# The Property of $\chi^2_{01}$-Concordance for Bayesian Confirmation Measures

Robert Susmaga and Izabela Szczęch

Institute of Computing Science, Poznań Univesity of Technology,
Piotrowo 2, 60-965 Poznań, Poland

**Abstract.** The paper considers evaluation of rules with particular interestingness measures being Bayesian confirmation measures. It analyses the measures with regard to their agreement with a statistically significant dependency between the evidence and the hypothesis. As it turns out, many popular confirmation measures were not defined to possess such a form of agreement. As a result, even in situations when there is only a weak dependency in data, measures could indicate strong confirmation (or disconfirmation), encouraging the user to take some unjustified actions. The paper employs a $\chi^2$-based coefficient allowing to assess the level of dependency between the evidence and hypothesis in experimental data. A method of quantifying the level of agreement (concordance) between this coefficient and the measure being analysed is introduced. Experimental results for 12 popular confirmation measures are additionally visualised with scatter-plots and histograms.

**Keywords:** Interestingness measures, Bayesian confirmation, statistical dependency

## 1 Introduction

Regardless of the application domain, a crucial step in discovering knowledge from data is the evaluation of induced patterns [2, 10, 17, 23]. Evaluation of patterns in form of *if-then* rules is often done using quantitative measures of interest (e.g. rule support, confidence, gain, lift) [10, 23]. Among such interestingness measures, an important role is played by a group called *Bayesian confirmation measures*. Generally, they express the degree to which a rule's premise (also referred to as the conditional part or evidence) confirms its conclusion (also referred to as the decision part or hypothesis) [5, 9]. To narrow down the field of available confirmation measures, various properties of such measures are introduced and analysed. Popular properties of confirmation measures include monotonicity property, $Ex_1$ property and its generalization to weak $Ex_1$, logicality $L$ property and its generalization to weak $L$, and a group of symmetry properties (for a survey refer to [5, 7, 12, 13]).

Let us stress that the property analysis becomes much more complex when we assume that it is conducted upon data that may be error-prone. But in practice, the existence of possible data errors is a real phenomenon and must be taken into account, so that insignificant, accidental conclusions could be eliminated [14]. Unfortunately, at times the popular confirmation measures may indicate strong confirmation or strong disconfirmation, while there is only a weak dependency in data [22]. Such indications are potentially dangerous, since they may lead to unjustified, and thus inappropriate, user actions. To examine this aspect of the confirmation measures, the paper assesses the significance of the dependency between the evidence and the hypothesis in experimental data, and introduces a method of quantifying the level of agreement (referred to as concordance) between this assessment and the measure being analysed.

The rest of the paper is organized as follows. Section 2 describes the concept of Bayesian confirmation and defines popular measures. An overview of common measure properties is presented in Section 3. Section 4 discusses hazards of using the confirmation measures under observational errors, including methodology aimed at assessing the ($\chi^2$-based) level of dependency between the evidence and the hypothesis in data. Moreover, it introduces concordance between the $\chi^2$-based coefficient and confirmation measures. Last but not least, it provides experimental evaluations of the selected confirmation measures. Final remarks and conclusions are contained in Section 5.

## 2 Bayesian Confirmation Measures

In this paper, we consider evaluation of patterns represented in the form of rules. The starting point for such rule induction process (rule mining) is a sample of a larger reality, often represented in the form of a data table. Formally, a data table (dataset) is a pair $S = (U, A)$, where $U$ is a non-empty finite set of objects, called the *universe*, and $A$ is a non-empty finite set of *attributes* providing descriptions to the objects.

A rule induced from the dataset consists of a *premise* "*if E*" (referring to an existing piece of evidence, $E$) and a *conclusion* "*then H*" (referring to a hypothesised piece of evidence, $H$). Below, we shall use the common, shortened denotation $E \to H$ (read as "*if E, then H*").

To evaluate the patterns induced from datasets with respect to their relevance and utility, quantitative interestingness measures have been proposed and analysed [10]. This paper concentrates on a group of interestingness measures called Bayesian confirmation measures. They quantify the degree to which the evidence in the rule's premise $E$ provides support *for* or *against* the hypothesised piece of evidence in the rule's conclusion $H$ [9].

In the context of a particular dataset, the relation between $E$ and $H$ may be quantified by four non-negative frequencies $a$, $b$, $c$ and $d$, briefly represented in a $2 \times 2$ contingency table (Table 1). As an illustration, let us recall a popular folk statement that "*all ravens are black*", formalized as a rule "*if x is a raven, then x is black*", often used by Hempel [15]. Regarding that rule, the frequencies may be interpreted as follows: $a$ is the number of black ravens, $b$ is the number of black non-ravens, $c$ is the number of non-black ravens, and $d$ is the number of non-black non-ravens. Observe that $a$, $b$, $c$ and $d$ can thus be used to estimate probabilities: e.g. the probability of the premise is expressed as $P(E) = (a+c)/n$, the conditional probability of the conclusion given the premise is $P(H|E) = P(H \cap E)/P(E) = a/(a+c)$, and so on.

**Table 1.** An exemplary contingency table of the rule's premise and conclusion.

| | $H$ | $\neg H$ | $\Sigma$ |
|---|---|---|---|
| $E$ | $a$ | $c$ | $a+c$ |
| $\neg E$ | $b$ | $d$ | $b+d$ |
| $\Sigma$ | $a+b$ | $c+d$ | $n$ |

The group of confirmation measures that we shall present and analyse consists of interestingness measures that satisfy the property of Bayesian confirmation. Formally, for a rule $E \to H$, an interestingness measure $c(H,E)$ has the property of Bayesian confirmation when it satisfies the following conditions:

$$c(H,E) \begin{cases} > 0 \text{ when } P(H|E) > P(H) & \text{(confirmation)}\,, \\ = 0 \text{ when } P(H|E) = P(H) & \text{(neutrality)}\,, \\ < 0 \text{ when } P(H|E) < P(H) & \text{(disconfirmation)}\,. \end{cases} \tag{1}$$

Thus, the confirmation is interpreted as an increase in the probability of the conclusion $H$ provided by the premise $E$ (similarly for the neutrality and the disconfirmation).

Let us stress that the list of alternative, non-equivalent measures of Bayesian confirmation is quite large [5, 8]. The commonly used confirmation measures are presented in Table 2 (for brevity, some definitions are only formulated for two of the main defined situations: confirmation and disconfirmation; in the case of neutrality their values default to zero).

## 3 Properties of Bayesian Confirmation Measures

To discriminate between interestingness measures and help to choose a suitable one for a particular application, many properties have been proposed and compared in the literature [7, 10, 17, 11]. Properties group the measures according to similarities in their behaviour. Among commonly used properties of confirmation measures there are such properties as:

- *Property* $M$, ensuring monotonic dependency of a measure on the number of objects satisfying (supporting) or not the premise and/or the conclusion of the rule [12, 23], so that the measure is non-decreasing with respect to $a$ and $d$, and non-increasing with respect to $b$ and $c$. Thus, e.g. arrival of new objects supporting the rule (or counterexamples, respectively) to the dataset cannot lower (increase) the value of the measure.
- *Property* $Ex_1$, and its generalization *weak* $Ex_1$, assuring that any conclusively confirmatory rule is assigned a higher value than any rule which is not

**Table 2.** Popular confirmation measures.

| | |
|---|---|
| $D(H,E) = P(H\|E) - P(H) = \dfrac{a}{a+c} - \dfrac{a+b}{n}$ | [6] |
| $M(H,E) = P(E\|H) - P(E) = \dfrac{a}{a+b} - \dfrac{a+c}{n}$ | [18] |
| $S(H,E) = P(H\|E) - P(H\|\neg E) = \dfrac{a}{a+c} - \dfrac{b}{b+d}$ | [4] |
| $N(H,E) = P(E\|H) - P(E\|\neg H) = \dfrac{a}{a+b} - \dfrac{c}{c+d}$ | [19] |
| $C(H,E) = P(E \wedge H) - P(E)P(H) = \dfrac{a}{n} - \dfrac{(a+c)(a+b)}{n^2}$ | [3] |
| $F(H,E) = \dfrac{P(E\|H) - P(E\|\neg H)}{P(E\|H) + P(E\|\neg H)} = \dfrac{ad-bc}{ad+bc+2ac}$ | [16] |
| $Z(H,E) = \begin{cases} 1 - \dfrac{P(\neg H\|E)}{P(\neg H)} = \dfrac{ad-bc}{(a+c)(c+d)} & \text{in case of confirmation} \\ \dfrac{P(H\|E)}{P(H)} - 1 = \dfrac{ad-bc}{(a+c)(a+b)} & \text{in case of disconfirmation} \end{cases}$ | [5] |
| $A(H,E) = \begin{cases} \dfrac{P(E\|H) - P(E)}{1-P(E)} = \dfrac{ad-bc}{(a+b)(b+d)} & \text{in case of confirmation} \\ \dfrac{P(H) - P(H\|\neg E)}{1-P(H)} = \dfrac{ad-bc}{(b+d)(c+d)} & \text{in case of disconfirmation} \end{cases}$ | [13] |

conclusively confirmatory, and any conclusively disconfirmatory rule is assigned a lower value than any rule which is not conclusively disconfirmatory [5, 13].

– *Logicality* $L$, and its generalization *weak* $L$, indicating conditions under which measures should obtain their maximal/minimal values [5, 9, 13]. Another property closely related to $L$, $Ex_1$ and their generalizations is *maximality/minimality* proposed in [11].

Searching for measures that possess property $Ex_1$, Crupi et al. [5] have proposed measure $Z(H,E)$. Later, as its likelihoodist counterpart, measure $A(H,E)$ has been proposed in [13] (for definitions see Table 2). It has been proved in [13] that neither measure $Z(H,E)$ nor $A(H,E)$ satisfies weak $Ex_1$, however new measures enjoying weak $Ex_1$ can be derived from $Z(H,E)$ and $A(H,E)$. They are denoted as $c_1(H,E)$, $c_2(H,E)$, $c_3(H,E)$, and $c_4(H,E)$ (for definitions see Table 3; brevity comments similar to that of Table 2 apply here). Measures $c_1(H,E)$ and $c_2(H,E)$ are defined using parameters $\alpha$ and $\beta$, where $\alpha+\beta = 1$ and $\alpha > 0$, $\beta > 0$. Observe that parameters $\alpha$ and $\beta$ can be used to closen the new measure to $Z(H,E)$ or $A(H,E)$, i.e. to Bayesian or likelihoodist inspirations.

**Table 3.** Derived confirmation measures.

| | |
|---|---|
| $c_1(H,E) =$ | $\begin{cases} \alpha + \beta A(H,E) & \text{in case of confirmation when } c = 0 \\ \alpha Z(H,E) & \text{in case of confirmation when } c > 0 \\ \alpha Z(H,E) & \text{in case of disconfirmation when } a > 0 \\ -\alpha + \beta A(H,E) & \text{in case of disconfirmation when } a = 0 \end{cases}$ |
| $c_2(H,E) =$ | $\begin{cases} \alpha + \beta Z(H,E) & \text{in case of confirmation when } b = 0 \\ \alpha A(H,E) & \text{in case of confirmation when } b > 0 \\ \alpha A(H,E) & \text{in case of disconfirmation when } d > 0 \\ -\alpha + \beta Z(H,E) & \text{in case of disconfirmation when } d = 0 \end{cases}$ |
| $c_3(H,E) =$ | $\begin{cases} A(H,E)Z(H,E) & \text{in case of confirmation} \\ -A(H,E)Z(H,E) & \text{in case of disconfirmation} \end{cases}$ |
| $c_4(H,E) =$ | $\begin{cases} min(A(H,E), Z(H,E)) & \text{in case of confirmation} \\ max(A(H,E), Z(H,E)) & \text{in case of disconfirmation} \end{cases}$ |

# 4 Using Bayesian Confirmation Measures in Error-Prone Situations

## 4.1 The Property of Concordance

In real-life situations the existence of possible errors must be taken into account. Thus, we should look for a statistically significant dependency between the evidence and the hypothesis, which may be quantified and measured with different tools. A good and popular one is the two-dimensional $\chi^2$ test, often used to test for the independence of two discrete-valued variables. The popular alternatives to this test include the Cramer's V coefficient, the Yule's Q coefficient or the Fisher coefficient [20].

For $2 \times 2$-sized contingency tables, of the form $\left[\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right]$, as used in defining confirmation measures, a coefficient $\chi^2_0 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$ is defined. This coefficient is approximately $\chi^2$-distributed and ranges from 0 to $n$. To make it $n$-independent, it is scaled down (divided) by $n$, producing a value belonging to the interval $[0,1]$. This version of the coefficient will be further referred to as the "scaled-down $\chi^2_0$" and denoted as $\chi^2_{01}$.

In practice, two potentially unfavourable situations can concern the confirmation measure applied to a contingency table created from error-prone data:

– the value of $c(H,E)$ indicates either weak confirmation or weak disconfirmation, while there is a strong dependency between the evidence and the hypothesis,
– the value of $c(H,E)$ indicates either strong confirmation or strong disconfirmation, while there is only a weak dependency between the evidence and the hypothesis.

To counteract those, there arises a need to evaluate the concordance between confirmation measures and statistical significance of the evidence-hypothesis dependency. For such an evaluation to be useful, it should provide continuous measurements, the higher the more the measure $c(H, E)$ 'agrees' with the level of dependency between the evidence and the hypothesis. This evaluation may be performed using different statistical tools, and in this study we use linear Pearson correlation between $|c(H, E)|$ (the absolute value of $c(H, E)$) and $\chi^2_{01}$, denoted as $r(|c(H, E)|, \chi^2_{01})$. Taking $|c(H, E)|$ into account (thus ignoring the sign of $c(H, E)$) is essential, as it is the absolute value of the confirmation measure, and not its sign, that determines the 'strength' of $c(H, E)$ (i.e. the degree to which the premise of a rule evaluated by the measure confirms or disconfirms its conclusion). Potential alternatives to the linear Pearson correlation include the Spearman rank correlation coefficient [21] or mutual information measures [1].

What is specific about the property of concordance is that it is a representative of continuous-type properties: it can be quantified as the agreement with the level of dependency between $E$ and $H$.

The relation between $\chi^2_{01}$ coefficient and a given confirmation measure $c(H, E)$ may be additionally visualized, which is easily done with a scatter-plot of $c(H, E)$ against $\chi^2_{01}$. Each such scatter-plot will fit a $2 \times 1$-sized rectangular envelope, with its axes ranging from $-1$ to $+1$ (horizontal, $c(H, E)$) and from 0 to 1 (vertical, $\chi^2_{01}$), as illustrated in Figure 1, with lighter and darker regions and graded transitions between them. Given a measure $c(H, E)$, the points of the $c(H, E)$-versus-$\chi^2_{01}$ scatter-plot should possibly occupy the darker regions of the figure, while possibly avoiding any of the lighter ones.

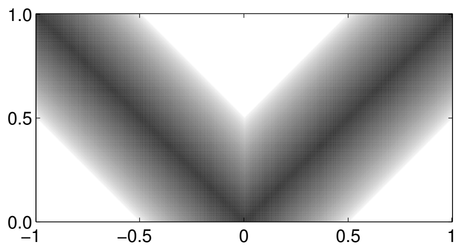


**Fig. 1.** The desirable (darker) and undesirable (lighter) regions of the $c(H, E)$-versus-$\chi^2_{01}$ scatter-plot of $c(H, E)$.

### 4.2 The Experimental Set-up

Given $n > 0$ (the total number of observations), the dataset is generated as the set of all possible $\left[\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right]$ contingency tables satisfying $a + b + c + d = n$. The set is thus exhaustive and non-redundant (i.e. it contains exactly one copy of each contingency table satisfying the above condition).

The exact number $t$ of tables in the set is $t = (n+1)(n+2)(n+3)/6$. This value grows quickly, although polynomially, not exponentially; e.g. the number of all tables for $n = 128$ equals $t = 366145$. Unfortunately, the number $t$ can become considerable: for $n$ about 1000 (a typical number of objects in a benchmark classification data set) $t$ exceeds hundreds of millions.

After having set the total number of observations $n$ to 128, the following operations were performed:

- the exhaustive and non-redundant set of $\left[\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right]$ contingency tables satisfying $a + b + c + d = n$ was generated,
- the values of the 12 selected confirmation measures (with $c_1(H,E)$ and $c_2(H,E)$ defined for $\alpha = \beta = 0.5$) for all the generated tables were calculated,
- the values of the $\chi^2_{01}$ coefficient for all the generated tables were computed,
- the correlations between the absolute values of each of 12 selected confirmation measures and the $\chi^2_{01}$ coefficient (i.e. concordances) were established.

Similar steps (but with $n$ decreased to 32 to facilitate the rendering process) led to the charts, i.e. scatter-plots of $c(H,E)$ against $\chi^2_{01}$ (Figure 2) and so called triple-region histograms of $c(H,E)$ (Figure 3). The triple-region histograms show the distribution of the measure, with each bar additionally displaying the number of points situated above (upper white region), on (dark region) or below (lower white region) the $|c(H,E)| = \chi^2_{01}$ line. Characteristically, the size of the lower region always exceeds considerably the size of the upper region, while the dark region is only a thin, horizontal strip (with the notable exception of $c_3(H,E)$, for which only the dark region exists).

**Table 4.** The coefficients of the $\chi^2_{01}$-concordance of the 12 selected confirmation measures.

| $c(H,E)$ | $r(\lvert c(H,E)\rvert, \chi^2_{01})$ | $c(H,E)$ | $r(\lvert c(H,E)\rvert, \chi^2_{01})$ |
|---|---|---|---|
| $D(H,E)$ | 0.713 | $Z(H,E)$ | 0.694 |
| $M(H,E)$ | 0.713 | $A(H,E)$ | 0.694 |
| $S(H,E)$ | 0.912 | $c1(H,E)$ | 0.697 |
| $N(H,E)$ | 0.912 | $c2(H,E)$ | 0.697 |
| $C(H,E)$ | 0.908 | $c3(H,E)$ | 1.000 |
| $F(H,E)$ | 0.711 | $c4(H,E)$ | 0.957 |

### 4.3 The Experimental Results

The conducted experiments revealed interesting results of both generic and specific nature [22]. The following remarks concern the $\chi^2_{01}$-concordance (as quantified by the Pearson correlation coefficient $r$) of the measures (see Table 4 and Figures 2 and 3):

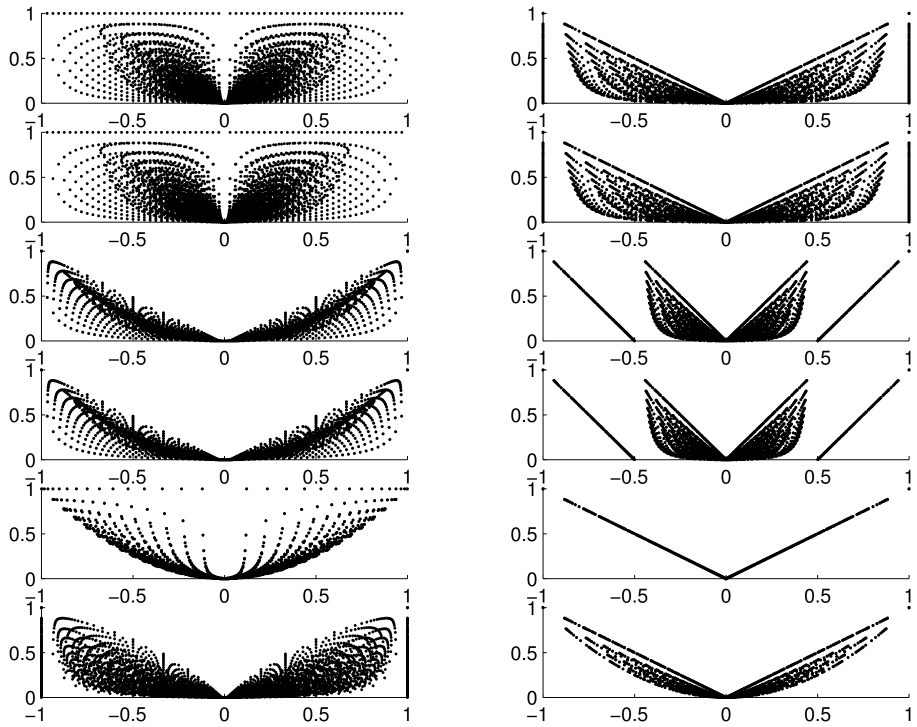

**Fig. 2.** Scatter-plots of the 12 selected confirmation measures against $\chi^2_{01}$ (left-hand column: measures $D(H,E)$, $M(H,E)$, $S(H,E)$, $N(H,E)$, $C(H,E)$, $F(H,E)$; right-hand column: measures $Z(H,E)$, $A(H,E)$, $c_1(H,E)$, $c_2(H,E)$, $c_3(H,E)$, $c_4(H,E)$; $c_1(H,E)$ and $c_2(H,E)$ defined with $\alpha = \beta = 0.5$).

- measure $c_3(H,E)$ enjoys an ideal $\chi^2_{01}$-concordance, which is due to the fact that $|c_3(H,E)| = \chi^2_{01}$,
- the concordance of the other measures ranges from 0.957 ($c_4(H,E)$) down to 0.694 ($Z(H,E)$ and $A(H,E)$), in result of which all of them can be referred to as approximately concordant,
- the absolute values of the approximately concordant measures tend to exceed those of $\chi^2_{01}$.

A conclusion is that not all of the measures possess ideal concordance. The less concordant measures should thus be used with some care, especially when applied to real-life, error-prone data, as the may express either strong confirmation or strong disconfirmation in statistically insignificant situations.

It is especially interesting that measures $c_1(H,E)$ and $c_2(H,E)$, which depend on the value of the $\alpha$ parameter, i.e. the free parameter that is used to define these measures (the $\beta$ parameter is, on the other hand, constrained, as $\beta = 1 - \alpha$), evince varying shapes of their corresponding scatter-plots, see Figure 4. This will necessarily influence their correlations with the $\chi^2_{01}$ coefficient.

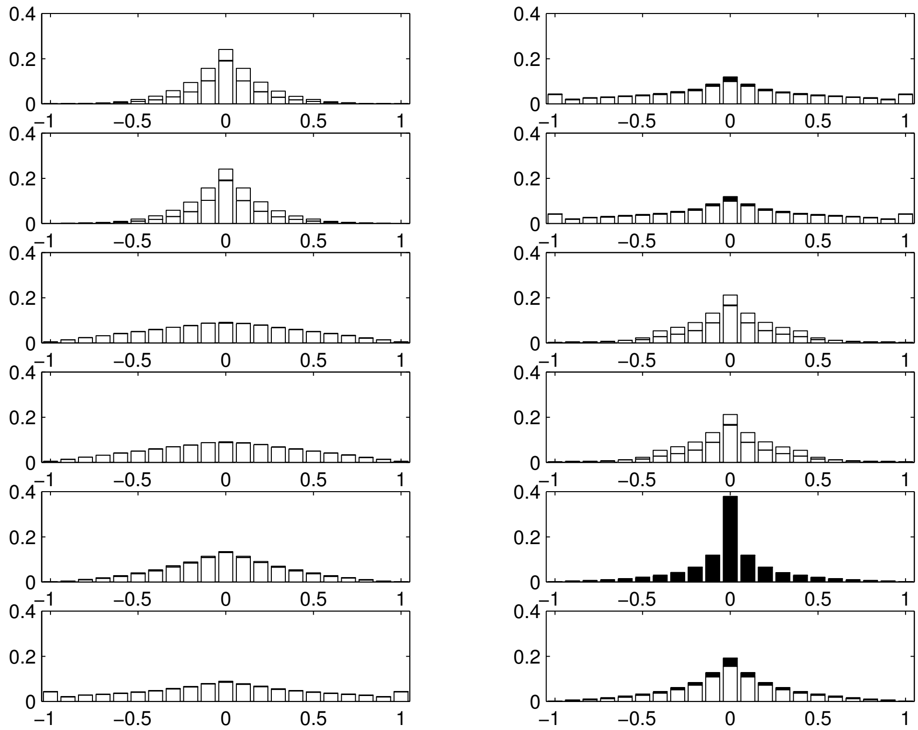

**Fig. 3.** Triple-region histograms of the 12 selected confirmation measures $c(H, E)$ in relation to $\chi^2_{01}$ (left-hand column: measures $D(H, E)$, $M(H, E)$, $S(H, E)$, $N(H, E)$, $C(H, E)$, $F(H, E)$; right-hand column: measures $Z(H, E)$, $A(H, E)$, $c_1(H, E)$, $c_2(H, E)$, $c_3(H, E)$, $c_4(H, E)$; $c_1(H, E)$ and $c_2(H, E)$ defined with $\alpha = \beta = 0.5$).

Because, by definition, most values of these measures belong to the interval $(-\alpha, +\alpha)$, see Figure 4 (more details can be found in [22]), their concordances are then also changed accordingly, see Figure 5. This means that the $\alpha$ parameter can be directly used to control this aspect of these two measures. In particular, when $\alpha \to 1.0$, measures $c_1(H, E)$ and $c_2(H, E)$ approach measures $Z(H, E)$ and $A(H, E)$, respectively, in which case they also acquire their corresponding concordances (which is, in both cases, 0.694).

For more detailed analyses of these (and other) properties of the confirmation measures see [22].

## 5 Conclusions

The paper considers Bayesian confirmation measures, which have become the subject of numerous, intensive studies. What is characteristic of these studies is that virtually all of them were confined to environments that had been explicitly

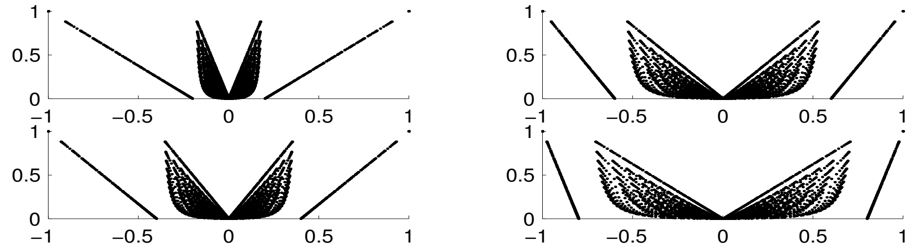


**Fig. 4.** Scatter-plots of measures $c_1(H, E)$ and $c_2(H, E)$ against $\chi^2_{01}$, defined for various values of $\alpha$ (left-hand column: $\alpha = 0.2$, $\alpha = 0.4$; right-hand column: $\alpha = 0.6$, $\alpha = 0.8$), see Figure 2 for $\alpha = 0.5$.

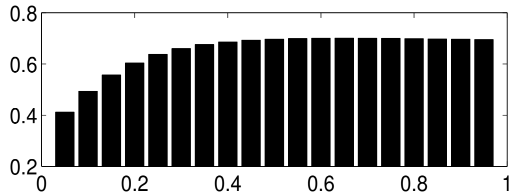


**Fig. 5.** The concordances of $c_1(H, E)$ and $c_2(H, E)$, as influenced by the changing $\alpha$.

or implicitly assumed to be free from observational errors. In real-life situations, however, the existence of such errors must be taken into account and properly approached. This goal is in this paper accomplished with the $\chi^2$ test, commonly used to examine for the dependence between two discrete-valued variables.

The actual amount of how concordant a confirmation measure is with the level of dependency between the evidence and the hypothesis is quantified with the Pearson correlation coefficient between the measure and an introduced $\chi^2_{01}$ coefficient. The relations between the measures and $\chi^2_{01}$ are additionally illustrated by scatter-plots and specialized, triple-region histograms.

The general conclusion is that most measures possess rather high, although not ideal, concordance. The scatter-plots and the triple-region histograms of these measures reveal particular situations in which they express either strong confirmation or strong disconfirmation in statistically insignificant situations. This means that they should be used with special care in error-prone environments. Interestingly enough, the concordance of the parametrized confirmation measures, $c_1(H, E)$ and $c_2(H, E)$, is influenced by the parameters used in their definitions, so it may be controlled to some extent. Measure $c_3(H, E)$, a notable exception amongst the 12 selected confirmation measures, enjoys full concordance, so its indications may assumed to be safest in this particular respect.

**Acknowledgment.** The work has been supported by the Polish Ministry of Science and Higher Education.

## References


1. Bell, C.: Mutual information and maximal correlation as measure dependence. The Annals of Mathematical Statistics 33, 587–595 (1962)
2. Brzeziński, D., Stefanowski, J.: Accuracy updated ensemble for data streams with concept drift. In: Proceedings of the 6th International Conference on Hybrid Artificial Intelligent Systems. pp. 155–163 (2011)
3. Carnap, R.: Logical Foundations of Probability, 2nd ed. University of Chicago Press (1962)
4. Christensen, D.: Measuring confirmation. Journal of Philosophy 96, 437–461 (1999)
5. Crupi, V., Tentori, K., Gonzalez, M.: On bayesian measures of evidential support: Theoretical and empirical issues. Philosophy of Science 74, 229–252 (2007)
6. Eells, E.: Rational Decision and Causality. Cambridge University Press, Cambridge (1982)
7. Eells, E., Fitelson, B.: Symmetries and asymmetries in evidential support. Philosophical Studies 107(2), 129–142 (2002)
8. Fitelson, B.: The plurality of bayesian measures of confirmation and the problem of measure sensitivity. Philosophy of Science 66, 362–378 (1999)
9. Fitelson, B.: Studies in Bayesian Confirmation Theory. Ph.D. thesis, University of Wisconsin, Madison (2001)
10. Geng, L., Hamilton, H.: Interestingness measures for data mining: A survey. ACM Computing Surveys 38(3) (2006)
11. Glass, D.H.: Confirmation measures of association rule interestingness. Knowlegde Based Systems 44, 65–77 (2013)
12. Greco, S., Pawlak, Z., Słowiński, R.: Can bayesian confirmation measures be useful for rough set decision rules? Engineering Applications of Artifficial Intelligence 17, 345–361 (2004)
13. Greco, S., Słowiński, R., Szczęch, I.: Properties of rule interestingness measures and alternative approaches to normalization of measures. Information Sciences 216, 1–16 (2012)
14. Hastie, T., Tibshirani, R., Friedman, J.: Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer-Verlag (2003)
15. Hempel, C.: Studies in the logic of confirmation (i). Mind 54, 1–26 (1945)
16. Kemeny, J., Oppenheim, P.: Degrees of factual support. Philosophy of Science 19, 307–324 (1952)
17. Lenca, P., Meyer, P., Vaillant, B., Lallich, S.: On selecting interestingness measures for association rules: User oriented description and multiple criteria decision aid. European Journal of Operational Research 184(2), 610–626 (2008)
18. Mortimer, H.: The Logic of Induction. Paramus, Prentice Hall (1988)
19. Nozick, R.: Philosophical Explanations. Clarendon Press, Oxford, UK (1981)
20. Rayner, J., Best, D.: A Contingency Table Approach to Nonparametric Testing. Taylor & Francis Group (2001)
21. Spearman, C.E.: The proof and measurement of association between two things. American Journal of Psychology 15, 72–101 (1904)
22. Susmaga, R., Szczęch, I.: Statistical significance of bayesian confirmation measures. Tech. rep., RA-010/12, Poznań University of Technology (2012)
23. Szczęch, I.: Multicriteria attractiveness evaluation of decision and association rules. Transactions on Rough Sets X, LNCS series 5656, 197–274 (2009)