Shapley Decomposition

European Doctoral School of Demography

Benjamin Elbers

2025-05-24

Overview

1. Introduction to decomposition
2. Shapley decomposition
3. Applications in demography
4. Advanced Topics: Many factors and hierarchical factors

Organization

- GitHub repository
- Exercise sheet available there as RMarkdown file
- Additional exercises to try later

Part 1: Introduction to Decomposition

Why decomposition?

- Loosely defined: understand a statistic (often a difference) through its constituting factors
- In *demography*: understanding the differences (e.g., in terms of death rates, fertility behavior) between two populations—either over time or across areas
- In *machine learning*: understanding the prediction of a complex model
- In *statistics*: splitting the variance into a within and a between part: $\text{Var}[X] = \text{E}[\text{Var}[X \mid Y]] + \text{Var}[\text{E}[X \mid Y]]$
- In *sociology*: understand the factors that contribute to an increase/decrease in segregation

A classic

Kitagawa (1955): "When comparing the incidence of some phenomenon in two or more groups, social researchers place much emphasis on the need for holding constant those related factors that would tend to distort the comparison. For example, before comparing the death rates for the residents of two areas, demographers frequently control the factors of differences between the areas in age, sex and race composition."

Example data (Clogg and Eliason 1998)

"Would you like to have another child?", by no. of children

Age Group	4+ children		One child	
	N	%	N	%
20 to 24	27	37.0	363	90.1
25 to 29	152	19.1	208	76.9
30 to 34	224	15.2	96	56.2
35 to 39	239	5.0	59	20.3
40 to 44	211	6.2	48	10.4
Overall	**853**	**11.5**	**774**	**72.1**

- The overall rate is a function of both the age distribution and the age-specific rates.

Notation

Let's introduce some notation:

- y_{ij} is the rate for the ith group in the jth population,
- n_{ij} is the size of the ith group in the jth population, and
- $N_j = \sum_i n_{ij}$ is the size of population j.

Notation (continued)

Age Group	4+ children		One child	
	N	%	N	%
20 to 24	n_{11}	y_{11}	n_{12}	y_{12}
25 to 29	n_{21}	y_{21}	n_{22}	y_{22}
30 to 34	n_{31}	y_{31}	n_{32}	y_{32}
35 to 39	n_{41}	y_{41}	n_{42}	y_{42}
40 to 44	n_{51}	y_{51}	n_{52}	y_{52}
Overall	N_1	$\bar{y}_1$	N_2	$\bar{y}_2$

- Let $p_{ij} = n_{ij}/N_j$ such that $\sum_i p_{ij} = 1$.
- $\bar{y}_j$ is the overall rate in population j.
- $\bar{y}_1 = \sum_{i=1}^{5} p_{i1} y_{i1}$, $\bar{y}_2 = \sum_{i=1}^{5} p_{i2} y_{i2}$

Standardization

- One approach is standardization: Remove the compositional effect by using the identical distribution for both overall rates.
- "What would the overall rate of the women with one child be if this population had the same distribution as the women with 4+ children?"
  - Answer: $\bar{y}_{\text{standardized(1)}} = \sum_{i=1}^{5} p_{i1} y_{i2}$
- "What would the overall rate of the women with 4+ children be if this population had the same distribution as the women with one child?"
  - Answer: $\bar{y}_{\text{standardized(2)}} = \sum_{i=1}^{5} p_{i2} y_{i1}$

Standardization results

Age Group	4+ children		One child		Difference
	N	%	N	%	
20 to 24	27	37.0	363	90.1	
25 to 29	152	19.1	208	76.9	
30 to 34	224	15.2	96	56.2	
35 to 39	239	5.0	59	20.3	
40 to 44	211	6.2	48	10.4	
Overall	**853**	**11.5**	**774**	**72.1**	**60.6**
Standardized(1)		**11.5**		**39.6**	**28.1**
Standardized(2)		**25.1**		**72.1**	**47.0**

Standardization (continued)

- Standardization requires a “reference” distribution—this can also be an artificial distribution.
- For instance, we could use

$$\hat{p}_i = \frac{p_{i1} + p_{i2}}{2}$$

  as the identical reference distribution for both populations.
- Different reference distributions will give different answers.

Decomposition

- While standardization is useful, Kitagawa (1955) introduced an even more useful idea: **decomposition**.
- The idea is to decompose the difference into two parts:

$$\bar{y}_2 - \bar{y}_1 = \underbrace{(\text{differences in composition})}_{\Delta \mathrm{p}} + \underbrace{(\text{differences in rates})}_{\Delta y}$$

Kitagawa (1955)

She proposed

$$\Delta \mathsf{p} = \sum_i \frac{y_{i1} + y_{i2}}{2} \left(p_{i2} - p_{i1}\right)$$

$$\Delta \mathsf{y} = \sum_i \frac{p_{i1} + p_{i2}}{2} \left(y_{i2} - y_{i1}\right)$$

such that

$$\bar{y}_2 - \bar{y}_1 = \Delta \mathsf{p} + \Delta \mathsf{y}$$

Exercises: Part 1

1. Proof $\bar{y}_2 - \bar{y}_1 = \Delta\text{p} + \Delta\text{y}$.
2. How does the term Δy relate to standardization?
3. In R, use the `children` dataset to decompose the difference in overall rates using Kitagawa's method.

Part 2: Shapley Decomposition

Decomposition challenges

- Decomposition techniques are widely used in demography
- Kitagawa's case was simple: just two populations and one 'factor' (age distribution), but generalizations exist
- Problem: The algebra can start getting somewhat involved

$$A_{ijklmn} = \left(\frac{N_{ijklmn}}{N_{.jklmn}}\right)^{\frac{1}{6}} \cdot \left(\frac{N_{ijklm.}}{N_{.jklm.}} \cdot \frac{N_{ijkl.n}}{N_{.jkl.n}} \cdot \frac{N_{ijk.mn}}{N_{.jk.mn}} \cdot \frac{N_{ij.lmn}}{N_{.j.lmn}} \cdot \frac{N_{i.klmn}}{N_{..klmn}}\right)^{\frac{1}{30}} \cdot$$

$$\left(\frac{N_{ijkl..}}{N_{.jkl..}} \cdot \frac{N_{ijk.m.}}{N_{.jk.m.}} \cdot \frac{N_{ijk..n}}{N_{.jk..n}} \cdot \frac{N_{ij.lm.}}{N_{.j.lm.}} \cdot \frac{N_{ij.l.n}}{N_{.j.l.n}} \cdot \frac{N_{ij..mn}}{N_{.j..mn}} \cdot \frac{N_{i.klm.}}{N_{..klm.}} \cdot \frac{N_{i.kl.n}}{N_{..kl.n}} \cdot \frac{N_{i.k.mn}}{N_{..k.mn}} \cdot \frac{N_{i..lmn}}{N_{...lmn}}\right)^{\frac{1}{60}} \cdot$$

$$\left(\frac{N_{i...mn}}{N_{....mn}} \cdot \frac{N_{i..l.n}}{N_{...l.n}} \cdot \frac{N_{i..lm}}{N_{...lm}} \cdot \frac{N_{i.k..n}}{N_{..k..n}} \cdot \frac{N_{i.k.m.}}{N_{..k.m.}} \cdot \frac{N_{i.kl..}}{N_{..kl..}} \cdot \frac{N_{ij...n}}{N_{.j...n}} \cdot \frac{N_{ij..m.}}{N_{.j..m.}} \cdot \frac{N_{ij.l..}}{N_{.j.l..}} \cdot \frac{N_{ijk...}}{N_{.jk...}}\right)^{\frac{1}{60}} \cdot$$

$$\left(\frac{N_{i....n}}{N_{.....n}} \cdot \frac{N_{i...m.}}{N_{....m.}} \cdot \frac{N_{i..l..}}{N_{...l..}} \cdot \frac{N_{i.k...}}{N_{..k...}} \cdot \frac{N_{ij....}}{N_{.j....}}\right)^{\frac{1}{30}} \cdot \left(\frac{N_{i.....}}{N_{......}}\right)^{\frac{1}{6}} .$$

Decomposition challenges (continued)

- Usual decomposition procedure: Define your statistic of interest, then creatively try to find algebraic expressions that "isolate" certain factors
- Often, this will be complex or impossible with (a) many factors, (b) if the statistic of interest is a non-linear function of the factors.
    - Generally, if the statistic of interest is a linear function, the decomposition is trivial.
- Some decompositions results include a "residual" term that is hard to interpret.

Shapley decomposition

- The Shapley decomposition is a *principled* way of defining a decomposition in a framework of counterfactuals—there is some arbitrary element to every decomposition.
- Standardization is a form of counterfactual thinking: What would population A look like if it had population's B age structure?
- The statistic can be an arbitrarily complex function of the factors—the Shapley decomposition will never fail to decompose the statistic.

Shapley values

- A concept from game theory, introduced by Lloyd Shapley (1953)
- **Example**: Define the following 'game' of three players. Players 1 and 2 supply right-hand gloves, while Player 3 supplies a left-hand glove. Any player can enter a coalition or not. The game is only successful if players with both types of gloves enter into a coalition.
- Clearly, player 3 supplies more value because they are the only one supplying a left-hand glove—but how much more value compared to players 1 and 2?

Shapley values (continued)

- Let $N = \{1, 2, 3\}$ (the three players)
- The value function $v(\cdot)$ takes a set of players C ("coalition"), and outputs either zero or one, depending on whether the coalition is successful.
- Examples:
  - $v(\emptyset) = 0$
  - $v(\{1\}) = 0$
  - $v(\{3\}) = 0$
  - $v(\{1, 2\}) = 0$
  - $v(\{1, 3\}) = 1$
  - $v(\{1, 2, 3\}) = 1$

Shapley values (continued)

- The fully-defined value function for this game is

$$v(C) = \begin{cases} 1 & \text{if } |C| \geq 2 \text{ and } 3 \in C. \\ 0 & \text{otherwise} \end{cases}$$

  where C is the set of players that entered the coalition.

- In words: At least 2 players need to enter the coalition, and player 3 needs to be part of it.

Shapley value formula

- The Shapley value for each player k is defined as

$$\varphi_k(N,v) = \sum_{S \subseteq N \setminus \{k\}} \frac{|S|!(m-1-|S|)!}{m!} \left[v(S \cup \{k\}) - v(S)\right]$$

  where m is the number of players.
- The notation $S \subseteq N \setminus \{k\}$ means: all possible subsets of N excluding factor k (**power set** of $N \setminus \{k\}$)
- In words, the formula says: For all possible coalitions (subsets) that exist, compute the contribution of player k, and take a weighted average of the contribution of k.
- The difference $v(S \cup \{k\}) - v(S)$ asks: what difference does it make to add player k?

Shapley values calculation

- Let's calculate the Shapley value of player 1
- Good to know: A set of cardinality p has a power set of 2^p elements
- For player 1 ($k = 1$):

$S \subseteq N \setminus \{k\}$	$\frac{\lvert S\rvert!(m-1-\lvert S\rvert)!}{m!}$	$v(S \cup \{k\})$	$v(S)$	$v(S \cup \{k\}) - v(S)$
$\{\emptyset\}$				
$\{2\}$				
$\{3\}$				
$\{2, 3\}$				

Shapley values calculation

Let's calculate the Shapley value of player 1:

$S \subseteq N \setminus \{k\}$	$\frac{\lvert S\rvert!(m-1-\lvert S\rvert)!}{m!}$	$v(S \cup \{k\})$	$v(S)$	$v(S \cup \{k\}) - v(S)$
$\{\emptyset\}$	$\frac{0!(3-1-0)!}{3!} = \frac{1}{3}$	$v(\{1\}) = 0$	$v(\{\emptyset\}) = 0$	0
$\{2\}$	$\frac{1!(3-1-1)!}{3!} = \frac{1}{6}$	$v(\{1,2\}) = 0$	$v(\{2\}) = 0$	0
$\{3\}$	$\frac{1!(3-1-1)!}{3!} = \frac{1}{6}$	$v(\{1,3\}) = 1$	$v(\{3\}) = 0$	1
$\{2,3\}$	$\frac{2!(3-1-2)!}{3!} = \frac{1}{3}$	$v(\{1,2,3\}) = 1$	$v(\{2,3\}) = 1$	0

- Player 1 is only helpful in *one* situation
- Shapley value is $\varphi_1 = \frac{1}{3} \times 0 + \frac{1}{6} \times 0 + \frac{1}{6} \times 1 + \frac{1}{3} \times 0 = \frac{1}{6}$

Shapley values results

- $\varphi_1 = \frac{1}{6}$, $\varphi_2 = \frac{1}{6}$, $\varphi_3 = \frac{2}{3}$
- Note that
  $$\sum_k \varphi_k = 1,$$
  as
  $$v(\{1,2,3\}) - v(\{\emptyset\}) = 1$$
- The Shapley value has many desirable properties—which we won't go into.

R Example 1

```r
glove <- function(factors) {
  if (length(factors) > 1 & 3 %in% factors) return (1)
  return (0)
}

glove(c())
```

```
#> [1] 0
```

```r
glove(c(1))
```

```
#> [1] 0
```

```r
glove(c(1, 3))
```

```
#> [1] 1
```

```r
glove(c(1, 2, 3))
```

```
#> [1] 1
```

R Example 1 (continued)

```r
library("shapley")
# install via: remotes::install_github("elbersb/shapley")
shapley(glove, c(1, 2, 3), silent = TRUE)
```

```
#>   factor  value
#> 1      1 0.1667
#> 2      2 0.1667
#> 3      3 0.6667
```

Exercises: Part 2

1. What happens if the value function is additive in the factors?
   - e.g., let $N = \{a, b, c\}$, and $v(N) = a + b + c$, $v(\{a, c\}) = a + c$, $v(\{\emptyset\}) = 0$ etc.

Difficult take-home exercises (if you really want to!)

2. Proof that $v(N) - v(\emptyset) = \varphi_1 + \varphi_2 + \cdots + \varphi_m$. (This is called the **efficiency** property.)
3. Proof $\sum_{S \subseteq N \setminus \{k\}} \frac{|S|!(m-1-|S|)!}{m!} = 1$. (This is to see that $\varphi_i(N, v)$ is truly a weighted average, with the weights summing to 1.)

Part 3: Applications in demography

Back to demography

- What does this have to do with decomposition and demography?
- The Shapley value approach allows us to decompose any statistic (= value function) into additive components.
- Let's say we have factors $N = \{1, 2, \ldots, m\}$, then the Shapley framework gives us:

$$\underbrace{v(N)}_{\text{outcome when all factors are "on"}} - \underbrace{v(\emptyset)}_{\text{when factors are "off"}} = \varphi_1 + \varphi_2 + \cdots + \varphi_m.$$

R Example 2: Kitagawa

- Remember Kitagawa:

$$\bar{y}_2 - \bar{y}_1 = \underbrace{(\text{differences in composition})}_{\Delta \text{p}} + \underbrace{(\text{differences in rates})}_{\Delta y}$$

- We have two factors: Δp and Δy, i.e. $N = \{\Delta p, \Delta y\}$

R Example 2: Kitagawa (continued)

- Define the value function as follows

$$v(C) = \begin{cases} \sum_i p_{i1} y_{i1} & \text{if } C = \{\emptyset\} \\ \sum_i p_{i2} y_{i1} & \text{if } C = \{\Delta p\} \\ \sum_i p_{i1} y_{i2} & \text{if } C = \{\Delta y\} \\ \sum_i p_{i2} y_{i2} & \text{if } C = \{\Delta p, \Delta y\} \end{cases}$$

- Hence $v(\{\emptyset\}) = \bar{y}_1$ and $v(N) = v(\{\Delta p, \Delta y\}) = \bar{y}_2$.
- $v(\{\Delta p\})$ and $v(\{\Delta y\})$ provide counterfactuals

R Example 2: Implementation

```
rates <- function(factors,  data) {
  if ("weight" %in% factors) {
    weights <- data$n2 / sum(data$n2)
  } else {
    weights <- data$n1 / sum(data$n1)
  }
  if ("rate" %in% factors) {
    rate <- data$rate2
  } else {
    rate <- data$rate1
  }
  return(sum(weights * rate))
}
```

R Example 2: Results

```r
rates(c(), children)
```

```
#> [1] 11.49
```

```r
rates(c("weight"), children)
```

```
#> [1] 25.14
```

```r
rates(c("rate"), children)
```

```
#> [1] 39.61
```

```r
rates(c("weight", "rate"), children)
```

```
#> [1] 72.09
```

These numbers are familiar!

R Example 2: Shapley decomposition

```r
shapley(rates,  c("weight", "rate"),
            silent = TRUE, data = children)
```

```
#>   factor value
#> 1 weight 23.07
#> 2   rate 37.53
```

Kitagawa and Shapley

With two factors, the Shapley value function simplifies a lot!

For Δp:

$$\begin{aligned}
&\varphi_{\Delta p}(N, v) \\
&= \sum_{S \subseteq N \setminus \{\Delta p\}} \tfrac{|S|!(2-1-|S|)!}{2!} \left[v(S \cup \{\Delta p\}) - v(S) \right] \\
&= \tfrac{0!(1-0)!}{2} \left[v(\{\Delta p\}) - v(\{\emptyset\}) \right] + \tfrac{1!(1-1)!}{2} \left[v(\{\Delta p, \Delta y\}) - v(\{\Delta y\}) \right] \\
&= \tfrac{1}{2} \left[v(\{\Delta p\}) - v(\{\emptyset\}) \right] + \tfrac{1}{2} \left[v(\{\Delta p, \Delta y\}) - v(\{\Delta y\}) \right]
\end{aligned}$$

Kitagawa and Shapley (continued)

Plug in:

$$
\begin{aligned}
\ldots &= \tfrac{1}{2}\left[\sum_i p_{i2}y_{i1} - \sum_i p_{i1}y_{i1}\right] + \tfrac{1}{2}\left[\sum_i p_{i2}y_{i2} - \sum_i p_{i1}y_{i2}\right] \\
&= \tfrac{1}{2}\sum_i y_{i1}(p_{i2} - p_{i1}) + \tfrac{1}{2}\sum_i y_{i2}(p_{i2} - p_{i1}) \\
&= \sum_i \tfrac{y_{i1}+y_{i2}}{2}\,(p_{i2} - p_{i1})
\end{aligned}
$$

The Shapley approach leads to the same solution as Kitagawa's approach!

The value function

- The value function $v(\cdot)$ is key in the Shapley value approach
- In game theory, the value function is given
- In demography, thinking about the value function is the key step
- When there are m factors, $v(\cdot)$ needs to return something sensible for each combination of the m factors—we need to think about the counterfactuals

R Example 3

Das Gupta (1993) discusses the change in the illegitimacy ratio for white women between 1963 and 1983.

Let α be the proportion unmarried, β the nonmarital fertility rate, and γ the marital fertility rate.

Then

$$R = \frac{\alpha\beta}{\alpha\beta + (1-\alpha)\gamma}$$

Data:

	1963	1983
α	0.296	0.417
β	0.011	0.019
γ	0.139	0.095
R	**0.032**	**0.125**

R Example 3 (continued)

- We have three factors: $N = \{\alpha, \beta, \gamma\}$
- The value function is $v(C) \rightarrow R$, where C can be any element of the powerset of N
- If a factor is absent, we use the 1963 value; if a factor is present, use 1983 value
- Examples:
  - $v(\emptyset)$ should return the R in 1963
  - $v(N)$ should return the R in 1983
  - $v(\{\alpha\})$ should return a 'counterfactual' R, using α from 1983, but β and γ from 1963

Exercise: What is $v(\{\alpha, \gamma\})$? What assumptions are we making here?

R Example 3: Implementation

```r
il <- function(factors = c()) {
  alpha <- ifelse("prop_unmarried" %in% factors, .417, .296)
  beta <- ifelse("nonm_fertility" %in% factors, .019, .011)
  gamma <- ifelse("m_fertility" %in% factors, .095, .139)
  alpha * beta / ((alpha * beta) + (1 - alpha) * gamma)
}
```

R Example 3: Testing

```r
il() # 1963
```

```
#> [1] 0.0322
```

```r
il(c("prop_unmarried", "nonm_fertility")) # mix of 1963 and 1983
```

```
#> [1] 0.08906
```

```r
il(c("m_fertility")) # mix of 1963 and 1983
```

```
#> [1] 0.04642
```

```r
il(c("prop_unmarried", "nonm_fertility", "m_fertility")) # 1983
```

```
#> [1] 0.1252
```

R Example 3: Results

```r
il(c("prop_unmarried", "nonm_fertility", "m_fertility")) - il()
```

```
#> [1] 0.09295
```

```r
shapley(il, c("prop_unmarried", "nonm_fertility", "m_fertility"),
        silent = TRUE)
```

```
#>            factor   value
#> 1 prop_unmarried 0.03378
#> 2 nonm_fertility 0.03471
#> 3    m_fertility 0.02446
```

- The illegitimacy ratio for white women increased by about 9 pp between 1963 and 1983
- What are we concluding from the decomposition?
- Note: The decomposition can also produce negative numbers.

Exercises: Part 3

1. Try the Shapley approach to Kitagawa's decomposition on your own. We defined $v(\{\emptyset\}) = \sum_i p_{i1} y_{i1}$ and $v(\{\Delta p, \Delta y\}) = \sum_i p_{i2} y_{i2}$. What happens when you switch these definitions?

2. Rewrite example 3 such that the value function uses a data frame. Try the decomposition on the `austria_chile` data.

3. Bongaarts (1978) expressed the total fertility rate (TFR) as the product of five factors: the total fecundity rate, and the indices of proportion married, noncontraception, induced abortion, and lactational infecundability. The dataset `bongaarts` contains data for South Korea, 1960 and 1970. Use the Shapley value approach to decompose the difference in the TFR between 1960 and 1970 into the contribution of the five factors.

Part 4: Advanced Topics

Two problems with Shapley decomposition

1. What do we do when we have many factors?
2. What if our factors have a hierarchical ordering?

Exponential growth

- The Shapley approach requires the evaluation of the value function on each of the elements of the power set of N
- e.g., with $N = \{1, 2, 3\}$, we need to compute $v(\emptyset)$, $v(\{1\})$, $v(\{2\})$, $v(\{3\})$, $v(\{1, 2\})$, $v(\{1, 3\})$, $v(\{2, 3\})$, $v(\{1, 2, 3\})$ $(2^3 = 8)$.
- This exponential growth can be a problem when you have many factors and the value function is reasonably complex

Solutions

- Use Julia, C++, ...—R is (often) slow
- Use Owen values if your factors have some kind of hierarchical ordering (see later slides)
- Sample from the power set instead of computing all differences $v(S \cup \{\Delta p\}) - v(S)$ for all S
  - Try `shapley_sampled()` in the package

Sampling Algorithm

The algorithm requires two parameters: t is the minimum number of iterations, and s is the desired standard error. The algorithm to approximate φ_i for a factor i is as follows:

1. Repeat the following steps for $j = 1, 2, ...$
   1.1 Sample an integer w between 0 and $m - 1$.
   1.2 Sample w elements from $N \setminus \{i\}$ without replacement, call the resulting set R.
   1.3 Calculate $\hat{\varphi}_i^j = v(R \cup \{i\}) - v(R)$.
   1.4 If $j > t$, calculate $\hat{s}$, the standard error of the $\hat{\varphi}_i^j$'s, and stop if $s < \hat{s}$.
2. Let $\hat{\varphi}_i \equiv \frac{1}{M} \sum_{j=1}^{M} \hat{\varphi}_i^j$ where M is the number of $\hat{\varphi}_i^j$ sampled.

Example: Decomposing segregation

- Segregation indices measure how unequally different groups are distributed over neighborhoods, schools
  - In the U.S., often racial groups
- Segregation is minimized at 0 when all neighborhoods contain identical group proportions; segregation is maximized at 1 when every neighborhood contains just a single group.
- Why does segregation change? Use Shapley decomposition to understand which groups drive increase/decline in segregation.

Example: Decomposing segregation change

nature cities

Article https://doi.org/10.1038/s44284-024-00032-w

Explaining changes in US residential segregation through patterns of population change

Received: 21 August 2023

Accepted: 11 January 2024

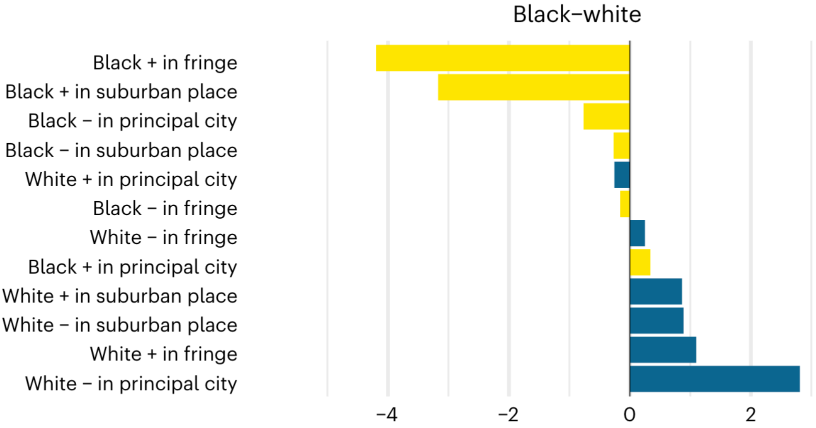
Published online: 09 February 2024

Check for updates

Benjamin Elbers

While overall residential segregation in US cities has declined in the past 30 years—especially between the Black and white populations—relatively little is known about the patterns of population change that caused these

Example: Decomposing segregation



Hierarchical ordering of factors

Often, decomposition factors will have some form of hierarchical ordering:

- e.g., categorical variables in R^2 decomposition
- e.g., individual age aggregates to age groups
- e.g., group-specific factors in segregation decomposition

Example: Extension of Kitagawa

- Kitagawa's decomposition yields two factors: one accounting for differences in composition and one accounting for differences in rates
- With Shapley decomposition, it's easy to extend this decomposition by splitting up the 'rate' factor into as many factors as they are age groups
- Before we had

$$\bar{y}_2 - \bar{y}_1 = \Delta\mathsf{p} + \Delta\mathsf{y}$$

- Now we want

$$\bar{y}_2 - \bar{y}_1 = \Delta p + \Delta y_1 + \Delta y_2 + \cdots + \Delta y_n,$$

  i.e. we have one rate factor for each age group

Hierarchical ordering problem

- Often a problem:

$$\Delta y \neq \Delta y_1 + \Delta y_2 + \cdots + \Delta y_n$$

- The two value functions don't (necessarily) yield "aggregation-consistent" results
- This is a downside of Shapley values

Solutions for hierarchical factors

Two proposals have been made:

1. Nested Shapley values (Sastre and Trannoy 2002)—not recommended
2. Owen values (Owen 1977)

Both also help with easing computational burden

Owen values

- Owen values are a generalization of the Shapley value
- Instead of iterating over all permutations, we take into account the grouping structure
- To factors belonging to other groups, the factors of a particular group can only appear together
- The factors within a group “play” against themselves—in this situation, the other groups are either completely present or completely absent

Example: Owen values

- Let's say we have the grouping structure $G = \{\{1,2\},\{3,4\}\}$
- With four factors, the Shapley decomposition would require $2^3 = 8$ terms for the computation of each factor
- For factor 1:
  1. $v(\{1\}) - v(\emptyset)$
  2. $v(\{1,2\}) - v(\{2\})$
  3. $v(\{1,3\}) - v(\{3\})$
  4. $v(\{1,4\}) - v(\{4\})$
  5. $v(\{1,2,3\}) - v(\{2,3\})$
  6. $v(\{1,2,4\}) - v(\{2,4\})$
  7. $v(\{1,3,4\}) - v(\{3,4\})$
  8. $v(\{1,2,3,4\}) - v(\{2,3,4\})$

Example: Owen values

- Owen values for factor 1:
    1. $v(\{1\}) - v(\emptyset)$
    2. $v(\{1,2\}) - v(\{2\})$
    3. $\cancel{v(\{1,3\}) - v(\{3\})}$
    4. $\cancel{v(\{1,4\}) - v(\{4\})}$
    5. $\cancel{v(\{1,2,3\}) - v(\{2,3\})}$
    6. $\cancel{v(\{1,2,4\}) - v(\{2,4\})}$
    7. $v(\{1,3,4\}) - v(\{3,4\})$
    8. $v(\{1,2,3,4\}) - v(\{2,3,4\})$
- For the computation of factor 1, the other group $\{3,4\}$ is either completely present or completely absent

Example: Extension of Kitagawa implementation

```
# new value function
rates_detailed <- function(factors, data) {
  if ("weight" %in% factors) {
    weights <- data$n2 / sum(data$n2)
  } else {
    weights <- data$n1 / sum(data$n1)
  }

  rate <- data$rate1
  for (i in 1:5) {
    if (paste0("rate", i) %in% factors) {
      rate[i] <- data$rate2[i]
    }
  }
  return(sum(weights * rate))
}
```

Example: Extension of Kitagawa results

```
shapley(rates, c("weight", "rate"),
  silent = TRUE, data = children)
```

```
#>   factor value
#> 1 weight 23.07
#> 2   rate 37.53
```

```
owen(rates_detailed,
  list("weight",c( "rate1", "rate2", "rate3", "rate4", "rate
  silent = TRUE, data = children)
```

```
#>   group factor   value
#> 1     1 weight 23.0717
#> 2     2  rate1 13.2786
#> 3     2  rate2 12.9261
#> 4     2  rate3  7.9397
#> 5     2  rate4  2.7298
#> 6     2  rate5  0.6584
```

Conclusion

Conclusion

- It's all in the counterfactuals: The Shapley value forces you to think about counterfactuals regarding each possible combination of factors
  - Causality: Potential outcomes framework
- Once you have decided on the value function $v(\cdot)$, the Shapley value formula takes care of the rest
- Shapley decomposition provides a linearization—this can be useful or not. By definition, there are no interaction effects/residuals.

Final remarks

- There is a decomposition literature in Economics (look up Blinder–Oaxaca decomposition for applying Kitagawa's decomposition in a regression context)
- Shapley values have recently become very popular in machine learning to understand the outcome of complex, non-linear models
- Some decomposition problems can also be understood in a regression framework—a regression is also a form of standardization

References

- Preston et al. (2001), *Demography*, chapters 2.2 (standardization) and 2.3 (Kitagawa decomposition)
- Kitagawa, Evelyn M. (1955), "Components of a Difference Between Two Rates." *Journal of the American Statistical Association* 50(272): 1168-1194.
- Shorrocks, Anthony F. (2013), "Decomposition procedures for distributional analysis: a unified framework based on the Shapley value." *Journal of Economic Inequality* 11:99–126.