

# 絶滅危惧種保護法の評価

今村隼人

# 絶滅危惧種保護法

- 1973 年 アメリカで**絶滅危惧種保護法**が制定される (Endangered Species Act: **ESA**)
- 絶滅危惧種の捕獲や生息地の土地開発に対する**罰則**を定めた法律
- ESA の有効性については**様々な問題点**が指摘されていた

# ESA が開発を促進している

- 絶滅危惧種の多くは **私有地** に生息している
- 開発規制はあるが、土地所有者への **十分な補償がない**
- 結果、所有地に「逆のインセンティブ」が働く
- 「規制される前に開発してしまおう」という **駆け込み開発** が発生

# 分析方法：Ferraro et al. (2007)

- データと結果変数  $Y_i$ 
  - アメリカ国内の絶滅の恐れのある生物  $i$  の **絶滅危険度指数**
- トリートメント変数  $D_i$  と比較対象
  1. リストへの **指定のみ**
  2. 指定 + **高額** 助成金
  3. 指定 + **少額** 助成金
- **コントロール群：ESA 未指定**

# 分析方法：Ferraro et al. (2007) (ii)

## – 推定手法：なぜマッチング法か？

- もし割り当てが完全ランダムなら、平均比較でよい
- しかしトリートメントの定義上、その仮定は正しくない (例: 危険度が  
高い種ほど指定されやすい等のバイアス)
- ⇒ したがって **マッチング法** を用いるのが適切

# 推定結果

	指定のみ (1)	高額助成 (2)	少額助成 (3)
標準化ユークリッド距離	-0.019 (0.839)	<b>0.454</b> (0.001)	<b>-0.213</b> (0.027)
マハラノビス	-0.019 (0.823)	<b>0.409</b> (0.001)	<b>-0.181</b> (0.047)
観測数	430	329	396

## 推定結果 (ii)

- (2)のみが **有意に正** (=危険度が改善)
- (3)は有意に負、(1)は効果なし、または負の傾向

**考察:** リストの拡大（指定）だけでは不十分。開発規制に対する補償など、**十分な助成金の拠出** こそが重要である。

# Rによるデータ演習



# 演習内容

企業で社員の能力向上のためのプログラムを実施。参加は任意。研修に参加することの賃金への因果的効果 (ATET) を調べる。

## 4 つの変数に注目:

**wagea** 研修期間後の賃金 (単位: 万円)

**D** 研修参加の有無 (参加=1, 不参加=0)

**years** 経験年数

**wageb** 研修以前の賃金 (単位: 万円)

# トリートメント変数による平均値の差

単純な平均比較を行う：

```
wagedata %>%  
  summarise(mean.treated = mean(wagea[D==1]),  
            mean.controlled = mean(wagea[D==0]),  
            mean.diff = mean.treated - mean.controlled)
```

```
# A tibble: 1 × 3  
  mean.treated mean.controlled mean.diff  
    <dbl>         <dbl>         <dbl>  
1      25.7         26.8         -1.13
```

→ 研修に参加すると、逆に賃金が 1.13 万円下がっている？

# 研修参加と共変量の相関

なぜ下がったのか？ 参加有無  $D$  と、経験年数・研修前賃金の相関を見る。

```
wagedata %>%  
  summarise(cor_years = cor(D, years),  
            cor_wageb = cor(D, wageb))
```

```
# A tibble: 1 × 2  
  cor_years cor_wageb  
    <dbl>    <dbl>  
1   -0.191   -0.184
```

→ どちらも **負の値** となっている。

# 分析結果の解釈

- **事実 1:** 単純比較では、研修によって収入が減っている。
- **事実 2:** 研修参加  $D$  と「勤続年数」「研修前の収入」の相関は **負**。
  - 「元々賃金が低く、経験が浅い人」ほど研修に参加していた。
  - 研修への参加と潜在的な賃金がランダムであるという仮定を満たしていない（セレクションバイアスがある）。

したがって、共変量を調整できる **マッチング法を用いるのが適切である。**

# マッチング法の実施方法

R の MatchIt パッケージを用いてマッチング法を実装する。 `matchit()` 関数により、傾向スコアの推定とマッチングを同時に行える。

```
library(MatchIt)
m.out <- matchit(D ~ years + wageb,
  data = wagedata,
  method = "nearest",
  distance = "scaled_euclidean",
  replace = TRUE)

m.out %>% summary()
```

# 推定結果 (1): マッチング前

Summary of Balance for All Data:

	Means Treated	Means Control	Std. Mean Diff.	Var. Ratio	eCDF Mean
years	8.2311	10.9377	-0.4970	0.6556	0.1044
wageb	24.4874	26.6637	-0.4734	0.6734	0.0989
years	0.1722				
wageb	0.2094				

- **Mean Diff:** 各グループの平均値の差を、トリートメントグループの標準偏差で割った値（標準化平均差）。
- **現状:** Mean Diff の絶対値が大きい（約 0.5）ため、**共変量はバランスできていない。**

## 推定結果 (2): マッチング後

Summary of Balance for Matched Data:

	Means Treated	Means Control	Std. Mean Diff.	Var. Ratio	eCDF Mean
years	8.2311	8.1933	0.0069	1.0150	0.0044
wageb	24.4874	24.4958	-0.0018	0.9934	0.0008
years	0.0210	0.0239			
wageb	0.0042	0.0037			

- **改善**: Mean Diff の値が 0 に近づいており、バランスが改善した。
- **結論**: 両グループは、トリートメントを受けたかどうか以外は **平均的に同じ性質を持つ** といえる。

# 図表の作成（ラブプロット）

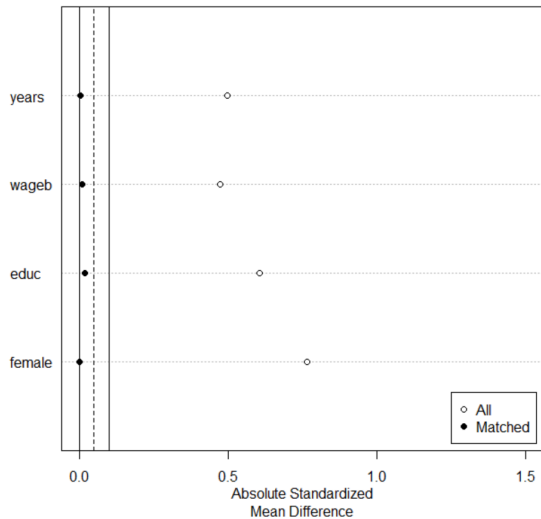
共変量のバランスの変化を可視化するために **ラブプロット** (Love Plot) を作成する。

```
# 共変量を増やして再度マッチング
m.out_full <- matchit(D ~ years + wageb + educ + female,
                      data = wagedata,
                      method = "nearest",
                      distance = "scaled_euclidean",
                      replace = TRUE)

# summary() の結果を plot() に渡すだけ
m.out_full %>%
  summary() %>%
  plot(xlim = c(0, 1.5))
```



## 図表の作成（ラブプロット） (ii)



## 図表の作成（ラブプロット） (iii)

- 縦軸に共変量、横軸に **標準化された平均の差の絶対値 (Mean Diff)** をプロットする。
- マッチング前後（Before/After）でのバランスの改善を一目で確認できる。
- 共変量が多い場合にマッチングがうまくいっているか確認するのに便利

# 実際に回帰を行う

マッチング済みのデータ `match.data(m.out)` を用いて、重み付き回帰分析を行う。

```
library(broom)

match.data(m.out) %>%
  lm(wagea ~ D, data = ., weights = weights) %>%
  tidy()
```

# 推定結果 (Scaled Euclidean)

```
# A tibble: 2 × 5
  term          estimate std.error statistic    p.value
<chr>         <dbl>     <dbl>     <dbl>    <dbl>
1 (Intercept)    24.7       0.367     67.1 1.01e-216
2 D              1.07       0.472     2.26 2.43e- 2
```

- $D$  の係数が **1.07** となっている。
- **結論:** 「研修プログラムに参加することの賃金への因果的効果 (ATET) は **およそ 1 万円** である」という結果が得られた。

# 傾向スコアによるマッチング

距離尺度を `scaled_euclidean` から `glm` (傾向スコア) に変更して比較する。

```
m.out <- matchit(D ~ years + wageb,  
  data = wagedata,  
  method = "nearest",  
  distance = "glm", # 傾向スコア  
  replace = TRUE)  
  
m.out %>%  
  match.data() %>%  
  lm(wagea ~ D, data = ., weights = weights) %>%  
  tidy()
```

# 推定結果 (傾向スコア)

```
# A tibble: 2 × 5
  term          estimate std.error statistic    p.value
<chr>         <dbl>     <dbl>     <dbl>    <dbl>
1 (Intercept)    25.1       0.359      70.0 2.25e-229
2 D              0.622       0.471      1.32 1.87e- 1
```

- ATET が **0.622** となっており、標準化ユークリッド距離の場合よりも推定値が減少した。

## 推定結果 (傾向スコア) (ii)

**考察:** `summary(m.out)` でバランスを確認すると、標準化平均差 (Mean Diff) が大きくなっており、標準化ユークリッド距離を用いた場合よりもマッチングの質が落ちていることがわかった。

どのマッチング法がうまくいくかはケースバイケース。 **最も共変量のバランスがよいマッチングを採用すべきである。**