

ゼミナール紹介

Keywords : 統計科学, 大規模データ解析, 因果推論

2024年12月27日

片山 翔太

本日のスライドはこちら →

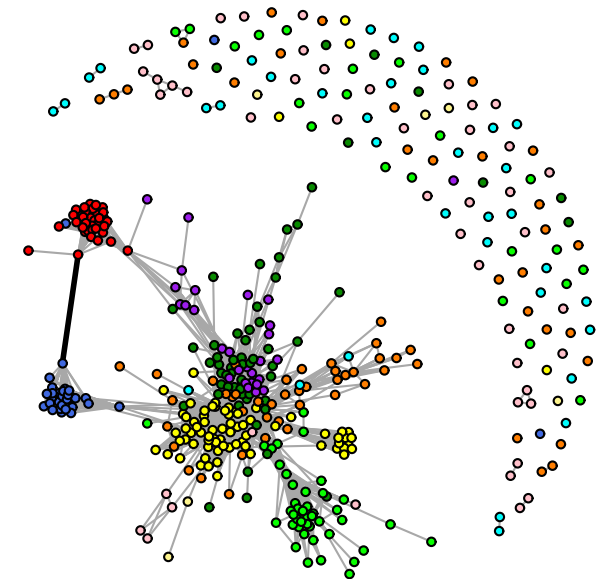


本日のアウトライン

- 自己紹介
- ゼミナール紹介
- 入ゼミに関する諸連絡
- Q and A

自己紹介

- 氏名：片山 翔太 (かたやま しょうた)
- 出身：熊本県荒尾市
- 担当授業：統計学(日吉), 確率・統計(三田)
- 専門：統計科学・機械学習
 - 高次元データ, 漸近分布論, モデル選択, ロバスト推測
 - グラフィカルモデリング, 因果推論



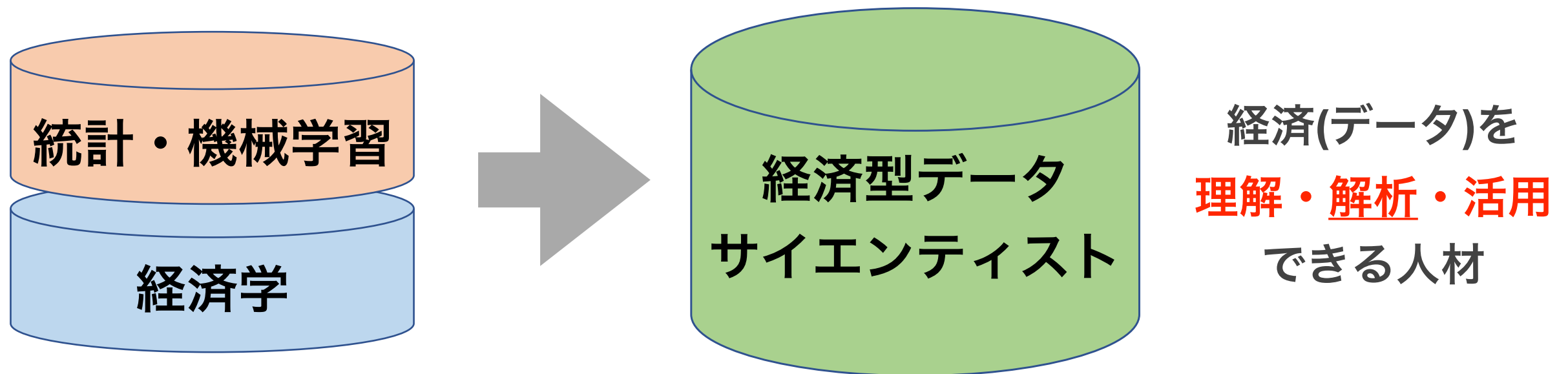
ゼミの特色

- 少人数制のゼミナール (個別ゼミ)
 - ひとつのテーマについて**2~3人**で深い議論を行う
 - ガウス過程, 合成コントロール法 (Synthetic control)
 - 基本的には**希望に応じて開催**
- プレゼンテーション能力の強化
 - 少なくとも**2週間に1回**は構成員の前で発表
 - 輪読 + 進捗報告 + α
- 大学院進学を推奨
 - 慶應・経済学研究科, 京大・経済学研究科
 - 一橋・経営管理研究科, 一橋・データサイエンス研究科

どんなことを扱うゼミ？



- 統計科学的な理論・方法論を詳しく研究！
 - 社会でもデータ解析スキルの需要大
 - 最近では「データ解析経験が当たり前」ということも...
 - データを解析できるだけでは不十分



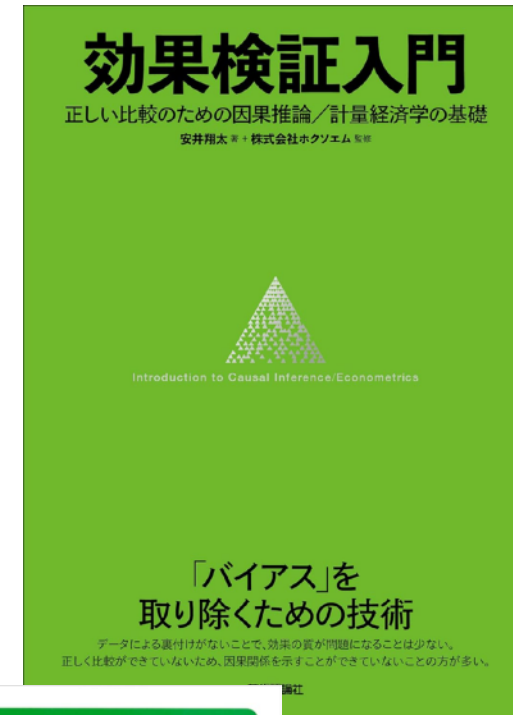
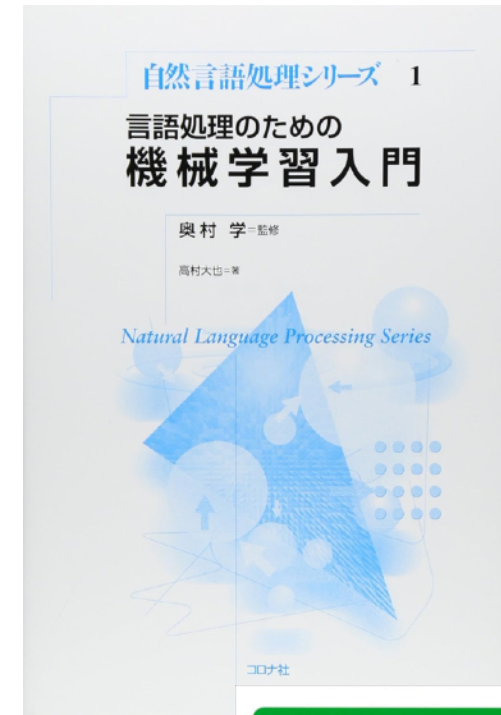
どんなことを扱うゼミ？

- To Do

- **解析** ← 理論と方法論の正しい理解
- **活用** ← 思考力と応用力

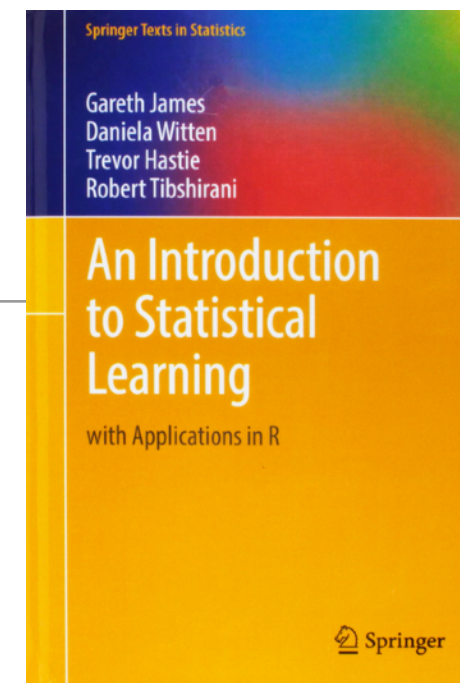
- 文献の輪読

- 数理統計や**因果推論**に関するテキスト
- 構成員全員が理解できるように説明
- **思考力 + ディスカッション力を磨く**



どんなことを扱うゼミ？

- データ解析の実践
 - 統計科学の手法を学んでRで実践する



```
> result<-randomForest(AHD~.,  
+                        data=heart)
```

```
> importance(result)  
  
             MeanDecreaseGini  
Age             12.656294  
Sex              4.892732  
ChestPain       19.772831  
RestBP          10.648762  
Chol            11.642479
```

```
> for(b in 1:200){  
+   tmp<-sort(sample(1:n,n,replace=T))  
+   b.sample<-heart[tmp,]  
+   oob.sample<-heart[-tmp,] ##oob sample  
+   result<-rpart(AHD~.,data=b.sample,control=rpart.control(minsplit=10,cp=0.02))  
+   pre<-as.numeric(predict(result,newdata=oob.sample,type="class"))  
+   y <- as.numeric(oob.sample[, "AHD"])  
+   oob.error.normal <- sum(y!=pre)/length(y) #misclassification rate  
+  
+   for (j in 1:13) {  
+     oob.sample<-heart[-tmp,]  
+     n.oob <- length(oob.sample[,1])  
+     oob.sample[,j]<-oob.sample[sample(1:n.oob,n.oob),j] ##sorted variable j  
+     randomized.pre<-as.numeric(predict(result,newdata=oob.sample,type="class"))  
+     oob.error.random[b,j] <- sum(y!=randomized.pre)/length(y) - oob.error.normal  
+   }  
+ }
```

Which do you like?

具体的な手法

- 春学期

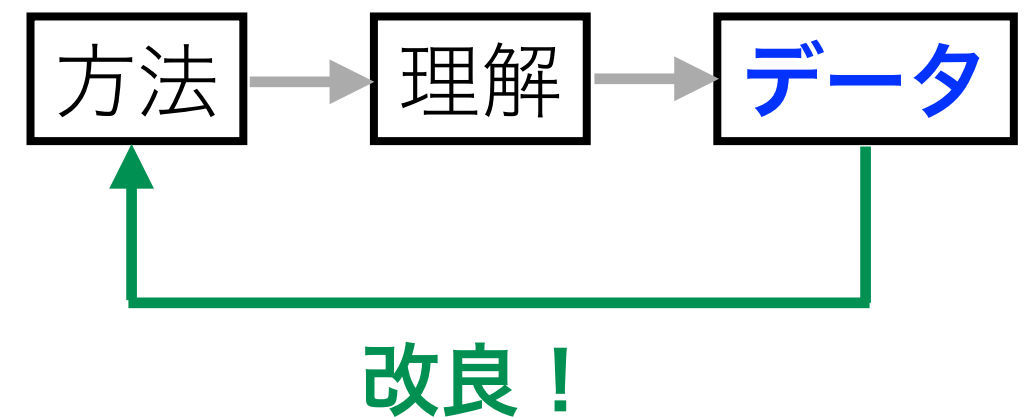
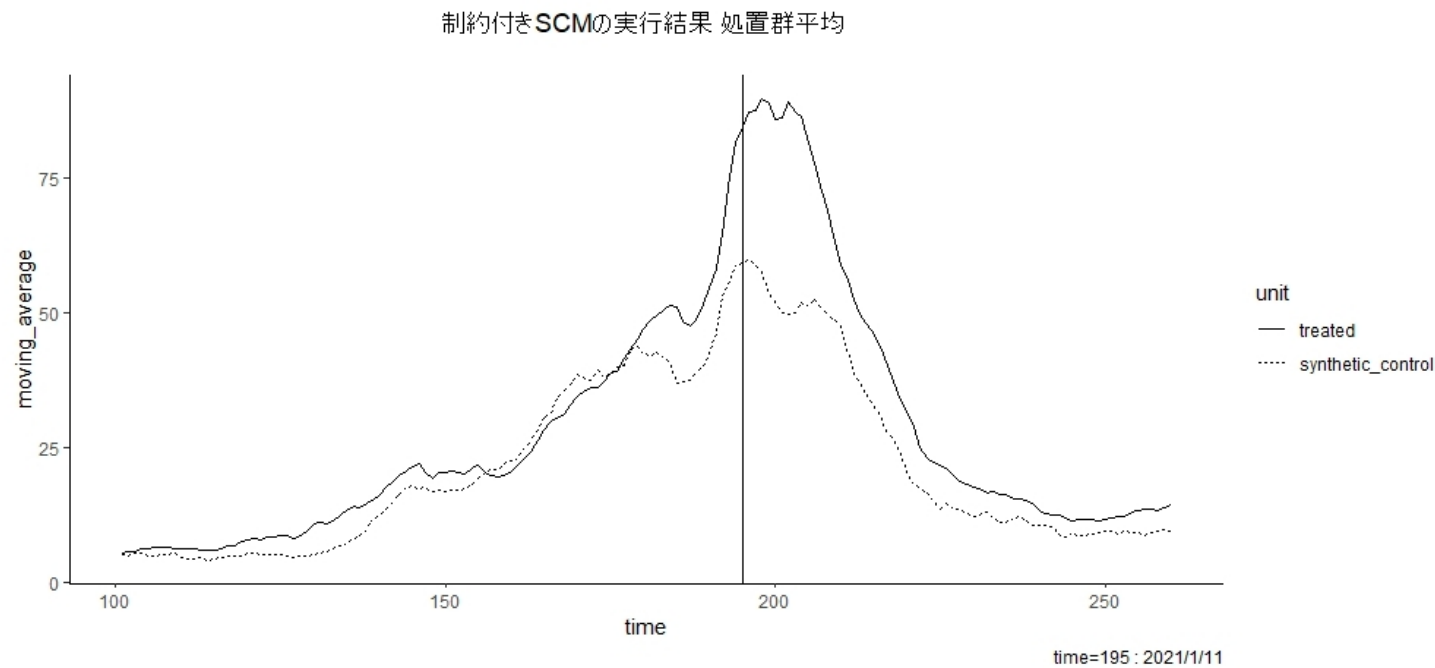
- 線形回帰分析
- ロジスティック回帰分析
- 主成分分析

- 秋学期

- 変数選択(forward, backward)
- 正則化回帰(ridge, lasso)
- **Synthetic control method**

どんなことを扱うゼミ？

- データ解析の実践 (2021年三田祭の内容)
 - 成人式開催によるCOVID-19感染拡大への影響を研究
 - 制約付きSynthetic control methodを新たに提案し解析



どんなことを扱うゼミ？

- 卒業研究のテーマを一部抜粋

- バランシングによるSynthetic control methodの改善
- 一般化線形混合モデルによる野球選手年棒の分析
- 機械学習を使ったリズムゲームの自動楽譜作成
- 最小角回帰についての考察
- Staggered Adoption を考慮したSCMの十分条件について
- 強化学習によるテトリスプレイとその評価
- パネルデータにおけるスパースモデリング
- 機械学習を使った三味線波形データの解析

具体的な活動内容

- **輪読ゼミ (月曜4限, 3年生)**
 - 基本的に3年生が中心
 - 統計科学の手法を演習 (片山, 反転学習)
- **全体ゼミ (月曜5限, 3年生+4年生)**
 - 卒業研究進捗報告
- **サブゼミ (任意)**
 - プログラミング言語Julia, 深層学習, 強化学習, 最適輸送
- **個別ゼミ (任意)**



募集要項

- ・ **こちらからのお願い**

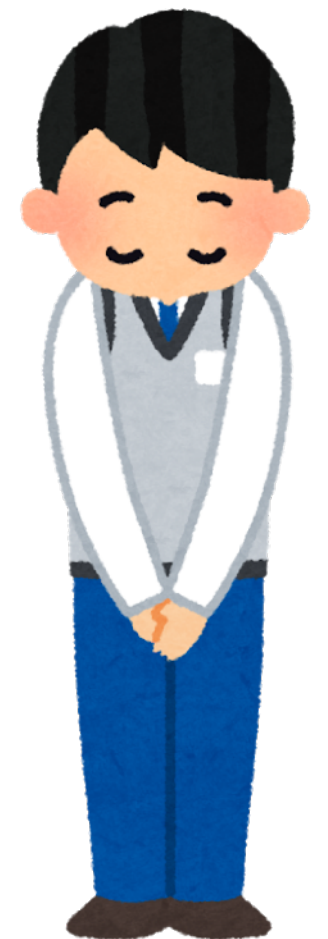
- ・ 考えること・議論が苦ではない方お待ちしております
- ・ 能動的な活動を望みます

- ・ **募集人数**

- ・ A・B日程合わせて**5名前後** (ご容赦下さい...)

- ・ **選考内容**

- ・ 事前課題, 面接, エントリーシート, 成績表
- ・ 詳細は追って提示します



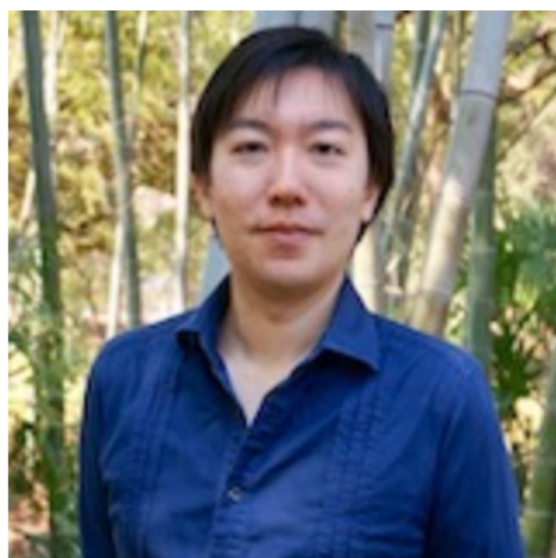
ホームページ

🔍 K-RIS 片山翔太



片山 翔太 (カタヤマ ショウタ)

Katayama, Shota



所属 (所属キャンパス)

経済学部 (三田)

職名

准教授

メールアドレス

katayama@econ.keio.ac.jp

HP

Personal Laboratory

勉強しておいた方がよいこと

線型代数学

行列

線型空間

固有値

解析学

極限

偏微分

重積分

最適化

凸関数

KKT条件

関連する経済学部授業

- 日吉

- 統計科学・機械学習に必要な数学

- 微分積分入門, 微分積分, 線形代数, 線形代数続論
 - 解析学入門, 確率論入門

- 最適化

- 経済数学

- 三田

- 実証研究

- 計量経済学中級, 時系列分析, ベイズ統計学

- 数理統計入門

- 確率統計, 解析学, 人工知能入門