PROBABILISTIC AND CAUSALITY TOOLS FOR SUPPORT OF MANAGERIAL DECISIONS

VLADISLAV BÍNA – LUCIE VÁCHOVÁ

Abstract: *The paper concerns a brief introduction to (or rather a mere degustation of) the theory of multidimensional probabilistic distributions and causal inference as a supportive tool for managerial decision making. More specifically, in vast majority of cases the managers face the need to make a decision under risk or uncertainty and need more or less sophisticated tools as a support of their decisions. A bunch of exact methods based on probability theory were developed, but it is apparent that the natural way of thinking often relies on the terms of cause and effect. This can be seen in the popularity of "soft" tools like cognitive maps and influence diagrams. The paper therefore brings an introduction to the methodology providing a possibility to perform exact probabilistic calculations in models representing causal relations and particularly focuses on the so called interventions, i.e., actions resulting in setting a variable to some value by a decision or regulation. A simplified introductory example is proposed and solved using the Pearl's calculus of intervention. To show the connection with "soft" methods, the author presents this simple problem as an influence diagram computationally handled by sophisticated software tool, namely Lumina Analytica. This software visually represents the problem as well known influence diagrams enriched by numerous mathematical tools (in our case probabilistic inference, but in general it can perform tasks like Monte Carlo methods in scenario analysis, sensitivity analysis, Bayesian networks, etc.).*

Keywords: *probability, causality, intervention, influence diagrams, decision making*

JEL Classification: *C63, C11*

1. INTRODUCTION AND MOTIVATION

From the philosophical perspective, the paper deals with the essential themes of human thinking: uncertainty and causality. Relations of dependency and causality rank among the crucial universal concepts of the philosophy and science. The first scientific utilization of the notion of cause dates back to the ancient times. Let us pick up, e.g., the Aristotle's four causes [2]: the material cause, the formal cause, the efficient cause and the final cause. Here the efficient (or moving) cause has properties which are closest to the modern understanding of the cause, i.e., thing that acts in such a way that it leads to an effect in another thing. The development of physics in the 17th century brought a shift of conception of cause from this active role of "mover" towards perceiving the cause rather as inactive article in the implication chain of the laws of nature. The notions of dependency and independency originate in nineteenth century in the field of differential and integral calculus and were often used in twentieth century within the rapidly developing domains of probability and statistics.

Though the notions of cause and effect appear to rank about the cornerstones of human reasoning, statistical methods almost completely fail to address it. The bunch of methods in the field of statistical analysis and probability cannot handle the causal relations among the analyzed variables. This deflection from the causation as a basic mean for the expressing of the relation between a pair of variables originated with the famous Galton's interpretation of correlation. The well-known assertion says "Correlation does not imply causation." Many examples are shown in the textbooks where some common cause (called confound) significantly affects the correlated variables. For explanation of the basic notions, see a book focused on causal inference in experimental design by William Shadish et al. [12].

As we mentioned, the causality can't be proved by means of statistics. Of course, there are cases when statistical tools give a strong evidence for the existence of causal relation. In natural sciences, the experimental design provides a possibility to search for causes of some effect. In strictly controlled tests scientists always can repeat the application of some drug to laboratory mice or compare use of different operating techniques on sufficiently large sample. But it is impossible to run some historical events repeatedly in different setting, the government cannot afford to repeat an intervention in economy of a state and managers cannot afford to try different types of assembly lines in their enterprise. Such impossibility leads us to the need to perform an analysis of collected data and investigate the causal relations.

Contemporary managers are used to perform plenty of analyses using the modern information tools and systems. These provide a possibility to easily generate enormous quantity of information based on the questions the manager is interested in. But often they misinterpret the results and confuse the operation of conditioning with the result of an intervention. Whereas the first technique is a general operation employed in probabilistic tools (see, e.g., Finn Jensen and Thomas Nielsen [5]) the latter possibility is available, e.g., in causal networks and their equivalents developed by Judea Pearl [9]. But the difference can be significant, in some situations really substantive. For instance, consider the historical situation when alternative bio-fuels were just a marginal choice. There was quite low volume of demand for the alternative types of fuel in the situation of open market. But apparently, the situation changed when the legal act prescribed the obligatory mixing with the conventional fuel. To model the difference correctly one needs to carefully distinguish between the operations of conditioning and intervention.

The Pearl's approach to causality modeling and the calculus of intervention have inspired many contemporary researchers in different fields of management, e.g., applications in strategic management by Michael Ryall [10], in psychology by York Hagmayer et al. [3], in educational research and other social sciences Richard Murnane [8] or an introduction to causal inference for the researchers in social sciences by Stephen Morgan and Christopher Winship [7]. The authors already published an economic application of the classical Pearl's graphical models [13], and, simultaneously, started the development of an original "algebraic" (non-graphical) theoretical tool based on compositional models [1].

2. METHODOLOGY AND BASIC NOTIONS

We said above that there is a significant difference between the correlation and causation. Correlation describes symmetrical dependence of variables. In contrast, causation is an asymmetric relation, and this is a frequent source of confusion when these two different relations are incorrectly exchanged and wrong deductions are obtained.

The first quantitative tools for handling the cause and effect appeared as an augmentation of the Galton's and Person's concept of correlation and later the approaches of multiple regression analysis into the tool called path analysis. This approach was developed by Sewall Wright in 1920s [14] and can be perceived as a special case of *structural equation model* (SEM) developed in 1940s by Trygve Haavelmo [4]. The SEM approach is said to combine the quantitative data with qualitative assumptions on causal relations and thus gives a tool to check or estimate the causality in data. Notice that the SEM methodology provides a possibility to employ latent (unmeasured) variables. The interpretation of structural equation as an algebraic object without any causal content was criticized by Judea Pearl who proposes to reestablish the original conception that the causality is an inherent content of SEM. And it was him who formally redefined the SEM methodology using his (nowadays famous) approach to causality [9].

Bayesian networks are often used to represent causal relationships, since its graph employs (oriented) arcs. But the graphical representation is ambiguous, some arcs can be reversed (see Ross Shachter in [11]). Actually, the acyclic directed graphs represent conditional independence relations but they do not necessarily embed causal relations. But the nature of human thinking usually employs the notions of cause and effect and finds it much easier to construct the model on causal assumptions. And this is the source of popularity of influence diagram models and Bayesian networks (for details see again, e.g., Jensen and Nielsen [5]). And this was also why Judea Pearl [9] has been developing his approach of causal networks, i.e., Bayesian networks with the relations explicitly defined as causal. The *causal Bayesian networks* thus enable us to perform an intervention (actively setting a value of variable X using the intervention operator). This methodology provides a possibility to estimate the impact of external actions from the data collected before the intervention.

Now let us clarify several basic notions concerning the modeling of causal relations. We will denote the variables by big roman letters and their particular values by small roman letters. A *directed acyclic graph* (DAG) is a graph with all edges directed and containing no directed cycles. A *causal structure* of a set of variables V is a DAG with each node corresponding to a distinct element of V and each of the arrows representing a direct functional relationship among corresponding variables. A *causal model* can be (according to [9]) specified as a pair $M = \langle D, \Theta_D \rangle$ (where D is a causal structure and Θ_D is a set of parameters compatible with D). The parameters Θ_D assign a function $x_i = f_i(pa_i; u_i)$ to each $X_i \in V$ and a probability measure $P(u_i)$ to each U_i, where pa_i are particular values of variables from PA_i, i.e., of the parents of X_i in D and where each U_i is a random disturbance (*unmeasured variable*) distributed according to $P(u_i)$, independently of all other U.

An *intervention* (or an *action*) denoted $do(X = x)$ is setting the fixed value by an external force on the specific set of variables in the causal diagram. This is often shortened to $do(x)$ or $\hat{x}$ (see [9]). Given two disjoint sets of variables X and Y, a *causal effect* of X on Y is a function from X to thespace of probability distributions on Y and is denoted by $P(y|\hat{x})$. For each realization $x$of X, $P(y|\hat{x})$ gives the conditional probability of $Y = y$ induced by deleting from the causal model all equations corresponding tovariables in set X and substituting $X = x$ in the remaining equations. For more details and definition of two following criteria see, e.g., [9].

A set of variables Z satisfies the *back-door criterion* relative to an ordered pair of variables (X, Y) in a DAG G if:

1. no node in Z is a descendant of X,
2. Z blocks every path between X and Y that contains an arrow into X (i.e., every *back-door path*).

The important result of *back-door criterion* [9] states that if a set of variables Z satisfies the *back-door criterion* relative to (X, Y), then the causal effect of X on Y is identifiable and can be calculated using the formula

$$P(y|\hat{x}) = \sum_z P(y|x, z)\, P(z). \tag{1}$$

A set of variables Z satisfies the *front-door criterion* relative to an ordered pair of variables (X, Y) if:

1. Z intercepts all directed paths from X to Y,
2. back-door path from X to Z does not exist,
3. all back-door paths from Z to Y are blocked (synonym: d-separated [9]) by X.

Then if Z satisfies the *front-door criterion* relative to (X, Y) and if $P(x, z) > 0$, then the causal effect of X on Y is identifiable and is given by the formula (see, e.g., [9])

$$P(y|\hat{x}) = \sum_z P(z|x) \sum_{x'} P(y|x', z) P(x'). \tag{2}$$

Let us summarize that the causal effect of X on $Y$is *identifiable* particularly in the following cases:

1. the back-door paths between X and Y does not exist, the causal effect can be computed directly as the corresponding conditional distribution,
2. observed variables from Z block every back-door path between X and Y,
3. elements of Z satisfy front-door criterion relative to the pair of variables (X, Y).

For an application of different conditions of identifiability together with economical applications see [13]. An example concerning the third case – the front-door criterion – can be found in [1].

3. DIRECT ESTIMATION OF CAUSAL EFFECT

We will demonstrate the above sketched apparatus on rather toy examples taking advantage of software Lumina Analytica [6] which provides the possibility to both graphically represent the causal relations and perform the calculations to identify the causal effect of an intervention.

First of all let us examine a case where our profit Z (e.g., on selling of some type of goods in a small shop) is influenced by the price P and demand D but those two factors are not independent. We believe that our setting of price changes significantly the demand. For sake of simplicity we categorize roughly all considered variables. Our expert knowledge of the situation in the form of (conditional) probabilities of considered variants is summarized in Tables 1, 2 and 3.

Table 1 Probability distribution $P(p)$

Price	Low	High
	0,4	0,6

Source: own example

Table 2 Conditional probability distribution $P(d|p)$

	Price	Low	High
Demand	Low	0,1	0,6
	Medium	0,4	0,3
	High	0,5	0,1

Source: own example

Table 3 Conditional probability distribution $P(z|p,d)$

	Price	Low		High	
	Profit	Loss	Gain	Loss	Gain
Demand	Low	0,95	0,05	0,7	0,3
	Medium	0,9	0,1	0,4	0,6
	High	0,6	0,4	0,1	0,9

Source: own example

The above described situation can be easily depicted as a simple causal network like the one in Figure 1 (the index nodes contain levels of considered variables and thus are only auxiliary).

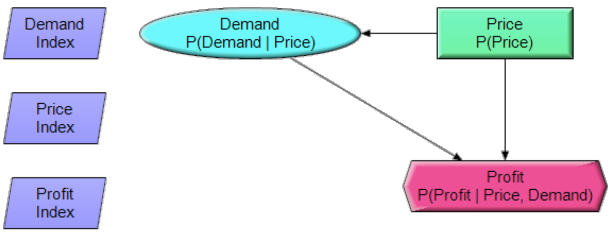

Figure 1 A causal dependence of profit on price
Source: own work with Lumina Analytica software

Now we would like to fix the price on the level "High" and see the effect on our profit. Since there does not exist any back-door path heading into the Price node we can compute the causal effect directly as the conditional probability. The conditional probability is computed from the joint distribution of profit, price and demand marginalizing out the demand and using the conditioning by price in the following formula

$$P(z|\hat{p}) = P(z|p) = \frac{\sum_d P(p)P(d|p)P(z|p,d)}{P(p)} = \\ = \sum_d P(d|p)P(z|p,d). \tag{3}$$

Analytica software can easily evaluate this effect (Figure 2) and give us the probability of positive profit as an effect of setting high price $P\big(z = \text{Gain}\big|p = \widehat{\text{High}}\big) = 0{,}45$.

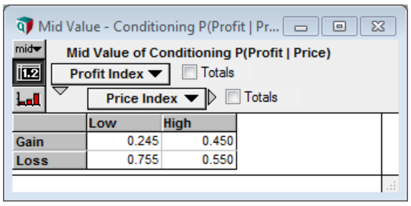

Figure 2 An evaluation of the causal effect
Source: own work with Lumina Analytica software

4. IDENTIFICATION OF CAUSAL EFFECT USING BACK-DOOR CRITERION

Let us now modify the situation and assume that the supply is constant and the demand strongly influences the price. Again we want to evaluate the effect of price P change on the profit Z in the situation when both price and profit are influenced by the demand D. The whole causal model in Analytica software is depicted on Figure 3.

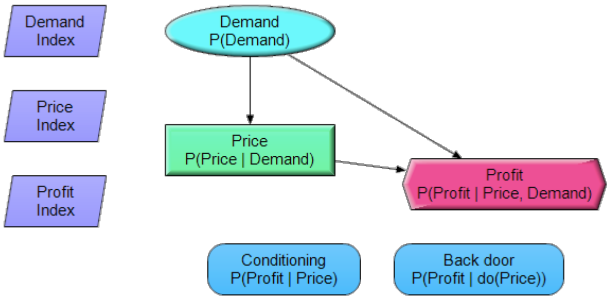

Figure 3 A causal dependence of profit on price
Source: own work with Lumina Analytica software

In the modified example we again categorize roughly all considered variables and our (expert) knowledge is again given in the form of (conditional) probabilities summarized in Tables 4, 5 and 6.

Table 4 Probability distribution $P(d)$

Demand	Low	Medium	High
	0,3	0,5	0,2

Source: own example

Table 5 Conditional probability distribution $P(p|d)$

	Demand	Low	Medium	High
Price	Low	0,9	0,4	0,05
	High	0,1	0,6	0,95

Source: own example

Table 6 Conditional probability distribution $P(z|p,d)$

	Demand	Low		Medium		High	
	Profit	Loss	Gain	Loss	Gain	Loss	Gain
Price	Low	0,95	0,05	0,9	0,1	0,6	0,4
	High	0,7	0,3	0,4	0,6	0,1	0,9

Source: own example

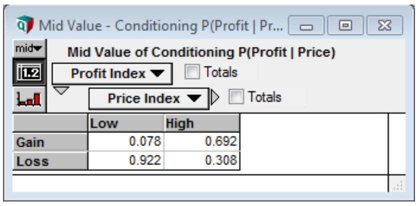

Figure 4 Effect of conditioning in the model
Source: own work with Lumina Analytica software

A simple and generally used estimation of causal effect is based on the process of conditioning. I.e., we evaluate

$$P(z|\hat{p}) = P(z|p) = \frac{\sum_d P(d)P(p|d)P(z|p,d)}{\sum_d P(d)P(p|d)}, \qquad (4)$$

which can be computed with Analytica (see Figure 4).

But this result is not correct since we do not take into account the change in structure implied by the process of intervention (again see, e.g., [9]). In order to get correct result it should be computed according to Formula 1

$$P(z|\hat{p}) = \sum_d P(z|p,d)\,P(d). \qquad (5)$$

Again we can employ Analytica with the result shown on Figure 5. We can see that the correct evaluation of causal effect using the back-door criterion provides much lower estimate of probability of positive profit as an effect of setting high price, i.e., $P\big(z = \mathrm{Gain}\big|p = \widehat{\mathrm{High}}\big) = 0{,}57$.

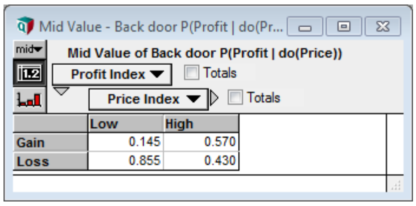

Figure 5 Evaluation of causal effect using back-door criterion
Source: own work with Lumina Analytica software

5. CONCLUSION

We presented a methodology for evaluation of the causal effect on a pair of simple examples. The results in the latter example show that the correct computation using causal approach provides significantly different probability estimates which can generally lead to different decision in comparison to the recommendations of usually – but incorrectly – used method of conditioning.

Authors also sketch the features of decision support tool Lumina Analytica which provide graphically attractive and easy-to-use way to evaluate the effects of intervention.

Notice that the situation can be enriched by employing a latent variable influencing both demand and profit variables in which case the causal effect still remains identifiable (for details refer again to Judea Pearl [9]).

Acknowledgement

The research was supported by GA ČR under the project number GA 403/12/2175.

REFERENCES

[1] BÍNA, V., JIROUŠEK, R. 2014. ***Causality and Intervention as Decision Support Tool in Business Process Management.*** Submitted to: The 12th International Conference on Business Process Management, Haifa, Israel, September 2014.

[2] FALCON, A. 2014. ***Aristotle on Causality.*** The Stanford Encyclopedia of Philosophy (Spring 2014 Edition), ZALTA, E. N. (Ed.). [online] [cit. 2014-04-13]. Available on Internet: http://plato.stanford.edu/archives/spr2014/entries/aristotle-causality/.

[3] HAGMAYER, Y., SLOMAN, S. A., LAGNADO, D. A., WALDMANN, M. R. 2007. ***Causal reasoning through intervention***. Causal learning: Psychology, philosophy, and computation, Oxford University Press, New York, 2007. s. 86–100. ISBN 0195176804.

[4] HAAVELMO, T. 1943. ***The statistical implications of a system of simultaneous equations.*** Econometrica Vol. 11 (1–2), 1943. s. 1–12. ISSN 1468-0262.

[5] JENSEN, F. V., NIELSEN, T. D. 2011. ***Probabilistic decision graphs for optimization under uncertainty.*** 4OR-Q Journal of Operation Research, Vol. 9 (1), Springer Verlag, 2011, s. 1–28. ISSN 1619-4500.

[6] ***Lumina decision systems Inc.*** 2009. Analytica, Lumina. Los Gatos, CA (USA). [online] [cit. 2014-04-13] Available on Internet: http://www.lumina.com/.

[7] MORGAN, S., WINSHIP, C. 2007. ***Counterfactuals and Causal Inference: Methods and Principles for Social Research***. Cambridge University Press, New York, 2007. 328 s. ISBN 0521829984.

[8] MURNANE, R. J., WILLETT, J. B. 2011. ***Methods Matter: Improving Causal Inference in Educational and Social Science Research.*** Oxford University Press, Oxford, 2011. 416 s. ISBN 0199780315.

[9] PEARL, J. 2009. ***Causality: Models, Reasoning, and Inference.*** Cambridge University Press, New York, 2009. 484 s. ISBN 052189560X.

[10] RYALL, M. 2009. ***Causal Ambiguity, Complexity, and Capability-Based Advantage***. Management Science Vol. 55 (3), 2009. s. 389–403. ISSN 0025-1909.

[11] SHACHTER, R. D. 1986. ***Intelligent probabilistic inference.*** In LEMMER, J. F., KANAL, L. N. (Eds.): Uncertainty in Artificial Intelligence, North Holland, Amsterdam, 1986. s. 371–382. ISBN 0-444-70058-7.

[12] SHADISH, W. R., COOK, T. D., CAMPBELL, D. T. 2002. ***Experimental and quasi-experimental designs for generalized causal inference.*** Boston, MA: Houghton Mifflin, 2002. 623 s. ISBN 0395615569.

[13] VÁCHOVÁ, L., BÍNA, V. 2013. ***A Note on Computation of Causal Effect in Economic Models***. In VOJÁČKOVÁ H. (Ed.) Proceedings of 31th International Conference Mathematical Methods in Economics, Jihlava, 2013. s. 974–979. ISBN 978-80-87035-76-4.

[14] WRIGHT, S. 1921. ***Correlation and causation***. J. Agricultural Research Vol. 20, 1921. s. 557–585. ISSN 0095-9758.

Ing.Vladislav BÍNA, Ph.D.
Ing. Lucie VÁCHOVÁ
University of Economics in Prague, Faculty of Management in Jindřichův Hradec, Department of Exact Methods
Jarošovská 1117/II, Jindřichův Hradec, 37701, Czech Republic
e-mail: bina@fm.vse.cz, vachova@fm.vse.cz