その他), [中谷 朋昭](#) (その他), [蓮見 亮](#) (その他), [林 真広](#) (その他), [樋口 千洋](#) (その他), [福島 真太郎](#) (その他), [牧山 文彦](#) (その他), [横山 貴央](#) (その他), [akira](#) (その他), [mickey24](#) (その他)

[この商品の最初のレビューを書き込んでください。](#) いいね (6)

価格: **¥ 3,990** 通常配送無料 [詳細](#)

通常2~4週間以内に発送します。 [在庫状況](#)について

この商品は、[Amazon.co.jp](#) が販売、発送します。ギフトラッピングを利用できます。

[中古品1点](#) ¥ 6,133より

動画レコメンド

動画レコメンド

関連オススメ動画リスト

次のリストを見る

6 / 30

SF家電 洗濯機編

みんなの評価

評価する

000 / 000

関連オススメ動画リスト

からいこい 家電の選び方

SF家電 電子レンジ編

この洗濯機は だれに

男の家事 (洗濯編)

回し回しで もこんな風に遊ぶ

視聴履歴を見る

全件クリア

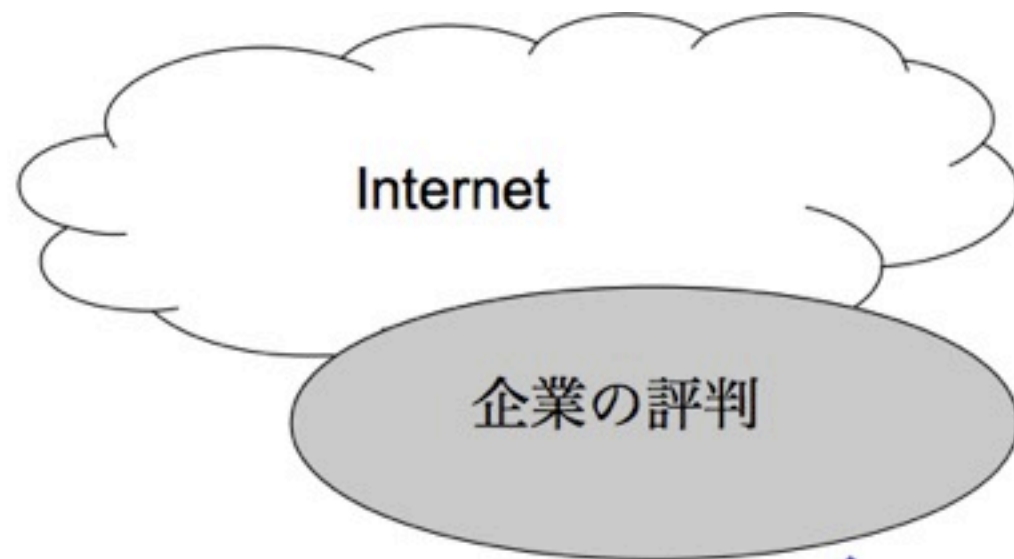
利用環境の確認

Get Adobe FLASHPLAYER

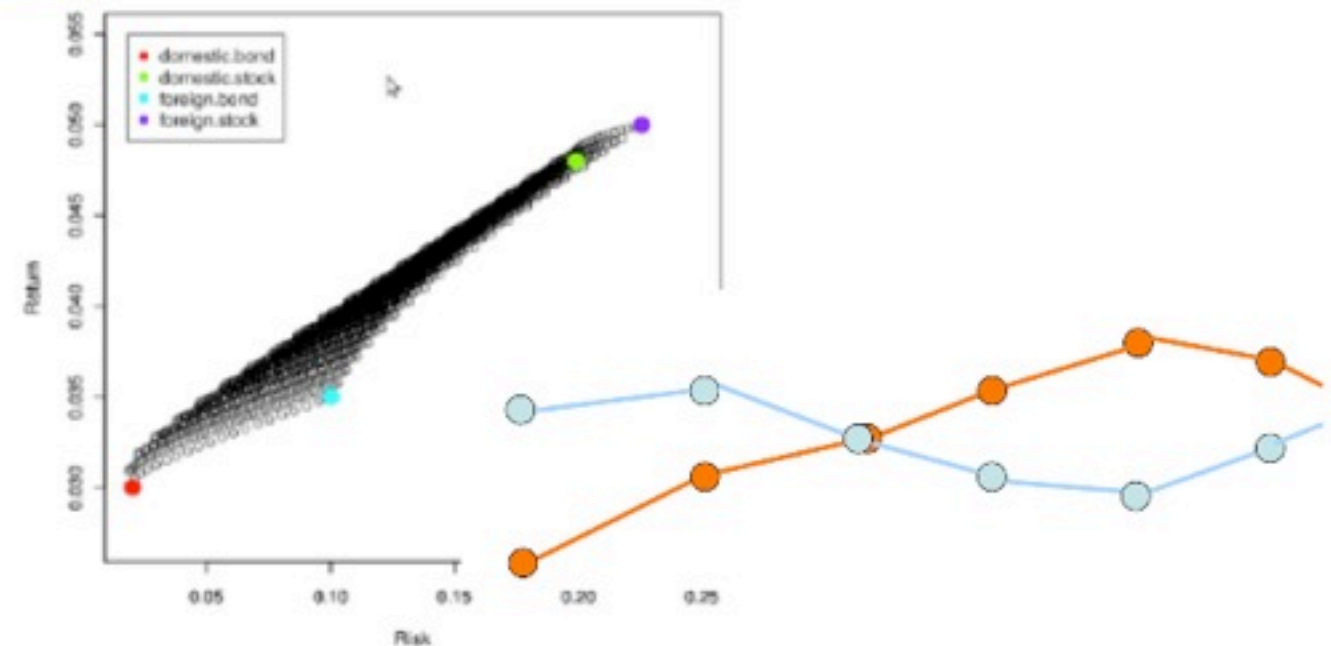
動画と音声の再生には、プレイヤーソフトウェア(Adobe Flash Player)が必要です。左のボタンを押して移動したサイトへ入手できます。さらに詳しくは...

閲覧されてる動画に関連する
おすすめ動画を表示

テキストマイニング+金融工学



Web上にある評判情報から
市場予測！

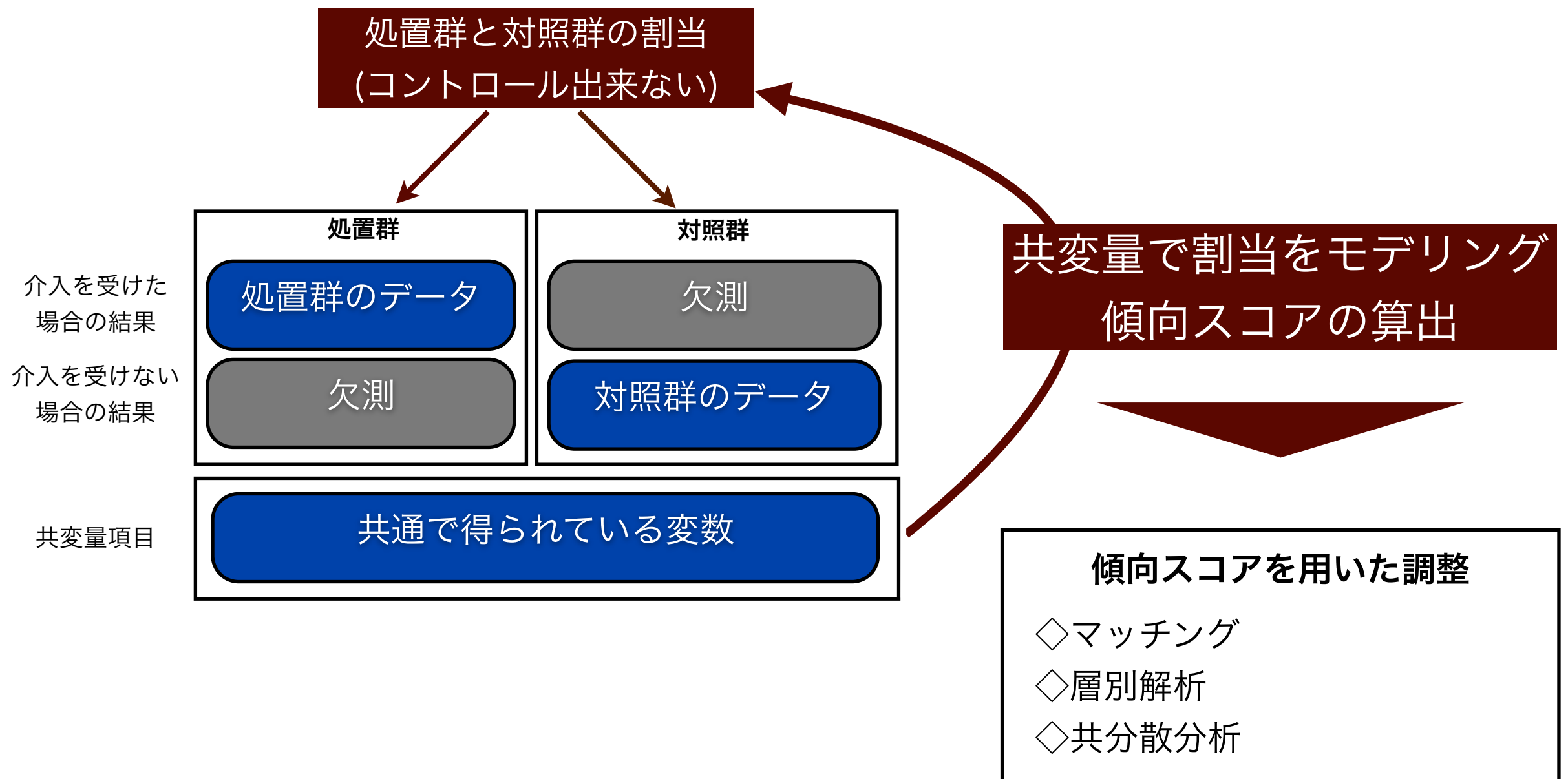


AGENDA

- 自己紹介
- 傾向スコア解析
 - 実験出来るデータ
 - 実験出来ないデータ
 - 共変量調整
 - 傾向スコア推定
 - 傾向スコアを用いた調整
- Rによる実行
- 最後に

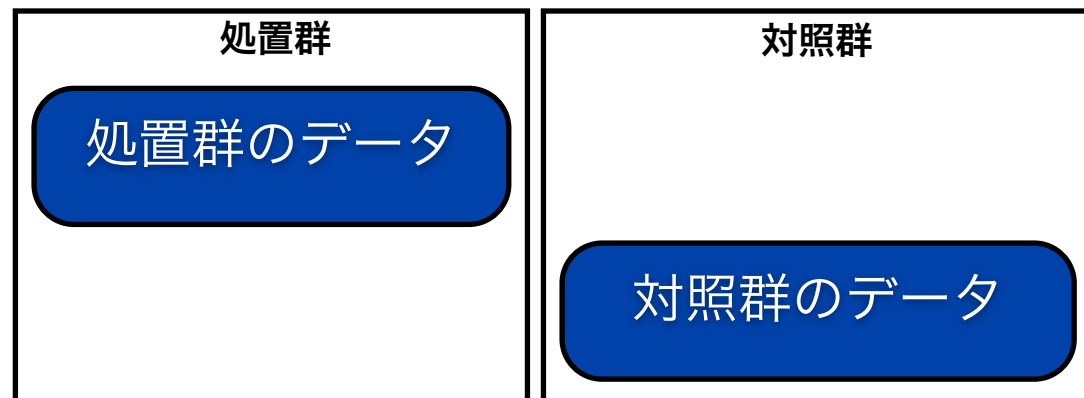
傾向スコア解析

実験出来ないデータの因果関係を解析する



実験出来るデータ

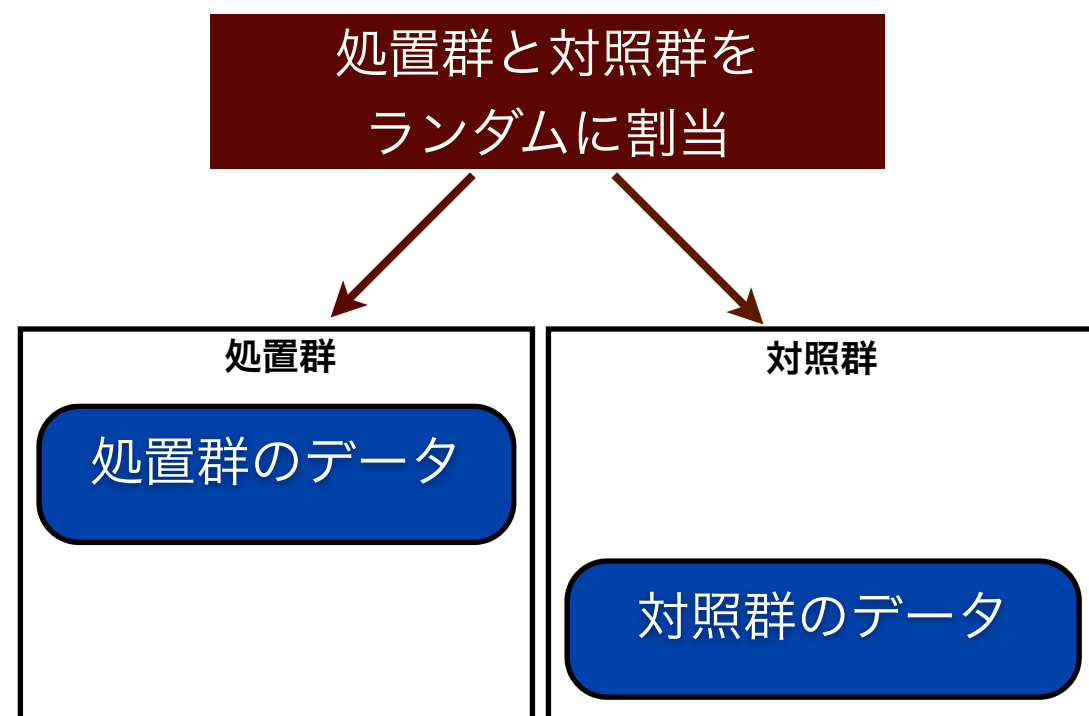
因果効果は単純な処置群と対照群の差になる



因果効果 = 処置群の平均 - 対照群の平均

実験出来るデータ

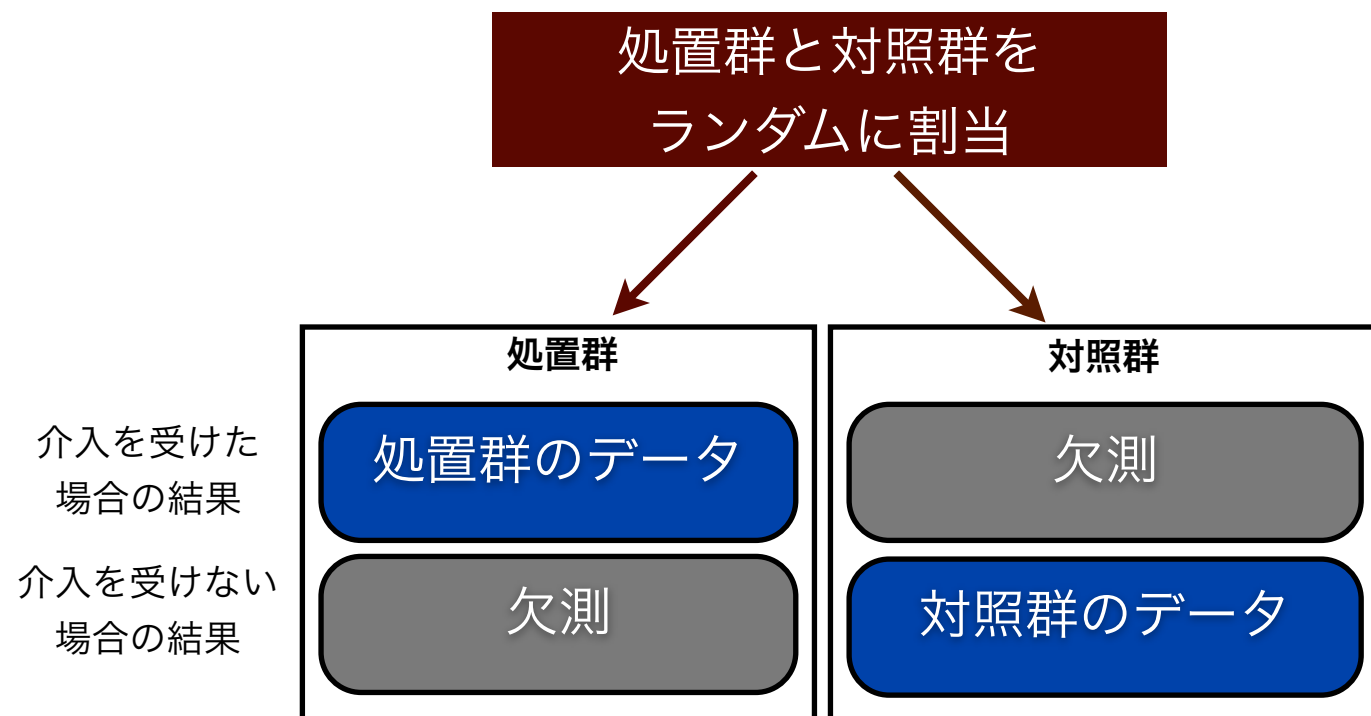
因果効果は単純な処置群と対照群の差になる



因果効果 = 処置群の平均 - 対照群の平均

実験出来るデータ

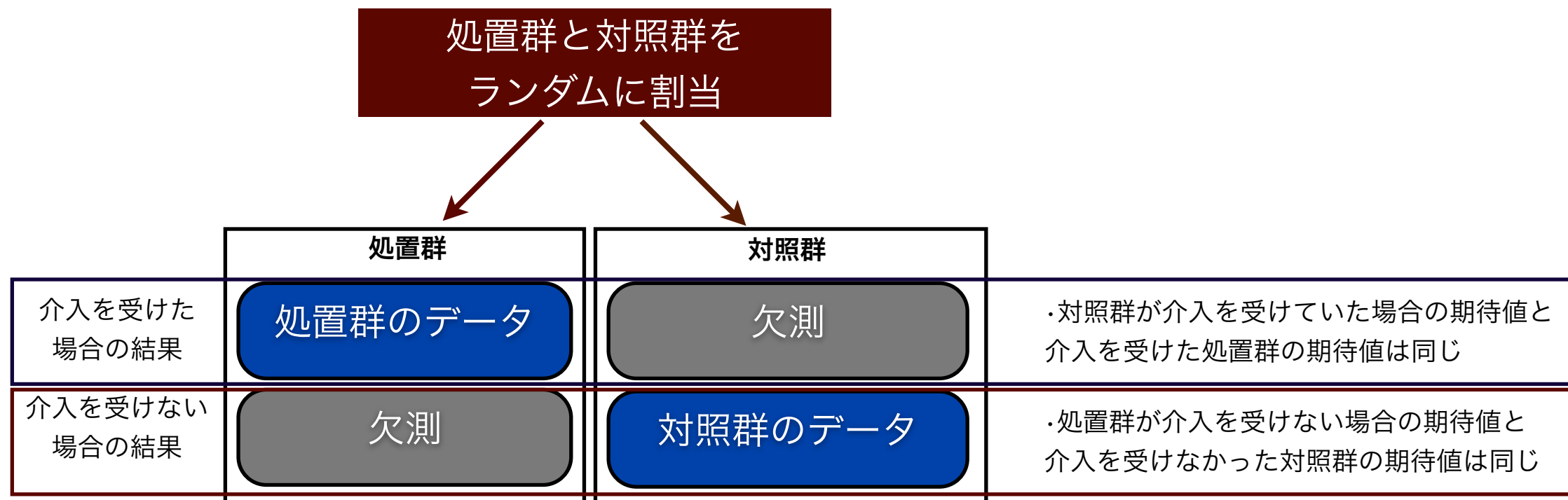
因果効果は単純な処置群と対照群の差になる



因果効果 = 処置群の平均 - 対照群の平均

実験出来るデータ

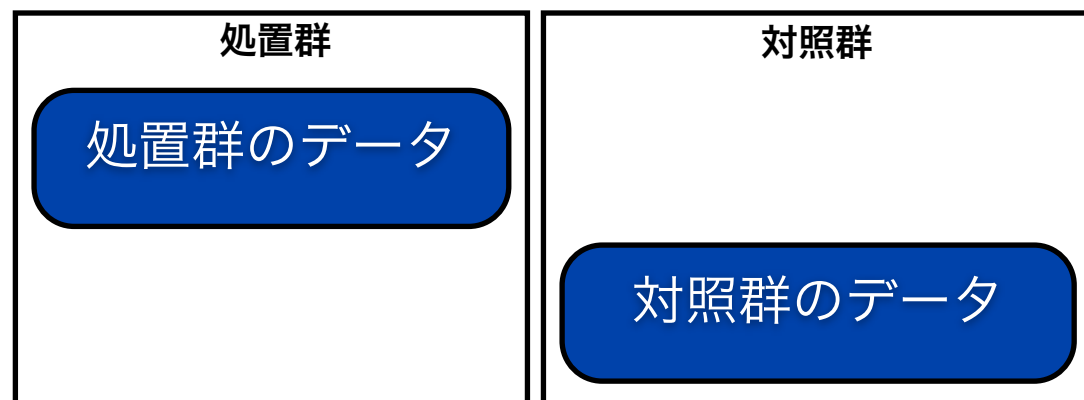
因果効果は単純な処置群と対照群の差になる



因果効果 = 処置群の平均 - 対照群の平均

実験出来ないデータ

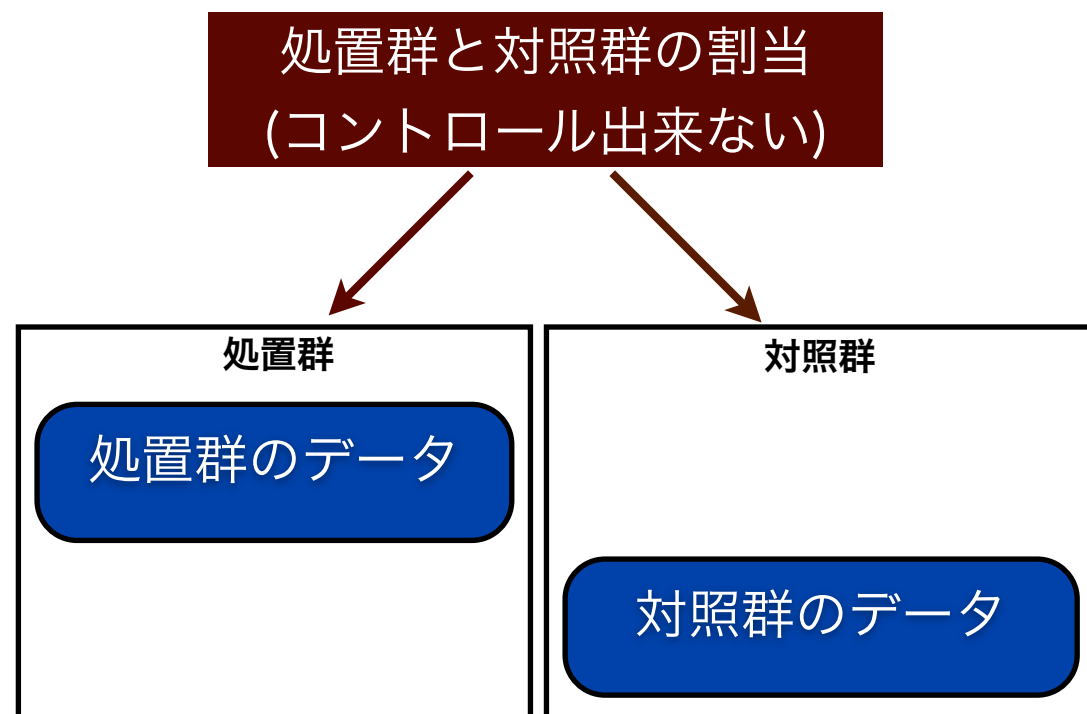
割当によって処置群と対照群に差が生じるため
単純に比較することが出来ない



因果効果 \neq 処置群の平均 - 対照群の平均

実験出来ないデータ

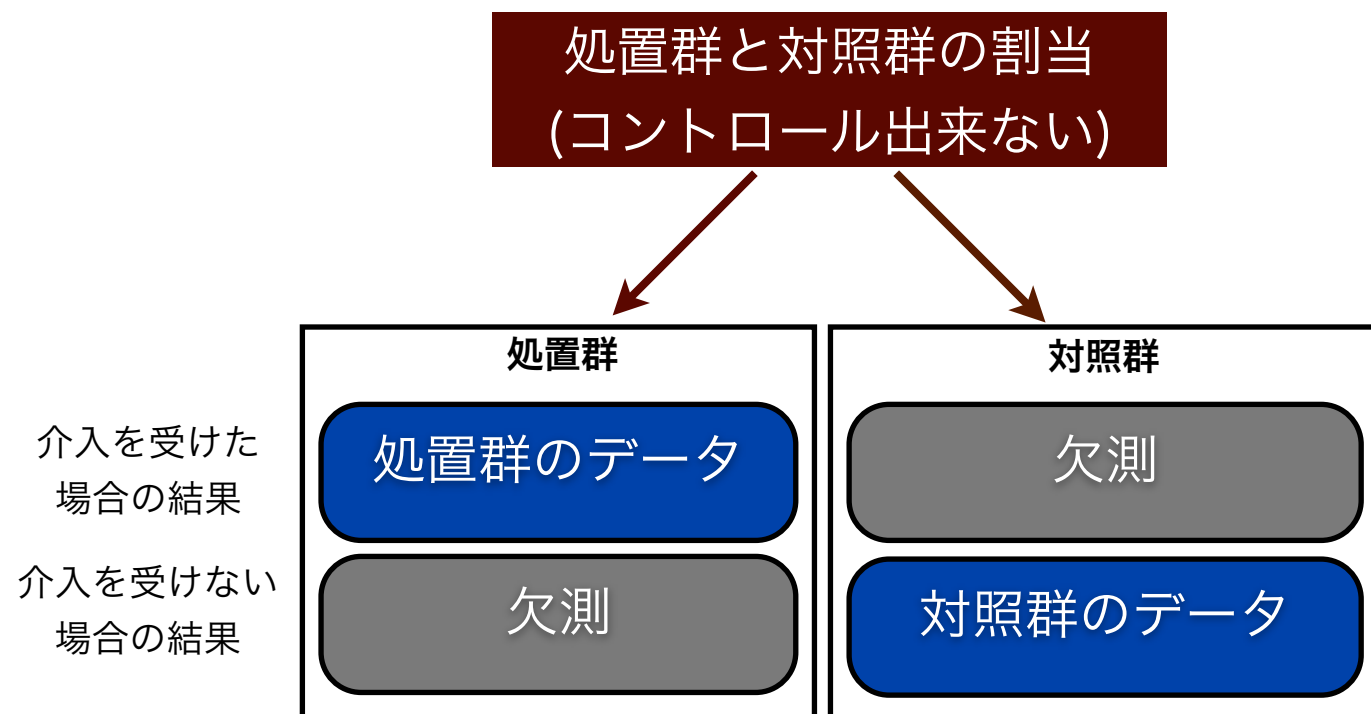
割当によって処置群と対照群に差が生じるため
単純に比較することが出来ない



因果効果 \neq 処置群の平均 - 対照群の平均

実験出来ないデータ

割当によって処置群と対照群に差が生じるため
単純に比較することが出来ない



因果効果 \neq 処置群の平均 - 対照群の平均

実験出来ないデータ

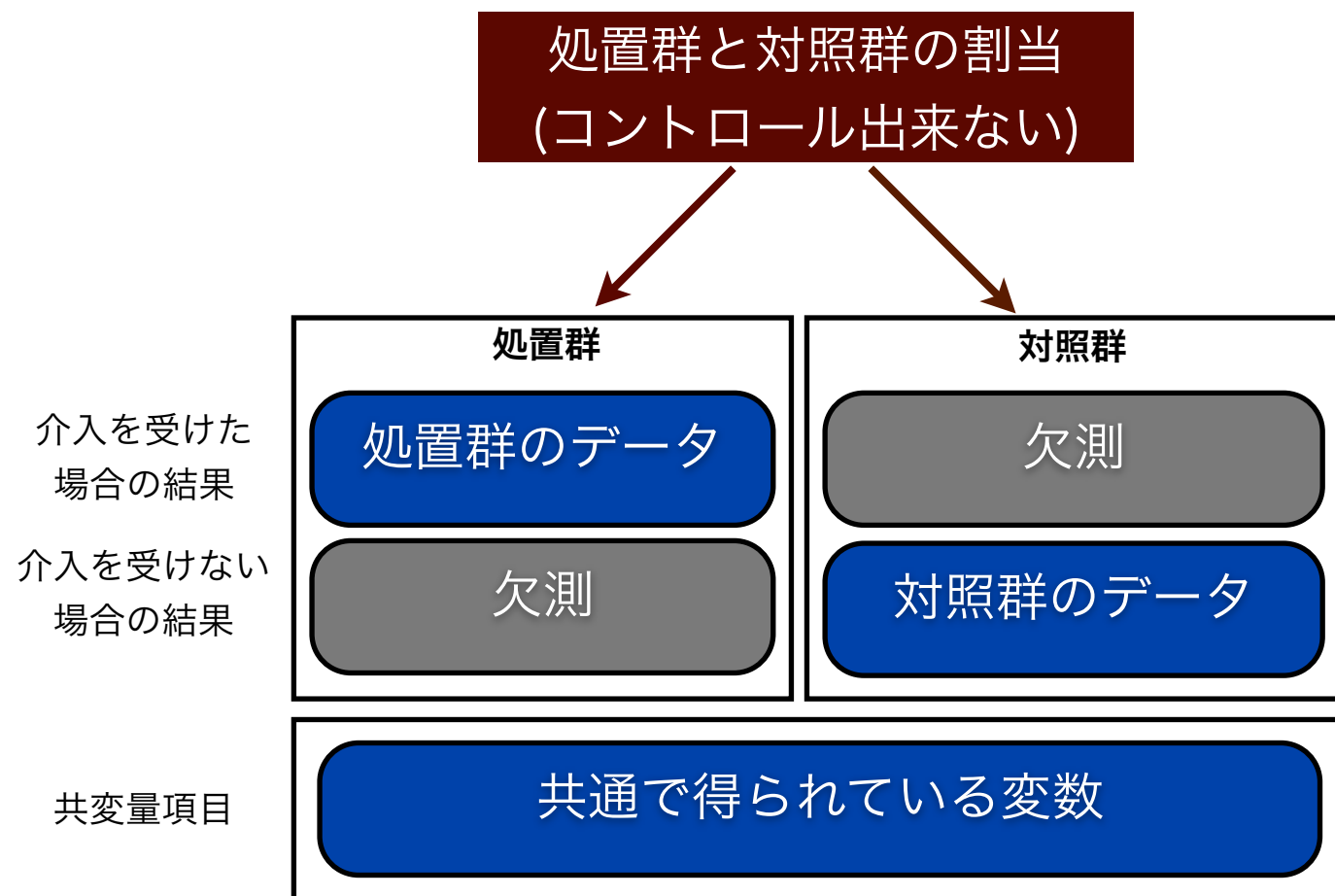
割当によって処置群と対照群に差が生じるため
単純に比較することが出来ない

処置群と対照群の割当 (コントロール出来ない)			
	処置群	対照群	
介入を受けた 場合の結果	処置群のデータ	欠測	・対照群が介入を受けていた場合の期待値と 介入を受けた処置群の期待値が異なる
介入を受けない 場合の結果	欠測	対照群のデータ	・処置群が介入を受けない場合の期待値と 介入を受けなかった対照群の期待値が異なる

因果効果 \neq 処置群の平均 - 対照群の平均

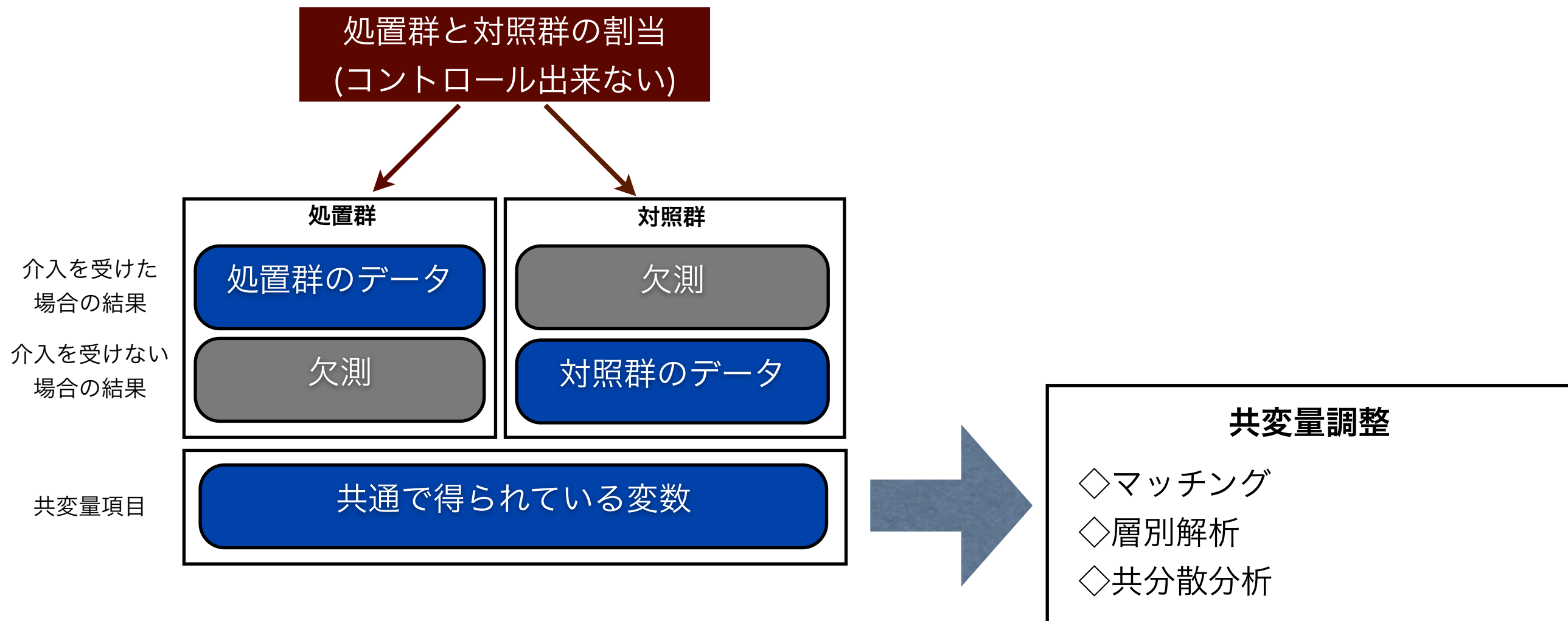
共変量調整

割当や結果変数に影響している共通の変数を用いて
因果効果以外の効果を除去する



共変量調整

割当や結果変数に影響している共通の変数を用いて
因果効果以外の効果を除去する



共変量調整：マッチング

共変量調整：層別解析

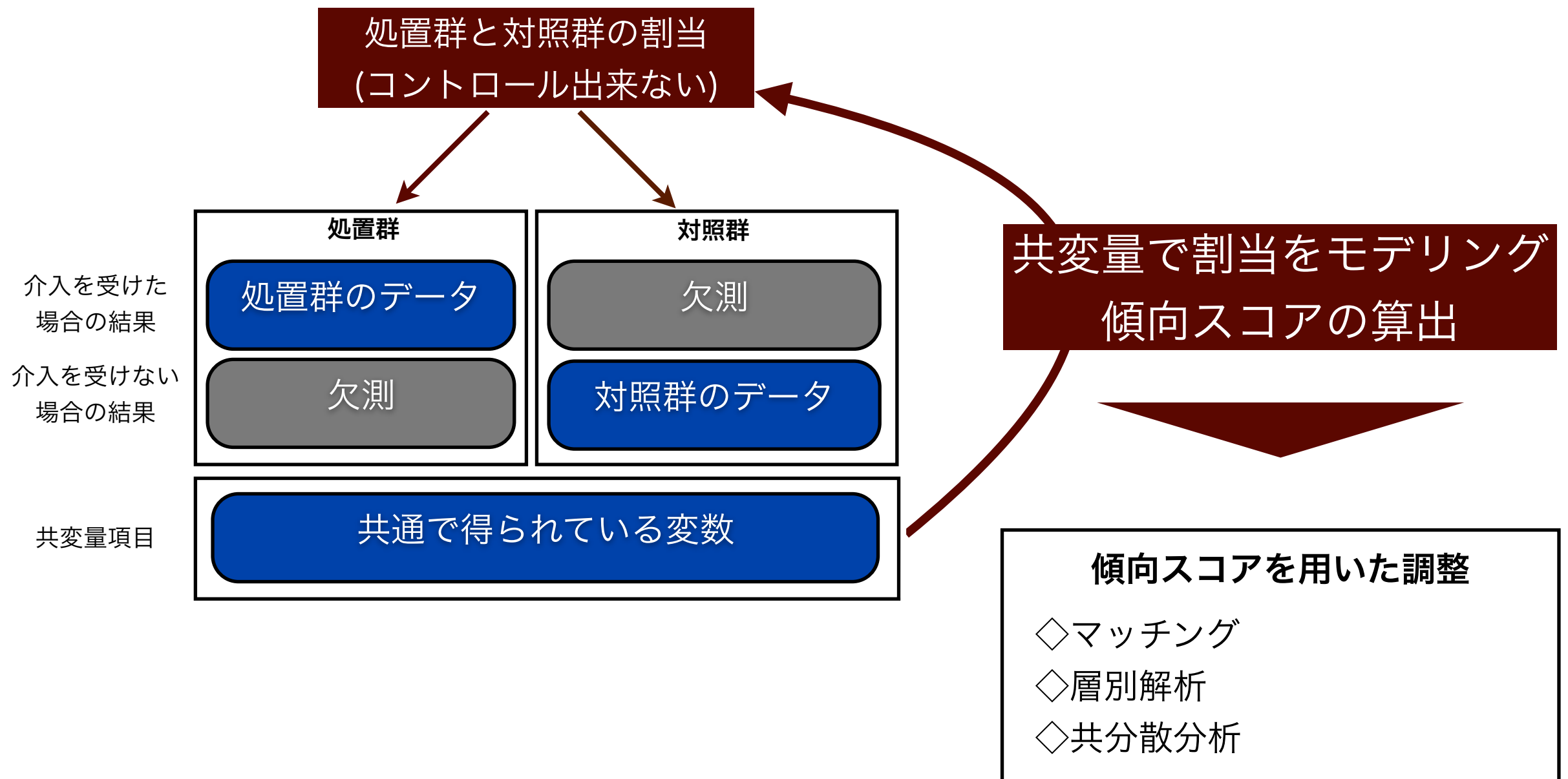
共変量調整：回帰モデル

共変量調整の問題点

- マッチング・層別解析での問題
 - 共変量に連続変数があると完全一致のペアは作れない
 - 次元問題
 - サポート問題
- 回帰モデルでの問題点
 - 結果変数と共変量のモデリングが必要
 - 直接因果効果の推定値は得られない

傾向スコア解析

実験出来ないデータの因果関係を解析する



傾向スコアとは

傾向スコアの推定

傾向スコアを用いた調整

- マッチング
- 層別解析
- 共分散分析

AGENDA

- 自己紹介
- 傾向スコア解析
 - 実験出来るデータ
 - 実験出来ないデータ
 - 共変量調整
 - 傾向スコア推定
 - 傾向スコアを用いた調整
- **Rによる実行**
- 最後に

Rによる実行

AGENDA

- 自己紹介
- 傾向スコア解析
 - 実験出来るデータ
 - 実験出来ないデータ
 - 共変量調整
 - 傾向スコア推定
 - 傾向スコアを用いた調整
- Rによる実行
- **最後に**

次回以降の
発表者・LTを募集しています！

ご清聴ありがとうございました

AGENDA

- 自己紹介
- 傾向スコア解析
 - 実験出来るデータ
 - 実験出来ないデータ
 - 共変量調整
 - 傾向スコア推定
 - 傾向スコアを用いた調整
- Rによる実行
- 最後に

付録

参考文献