

Professors and other students hate it when students cheat on exams, and some professors have devised methods that make it relatively easy to detect cheating on essays and similar questions. But detecting cheating on multiple choice exams is not as easy. That's where probability and statistics can help. Professor Robert Mogull thought he detected cheating by two students on multiple choice exams in his statistics class at Sacramento State University because they had identical questions wrong on all four 25-question multiple choice exams. He calculated the probability of that happening to be extremely small, failed the two students, and published a paper explaining his method (Mogull, 2003). But others have criticized his probability calculations because he assumed that all students were equally likely to miss any particular question (Actuarial Outpost, 2013). The critics pointed out that there are all sorts of reasons why two particular students might miss the same questions as each other, especially if they were friends. Perhaps they studied together, they had the same major and thus had similar knowledge, had the same statistics course in high school, and so on.

In a similar case, Klein (1992) described a situation in which two students were accused of cheating on a multiple-choice medical licensing exam. They had been observed whispering during one part of the 3-day exam and their answers to the questions they got wrong very often matched each other. The licensing board determined that the statistical evidence for cheating was overwhelming. They estimated that the odds of two people having answers as close as these two did were less than 1 in 10,000. Further, the students were husband and wife. Their tests were invalidated.

The case went to trial, and upon further investigation, the couple was exonerated. They hired a statistician who was able to show that the agreement in their answers during the session in which they were whispering was no higher than it was in the other sessions. What happened? The board assumed students who picked the wrong answer were simply guessing among the other choices. This couple had grown up together and had been educated together in India. Answers that would have been correct for their culture and training were incorrect for the American culture (for example, whether a set of symptoms was more indicative of tuberculosis or a common cold). Their common mistakes often would have been the right answers for India. So, the licensing board erred in calculating the odds of getting such a close match by using the assumption that they were just guessing. And, according to Klein, "with regard to their whispering, it was very brief and had to do with the status of their sick child" (p. 26). ■