Philosophy of Statistics: Homework 3

due on Gradescope by 11am on Thursday May 6

Guidelines. Some questions ask you to justify your answers. For these questions, credit will be based on how well you justify your answers, not on whether your answers are correct. (There's often no consensus on the correct answers, even among statisticians.) However, that doesn't mean that anything goes: some answers will be hard to justify well. I give suggested word counts but these are just ballpark numbers. Don't sweat them too much. Collaboration is encouraged, but make sure to write up your answers by yourself and list your collaborators.

Problem 1 (15 points). Can you come up with a crate-and-boxes problem and a decision rule with all the following properties:

- the true red probability is 3/5
- the accuracy is 7/10
- the false red rate is higher among blue cube boxes than yellow cube boxes
- the green prediction value is equal among blue and yellow cube boxes

Your answer should consist of frequency tables and a decision rule, as well as calculations to show that your example has all these properties.

Problem 2 (20 points). In one of the readings, Brian Hedden described a prediction problem and a decision rule such that: (i) the decision rule is (so he argued) perfectly fair, but (ii) the decision rule does not have equal true positive and true negative rates across groups. Here's the example:

> Room A contains 12 people with coins labeled '0.75' and 8 people with coins labeled '0.125'. The former are all predicted to be heads people (positive), and nine of them are in fact heads people. The latter are all predicted to be tails people (negative), and seven of them are in fact tails people. Room B contains 10 people with coins labeled '0.6' and 10 people with coins labeled '0.4'. The former are all predicted to be heads people, and six of them are in fact heads people. The latter are all predicted to be tails people, and six of them are in fact tails people.

As you can see, instead of using the crate-and-boxes framework, like we've been doing, Hedden's example involves people, rooms and coins. But can you translate Hedden's example into our framework? In other words, write down some frequency tables and a decision rule which have the same structure as Hedden's example, and then calculate the overall, blue and yellow confusion tables.

Problem 3 (20 points). Let's think more about the *accuracy* of a decision rule in a crate-and-boxes problem.

The red prediction value is the probability of being right, given that you predict red. The green prediction value is the probability of being right, given that you predict green. The accuracy is the probability of being right, unconditionally. So these are three related measures of accuracy. And, as we've seen, we can calculate these either overall, or among blue cube boxes, or among yellow cube boxes.

OK, now imagine the following scenario: the accuracy is higher for yellow cube boxes than blue cube boxes, but the red prediction value and green prediction value are both higher for blue cube boxes than yellow cube boxes. That would be an awkward situation. Our three measures of accuracy would be pointing in different directions: we'd be more accurate for yellow than blue cube boxes overall, but more accurate for blue than yellow cube boxes given each prediction.

First, show that, surprising as it might seem, that scenario is in fact possible. Your answer should consist of frequency tables and a decision rule, along with a calculation of the relevant probabilities.

Second, discuss what we should make of such a situation. For example, which measure(s) of accuracy, if any, should we care about?

Problem 4 (15 points). Some theorists have claimed that a decision rule is fair only if it has all four properties we talked about in class. Others have claimed that all that's required is to balance the positive and negative prediction values, not the true positive and true negative rates. Yet others have claimed the reverse. And others still have claimed that what matters is some other property, such as equal accuracy across groups. Or perhaps all these properties are just red herrings: a fair decision rule needn't balance *any* of them. Who is right?

In your answer, make sure to support your claims by reference to real-life prediction problems (e.g. the examples on Slide 15 of Class 6). (300 words)

Problem 5 (15 points). Judy Thomson contrasted cases such as the following:

> EYEWITNESS. Mrs. Smith was driving home late one night. A cab came towards her, weaving wildly from side to side across the road. She had to swerve to avoid it. Her swerve took her into a parked car. In the crash, she suffered two broken legs. Mrs. Smith therefore sued Red Cab Company. Her evidence is as follows: there are only two cab companies in town, Red Cab (all of whose cabs are red) and Green Cab (all of whose cabs are green), and an eyewitness testified that the cab which caused the accident was red.
>
> MARKET SHARE. Mrs. Smith was driving home late one night. A cab came towards her, weaving wildly from side to side across the road. She had to swerve to avoid it. Her swerve took her into a parked car. In the crash, she suffered two broken legs. Mrs. Smith therefore sued Red Cab Company. Her evidence is as follows: there are only two cab companies in town, Red Cab (all of whose cabs are red) and Green Cab (all of whose cabs are green), and of the cabs in town that night, nine out of ten were operated by Red Cab.

She claimed that in EYEWITNESS the evidence against Red Cab is "individualized" but in MARKET SHARE it is "merely statistical".

First, can you come up with your own pair of cases such that Thomson would say the evidence is individualized in the first case but merely statistical in the second? Your cases should be original, not just minor variations on cases we discussed in class.

Second, either try to characterize what the key difference is between the evidence in your pair of cases, or, if you think there is no important difference, because Thomson's distinction between individualized and merely statistical evidence is spurious, then explain why. (200 words)

Problem 6 (15 points). Recall an example from class:

> PRISONERS. 25 prisoners are exercising in a prison yard, when 24 of them suddenly join together in a planned attack on the prison guards. The remaining prisoner tries to stop the attack. There is no available evidence distinguishing the innocent prisoner from the rest. Local prosecutors randomly select one of the prisoners and bring him to trial for participating in the attack. (Nesson 1979)

Consider the following argument:

(P1) Given the evidence, you should be 96% sure that the defendant attacked the guard.

(P2) You should convict the defendant just if, given the evidence, it is beyond a reasonable doubt that the defendant is guilty.

(C) So, you should convict the defendant.

Evaluate this argument. If a premise is false, which and why? If the conclusion doesn't follow from the premises, then why not? If the conclusion is true, then what explains people's general reluctance to convict on the basis of such evidence? Try to anticipate how someone might object to your evaluation. (300 words)