Collusion in Multiple Choice Examinations

S. J. KILPATRICK, JR.

Department of Biometry,
Medical College of Virginia,
Richmond, 23219

Many medical schools have now adopted multiple choice examinations and computer grading. The computer may also be used to detect or confirm collusion between students in these examinations. In the following, a method is given for calculating the probability that two given answer sheets by chance agree to the extent observed.

Method

Consider a multiple choice examination. It is not necessary to assume that the number of alternatives in each question is constant. For a given pair of students,

let z = the number of questions which both students got wrong,
let x = the number of these z questions on which both students agreed,
and y = the number of these z questions on which the students disagreed.
Note $x + y = z$

The computer is programmed to get $(x/z)\%$, the % agreement between the two suspected students among those questions jointly wrong; and $(X/Z)\%$, the maximum % agreement between each of these students and every other member of the class. We can now tabulate these figures as follows:-

	No. of these with same (wrong) answer	No. of these with different (wrong) answer	Total No. of questions jointly wrong
Suspected Pair	x	y	z
Next Highest pair involving one of the above	X	Y	Z
Total	x + X	y + Y	z + Z

$$\chi^2 = \frac{(xY - Xy)^2(z + Z)}{(x + X)(y + Y).z.Z}$$

is then evaluated as a single tailed χ^2 with 1 d.f. to get P, the probability that these erroneous agreements occurred by chance.

Example

In an examination consisting of 200 multiple choice questions, two students, S_1 and S_2, are suspected of being in collusion. S_1 scored 127 (63.5%) out of 200 whereas S_2 had 105 (52.5%). Sixty-five questions were jointly wrong in these two papers and of these S_1 and S_2 agreed in 51, ie, 78%. A listing of all possible pairs including either S_1 or S_2 showed that the next highest percentage agreement among jointly wrong questions occurred with S_1 and S_{35} for which there were 53 questions jointly wrong and 26 of these agreed. The fourfold table is then:

	No. Agreeing	No. disagreeing	Total No. jointly wrong
S_1 vs S_2	51 (78%)	14	65
S_1 vs S_{35} next highest	26 (49%)	27	53
Totals	77	41	118

$$\chi^2 = \frac{(51 \times 27 - 26 \times 14)^2.118}{77 \times 41 \times 65 \times 53} = 11.13$$

$$P(\chi_1^2 \geq 11.13) < .001$$

On this basis there is less than one chance in 1000 that the two students, S_1 and S_2, accidentally agreed with each other to the extent observed or greater, ie, 51 out of 65 or 78%.

Discussion

This approach may be criticized on a number of statistical grounds but the method advocated is simple, makes few assumptions, and gives the suspected students the benefit of the doubt. The resultant probability is the chance that the two students fortuitously agree with respect to their wrong answers to the extent observed or greater as compared against the next closest pair. This of course assumes that only two students are in collusion. If there are more than two students suspected, the above test needs to be modified in an obvious way.