Final Project: Judge-Jury Agreement

Due on November 25, 2015, 2:00PM

The table below shows the summarized results of a recent study of criminal cases in the United States by the National Center for State Courts (NCSC). The NCSC study produced judge and data on 290 nonhung-jury trials for noncapital criminal cases in 2000-2001. In the table below, column A is judge's classification as guilty (1) or not guilty (0), column B is jury's classification, column C is judge's assessment of strength of evidence (1 = weak, 2 = medium, 3 = strong), and D is jury's assessment.

Count	A	B	C	D	Count	A	B	C	D
7	0	0	1	1	0	1	0	1	1
6	0	0	1	2	0	1	0	1	2
0	0	0	1	3	0	1	0	1	3
8	0	0	2	1	6	1	0	2	1
15	0	0	2	2	26	1	0	2	2
0	0	0	2	3	1	1	0	2	3
0	0	0	3	1	2	1	0	3	1
0	0	0	3	2	11	1	0	3	2
0	0	0	3	3	1	1	0	3	3
1	0	1	1	1	0	1	1	1	1
2	0	1	1	2	3	1	1	1	2
2	0	1	1	3	2	1	1	1	3
1	0	1	2	1	1	1	1	2	1
7	0	1	2	2	41	1	1	2	2
1	0	1	2	3	36	1	1	2	3
0	0	1	3	1	2	1	1	3	1
1	0	1	3	2	36	1	1	3	2
0	0	1	3	3	52	1	1	3	3

Goal of data analysis: To understand the disagreement between the judge and the jury on (1) strength of evidence and (2) guilty/not guilty. In your nontechnical report, first assume that your client is a judge and provide suggestions/comments; then assume your client is a suspect and provide suggestions/comments.

Required questions:

1. Based on 290 trials, make comments on chances of disagreement on the strength of evidence according to the judge and the jury.

2. Given the strength of evidence according to the judge and the jury (both column C and column D), construct your best model to predict guilty/not guilty according to the judge (column A) and the jury separately. Estimate the chance of disagreement too.

3. Use 5-fold cross-validation to estimate the prediction error of your model in **2.**