Attribution problems and their applications. A game theoretical approach

Joaquín Sánchez-Soriano

U.I. CIO. Universidad Miguel Hernández de Elche (SPAIN)

November 24, 2020

Kempten Autumn Talks, November 25th, 2020

Table of Contents

Table of Contents

Motivation

Roughly speaking, an attribution problem consists of determining what is the relevance of a series of factors for a particular result to be produced. For example:

- Risk factors in a disease
- Advertising channels to convert ads into purchases
- Museum passes
- Combined tickets in public transport
- Attributes selection in classification problems
- Gene expressions for a disease
- ...

Motivation

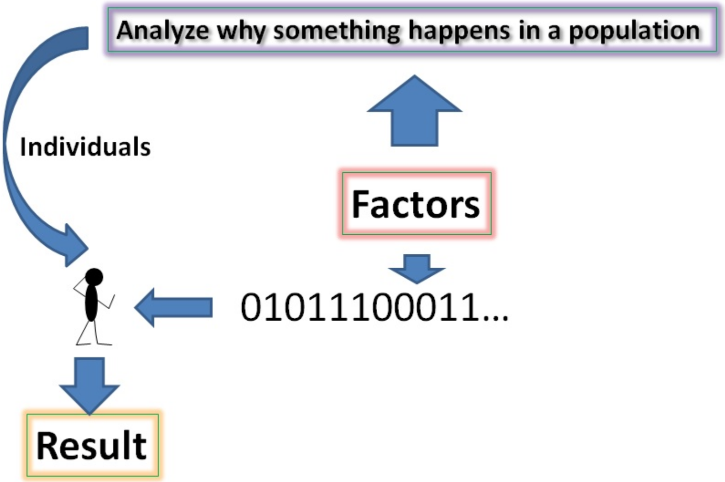

Analyze why something happens in a population
Individuals
Factors
01011100011...
Result

Table of Contents

Basic mathematical model

Basic model. $A = (N, M, F, E)$

- $N = \{1, \ldots, n\}$ is the set of attributes or factors.
- $M = \{1, \ldots, m\}$ is a population
- $F : M \rightarrow N$, such that for each $j \in M$, $F(j) \subset N$ is the set of factors satisfied by individual j
- Obviously, all information about the population M can be summarized in a matrix $|M| \times |N|$, with 0's and 1's entries.

$$E = \begin{pmatrix} e_{11} & \cdots & e_{1n} \\ \vdots & \ddots & \vdots \\ e_{m1} & \cdots & e_{mn} \end{pmatrix}$$

Basic mathematical model

The basic game

How do we associate a game with this problem? The simplest answer is

$$v^A(S) = \sum_{j \in M : F(j) \subset S} 1, \quad S \in N,$$

i.e., counting the number of individuals that have all factors in S.

This game is the well-known museum pass game (Ginsburg and Zang (2003)).

Some applications

The museum pass game (Ginsburng-Zang, 2003)

Several museums in a region (city, country...) collaborate to offer passes which give visitors (tourists as well as residents) unlimited access to visit all collaborating museums during a limited period time. The problem is how to share the net income from the sale of passes among the participating museums.

- $N = \{1, \ldots, n\}$ is the set of collaborating museums.
- $M = \{1, \ldots, m\}$ is the set of visitors
- $F : M \to N$, such that for each $j \in M$, $F(j) \subset N$ is the set of visited museums by user j
- E is the information about all visits to the museums.
- $v^M(S)$ is the net income obtained by museums in S

Some applications

Factors on myocardial infarction (Land-Gefeller, 2000)

Application aimed at quantifying the influence of smoking and three types of cholesterol levels (namely, LDL, HDL, and VLDL) on myocardial infarction. The occurrence of infarction was related to cholesterol levels and smoking habit.

- $N = \{1, \ldots, n\}$ is the set of risk factors.
- $M = \{1, \ldots, m\}$ is the set of infarcted people with at least one risk factors
- $F : M \to N$, $F(j) \subset N$ is the set of risk factors in infarcted individual j
- E information about all infarcted people. Data of 6029 male industrial workers aged 40-60, in 5 year, in Gottingen.
- $v^R(S)$ number of infarcted people with all risk factors in S

Some applications

Microarray technology (Moretti-Patrone-Bonassi, 2007)

Relevance of genes keeping into account their individual behaviors and their interactions when the biological system is studied under a condition of interest (e.g., a disease state, the exposure to environmental or therapeutic agents, etc.).

- $N = \{1, \ldots, n\}$ is the set of genes.
- $M = \{1, \ldots, m\}$ is the set of individuals
- $F : M \to N$, $\emptyset \neq F(j) \subset N$ set of genes with an abnormal expression in individual j
- E information about all individuals (microarray data).
- $v^G(S)$ number of individuals with all genes with abnormal expression in S .

Some applications

Publication credits in bibliometrics (Karpov, 2014)

The publication credit allocation problem is one of the fundamental problems in bibliometrics. The problem is to share "quality" among co-authors.

- $N = \{1, \ldots, n\}$ is the set of authors.
- $M = \{1, \ldots, m\}$ is the set of publications
- $F : M \to N$, $\emptyset \neq F(j) \subset N$ set of co-authors of publication j
- E information about all publications.
- $v^P(S)$ number of publications with all its co-authors in S.

Some applications

Approval voting (Brams-Fishburn, 1978) (Brams-Kilgour,2014)(Ginsburgh-Zang, 2012)

Approval voting is well suited to electing a single candidate but how can it be used to reflect the diversity of interests in the electorate for multiwinner elections?

- $N = \{1, \ldots, n\}$ is the set of candidates.
- $M = \{1, \ldots, m\}$ is the set of voters
- $F : M \to N$, $\emptyset \neq F(j) \subset N$ set of candidates voted by j
- E information about all votes.
- $v^{AV}(S)$ number of votes with all the voted candidates in S.

Some applications

The English after Brexit (Ginsburgh-MorenoTernero-Weber, 2017)

Which are the consequences of Brexit on the ranking of languages in the European Union.

- $N = \{1, \ldots, n\}$ is the set of languages.
- $M = \{1, \ldots, m\}$ is the set of citizens
- $F : M \to N$, $\varnothing \neq F(j) \subset N$ set of languages with a non-zero knowledge spoken by j
- E information about all citizens.
- $v^{AV}(S)$ number of citizens who only have knowledge of languages in S.

Some applications

Traffic in networks (Algaba-Beal-Fragnelli-Llorca-SS, 2019)

Different network operators collaborate to provide a better service and attract more users.

- $N = \{1, \ldots, n\}$ is the set of owner of different parts of the network.
- $M = \{1, \ldots, m\}$ is the set of units of traffic
- $F : M \to N$, $\emptyset \neq F(j) \subset N$ set of parts of the network used by j
- E information about all traffic units.
- $v^T(S)$ number of units of traffic that go only through S.

Some applications

Impact of marketing campaigns (Google)

A company launches a new marketing campaign using different marketing touchpoints or channels (e.g., organic, search, display, e-mail...) and the result of this campaign is the number of conversions (purchases).

- $N = \{1, \ldots, n\}$ is the set of touchpoints or channels.
- $M = \{1, \ldots, m\}$ is the set of clients who purchase
- $F : M \to N$, $\emptyset \neq F(j) \subset N$ set of touchpoints used by client j
- E information about all clients.
- $v^T(S)$ number of conversions through only touchpoints in S.

The basic solution

The Shapley value (Shapley, 1953)

Given (N, v), for each $i \in N$

$$\phi_i(N, v) = \sum_{S \subseteq N \setminus \{i\}} \frac{s!(n-s-1)!}{n!} \left(v(S \cup \{i\}) - v(S) \right),$$

where $s = |S|$ and $n = |N|$.

The basic solution

The Shapley value in the basic game

Given (N, v^M), for each $i \in N$

$$\phi_i(N, v^M) = \sum_{j \in M : i \in F(j)} \frac{1}{|F(j)|}.$$

Therefore, the ratio $\frac{\phi_i(N, v^M)}{|M|}$ reflects the impact of attribute i within the population M.

The Shapley value implies that the value derived from each individual in the population is equally distributed among the factors that this individual has.

Some remarks

- In all cases $F(j) \neq \varnothing$
- Once we have a game defined we can apply any solution in cooperative game, but the Shapley value is usually used.
- Properties of the solution used in each case must be relevant for the problem at hand
- As the games have been defined, in all cases, mainly the relevance of the factors in the problem at hand is measured

Some remarks

- In each problem other elements could be taken into account.
- Other characteristic functions could be defined according to the information available. For example, in the museum pass problem, the prices of the single tickets.
- Other rules could be used. For example, the proportional rule or even any bankruptcy rule.

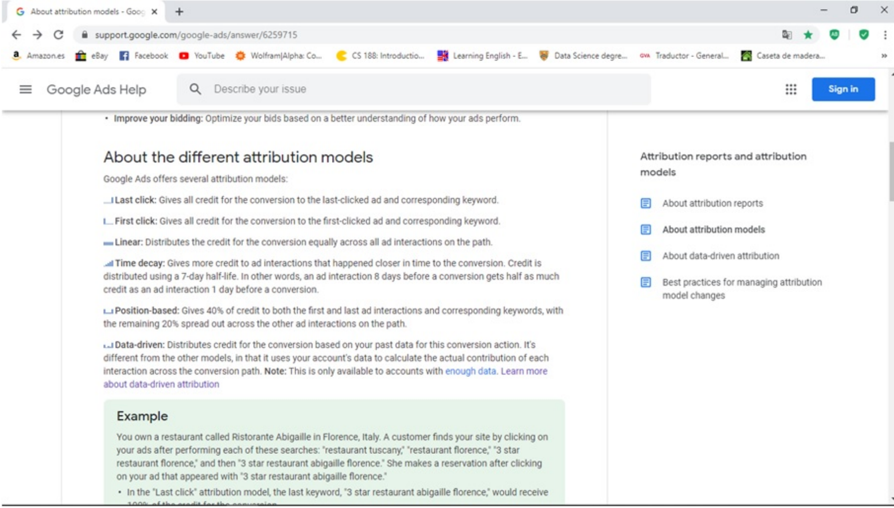
Some remarks

Some asymmetries

In the basic model all attributes are considered symmetric, but possibly in some cases there exist asymmetries. Some asymmetries can arise from:

- Particular characteristics of the attributes.
- The order in which the attributes appear.
- Number of times that an attribute can be present.
- Others

A real example of use



About attribution models - Goog

support.google.com/google-ads/answer/6259715

Amazon.es eBay Facebook YouTube Wolfram|Alpha: Co... CS 188: Introductio... Learning English - E... Data Science degre... Traductor - General... Caseta de madera...

Google Ads Help

Describe your issue

Sign in

- **Improve your bidding:** Optimize your bids based on a better understanding of how your ads perform.

About the different attribution models

Google Ads offers several attribution models:

Last click: Gives all credit for the conversion to the last-clicked ad and corresponding keyword.

First click: Gives all credit for the conversion to the first-clicked ad and corresponding keyword.

Linear: Distributes the credit for the conversion equally across all ad interactions on the path.

Time decay: Gives more credit to ad interactions that happened closer in time to the conversion. Credit is distributed using a 7-day half-life. In other words, an ad interaction 8 days before a conversion gets half as much credit as an ad interaction 1 day before a conversion.

Position-based: Gives 40% of credit to both the first and last ad interactions and corresponding keywords, with the remaining 20% spread out across the other ad interactions on the path.

Data-driven: Distributes credit for the conversion based on your past data for this conversion action. It's different from the other models, in that it uses your account's data to calculate the actual contribution of each interaction across the conversion path. **Note:** This is only available to accounts with enough data. Learn more about data-driven attribution

Example

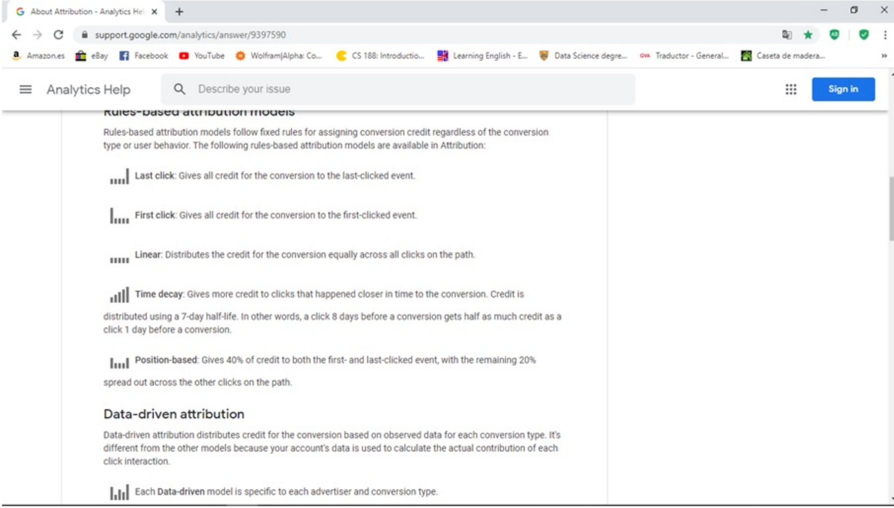
You own a restaurant called Ristorante Abigaille in Florence, Italy. A customer finds your site by clicking on your ads after performing each of these searches: "restaurant tuscany," "restaurant florence," "3 star restaurant florence," and then "3 star restaurant abigaille florence." She makes a reservation after clicking on your ad that appeared with "3 star restaurant abigaille florence."

- In the "Last click" attribution model, the last keyword, "3 star restaurant abigaille florence," would receive

Attribution reports and attribution models

- About attribution reports
- **About attribution models**
- About data-driven attribution
- Best practices for managing attribution model changes

A real example of use



Rules-based attribution models

Rules-based attribution models follow fixed rules for assigning conversion credit regardless of the conversion type or user behavior. The following rules-based attribution models are available in Attribution:

Last click: Gives all credit for the conversion to the last-clicked event.

First click: Gives all credit for the conversion to the first-clicked event.

Linear: Distributes the credit for the conversion equally across all clicks on the path.

Time decay: Gives more credit to clicks that happened closer in time to the conversion. Credit is distributed using a 7-day half-life. In other words, a click 8 days before a conversion gets half as much credit as a click 1 day before a conversion.

Position-based: Gives 40% of credit to both the first- and last-clicked event, with the remaining 20% spread out across the other clicks on the path.

Data-driven attribution

Data-driven attribution distributes credit for the conversion based on observed data for each conversion type. It's different from the other models because your account's data is used to calculate the actual contribution of each click interaction.

Each **Data-driven** model is specific to each advertiser and conversion type.

A real example of use

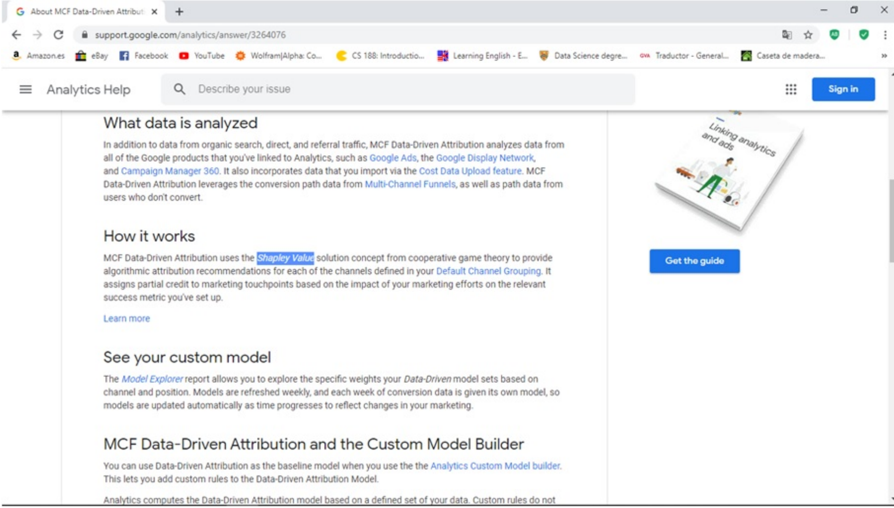


Table of Contents

Literature review 1

- Land-Gefeller 00: Myocardial infarction & Risk & Sh
- Gisnburng-Zang 03: Museum pass & Basic model & Sh
- Gisnburng-Zang 04: Museum pass & Basic model-prices & Sh-Others & Properties
- Naor 05: Sharing cars & Basic model & Sh & Charaterization
- Moretti-Patrone-Bonassi 07: Microarray technology & Basic model & Sh & Characterization
- Albino et al. 08: Genes in tumor & Basic model & Sh
- Wang 11: Museum pass & Costs & Sh & Characterization
- Esteban-Wall 11: genes in autism & Basic model & Sh

Literature review 2

- Ginsburgh-Zang 12: wines & Basic model & Sh & Characterization
- EstevezFernandez-Borm-Maers 12: Museum pass & Basic model-prices & BR & Characteriztions
- CasasMendez-Fragnelli-GarciaJurado 14: Museum pass & Prices & Sh-PropR-BR & Inventory of Properties
- Karpov 14: Bibliometrics & Several models & Equivalence theorem
- Bergantiños-MorenoTernero 15: Museum pass & Basic model-prices & Sh & Characterizations
- Ginsburgh-MorenoTernero-Weber 17: Languages & Levels of knowledge & Sh-Other & Characterizations
- Cesari-Algaba-Moretti-Nepomuceno 18: genes in co-expression networks & weights & Sh

Literature review 3

- LopezNavarrete-SS-Bonastre 19: Multimedia & revenues & Sh-Others
- Algaba-Fragnelli-Llorca-SS 19a: Transport modes & prices and costs & Sh-Others & Characterizations
- Algaba-Beal-Fragnelli-Llorca-SS 19b: networks and museum pass & basic models
- LopezNavarrete-SS-Bonastre 20: Youtube & order-repetitions-revenues & Sh-Others & Properties
- Martinez-SS 20a: COVID19 & Extended model & Others & Characterizations
- Martinez-SS 20b: Museum pass & Extended model & Others & Characterizations
- Google: Marketing & Basic model-revenues-order & Sh-Others

Table of Contents

Mathematical model (Martinez-SS, 2020a)

- $P = \{1, \ldots, p\}$ is the set of **pathologies**,
- Group of **individuals**, $N = \{1, \ldots, n\}$.
- A **lethality matrix** is a matrix $X \in \{0,1\}^{n \times p}$ of n rows and p columns, where

$$x_{ia} = \begin{cases} 1 & \text{if } i \text{ dies with pathology } a \\ 0 & \text{otherwise} \end{cases}$$

- $x_{i\cdot} \rightarrow i$-th row of X, pathologies that i had when died.
- $x_{\cdot a} \rightarrow a$-th column of X, individuals that died with pathology a.
- $\mathcal{D}^N$, domain of all possible matrices with individuals in N. $\mathcal{N}$ is the set of finite subsets of $\mathbb{N}$. $\mathcal{D} \equiv \bigcup_{N \in \mathcal{N}} \mathcal{D}^N$ is the class of all possible matrices with variable population.

Rules

Example

$N = \{1, \ldots, 6\}$ individuals who die with one or several pathologies in $P = \{\text{diabetes, high blood pressure, bronchitis}\}$. Data are:

$$X = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

Q: Which is the relevance (or influence) of each pathology on the lethality of this society?

Rules

Definition of lethality relevance rule

A lethality relevance **rule**. It is a mapping $\lambda : \mathcal{D} \longrightarrow \mathbb{R}_{+}^{p}$ that assigns, to each lethality matrix $X \in \mathcal{D}^{N}$, a vector $\lambda(X)$, where $\lambda_a(X)$ is the relevance of pathology $a \in P$ on the lethality in N population .

Rules

Count rule

For each $X \in \mathcal{D}^N$ and each $a \in P$,

$$\lambda_a^C(X) = \sum_{i=1}^{n} x_{ia}.$$

Share rule

For each $X \in \mathcal{D}^N$ and each $a \in P$,

$$\lambda_a^S(X) = \frac{1}{n} \sum_{i=1}^{n} x_{ia}.$$

Rules

Ratio rule

For each $X \in \mathcal{D}^N$ and each $a \in P$,

$$\lambda_a^R(X) = \begin{cases} \dfrac{1}{\|X\|_1} \displaystyle\sum_{i=1}^{n} x_{ia} & \text{if } \|X\|_1 > 0 \\ 0 & \text{if } \|X\|_1 = 0 \end{cases},$$

where $\|X\|_1 = \sum_{a \in P} \sum_{i=1}^{n} x_{ia}$.

Since it may eventually occur that none of the individuals in the society N dies having some of the particular pathologies included in P, we also have to consider the case of X being the null matrix.

Rules

Equal attribution rule

For each $X \in \mathcal{D}^N$ and each $a \in P$,

$$\lambda_a^E(X) = \begin{cases} \sum_{i \in N_a^*} \frac{x_{ia}}{|x_{i\cdot}|} & \text{if } \|X\|_1 > 0 \\ 0 & \text{if } \|X\|_1 = 0 \end{cases},$$

where $|x_{i\cdot}| = \sum_{a \in P} x_{ia}$ and $N_a^* = \{i \in N : |x_{ia}| \geq 1\}$.

Rules

Example (cont'ed)

We first illustrate the functioning of the equal attribution rule. From matrix X we construct the following matrix:

$$X^E = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

Thus,

$$\lambda^E(X) = \left(1 + \frac{1}{3} + \frac{1}{2}, \frac{1}{2} + \frac{1}{2} + \frac{1}{3}, \frac{1}{2} + \frac{1}{2} + \frac{1}{3} + \frac{1}{2}\right) = \left(\frac{11}{6}, \frac{8}{6}, \frac{11}{6}\right)$$

Rules

Example (cont'ed)

The count rule, for the same example, is

$$\lambda^{C}(X) = (3, 3, 4)$$

If we apply the share rule we get

$$\lambda^{S}(X) = \left(\frac{1}{2}, \frac{1}{2}, \frac{2}{3}\right)$$

Finally, the lethality relevance according to the ratio rule is

$$\lambda^{R}(X) = \left(\frac{3}{10}, \frac{3}{10}, \frac{4}{10}\right)$$

Rules

Aggregate impact

Given a lethality matrix $X \in \mathcal{D}$, we define the aggregate impact of the pathologies in P as the aggregation of lethality relevances, $\Lambda(X) = \sum_{a \in P} \lambda_a(X)$.

It is worth noting that, as the previous example illustrates, the aggregate impacts of different rules are, in general, different. Indeed, in Example, $\Lambda^E(X) = 5$, $\Lambda^C(X) = 10$, $\Lambda^S(X) = \frac{5}{3}$, and $\Lambda^R(X) = 1$.

Properties

Neutrality

For each $X \in \mathcal{D}^N$,

$$\lambda_a(\pi(X)) = \lambda_{\pi(a)}(X),$$

where $\pi(X)$ is a permutation of the columns of X.

Irrelevance

For each $X \in \mathcal{D}^N$ and each $a \in P$, if $x_{ia} = 0$ for all $i \in N$ then $\lambda_a(X) = 0$.

Properties

Composition

Let $N, M \in \mathcal{N}$ such that $N \cap M = \emptyset$. For each $X \in \mathcal{D}^N$ and each $Y \in \mathcal{D}^M$,

$$\lambda(X \oplus Y) = \lambda(X) + \lambda(Y),$$

where $X \oplus Y \in \mathcal{D}^{N \cup M}$ is the matrix resulting from stacking X above Y (by rows).

Main results

Theorem. Characterization

Let $X \in \mathcal{D}^N$. A rule λ satisfies neutrality, irrelevance, and composition if and only if, for each $a \in P$,

$$\lambda_a(X) = \sum_{i=1}^{n} w_i(x_{i\cdot})x_{ia}$$

for some neutral functions $w_i : \{0,1\}^N \longrightarrow \mathbb{R}_+$, $\forall i \in N$.

Main results

Proposition. Logical independence of the properties

1. The share rule satisfies neutrality and irrelevance, but violates composition. For each $X \in \mathcal{D}^N$ and each $a \in P$,

$$\lambda_a^S(X) = \frac{1}{n}\sum_{i=1}^{n} x_{ia}$$

2. For each $X \in \mathcal{D}^N$ and each $a \in P$, $\lambda_a^1(X) = \|X\|_1$ satisfies neutrality and composition, but violates irrelevance.
3. The following rule satisfies irrelevance and composition, but violates neutrality. For each $X \in \mathcal{D}^N$ and each $a \in P$,

$$\lambda_a^2(X) = \begin{cases} \sum_{i=1}^{n} x_{ia} & \text{if } a = 1 \text{ and } a \text{ is not irrelevant} \\ 0 & \text{otherwise} \end{cases}$$

Main results

Remarks

1. The share and ratio rules do not belong to the family
2. The count rule is an element of the family, with $w_i(x_{i\cdot}) = 1$, $\forall i \in \mathbb{N}$.
3. The equal attribution rule also belongs to this family, with

$$w_i(x_{i\cdot}) = \begin{cases} \frac{1}{|x_{i\cdot}|}, & \textit{if} \quad |x_{i\cdot}| \geq 1, \\ 0, & \textit{otherwise} \end{cases}, \forall i \in \mathbb{N}.$$

Note that when $|x_{i\cdot}| = 0$, any function w_i is possible, but we use $w_i \equiv 0$ for simplicity.

Equal attribution rule and the Shapley value

The set of players is the set of pathologies P, and the characteristic function for each coalition $Q \subset P$ is given by

$$v(Q) = |\{i \in N : x_{ia} = 1 \text{ for some } a \in Q\}|$$

For a given cooperative game (P, v), a solution is a vector $s \in \mathbb{R}^n_+$ such that $\sum_{a \in P} s_a = v(P)$, where s_i represents the allocation to player i.

Recall. The Shapley value

For each $a \in P$,

$$\text{Sh}_a(v) = \sum_{Q \subset P \setminus \{a\}} \gamma(P, Q)(v(Q \cup \{a\}) - v(Q)),$$

where $\gamma(P, Q) = \frac{|Q|!(p-|Q|-1)!}{p!}$ and $|Q|$ is the cardinal of Q.

Equal attribution rule and the Shapley value

Example (cont'ed)

For the lethality matrix of the Example, the associated characteristic function is:

$$\begin{array}{llll} v(\emptyset)=0, & v(\{1\})=3, & v(\{1,2\})=5, & v(\{1,2,3\})=5, \\ & v(\{2\})=3, & v(\{1,3\})=5, & \\ & v(\{3\})=4, & v(\{2,3\})=4 & \end{array}$$

And the corresponding Shapley value

$$\mathrm{Sh}_a(v)=\left(\frac{11}{6},\frac{8}{6},\frac{11}{6}\right)$$

Equal attribution rule and the Shapley value

Theorem

The equal attribution rule and the Shapley value of the associated cooperative game coincide.

Thus we have showed that the justification to use the equal attribution rule is twofold:

- it belongs to the family of rules uniquely determined by a suitable combination of properties.
- it corresponds to the Shapley value of the associated natural game.

An application to COVID-19 disease

- The coronavirus disease 2019 (COVID-19) is an infectious disease caused by severe acute respiratory syndrome coronavirus.
- One of the first studies was published by the *The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team*. This is a descriptive and exploratory analysis of all cases of COVID-19 diagnosed nationwide in China in February 2020.
- Among the several aspects analyzed in this work, the authors identify hypertension, cardiovascular disease, diabetes, and chronic respiratory disease as the main sources of co-morbidity among those patients whose death was caused by COVID-19 (32.8% of deaths did not have a preexisting disease.)

An application to COVID-19 disease

Pathology	% of deaths
Hypertension (HYP)	39.7%
Cardiovascular disease (CAR)	22.7%
Diabetes (DIA)	19.7%
Chronic respiratory disease (RES)	7.9%

Table : Co-morbid conditions reported by *The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team*.

An application to COVID-19 disease

Let N be the set of patients who died with diagnosed COVID-19, and let $P = \{\text{HYP}, \text{CAR}, \text{DIA}, \text{RES}\}$ the set of the the four pathologies used. The lethality matrix X would be obtained from the micro data, and identifies the preexisting diseases (other than the coronavirus).

Patient	HYP	CAR	DIA	RES
1	1	0	0	0
2	0	0	0	0
3	0	1	1	0
4	1	0	1	1
5	0	1	0	0
⋮	⋮	⋮	⋮	⋮

Table : Each row corresponds to each patient in the database. If the patient has died with the pathology in the column, the value is 1, and 0 otherwise.

An application to COVID-19 disease

I

f the micro data were public, the computation of the equal attribution rule would be straightforward, and the medical implications of the results could be analyzed. This is easily doable by a health authority with access to the information. Unfortunately, these micro data are not made public. However, we can circumvent the lack of data at individual level by means of simulations. We know that the results are not as reliable as if we use real data, but we believe they are sufficiently valid to illustrate the application of the theoretical model.

An application to COVID-19 disease

Simulation conditions 1

- We may ignore the entries of the lethality matrix X, but we know that the sum of the entries in the first column must coincide with the numbers of deaths with hypertension. The same reasoning applies to the rest of columns/pathologies.
- Let us assume that N has 50 individuals whose pathologies are distributed similarly to the Table. Larger numbers of individuals make the simulation computationally intractable. In particular, we suppose that 20 individuals died with hypertension, 11 individuals died with cardiovascular disease, 10 individuals died with diabetes, and 4 individuals died with chronic respiratory disease

An application to COVID-19 disease

Simulation conditions 2

- Let $\mathcal{X}$ be the set of all possible lethality matrices that are compatible the distribution of deaths in the previous step, that is, $X \in \mathcal{X}$ if and only if $\sum_{i \in N} x_{ia} = \text{obs}(a)$ for all $a \in P$, where $\text{obs}(a)$ denotes the observations of pathology a.
- We computationally generate all the matrices in $\mathcal{X}$.
- For each $X \in \mathcal{X}$ obtained in the previous step, we compute the equal attribution rule $\lambda^E(X)$.
- Finally, we average across $\mathcal{X}$, obtaining $\overline{\lambda}^E$.

An application to COVID-19 disease

Simulation results

Pathology	$\overline{\lambda}_a^E$	ν_a^E
Hypertension	44.78	0.512
Cardiovascular disease	19.42	0.222
Diabetes	17.23	0.197
Chronic respiratory disease	6.04	0.069

Table : Equal attribution rule for COVID-19

Table of Contents

Final comments

- The museum pass model is behind many attribution problems.
- The Shapley value can be considered a benchmark solution for attribution problems.
- An application in which one of the conditions of the basic model of attribution is removed is analyzed.
- In Martinez-SS (2020b), the previous model is analyzed in deep for the museum pass problem.

Table of Contents

References 1

- Shapley, L. S., 1953. A value for n-person games. In: Contribution to the Theory of Games vol. II (H.W. Kuhn and A.W. Tucker eds). Annals of Mathematics Studies 28. Princeton University Press, Princeton.
- Brams, S. J., Fishburn, P. C., 1978. Approval voting. The American Political Science Review 72, 931-847.
- Land M, Gefeller O (2000) A multiplicative variant of the Shapley value for factorizing the risk of disease. In: Patrone F, Garcia-Jurado I, Tijs S (eds) Game practice: contributions from applied game theory. Kluwer Academic, Dordrecht, pp 143-158
- Ginsburgh, V. and Zang, I. (2003). The museum pass game and its value. Games and Economic Behavior, 43:322-325.
- Ginsburgh, V., and Zang, I. (2004). Sharing the income of a museum pass program. Museum Management and Curatorship, 19(4), 371-383.

References 2

- Naor, M., 2005. On fairness in the carpool problem. Journal of Algorithms 55, 93-98.
- Moretti S, Patrone F, Bonassi S (2007) The class of microarray games and the relevance index for genes. Top 15:256-280
- Albino, D., Scaruffi, P., Moretti, S., Coco, S., Truini, M., Cristofano, C. D., Cavazzana, A., Stigliani, S., Bonassi, S., and Tonini, G. P. (2008). Identification of low intratumoral gene expression heterogeneity in neuroblastic tumors by genome-wide expression analysis and game theory. Cancer, 113:1412-1422.
- Wang, Y., 2011. A museum cost sharing problem. American Journal of Operations Research 1, 51-56.
- Esteban, F. J. and Wall, D. P. (2011). Using game theory to detect genes involved in Autism Spectrum Disorder. TOP, 19:121-129.

References 3

- Ginsburgh, V., Zang, I., 2012. Shapley ranking of wines. Journal of Wine Economics 7, 169-180.
- Estevez-Fernandez, A., Borm, P., and Maers, H. (2012). A note on passepartout problems. International Game Theory Review, 14:1-9.
- Casas-Mendez, B., Fragnelli, V., and Garcia-Jurado, I. (2013). A survey of allocation rules for the museum pass problem. Journal of Cultural Economics, 38:191-205.
- Brams, S. J., Kilgour, M., 2014. Satisfaction approval voting. In: Fara, R., Leech, D., Salles, M. (Eds.), Voting Power and Procedures Essays in Honour of Dan Felsenthal and Moshe Machover. Springer, Heidelberg, pp. 323-346.
- Karpov, A., 2014. Equal weights co-authorship sharing and the Shapley value are equivalent. Journal of Informetrics 8, 71-76.

References 4

- Bergantiños, G. and Moreno-Ternero, J. D. (2015). The axiomatic approach to the problem of sharing the revenue from museum passes. Games and Economic Behavior, 89:78-92.
- Ginsburgh, V., Moreno-Ternero, J. D., Weber, S., 2017. Ranking languages in the European Union: Before and after Brexit. European Economic Review 93, 139-151.
- Cesari, G., Algaba, E., Moretti, S., and Nepomuceno, J. A. (2018). An application of the Shapley value to the analysis of co-expression networks. Applied Network Science, 3:3-35.
- F. Lopez-Navarrete, J. Sanchez-Soriano, O.M. Bonastre. (2019). Allocating revenues in a smart TV ecosystem. International Transactions in Operational Research vol. 26:1611-1632.

References 5

- E. Algaba, V. Fragnelli, J. Sanchez-Soriano. (2019) Horizontal cooperation in a multimodal public transport system: The profit allocation problem. European Journal of Operational Research vol. 275:659-665.
- E. Algaba; S. Beal; V. Fragnelli, N. Llorca; J. Sanchez-Soriano. 2019. Relationship between labeled network games and other cooperative games arising from attributes situations. Economic Letters vol. 185: art.108708.
- Martinez, R., Sanchez-Soriano, J. (2020) A measure of lethality relevance. An application to COVID-19. ThE. Submitted (under revision)
- Martinez, R., Sanchez-Soriano, J. (2020) A generalized domain for the museum pass problem. Mimeo. Submitted (under revision)

References 6

- F. Lopez-Navarrete, J. Sanchez-Soriano, O.M. Bonastre. (2020). Dynamic generation and allocation of revenues in a video website. Mimeo. Submitted (under revision)
- Google. https://support.google.com/google-ads/topic/7279627?hl=en&ref_topic=3119145
- jie Guan, et al. (2020). Comorbidity and its impact on 1590 patients with COVID-19 in china: a nationwide analysis. European Respiratory Journal, 55:2000547.

THANKS FOR YOUR ATTENTION!!!