Interestingness/Evaluation Measures of Association Rules

Marzena Kryszkiewicz
Institute of Computer Science
Warsaw University of Technology

1

Relative support of an Association Rule

- The *relative support* of an association rule $X \rightarrow Y$ is defined as the probability of the co-occurrence of X and Y:

$$rSup(X \rightarrow Y) = P(XY).$$

- **Property.** The relative support depends on the probability of the base of a rule, but depends neither on the probability of its antecedent nor the probability of its consequent.

2

Confidence of an Association Rule

- The *confidence* of an association rule $X \rightarrow Y$ is defined as the conditional probability that Y occurs provided X occurs:

$$conf(X \rightarrow Y) = P(XY) / P(X).$$

- **Property.** The confidence depends on the probability of the antecedent of a rule and the probability of its base, but does not depend on the probability of its consequent.

3

(In)dependence of Events

- X and Y called *independent* if:

$$P(XY) = P(X) * P(Y).$$

- Otherwise, they are called *dependent*.
- **Natural interpretation:**
  - If $P(XY) > P(X) * P(Y)$, then X and Y are dependent positively.
  - If $P(XY) = P(X) * P(Y)$, then X and Y are independent.
  - If $P(XY) < P(X) * P(Y)$, then X and Y are dependent negatively.
- **Property.** $P(XY) \neq P(X) * P(Y)$ does not determine the degree to which X and Y are dependent.

4

Confidence and event (in)dependence…

TId	X	Y
1		x
2		x
3		x
4		x
5		x
6		x
7		x
8		x
9	x	x
10	x	

- $P(X) = \frac{2}{10}, P(Y) = \frac{9}{10} \Rightarrow P(X) \times P(Y) = 0.18.$
- $P(XY) = \frac{1}{10} = 0.1 < 0.18 = P(X) \times P(Y)$, czyli **$X$ and Y are dependent negatively.**
- ***$conf(X \rightarrow Y) = P(XY) / P(X) = \frac{1}{2}$.***

- $P(X) = \frac{2}{10}, P(\bar{Y}) = \frac{1}{10} \Rightarrow P(X) \times P(\bar{Y}) = 0.02.$
- $P(X\bar{Y}) = \frac{1}{10} = 0.1 > 0.02 = P(X) \times P(\bar{Y})$, czyli **$X$ and $\bar{Y}$ are dependent positively.**
- ***$conf(X \rightarrow \bar{Y}) = P(X\bar{Y}) / P(X) = \frac{1}{2}$.***

5

Confidence and event (in)dependence

TId	X	Y
1		x
2		x
3		x
4		x
5		x
6		x
7		x
8	x	x
9	x	x
10	x	

- $P(X) = \frac{3}{10}, P(Y) = \frac{9}{10} \Rightarrow P(X) \times P(Y) = 0.27.$
- $P(XY) = \frac{2}{10} = 0.2 < 0.27 = P(X) \times P(Y)$, czyli **$X$ and Y are dependent negatively.**
- ***$conf(X \rightarrow Y) = P(XY) / P(X) = \frac{2}{3}$.***

- $P(X) = \frac{3}{10}, P(\bar{Y}) = \frac{1}{10} \Rightarrow P(X) \times P(\bar{Y}) = 0.03.$
- $P(X\bar{Y}) = \frac{1}{10} = 0.1 > 0.03 = P(X) \times P(\bar{Y})$, czyli **$X$ and $\bar{Y}$ are dependent positively.**
- ***$conf(X \rightarrow \bar{Y}) = P(X\bar{Y}) / P(X) = \frac{1}{3}$.***

6

Lift of an Association Rule

- The *lift* of an association rule $X \rightarrow Y$ is defined as the ratio of the conditional probability of the occurrence of Y given X occurs to the probability of the occurrence of Y:

$$lift(X \rightarrow Y) = \frac{conf(X \rightarrow Y)}{P(Y)} = \frac{P(XY)}{P(X) \times P(Y)}.$$

- **Property.**

Y and X are dependent pos.	$P(XY) > P(X) \times P(Y)$	$lift(X \rightarrow Y) > 1$	$lift(Y \rightarrow X) > 1$
Y and X are independent	$P(XY) = P(X) \times P(Y)$	$lift(X \rightarrow Y) = 1$	$lift(Y \rightarrow X) = 1$
Y and X are dependent neg.	$P(XY) < P(X) \times P(Y)$	$lift(X \rightarrow Y) < 1$	$lift(Y \rightarrow X) < 1$

- **Property.** $lift(X \rightarrow Y) = lift(Y \rightarrow X)$.

7

Lift and event (in)dependence

TId	X	Y	Z	V
1	x	x		
2	x	x		
3	x	x		
4	x	x		
5	x	x		
6	x	x		
7	x	x		
8	x	x		
9	x	x		
10			x	x

- $P(X) = \frac{9}{10}$, $P(Y) = \frac{9}{10} \Rightarrow P(X) \times P(Y) = 0.81$.
- $P(XY) = \frac{9}{10} = 0.9 > 0.81 = P(X) \times P(Y)$, czyli **$X$ and Y are dependent positively.**
- $lift(X \rightarrow Y) = \frac{P(XY)}{P(X) \times P(Y)} = \frac{0.9}{0.9 \times 0.9} = \frac{10}{9} = 1\frac{1}{9}$.

- $P(Z) = \frac{1}{10}$, $P(V) = \frac{1}{10} \Rightarrow P(Z) \times P(V) = 0.01$.
- $P(ZV) = \frac{1}{10} = 0.1 > 0.01 = P(Z) \times P(V)$, czyli **$Z$ and V are dependent positively.**
- $lift(Z \rightarrow V) = \frac{P(ZV)}{P(Z) \times P(V)} = \frac{0.1}{0.1 \times 0.1} = \frac{10}{1} = 10$.

8

Certainty Factor of an Association Rule…

- The *certainty factor* of an association rule $X \rightarrow Y$ is defined as the degree to which the probability of the occurrence of Y can change when X occurs:

$$cf(X \rightarrow Y) = \begin{cases} \frac{conf(X \rightarrow Y) - P(Y)}{1 - P(Y)} & \text{if } conf(X \rightarrow Y) > P(Y), \\ 0 & \text{if } conf(X \rightarrow Y) = P(Y), \\ -\frac{P(Y) - conf(X \rightarrow Y)}{P(Y) - 0} & \text{if } conf(X \rightarrow Y) < P(Y). \end{cases}$$

9

Certainty Factor

$$cf(X \rightarrow Y) = \begin{cases} \frac{conf(X \rightarrow Y) - P(Y)}{1 - P(Y)} & \text{if } conf(X \rightarrow Y) > P(Y), \\ 0 & \text{if } conf(X \rightarrow Y) = P(Y), \\ -\frac{P(Y) - conf(X \rightarrow Y)}{P(Y) - 0} & \text{if } conf(X \rightarrow Y) < P(Y). \end{cases}$$

- **Property.**

asymmetric

$$cf(X \rightarrow Y) = \begin{cases} \frac{P(XY) - P(X) \times P(Y)}{P(X) - P(X) \times P(Y)} & \text{if } P(XY) > P(X) \times P(Y), \\ 0 & \text{if } P(XY) = P(X) \times P(Y), \\ -\frac{P(X) \times P(Y) - P(XY)}{P(X) \times P(Y) - 0} & \text{if } P(XY) < P(X) \times P(Y). \end{cases}$$

symmetric

10

Extreme Values of Dependence Measures

measure	max	min
$P(XY)$	1	0
$conf(X \rightarrow Y)$	1	0
$lift(X \rightarrow Y)$	∞	0
$cf(X \rightarrow Y)$	1 if X and Y are dependent positively	-1 if X and Y are dependent negatively

11

Extreme Values of $P(XY)$ under Given $P(X)$ & $P(Y)$

- **Proposition.**

a) $max\_P(XY|_{P(X),\, P(Y)}) = \min\{P(X), P(Y)\}$

b) $min\_P(XY|_{P(X),\, P(Y)}) = \begin{cases} 0 & \text{if } P(X) + P(Y) \leq 1 \\ P(X) + P(Y) - 1 & \text{if } P(X) + P(Y) > 1 \end{cases}$

$= \max\{0,\ P(X) + P(Y) - 1\}$

X	Y
x	x
x	x
x	

a)

X	Y
x	
x	
x	
	x
	x

b)

X	Y
x	
x	
x	x
x	x
x	x
	x

c)

12

Extreme Values of Measures under $P(X)$ & $P(Y)$

- **Proposition.**

instead of 1 instead of 0

measure	max under the assumption that $P(X)$ and $P(Y)$ are fixed	min under the assumption that $P(X)$ and $P(Y)$ are fixed
$P(XY)$	$\min\{P(X), P(Y)\}$	$\max\{0,\ P(X)+P(Y)-1\}$
$conf(X\to Y)$	$\frac{\min\{P(X), P(Y)\}}{P(X)}$	$\frac{\max\{0, P(X)+P(Y)-1\}}{P(X)}$
$lift(X\to Y)$	$\frac{\min\{P(X), P(Y)\}}{P(X)\times P(Y)}$	$\frac{\max\{0, P(X)+P(Y)-1\}}{P(X)\times P(Y)}$
$cf(X\to Y)$	$\frac{\min\{P(X), P(Y)\}-P(X)\times P(Y)}{P(X)-P(X)\times P(Y)}$ if X and Y are dependent positively	$-\frac{P(X)\times P(Y)-\max\{0, P(X)+P(Y)-1\}}{P(X)\times P(Y)-0}$ if X and Y are dependent negatively

Dependence Factor of an Association Rule

- The *dependence factor* of an association rule $X \to Y$ is defined as the degree to which the probability of the occurrence of Y can change when X occurs for given values of $P(X)$ and $P(Y)$:

instead of 1

$$df(X\to Y)=\begin{cases}\frac{conf(X\to Y)-P(Y)}{max\_conf(X\to Y\,|_{P(X),P(Y)})-P(Y)} & \text{if } conf(X\to Y)>P(Y),\\ 0 & \text{if } conf(X\to Y)=P(Y),\\ -\frac{P(Y)-conf(X\to Y)}{P(Y)-min\_conf(X\to Y\,|_{P(X),P(Y)})} & \text{if } conf(X\to Y)<P(Y).\end{cases}$$

instead of 0

Dependence Factor in Terms of Unconditional Probabilities…

instead of 1

$$df(X\to Y)=\begin{cases}\frac{P(XY)-P(X)\times P(Y)}{max\_P(XY\,|_{P(X),P(Y)})-P(X)\times P(Y)} & \text{if } P(XY)>P(X)\times P(Y),\\ 0 & \text{if } P(XY)=P(X)\times P(Y),\\ -\frac{P(X)\times P(Y)-P(XY)}{P(X)\times P(Y)-min\_P(XY\,|_{P(X),P(Y)})} & \text{if } P(XY)<P(X)\times P(Y).\end{cases}$$

instead of 0

Dependence Factor in Terms of Unconditional Probabilities

- **Property.** $df(X\to Y) = df(Y\to X)$.

instead of 1

$$df(X\to Y)=\begin{cases}\frac{P(XY)-P(X)\times P(Y)}{\min\{P(X),P(Y)\}-P(X)\times P(Y)} & \text{if } P(XY)>P(X)\times P(Y),\\ 0 & \text{if } P(XY)=P(X)\times P(Y),\\ -\frac{P(X)\times P(Y)-P(XY)}{P(X)\times P(Y)-\max\{0,P(X)+P(Y)-1\}} & \text{if } P(XY)<P(X)\times P(Y).\end{cases}$$

instead of 0

Dependence Factor versus Certainty Factor

- **Theorem.**
  - $df(X\to Y) = cf(X\to Y) = 0$ if $P(XY) = P(X)\times P(Y)$.
  - $|df(X\to Y)| \geq |cf(X\to Y)|$ if $P(XY) \neq P(X)\times P(Y)$.
  - If $P(XY) > P(X)\times P(Y)$, then:

    $df(X\to Y) = \max\{cf(X\to Y), cf(Y\to X)\}$.
  - If $P(XY) < P(X)\times P(Y)$, then:

    $df(X\to Y) = cf(X\to Y)$ if $P(X) + P(Y) \leq 1$.

    $|df(X\to Y)| > |cf(X\to Y)|$ if $P(X) + P(Y) > 1$.

Dependence Factor: $P(X) + P(Y) < 1$

- **Table.** Comparison of values of $lift(X\to Y)$, $cf(X\to Y)$ and $df(X\to Y)$ when $P(X) + P(Y) < 1$

$P(X)$	$P(Y)$	$P(XY)$	$P(X)\times P(Y)$	$lift(X\to Y)$	$cf(X\to Y)$	$cf(Y\to X)$	$df(X\to Y)=df(Y\to X)$
0.60	**0.30**	**0.30**	**0.18**	**1.67**	**0.29**	**1.00**	**1.00**
0.60	0.30	0.25	0.18	1.39	0.17	0.58	0.58
0.60	0.30	0.20	0.18	1.11	0.05	0.17	0.17
0.60	0.30	0.18	0.18	1.00	0.00	0.00	0.00
0.60	0.30	0.15	0.18	0.83	-0.17	-0.17	-0.17
0.60	0.30	0.10	0.18	0.56	-0.44	-0.44	-0.44
0.60	0.30	0.00	0.18	0.00	-1.00	-1.00	-1.00

Dependence Factor: $P(X) + P(Y) > 1$

- **Table.** Comparison of values of *lift*(X→Y), *cf*(X→Y) and *df*(X→Y) when $P(X) + P(Y) > 1$

P(X)	P(Y)	P(XY)	P(X)×P(Y)	lift(X→Y)	cf(X→Y)	cf(Y→X)	df(X→Y)=df(Y→X)
0.80	0.60	0.60	0.48	1.25	0.38	1.00	1.00
0.80	0.60	0.55	0.48	1.15	0.22	0.58	0.58
0.80	0.60	0.50	0.48	1.04	0.06	0.17	0.17
0.80	0.60	0.48	0.48	1.00	0.00	0.00	0.00
0.80	0.60	0.45	0.48	0.94	-0.06	-0.06	-0.37
0.80	**0.60**	**0.40**	**0.48**	**0.83**	**-0.17**	**-0.17**	**-1.00**

19

Dependence Factor in Terms of Lift

instead of ∝

$$df(X \to Y) = \begin{cases} \dfrac{lift(X \to Y) - 1}{max\_lift(X \to Y \mid_{P(X),P(Y)}) - 1} & \text{if } lift(X \to Y) > 1, \\ 0 & \text{if } lift(X \to Y) = 1, \\ -\dfrac{1 - lift(X \to Y)}{1 - min\_lift(X \to Y \mid_{P(X),P(Y)})} & \text{if } lift(X \to Y) < 1. \end{cases}$$

instead of 0

20

Conclusions

- *df*(X→Y) always reaches 1 when the dependence between *X* and *Y* is strongest possible positive and −1 when the dependence between *X* and *Y* is strongest possible negative for any given values of *P*(*X*) and *P*(*Y*).
- Unlike the dependence factor, the *certainty factor* and *lift* are misleading in expressing the strength of the dependence as each of them may take a value close to the value characteristic for independence even when two events are maximally strongly positively or negatively dependent.

21

Further results

$$df(X \to Y) = df(\overline{X} \to \overline{Y}) = -df(X \to \overline{Y}) = -df(\overline{X} \to Y).$$

22

References

- Sergey Brin, Rajeev Motwani, Craig Silverstein: Beyond Market Baskets: Generalizing Association Rules to Correlations. SIGMOD Conference 1997: 265-276
- Marzena Kryszkiewicz: Dependence Factor for Association Rules. ACIIDS (2) 2015: 135-145
- Marzena Kryszkiewicz: Dependence Factor as a Rule Evaluation Measure. Challenges in Computational Statistics and Data Mining 2016: 205-223
- Marzena Kryszkiewicz: Virtual Balancing of Decision Classes. ACIIDS (1) 2017: 673-684
- Shortliffe, E. and Buchanan, B.: A model of inexact reasoning in medicine. Mathematical Biosciences, vol. 23, 351–379 (1975)

23