Alternative normalization schemas for Bayesian confirmation measures

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

[1] Faculty of Economics, University of Catania, Corso Italia, 55, 95129 Catania, Italy
salgreco@unict.it
[2] Institute of Computing Science, Poznan University of Technology, 60-965 Poznan, Poland
{Izabela.Szczech, Roman.Slowinski}@cs.put.poznan.pl
[3] Systems Research Institute, Polish Academy of Sciences, 01-447 Warsaw, Poland

Abstract. Analysis of rule interestingness measures with respect to their properties is an important research area helping to identify groups of measures that are truly meaningful. In this article, we analyze property Ex_1, of preservation of extremes, in a group of confirmation measures. We consider normalization as a mean to transform them so that they would obtain property Ex_1 and we introduce three alternative approaches to the problem: an approach inspired by Nicod, Bayesian, and likelihoodist approach. We analyze the results of the normalizations of seven measures with respect to property Ex_1 and show which approaches lead to the desirable results. Moreover, we extend the group of ordinally non-equivalent measures possessing valuable property Ex_1.

Keywords:Normalization, Bayesian confirmation measures, property Ex_1.

1 Introduction

One of the main objectives of data mining process is to identify "*valid, novel, potentially useful, and ultimately comprehensible knowledge from databases*" [6]. When mining large datasets, the number of knowledge patterns, often expressed in a form of "*if..., then...*" rules, can easily be overwhelming rising an urgent need to identify the most useful ones. Addressing this issue, various quantitative measures of rule interestingness (attractiveness) have been proposed and studied, e.g., *support*, $confidence$, $lift$ (for a survey on interestingness measures see [1], [9], [13]). The literature is a rich resource of ordinally non-equivalent measures that reflect different characteristics of rules. There is no agreement which measure is the best. To help to analyze objective measures, some `properties` have been proposed, expressing the user's expectations towards the behavior of measures in particular situations. Properties of measures group the measures according to similarities in their characteristics. Using the measures which satisfy the desirable properties, one can avoid considering unimportant rules. Different properties were surveyed in [5], [9], [10], [19]. In this paper, we focus on two desirable properties: *property of confirmation* quantifying the degree to which the premise of the rule provides evidence for or against the conclusion [8], [2], and *property* Ex_1 guaranteeing that any conclusively confirmatory rule, for which the premise ϕ entails the conclusion ψ (i.e. such that $\phi \models \psi$), is assigned a higher value of measure than any rule which is not conclusively confirmatory, and that any conclusively disconfirmatory rule, for which ϕ refutes ψ (i.e. such that $\phi \models \neg\psi$), is assigned a lower value than any rule which is not conclusively disconfirmatory [4], [11]. Though property Ex_1 is so intuitively clear and required, it is not satisfied by many popular measures. Looking for a way of transforming seven chosen confirmation measures, so they would fulfill Ex_1, Crupi et al. [4] proposed to normalize them. Their approach, however, is only one of many ways to handle this issue. In this paper, we extend their analysis and propose three other alternative normalization schemas. Moreover, we analyze them with respect to property Ex_1 presenting and commenting the results of application of different normalizations to the chosen measures. Furthermore, as the result of our work, there also emerges a set of interestingness measures (alternative to one of Crupi et al.) that satisfy desirable properties and thus extend the family of valuable measures.

2 Preliminaries

A *rule* induced from a dataset U shall be denoted by $\phi \rightarrow \psi$. It consists of a premise (evidence) ϕ and a conclusion (hypothesis) ψ. A rule is a logical sentence in the sense that elementary conditions on attributes are connected by logical "and", on both sides of the rules. However, on a particular attribute they can concern evaluations expressed on nominal, ordinal or cardinal scales. For each rule we consider the number of objects which satisfy both the premise and the conclusion, only the premise, only the conclusion, neither the premise nor the conclusion. However, this does not mean that in our data each object can assume only values e.g., ψ or $\neg\psi$. It simply means that when we evaluate a rule of the type "if ϕ, then ψ" we take into account set of objects that satisfy ψ and a set of objects that do not satisfy ψ.

In general, by $sup(\gamma)$ we denote the number of objects in the dataset for which γ is true. Thus, $sup(\phi \rightarrow \psi)$ is the number of objects satisfying both the premise and the conclusion of a $\phi \rightarrow \psi$ rule. Moreover, the following notation shall be used throughout the paper: $a = sup(\phi \rightarrow \psi)$, $b = sup(\neg\phi \rightarrow \psi)$, $c = sup(\phi \rightarrow \neg\psi)$, $d = sup(\neg\phi \rightarrow \neg\psi)$. Observe that b can be interpreted as the number of objects that do not satisfy the premise but satisfy the conclusion of the $\phi \rightarrow \psi$ rule. Analogous observations hold for c and d. Moreover, the following relations occur: $a+c = sup(\phi)$, $a+b = sup(\psi)$, $b+d = sup(\neg\phi)$, $c+d = sup(\neg\psi)$, and the cardinality of the dataset U, denoted by $|U|$, is the sum of a, b, c and d.

3 Property of Bayesian confirmation

Formally, an interestingness measure $c(\phi \rightarrow \psi)$ has the property of Bayesian confirmation (or simply confirmation) iff it satisfies the following conditions:

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

where $Pr(\psi)$ denotes the probability of ψ, and $Pr(\psi|\phi)$ is the conditional probability of ψ given ϕ.

This definition identifies confirmation with an increase in the probability of the conclusion provided by the premise, neutrality with the lack of influence of the premise on the probability of conclusion, and disconfirmation with a decrease of probability of the conclusion imposed by the premise [2]. Under the "closed world assumption" adopted in inductive reasoning, and because U is a finite set, it is legitimate to estimate probabilities in terms of frequencies, e.g., $Pr(\psi) = sup(\psi)/|U| = (a + b)/|U|$. In consequence, we can define the conditional probability as $Pr(\psi|\phi) = Pr(\phi \wedge \psi)/Pr(\phi)$, and it can be regarded as $sup(\phi \to \psi)/sup(\phi)$ (i.e. $a/(a+c)$). Thus, the above condition can be re-written:

$$c(\phi \to \psi) \begin{cases} > 0 & if \quad \frac{a}{a+c} > \frac{a+b}{|U|}, \\ = 0 & if \quad \frac{a}{a+c} = \frac{a+b}{|U|}, \\ < 0 & if \quad \frac{a}{a+c} < \frac{a+b}{|U|}. \end{cases} \tag{2}$$

Measures that possess the property of confirmation are referred to as *confirmation measures* or *measures of confirmation*. They quantify the degree to which the premise ϕ provides "support for or against" the conclusion ψ [8]. By using the attractiveness measures that possess this property one can filter out rules which are misleading and disconfirm the user, and this way, limit the set of induced rules only to those that are meaningful [18]. The only constraints (2) that the property of confirmation puts on a measure are that it assigns positive values in the situation when confirmation occurs, negative values in case of disconfirmation and zero otherwise. As a result, many alternative, non-equivalent measures of confirmation have been proposed. Most commonly used ones are gathered in Table (1) (selection provided in [4]):

4 Property Ex_1 of preservation of extremes

To handle the plurality of alternative confirmation measures Crupi et al. [4] have proposed a property (principle) Ex_1 resorting to considering inductive logic as an extrapolation from classical deductive logic. On the basis of classical deductive logic they construct a function v:

$$v(\phi, \psi) = \begin{cases} the \quad same \quad positive \quad value \quad if \quad \phi \models \psi, \\ the \quad same \quad negative \quad value \quad if \quad \phi \models \neg\psi, \\ 0 \quad otherwise. \end{cases} \tag{3}$$

For any argument (ϕ, ψ) v assigns it the same positive value (e.g., +1) if and only if the premise ϕ of the rule entails the conclusion ψ (i.e. $\phi \models \psi$). The same

Table 1. Common confirmation measures

$D(\phi \to \psi) = Pr(\psi\|\phi) - Pr(\psi) = \frac{a}{a+c} - \frac{a+b}{\|U\|}$	Carnap [2]
$S(\phi \to \psi) = Pr(\psi\|\phi) - Pr(\psi\|\neg\phi) = \frac{a}{a+c} - \frac{b}{b+d}$	Christensen [3]
$M(\phi \to \psi) = Pr(\phi\|\psi) - Pr(\phi) = \frac{a}{a+b} - \frac{a+c}{\|U\|}$	Mortimer [14]
$N(\phi \to \psi) = Pr(\phi\|\psi) - Pr(\phi\|\neg\psi) = \frac{a}{a+b} - \frac{c}{c+d}$	Nozick [16]
$C(\phi \to \psi) = Pr(\phi \wedge \psi) - Pr(\phi)Pr(\psi) = \frac{a}{\|U\|} - \frac{(a+c)(a+b)}{\|U\|^2}$	Carnap [2]
$R(\phi \to \psi) = \frac{Pr(\psi\|\phi)}{Pr(\psi)} - 1 = \frac{a\|U\|}{(a+c)(a+b)} - 1$	Finch [7]
$G(\phi \to \psi) = \frac{Pr(\neg\psi\|\phi)}{Pr(\neg\psi)} = 1 - \frac{c\|U\|}{(a+c)(c+d)}$	Rips [17]

value but of opposite sign (e.g., -1) is assigned if and only if the premise ϕ refutes the conclusion ψ (i.e. $\phi \models \neg\psi$). In all other cases (i.e. when the premise is not conclusively confirmatory nor conclusively disconfirmatory for the conclusion) function v obtains value 0.

From definition, any confirmation measure $c(\phi \to \psi)$ agrees with function $v(\phi, \psi)$ in the way that if $v(\phi, \psi)$ is positive then the same is true for $c(\phi \to \psi)$, and when $v(\phi, \psi)$ is negative, so is $c(\phi \to \psi)$. According to Crupi et al., the relationship between the logical implication or refutation of ψ by ϕ, and the conditional probability of ψ subject to ϕ should go further and demand fulfillment of the principle Ex_1 [4]:

$$if \quad v(\phi_1, \psi_1) > v(\phi_2, \psi_2), \quad then \quad c(\phi_1 \to \psi_1) > c(\phi_2 \to \psi_2). \tag{4}$$

Property Ex_1 is desirable for any interestingness measure as it guarantees that the measure will assign a greater value to any conclusively confirmatory rule (i.e. such that $\phi \models \psi$, e.g., *if x is seven of spades then x is black*) than to any rule which is not conclusively confirmatory (e.g., *if x is black then x is seven of spades*). Moreover, rules that are conclusively disconfirmatory (i.e. such that $\phi \models \neg\psi$, e.g., *if x is seven of spades then x is red*) will obtain smaller values of interestingness measures than rules which is not conclusively disconfirmatory (e.g., *if x is black then x is seven of spades*).

5 Normalization of confirmation measures

Having observed that confirmation measures D, S, M, N, C, R, G (defined earlier on) are contrary to Ex_1, Crupi et al. [4] proposed to normalize them by dividing each of them by the maximum (minimum, respectively) the measure obtains when $\phi \models \psi$, i.e. when the rule's premise entails its conclusion ($\phi \models \neg\psi$, respectively). Determining the maximum or minimum that a confirmation measure obtains in case of confirmation or disconfirmation has, however, no unique interpretation, and the approach applied by Crupi et al. is only one of many ways to handle this issue. We shall now propose and analyze four (including the approach of Crupi et al.) alternative schemas allowing to determine the maximum (or minimum) of any confirmation measure in those two situations. We denote by a', b', c' and d' the values of a, b, c and d, respectively, in case of maximizing or minimizing the confirmation. Each of the analyzed schemas eventually leads to a different normalization, as we divide the original measures by their maximum or minimum calculated using alternative schemas. Therefore next, we will present and discuss results of normalization of measures D, S, M, N, C, R, G using those approaches.

5.1 Approach inspired by Nicod

The Nicod's criterion presented in [15] says that an evidence confirms a rule $\phi \to \psi$ if and only if it satisfies both the premise and the conclusion of the rule, and disconfirms it if and only if it satisfies the premise but not the conclusion of the rule. Thus, objects for which the premise and the conclusion is supported are considered as positive examples for the rule and objects satisfying the premise but not the conclusion are counter-examples. Moreover, according to Nicod's criterion, an evidence that does not satisfy the premise is neutral with respect to the rule. It means that objects for which the premise is not satisfied are irrelevant to the rule, no matter whether they support the conclusion or not. Now, let us propose a schema, based on Nicod's criterion, for determination of maximum (or minimum) of a confirmation measure. Following Nicod's directives, the only objects that are relevant to a rule are positive examples and counter-examples. It brings us to an observation that a measure will obtain its maximum when all counter-examples change into positive examples. It means that the number of positive examples should take over all counter-examples (i.e. $a' = a + c$), and the number of counter-examples should drop to 0 (i.e. $c' = 0$). The number of evidence which are irrelevant to the rule should remain unchanged (i.e. $b' = b$ and $d' = d$). The schema for determination of the minimal value is analogous. Putting all the considerations together we obtain the approach, inspired by Nicod, to determine the extremes of any measure (Table 2).

Table 2. Schemas for determination of the extremes of any measure

Nicod's		Bayesian		Likelihoodist		Crupi's et al.	
Max	Min	Max	Min	Max	Min	Max	Min
$a' = a + c$	$a' = 0$	$a' = a + b$	$a' = 0$	$a' = a + c$	$a' = 0$	$a' = a + c$	$a' = 0$
$b' = b$	$b' = b$	$b' = 0$	$b' = a + b$	$b' = 0$	$b' = b + d$	$b' = b - c$	$b' = a + b$
$c' = 0$	$c' = a + c$	$c' = 0$	$c' = c + d$	$c' = 0$	$c' = a + c$	$c' = 0$	$c' = a + c$
$d' = d$	$d' = d$	$d' = c + d$	$d' = 0$	$d' = b + d$	$d' = 0$	$d' = c + d$	$d' = d - a$

5.2 Bayesian approach

Bayesian approach is related to the idea that the evidence confirms the hypothesis, if the hypothesis is more frequent with the evidence rather than without the evidence. Analogously, the evidence disconfirms the hypothesis, if ¬ hypothesis is more frequent with the evidence rather than without the evidence. Thus, determination of measure's extremes based on this approach should consider a rule from the perspective of its conclusion. Following Bayesian approach, let us observe that for a rule *if x is a raven then x is black* [12] a measure will obtain its maximum if *all black non-ravens* change into *black ravens* (i.e. $a' = a + b$ and $b' = 0$), and all *non-black ravens* change into *non-black non-ravens* (i.e. $d' = c+d$ and $c' = 0$). It is due to the fact that when there are no *black non-ravens* (i.e. $b' = 0$), the hypothesis of being *black* is more frequent with the premise of being a *raven* rather than with ¬premise of being a *non raven*, which means that the premise confirms the rule's conclusion. Moreover, when there are no *non-black ravens* (i.e. $c' = 0$), the ¬hypothesis of being *non-black* is disconfirmed as it is more frequent with the ¬premise of being a *non-raven* rather than with the premise of being a *raven*. Disconfirmation of ¬hypothesis is desirable as it results in confirmation of the hypothesis. The considerations about determination of the minimal value are analogous. The Bayesian approach to determination of a measure's maximum or minimum is summarized in Table 2.

5.3 Likelihoodist approach

The likelihoodist approach is based on the idea that the evidence confirms the hypothesis, if the evidence is more frequent with the hypothesis rather than without the hypothesis, and in this context, analogously, the evidence disconfirms the hypothesis, if the evidence is more frequent without the hypothesis rather than with the hypothesis. Thus, one can informally say that likelihoodists look at the rule from the perspective of its premise. According to likelihoodist approach, for a rule *if x is a raven then x is black* [12] a measure will obtain its maximum if all *non-black ravens* change into *black ravens* (i.e. $a' = a + c$ and $c' = 0$), and all *black non-ravens* change into *non-black non-ravens* (i.e. $d' = b + d$ and $b' = 0$). It results from the fact that when there are no *non-black ravens* (i.e. $c' = 0$), the evidence of being a *raven* is more frequent with the hypothesis of being *black* rather than with ¬hypothesis of being *non black*, which means that the premise confirms the rule's conclusion. Moreover, when there are no *black non-ravens* (i.e. $b' = 0$), the ¬evidence of being a *non-raven* is more frequent with the ¬hypothesis of being *non-black* rather than with the hypothesis of being *black*. Thus, we can conclude that hypothesis is disconfirmed by the ¬premise and as a result of that the hypothesis is confirmed by the premise. Determination of the minimal value of confirmation measure is analogous. The whole likelihoodist approach to calculating the measure's extremes is presented in Table 2.

5.4 Approach of Crupi et al.

Having proved that none of the measures: D, S, M, N, C, R nor G satisfies the desirable property Ex_1, Crupi et al. [4] showed an easy way to transform them into measures that do fulfill Ex_1. They presented formulas to which the considered measures boil down when $\phi \models \psi$ and when $\phi \models \neg\psi$, and proposed to normalize the measures by dividing them by the obtained formulas. Their article, however, does not provide any methodological schema to determine the measure's extremes - only the calculated formulas are given. Since, the approach of Crupi et al. brings such interesting results, we have analyzed it thoroughly in terms of our notation, i.e. a, b, c and d, and came up with a clear schema (see Table 2) that can be used to determine the extremes of any measure.

According to Crupi et al., dividing a measure by the formula obtained when $\phi \models \psi$ produces the normalized measure in case of confirmation (i.e. when $Pr(\psi|\phi) \geq Pr(\psi)$), and the division by absolute value of the formula obtained when $\phi \models \neg\psi$ gives the normalized measure in case of disconfirmation (i.e. when $Pr(\psi|\phi) < Pr(\psi)$). Interestingly, it turned out that the considered measures all gave the same result after that transformation, i.e. $D_{norm} = S_{norm} = M_{norm} = N_{norm} = C_{norm} = R_{norm} = G_{norm}$. Crupi et al. labeled the newly obtained measure of confirmation Z. In case of confirmation $Z = G$ and in case of disconfirmation $Z = R$. Crupi et al. [4] have proved that measure Z and all confirmation measures ordinally equivalent to Z satisfy property Ex_1.

6 Results of applying normalization schemas to measures

Each of the schemas presented by us to determine the extremes of measures eventually results in a different normalization. Table 3 presents them all. For the sake of the presentation, the definitions of the analyzed measures were simplified by basic mathematical transformations (column 1). The next four columns contain results for different normalization schemas, for each measure there are two rows containing the normalized measure in case of confirmation (the first row) and disconfirmation (the second row). The notation we used assumes that lower indexes signify the applied normalization (N stands for Nicod, B for Bayesian, L for likelihoodist, and C for Crupi et al.), and that the case of confirmation is marked by a "+" and the case of disconfirmation by a "-" (e.g., D_{N+} stands for measure D normalized in case of confirmation, using the approach inspired by Nicod).

Since the normalization of Crupi et al. was introduced as a tool for transforming the measures so they would satisfy the property Ex_1, we have analyzed the results of different normalizations of measures D, S, M, N, C, R, G from the view point of this property. Let us observe, that Ex_1 is satisfied by any confirmation measure that obtains its maximal value when there are no counterexamples to the rule and its minimal value when there are no positive examples to the rule. These two conditions can be regarded as sufficient for proving the possession of Ex_1 by measure $c(\phi \rightarrow \psi)$.

Theorem 1. *All confirmation measures D, S, M, N, C, R, G normalized using approach inspired by Nicod or approach of Crupi et al. satisfy property Ex_1. Moreover normalization using Bayesian approach gives measures satisfying Ex_1 only in case of measure D, R and G, whereas using likelihoodist approach, Ex_1 does not hold for any of the considered measures.*

Table 3. Results of alternative normalization approaches

Original measure	Nicod's norm.	Bayesian norm.	Likelihoodist norm.	Crupi et al norm.
$D(\phi \to \psi) = \frac{ad-bc}{\lvert U\rvert(a+c)}$	$D_{N^+} = \frac{ad-bc}{d(a+c)}$	G	S	G
	$D_{N^-} = \frac{ad-bc}{b(a+c)}$	R	S	R
$S(\phi \to \psi) = \frac{ad-bc}{(a+c)(b+d)}$	D_{N^+}	S	S	G
	D_{N^-}	S	S	R
$M(\phi \to \psi) = \frac{ad-bc}{\lvert U\rvert(a+b)}$	$M_{N^+} = \frac{(ad-bc)(a+b+c)}{d(a+b)(a+c)}$	N	$M_{L^+} = \frac{ad-bc}{(a+b)(b+d)}$	G
	R	N	R	R
$N(\phi \to \psi) = \frac{ad-bc}{(a+b)(c+d)}$	$N_{N^+} = \frac{(ad-bc)(a+b+c)}{(a+b)(a+c)(c+d)}$	N	N	G
	$N_{N^-} = \frac{(ad-bc)(a+c+d)}{(a+b)(a+c)(c+d)}$	N	N	R
$C(\phi \to \psi) = \frac{ad-bc}{\lvert U\rvert^2}$	D_{N^+}	N	S	G
	D_{N^-}	N	S	R
$R(\phi \to \psi) = \frac{ad-bc}{(a+b)(a+c)}$	M_{N^+}	G	M_{L^+}	G
	R	R	R	R
$G(\phi \to \psi) = \frac{ad-bc}{(a+c)(c+d)}$	G	G	G	G
	$G_{N^-} = \frac{(ad-bc)(a+c+d)}{b(a+c)(c+d)}$	R	$G_{L^-} = \frac{ad-bc}{(c+d)(b+d)}$	R

Proof. Possession of property Ex_1 can be verified by putting $c = 0$ and $a = 0$ in the normalized measure and checking whether it's formula boils down to 1 in case $c = 0$ and to -1 in case $a = 0$. The considered measures normalized using approach inspired by Nicod or approach of Crupi et al. are equal to 1 (or -1) when $c = 0$ (or $a = 0$).

The new measures obtained during normalization inspired by Nicod can be regarded as alternative ones with respect to measure Z advocated by Crupi et al. [4]. D_N, S_N, M_N, N_N, C_N, R_N, and G_N are as valuable as Z in terms of possession of Ex_1 and, generally, produce different rankings on rules than Z. It is an important result widening the group of non-equivalent measures satisfying property Ex_1.

Theorem 2. *Measures D_N, S_N, M_N, N_N, C_N, R_N, and G_N (resulting from application of normalization inspired by Nicod) are ordinally non-equivalent to measure Z.*

Proof. Measure f is ordinally equivalent to measure g iff for any rules r_1, r_2:

$$f(r_1)\left\{\begin{matrix} > \\ = \\ < \end{matrix}\right\} g(r_1) \quad iff \quad f(r_2)\left\{\begin{matrix} > \\ = \\ < \end{matrix}\right\} g(r_2). \tag{5}$$

The above condition needs to be fulfilled both in case of confirmation and disconfirmation. For Table 3 it is enough to consider measures D_{N^+}, M_{N^+}, N_{N^+} and G_{N^-}. The situation in which the number of objects in U is distributed over a, b, c and d is called scenario α. In scenario α, rule $r : \phi \to \psi$ is supported by a objects from U. Table 4 contains a counterexample proving that in two exemplary scenarios α_1 and α_2 measures D_{N^+}, and M_{N^+} produce rankings different than measure G. Measure G assigns r_2 greater value than to r_1, whereas measures D_{N^+}, and M_{N^+} rank those rules the other way round. Thus, D_N and M_N are ordinally non-equivalent to measure Z. By the next counterexample in Table 4, let us show that in scenarios α_3 and α_4 measure N_{N^+} produces different ranking than measure G. Observe that measure G assigns r_1 greater value, whereas measures N_{N^+} favors r_2, thus we can conclude that N_N is ordinally non-equivalent to Z. Finally, scenarios α_1 and α_2 from Table 4 prove that measure G_{N^-} produces different ranking than measure R. Here, G_{N^-} assigns r_1 greater value, whereas R favors r_2. Thus, G_N is ordinally non-equivalent to Z.

Table 4. Counterexamples showing ordinal non-equivalence of measures D_N, M_N, N_N, G_N and measure Z

Counterexample concerning measures D_{N^+} and M_{N^+}								
α_1	$a = 90$	$b = 8$	$c = 1$	$d = 1$	$\|U\| = 100$	$D_{N^+}(r_1)$=**0.90**	$M_{N^+}(r_1)$=**0.91**	$G(r_1) = 0.45$
α_2	$a = 70$	$b = 16$	$c = 4$	$d = 10$	$\|U\| = 100$	$D_{N^+}(r_2) = 0.86$	$M_{N^+}(r_2) = 0.90$	$G(r_1)$=**0.61**
Counterexample concerning measure N_{N^+}								
α_3	$a = 70$	$b = 1$	$c = 19$	$d = 1$	$\|U\| = 100$	$N_{N^+}(r_1) = 0.33$		$G(r_1)$=**0.26**
α_4	$a = 55$	$b = 2$	$c = 26$	$d = 17$	$\|U\| = 100$	$N_{N^+}(r_2)$=**0.37**		$G(r_1) = 0.25$
Counterexample concerning measure G_{N^-}								
α_1	$a = 90$	$b = 8$	$c = 1$	$d = 1$	$\|U\| = 100$	$G_{N^-}(r_1)$=**5.18**		$R(r_1) = 0.009$
α_2	$a = 70$	$b = 16$	$c = 4$	$d = 10$	$\|U\| = 100$	$G_{N^-}(r_2) = 3.22$		$R(r_1)$=**0.099**

7 Conclusions

Analysis of interestingness measures with respect to their properties is an important research area helping to identify groups of measures that are truly meaningful. In this article, we have focused on possession of property Ex_1 in a group of popular confirmation measures. Normalization of measures as a way to transform them so that they would obtain property Ex_1 has been considered. A crucial step of such normalization is determination of the extremes of the measures in case of confirmation and disconfirmation. In this article, we have introduced three alternative approaches to this problem, i.e. an approach inspired by Nicod, Bayesian, and likelihoodist approach. All these approaches, as well as that of Crupi et al., lead to different results and normalizations, as they consider the concept of confirmation from different perspectives. A set of seven confirmation measures, earlier analyzed by Crupi et al., has been normalized using those four schemas. We have analyzed the results of the normalizations with respect to property Ex_1. The conclusions that we obtained show that approach inspired by Nicod, as well as approach of Crupi et al., give normalized measures with property Ex_1 in cases of all of the considered measures. Moreover, we have proved that measures obtained through those normalizations are ordinally non-equivalent. Thus, we have extended the group of measures possessing valuable property Ex_1.

References

1. Bramer, M., 2007. *Principles of Data Mining*, Springer-Verlag, New York Inc.
2. Carnap, R., 1962. *Logical Foundations of Probability* 2nd ed. University of Chicago Press, Chicago.
3. Christensen, D., 1999. Measuring confirmation. *Journal of Philosophy* 96, 437-461.
4. Crupi, V., Tentori, K., Gonzalez, M., 2007. On Bayesian measures of evidential support: Theoretical and empirical issues. *Philosophy of Science*.
5. Eells, E., Fitelson, B., 2002. Symmetries and asymmetries in evidential support. *Philosophical Studies*, 107 (2): 129-142.
6. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., 1996. From data mining to knowledge discovery: an overview. [In]: Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., Uthursamy, R. (eds) *Advances in Knowledge Discov. and Data Mining*, AAAI Press 1-34.
7. Finch, H.A., 1999. Confirming Power of Observations Metricized for Decisions among Hypotheses. *Philosophy of Science*, 27, 293-307 and 391-404.
8. Fitelson, B., 2001. *Studies in Bayesian Confirmation Theory*. Ph.D. Thesis, University of Wisconsin, Madison.
9. Geng, L., Hamilton, H.J., 2006. Interestingness Measures for Data Mining: A Survey. *ACM Computing Surveys*, vol. 38, no. 3, article 9. ACM Inc.
10. Greco, S., Pawlak, Z., Słowiński, R., 2004. Can Bayesian confirmation measures be useful for rough set decision rules? *Eng. Application of Artif. Intelligence* 17: 345-361.
11. Greco, S., Sowiski, R., Szczch, I., 2008. Assessing the quality of rules with a new monotonic interestingness measure Z, in: *Artificial Intelligence and Soft Computing (ICAISC 2008)*, LNAI, vol. 5097, pp. 556-565. Springer, Heidelberg.
12. Hempel, C.G., 1945. Studies in the logic of confirmation (I). *Mind* 54, 1-26.
13. McGarry, K., 2005. A survey of interestingness measures for knowledge discovery. *The Knowledge Engineering Review*, vol. 20:1, Cambridge University Press, pp.39-61..
14. Mortimer, H., 1988. *The Logic of Induction*, Paramus, Prentice Hall.
15. Nicod, J., 1923. *Le probleme de la logique de l'induction*. Alcan, Paris
16. Nozick, R., 1981. *Philosophical Explanations*, Clarendon Press Oxford (UK).
17. Rips, L.J., 2001. Two Kinds of Reasoning. *Psychological Science*, 12, 129-134.
18. Szczęch I., 2009. *Multicriteria Attractiveness Evaluation of Decision and Association Rules*, TRS X, vol. 5656/2009, pp. 197-274, Springer, Heidelberg.
19. Tan, P.-N., Kumar, V., Srivastava, J., 2002. Selecting the right interestingness measure for association patterns. In: *Proc. of the 8th international Conf. on Knowledge Discovery and Data Mining (KDD 2002)*. Edmonton, Canada, pp.32-41.