Analysis of symmetry properties for Bayesian confirmation measures

Salvatore Greco[1], Roman Słowiński[2,3], and Izabela Szczęch[2]

[1] Department of Economics and Business, University of Catania, Italy
salgreco@unict.it
[2] Institute of Computing Science Poznan University of Technology, Poland
{Izabela.Szczech, Roman.Slowinski}@cs.put.poznan.pl
[3] Systems Research Institute, Polish Academy of Sciences,Warsaw, Poland

Abstract. The paper considers symmetry properties of Bayesian confirmation measures, which constitute an important group of interestingness measures for evaluation of rules induced from data. We demonstrate that the symmetry properties proposed in the literature do not fully reflect the concept of confirmation. We conduct a thorough analysis of the symmetries regarding that the confirmation should express how much more probable the rule's hypothesis is when the premise is present rather than when the premise is absent. As a result we point out which symmetries are desired for Bayesian confirmation measures and which are truly unattractive. Such knowledge is a valuable tools for assessing the quality and usefulness of measures.

Keywords: Bayesian confirmation measures, symmetry properties, rule evaluation

1 Introduction

Discovering knowledge from data is the domain of inductive reasoning. Knowledge patterns induced from data are often expressed in a form of "*if*, *then*" rules. They are consequence relations representing correlation, association, causation between independent and dependent attributes. To measure the relevance and utility of the discovered rules many measures of interestingness have been proposed and studied [7], [9], [11]. Among these measures, an important role is played by Bayesian confirmation measures, which express in what degree a premise confirms a conclusion [1], [10], [12], [13]. Such measures are extremely valuable due to the fact that they indicate disconfirmatory rules, i.e. rules which are completely misleading and should be discarded from further use. In this context, the group of confirmation measures should be regarded as a useful tool for evaluation of rules. Analysis of confirmation measures with respect to their properties is an active research area. Properties express the user's expectations towards the behavior of measures in particular situations. They group the measures according to similarities in their characteristics. Using the measures which satisfy the desirable properties one can avoid considering unimportant rules.

Therefore, knowledge of which commonly used interestingness measures satisfy certain valuable properties, is of high practical and theoretical importance [14]. Among widely studied properties for Bayesian confirmation measures, there is a group of symmetry properties [2-6], monotonicity properties [1], [8], [14], weak Ex_1 and weak logicality L property [15]. In this article we consider symmetry properties for Bayesian confirmation measures. Though this issue has been taken up by many authors before (e.g., [2], [3], [4]), we propose a new set of desirable symmetry properties that exploits the deep meaning of the confirmation concept. The confirmation measures should express how much more probable the hypothesis is when the premise is present rather than when the premise is absent. Regarding such an interpretation, we justify that evidence symmetry, hypothesis symmetry and the combination of them both, are the only truly desirable symmetry properties. The paper is organized as follows. In the next section there are preliminaries on rules and their quantitative description. In section 3, we discuss the concept of confirmation and its interpretation. Section 4 describes the approaches to symmetry properties in the literature. The proposition of a new set of desirable symmetries is introduced in section 5. Finally, the last section provides conclusions.

2 Preliminaries

A rule induced from a dataset U shall be denoted by $E \to H$ (read as "if E, then H"). It consists of a premise (evidence) E and a conclusion (hypothesis) H. Throughout the paper we shall use the following notation corresponding to a 2x2 contingency table of the premise and the conclusion (Table (1)):

- a is the number of positive examples to the rule, i.e., the number of objects in U satisfying both the premise and the conclusion of the rule,
- b is the number of objects in U not satisfying the rule's premise, but satisfying its conclusion,
- c is the number of counterexamples i.e. objects in U satisfying the premise but not the conclusion of the rule,
- d is the number of objects in U that do not satisfy neither the premise nor the conclusion of the rule.

The cardinality of the dataset U, denoted by $|U|$, is the sum of a, b, c and d.

Table 1. Contingency table of E and H

	H	$\neg H$	Σ
E	a	c	$a+c$
$\neg E$	b	d	$b+d$
Σ	$a+b$	$c+d$	$\lvert U \rvert$

Reasoning in terms of a, b, c and d is natural and intuitive for data mining techniques since all observations are gathered in some kind of an information table describing each object by a set of attributes. However, a, b, c and d can also be regarded as frequencies that can be used to estimate probabilities: e.g., the probability of the premise is expressed as $Pr(E) = (a + c)/|U|$, and the probability of the conclusion as $Pr(H) = (a + b)/|U|$. Moreover, conditional probability of the conclusion given the premise is $Pr(H|E) = a/(a + c)$.

3 Property of Bayesian confirmation

Formally, an interestingness measure $c(H, E)$ has the property of Bayesian confirmation if and only if it satisfies the following BC (1) conditions:

$$c(H, E) \begin{cases} > 0 & if \quad Pr(H|E) > Pr(H), \\ = 0 & if \quad Pr(H|E) = Pr(H), \\ < 0 & if \quad Pr(H|E) < Pr(H). \end{cases} \tag{1}$$

The BC definition identifies confirmation with an increase in the probability of the conclusion H provided by the premise E, neutrality with the lack of influence of the premise E on the probability of conclusion H, and finally disconfirmation with a decrease of probability of the conclusion H imposed by the premise E [6], [10].

A logically equivalent way to express the BC conditions is [6], [10]:

$$c(H, E) \begin{cases} > 0 & if \quad Pr(H|E) > Pr(H|\neg E), \\ = 0 & if \quad Pr(H|E) = Pr(H|\neg E), \\ < 0 & if \quad Pr(H|E) < Pr(H|\neg E). \end{cases} \tag{2}$$

To avoid ambiguity, we shall denote the above formulation (2) as BC'. Notice that according to BC, E confirms H when E raises the probability of H, while, according to BC', E raises the probability of H if the probability of H given E is higher than the probability of H given non E.

Measures that possess the property of Bayesian confirmation are referred to as *confirmation measures* or *measures of confirmation*. For a given rule $E \to H$, interestingness measures with the property of confirmation express the credibility of the following proposition: H is satisfied more frequently when E is satisfied, rather than when E is not satisfied.

Let us stress that the BC conditions (or BC' equivalently) do not impose any constraints on the confirmation measures except for requiring when the measures should obtain positive or negative values. As a result many alternative, non-equivalent measures of confirmation have been proposed [3], [5], [8], [14].

To help to handle the plurality of Bayesian confirmation measures, many authors have considered desirable properties of such measures. Analysis of measures with respect to their properties is a way to distinguish measures that behave according to user's expectations. An important group of properties constitute symmetry properties considered by many authors e.g., Carnap [2], Eells and Fitelson [4], Crupi et al. [3]. Though the literature on symmetries is rich, we claim that there is a need to propose a new approach to the analysis of the symmetry properties that exploits the deep meaning of the confirmation concept. In fact, a confirmation measure should give an account of the credibility that it is more probable to have the conclusion when the premise is present, rather than when the premise is absent. Following that interpretation we propose a new set of desirable symmetry properties for Bayesian confirmation measures.

4 Properties of symmetry

The work of Carnap [2] had inspired many authors that took up the symmetry topic. In particular, Eells and Fitelson have analysed in [4] a set of well-known confirmation measures from the viewpoint of the following four properties of symmetry:

- evidence symmetry ES: $c(H,E) = -c(H,\neg E)$
- hypothesis symmetry HS: $c(H,E) = -c(\neg H,E)$
- commutativity (inversion) symmetry IS: $c(H,E) = c(E,H)$
- total (evidence-hypothesis)symmetry EHS: $c(H,E) = c(\neg H,\neg E)$

Let us observe that the above symmetries are formed by applying the negation operator to the evidence (ES), to the hypothesis (HS), or both (EHS), as well as switching the position of the evidence and the hypothesis (IS). The research of Eells and Fitelson [4] implies that only hypothesis symmetry HS is a desirable property, while evidence symmetry ES, commutativity symmetry IS and total symmetry EHS are not. As an illustration of their reasoning, they used rules concerning drawing cards from a standard deck. They claim e.g. that the evidence symmetry should be discarded due to the following counterexample: *if the drawn card is the seven of spades then the card is black*. Obviously, the *seven of spades* confirms that *the card is black* to a greater extent than *the not-seven of spades* disconfirms the same hypothesis. As a result the equality in evidence symmetry is found unattractive by Eells and Fitelson. Thus, in their opinion, an acceptable measure of Bayesian confirmation should not satisfy the evidence symmetry (i.e. for some situation $c(H,E) \neq -c(H,\neg E)$). Analogous reasoning can be conducted with respect to the other symmetries analyzed in [4].

Recently, Crupi et at. [3] have argued for an extended and systematic treatment of the issue of symmetry properties. They propose to analyse a confirmation measure $c(H,E)$ with respect to seven symmetries being all combinations obtained by applying the negation operator to the premise, hypothesis or both, and/or by inverting E and H:

- $ES(H,E) : c(H,E) = -c(H,\neg E)$
- $HS(H,E) : c(H,E) = -c(\neg H,E)$
- $EIS(H,E) : c(H,E) = -c(\neg E,H)$
- $HIS(H,E) : c(H,E) = -c(E,\neg H)$
- $IS(H,E) : c(H,E) = c(E,H)$
- $EHS(H,E) : c(H,E) = c(\neg H,\neg E)$
- $EHIS(H,E) : c(H,E) = c(\neg E,\neg H)$

Moreover, Crupi et al. [3] go even further as the analysis is conducted separately for the case of confirmation (i.e. when $Pr(H|E) > Pr(H)$) and for the case of disconfirmation (i.e. when $Pr(H|E) < Pr(H)$). Such approach results in 14 symmetry properties. Using examples (analogical to Eells and Fitelson) of drawing cards from a standard deck, Crupi et al. point out which of the symmetries are desired and which are definitely unwanted. For instance, they concur with the results of Eells and Fitelson regarding the inversion symmetry only in case of confirmation. Crupi et al. claim that IS is desirable in case of disconfirmation, as for an exemplary rule: *if the drawn card is an Ace, then it is a face*, the strength with which an *Ace* disconfirms *face* is the same as the strength with which the *face* disconfirms an *Ace*, i.e. $c(H,E) = c(E,H)$. The results obtained by Crupi et al. point that in case of confirmation only the HS, HIS and $EHIS$ are the desirable properties. In case of disconfirmation, they favour HS, EIS and IS properties, finding all other symmetries as unattractive.

5 A new set of symmetry properties

Let us observe that the approaches of Eells and Fitelson [4] as well as Crupi et al. [3] mainly concentrate on entailment and refutation of the hypothesis by the premise. This, however, boils the concept of confirmation down only to situations where there are no counterexamples (entailment) and where there are no positive examples to a rule (refutation).

In our opinion, the concept of confirmation is much broader than a simple analysis whether there are counterexamples to a rule or not. In fact, according to the BC' interpretation of the confirmation concept, a confirmation measure should give an account of the credibility that it is more probable to have the conclusion when the premise is present, rather than when the premise is absent. This means that we should look at confirmation from the perspective of passing from a situation where the premise is absent to the situation where the premise is present. Then, the increase of confirmation (i.e. the difference in conditional probabilities $Pr(H|E)$ and $Pr(H|\neg E)$) becomes important, not just the absence or presence of counterexamples.

Analogically, for disconfirmation a confirmation measure $c(H,E)$ should express how much it is less probable to have the hypothesis when the premise is present rather than when the premise is absent. Again, we should, thus, pass from the situation where the premise is absent to the situation where the premise is present, just like we did in case of confirmation. Therefore, we postulate to consider the symmetry properties together for cases of confirmation and disconfirmation. There is no need to treat them differently as they both consider passing from $Pr(H|\neg E)$ to $Pr(H|E)$.

Let us now conduct the analysis aiming at determining which symmetry properties are desirable and which are unattractive regarding the deep meaning of confirmation concept. First, let us consider the evidence symmetry (ES). Analyzing ES we need to verify whether the equation $c(H,E) = -c(H,\neg E)$ is desirable or not. Let us examine both sides of this equation using an exemplary scenario α where the values of contingency table of E and H are: $a = 100$, $b = 99$, $c = 0$, $d = 1$. Let us observe, that for $c(H, E)$ we have that $Pr(H|\neg E) = 0.99$ and $Pr(H|E) = 1$, which gives us a 1% increase of confirmation. On the other hand, for $c(H, \neg E)$ we get exactly the same components but the other way around: $Pr(H|E) = 1$ and $Pr(H|\neg E) = 0.99$, which results in 1% decrease of confirmation. Thus, clearly the confirmation of a rule $E \rightarrow H$ should be of the same value but of the opposite sign as the confirmation of a $\neg E \rightarrow H$ rule. Therefore, we can conclude that the evidence symmetry is desirable.

This result is in opposition to what Eells and Fitelson [4], and Crupi et al. [3] advocated for. It is due to the fact that they treat the entailment of the conclusion by the premise (i.e. situation where there are no counterexamples to the rule) as the maximal confirmation, whereas we consider the increase of confirmation when passing from the absence of the premise to its presence. For the exemplary rule of Eells and Fitelson: *if the drawn card is the seven of spades then the card is black*, the conditional probabilities are the following $Pr(H|E) = 1$ and $Pr(H|\neg E) = 51/103$. They claim that just because E definitely confirms H, while E does not definitely disconfirms H, we should regard the ES as unattractive. However, for the rules $E \rightarrow H$ and $\neg E \rightarrow H$, if we interpret the concept of confirmation as expressing how much more probable is the hypothesis (to have *the black card* both for $E \rightarrow H$ and $\neg E \rightarrow H$), when the evidence (we have drawn *the seven of spades* for $E \rightarrow H$ and we have drawn *not-the seven of spades* for $\neg E \rightarrow H$) is realized rather than when it is not realized (we have drawn *not-the seven of spades* for $E \rightarrow H$ and we have drawn *the seven of spades* for $\neg E \rightarrow H$), the confirmation has the same absolute value but opposite sign. Thus, we claim that using the deep meaning and interpretation of the confirmation concept, evidence symmetry is a desirable property for Bayesian confirmation measures.

We have conducted analogous analysis for all other symmetries. The results are gathered in Table 2. The set of desirable properties contains only the evidence symmetry, the hypothesis symmetry and their composition i.e. the evidence-hypothesis symmetry. This implies that a valuable Bayesian confirmation measure should satisfy only those symmetry properties. By defining the new set of symmetry properties we gain a tool for assessing the quality of confirmation measures.

6 Conclusions

Bayesian confirmation measures constitute an important group of measures for evaluation of rules induced from data. A valid research area concerns the properties of confirmation measures. Analysis of measures with respect to their properties allows to determine measures that behave according to the user's expectations. It is also a way to handle the plurality of measures and point the most appropriate ones for particular applications.

This article concentrated on the group of symmetry properties. Our analysis was conducted regarding that confirmation measures should reflect how much

Table 2. New symmetry properties

ES	YES *for any* (H,E) $c(H,E) = -c(H,\neg E)$
HS	YES *for any* (H,E) $c(H,E) = -c(\neg H,E)$
EIS	NO *for some* (H,E) $c(H,E) \neq -c(\neg E,H)$
HIS	NO *for some* (H,E) $c(H,E) \neq -c(E,\neg H)$
IS	NO *for some* (H,E) $c(H,E) \neq c(E,H)$
EHS	YES *for any* (H,E) $c(H,E) = c(\neg H,\neg E)$
$EHIS$	NO *for some* (H,E) $c(H,E) \neq c(\neg E,\neg H)$

more it is probable to have the conclusion H when the premise E is present rather than when it is absent. Such interpretation of the confirmation concept led to our proposition of a new set of desirable symmetry properties. We claim that only symmetries formed by applying the negation operator to the rule's premise, conclusion or both (i.e. ES, HS and EHS) are desirable. Properties IS, EIS, HIS, $EHIS$ are unattractive. Thus, valuable confirmation measures should only satisfy ES, HS and EHS.

Consequently, our future research will concentrate on verification which of the commonly used confirmation measures satisfy ES, HS and EHS, not enjoying the other symmetries at the same time. Moreover, experiments on real datasets shall be performed to show the advantages of using such measures.

References

1. Brzezinska, I., Greco, S., Slowinski, R.: Mining Pareto-optimal rules with respect to support and anti-support. Engineering Applications of Artificial Intelligence Vol. 20(5) 587-600 (2007)
2. Carnap, R.: Logical Foundations of Probability, 2nd ed., University of Chicago Press, Chicago (1962)
3. Crupi, V., Tentori, K., Gonzalez, M.: On Bayesian measures of evidential sup-port: Theoretical and empirical issues. Philosophy of Science, vol. 74, 229-252 (2007)
4. Eells, E., Fitelson, B.: Symmetries and asymmetries in evidential support. Philosophical Studies, vol. 107, no. 2, 129-142 (2002)

5. Fitelson, B.: The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity. Philosophy of Science, vol. 66, 362-378 (1999)
6. Fitelson, B.: Studies in Bayesian Confirmation Theory. Ph.D. Thesis, University of Wisconsin, Madison (2001)
7. Geng, L., Hamilton, H.J.: Interestingness Measures for Data Mining: A Survey. ACM Computing Surveys, vol. 38, no. 3, article 9 (2006)
8. Greco, S., Pawlak, Z., Slowinski, R.: Can Bayesian confirmation measures be useful for rough set decision rules? Engineering Applications of Artificial Intelligence, 17: 345-361 (2004)
9. Hilderman, R. and Hamilton, H.: Knowledge Discovery and Measures of Interest. Kluwer Academic Publishers (2001)
10. Maher, P.: Confirmation Theory. The Encyclopedia of Philosophy (2nd ed.). Macmillan Reference, USA (2005)
11. McGarry, K.: A survey of interestingness measures for knowledge discovery. The Knowledge Engineering Review, vol. 20, no.1, 39-61 (2005)
12. Slowinski, R., Brzezinska, I., Greco, S.: Application of Bayesian confirmation measures for mining rules from support-confidence Pareto-optimal set. [Invited paper in]: L.Rutkowski, R.Tadeusiewicz, L.A.Zadeh, J.Zurada (eds.), Artificial Intelligence and Soft Computing (ICAISC 2006). LNAI 4029, Springer-Verlag, Berlin, pp. 1018-1026 (2006)
13. Slowinski, R., Szczech, I., Urbanowicz, M., Greco, S.: ining Association Rules with respect to Support and Anti-support - Experimental Results, in: Proc. of RSEISP2007, TheRough Sets and Emerging Systems Paradigms 2007 International Conference. LNAI, vol. 4585, Springer-Verlag, Berlin, 534-543 (2007)
14. Szczech, I.: Multicriteria Attractiveness Evaluation of Decision and Association Rules, Transactions on Rough Sets X, LNCS 5656, pp.197-274 (2009)
15. Szczech, I., Greco, S. Slowinski, R.: New property for rule interestingness measures In: Ganzha, M., Maciaszek, L., Paprzycki, M. (eds), Proceedings of the Federated Confer-ence on Computer Science and Information Systems FedCSIS, pp.103-108 (2011)