1. The reason why it is a drawback is because Pr(B)may be too high so that the data of Pr(B|A) have little value. For example, most people have drunk milk so that the information that most people drinking juice drinks milk is meaningless. Considering lift, if there is too much B, lift will not be too high, because the support will be high. And So does the conviction.
2. Conf: Since pr(a) isn't equal to pr(b), so pr(ab)/p(a) isn't equal to pr(ab)/p(b). so conf isn't symmetrical.
   Lift: lift(a|b)=n*pr(ab)/pr(b)*pr(a)=lift(b|a),so it is symmetrical
   Conv=(n*pr(a)-pr(a)*pr(b))/(n*pr(a))-n*pr(ab)
3. Confidence=p(ab)/p(a)=1, correct
   Lift=conf(a->b)/p(b)=1/p(b)
   Incorrect
   Conv(a->b)=(1-s(b))/1-conf(a->b)),since the denominator=0,so correct