

# 構造推定入門: 演習課題

## 応用社会科学RAブートキャンプ<sup>o</sup>

講師: 遠山祐太 (早稲田大学)

最終更新: 2023-08-30 02:04:31

# 演習課題

# はじめに

- 前半の講義内容に基づいた演習問題をいくつか紹介します。
- 自分で解けそう & 気になった問題から取り掛かってください。
- 注1：問題の指示が若干ふわっとしているところがあります。
  - 曖昧な箇所はご自身で考えて回答してください。
- 注2：模範解答は特に用意していません
  - **実際のRA Workに「模範解答」はありません！**
  - 皆さんと同時に、私も手元で解いていきます。
- **質問大歓迎！！**
  - 今夜は一泊する予定ですので、授業後・夕食後などにもお気軽にどうぞ。

# 課題 1 : 利潤最大化価格

- きのこの山とたけのこの里の限界費用（単位費用）を120円とする。
- 講義パートにおけるデータ&ランダム係数ロジットの推定結果に基づいて、利潤最大化するような価格について検討しよう。
- 問1：たけのこの里の価格を200円に固定したとき、きのこの山の利潤が最大になるようなきのこの山の価格を求めよ。
- 問2：きのこの山の価格を200円に固定したとき、たけのこの里の利潤が最大になるようなたけのこの里の価格を求めよ。
- 【難しめ】問3：きのこの山・たけのこの里の価格を同時に動かせるとき、きのこの山とたけのこの里の**合計利潤**が最大になるような価格を求めよ。
- ヒント：利潤を計算する関数を定義し、それについて数値最適化しよう。

## 課題 2 : RA ブートキャンプのサーベイデータ

- 皆さんに到着初日に回答してもらったきのこ・たけのこサーベイの結果が `data/KTSurvey_RABootcamp2023_cleaned.csv` にあります。
- このデータを用いて、講義パートで行った分析(`mlogit`での推定、フルスクラッチでの推定)を実行せよ。また、結果の差について検討 & 議論せよ。
- コメント : データは最低限のクリーニングをしたもの。
  - 変数名は講義パートのデータ(`data/KTSurvey_Spring2023_cleaned.csv`)と同じ
  - 変数の中身は、講義パートとは異なっている。
  - データ共有についての不同意者の回答を落としている。

## 課題3：個人属性の影響

- サーベイでは消費者属性（性別、年齢、出身地、などなど）についても聞いている。
- きのこの山・たけのこの里への選好が、消費者属性に依存するようなモデルを考えて推定せよ。
- 例として：

$$\begin{aligned}U_{i,k,Kinoko} &= \alpha_{Kinoko} + \gamma_{Kinoko}^{age} \cdot age_i - \beta \cdot p_{k,Kinoko} + \epsilon_{i,k,Kinoko} \\U_{i,k,Takenoko} &= \alpha_{Takenoko} + \gamma_{Takenoko}^{age} \cdot age_i - \beta \cdot p_{k,Takenoko} + \epsilon_{i,k,Takenoko} \\U_{i,k,outside} &= \epsilon_{i,k,outside}\end{aligned}$$

- ヒント1：`mlogit`で推定する前に、消費者属性と選択パターンの相関を見てみましょう。

- ヒント2 : `mlogit` の [解説](#) が参考に。
- 上の例を一部改造 :

```
library("mlogit")
```

```
## Warning: package 'mlogit' was built under R version 4.0.5
```

```
## Warning: package 'dfidx' was built under R version 4.0.5
```

```
data("ModeCanada", package = "mlogit")
MC <- dfidx(ModeCanada, subset = noalt == 4)
ml.MC1 <- mlogit(choice ~ cost + freq | income | 0, MC,
                 alt.subset = c("car", "train", "air"), reflevel = "car")
```

- この例では
  - 交通手段選択 : Train, Air, Car。なお、Carを基準オプションとしている。
  - 選択肢ごとに変わる変数として、`cost, freq, ovt`
  - 消費者属性は`income`。この属性が、選択肢によって影響が異なる（つまり係数が異なる）。

```
summary(ml.MC1)
```

```
##
## Call:
## mlogit(formula = choice ~ cost + freq | income | 0, data = MC,
##       alt.subset = c("car", "train", "air"), reflevel = "car",
##       method = "nr")
##
## Frequencies of alternatives:choice
##      car   train    air
## 0.45757 0.16721 0.37523
##
## nr method
## 5 iterations, 0h:0m:0s
## g'(-H)^-1g = 0.000617
## successive function values within tolerance limits
##
## Coefficients :
##              Estimate Std. Error  z-value  Pr(>|z|)
## (Intercept):train -1.4543116   0.1695671  -8.5766 < 2.2e-16 ***
## (Intercept):air   1.3797110   0.3574951   3.8594 0.0001137 ***
## cost              -0.0643521   0.0032226 -19.9691 < 2.2e-16 ***
## freq              0.1003497   0.0040576  24.7311 < 2.2e-16 ***
## income:train      -0.0088747   0.0031338  -2.8320 0.0046263 **
## income:air         0.0267426   0.0033714   7.9321 2.22e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



## 【難】 課題 4 : ランダム係数ロジットモデル

- (これは今日やる課題ではなく、ブートキャンプ後の発展課題として考えてください。)
- ランダム係数ロジットモデルについて、`mlogit`パッケージを使わずに、フルスクラッチで推定せよ。
- 尤度関数の定義については補足資料を参照。また、Train Chapter 6も参考に。