Aufgabe 4

1

$$c((X \setminus \hat{X}) \Rightarrow X) = \frac{n_{(X\setminus\hat{X}) \cup \hat{X}}(T)}{n_{(X\setminus\hat{X})}(T)} = \frac{n_X(T)}{n_{(X\setminus\hat{X})}(T)}$$

$$\Updownarrow$$

✓

$$s(X) = \frac{n_X(T)}{n(T)}$$

$$s(X\setminus\hat{X}) = \frac{n_{X\setminus\hat{X}}(T)}{n(T)}$$

$$\frac{n_X(T)}{n(T)} \cdot \frac{n(T)}{n_{(X\setminus\hat{X})}(T)} = \frac{n_X(T)}{n_{(X\setminus\hat{X})}(T)}$$

2

1

Man hat z. B folgendes Itemset:

{Mehl, Zucker, Butter, Eier, Backpulver}

Wobei B+ = {Mehl, Zucker, Butter} und H- = {Eier, Backpulver}. Daraus ergibt sich dann z. B. B = {Mehl, Zucker} und H = {Butter, Eier, Backpulver}. Daraus lässt sich dann z. B. folgende Assoziation formulieren:

{Mehl, Zucker, Butter} -> {Eier, Backpulver}; B+ -> H-

{Mehl, Zucker} -> {Butter, Eier, Backpulver}; B -> H

Da in B+ jetzt mehr Elemente und in H- weniger Elemente vorhanden sind, kann man davon ausgehen, dass die Confidence dort höher ist, da n_L(T) vermutlich dort eine Menge mit weniger Tupeln ist und somit nach der Formel für die Confidence diese dort größer ist.

2

Man setze D auf die Differenz $B^+ - B$ bzw. $H - H^-$, welche hier die gleiche Menge definieren, was sich aus F ergibt:

$$\begin{aligned} B + H &= B^+ + H^- \qquad | - B - H^- \\ H - H^- &= B^+ - B \end{aligned}$$

$confidence(B^+ \Rightarrow H^-)$ ist nach Definition

$$\frac{n_{B^+ \wedge H^-}(T)}{n_{B^+}(T)} = \frac{n_{B+D \wedge H-D}(T)}{n_{B+D}(T)} = \frac{n_{B \wedge H}(T)}{n_{B+D}(T)}$$

Damit kann nun die Abschätzung stattfinden, dass die Anzahl der Tupel, in welchen B + D erfüllt ist, geringer-gleich, als die, in denen nur B erfüllt ist, also $n_B(T) \geq n_{B+D}(T)$ und daraus folgt dann

$$\frac{n_{B \wedge H}(T)}{n_{B+D}(T)} \geq \frac{n_{B \wedge H}(T)}{n_B(T)}$$

Daraus ergibt sich dann

$$\frac{n_{B \wedge H}(T)}{n_B(T)} = confidence(B \Rightarrow H)$$

Somit gilt also

$$confidence(B \Rightarrow H) \leqslant confidence\left(B^+ \Rightarrow H^-\right)$$