XIII CONGRESSO NAZIONALE DEGLI ATTUARI

INNOVAZIONE TECNOLOGICA E RISCHI SISTEMICI: L'ATTUARIO VALUTATORE GLOBALE DELL'INCERTEZZA

ROMA
10-12 Novembre 2021

ORDINE NAZIONALE DEGLI ATTUARI

CONSIGLIO NAZIONALE DEGLI ATTUARI

Shapley values: how to price single risks in an insurance portfolio

Giovanni Rabitti
Department of Actuarial Mathematics and Statistics
Heriot-Watt University, Edinburgh

Motivation

- Consider a function $Y = g(X_1, X_2, ..., X_n)$, which can be
  - a portfolio evaluation model
  - a mathematical model (actuarial pricing model,...)
  - a statistical model (regression, neural network,...)
- We want to find a way to fairly allocate quantities of interest related to Y into the individual contributions of variables $X_1, X_2, ..., X_n$.
- These can be: $\mathbb{E}[Y], \mathbb{V}[Y], \rho(Y), ...$

A game theory approach: the Shapley value

- Shapley values come from cooperative game theory [Shapley, 1953].
- A coalition of n players generate a value and we have to divide it so that each member receives his/her fair part of the value that he/she contributed to generate
- Consider a value function ν for any possible coalition T, then the Shapley value for the i-th player is:

$$\phi_i^{\nu} = \sum_{T \subseteq \{1,2,\ldots,n\} \setminus \{i\}} \frac{|T|!(n-|T|-1)!}{n!} \left(\nu(T \cup \{i\}) - \nu(T)\right).$$

Properties of the Shapley value

The Shapley value can be characterized by several interesting properties. Denote by $N = \{1, 2, ..., n\}$. Then,

- (Efficiency) $\sum_{i=1}^{n} \phi_i(\nu) = \nu(N)$.
- (Symmetry) If $\nu(J \cup \{i\}) = \nu(J \cup \{j\})$ for all $J \subseteq N \setminus \{i, j\}$, then $\phi_i(\nu) = \phi_j(\nu)$.
- (Dummy player) If $\nu(J \cup \{i\}) = \nu(J)$ for all $J \subseteq N$, then $\phi_i(\nu) = 0$.
- (Linearity) If two value functions ν and μ have Shapley values, respectively, $\phi(\nu)$ and $\phi(\mu)$, then the game with value $\alpha\nu + \beta\mu$ has Shapley value $\alpha\phi(\nu) + \beta\phi(\mu)$ for all $\alpha, \beta \in \mathbb{R}$.

Shapley (1953) proved that the Shapley value is the unique attribution method satisfying these four properties.

Variance Games (Scarsini et al.,2018, Galeotti and R., 2021)

- Consider the risk analysis problem:

$$g(X_1, X_2, ..., X_n) = S = \sum_{i=1}^{n} X_i$$

without any (!) distributional assumption on X_i.

- Define the value function

$$\nu(J) = \mathbb{V}[S_J] = \mathbb{V}\left[\sum_{j \in J} X_j\right]$$

- Scarsini et al. (2018) prove that

$$\phi_i = Cov[X_i, S]$$

- Interpretation: If a random variable contributes to hedge a risk, it is rewarded with a negative Shapley value.

Tail Variance Games and Premium Principles

- With Marcello Galeotti (University of Florence) we have consider the tail variance game
$$\nu_s(J) = \mathbb{V}\left[S_J | S > s\right]$$
- We have proved that in this case
$$\phi_i = Cov[X_i, S | S > s]$$
- Then, the tail covariance premium principle is
$$TCP(X_i | S > s) = \mathbb{E}[X_i | S > s] + \alpha \frac{\phi_i}{\sqrt{\sum_{i=1}^{n} \phi_i}}$$

In this case the grand total value of the game is
$$\sum_{i=1}^{n} TCP(X_i | S > s) = \mathbb{E}[S | S > s] + \alpha \sqrt{\mathbb{V}\left[S | S > s\right]}$$

Shapley effects

- But models are not always linear...
- Art Owen's Idea [Owen, 2014]: what if value = explained variance by a input/player?

$$\nu(T) = \mathbb{V}\left[\mathbb{E}\left(f(\mathbf{X})|X_T\right)\right]$$

- This idea has opened the door to Shapley effects for global sensitivity analysis when Sobol' analysis has conceptual difficulties, for instance:
  1. dependent inputs;
  2. domain with holes.

Given data setting: individual medical costs billed by health insurance

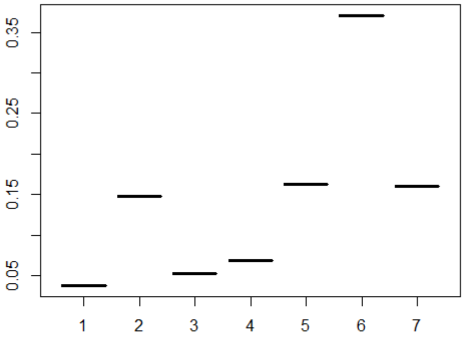


Figure 1: "age" "sex" "bmi" "steps" "children" "smoker" "region"

Car insurance premium determinants

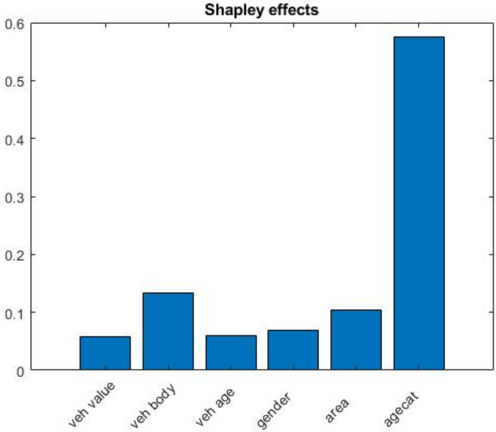
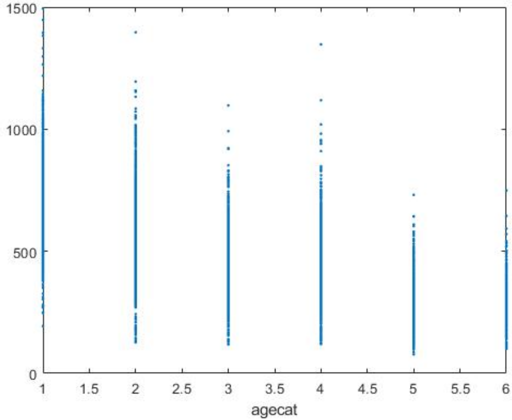


Figure 2: (In collaboration with Amir Khorrami Chokami and Arianna Vallarino, University of Turin)



1500
1000
500
0
1
1.5
2
2.5
3
3.5
4
4.5
5
5.5
6
agecat

Skewness allocation indices and higher order Shapley

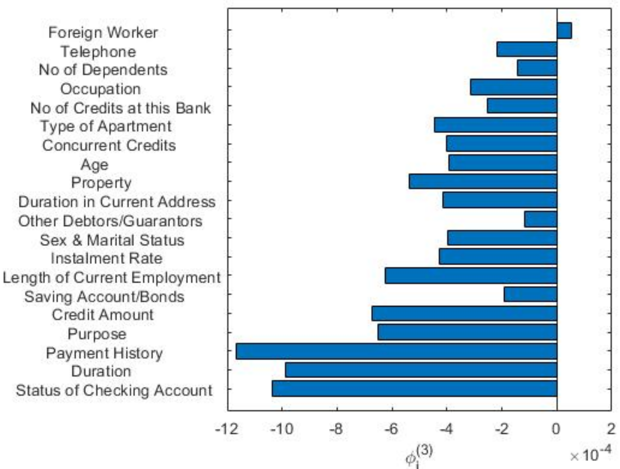
- Consider the higher order index

$$h_u^{(p)} = \mathbb{E}\left[\left(\mathbb{E}\left[Y|\mathbf{X}_u\right] - \mathbb{E}\left[Y\right]\right)^p\right].$$

- If $h_1^{(3)} < 0$ and $h_2^{(3)} > 0$, this may indicate that controlling x_1 is more important for attaining (or avoiding) very small values of f while x_2 is more important for large values of f.
- With Borgonovo (Bocconi) and Plischke (Clausthal TU) we consider the value function $\nu(u) = h_u^{(3)}$ and call the Shapley values the higher order Shapley effects.

German credit scoring dataset

We used a logistic model to find the predicted probability that a credit applicant belongs to a good risk class (1) or not (0).



UniCredit bank operational losses

- This dataset contains 40,872 operational losses registered between January 2005 and June 2014.
- Hambuckers et al. (2018) wanted to identify the most important economic and financial variables determining the severity of the operational losses.
- They fitted a generalized Pareto regression using a penalized-likelihood procedure and selected the following variables according to the model fit indicators:

Selected variables	Description
GDP EU	EU GDP growth rate
UNEMP IT	Italian unemployment rate
VIX	Volatility index computed from a panel of S&P500 option prices
VFTSE	Volatility index based on options of the FTSE 100 index
LR	Leverage ratio: the total value of the assets divided by the value of the equity
HPI	Growth rate of residential property price index in Europe
ST	Short-term (3 months) interbank interest rate for Italy
DGR	Deposit growth rate variations
MIB	Milano Italia Borsa return index

Operational loss event type	Description
IFRAUD	Internal frauds
EFRAUD	External frauds, related to credit card payments and others
EPWS	Employment practices and workplace safety
CPBP	Clients, products, and business practices, related to derivatives and financial instruments
EDPM	Execution, delivery, and process management, related to financial instruments and payments

	$h_N^{(3)}/10^{16}$	GDP EU	UNEMP IT	VIX	VFTSE	LR	HPI	ST	DGR	MIB
IFRAUD	2.7323	0.0331	0.0949	0.1117	0.0470	0.1001	0.0982	0.0831	**0.3411**	0.0907
EFRAUD	0.0194	**0.2623**	0.0254	0.1084	0.0770	0.1194	0.1222	0.0477	0.1335	0.1040
EPWS	0.1601	0.0256	0.1777	0.0129	0.0143	0.1262	0.0445	0.0570	**0.3157**	0.2262
CPBP	0.2548	0.0557	0.1277	0.0965	0.0704	0.1053	0.1207	0.0737	**0.2032**	0.1468
EDPM	0.0020	0.0428	0.1104	0.0953	0.0627	0.1805	0.0856	0.0573	0.1410	**0.2243**

- IFRAUD: (Povel et al. 2007, RFS) higher size of transactions, higher internal frauds
- EFRAUD: (Cope et al. 2012, JBF) booming economy, higher external frauds
- EPWS: (Hambuckers et al. 2018, JAE) severance payments depend on economic conditions and on the wages
- EDPM, CPBP: (Hambuckers et al. 2018) financial markets

- Borgonovo, Plischke, R., *Interactions and Computer Experiments*, Accepted, Scandinavian Journal of Statistics
- Galeotti, R., *On the comparison of Shapley values for the variance and the standard deviation games*, Journal of Applied Probability, 2021
- Galeotti, R., *Tail variance allocation, Shapley value and the majorization problem*, In Preparation
- Plischke, R., Borgonovo, *Computing Shapley effects for Sensitivity Analysis*, SIAM/ASA Journal on Uncertainty Quantification, 2021
- Plischke, R., Borgonovo, *Has the Spell Been Broken? Estimating Global Sensitivity Measures via Nearest Neighbors*, In Preparation
- R., Borgonovo, Plischke, *Higher order Shapley effects for the sensitivity analysis of extremes*, In Preparation

Thank you!