Comparison of classifers: the McNemar test and the proportions test

When doing classification, we evaluate the classifiers by counting the number of true/false on a test set, and the outcome is an estimate p-hat of a probability p of correct prediction.

If we have two classifers, each classifier has its own probability.

Proportions test	Classifier 1	Classifier 2
Correct		
False		

McNemar test	Classifier 1 right	Classifier 1 wrong
Classifier 2 right		
Classifier 2 wrong		

1) How do the two tests differ? Do they have the same inputs?
2) What is the assumptions behind the proportions test?
3) Why does (could) this assumption go wrong for comparison of classifiers?
4) Think of a situation with classifiers where the proportions test would be correct.